

N O T I C E

THIS DOCUMENT HAS BEEN REPRODUCED FROM
MICROFICHE. ALTHOUGH IT IS RECOGNIZED THAT
CERTAIN PORTIONS ARE ILLEGIBLE, IT IS BEING RELEASED
IN THE INTEREST OF MAKING AVAILABLE AS MUCH
INFORMATION AS POSSIBLE



80-10030
TM-80811

(E80-10030) PROCEEDINGS OF TECHNICAL SESSIONS, VOLUMES 1 AND 2: THE LACIE SYMPOSIUM (NASA) 1095 p HC A99/MF A01

CSCL 02C

N80-15448
THRU
N80-15525
Unclas
00030

G3/43

PROCEEDINGS OF TECHNICAL SESSIONS VOLUME 1 + 2

THE LACIE SYMPOSIUM

OCTOBER 1978

Made available under the sponsorship of the National Aeronautics and Space Administration. For any use, please refer to the original source.

Original photography may be purchased from ERUS Data Center

Sioux Falls, SD 57198



National Aeronautics and Space Administration

Lyndon B. Johnson Space Center
Houston, Texas 77058

JULY 1979

A TECHNICAL DESCRIPTION OF THE LARGE AREA CROP INVENTORY EXPERIMENT (LACIE)

THE LACIE SYMPOSIUM

OCTOBER 23 TO 26, 1978
NASA JOHNSON SPACE CENTER

Sponsored by

NASA

—National Aeronautics and Space Administration,
Lyndon B. Johnson Space Center

USDA

—United States Department of Agriculture

NOAA

—National Oceanic and Atmospheric Administration

Foreword

Accurate and timely global crop forecasts are of the utmost importance to all countries in managing their most important resource — food. Information about agricultural production is crucial to a wide range of decisions by agribusiness, policymakers, resource planners, and agriculture technologists. Decisions made on the basis of inadequate information regarding the food supply, its distribution, and the expectation from new harvests can have severe economic and social impact.

The United States is a major partner in a large and interdependent worldwide network of importing and exporting nations. The United States is the foremost grain exporter to a global market increasing to meet an expanding world demand. The resupply from new harvests is extremely variable within each year and between years. The world's most important food grain, wheat, is in the process of being planted and harvested in different regions of the world throughout each year. This crop is grown mostly in semiarid regions that have marginal weather conditions, where disaster years followed by years of bumper crops are common. Although such organizations as the United Nations Food and Agriculture Organization (FAO) and the United States Department of Agriculture (USDA) were chartered to provide information on global food production, their reports have been heavily reliant on information generated by the countries themselves. This information is derived from crop survey systems that are often inadequate or, in some cases, nonexistent.

Aerospace remote-sensing technology emerging from several decades of research is beginning to provide a means to economically provide better crop forecasts.

In 1974, the Large Area Crop Inventory Experiment (LACIE) — a joint effort of the National Aeronautics and Space Administration (NASA), the USDA, and the National Oceanic and Atmospheric Administration (NOAA) — began to apply this technology on an experimental basis to forecasting harvests in important wheat-production areas. Following completion of the analysis of data acquired over 3 global crop years, the results were documented and reported in a 4-day symposium held in October 1978 at the NASA Johnson Space Center in Houston, Texas. Prior to the symposium, a team consisting of approximately 40 independent university, industry, and government scientists and researchers assembled periodically from March through July 1978 at the Johnson Space Center to review the LACIE results in considerable detail. These peer-review results were also reported at the symposium.

The material contained in this document consists of the proceedings of the technical sessions of that symposium. The overview and peer-review papers are published in other volumes.

ROBERT B. MACDONALD
Manager, LACIE Project

July 1979

Contents

Page

VOLUME I

EXPERIMENT DESIGN SESSION

FOREWORD	1
SAMPLING, AGGREGATION, AND VARIANCE ESTIMATION FOR AREA, YIELD, AND PRODUCTION IN LACIE	3
<i>C. R. Hallum, R. S. Chhikara, A. H. Feiveson, and A. G. Houston</i>	
ECONOMETRIC MODELS FOR PREDICTING CONFUSION CROP RATIOS	19
<i>D. E. Umberger, M. H. Proctor, J. E. Clark, L. M. Eisgruber, and C. B. Braschler</i>	
LACIE SAMPLING DESIGN	47
<i>A. H. Feiveson, R. S. Chhikara, and C. R. Hallum</i>	
LACIE AREA SAMPLING FRAME AND SAMPLE SELECTION	53
<i>C. J. Liszcz</i>	
LACIE LARGE-AREA ACREAGE ESTIMATION	57
<i>R. S. Chhikara and A. H. Feiveson</i>	
LARGE-AREA AGGREGATION AND MEAN-SQUARED PREDICTION ERROR ESTIMATION FOR LACIE YIELD AND PRODUCTION FORECASTS	67
<i>R. S. Chhikara and A. H. Feiveson</i>	
CLASSIFICATION AND MENSURATION OF LACIE SEGMENTS	73
<i>R. P. Heydorn, R. M. Bizzell, J. A. Quirein, K. M. Abotteen, and C. A. Sumner</i>	
LACIE REGISTRATION PROCESSING	87
<i>Gerald J. Grebowsky</i>	
DEVELOPMENT OF LACIE CCEA-I WEATHER/WHEAT YIELD MODELS	99
<i>N. D. Strommen, C. M. Sakamoto, S. K. LeDuc, and D. E. Umberger</i>	
GROWTH STAGE ESTIMATION	109
<i>V. S. Whitehead, D. E. Phinney, and W. E. Crea</i>	
ACCURACY ASSESSMENT: THE STATISTICAL APPROACH TO PERFORMANCE EVALUATION IN LACIE	115
<i>A. G. Houston, A. H. Feiveson, R. S. Chhikara, and E. M. Hsu</i>	
MANUAL INTERPRETATION OF LANDSAT DATA	131
<i>C. M. Hay</i>	

✓

III

	Page
SYSTEM IMPLEMENTATION AND OPERATIONS SESSION	
FOREWORD	145
ACQUISITION AND PREPROCESSING OF LANDSAT DATA	147
<i>T. N. Horn, L. E. Brown, and W. H. Anonsen</i>	
ANCILLARY DATA ACQUISITION FOR LACIE	157
<i>B. E. Spiers and R. L. Patterson</i>	
LACIE DATA-HANDLING TECHNIQUES	163
<i>G. H. Waits</i>	
THE ACQUISITION, STORAGE, AND DISSEMINATION OF LANDSAT AND OTHER LACIE SUPPORT DATA	169
<i>L. F. Abbotts and R. M. Nelson</i>	
THE CLASSIFICATION AND MENSURATION SUBSYSTEM	179
<i>K. M. Abotteen and R. M. Bizzell</i>	
CONCEPTS LEADING TO THE IMAGE-100 HYBRID INTERACTIVE SYSTEM.....	203
<i>T. F. Mackin and J. M. Sulester</i>	
USDA ANALYST REVIEW OF THE LACIE IMAGE-100/HYBRID SYSTEM TEST.....	211
<i>P. Ashburn, K. Bülow, H. L. Hansen, and G. A. May</i>	
OPERATION OF THE YIELD ESTIMATION SUBSYSTEM.....	217
<i>D. G. McCrary, J. L. Rogers, and J. D. Hill</i>	
THE CROP ASSESSMENT SUBSYSTEM: SYSTEM IMPLEMENTATION AND APPROACHES USED FOR THE GENERATION OF CROP PRODUCTION REPORTS...	227
<i>W. E. McAllum, R. E. Hatch, S. M. Boatright, C. J. Liszcz, and S. M. Evans</i>	
LACIE STATUS AND TRACKING.....	243
<i>V. M. Dauphin, C. H. Jeffress, and J. M. Everette</i>	
LACIE QUALITY ASSURANCE	249✓
<i>G. L. Gutschewski</i>	
OPERATIONS REPORTING.....	257°
<i>R. G. Musgrove and Dale R. Marquis</i>	
EOD FACILITIES CONFIGURATION MANAGEMENT OFFICE.....	261°
<i>V. M. Dauphin and R. F. Palmer</i>	
ACCURACY ASSESSMENT SYSTEM AND OPERATION.....	265
<i>D. E. Pitts, A. G. Houston, G. Badhwar, M. J. Bender, M. L. Rader, W. G. Eppler, C. W. Ahlers, W. P. White, R. R. Vela, E. M. Hsu, J. F. Potter, and N. J. Clinton</i>	
LACIE APPLICATIONS EVALUATION SYSTEM EFFICIENCY REPORT	281
<i>Timothy T. White</i>	

✕
 ✕

	Page
DATA PROCESSING SYSTEMS DESIGN SESSION	
FOREWORD	289
CARTOGRAPHY: LACIE'S SPATIAL PROCESSOR	291
<i>M. L. Rader and R. R. Vela</i>	
THE LACIE DATA BASES: DESIGN CONSIDERATIONS.....	297
<i>L. E. Westberry</i>	
MAN-MACHINE INTERFACES IN LACIE/ERIPS	307
<i>Barbara B. Duprey</i>	
LACIE/ERIPS SOFTWARE SYSTEM SUMMARY	317
<i>C. I. Johnson</i>	
CONSIDERATIONS FOR DESIGN OF FUTURE RESEARCH AND DEVELOPMENT INTERACTIVE IMAGE ANALYSIS SYSTEMS.....	333
<i>T. B. Wilkinson</i>	
VERY HIGH SPEED PROCESSING: APPLICABILITY OF PERIPHERAL DEVICES TO PIXEL-DEPENDENT TASKS.....	345
<i>J. C. Lyon</i>	
A LOOK AT COMPUTER SYSTEM SELECTION CRITERIA.....	355
<i>E. W. Poole, F. L. Flowers, and W. I. Stanley</i>	
COST AND PERFORMANCE CHARACTERISTICS OF DATA SYSTEM CONFIGURATIONS FOR PROCESSING REMOTELY SENSED DATA	397
<i>P. J. Gregor and J. F. Spitzer</i>	
EXPERIMENT RESULTS AND ACCURACY SESSION	
FOREWORD	409
THE LACIE CROP YEARS: AN ASSESSMENT OF THE CROP CONDITIONS EXPERIENCED IN THE 3 YEARS OF LACIE.....	411
<i>J. D. Hill and D. R. Thompson</i>	
APPLICATION OF LANDSAT DIGITAL DATA FOR MONITORING DROUGHT	431
<i>D. R. Thompson and O. A. Wehmanen</i>	
LACIE AREA, YIELD, AND PRODUCTION ESTIMATE CHARACTERISTICS: U.S. GREAT PLAINS.....	439
<i>Duane L. Marquis</i>	
LACIE AREA, YIELD, AND PRODUCTION ESTIMATE CHARACTERISTICS: U.S.S.R.	481
<i>J. R. Hickman</i>	



	Page
LACIE AREA, YIELD, AND PRODUCTION ESTIMATE CHARACTERISTICS: CANADA	513
<i>Delanne Conte, A. G. Houston, and L. O. Lovfald</i>	
ACCURACY AND PERFORMANCE OF LACIE AREA ESTIMATES	527
<i>J. F. Potter, E. M. Hsu, A. G. Houston, and D. E. Pitts</i>	
ACCURACY AND PERFORMANCE OF LACIE YIELD ESTIMATES IN MAJOR WHEAT PRODUCING REGIONS OF THE WORLD	575
<i>D. E. Phinney, R. G. Stuff, A. G. Houston, E. M. Hsu, and M. H. Trenchard</i>	
ACCURACY AND PERFORMANCE OF LACIE CROP DEVELOPMENT MODELS	589
<i>S. K. Woolley, V. S. Whitehead, R. G. Stuff, and W. E. Crea</i>	
ECONOMIC EVALUATION: CONCEPTS, SELECTED STUDIES, SYSTEM COSTS, AND A PROPOSED PROGRAM	605
<i>Frank H. Osterhoudt</i>	
SUPPORTING RESEARCH AND TECHNOLOGY SESSION	
FOREWORD	617
METHODS FOR SEGMENT WHEAT AREA ESTIMATION	621
<i>R. P. Heydorn, M. C. Trichel, and J. D. Erickson</i>	
ESTIMATING CROP PROPORTIONS FROM REMOTELY SENSED DATA	633
<i>A. H. Feiveson</i>	
ON THE CLUSTERING OF MULTIDIMENSIONAL PICTORIAL DATA	647
<i>J. D. Bryant</i>	
ON EVALUATING CLUSTERING PROCEDURES FOR USE IN CLASSIFICATION	661
<i>M. D. Pore, T. E. Moritz, D. T. Register, S. S. Yao, and W. G. Eppler</i>	
CLASSY—AN ADAPTIVE MAXIMUM LIKELIHOOD CLUSTERING ALGORITHM	671
<i>R. K. Lenington and M. E. Rassbach</i>	
LINEAR FEATURE SELECTION WITH APPLICATIONS	691
<i>H. P. Decell, Jr., and L. F. Guseman, Jr.</i>	
FEATURE EXTRACTION APPLIED TO AGRICULTURAL CROPS AS SEEN BY LANDSAT	705
<i>R. J. Kauth, P. F. Lambeck, W. Richardson, G. S. Thomas, and A. P. Pentland</i>	
COMPENSATION FOR ATMOSPHERIC EFFECTS IN LANDSAT DATA	723
<i>P. F. Lambeck and J. F. Potter</i>	
DEVELOPMENT OF PARTITIONING AS AN AID TO SPECTRAL SIGNATURE EXTENSION	739
<i>R. W. Thomas, C. M. Hay, and J. C. Claydon</i>	

	Page
METHODS OF EXTENDING CROP SIGNATURES FROM ONE AREA TO ANOTHER.....	757
<i>T. C. Minter</i>	
SIGNATURE EXTENSION METHODS IN CROP AREA ESTIMATION.....	801
<i>R. J. Kauth and W. Richardson</i>	
AN EVALUATION OF PROCEDURE 1.....	825
<i>S. G. Wheeler, R. P. Heydorn, P. N. Misra, W. Lee, Jr., and R. T. Smart</i>	
THE VEGETATIVE INDEX NUMBER AND CROP IDENTIFICATION.....	843
<i>P. Ashburn</i>	
MANUAL LANDSAT DATA ANALYSIS FOR CROP TYPE IDENTIFICATION.....	857
<i>C. M. Hay</i>	
LACIE ANALYST INTERPRETATION KEYS.....	867
<i>J. G. Baron, R. W. Payne, and W. F. Palmer</i>	
COLORIMETRIC CONSIDERATION OF TRANSPARENCIES FOR A TYPICAL LACIE SCENE.....	887
<i>R. D. Juday</i>	
GENERATION OF UNIFORM CHROMATICITY SCALE IMAGERY FROM LANDSAT DATA.....	899
<i>R. D. Juday, F. Johnson, R. A. Abotteen, and M. D. Pore</i>	
IMAGE AND NUMERICAL DISPLAY AIDS FOR MANUAL INTERPRETATION.....	911
<i>R. A. Abotteen</i>	
A PROGRAMED LABELING APPROACH TO IMAGE INTERPRETATION.....	923
<i>M. D. Pore and R. A. Abotteen</i>	
STATUS OF YIELD ESTIMATION TECHNOLOGY: A REVIEW OF SECOND-GENERATION MODEL DEVELOPMENT AND EVALUATION.....	937
<i>R. G. Stuff, T. L. Barnett, G. O. Bootwright, D. E. Phinney, and V. S. Whitehead</i>	
A UNIVERSAL MODEL FOR ESTIMATING WHEAT YIELDS.....	951
<i>A. M. Feyerherm and G. M. Paulsen</i>	
THE LAW OF THE MINIMUM AND AN APPLICATION TO WHEAT YIELD ESTIMATION.....	961
<i>R. B. Cate, D. E. Phinney, and M. H. Trenchard</i>	
CCEA SECOND-GENERATION WHEAT YIELD MODEL FOR HARD RED WHEAT IN NORTH DAKOTA.....	971
<i>S. K. LeDuc</i>	
PREDICTION OF WHEAT PHENOLOGICAL DEVELOPMENT: A STATE-OF-THE-ART REVIEW.....	981
<i>M. W. Seeley, M. H. Trenchard, D. E. Phinney, J. R. Baker, and R. G. Stuff</i>	
NEW DEVELOPMENTS IN SAMPLING AND AGGREGATION FOR REMOTELY SENSED SURVEYS.....	991
<i>A. H. Feiveson</i>	

ix

vii

	Page
NATURAL SAMPLING STRATEGY	995
<i>C. R. Hallum and J. P. Basu</i>	
MULTIYEAR ESTIMATES FOR THE LACIE SAMPLING PLANS	1015
<i>H. O. Hartley</i>	
WEIGHTED AGGREGATION	1029
<i>A. H. Felteson</i>	
DESIGN, IMPLEMENTATION, AND RESULTS OF LACIE FIELD RESEARCH	1037
<i>M. E. Bauer, M. C. McEwen, W. A. Malila, and J. C. Harlan</i>	
THE USDA APPLICATION TEST SYSTEM SESSION	
FOREWORD	1067
THE APPLICATION TEST SYSTEM: AN APPROACH TO TECHNOLOGY TRANSFER	1069
<i>A. C. Aaronson, K. Bülow, F. C. David, R. L. Packard, and F. W. Ravet</i>	
FUNCTIONAL DEFINITION AND DESIGN OF A USDA SYSTEM	1075
<i>S. M. Evans, E. R. Dario, and G. L. Dickinson</i>	
DATA BASE DESIGN FOR A WORLDWIDE MULTICROP INFORMATION SYSTEM ...	1085
<i>W. G. Driggers, J. M. Downs, J. R. Hickman, and R. L. Packard</i>	
THE APPLICATION TEST SYSTEM: TECHNICAL APPROACH AND SYSTEM DESIGN	1097
<i>J. L. Benson, D. R. McClelland, J. D. Tarbet, and R. T. Purnell</i>	
RESOURCE MODELING: A REALITY FOR PROGRAM COST ANALYSIS	1103
<i>L. D. Fouts and R. L. Hurst</i>	
THE APPLICATION TEST SYSTEM: EXPERIENCES TO DATE AND FUTURE PLANS	1121
<i>G. A. May, P. Ashburn, and H. L. Hansen</i>	



 VIII

Experiment Design

FOREWORD

In general, experiment design refers to the design used to ensure that information collected in an experiment will be relevant to the problems under investigation. It is the complete sequence of steps taken ahead of time to ensure that appropriate data will be obtained to permit valid inferences concerning these problems. Ideally, an experiment design should arrange the data collection and analysis to answer the experimental questions as efficiently as possible. The principles of good experimental design were given heavy emphasis in LACIE planning activities, and every effort was made to learn as much as possible given the available time, money, personnel, and experimental material.

The general theme of the LACIE design consisted of (1) identifying the state-of-the-art technology that would permit inventory of crops using satellite and meteorological data, (2) making necessary tests and evaluations to develop procedures that used state-of-the-art technology and to determine how well these procedures worked, and (3) subjecting the designed system to use in a quasi-operational environment for final performance assessment and identification of needed refinements and improvements.

The LACIE design activities were structured into five major technical components: (1) sampling and aggregation, (2) growth stage estimation, (3) classification and mensuration, (4) yield estimation, and (5) accuracy assessment.

The complexity and interdependency of these components and their supporting systems necessitated that a major part of the overall design be composed of the structuring and dovetailing of the constituent parts into a quasi-operational system that supported the LACIE objectives while making every effort to conserve time, money, personnel, and experimental material. The purpose of the Experiment Design Session is to detail the LACIE technical

design, including the relationships between components and (for each component) (1) the initial state-of-the-art methodology in LACIE, (2) the test and evaluation procedures used to identify or improve the state of the art, (3) the performance assessment procedures applied to the overall system, and (4) the chronology of the state of the art as it evolved during the 3 years of LACIE.

Sampling and Aggregation

Throughout LACIE, the only feasible cost-effective way to meet project objectives was to determine total wheat production over an area by looking at only a subset of the area. Consequently, LACIE technology drew heavily from statistical survey methodology (supported by a broad base of remote-sensing technology). The paper by Hallum et al. entitled "Sampling, Aggregation, and Variance Estimation for Area, Yield, and Production in LACIE" provides an overview of the LACIE sampling technology used in a quasi-operational mode throughout Phases I, II, and III. A general description, the rationale, design restrictions, and other design characteristics are given for the sampling design and the aggregation procedures for estimating wheat area, yield, and production, along with an overview of their associated prediction error estimates. Specific details are provided in the supporting papers.

Crop Development Stage Estimation

In the early preparation for LACIE, it was apparent that year-to-year variations in the seasons made the use of unadjusted crop calendars to distinguish wheat from other crops a questionable procedure. It was further recognized that because yields could be drastically affected by unusual events at

critical times in wheat development (i.e., high temperatures at heading) yield models to be developed would most likely require a good estimation of the true or actual development stage of the crop throughout the year. The paper by Whitehead and Phinney entitled "Growth Stage Estimation" provides an overview of the LACIE growth stage estimation technology, including a general description, the rationale, design restrictions, and other design characteristics.

Classification and Mensuration

One of the LACIE goals was to estimate wheat acreage using Landsat as the primary data source and without using ground enumerative data. A fundamental approach resulting from the LACIE design uses a machine classification technique to separate the wheat area in each of a number of 5- by 6-nautical-mile segments. It was apparent at the outset of LACIE that a limited amount of manual interpretation would be required for the method to work. The paper by Heydorn et al. entitled "Classification and Mensuration of LACIE Segments" provides the details of that part of the LACIE design. A vastly improved technology for the classification of complex data structures inherent in multirate acquisition of multispectral data has evolved from use of this approach. The major result of this evolution is the availability of a nearly optimum man/machine-processing procedure.

Yield Estimation

The paper by Strommen et al. entitled "Development of LACIE CCEA-I Weather/Wheat Yield Models" provides details of the design used to identify and evaluate yield estimation models oriented toward supporting project objectives. Included is a discussion of the rationale, the design restrictions, and the chronology of the evolving yield model development during Phases I, II, and III of LACIE.

Accuracy Assessment

An important function in LACIE is the evaluation of results obtained at various stages of the experiment. The objective of LACIE is not only to demonstrate the technological feasibility for estimating large-area wheat production using the LACIE approach but also to produce estimates which satisfy certain accuracy and reliability goals. The accuracy assessment effort is designed to check the accuracy of the products of the experimental operations throughout the crop growing season and to determine whether the procedures used are adequate to accomplish the desired accuracy and reliability goals. The paper by Houston et al. entitled "Accuracy Assessment: The Statistical Approach to Performance Evaluation of LACIE" describes the methodology for assessing the accuracy for area, yield, and production.

Sampling, Aggregation, and Variance Estimation for Area, Yield, and Production in LACIE

C. R. Hallum,^a R. S. Chhikara,^b A. H. Feiveson,^a and A. G. Houston^a

CROP INVENTORY: A STATISTICAL SURVEY

Classical Characteristics

Limited resources, a strong demand for breadth and timeliness of coverage, and recent advances in sample survey methodology are but a few reasons that crop inventory efforts have become heavily reliant on statistical survey methodology. This fact comes as no surprise since the task in the majority of such efforts is to determine the total crop area from information obtained over a subset of the area—a special case of the classical definition of a sample survey (ref. 1). In this case, two questions arise: (1) how to select the "part" from the "whole" and (2) how to generalize from the selected part to the whole. The problem is one of finding that combination of selection and estimation procedures which minimizes the cost, ensuring at the same time a specified accuracy for the inference from a part to the whole.

Until the last 40 years, little attention had been given to the problems of how to obtain a good sample and how to draw sound conclusions from the results. If the distribution from which one is sampling is uniform, then practically any sample will suffice; however, in the case of a crop inventory where the distribution is far from uniform, the method by which the sample is obtained is critical, and the study of techniques that ensure a trustworthy sample becomes extremely important.

In some cases, it may seem feasible to obtain the information desired about a particular population by taking a complete enumeration or census. Administrators accustomed to dealing with censuses tended

to be suspicious of samples and reluctant to use them in place of censuses. Although this attitude no longer persists, it would be worthwhile to list the principal advantages of sampling as compared with complete enumeration.

1. Reduced cost—Securing data from a small fraction of the population costs less than making a complete enumeration.

2. Greater speed—Data can be collected and summarized more quickly with a sample than with a census.

3. Greater scope—Surveys that rely on sampling have more scope and flexibility regarding the types of information that can be obtained; the area of coverage can be more extensive than in a census.

4. Greater accuracy—Because personnel of higher quality can be employed and given intensive training and because more careful supervision of the fieldwork and processing of results becomes feasible when the volume of work is reduced, a sample may actually produce more accurate results than a complete enumeration.

Upon examining the various steps required to perform a sample survey (ref. 1), it becomes quite clear that sampling is a practical business that calls for several different types of skills. In making a crop inventory, before sampling theory can be applied, it is necessary to determine which crops are to be considered, which geographical areas are to be surveyed, how data measurements are to be made, and how fieldwork is to be organized. Although these topics are not discussed further in this paper, their importance should be emphasized. Sampling demands attention to all phases of the activity—poor work in one phase may ruin a survey in which everything else is done well.

The purpose of sampling theory is to make sampling more efficient. It attempts to develop methods of sample selection and of estimation that provide, at minimum cost, estimates that are precise enough to satisfy project objectives. In order to apply this prin-

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Houston, Texas.

ciple, one must be able to predict the precision and the expected cost. So far as precision is concerned, the magnitude of estimation error in any specific situation cannot be foretold since this would require a knowledge of the true value for the population. One of the more standard ways of judging the precision, however, is to examine the frequency distributions of the estimates from a sampling procedure that has been applied repeatedly to the same population. A further simplification may be introduced in situations where the sample sizes are large enough that there is good reason to expect that sample estimates are approximately normally distributed (e.g., for accuracy assessment purposes, this is a key assumption—and, seemingly, not a bad one—made in regard to the distribution of the production estimator discussed later in this paper). In this case, the frequency distribution can be established precisely from the mean μ and the standard deviation σ , both of which can be estimated from sample survey theory. If a sample is taken by a procedure known to give an unbiased estimate $\hat{\mu}$ of μ with standard deviation $\sigma_{\hat{\mu}}$, then, although the exact value of the error $\hat{\mu} - \mu$ is unknown, from the properties of the normal distribution, the chances (probabilities) are 0.32 (about 1 in 3) that the absolute error $|\hat{\mu} - \mu|$ exceeds $\sigma_{\hat{\mu}}$; 0.05 (1 in 20) that the absolute error $|\hat{\mu} - \mu|$ exceeds $1.96 \sigma_{\hat{\mu}}$; and 0.01 (1 in 100) that the absolute error $|\hat{\mu} - \mu|$ exceeds $2.58 \sigma_{\hat{\mu}}$.

The preceding discussion assumes $\sigma_{\hat{\mu}}$, as computed from the sample, is known exactly. Actually $\sigma_{\hat{\mu}}$, like μ , is estimated subject to a sampling error, but for large sample sizes, the preceding results still hold.

When estimates are biased, a useful criterion is the mean squared error (MSE) of the estimate. In particular, the MSE of $\hat{\mu}$, denoted by $MSE(\hat{\mu})$ is given by

$$MSE(\hat{\mu}) = (\text{variance of } \hat{\mu}) + (\text{bias})^2 \quad (1)$$

In statistical terminology, the term "precision" refers to the repeatability of a measurement. "Low precision" means that there is wide variation in repeated measurements of the same object, whereas "high precision" means that there is little variation between repeated measurements. "Accuracy," however, refers to the MSE; the smaller the value of the MSE, the more accurate the estimator. Thus, low accuracy could result from a large-bias term with either low or high precision, or from a small-bias term coupled with low precision, as illustrated in

figure 1. High accuracy results when the bias term is small or zero and precision is high.

In the case of LACIE, where the level of accuracy is stipulated, the best sample design is that for which the cost of the survey is minimum. However, if the cost of the survey is specified, the best sample design is that which gives the highest accuracy. This rule is the guiding principle of classical sample survey methodology.

Overall Philosophy/Aspects Unique to LACIE

The first systematic attempt to collect agricultural statistics dates back more than a century to the Census of 1840 (ref. 2). From that date forward, an increasing volume of agricultural statistics has been collected periodically in the U.S. Census enumerations every 10 years to 1920 and every 5 years thereafter. A rudimentary system of annual agricultural estimation was also begun about 1840 in the Patent Office. Upon Commissioner Ellsworth's resignation in 1845, however, interest in agricultural statistics subsided in the Patent Office, and it was not until after the U.S. Department of Agriculture (USDA) was organized in 1862 that annual intercensus estimates were again revived (ref. 3).

Current monthly reports on crop conditions also predated the establishment of the Department of Agriculture by a few months. Orange Judd, editor of the American Agriculturalist, published summaries of crop condition reports submitted voluntarily by subscribers to his paper for 5 months, May through September 1862 (ref. 3). Judd's efforts were the forerunner to the USDA program of monthly reports on crop prospects. These reports have been issued

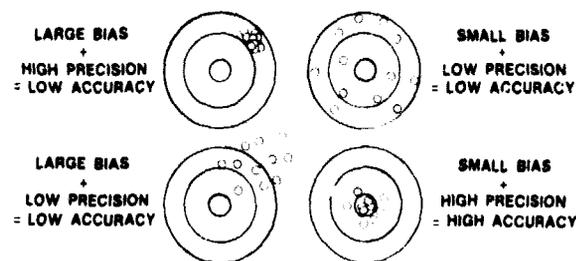


FIGURE 1.—Graphical representation of precision and accuracy. Circles denote estimates from repeated trials with the center of the target representing the location of the true quantity being estimated.

regularly during the growing season since the first publication in July 1863.

Since 1863, crop surveys in the Department of Agriculture have expanded greatly until today a large volume of agricultural estimates is published on a periodic basis. This substantial expansion in the volume of agricultural data has not been paralleled by major improvements in estimation methods, which is somewhat distressing in view of the significant developments in the theory of sample design—particularly in the past 40 years. Until recent efforts of the Statistical Reporting Service (SRS) and LACIE, the typical procedure has used mailed inquiries for collecting basic data and an assortment of techniques for removing bias in the transformation of raw data into published estimates.

A complete evaluation of agricultural statistics must embrace such characteristics as breadth of coverage, geographical detail of estimates, timeliness, and frequency of releases. However, the criterion of reliability should probably have top priority over all other criteria of evaluation. The choice of methods of estimation when the estimating agency is faced with limited resources and a strong demand for breadth and timeliness of coverage may well dictate some sacrifice of precision in the estimates.

At the outset of LACIE, remote-sensing technology without supporting ground data appeared to offer a cost-effective approach to a global crop estimation system that could provide improved information to USDA and NASA. More specifically, LACIE was the first attempt to survey an important crop (wheat) on a large (quasi-global) scale at repeated intervals over a wide range of conditions.

Given the project objectives, with specific emphasis on large-area estimation and the relative importance of more timely and more accurate estimates of foreign production, the LACIE design was restricted to readily available foreign data (e.g., Landsat imagery, meteorological data, National Oceanic and Atmospheric Administration (NOAA) satellite data, and published historical data for political subdivisions within countries). In addition, the LACIE system was designed within a framework of constraints originating from several sources. Examples of these constraints are the acquisition frequency restrictions specified by the NASA Goddard Space Flight Center (GSFC) and the requirement that implemented classification technology be used. Additional considerations were costs, schedule milestones, available resources for system implementation, the specified performance criterion,¹ and

the volume of Landsat data that could be stored and processed.

Some of the questions to be answered at the outset of LACIE were

1. Can a sampling strategy for acquisition of Landsat data be designed to achieve the required accuracy with a manageable data load?

2. How can the geographic wheat distribution characteristics (e.g., within-strata variances) best be determined so as to efficiently sample?

3. What is a good configuration for the primary sampling unit and what should the sampling frame be?

4. Does loss of the sampling unit due to cloud cover cause excessive errors, such as bias?

To develop and evaluate the LACIE survey system, the experiment was planned to consider first the wheat-growing regions of the United States, where reliable, independent survey estimates and ground data would be available. In Phase I, major emphasis was devoted to identifying significant problems and incorporating necessary changes into the on-line system.

To simplify the explanation of the LACIE sampling and aggregation approach, it will be worthwhile at this point to define the hierarchical structure of the units into which each country is subdivided. Each country is considered to be subdivided, first of all, into regions. Regions are the most "coarse" political subdivisions of a country (e.g., the U.S. Great Plains (USGP) is a U.S. "region"). Regions are further subdivided into zones. Zones are states, for example, in the United States and subcollections of oblasts in the U.S.S.R.; in any case, they are elements of the regional-level subdivisions. Zones are further subdivided into strata. In the United States, crop-reporting districts (CRD's) are the strata; they are subdivisions of the states. In the U.S.S.R., the strata are the oblasts. Finally, in countries with detailed historical data (i.e., in countries with historical data available at a level below the strata), the strata are further subdivided into units referred to as substrata. No further subdivision is made below the substrata

¹The LACIE performance goal for accuracy is to obtain country at-harvest production estimates which are within 10 percent of the actual production 90 percent of the time. This is referred to as the 90/90 criterion and is used to determine the LACIE error budget. The major requirement for the system was an ability to produce current estimates of wheat area, yield, and production for selected major wheat-producing regions on a scheduled basis throughout the crop season.

level. Examples of substrata are counties in the United States, shires in Australia, and municipalities in Canada. Figure 2 depicts the overall hierarchical structure.

Based on known constraints, a total allocation of 4800 sample segments was divided among the selected countries using the criteria and procedures discussed in the paper by Feiveson entitled "LACIE Sample Design."

Although certain engineering constraints affected the implementation of the initial LACIE sampling strategy, the evidence examined in Phase I indicated that these factors did not significantly affect accuracy. The majority of these constraints were removed from the Phase III design and are no longer inherent in the system. Various sampling and aggregation problems encountered throughout Phases I, II, and III of LACIE are discussed later in this paper.

In summary, LACIE's task has been to determine the total wheat production in an area by looking at only a subset of that area; consequently, LACIE technology draws heavily from statistical survey methodology which, in turn, is supported by a broad

base of remote-sensing technology. The remainder of this paper is an overview of the initial LACIE sampling technology used in a quasi-operational mode throughout Phases I, II, and III. A general description will be given of the sampling scheme and the aggregation procedures for estimating wheat area, yield, and production, along with a brief discussion of their associated prediction error estimates. Specific details are restricted to the supporting papers in this session.

SAMPLING, ESTIMATION, AND AGGREGATION FOR AREA

The LACIE sampling technology is designed to cost-effectively estimate wheat area and production in countries of interest with a predesignated precision level. The level of precision is dependent on many factors including (1) the configuration and geographical extent of the basic sampling unit (sample segment), (2) the sample selection procedure, (3) the number and distribution of the sample segments (i.e., the allocation procedure), and (4) the aggrega-

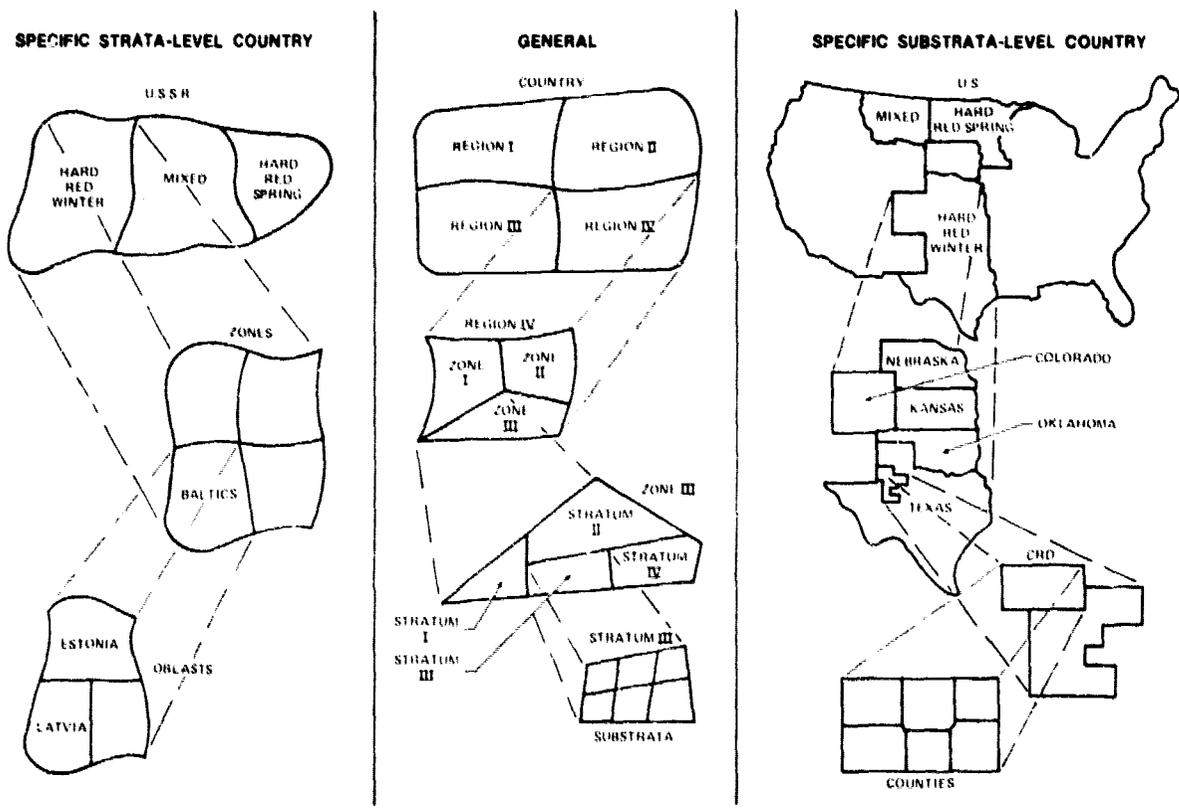


FIGURE 2.—Hierarchical relationship between country subdivision elements.

tion procedures for estimating wheat area, yield, and production. The following sections briefly summarize the major aspects of the LACIE sampling technology.

Configuration and Geographical Extent of the Sampling Unit

The sampling unit used in LACIE is a 5- by 6-nautical-mile rectangle. A key factor leading to this choice originated from GSFC engineering constraints. In particular, initial considerations were limited to areas no larger than 25 miles on a side since GSFC could not register areas larger than this to within ± 1 -pixel accuracy. Furthermore, GSFC's ability to handle a maximum of 4800 segments placed initial restrictions on the minimum segment size that could be tolerated. The two most important considerations, however, leading to this choice are attributable to the analysts' needs and to the required sampling precision. Initial indications were that an area of 30 square nautical miles was sufficiently large to provide the analyst with a good perspective on the variety and distribution of crops within a given locality. A smaller segment size tended to make the classification task more difficult to perform without evidencing any significant benefits. Retrospectively, the 5- by 6-nautical-mile segment has proved to be satisfactory in terms of both serving the analysts' needs and permitting required sampling precision without creating an unmanageable data load. Finally, the rectangular configuration was especially amenable to computer storage and manipulation.

Sample Selection Procedure

The LACIE sampling strategy depends on whether a country has detailed historical data (e.g., United States, Canada, and Australia) or whether it has data at only one level smaller than the country itself (e.g., U.S.S.R., China, Argentina, Brazil, and India). In the latter case, a standard stratified sampling scheme is employed, whereas in the first situation, the sampling strategy consists of a two-stage stratified random sample in which "substrata" (smallest political area for which acreage, yield, and crop calendars are published) are the primary sampling units. The 5- by 6-nautical-mile segments are the secondary units.

The sampling frame consists of the agricultural area within the major wheat-producing regions of a

country. It is a collection of 5- by 6-nautical-mile segments in agricultural areas as determined by an "ag/non-ag" delineation created from Landsat imagery and/or USDA Foreign Agricultural Service (FAS) land use maps. Each stratum/substratum, then, is a collection of segments. Figures 3 and 4 illustrate typical strata in countries with and without detailed historical data, respectively. In Phase I, the Landsat imagery over some areas to be sampled was of insufficient quality to support the sampling frame generation; over such areas, use was made of existing maps (e.g., land use, topographical) to help determine the ag/non-ag areas. The tendency in areas not having quality Landsat coverage was to be conservative and to retain areas that were questionable as agricultural. (In some cases, only cities and mountainous areas were excluded.) Although this approach increases the chances of including all the wheat in the sampling frame, it can result in a higher percentage of segments with little or no wheat as well as a higher wheat area variance. This situation im-

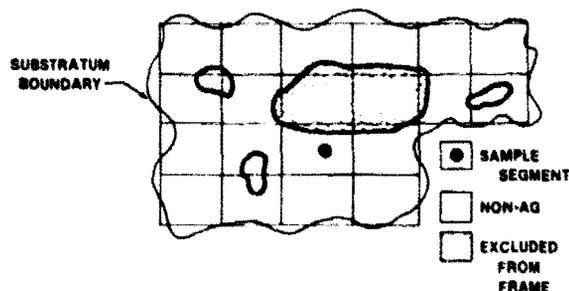


FIGURE 3.—Substratum sampling frame.

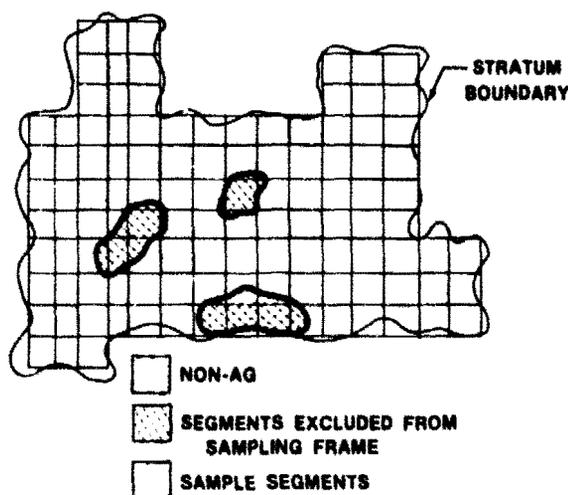


FIGURE 4.—Stratum sampling frame.

proved in Phase III with the availability of improved Landsat data coverage over such areas.

Allocation of Segments

Background for estimation of total country sample size.—At the outset of LACIE, a determination was made that more than 600 sample segments would be required in the United States to achieve an expected country-level sampling error of approximately 2.5 percent for wheat. On the basis of this determination, a proportional (to the wheat acreage from an epoch year) allocation was performed to seven other major wheat-producing countries (U.S.S.R., Brazil, India, Canada, Australia, China, and Argentina) to determine a worldwide (i.e., eight country) allocation. The resulting segment total was slightly less than the determined system capacity (because of hardware, data manipulation capabilities, etc.) of approximately 4800 segments. Consequently, early in the project, a decision was made to perform a worldwide allocation (to the previously mentioned eight countries) of 4800 segments, divided among the countries in proportion to their wheat acreages in an epoch year. Specifically, using the 1972 wheat area for the eight countries (obtained from FAS agricultural attaches and others), the country-level allocations, based on wheat area in thousands of hectares, were as follows.

Country	Sample segments	Wheat area
1. United States	637	19 138
2. U.S.S.R.	1949	58 492
3. Brazil	47	1 500
4. India	626	19 139
5. Canada	283	8 640
6. Australia	257	7 776
7. China	810	24 400
8. Argentina	165	4 965

Alterations to this allocation were made beginning in Phase II which resulted in different country totals for the United States in Phases II and III and for the United States and the U.S.S.R. in the Transition Year. In particular, during Phase II, a significant underestimate of the wheat area was observed in North Dakota. Further analysis indicated that the major problem was with the sample placement rather than with the classification. Indicated solutions were the allocation of additional segments or improved stratification to reduce agricultural area variability, or both. Consequently, 20 additional segments were

allocated to North Dakota in Phase II, resulting in a significantly improved wheat area estimate for that zone. These results were the primary driver for the decision to make a revised allocation over the yardstick region (the USGP in this case) for Phase III. Additional motivation for performing a revised allocation included the following:

1. To reduce the sampling error to approximately 2.2 percent (i.e., to a point of relative insignificance compared to the classification error, then shift emphasis to improving the classification procedure)
2. To make use of the improved Landsat imagery in an updated sampling frame
3. To employ a set of segments allocated such that the LACIE production estimate could be expected to satisfy the 90/90 criterion after allowing for errors due to sampling, classification, yield prediction, and loss of data

The revised allocation resulted in an increase in the total number of sample segments in the USGP from 431 in Phase II to 601 in Phase III.

From Phase III to the Transition Year, the allocations changed for the U.S.S.R. and the United States. The U.S.S.R. difference resulted from the verification in Phase III of what had been suspected for some time—the U.S.S.R. had been oversampled in the initial allocation. Consequently, a revised allocation was made, oriented toward achieving the 90/90 criterion for the LACIE production estimate, allowing for the same errors as in the yardstick region. The result was a reduction from 1949 to 1111 segments in the U.S.S.R.

The alteration in the USGP allocation from Phase III to the Transition Year was made for purposes of testing a "natural" sampling strategy (the details of this strategy are included in Feiveson's paper). More specifically, it is well known that the level of accuracy of an estimator such as that utilized in LACIE depends on the sample size and is adversely affected by the variability or heterogeneity of the characteristic(s) being measured (in this case, the heterogeneity of wheat density and yield). Apart from increasing the sample size, another means of effectively reducing this heterogeneity is by improving the stratification. During Phase II of LACIE, a methodology was developed to use Landsat imagery and agrophysical data to improve stratification in foreign areas. This method ignored political boundaries and re-stratified along boundaries of areas that are more homogeneous in agricultural density, soil characteristics, and average climatic conditions. The use of this natural sampling strategy domestically in

the Transition Year was intended to provide better applicability of the yardstick region as a quantifier of foreign results. The allocation to support this strategy resulted in a change in the USGP from 601 segments in Phase III to 487 segments in the Transition Year.

Within-country sample allocation.—Sample segments were allocated to the strata/substrata within a country based on weights which, in Phase I, were a function of (1) the agricultural area in the stratum/substratum and (2) the within-stratum/substratum standard deviation of wheat area from segment to segment. Estimates of the former were obtained from the sampling frame generated from USDA land use maps and from Landsat imagery over those areas with sufficient quality coverage. Estimates of the latter were determined by assuming that the per-segment numbers of pixels classified as wheat followed the binomial distribution. Under this assumption, the within-stratum/substratum standard deviation of wheat area from segment to segment is a function of the proportion of wheat in the stratum/substratum, which was obtained from historical data for an epoch year.

In the revised allocations made in the USGP and the U.S.S.R. in Phase III and the Transition Year, the allocation weights were a function of (1) the agricultural area in the substratum/stratum, (2) the within-substratum/stratum standard deviation of small-grains (wheat) area from segment to segment in the USGP and in the U.S.S.R., (3) the classification error variance, (4) the substratum/stratum yield estimate, and (5) the substratum/stratum yield prediction error.

The agricultural area in the substratum/stratum was obtained by the same procedure used in Phase I; however, the availability of higher quality Landsat imagery over more extensive areas permitted the use of a more refined sampling frame for Phase III.

Direct estimation of the within-substratum standard deviation of wheat area from segment to segment in the U.S. yardstick region (and other substratum-level countries) was not possible since most substrata had insufficient segments. Although the approach taken in Phase I of resorting to a binomial distribution assumption and using epoch-year historical wheat data to estimate these variances seemed to work reasonably well, there were indications from the Phase II North Dakota study that an improvement was attainable by taking a slightly different approach. As a part of the North Dakota study, approximately 40 counties were selected

throughout the USGP in an attempt to obtain an improved estimator of the within-substratum standard deviation of small-grains area. The proposed approach consisted of modeling the relation between the segment-to-segment standard deviation of small-grains area at the substratum-level to that of the segment-to-segment standard deviation of agricultural area. The rationale for concentrating on small grains as opposed to wheat included the following.

1. At the time, it was impossible to distinguish wheat from other small grains in Landsat imagery.

2. Because of the predominance of wheat in the areas to be sampled, the belief was that, for allocation purposes, replacing unobservable wheat information with small-grains data would be a reasonable substitution. (Areas outside the range of variability over which the model was developed offered the greatest potential for degradation.)

3. The procedure was repeatable in other substratum-level countries.

Estimates of the other quantities that were input to the allocation procedure are detailed in Feiveson's paper. Depending on the resulting allocation weights, substrata in countries with substratum-level historical data (United States, Canada, Australia) were designated as Group I (high sampling rate), Group II (low sampling rate), and Group III (not sampled). Stratum-level countries (i.e., those having historical data at only one level smaller than the country itself—U.S.S.R., China, Argentina, India, Brazil) had their strata assigned to Group I or Group III only. In the United States, where substrata are counties, the range of variation of the allocation weights was such that to meet the 90/90 criterion in the USGP, it was necessary to allocate anywhere from 0 to 5 segments to a county. In stratum-level countries, several times that number were assigned to a stratum. (For specific details of the allocation procedure, see the paper by Feiveson.)

Area Estimation

Until the Transition Year, the methodology for the direct estimation of wheat area at the segment level did not exist. Instead, a segment-level winter, spring, or total small-grains estimate was made for each aggregable segment.² The area of wheat in each

²Certain criteria had to be met before a given segment was approved for aggregation purposes. These criteria are detailed in the paper by Heydorn et al. entitled "Classification and Mensuration of LACIE Segments."

such segment was then estimated by using a "confusion crop" ratio. In particular, in Phases I and II (and Phase III in the U.S.S.R.), the substratum/stratum-level historical ratio of wheat to small grains, estimated from epoch-year data, was applied to the raw small-grains estimate from the Classification and Mensuration Subsystem (CAMS).

During Phase II, development was initiated on an econometric confusion crop ratio model that was later evaluated and implemented in the USGP early in Phase III. This model provided confusion crop ratio estimates at the CRD level and made use of a considerable amount of near-real-time economic information. The development and evaluation of this model are detailed further in the paper by Umberger et al. entitled "Econometric Models for Predicting Confusion Crop Ratios."

Given the set $\{p_i\}_{i=1}^n$ of aggregable segment-level wheat proportion estimates covering the area of interest, the LACIE wheat area estimate \hat{A} of the given area of interest is expressible in the most simplified form as

$$\hat{A} = \sum w_i \hat{p}_i \quad (2)$$

where \hat{p}_i = the estimate of the proportion of wheat in the i th aggregable sample segment and w_i = the aggregation weight associated with the i th aggregable sample segment. In general, w_i is a function of the epoch-year historical data, the agricultural area in the stratum containing the segment, and whether or not the segment is used in part of a Group III ratio (see "LACIE Sample Design"). With the exception of the Group III ratio case, the form of the estimator in equation (2) is the same as the standard stratified sampling estimator (ref. 1).

Stratum/substratum-level estimates.—Stratified area estimation is performed on a base of stratum or substratum elements (depending on whether the country is a stratum- or a substratum-level country). In the LACIE framework, substrata are designated as Group I (high sampling rate), Group II (low sampling rate), and Group III (not sampled). The designations are made on the basis of a threshold value (detailed in the "LACIE Sample Design" paper). However, the qualitative definitions are as follows:

1. Group I substrata—Intensive wheat-producing areas which are allocated one or more sample segments.

2. Group II substrata—Areas which produce some wheat; one sample segment is allocated with

probabilities proportional to size (PPS) (as determined from the wheat area in an epoch year); thus, some, but not all, Group II substrata receive a segment.

3. Group III substrata—Areas historically having very little wheat; thus, no sample segments are allocated.

The Group I substratum wheat area estimator \hat{A}_I is simply the standard stratified sampling estimator

$$\hat{A}_I = \sum \bar{p}_i a_i \quad (3)$$

where \bar{p}_i = the estimate of the mean proportion of wheat in the i th Group I substratum agricultural area and a_i = the i th Group I substratum agricultural area.

The Group II stratum wheat area estimator \hat{A}_{II} is the standard PPS-type estimator (ref. 1), which has the general form

$$\hat{A}_{II} = \sum_{i=1}^m \hat{A}_i / \Pi_i \quad (4)$$

where \hat{A}_i = the estimate of the wheat area in the i th sampled substratum and Π_i = the probability with which the i th Group II substratum was selected to receive a segment.

The Group III estimator \hat{A}_{III} of the collection of Group III substrata within a given stratum is a ratio estimator having the form

$$\hat{A}_{III} = \left(\frac{\hat{A}_I + \hat{A}_{II}}{W_I + W_{II}} \right) W_{III} \quad (5)$$

where W_I , W_{II} , and W_{III} denote the historical wheat acreages, from an epoch year, over the associated areas that \hat{A}_I , \hat{A}_{II} , and \hat{A}_{III} estimate at some level. The level (i.e., either stratum or zone) at which this ratio is applied is dependent on the availability of aggregable segments in the stratum. Specifically, if there is less than a certain threshold (i.e., to date, a threshold of three has been employed—see the paper by Chhikara and Feiveson entitled "LACIE Large-Area Acreage Estimation") of aggregable segments at the stratum level, this ratio is applied at the zone level; otherwise, it is applied at the stratum level. In the former case, this simply means \hat{A}_I and \hat{A}_{II} are the total estimates of wheat area for all Group I and Group II substrata, respectively, in the parent zone. (Of course, W_I and W_{II} are the associated historical wheat acreages over the corresponding areas.) Otherwise, \hat{A}_I and \hat{A}_{II} are the total estimates of wheat area

for all Group I and Group II substrata, respectively, in the parent stratum. (Again, W_I and W_{II} are the associated historical acreages.) In either case, \hat{A}_{III} is the total wheat area estimator of all Group III substrata (having historical wheat area W_{III} in the epoch year) in the parent stratum.

In stratum-level countries, recall there are Group I and Group III strata only. In this case, the Group I and Group III estimators are of the same form as explained previously; however, the stratified area estimation is performed on the base of stratum elements. Moreover, the Group III ratio estimator is always applied at the zone level with one exception: the situation has occurred (in Phases II and III of LACIE in the U.S.S.R.) wherein one or more zones had little to no acquired segments at the time of aggregation. In this situation, such zones are ratio estimated using a ratio estimator not unlike that discussed previously, whereby stratum wheat area estimates from surrounding zones (having sufficient segment coverage) were employed as the basis for the ratio estimation.

Higher level estimates.—Wheat area estimates above the stratum/substratum base level, such as at the zone, region, or country level, are obtained simply by adding the estimates for the strata included in the area of interest (i.e., zone, region, country, etc.). In mixed wheat areas, the estimation/aggregation procedure is performed separately for each crop type (winter and spring wheat). To obtain total wheat, the separate estimates for each crop type are aggregated.

YIELD AND PRODUCTION ESTIMATION

Yield model development and evaluation in LACIE has been primarily the responsibility of the Center for Climatic and Environmental Assessment (CCEA) (a branch of the National Oceanic and Atmospheric Administration (NOAA)). The yield models used in LACIE have been of an "agromet" type (i.e., have utilized agronomic and meteorological data as inputs to their development and use) and have been polynomial functions of such weather variables as monthly precipitation and potential evapotranspiration. Weather variables are entered as departures from long-term normals. In general, the assumed model form is

$$Y = X\beta + \epsilon \quad (6)$$

where Y is the vector of historical yields over the stratum of interest, X is the matrix of weather data, β is the weather coefficient to be estimated, and ϵ is a vector of random errors. Consequently, the weather coefficients were estimated, by $\hat{\beta}$, in the standard least squares manner, i.e.,

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (7)$$

Yield prediction for the i th yield stratum is then computed using

$$\hat{Y}_i = \hat{\beta} X_i \quad (8)$$

where X_i is the current-year weather observation vector for the i th stratum. Coefficients were estimated separately for each yield stratum because agrophysical differences between regions are such that it is unlikely the same yield model would hold.

Some of the assumptions/shortcomings of this model are

1. The assumption of the adequacy of the model form to predict yield.

2. The historical yields utilized in the model development were those reported by the SRS domestically and the FAS in foreign areas; however, the weather data came from the various meteorological stations and, hence, were not sampled in the same manner as the yield data.

3. Initially in LACIE, some yield strata utilized common weather data which induced unaccounted-for correlations. Prior to the beginning of Phase III, an adjustment was made to the yield strata in order to correct for this situation. The resulting yield strata were no longer necessarily state-level strata; figures 5 and 6 illustrate the new yield strata in the USGP and the U.S.S.R.

An "average" yield estimate is obtained for a yield stratum s by

$$\bar{Y}_s = \frac{P_s}{A_s} \quad (9)$$

where A_s and P_s are the area and production, respectively, for the stratum. The specifics of the yield modeling and evaluation effort are documented in the paper by Strommen et al. entitled "Development of LACIE CCEA-I Weather/Wheat Yield Models" and will not be detailed further herein; the remainder of this discussion considers the estimated yields and their associated prediction error estimates as given.

The LACIE production estimator (depicted in fig. 7) is simply the product of the yield and acreage (ag-

gregated to the "pseudozone" level³ to account for the lack of coincidence between the yield and area strata) estimators. An estimate of the production in a pseudozone is obtained by the product of its area estimate with its yield prediction, and these estimates are aggregated to predict zone and higher level production. The form, therefore, of the production estimator \hat{P} for a given area of interest is as follows

$$\hat{P} = \sum \hat{Y}_i \hat{A}_i \quad (10)$$

where \hat{Y}_i = the yield estimator for the i th pseudozone and \hat{A}_i = the estimator of wheat area for the i th pseudozone. (Specifically, \hat{A}_i is the aggregate

³A "pseudozone" is the area resulting from the intersection of a yield stratum with the area strata in a zone.

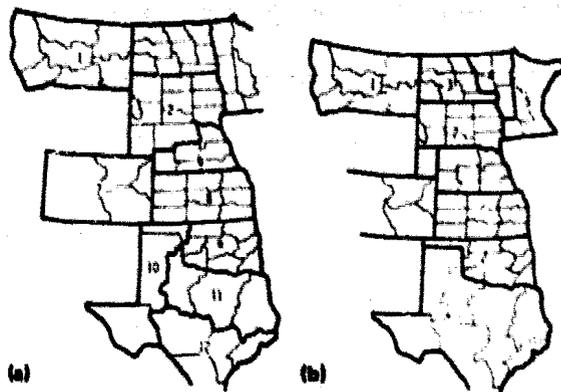


FIGURE 5.—Wheat yield model (weather regression) coverage for the U.S. Great Plains. (a) Winter wheat model boundaries. (b) Spring wheat model boundaries.

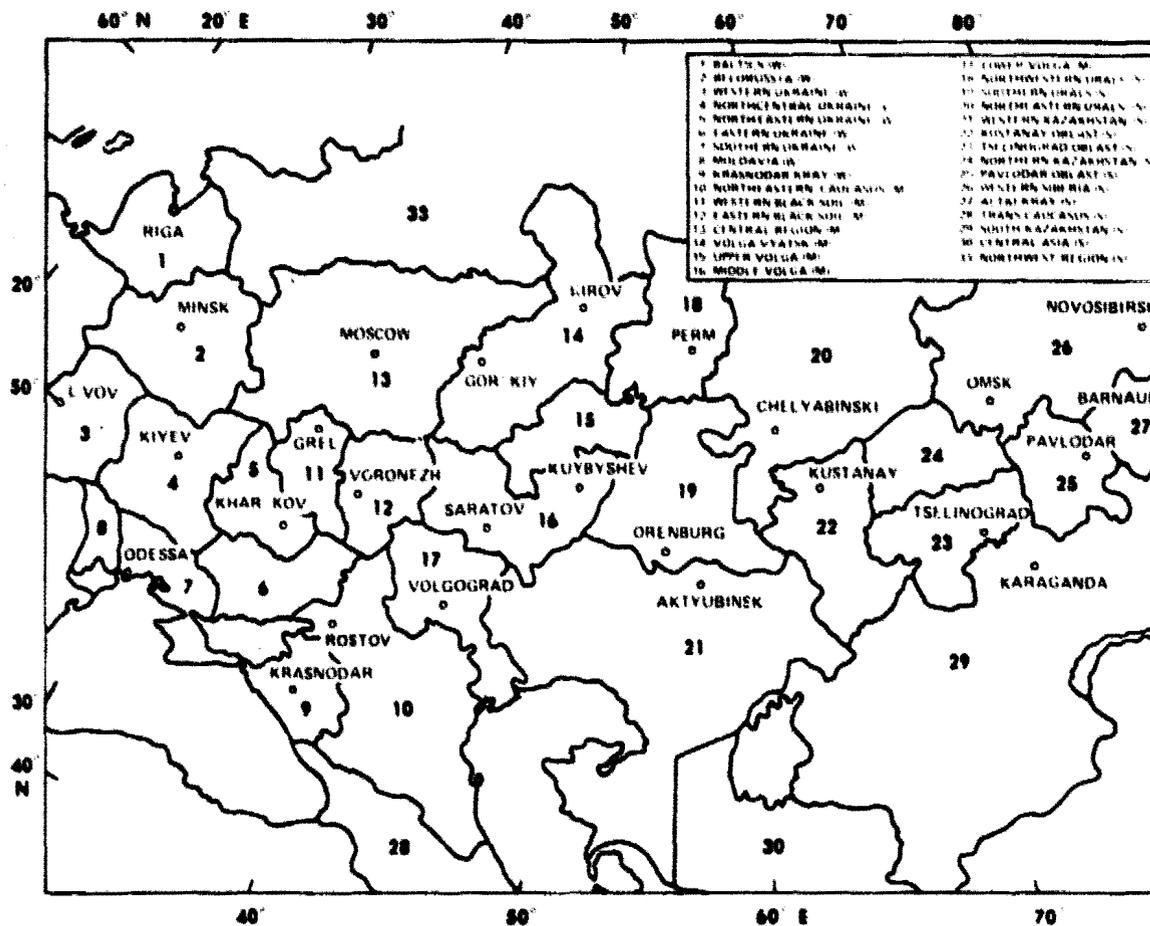


FIGURE 6.—U.S.S.R. crop regions covered by spring and winter wheat yield regression models (regions 31 and 32 are not shown on this map).

of those stratum/substratum wheat area estimates in the parent pseudozone.)

In mixed-wheat areas, this procedure is performed separately for each crop type. The total production is obtained by adding the estimates for each crop type. Assessment of the LACIE area, yield, and production estimators requires estimation of their respective variances. (The accuracy assessment details of LACIE are included in the paper by Houston et al. entitled "Accuracy Assessment: The Statistical Approach to Performance Evaluation in LACIE.") The following section summarizes this procedure.

VARIANCE ESTIMATION

Area Variance Estimation

In stratum-level countries, there are frequently enough segments available to permit the direct computation of the sample variance from segment to segment in each stratum. Recall that the Group III

strata are ratio estimated from various Group I strata estimates. Consequently, the wheat area variance of a Group III stratum is a function of the Group I variances. The wheat area variances for higher levels (e.g., zone, region, or country) are essentially aggregates of the stratum-level variances. The Group III variances are appropriate linear combinations of Group I variances.

For substratum-level countries, the within-substratum wheat area variance estimation procedure is somewhat more complicated since there are many substrata with only one segment. A summary of the procedure follows; the specific details are included in the paper by Chhikara and Feiveson entitled "Large Area Aggregation and Mean Square Prediction Error Estimation for LACIE Yield and Production Estimates."

1. Divide all the substrata within a zone into "collections" based on prior within-substratum variances computed at the time of allocation.

2. Within the i th collection, estimate a variance S_i^2 by regressing substratum estimates against historical data and computing the residual variance.

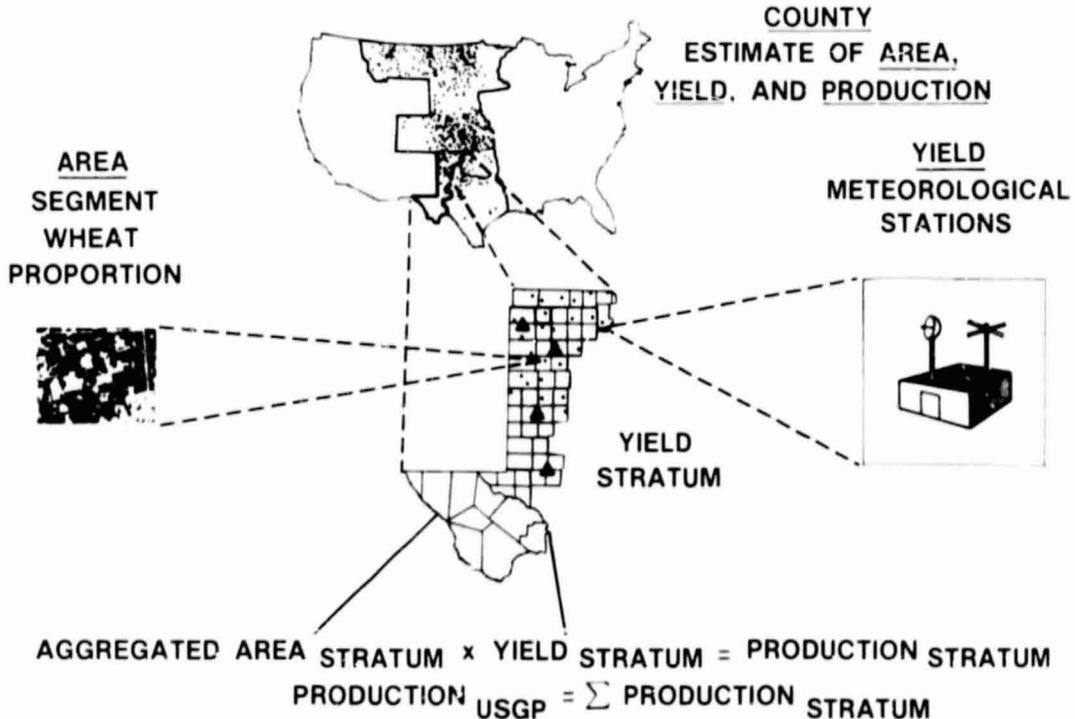


FIGURE 7.—Production estimation from sampling.

3. Assign the value S_r^2 for the within-stratum variance for all substrata in the r th collection.

The model assumed in the regression fit carried out in step 2 is as follows.

$$C_i = \alpha + \beta X_i + \delta_i \quad (11)$$

where C_i = the true wheat acreage in the i th substratum, X_i = the historical wheat acreage in the i th substratum for an epoch year, and δ_i = random fluctuation. Also

$$\hat{C}_i = C_i + \epsilon_i \quad (12)$$

$$\hat{C}_i = \alpha + \beta X_i + \epsilon_i + \delta_i \quad (13)$$

where \hat{C}_i = the LACIE wheat area estimate in the i th substratum and ϵ_i = the sampling plus classification error.

It is also assumed that the variance of δ_i is considerably smaller than that of ϵ_i . (Based on previous observations, this appears to be a good assumption.)

Yield Mean-Square Prediction Error Estimation

Referring to the form of the CCEA yield model estimator (see eqs. (6) to (8)), the yield mean-square prediction error (MSPE) estimator for a given yield stratum is the standard

$$E(\hat{Y}_i - Y_i)^2 = \hat{\sigma}^2 \left[1 + X_i^T (X^T X)^{-1} X_i \right] \quad (14)$$

where $E(\cdot)$ denotes the expectation operator.

The variance of the "average" yield given by equation (9) is obtained using the approximate variance of a ratio between two correlated random variables (detailed in the paper by Chhikara and Feiveson).

Production Variance Estimation

Two basic assumptions are made in arriving at the final form of the LACIE production variance estimator.

1. Segment-level wheat area estimates are mutually independent and unbiased.
2. Yield estimates are unbiased, are mutually independent (at the yield stratum level as opposed to

the pseudozone level), and are independent of the acreage estimates.

Under these assumptions, if the summation in equation (10) is taken over yield strata (as opposed to pseudozones) the variance of the production estimator \hat{P} is expressible as

$$V(\hat{P}) = \sum_i \left[V(\hat{Y}_i) A_i^2 + Y_i^2 V(\hat{A}_i) + V(\hat{A}_i) V(\hat{Y}_i) \right] \quad (15)$$

where $V(\hat{A}_i)$ and $V(\hat{Y}_i)$ denote the variances of the acreage and yield estimators, respectively, of the i th yield stratum, and A_i and Y_i denote the estimator means for acreage and yield, respectively, for the i th yield stratum. The LACIE production variance estimator, then, is approximately that obtained by replacing the parameters on the right side of equation (15) with their respective estimates. (An additional adjustment of changing the sign preceding the term $V(\hat{A}_i) V(\hat{Y}_i)$ to a minus results in an unbiased variance estimator of \hat{P} .) Further details are included in the paper by Chhikara and Feiveson.

SPECIAL PROBLEMS ENCOUNTERED IN LACIE SAMPLING AND AGGREGATION

In an experiment having as many constraints and complexities as LACIE had, there is the expectation from the outset of encountering many problems. In LACIE, an Accuracy Assessment Subsystem was created not only to closely monitor the LACIE estimates and assess their accuracy but also to expedite the surfacing of various problems and their subsequent resolutions in order that system impacts could be held to a minimum (see the paper by Houston et al. for the specifics of the accuracy assessment functions in LACIE). During Phases I, II, and III of LACIE, a number of sampling and aggregation problems surfaced. For example, by the end of Phase I, it was clear that the design as it existed at that time had certain disadvantages such as

1. The domestic approach did not appear to be entirely adequate as an indicator of the expected performance levels in foreign regions; i.e., the U.S. county is a substratum of much smaller size than the areas for which data were available in most foreign regions.
2. Considerable effort was required to establish the degree to which all assumptions were sufficiently satisfied; moreover, extensive data were required to

evaluate the precision of area, yield, and production estimates.

Other problems related to sampling and aggregation that surfaced during Phases I, II, or III are discussed in the following paragraphs.

Crop Type Estimation in Mixed-Wheat Areas

In areas having significant amounts of both spring and winter wheat, there was the question of whether analysts should provide estimates of both spring and winter wheat from every segment or predesignate each segment as either a spring or a winter wheat segment and provide only one crop type estimate from each segment. In Phases I and II, the latter procedure was employed; predesignation of segments was initially (in Phase I) performed as follows. For each stratum (CRD in the United States, oblast in the U.S.S.R.), the proportion of allocated segments predesignated as spring (winter) was the same as the proportion of spring (winter) wheat area in the stratum in an epoch year; segment labels at this proportion were levied randomly. The results from this approach appeared favorable.

A decision was made at the end of Phase II to implement the first procedure; i.e., the one in which analysts are required to pass both a spring and a winter wheat estimate from each segment in a mixed-wheat area. Unfortunately, there were several mixed-wheat areas in which segments had almost no spring wheat or almost no winter wheat, thus forcing the analyst to look for "a needle in a haystack." Indications were that more care should be exercised in designating which strata should be "mixed."

During Phase III, the Accuracy Assessment group conducted a study which indicated that a reasonable guideline is to designate a stratum as mixed if neither crop type's presence (historically) is below approximately 20 percent of the total wheat area (i.e., winter plus spring). To date, this procedure appears to be working satisfactorily; however, this issue requires further investigation.

Nonresponse Because of Cloud Cover

Because of atmospheric effects such as haze and cloud cover, LACIE does not get coverage over every segment on every pass. To counteract this problem,

the aggregation logic uses a ratio estimator to provide estimates of nonresponse areas in the same manner as previously discussed for Group III areas. The magnitude of the bias induced by nonresponse has been monitored in the Accuracy Assessment program. Indications are that the loss of acquisitions from cloud cover was a problem in Phase I; however, tests conducted to date indicate that error arising from this loss is probably random with no significant bias being introduced. In foreign areas where the strata are considerably larger than counties, the bias induced by nonresponse is believed to be somewhat more pronounced (particularly at the stratum level) but not to such a degree as to warrant alarm, although this conclusion has not been rigorously verified.

Classification and Yield Prediction Bias

As indicated in the first section of this paper, an assessment of the accuracy of the LACIE production estimator requires knowledge of its variance and bias. The variance is computed as has been indicated. Although the sampling scheme employed in LACIE is oriented toward having no sampling bias (except that induced by cloud cover and ratio estimation), Accuracy Assessment still has the chore of estimating the bias induced by classification and yield prediction. As a result, "wall-to-wall" ground data inventories have been taken over a subsample of the segments in the yardstick area for use in estimating bias and for general diagnostic purposes. The approach has been to estimate the bias from ground data taken from a random subsample of approximately one-third of the allocated segments. (The resulting segments are referred to as "blind sites.") However, until a larger subsample of blind sites is available (not a cost-effective approach) or until a more clever procedure for bias estimation is found, reliable estimates of the magnitude of the bias will continue to be lacking.

Instability of Group III Ratios

The information in the section on area estimation indicates that the Group III estimator \hat{A}_{III} for a given Group III stratum/substratum has the general form

$$\hat{A}_{III} = R\hat{A} \quad (16)$$

where \hat{A} is an estimator of a nearby area having sufficient data for a direct estimate and R is the ratio of the historical wheat area in the Group III area to the historical wheat area in the area estimated by \hat{A} . Consequently, the variance of \hat{A}_{III} is

$$V(\hat{A}_{III}) = R^2 V(\hat{A}) \quad (17)$$

Examination of the graph of $V(\hat{A}_{III})$ (fig. 8) for various values of $V(\hat{A})$ suggests that one should expect instability in the Group III ratio estimator over areas wherein R and/or $V(\hat{A})$ are large. This situation was not rigorously investigated until Phase III of LACIE. At that time, Accuracy Assessment conducted a study which led to the conclusion that areas to be used as a base for ratio estimation of a particular stratum should be selected such that the magnitude of R is between 1 and 1.5. Although this approach seems to perform satisfactorily, further work is warranted.

Poor Estimation of Group II "Sizes"

Shortly after completion of the revised allocation in the USGP and before the first Phase III aggregation, the 1974 U.S. Agricultural Census data became available and were used in place of the 1969 census data to update the Crop Assessment Subsystem (CAS) aggregation data base. No historical data used in performing the revised allocation were retained in the data base. As a result, sufficient disagreement existed in a number of the Group II substrata between the allocation and the aggregation "sizes" (recall that a PPS allocation was performed on the collection of Group II substrata—as a result, a Group II bias arises unless the Group II allocation and aggregation "sizes" (weights) agree) to induce considerable bias in the first two aggregations in Phase III—particularly in the state of South Dakota.

Further investigation determined that the Phase III allocation of segments (based on small grains—see "LACIE Sample Design" for details) led to the allocation of segments to some counties which historically contain little or no wheat. Specifically, it was found that 32 counties designated as Group III counties should probably have been Group I or II counties, whereas 12 Group I and II counties should have been Group III counties. The main effect of the first error was to increase bias and variance through the use of Group III ratios. The second error was corrected by redesignating the appropriate Group I and Group II as Group III for the remainder of Phase III.

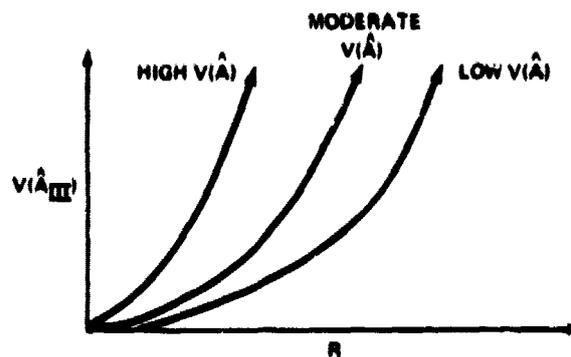


FIGURE 8.—Relationship between $V(\hat{A}_{III})$ and R for low, moderate, and high values of $V(\hat{A})$.

Thresholding/Screening of Outliers

The accuracy of segment-level estimates varies considerably with the segment as a result of such factors as crop development stage at the time of analysis (early season, midseason, harvest), acquisition pattern (timing relative to crop biostages and number of missing acquisitions), analyst effect, and segment effect (e.g., field size). As a result, various segment estimates received by CAS are sufficiently inaccurate as to not warrant inclusion in the aggregation. During Phase III of LACIE, a procedure for thresholding early-season estimates was developed to eliminate segments the estimates for which were made much earlier in the year and which, consequently, were not reliable indicators of the actual wheat area. In addition, a screening procedure was developed to detect segment outliers that may be due to other reasons described in more detail in the following paragraphs.

The LACIE began data processing for crop year 1977 winter wheat when the normal crop calendar reached 2.0 on the Robertson growth scale. These early-season classifications by CAMS resulted in segment proportion estimates made before all wheat in a segment was detected by LACIE. These early-season estimates remain in the data base and are aggregated until replaced by a later acquisition. If these suspect estimates are eliminated from the aggregation as more data become available, the subsequent aggregated area estimates should more accurately reflect the actual wheat area (provided, of course, enough segments remain to permit an aggregation with tolerable variance). During Phase III, an empirical approach was developed to arrive objectively at a threshold date for each state. Although the thresholding of the early-season acquisitions has the

effect of reducing the early-season bias, the loss of segments also increases the variance. Consequently, the empirical approach attempted to select the optimal biostage for a threshold date such that the combined contributions of variance and bias yielded a minimal mean-squared error of the acreage estimate. Figure 9 depicts this approach.

At about the same time in Phase III, a screening procedure was developed for the USGP to flag segments having highly questionable estimates as a result of such influences as bad classification or application of an erroneous small-grains-to-wheat ratio. It was clear from the problems experienced in South Dakota, for example, that the aggregation logic is very sensitive to acreage estimation errors for counties located in low to marginal wheat-growing areas. To achieve an accurate area estimate, it is desirable to screen segments whose estimates are probable outliers. Following is the procedure developed and applied in Phase III in the USGP.

1. Counties were grouped into four categories: low, marginal, medium, and high wheat density counties.

2. For each aggregable segment, the log of the ratio of the CAMS estimate of wheat proportion in the segment to that of the county containing the segment (estimated from the epoch-year data) is computed.

3. The distribution of the data generated in step 2 is investigated within each category and confidence intervals are constructed about the category means.

4. Those segments in a category that lie outside the category confidence interval computed in step 3 are declared outliers.

The thresholding and screening procedures were applied in the USGP in Phase III with quite satisfactory results.

Additional problems that have not received the degree of attention demanded by those discussed previously and are thus awaiting resolution or further refinement include the following.

1. An investigation is needed to assess the degree of dependency between acreage and yield model errors followed by the incorporation of results into the production prediction mean-squared error estimator. (Currently, the acreage and yield errors are assumed to be independent.)

2. Crop acreage variances are not computable for strata having only one aggregable segment. Currently, the variances computed at the time of allocation are used for such strata; however, in some cases,

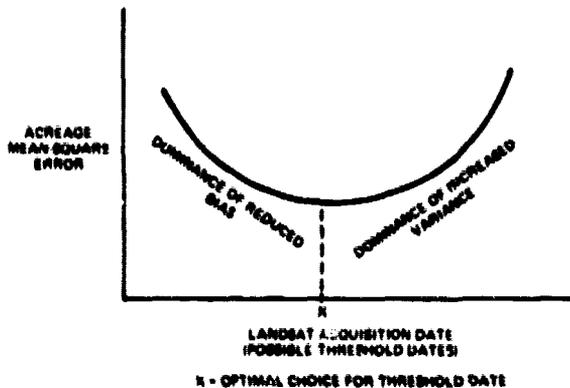


FIGURE 9.—Relationship between the acreage mean-square error and candidate threshold dates.

these estimates are regression estimates. This situation warrants further investigation for potential improvements.

3. There is a problem in estimating crop acreages over nonresponse areas; i.e., areas for which little or no satellite data are available and historical data are either very poor or nonexistent. Currently, the approach taken in LACIE is to use a ratio estimator (the Group III ratio estimation procedure) that estimates the trend, relative to an epoch year using nearby, satellite-acquired data as a base in the ratio estimation. Further research is needed to improve this situation, particularly in foreign areas where little or no historical data are available.

4. Related efforts to date in LACIE have produced a single-crop sampling strategy and the associated aggregation and variance estimation formulations. There is now a need to generalize to the case of multiple crops (work is currently being initiated to this end). In particular, there is a need for research to develop and test a multicrop production estimation procedure that incorporates various constraints and interrelations between crops including the correlations between separate crop type acreage estimates, the correlation between acreage and yield, and inherent constraints such as the inequality constraint that disallows total crop area to exceed the total agricultural area.

5. Results from LACIE indicate that, apart from increasing the sample size, an improved stratification is another means of improving precision of crop production estimators. Recently, the methodology has been developed for interpreting and synthesizing satellite data (particularly Landsat), soils information, and meteorological data to enable an improved

stratification. The resulting strata are referred to as "agrophysical units" (geographic areas having definable/comparable agronomic and physical parameters which reflect a certain range of agricultural use and management). There is a need to investigate further the relationship between agrophysical unit development and both conventional soils mapping and soils mapping using satellite data to develop methods of improving the utility of agrophysical units in large area crop production estimation. Emphasis should be placed on the development and testing of techniques applicable in foreign areas where suitable soils maps and other pertinent data are generally not available.

SUMMARY/CONCLUSIONS

It was essential that the sampling strategy used in LACIE be good enough to support a cost-effective system. This has been particularly evidenced by the following.

1. An approximately 2-percent sampling error has been achieved in LACIE by sampling only approximately 2 percent of the sampling frame.
2. The sample design in the yardstick region for which historical data were available down to a substratum level to support missing data resulting from cloud cover provided the most accurate estimate possible.
3. The implemented strategy provided data of sufficient quantity and quality to support required

performance levels and also to satisfy the existing constraints.

4. The allocation scheme appeared to provide the most efficient usage of the available data and gave efficient segment coverage of major producing areas and thus improved the probability of an accurate estimate.

ACKNOWLEDGMENT

The LACIE is particularly indebted to Dr. H. O. Hartley, Professor of Statistics at Texas A & M University, for his many contributions to LACIE, particularly in the area of sampling and aggregation. Dr. Hartley has been instrumental in expediting the solutions to many key problems over the 3 years of LACIE.

REFERENCES

1. Cochran, William G.: Sampling Techniques. Third ed., John Wiley & Sons, 1977.
2. Benedict, M. R.: Development of Agricultural Statistics in the Bureau of the Census. J. Farm Economics, vol. 21, Nov. 1939, pp. 735-760.
3. Ebhling, Walter H.: Why the Government Entered the Field of Crop Reporting and Forecasting. J. Farm Economics, vol. 21, Nov. 1939, pp. 718-734.

Econometric Models for Predicting Confusion Crop Ratios

D. E. Umberger,^a M. H. Proctor,^a J. E. Clark,^b L. M. Eisgruber,^b and C. B. Braschler^c

INTRODUCTION

Using econometric models to predict annual adjustments in the ratios of crop acreages is unique in economic literature. The more common approach is to develop models which attempt to directly predict annual adjustments in the planted (or harvested) acreage of an individual crop within an area. The need for the LACIE was partly derived from the past lack of success in devising ways (including econometric models) to predict the acreage of wheat and other crops in foreign areas. However, the impetus for developing ratio models arose from a practical problem in the LACIE—the inability to accurately classify spring wheat signatures in certain areas where “confusion” crops are grown. The question the ratio modeling effort attempted to answer was “Assuming that accurate and reliable confusion crop acreage estimates for an area were made available from remote-sensing sources, could ratio models be developed to provide accurate and reliable estimates of the acreages of individual crops?”

The study which is the subject of this paper was initiated by LACIE management to determine the feasibility of developing econometric models capable of improving the predictive characteristics of historical ratios to a level that would support established LACIE accuracy goals. A plan was developed outlining LACIE requirements for ratios, area, coverage, prediction timeliness, modeling approach, data needs, and resource requirements for developing

operational models on a time schedule consistent with providing confusion crop ratio estimates for Phase III. The task, which was undertaken by scientists from the University of Missouri and Oregon State University and by LACIE personnel at Columbia, Missouri, was begun in January 1977 and completed in July 1977. The results for the United States and Canada are reported separately (refs. 1 and 2).

The purpose of this study was to test the feasibility of using econometric methods to predict annual adjustments in the relative acreages of wheat and “confusion” crops in the States of North Dakota, South Dakota, Montana, and Minnesota and in the Canadian Province of Saskatchewan. “Confusion” crops were defined as those crops having similar crop calendars, including such spring-planted crops as spring wheat, barley, oats, and flax and such fall-planted crops as winter wheat and rye.

Confusion Crop Problems

Initially in the LACIE project, it was planned to use area samples of Landsat imagery to estimate wheat acreage for a test region (ref. 3). Regional wheat acreage estimates were to be derived from Landsat imagery by estimating the proportion of wheat acreage to total agricultural land in each segment and aggregating these estimates to the appropriate regional level. During LACIE Phases I and II (1975 and 1976 crops), however, operational difficulties were encountered in providing a reliable wheat-proportion estimate for sample segments in areas where confusion crops were grown. During Phase I, it became apparent that LACIE analysts, using existing classification procedures, were unable

^aU.S. Department of Agriculture, Columbia, Missouri.

^bOregon State University, Corvallis, Oregon.

^cUniversity of Missouri, Columbia, Missouri.

to reliably separate wheat signatures from the signatures of certain other crops in certain regions (ref. 3, pp. 1-11 and 2-60). This problem was particularly troublesome in those areas where several crops with similar stages of plant development are found. For example, potential confusion was suspected among winter wheat, spring wheat, rye, barley, oats, and flax signatures in the spring or mixed wheat states (Minnesota, Montana, North Dakota, and South Dakota) of the U.S. Great Plains. Rye, barley, and oats were the major crops being confused with spring wheat in the Saskatchewan Province of Canada. Similar classification-related problems were also suspected in the spring and mixed wheat areas of the U.S.S.R.

During Phase II, the accepted procedure was for the LACIE analyst to identify wheat for each sample segment where this could be done with a high degree of accuracy. Where the existence of confusion crops in a segment made the classification of wheat impossible or of questionable accuracy, the analyst would instead provide a proportion estimate of either winter or spring confusion crop acreage. Where there was possible confusion between both winter and spring small grains, a proportion estimate for total small grains was provided for the sample segment. Thus, the total set of classification results for use as input in acreage aggregations was one of the following: winter wheat (WW), spring wheat (SW), winter small grains (WG), spring small grains (SG), or total small grains (GR).

In cases where a WG, SG, or GR ratio was estimated for a sample segment, some procedure was required to obtain an estimate of the desired WW or SW ratio prior to aggregating the sample segment to the strata level (crop reporting district (CRD) in the United States). In Phase II, the procedure was to apply to each segment (with a WG, SG, or GR ratio) a substrata level (county in the United States) historical ratio of the proportion of winter wheat or spring wheat. The historical ratio was WW/WG or WW/GR for winter wheat and SW/SG or SW/GR for spring wheat. In general, historical ratios for estimating the proportion of winter wheat or spring wheat at the substrata level were based on the most recent single year of historical data available. County-level Statistical Reporting Service (SRS) data for 1975 were used to develop ratios in Minnesota, North Dakota, and South Dakota; 1973 data were used for Montana since more recent estimates were not available. In Canada, 1971 census data were used.

Because historical ratios proved to be a source of

error in estimating acreages, more accurate estimates of current-year confusion crop ratios were desired. Econometric models were considered as a possible method to reduce the error in historical ratios until classification techniques and procedures which could use Landsat imagery to reliably differentiate between wheat and other small grains were developed, tested, and implemented.

Objective

The overall objective of the ratio modeling effort was to develop, test, evaluate, and recommend for implementation a method of projecting confusion crop ratios for the 1977 crop year to a level of accuracy that would support the attainment of stated LACIE performance criteria.

LACIE CONFUSION CROP RATIO REQUIREMENTS

The nature of the LACIE operational system and aggregation procedures led to several unique requirements for econometric modeling. The requirements of the LACIE aggregation software often led to compromises with the preferred methodology for ratio model development. These compromises are addressed in the following sections.

Geographic Regions Considered

Potential confusion crops and required ratios varied among regions. Based on procedures used in Phase III, ratioing methodology was required for all active LACIE countries (the United States, Canada, and the U.S.S.R.). However, because the U.S.S.R. is a planned economy and the set of variables that might explain significant changes in crop ratios differs from that of the other countries, the initial ratio modeling effort was concentrated on specified U.S. and Canadian areas. Resource limitations and data availability were also factors contributing to this decision.

Geographic areas for which ratio models were developed to support Phase III estimates were the States of Minnesota, Montana, North Dakota, and South Dakota (fig. 1) and the Canadian Province of Saskatchewan (fig. 2).

Size of Geographic Area

Ideally, the LACIE needed a specific estimate of confusion crop ratios for each sample segment in a region; however, data did not exist to develop models that were capable of providing estimates at the segment level. The next larger area size was a county or substrata. However, the questionable accuracy of county acreage data and the large number of models needed presented practical problems for developing ratio models at the county level.

As a compromise between LACIE requirements and practical modeling constraints, confusion crop

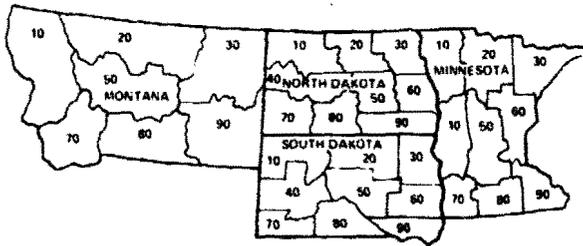


FIGURE 1.—Map of U.S. geographic areas showing state crop reporting districts modeled.

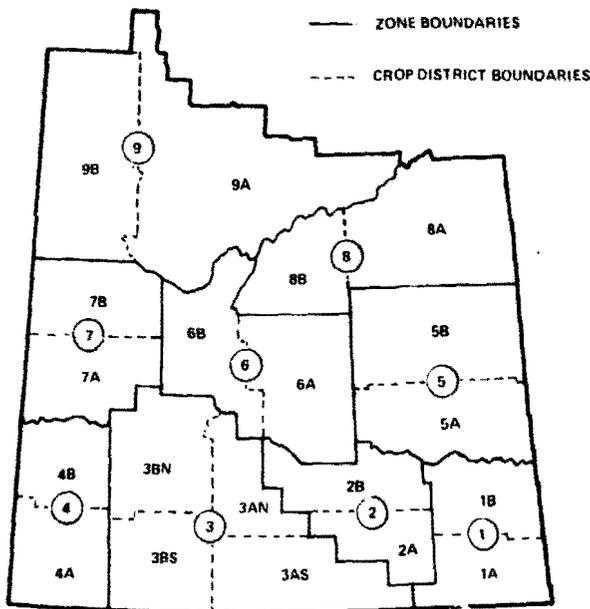


FIGURE 2.—Map of Saskatchewan showing geographic areas (zones and crop districts) modeled.

models were developed for LACIE strata; that is, for CRD's in the United States and for crop districts (CD's) in Saskatchewan. As a guide to independent variable selection, state-level models were also developed in the United States and a province model was developed for Saskatchewan.

Although preliminary state-level models were tested for the United States, all final models were estimated at the CRD level (fig. 1). Canadian models were reported for 20 CD's, 9 zones (zones being composed of 2 to 4 CD's), and the province (fig. 2).

Specification of Confusion Crop Ratios

Confusion crops were defined as those crops which could not be reliably differentiated from winter wheat or spring wheat using Landsat imagery. Although LACIE analysts were uncertain about the precise identity of confusion crops in all areas, they agreed that wheat, rye, barley, oats, and flax would provide a useful basis for testing the capability of econometric models in Phase III. Definition of specific needs in terms of planted versus harvested ratios was not clear at the beginning of the task. Where possible, models were developed for predicting both harvested and planted ratios. For Minnesota and Saskatchewan, where historical data on planted acreage were not available at the CRD level, only harvested acreage ratio models were developed.

The required product for input into the LACIE aggregation system was a ratio of winter wheat or spring wheat acreage to the appropriate small grains acreage classes. These classes were (1) winter grains (including winter wheat and winter rye); (2) spring small grains (including spring wheat, barley, oats, and flax); and (3) total grains (including both spring and winter small grains). In Minnesota, North Dakota, and Saskatchewan, it was necessary to project only two ratios: SW/SG and SW/GR. These two spring wheat ratios also were required for the mixed wheat states of Montana and South Dakota in those CRD's where both winter and spring grains were grown. In addition, two winter grain ratios (WW/WG and WW/GR) were required for South Dakota and one winter grain ratio (WW/GR) was required for Montana. Because no CRD-level acreage estimates were available for winter rye in Montana, no WW/WG ratio was estimated in that state. Thus, 168 CRD and 20 CD confusion crop ratio models were required (table I).

TABLE I.—Number of Confusion Crop Ratio Models Developed for Crop Reporting Districts in the United States and for Crop Districts in Canada

State or province	Number of CRD or CD models								Total
	Planted acreage				Harvested acreage				
	SW/SG	SW/GR	WW/WG	WW/GR	SW/SG	SW/GR	WW/WG	WW/GR	
Minnesota	0	0	0	0	9	9	0	0	18
Montana	7	7	7	0	7	7	7	0	42
North Dakota	9	9	0	0	9	9	0	0	36
South Dakota	9	9	9	9	9	9	9	9	72
Saskatchewan	0	0	0	0	20	0	0	0	20
Total	25	25	16	9	54	34	16	9	188

TECHNICAL BACKGROUND

This section analyzes the historical adjustments in confusion crop acreage and reviews past efforts to develop econometric models for predicting acreage adjustments of individual crops.

Analysis of Historical Confusion Crop Data

The SW/SG ratios for a CRD or CD are a function of spring wheat acreage and acreages of barley, oats, and flax in the area. Changes in the acreages of any

of these crops alter the SW/SG ratios. As shown in tables II to VI, annual changes in the SW/SG ratios historically have been large. For example, in Minnesota, the size of the annual changes in the SW/SG ratios historically ranged from -53.3 percent (in CRD 90) to 353.3 percent (in CRD 20). Changes in the SW/SG ratio from 1975 to 1976 ranged from -10.0 percent in CRD 30 to 150.0 percent in CRD 70. Although the range of year-to-year adjustments in the SW/SG ratios was generally larger in the Minnesota CRD's than in other states or in Saskatchewan, annual changes greater than ± 10 percent were nevertheless common in all CRD's and CD's. These

TABLE II.—Minnesota: Harvested Acreage of Spring Wheat and Spring Grains

[Calculated from ref. 4]

Geographic area	1976 data (range for 1963-76)			
	SW area, percent of state total	SG area, percent of state total	SW/SG area ratio	Annual change in SW/SG area, percent
CRD 10	49.2 (47.0 to 68.8)	42.3 (35.5 to 47.3)	0.640 (0.160 to 0.640)	10.3 (-38.5 to 62.5)
CRD 20	.4 (0.2 to 0.7)	1.0 (0.7 to 1.3)	.250 (0.030 to 0.300)	-16.6 (-50.0 to 366.7)
CRD 30	0 (0)	.1 (0.0 to 0.2)	.090 (0.030 to 0.110)	-10.0 (-50.0 to 133.3)
CRD 40	30.8 (23.8 to 34.7)	28.0 (25.3 to 31.9)	.600 (0.140 to 0.600)	15.4 (-16.7 to 60.0)
CRD 50	7.3 (1.6 to 8.3)	10.4 (9.3 to 11.9)	.390 (0.030 to 0.390)	25.8 (-62.5 to 200.0)
CRD 60	.3 (0.1 to 0.3)	1.5 (1.5 to 2.4)	.090 (0.010 to 0.090)	12.5 (-50.0 to 200.0)
CRD 70	5.4 (0.5 to 5.4)	7.4 (5.1 to 9.4)	.400 (0.010 to 0.400)	150.0 (-66.7 to 150.0)
CRD 80	4.8 (0.6 to 12.1)	4.6 (3.0 to 7.2)	.580 (0.060 to 0.580)	56.8 (-65.0 to 266.7)
CRD 90	1.8 (0.5 to 4.0)	4.6 (4.6 to 7.2)	.210 (0.030 to 0.210)	31.2 (-62.5 to 180.0)

changes show the problem of using last year's ratio as a predictor for the current year.

The size of the SW/SG ratio for a CRD or CD is not necessarily related to the percentage of the state spring wheat acreages accounted for by that CRD or CD, but the largest SW/SG ratios tend to be in those CRD's or CD's with the largest spring wheat acreages (tables II to VI). The CRD's with the largest acreages of spring wheat also tended to have the most stable SW/SG ratios. A small error in estimating the SW/SG ratio for a CRD with a major share of the state's wheat acreage can have a large impact on

the error in the spring wheat acreage estimate for that state. Thus, in developing ratio models, it was important that efforts be focused on those CRD's with the large spring wheat acreages. In 1976, most of Minnesota's spring wheat and spring grain acreages were concentrated in CRD's 10 and 40 (table II). CRD 10 alone accounted for 47.0 to 68.8 percent of the state's spring wheat acreage (excluding durum) during the 1963-76 period. In Montana, CRD's 20 and 30 together typically account for about 90 percent of the spring wheat acreage with the latter CRD alone accounting for more than one-half of the total

TABLE III.—Montana: Harvested Acreage of Spring Wheat and Spring Grains

[Calculated from ref. 5]

Geographic area	1976 data (range for 1963-76)			
	SW area, percent of state total	SG area, percent of state total	SW/SG area ratio	Annual change in SW/SG area, percent
CRD 10	0.6 (0.4 to 1.0)	2.3 (1.6 to 2.4)	0.170 (0.114 to 0.243)	-15.8 (-36.6 to 66.5)
CRD 20	30.8 (20.6 to 39.6)	35.3 (32.5 to 44.2)	.543 (0.229 to 0.543)	25.4 (-50.9 to 71.4)
CRD 30	57.8 (48.8 to 64.8)	43.3 (34.4 to 47.4)	.830 (0.654 to 0.830)	5.7 (-14.6 to 23.1)
CRD 50	4.5 (1.9 to 5.3)	8.3 (5.8 to 9.4)	.336 (0.119 to 0.336)	83.6 (-55.0 to 83.6)
CRD 70	.8 (0.5 to 1.0)	2.3 (1.7 to 2.3)	.222 (0.083 to 0.287)	69.5 (-47.9 to 112.9)
CRD 80	.9 (0.5 to 1.3)	3.0 (2.5 to 4.5)	.184 (0.078 to 0.184)	53.3 (-48.7 to 61.4)
CRD 90	4.7 (3.2 to 5.7)	5.6 (4.0 to 6.1)	.526 (0.369 to 0.594)	.4 (-28.5 to 60.9)

TABLE IV.—North Dakota: Harvested Acreage of Spring Wheat and Spring Grains

[Calculated from ref. 6]

Geographic area	1976 data (range for 1963-76)			
	SW area, percent of state total	SG area, percent of state total	SW/SG area ratio	Annual change in SW/SG area, percent
CRD 10	16.4 (15.3 to 18.3)	13.0 (12.5 to 13.6)	0.877 (0.579 to 0.877)	0.8 (-8.9 to 20.1)
CRD 20	11.6 (10.8 to 12.5)	11.1 (10.7 to 11.7)	.728 (0.502 to 0.728)	5.4 (-13.4 to 20.2)
CRD 30	11.3 (11.3 to 19.6)	19.3 (18.3 to 20.1)	.708 (0.419 to 0.708)	3.2 (-8.9 to 23.5)
CRD 40	7.5 (7.5 to 9.6)	6.4 (6.4 to 7.9)	.818 (0.576 to 0.818)	4.6 (-11.9 to 12.7)
CRD 50	12.5 (9.3 to 12.5)	11.2 (10.6 to 11.7)	.775 (0.401 to 0.775)	9.3 (-9.8 to 35.7)
CRD 60	13.2 (8.5 to 13.2)	13.5 (11.7 to 13.5)	.679 (0.293 to 0.679)	6.1 (-10.5 to 36.6)
CRD 70	8.1 (6.6 to 10.8)	6.7 (6.5 to 7.3)	.837 (0.584 to 0.837)	3.0 (-18.8 to 33.1)
CRD 80	7.1 (5.8 to 7.8)	6.6 (6.2 to 7.5)	.747 (0.471 to 0.747)	13.0 (-12.3 to 27.9)
CRD 90	12.3 (7.0 to 12.3)	12.2 (11.1 to 12.5)	.698 (0.322 to 0.698)	15.6 (-11.9 to 45.1)

TABLE V.—South Dakota: Harvested Acreage of Spring Wheat and Spring Grains

[Calculated from ref. 7]

Geographic area	1976 data (range for 1963-76)			
	SW area, percent of state total	SG area, percent of state total	SW/SG area ratio	Annual change in SW/SG area, percent
CRD 10	17.1 (13.1 to 18.3)	10.8 (7.7 to 10.8)	0.795 (0.535 to 0.795)	19.4 (-22.3 to 23.9)
CRD 20	38.2 (38.2 to 49.0)	24.6 (24.6 to 29.1)	.780 (0.460 to 0.780)	26.0 (-12.1 to 26.0)
CRD 30	26.7 (14.4 to 26.7)	23.3 (22.0 to 25.1)	.575 (0.178 to 0.575)	44.5 (-12.5 to 53.7)
CRD 40	.8 (0.8 to 2.2)	1.0 (1.0 to 2.4)	.384 (0.159 to 0.384)	29.3 (-29.8 to 67.7)
CRD 50	6.2 (6.2 to 14.7)	4.4 (4.4 to 12.1)	.708 (0.294 to 0.708)	43.3 (-20.6 to 26.4)
CRD 60	5.0 (1.6 to 5.0)	18.0 (13.2 to 18.0)	.140 (0.038 to 0.140)	29.6 (-24.6 to 53.5)
CRD 70	.5 (0.1 to 0.5)	.7 (0.6 to 1.0)	.403 (0.038 to 0.403)	38.5 (-67.0 to 112.5)
CRD 80	2.5 (1.0 to 2.8)	2.5 (2.5 to 4.0)	.494 (0.082 to 0.494)	80.3 (-34.5 to 108.4)
CRD 90	3.0 (1.2 to 3.0)	14.6 (8.8 to 14.6)	.102 (0.042 to 0.102)	30.8 (-28.9 to 57.8)

TABLE VI.—Saskatchewan: Harvested Acreage of Spring Wheat and Spring Grains

[Calculated from ref. 2]

Geographic area	1976 data (range for 1963-76)			
	SW area, percent of state total	SG area, percent of state total	SW/SG area ratio	Annual change in SW/SG area, percent
CD 1A	4.4 (4.4 to 5.4)	4.5 (4.4 to 5.1)	0.766 (0.622 to 0.865)	8.0 (-20.2 to 8.0)
CD 1B	3.3 (2.2 to 3.6)	3.8 (3.4 to 4.0)	.677 (0.413 to 0.830)	8.1 (-40.0 to 35.6)
CD 2A	4.7 (4.1 to 5.4)	4.0 (3.8 to 4.5)	.909 (0.713 to 0.911)	3.8 (-17.5 to 6.3)
CD 2B	6.7 (5.3 to 6.9)	5.8 (5.0 to 6.4)	.904 (0.646 to 0.932)	5.5 (-23.3 to 9.4)
CD 3AS	6.7 (6.1 to 10.2)	5.6 (5.6 to 7.2)	.935 (0.793 to 0.935)	2.6 (-8.9 to 6.5)
CD 3AN	3.4 (3.4 to 4.6)	3.0 (3.0 to 3.6)	.891 (0.762 to 0.939)	3.8 (-13.4 to 7.9)
CD 3BS	4.6 (4.4 to 8.2)	3.8 (3.8 to 5.9)	.932 (0.792 to 0.932)	1.7 (-7.5 to 10.2)
CD 3BN	6.4 (6.1 to 9.0)	5.5 (5.5 to 7.1)	.913 (0.731 to 0.944)	2.0 (-11.3 to 11.5)
CD 4A	2.4 (2.1 to 4.2)	2.1 (2.1 to 3.3)	.885 (0.731 to 0.885)	3.8 (-8.7 to 8.3)
CD 4B	3.3 (3.5 to 5.9)	2.8 (2.8 to 4.1)	.930 (0.797 to 0.960)	1.0 (-7.6 to 5.9)
CD 5A	6.5 (5.5 to 6.7)	6.9 (6.0 to 7.2)	.745 (0.526 to 0.828)	9.6 (-29.6 to 10.6)
CD 5B	6.4 (3.9 to 6.4)	7.7 (6.3 to 7.7)	.649 (0.379 to 0.767)	15.7 (-43.6 to 36.1)
CD 6A	8.1 (7.5 to 9.1)	7.5 (7.3 to 8.4)	.843 (0.615 to 0.931)	6.8 (-26.0 to 8.8)
CD 6B	6.0 (5.9 to 7.1)	5.8 (5.8 to 6.9)	.816 (0.606 to 0.888)	5.8 (-23.4 to 16.0)
CD 7A	6.1 (5.4 to 6.6)	5.4 (5.3 to 6.0)	.870 (0.613 to 0.944)	2.6 (-21.7 to 16.2)
CD 7B	4.9 (3.2 to 5.0)	5.0 (4.2 to 5.0)	.772 (0.571 to 0.870)	2.7 (-39.5 to 27.7)
CD 8A	3.8 (1.9 to 3.8)	5.0 (2.9 to 5.0)	.607 (0.390 to 0.767)	19.7 (-36.0 to 19.7)
CD 8B	4.3 (3.0 to 4.8)	4.8 (4.2 to 5.2)	.704 (0.410 to 0.807)	12.3 (-41.1 to 24.6)
CD 9A	4.6 (3.3 to 5.4)	6.4 (5.5 to 6.6)	.568 (0.344 to 0.741)	18.8 (-40.4 to 23.8)
CD 9B	3.2 (1.7 to 3.2)	4.5 (3.5 to 4.5)	.548 (0.264 to 0.729)	17.3 (-50.1 to 47.7)

acreage (table III). In North Dakota, spring wheat acreage was more evenly distributed with no CRD accounting for as much as 20 percent of the state acreage (table IV). Most of South Dakota's spring wheat acreage was grown in the three northern CRD's (10, 20, and 30) with CRD 50 accounting for another 6.2 percent of the acreage in 1976 (table V). In Saskatchewan, wheat acreage is distributed fairly evenly among the CD's with no CD accounting for as much as 10 percent of the total spring wheat acreage in the province (table VI). (A more complete set of data for the 1963-76 period is available from the authors on request as a statistical appendix.)

Review of Acreage Response Studies

The need for models capable of predicting ratios or relative acreages of wheat and other small grains is unique to the LACIE project, and no direct precedent for ratio modeling was found in the literature. Nonetheless, the ratio problem can be usefully viewed as a special case of the general problem of predicting how farmers adjust crop acreages in response to changing economic and physical signals. Several studies seeking to explain and predict wheat acreage response in the United States and Canada have appeared in recent years. A review of these studies provided useful background information in developing ratio models. All these studies employed single-equation multiple-regression analysis on time series data.

U.S. wheat acreage response studies.— Of three recent studies of U.S. wheat acreage responses reviewed, two were primarily concerned with assessing the impact of government policy on wheat plantings. Lidman and Bawden (ref. 8) focused on government allotment programs and the various incentives for program participation (loan rates, direct payments, voluntary diversion payments, etc.). Using a model to predict national wheat acreage from 1954 through 1970, the authors concluded that agricultural programs exerted a significant influence on the amount of wheat acreage planted, whereas lagged market price was not an important determinant in the acreage planted to wheat during the 1954-70 period (ref. 8, p. 333). No prediction tests were reported, but it is not likely that their model would perform well outside the sample period because a

general cropland set-aside program was substituted for the commodity specific allotment program in 1971.

In an attempt to assess the impact of this change in policy, Garst and Miller (ref. 9) developed a new set of wheat acreage response models. These models are quite similar to those presented by Lidman and Bawden, regressing planted acreage on wheat allotment, wheat diversions, and lagged price, with two dummy variables intended to represent changes in marketing quota requirements as well as the wheat set-aside variable. In all, three sets of model results are reported: one for all-wheat states, one for winter-wheat states, and one for all-spring-wheat states (North Dakota, South Dakota, Minnesota, and Montana). In the spring wheat model allotments, diversions and wheat set-aside variables were statistically important, whereas lagged price was not. The model explained a large proportion of the historical variation in spring wheat plantings (coefficient of multiple determination R^2 greater than 98.0). Nonetheless, Garst and Miller conclude that "as with the use of most models of this type, predictions of future impacts should be examined with some skepticism, particularly when they rely on data outside the range of that used in the regressions" (ref. 9, p. 36).

The third study reviewed the development of a predictive model as a primary goal of the analysis. Using state-level data for North and South Dakota for 1948 through 1974, Weaver, Morzuch, and Helmberger (ref. 10) hypothesize that planted acreages of wheat are a function of expected prices for wheat and alternative crops. In a departure from other acreage response studies, they employed indexes of future prices as proxies for expected prices. A trend intended "to account for any systematic changes attributable to changes in technology, relative factor prices, and other disturbing influences" (ref. 10, p. 8) is also included in the model. Predictions for 1975 and 1976 were generated extra-sample and then compared with preliminary official estimates. The authors concluded that ". . . the equations do not yield good predictions of planted spring wheat acreages in 1975 and 1976" (ref. 10, p. 18). Predicted acreage for North Dakota equaled 78 percent of the actual acreage in 1975 and 74 percent in 1976; for South Dakota, it equaled 74 percent of the actual acreage in 1975 and 67 percent in 1976. The authors characterize one of their important findings as being that, in the absence of binding acreage allotments

(such as those in effect in the years 1950 and 1954 through 1964), the acreage planted to spring wheat responds positively to the ratio of the expected price of spring wheat to the expected price of other crops (ref. 10, p. 18).

The poor predictive ability was not surprising considering the period used to develop the model. During most of the period, government programs created incentives to limit wheat and feed-grain acreage. During 1971 through 1973, government programs became more neutral concerning which crops farmers planted on their allotted crop acreage but continued to provide incentives for limiting total wheat and feed-grain acreages. After 1973, the wheat and feed-grain programs had essentially no impact on the grain industry because market prices were substantially in excess of target prices. Farmers were not only free to allocate "normal" cropland acreage among competing crops according to expected relative crop prices, but they also had an incentive to increase "normal" crop acreages by reducing land in fallow and, in some cases, planting crops on land usually considered to be marginal cropland. Thus, events in the period after 1973 introduced a basic supply shifter in the acreage response function which was not included in any of the studies reviewed.

Canadian acreage response studies.— Four Canadian acreage response studies were reviewed. Schmitz (ref. 11) theorizes that acreage planted to wheat is a function of expected price for wheat and other crops and of several nonprice variables, such as moisture before and during seeding, farm-level wheat stocks, wheat export sales, technology, and capital availability. (The last two variables were represented by a trend proxy in the estimations.) Estimation was accomplished by ordinary least-squares regression on national-level annual data from 1947 through 1966. In all, some 24 variations on the two basic models are reported. In no case was the correct sign obtained for moisture before and during planting or for livestock prices.

With respect to alternative crops, barley prices showed the expected sign but were not statistically significant, whereas flax prices were of the correct sign and usually significant in the various alternative models. Wheat prices were significant in every case, as were export sales and, generally, on-farm stock levels.

A study by Capel (ref. 12), which was intended to provide an efficient low-cost forecast of wheat acreage, employed the simplest form of the Nerlove distributed lag model, expressing acreage as a func-

tion of expected price. In practice, this amounts to a regression on wheat price and last year's acreage. Annual data for the Canadian prairie provinces from 1950 through 1967 were used to estimate the model coefficients in a double-log transformation. A test of predictive capability was not presented.

In the third study reviewed, Meilke (ref. 13) hypothesized that Canadian producers react to two different prices in making their production decisions. The first of these is the Canadian Wheat Board (CWB) initial and adjustment payments combined. Initial payments are received on delivery and constitute a floor price. Adjustment payments are usually made in the spring, retroactive to the beginning of the crop year. Meilke noted that the two payments taken together are quite stable. Final payments (and advances on final payments, or interim payments) are usually made 18 to 24 months after planting and are much more variable. As a test of the two-price hypothesis, Meilke constructed acreage response models for wheat, barley, and oats; one set includes the two variables separately, whereas one combines initial and final payments into a single variable. Other variables included marketings of wheat as a percentage of production plus on-farm carry-in stocks lagged 1 year (a proxy for the anticipated restrictiveness of marketing quotas), and a dummy variable equal to one in 1970 (representing the Lower Inventories for Tomorrow (LIFT) program). Price expectations are accounted for by a distributed lag formulation (i.e., the lagged dependent variable is included as a regressor).

These equations were estimated using annual data for the prairie provinces from 1949 through 1974 by ordinary least-squares regression. All coefficients were of the expected sign and had large *t*-values. Meilke interpreted these results as showing that "final payments have an effect on acreage" (ref. 13, p. 574).

The fourth study reviewed, by Meilke and Kramar (ref. 14), analyzed individual acreage response equations for corn, oats, barley, soybeans, mixed grain, and winter wheat in Ontario. Lagged barley yields, lagged barley prices, and lagged acreage of winter wheat were the independent variables used in the winter wheat equation. Predictions of winter wheat acreage for 1972 and 1973 using the model were considered "adequate," although naive forecasts (made by regressing acreages in year *t* on acreages in year *t* - 1) gave estimates of actual winter wheat acreage that were about as accurate. The results for other crop models were similar.

RATIO MODEL DEVELOPMENT PROCEDURES

Procedures used to develop and select ratio models were similar for both the United States and Canada. However, because of basic differences in agricultural policies and other factors, the resulting models were quite dissimilar.

Candidate Variable Selection

The acreage response literature suggested alternative variables for initial inclusion in the ratio models. Four categories of variables were indicated, representing changes in (1) economic conditions, (2) government policies, (3) historical crop-livestock patterns, and (4) physical conditions such as annual climatic variations.

Economic variables.— All else being equal, changes in planted acreages of wheat and other small grains should depend primarily on the expected net income relationship of these confusion crops prior to planting. Of course, the net income per acre from each crop is equal to the gross income (yield per acre in bushels times the farm price received per bushel) minus costs. Unfortunately, the net income farmers expect to receive for their crops when marketed is not directly observable. Time-series cost data are generally not available for the different confusion crops, but the relative costs of producing the different confusion crops generally change little from year to year. Thus, changes in expected gross income relationships should be an acceptable substitute for net income. Further, because small grains yields tend to respond similarly to climatic conditions, year-to-year variations in yields are correlated. Hence, price is likely to be the major economic factor causing annual adjustments in confusion crop ratios.

Unfortunately, neither the yields nor the prices farmers expect to receive for their crops when marketed are directly observable. Future prices are a possible proxy; however, future prices were not available for all confusion crops. More commonly, some combination of past prices is used to represent expected prices, a procedure based on the highly plausible assumption that past experiences determine farmers' price expectations. After selecting historical prices as a basis for farmers' price expectations, considerable latitude remains in choosing the precise means by which past price information is

reflected in the acreage response model. An obvious approach is to include prices lagged one or more periods as independent variables; however, alternative specifications may be weighted or unweighted moving averages of prices for several periods, a measure of relative price changes between two periods, and nonlinear transformations such as price squared or the log of price. In the ratio model analysis, several alternative price and yield combinations were analyzed.

Agricultural program variables.— Both the U.S. and Canadian Governments have taken an active role in determining wheat production, though it might be fair to characterize this role as an unintended one. Most production adjustment actions were undertaken to support income maintenance programs rather than as production management programs per se. Nevertheless, these programs have historically played an important role in determining small grains acreage adjustments. The Canadian and U.S. programs were sufficiently different to justify separate treatment here.

Canadian policy: In contrast to the United States, Canada employs a marketing board as the central feature of its wheat program (ref. 15). The CWB enjoys monopolistic power in both the procurement and the disposition of all wheat grown in the prairie provinces and the Peace River area of British Columbia that is destined for export or sale among provinces (feed wheat excepted). Prior to August 1, 1974, the CWB had similar control over feed wheat, oats, and barley grown in the Canadian West, but its jurisdiction has since been limited to the export market for feed grains, while the private sector has been allowed to carry out domestic feed-grain trade. In an effort to provide price stability for wheat, oats, and barley, the CWB employs "price pooling." Under this system, producers receive an initial payment, usually well below current market prices, on delivery of their crop to country elevators which act as agents of the CWB. A final payment is made based on CWB net revenues from marketing the crop.

Jolly (ref. 16) has identified two major effects of price pooling. First, because final payments are not generally made until 6 months after the close of the marketing year, and payments are occasionally delayed an additional 6 months after that, producers must make production and consumption decisions on the basis of incomplete price information. Second, since the pooling system averages out within-season price variation, the market mechanism does not serve to distribute deliveries over the year. The

second effect gives rise to an alternative method of distributing crop deliveries, the Grain Delivery Quota System. Under this system, producers are allowed to market a portion of their total quota at successive intervals. Quotas were determined by several criteria during the study period and most recently by "assigned acreage," which is defined as (1) land seeded to wheat, oats, barley, rye, rapeseed, and flaxseed; (2) land in summer fallow; (3) land in miscellaneous crops; and (4) land seeded to perennial forages up to a maximum of one-third of the total land in the other three categories. Changes in quotas are expected to cause changes in the relative confusion crop acreages.

The LIFT program was an important departure from historical Canadian programs. Enacted in 1970-71 for 1 year, LIFT was designed to reduce burdensome wheat inventories which had accumulated during the late 1960's. Producers were given a strong incentive to reduce wheat acreage and, indeed, seeded acreage decreased by 50 percent from the 1969-70 acreage.

The attempt to include Canadian policy in the ratio models introduces a large set of candidate variables which might theoretically be used to explain Canadian ratios. Possible variables suggested for testing included initial, interim, and final CWB payments to producers for wheat, oats, barley, and rye. Minimum support prices for wheat for domestic needs; data on the various CWB marketing quotas, particularly the so-called general quotas and the acreage factors (variously referred to as specific acreage, assigned acreage, etc.) used to compute quotas; and information on diversion payments made under the LIFT program were also considered.

U.S. programs: The study period used in developing the U.S. ratio models (1963 through 1976) included two rather dramatically different agricultural policy environments (ref. 17). Within these policy periods, important annual adjustments in farm programs were made. Before 1970, wheat production was strongly influenced by the U.S. Department of Agriculture (USDA) through an acreage allotment program. This program was voluntary (unlike its predecessor prior to 1963) in that farmers who did not comply with their allotments were not fined or otherwise penalized except by being denied access to government price and income support programs. Economic incentives for complying with program provisions were nonetheless strong, since participating producers received price support loans, marketing certificates redeemable for cash, and payments

for diverting additional acreage from their allotments, whereas noncomplying producers received no direct benefits. Similar programs were applied to feed grains such as corn and barley and, at various times, to oats and flax.

Under the Agricultural Act of 1970, beginning with the 1971 crop, the use of acreage allotments was suspended for wheat and feed grains. Instead, producers were required to keep a certain percentage of their total cropland out of production in order to be eligible for price supports. However, they were free to plant whatever crops they desired on the remaining land. Allotments were retained only in the limited sense that they were used to apportion domestic marketing certificates, which, to producers, were worth the difference between the wheat parity price and the average market price during the first 5 months of the year. A factor here was the provision which required a producer to plant allotted wheat acreage to maintain the allotment for certificate purposes. Compared to earlier programs, increased substitution was allowed between wheat and feed grains. In fact, one of the principal aims of abandoning the old allotment system was to allow the market to allocate land among crops. The shift in policy was continued by the Agriculture and Consumer Protection Act of 1973, legislation written in the wake of a widely perceived shift from surplus to shortage in the world wheat market. The intent of this legislation was to encourage expanded production (ref. 15, p. 19). An innovation of the 1973 act was the introduction of the target price concept, under which producers were paid the difference between the target price and the average market price for the first 5 months of the marketing year. As in the 1970 act, allotments were retained as the basis for program payments.

Candidate variables designed to reflect the agricultural policy environment during the early years of the 1973-76 period should obviously include allotted acreages for wheat and competing crops, as well as acreage diverted for payment from those allotments. Price support loan rates, the dollar value of diversion payments, and the value of wheat marketing certificates are also relevant. Changes in acreage allotments and diversion incentives are expected to have a direct effect on planted acreages. Since much of the government's price support activity was conducted through the nonrecourse loan program, levels of Commodity Credit Corporation (CCC) grain stocks are another possible explanatory variable. These same variables (or their equivalents)

might also have a role in explaining ratio adjustments after 1973, but any model incorporating them should consider the emergence of a policy specifically intended to reduce the role of the government in producers' decisions.

Historical crop livestock patterns.— The emphasis so far has been on modeling farmers' responses to changing economic and governmental policy signals. It seems reasonable to assume, however, that these responses are conditioned by historical crop and livestock production patterns. Farmers may have difficulty adjusting to changing conditions as fully and as quickly as they would like to adjust. Annual adjustments are limited (1) because farmers are committed to a particular rotation schedule, (2) because they have invested in specialized equipment while changing cropping patterns may require additional machinery investments with unknown payoffs, or (3) because local conditions make the marketing of a particular crop more desirable than is readily apparent from CRD- or state-level data. Furthermore, farmers may occasionally continue a particular pattern of crop production from sheer force of habit, regardless of the current economic signals they may be receiving. Thus, it was expected that regional crop/livestock patterns might tend to constrain rather than lead crop adjustments. In terms of the ratio project, the primary means suggested for allowing for historical production patterns was to include the lagged dependent variable as a possible explanatory variable.

Physical factors.— Some of the most important factors determining acreage patterns are physically based rather than socially oriented. Two obvious examples are soils and climate. Interregional differences in normal climatic states are important factors explaining interregional differences in production patterns. Together with the factor of differing soils, interregional climatic differences help to justify the development of separate models for each region (CRD or CD) of interest.

However, although explaining interregional differences, physical factors are normally somewhat less important causes of year-to-year changes in crop acreage. Although varying among regions, for all practical purposes, soils are constant over time. Climate, on the other hand, does exhibit variation through time; and annual climatic changes could conceivably help explain annual changes in crop ratios. For example, in areas where it is possible to substitute spring wheat for winter wheat, the SW/GR ratio might very well be affected by the severity of

winterkill. Likewise, weather-delayed planting dates could affect production patterns as farmers substitute crops with a shorter growing season. Considerations such as soil moisture levels at planting could also induce producers to substitute one crop for another.

Data Collection and Limitations

Data needs and collection problems differed between Canada and the United States.

Canadian data.— Data were available for a much longer time period in Canada. The existence of data for all candidate variables allowed CD model coefficients to be estimated over the 1948-49 to 1976-77 period. Nearly all Canadian agricultural statistics are gathered on a marketing-year basis. The marketing year begins on August 1 and ends the following July 31. Thus, 1976-77 indicates the year beginning August 1, 1976, and ending July 31, 1977. However, in the 1976-77 crop year in Saskatchewan, crops are planted in late April and May of 1976.

Data were collected for two types of variables: those specific to particular regions and those which can be applied across regions (ref. 2). Region-specific data included CD- or zone-level acreage estimates for oats, barley, rye, mixed grains, flax, rapeseed, and wheat. Nonspecific data included on-farm stocks of wheat at the beginning of the crop year; total wheat production; prices for wheat, barley, flax, rapeseed, and rye; CWB exports of wheat; marketings of wheat by producers to the CWB from the beginning of the crop year to March 1; the CWB initial buying price for wheat; and the CWB selling price for wheat (ref. 2). The CD acreage, yield, and production data were collected from the Saskatchewan Office of Statistics (ref. 18). Other data are available from national publications of Statistics Canada (refs. 19 to 22).

U.S. data.— The most limiting and difficult data base to construct was the long-term historical information on CRD agricultural program factors. The construction of CRD-level agricultural program variables required aggregating county-level data on each of several program features. Program details varied from year to year, and the method of reporting varied by state. Basic data for constructing the various agricultural program variables were provided by the national and individual state offices of the USDA Agricultural Stabilization and Conservation Service (ASCS) for the 1966-76 period. The official policy is to maintain national, state, and county pro-

gram summaries for 10 years only. In most cases, however, ASCS personnel at each state office generously provided personal copies of published ASCS data for 1963 through 1966 (refs. 23 to 26). Nevertheless, county data were not available for Minnesota in 1964 and 1965 or for Montana in 1963. Because of major differences in agricultural programs prior to 1963 and the extreme difficulties or the impossibility of collecting earlier county-level data, it was decided to limit U.S. data bases to the post-1962 period.

Another data limitation was the lack of timely CRD-level price data on confusion crops for each of the four states. Timely data were particularly lacking for monthly price series at the CRD level. However, in those states where CRD-level crop prices were reported, the size of the differences among monthly CRD prices was generally small (less than 5 percent). Also, because only state-level data are published in time to meet operational requirements, state prices were used in developing CRD models. These prices are published by the USDA Economics, Statistics, and Cooperatives Service (ESCS, formerly the Statistical Reporting Service (SRS)).

State and national data on quarterly grain stocks for 1961 through 1976 were collected from ESCS sources (ref. 27). Also, CRD data on mean monthly precipitation and mean monthly temperatures for 1961 through 1976 were tabulated from National Weather Service (NWS) records.

The CRD data on acreage, yield, and production for the individual confusion crops were generally available from the published data of the ESCS state office. Also, CRD acreage data for the 1976 crop year were not available until the final stages of the study.

Primarily because of limited historical data, the preliminary CRD models were estimated using data for (1) Minnesota, 1966 through 1975; (2) Montana, 1964 through 1975; (3) North Dakota, 1963 through 1975; and (4) South Dakota, 1963 through 1975. Data for 1976 were added before developing the models for predicting 1977 ratios. Thus, final models were developed using 11 observations for Minnesota, 13 observations for Montana, and 14 observations each for North and South Dakota.

Model Selection Procedures

Multiple regression analysis was used for parameter estimation (ref. 28). Several statistical tests were used as aids in selecting the "best" model

from a number of alternatives. Since the purpose of the ratio modeling effort was to produce predictions, the "best" model is the one with the greatest predictive ability. Indicators of a model's predictive ability include the coefficient of multiple determination R^2 , the F-test for significance of the overall regression, the consistency of coefficient signs with economic theory and other a priori information, the t-value tests of coefficient significance, the mean square error, and the Durbin-Watson d-statistic. Unfortunately, none of these tests provide a completely conclusive indicator of a model's predictive capability, and the final choice of a "best" model remains partly a matter of judgment.

The general procedure for the ratio study was to specify several plausible alternative models using data through 1975 and then to evaluate them by the available statistical tests. This exercise was supplemented by generating extra-sample predictions for the 1976 crop ratios. This allowed 1 year, 1976, for comparing predicted and observed values. The models that showed the greatest promise were then subjected to further analysis and refinement until the apparent "best" model emerged.

Because of the large number of models required, the analysis was begun by developing preliminary state- or province-level models to use as guides in variable selection for CRD and CD models. These preliminary studies were intended to narrow the range of possible independent variables or model alternatives to be considered. A principal rationale for employing state- or province-level data was that the development of individual models tailored to each CRD or CD would have required more research resources than were available and might have resulted in a product that was operationally cumbersome.

Instead, the intention was to produce for each state or province a general model (for each ratio type) that would retain the same basic variables for all CRD's or CD's while allowing the estimates of the coefficients to vary. In the early stages of the study, this procedure was also necessary because CRD and CD data were unavailable for analysis. However, in both countries, modifications were made at the state/province and CRD/CD level as the researchers tested and compared model results. Although the procedures used in the development of the U.S. and Canadian models were similar, enough differences existed to justify the reporting of each separately.

Canadian model development.—In preliminary work, the ratio of spring wheat to spring grains was

hypothesized to be a linear function of several economic, technological, and policy variables (ref. 2, pp. 33-37). A crop acreage ratio response function for Saskatchewan was estimated from a set of variables consisting of the annual percentage change in (1) the price of wheat, (2) the price of rapeseed, (3) the price of rye, (4) the price of barley, (5) the total production of wheat lagged 1 year, (6) the total on-farm stocks of wheat, (7) the dummy variable to account for the effects of the LIFT program instituted in 1970, (8) wheat exports, and (9) the SW/SG confusion crop ratio lagged 1 year. A more detailed definition of these variables, the preliminary model parameters, and the estimated coefficients can be found in reference 2.

From this set of independent variables, a forward-selection stepwise estimation procedure (found in several computer multiple-regression packages) was used for further analysis to select those variables which increased the R^2 , decreased the variance, and had the theoretically expected sign. The significance of the t-values of the coefficients was ignored because a predictive model was desired as the end result. (The Canadian and U.S. models were developed separately (refs. 1 and 2). The researcher who developed the Canadian models argued that "since the purpose of this research is to generate a 'predictive' model and not an 'explanatory' model (where t-values for individual coefficients become highly relevant), there is no compelling reason to test for the significance of the coefficients" (ref. 2, p. 35). This contention is supported by Johnston (ref. 29). More emphasis should be placed on a significant F-test and the size of the overall variance.)

Additional analysis resulted in two important changes: the removal of the dummy variable and the dropping of four other variables (the percentage change in the price of wheat, rapeseed, rye, and the production variable). The dummy variable was included in the first model to account for the effects of the LIFT program. Although it was one of the most significant variables in that model, dummy variables provide problems in a predictive model. Because it was desired to be able "to predict the efforts of events such as the LIFT program without having prior information about parameters on variables representing them . . . it was deemed desirable to test whether a model could be developed which had satisfactory predictive capabilities without including policy variables" (ref. 2, p. 35).

Four other variables were deleted from the preliminary model as a result of additional work

done at the zone level. The forward-selection stepwise estimation procedure was applied to each of the nine zones (fig. 2), and only those variables which had the theoretically expected sign for the coefficients in all nine zones were included in the final model. As a result of this procedure, the final model included only four independent variables—wheat exports, on-farm stocks of wheat, the lagged dependent variable, and the percentage change in the price of barley.

The final model and a discussion of the relevance and the predictive ability of the model for Canada is presented in the section entitled "Results and Discussion."

U.S. model development.—Preliminary models for each of the four states were developed separately as a guide for developing CRD models in each state (ref. 1). The general procedure used in initial model formulations involved selecting four or five policy, economic, and dummy variables; running the regression; and then examining the computer run. In selecting a particular model formulation, emphasis was placed on the significance of individual t-values for regression coefficients, the coefficient of multiple determination R^2 , the mean square error, and the Durbin-Watson test statistic.

The more numerous agricultural program changes and the much more limited data base available in the United States necessitated a different development procedure than that followed for Canada. Because a maximum of 13 years (1963 through 1975) of CRD data were available, initial model formulation was based on the need for frugality in the use of independent variables to maintain a sufficient number of degrees of freedom for statistical validity.

One of the consequences of the short data base was the limited ability to make use of a stepwise selection procedure for choosing among parameters. Alternative model possibilities and variable formulations had to be tested largely without the benefit of this statistical tool. "Curve fitting," or selecting a model which fits the sample data well but has no underlying validity, is always a danger when using model selection routines such as the stepwise procedure; this is particularly true when there are a large set of candidate regressors and a small sample, as in the U.S. study.

Although several combination variables consisting of ratios of gross income from the confusion crops or ratios of confusion crop prices were tested in preliminary model analysis, the best economic variable appeared to be the difference between price

series on competing crops; for example, wheat price minus oats price. Also, this reduced the number of independent variables in the model.

The importance of agricultural adjustment programs in determining confusion crop acreage, the high frequency of changes in program provisions during the period, and the varying impact among states due to different farmer participation rates indicated the necessity for finding a succinct way to capture the impacts of program changes. The main task was to specify variables which could separate the past effects of agricultural adjustment programs from economic predictor variables. As the programs were generally not important in planting decisions after 1973, no effort was made to develop a variable which could predict the impacts of the programs. (In 1978, it is likely that a program variable would again be needed.) The limited information at the CRD level and the limited resources precluded a search for such predictive variables.

The preliminary model results for the United States were generally disappointing. Numerous combinations of price and gross income variables for the confusion crops were tested in these preliminary models; however, little progress in model development was made until the 1976 CRD data bases became available for analysis.

The major problem in the early stages of the study was the failure to fully recognize the nature of the economic adjustment that was occurring after 1972. Beginning in 1972, high export demands caused farm prices to substantially exceed price support levels for the first time during the 1963-76 period. Prior to this time, the most important factors determining acreage adjustments in the United States were policy related. Until the 1976 acreage data became available, the nature of the structural changes taking place was not recognized by the researchers. With the availability of 1976 acreage data, it became obvious that farmers were rapidly adjusting wheat production to the "free market" conditions existing since 1973. Their reaction to "free market" conditions became a major variable in the models developed to predict 1977 ratios. The nature of the farmers' response to "free market" prices was not readily apparent using the 3 years of data from 1973 through 1975. The adjustment during these 3 years had been unidirectional. The nature of the adjustment became known as farmers responded to altered price relationships in the 1976 crop year. This is in contrast to Canada, where pricing policies remained fairly consistent during the study period.

Although climatic variables and U.S. grain stock

variables were introduced into several of the models, these variables did not appear to cause improved predictive capability. However, some of the events explained by dummy variables appeared to be climate related. A more complete discussion of the alternative models tested can be found in reference 1; the model forms selected for predicting 1977 ratios are presented in section 5.2 of that report.

RESULTS AND DISCUSSION

Because the model variables for the SW/SG ratios were identical to the variables for the SW/SG ratios and because the model variables for predicting planted-acreage ratios (in those states where predicted) were identical to the variables in the ratio models for harvested acreage, only the final models used to predict the SW/SG, WW/WG, and WW/GR ratios for harvested acreage are reported. (The model coefficients and the ratio predictions for 1976 and 1977 for those results not reported here are available from the authors as a statistical appendix.) Although the statistical procedures used to develop and select ratio models were similar for both the United States and Canada, basic differences between the two countries in agricultural policies and other factors caused the model forms for each country to be different.

Results for Canada

Model form.— The general form of the SW/SG ratio model provided for each CD, for each zone, and for the province was as follows.

$$SW_t^i / SG_t^i = \beta_0 + \beta_1 EXP_{SW}^{t-1} + \beta_2 SW_t^{i-1} / SG_t^{i-1} + \beta_3 CI_{SW}^{t-2} + \beta_4 EP_{BAR}^{t-1}$$

where

SW_t^i = acres of spring wheat harvested in year t in geographic unit i . For example, if t equals the 1977-78 crop year which begins August 1, 1977, and ends July 31, 1978, then that crop is planted in April and May of 1977, and t can be any of 20 CD's, 9 zones, or

- the province.
- SG_i = total acres of spring grain confusion crops (spring wheat, spring rye, oats, barley, mixed grains, and buckwheat) harvested in year t in geographic unit i
- EXP_{SW}^{t-1} = exports of total wheat in bulk (millions of bushels) from August 1 to March 1 in crop year $t - 1$
- SW_i^{t-1}/SG_i^{t-1} = the previous year's spring confusion crop ratio; lagged dependent variable. (Following Nerlove's technique (ref. 30), this amounts to a distributed lag.)
- CI_{SW}^{t-2} = the total on-farm stocks of wheat (millions of bushels) carried into crop year $t - 1$; where, for predicting the 1977-78 crop year ratio, the average for CI_{SW}^{t-2} is equal to the July 31, 1976, on-farm storage.
- EP_{BAR}^{t-1} = the change in price (dollars per bushel) of grade 1 Canadian western six-row barley, from March 1 in crop year $t - 2$ to March 1 in crop year $t - 1$. Calculated by the formula $(p_{BAR}^{t-2} - p_{BAR}^{t-1})/p_{BAR}^{t-2}$ where p_{BAR}^{t-1} is the March 1 CWB export selling price at Thunder Bay in crop year $t - 1$.

Table VII reports—for each CD, zone, and province—the ordinary least-squares estimates of the coefficients for the variables, the t -values for the estimates of the coefficient, and the coefficient of multiple determination R^2 for each equation. Each of the four independent variables has the theoretically expected sign in every CD and zone and in the province model. The sign of the coefficients for the change in the price of barley and for wheat stocks carry-in are expected to be negative, as increases in either would be expected to decrease the economic returns to wheat production. For example, a rise in the price of barley would be expected to cause some movement into barley production. When wheat stocks are high, an inducement exists for the economically rational producer to shift to another crop. The signs on the coefficients for the lagged dependent variable SW_i^{t-1}/SG_i^{t-1} and the annual exports

EXP_{SW}^{t-1} were positive as expected. An increase in exports usually indicates a better price for wheat and a tendency for farmers to move into wheat production. The lagged dependent variable indicates a tendency on the part of some farmers to follow recent production practices, partly because farmers give less than full credibility to current market signals.

Based on an examination of t -values in table VII, the most significant variables were the export and the lagged dependent variables, which are generally significant at the 1-percent level of probability. The least significant variable was the annual percentage change in the price of barley.

The total regression for every CD, zone, and province was significant at the 1-percent level. However, the R^2 values were generally low, particularly at the zone and CD levels. A low R^2 indicates that the model does not explain a considerable amount of the variation found in the historical SW/SG ratio. However, the CD's with the smallest R^2 , CD 4A and CD 4B, were also the ones with the smallest variation in the historical SW/SG ratios. The variation in zone 4 during the last 10 years ranged between 0.775 and 0.910, compared to 0.312 to 0.656 in zone 9. Thus, although the model explained little of the total variation in zone 4, there was little variation to be explained.

Precision and accuracy tests.—Tests were made of the accuracy of the model developed to predict the SW/SG ratio. At the province level, 10 single-year predictions from crop year 1967-68 to crop year 1976-77 were made using the "best" model and the previous year's ratio (lagged ratio). In table VIII, these results are compared to the actual ratios for each of the 10 years. The predictions with the "best" model were made by using data only up to the 1966-67 crop year for the first year's prediction (1967-68). Parameter estimates were then updated yearly and a prediction made for each of the following years.

The "best" model gave a more accurate estimate of the actual SW/SG ratio than did the lagged ratio in 7 of the 10 years (table VIII). The "best" model was accurate within 90 percent of the actual ratio value in 9 out of 10 years. However, the lagged ratio was also at least 90-percent accurate in 9 of the 10 years. Neither the "best" model nor any of the alternatives was able to readily predict the effects of the LIFT policy with no prior information on program parameters (ref. 2).

Ten-year tests were not run at the CD level. However, a paired difference test (see ref. 2 for procedure) was made between the predicted ratio values

of the model for 1976-77 and the predictions based on using last year's SW/SG ratio. A calculated t-value, based on the absolute values of the weighted CD paired differences, shows that the predicted 1976 SW/SG ratios have less error associated with them, and tests indicate they are from a different population than those predictions based on last year's ratio (table IX). Therefore, the ratio predictions from the "best" model are concluded to be statistically more

accurate (at the 1-percent level of significance) than the present LACIE technique of using last year's ratio (ref. 2, p. 24).

Model predictions for 1977.— Table X shows the 1977 predicted SW/SG ratio for each CD, zone, and province using the ratio model. In almost all cases, the predicted ratio is slightly larger than the 1976 ratio. Three of the four predictor variables are favorable toward increased wheat acreage in 1977.

TABLE VII.—Saskatchewan: The Models Chosen to Predict the SW/SG Harvested Acreage Ratios for Each Crop District, Zone, and Province

[Based on ref. 2, table 1]

Geographic unit	Ordinary least-squares estimates of the model coefficients					R ² (a)
	Constant	CI_{SW}^{t-2}	EXP_{SW}^{t-1} (a)	EP_{BAR}^{t-1}	SW_i^{t-1}/SG_i^{t-1} (a)	
CD 1A	0.175 (1.95) ^b	-8.32E-5 (-0.76)	5.28E-4 (2.53)	-0.103 (-1.55) ^c	0.643 (4.91)	66.5
CD 1B	.133 (1.72) ^b	-7.23E-5 (-.64)	6.84E-4 (3.14)	-.108 (-1.57) ^c	.620 (5.14)	71.4
CD 2A	.401 (3.99) ^a	-1.38E-4 (-2.54) ^a	4.00E-4 (4.03)	-3.45E-2 (-1.08)	.463 (3.80)	71.4
CD 2B	.589 (4.66) ^a	-3.08E-4 (-4.25) ^a	4.01E-4 (3.90)	-3.68E-2 (-1.08)	.269 (1.93) ^b	75.4
CD 3AS	.299 (2.99) ^a	-1.19E-4 (-2.78) ^a	2.63E-4 (3.32)	-2.81E-2 (-1.09)	.627 (5.35)	73.0
CD 3AN	.407 (4.01) ^a	-1.39E-4 (-2.36) ^b	3.88E-4 (3.49)	-4.11E-2 (-1.16)	.462 (3.69)	66.0
CD 3BS	.250 (2.27) ^b	-4.84E-5 (-1.14)	2.36E-4 (3.02)	-1.53E-2 (-.58)	.665 (4.97)	65.9
CD 35N	.572 (5.60) ^a	-2.16E-4 (5.46) ^a	2.74E-4 (4.78)	-1.23E-2 (-.65)	.324 (2.89)	83.3
CD 4A	.149 (.94)	-8.17E-6 (-.45)	2.87E-4 (2.71)	-4.88E-2 (-1.27)	.750 (3.81)	45.8
CD 4B	.408 (3.04) ^a	-1.24E-4 (-2.72) ^a	1.76E-4 (2.26) ^b	-2.80E-3 (-1.09)	.533 (3.76)	62.2
CD 5A	.210 (2.76) ^a	-1.42E-4 (-1.94) ^b	6.15E-4 (4.69)	-7.45E-2 (-1.75) ^b	.574 (5.37)	78.3
CD 5B	.169 (2.14) ^b	-1.49E-4 (-1.37) ^c	6.94E-4 (3.45)	-1.02E-1 (-1.57) ^c	.541 (4.20)	68.3
CD 6A	.499 (4.72) ^a	-3.10E-4 (-4.52) ^a	3.96E-4 (3.63)	-6.91E-2 (-1.85) ^b	.367 (3.06)	78.0
CD 6B	.403 (4.30) ^a	-2.48E-4 (-4.07) ^a	4.45E-4 (4.69)	-6.33E-2 (-2.01) ^b	.432 (3.87)	80.4
CD 7A	.490 (4.06) ^a	-3.57E-4 (-4.21) ^a	4.14E-4 (3.99)	-7.88E-3 (-.23)	.388 (3.04)	83.8
CD 7B	.337 (3.92) ^a	-2.27E-4 (-2.83) ^a	7.22E-4 (5.23)	-4.16E-2 (-.93)	.403 (3.50)	76.0
CD 8A	.188 (2.34) ^b	-1.41E-4 (-1.54) ^c	5.21E-4 (3.21)	-9.35E-2 (-1.75) ^b	.546 (4.15)	66.7
CD 8B	.355 (3.23) ^a	-2.89E-4 (-3.14) ^a	4.40E-4 (3.02)	-1.00E-1 (-2.02) ^b	.440 (3.19)	72.3
CD 9A	.240 (2.54) ^a	-2.10E-4 (-2.27) ^b	4.43E-4 (3.02)	-1.20E-1 (-2.47) ^b	.506 (3.69)	69.7
CD 9B	.237 (2.13) ^b	-2.59E-4 (-2.07) ^b	4.92E-4 (2.72)	-9.66E-2 (-1.62) ^c	.505 (3.28)	68.2
Zone 1	.135 (1.78) ^b	-7.15E-5 (-.72)	5.17E-4 (3.01)	-1.02E-1 (-1.71) ^b	.671 (5.83)	74.1
Zone 2	.529 (4.49) ^a	-2.35E-4 (-3.76) ^a	4.12E-4 (4.25)	-3.72E-2 (-1.15)	.324 (2.41) ^b	72.3
Zone 3	.388 (4.01) ^a	-1.26E-4 (-3.41) ^a	2.84E-4 (4.32)	-2.14E-2 (-.99)	.518 (4.62)	75.8
Zone 4	.338 (2.24) ^b	-7.47E-5 (-1.61) ^c	2.20E-4 (2.62)	-1.55E-2 (-.55)	.571 (3.37)	49.4
Zone 5	.184 (2.39) ^b	-1.41E-4 (-1.57) ^c	6.54E-4 (3.96)	-8.89E-2 (-1.67) ^c	.565 (4.80)	73.5
Zone 6	.460 (4.64) ^a	-2.84E-4 (-4.42) ^a	4.20E-4 (4.16)	-6.62E-2 (-1.95) ^b	.390 (3.40)	79.4
Zone 7	.447 (4.57) ^a	-3.07E-4 (-4.26) ^a	5.67E-4 (5.59)	-2.36E-2 (-.70)	.357 (3.08)	82.8
Zone 8	.272 (2.91) ^a	-2.18E-4 (-2.52) ^a	4.78E-4 (3.33)	-9.94E-2 (-2.05) ^b	.496 (3.80)	70.8
Zone 9	.240 (2.37) ^b	-2.29E-4 (-2.21) ^b	4.61E-4 (2.91)	-1.11E-1 (-2.11) ^b	.505 (3.51)	69.2
Province	.323 (3.68) ^a	-1.73E-4 (-3.04) ^a	4.30E-4 (4.51)	-6.28E-2 (-1.98) ^b	.505 (4.51)	77.4

^aSignificant at the 1-percent level.
^bSignificant at the 5-percent level.
^cSignificant at the 10-percent level.

The price of barley declined from 1976 to 1977, the lagged variable is larger than in the previous year, and wheat stock carry-in is at a 20-year low. Only the wheat export level was less favorable for 1977 than for 1976. Data were not yet available to evaluate the model predictions for 1977.

Results for the United States

Model form.—For the United States, a substantial number of alternative models were analyzed after 1976 CRD data became available, but only the “best” models were reported (ref. 1). Because the model variables differed from state to state, the results for each individual state are presented separately. Models are presented for SW/SG, WW/WG, and WW/GR ratios for each state where an analysis was made.

North Dakota CRD models.—The North Dakota SW/SG models included the following variables.

$$SW_{i,j}^t / SG_{i,j}^t = \beta_0 + \beta_1 PWO_j^t + \beta_2 WAD_{i,j}^t + \beta_3 D71^t$$

where $SW_{i,j}^t$ = spring wheat acreage (harvested or planted) for year t in CRD i of state j

$SG_{i,j}^t$ = spring grain acreage (harvested or planted) for year t in CRD i of state j

PWO_j^t = a 3-year weighted moving average of the difference between the season average spring wheat price (cents per bushel) and the season average oats price (cents per bushel) received by farmers in state j ; that is,

$$PWO_j^t = 0.5(PW_j^{t-1} - PO_j^{t-1}) + 0.3(PW_j^{t-2} - PO_j^{t-2}) + 0.2(PW_j^{t-3} - PO_j^{t-3})$$

where PW_j^{t-1} and PO_j^{t-1} are the season average price of wheat and oats, respectively, received by farmers in the marketing year prior to planting. The marketing year runs from June 1 to May 31; consequently, prices for the complete marketing year $t - 1$ are not

TABLE VIII.—A Comparison of the Accuracy of Three Techniques to Predict the SW/SG Confusion Crop Ratio for the Province of Saskatchewan, 1967-68 to 1976-77

[Based on ref. 2, table 2]

Year	Actual SW/SG ratio	Lagged SW/SG ratio (a)	Percent difference actual vs. lagged	Prediction based on “best” model	Percent difference actual vs. prediction based on model
1976-77	0.782	0.740	-5.37	0.803	2.69
1975-76	.740	.704	-4.86	.744	.54
1974-75	.704	.713	1.28	.714	1.42
1973-74	.713	.688	-3.51	.743	4.21
1972-73	.688	.631	-8.28	.727	5.67
1971-72	.631	.602	-4.60	.614	-2.69
1970-71	.602	.778	29.24	.722	19.93
1969-70	.778	.813	4.50	.808	3.86
1968-69	.813	.833	2.46	.796	-2.09
1967-68	.833	.824	-1.08	.858	3.00

^aActual SW/SG ratio lagged 1 year.

available prior to planting. All prices were assumed to be zero prior to 1971.

$WAD'_{i,j}$ = the size of the wheat allotment (hundreds of acres) minus the wheat acreage diverted (hundreds of acres) in CRD i of state j in year t

$D71'$ = a dummy variable equal to 1 for the year 1971 and 0 otherwise

Table XI shows for each CRD the ordinary least-squares estimates of the coefficients for the vari-

ables, the t-values for the estimates of the coefficients, and the coefficient of multiple determination for each equation. Generally, the signs of the coefficients are consistent with theoretical expectations. As expected, an increase in the difference between historical wheat prices and oats prices PWO'_t led farmers to plant a larger share of wheat acreage to small grain acreage. It should be noted that PWO'_t was a 3-year weighted moving average of the difference between spring wheat prices and oats prices. The weights are based on a hypothesis about how farmers "learn" to form expectations about future prices. The 3-year moving average generally provided a more significant variable than a single-year lagged price

TABLE IX.—Saskatchewan: Paired Comparison Test of the Difference in Accuracy Between the "Best" Model and Last Year's Ratio as Predictors of the 1976 SW/SG Ratio for Each Crop District^a

[Based on ref. 2, tables 12 and 13]

CD	Percent difference (absolute value) between 1976 actual SW/SG ratio and —		Weight, percent SW acreage in crop district × 100	Difference between model-predicted error and last year's error weighted by percent SW acreage (b)
	Model-predicted 1976 SW/SG ratio	Last year's (1975) SW/SG ratio		
	d_i	d_i^2	w_i^{t-76}	$x_i^{t-76} = w_i^{t-76} (d_i^t - d_i^2)$
6A	4.63	6.41	8.1	-14.42
2B	.22	5.20	6.7	-33.37
3AS	.43	2.57	6.7	-14.34
5A	1.34	8.72	6.6	-48.71
3BN	.77	1.97	6.4	-7.68
5B	.00	13.56	6.4	-86.78
7A	4.25	2.53	6.1	10.50
6B	3.43	5.51	6.0	-12.48
7B	5.96	2.59	4.9	16.51
2A	.77	3.63	4.7	-13.44
3BS	.75	1.72	4.6	-4.46
9A	5.99	15.85	4.6	-45.36
1A	.65	7.44	4.4	-29.82
8B	5.40	10.94	4.3	-23.82
8A	1.81	16.47	3.8	-55.71
3AN	.79	3.70	3.4	-9.89
1B	3.84	7.53	3.3	-12.18
4B	.75	.97	3.3	-.72
9B	10.58	14.78	3.2	-13.44
4A	3.50	3.62	2.4	-.29

^aTest based on the absolute value of the percent error of the two predicted values from the actual SW/SG ratio.

^bFor these data, $\bar{X} = 20.00$, $\sum(X_i - \bar{X})^2 = 11\,278.99$, $S_p^2 = 5.45$. Under the hypothesis of no difference, $H_0: \bar{X} = 0$ against $H_a: \bar{X} \neq 0$ with $t = 3.67$ for $t_{(.01,19)} = 2.539$; H_0 is rejected and the two sets of predictions are from different populations.

difference variable. The sign of the policy variable $WAD_{i,j}^t$ was expected to be positive. This variable represents the net effect of the government wheat program on wheat acreages. The signs of $WAD_{i,j}^t$ were negative in CRD 50 and CRD 60 but the t-values of the coefficients were extremely small, indicating the variable could possibly be dropped in these two CRD models. The dummy variable represents a 1-year increase in spring wheat acreage due to a government policy shift from a diversion program (1962 through 1970) to a set-aside program (1971 through 1974). Spring wheat acreage rose substantially in 1971, and other variables were unable to account for this rise.

The only independent variable affecting changes in the ratio after 1971 was PWO_j^t . Other variables in the model serve simply to specify model structure in the earlier years of the 1973-76 period. The relatively

high R^2 values indicate that the models explain a large amount of the historical variation in the SW/SG ratio.

Minnesota CRD models.— The SW/SG model for harvested acreage in Minnesota differed from the North Dakota model in the use of a lagged dependent variable. The model equation was of the form

$$SW_{i,j}^t / SG_{i,j}^t = \beta_0 + \beta_1 PWO_j^t + \beta_2 WAD_{i,j}^t + \beta_3 (SW_{i,j}^{t-1} / SG_{i,j}^{t-1})$$

The estimated coefficients for the CRD models are shown in table XII.

In most of the Minnesota CRD's, the only significant t-values for variable coefficients were for the

TABLE X.—Saskatchewan: The 1977 Predicted Spring Grains Confusion Crop Ratios (SW/SG) for the Crop Districts and Province

[Based on ref. 2, table 14]

Geographic unit	Actual 1976 SW/SG ratio	Predicted 1970 SW/SG ratio	90-percent confidence limits for the predicted 1977 SW/SG ratio (a)	90-percent confidence limits for the predicted 1977 SW/SG ratio
CD 1A	.0766	.0791	0.719 to 0.879	0.634 to 0.947
CD 1B	.677	.711	.646 to .790	.549 to .873
CD 2A	.909	.910	.827 to 1.000	.843 to .986
CD 2B	.904	.913	.830 to 1.000	.832 to .994
CD 3AS	.935	.941	.855 to 1.000	.879 to 1.000
CD 3AN	.891	.904	.822 to 1.000	.819 to .988
CD 3BS	.932	.932	.847 to 1.000	.869 to .955
CD 3BN	.913	.922	.838 to 1.000	.877 to .968
CD 4A	.885	.880	.800 to .978	.786 to .975
CD 4B	.930	.939	.854 to 1.000	.879 to 1.000
CD 5A	.745	.776	.705 to .862	.675 to .877
CD 5B	.649	.678	.616 to .753	.526 to .831
CD 6A	.843	.889	.808 to .988	.803 to .975
CD 6B	.816	.850	.773 to .944	.776 to .925
CD 7A	.870	.908	.825 to 1.000	.826 to .989
CD 7B	.772	.806	.733 to .896	.700 to .911
CD 8A	.607	.638	.580 to .709	.512 to .764
CD 8B	.704	.759	.690 to .843	.645 to .872
CD 9A	.568	.627	.570 to .697	.512 to .742
CD 9B	.548	.621	.565 to .690	.477 to .764
Province	.782	.812	.738 to .902	.737 to .887

^aFor the predicted 1977 SW/SG ratio presented here to be within ± 10 percent of actual, the actual 1977 value must fall within this range.

price variable. However, except in CRD 30 where wheat acreage did not exceed 1000 acres, the R^2 values exceeded 0.83, indicating that the models explained a large part of the historical variation in SW/SG. When evaluating the results, it should be noted that almost 90 percent of all spring wheat

acreage is in CRD's 10 and 40. Again, the signs of the wheat allotment minus the diverted acreage variable were inconsistent among CRD's. It is worth noting that the barley allotment used on state data in preliminary model analysis showed a highly significant relationship to the SW/SG ratio. However, data

TABLE XI.—North Dakota: The Models Chosen to Predict the SW/SG Harvested Acreage Ratios for Each CRD

[Based in part on ref. 1, table 13]

CRD	Ordinary least-squares estimates of the model coefficients				R ² (a)
	Constant (a)	PWO _j ^d (a)	WAD _{ij} ^d	D71 ^d	
10	0.635 (15.24)	8.25E-4 (4.50)	6.03E-6 (1.31)	0.183 (2.86) ^b	78.7
20	.491 (21.73)	7.24E-4 (7.36)	1.06E-5 (3.23) ^a	.171 (5.06) ^a	89.2
30	.509 (12.29)	6.48E-4 (3.63)	5.30E-7 (.13)	.145 (2.38) ^b	81.0
40	.565 (20.86)	7.77E-4 (6.55)	1.45E-5 (3.00) ^b	.168 (4.09) ^a	85.7
50	.475 (12.49)	8.86E-4 (5.38)	-1.38E-6 (-.21)	.131 (2.33) ^b	90.6
60	.405 (9.49)	8.87E-4 (4.80)	-7.81E-6 (-.91)	.115 (1.82) ^c	90.6
70	.514 (24.28)	9.91E-4 (10.80)	3.03E-5 (9.39) ^a	.253 (8.09) ^a	92.4
80	.449 (14.81)	8.45E-4 (6.38)	-1.91E-5 (2.71) ^b	.165 (3.62) ^a	85.7
90	.337 (9.59)	1.02E-3 (6.67)	5.24E-6 (.75)	.101 (1.95) ^c	91.9

^aSignificant at the 1-percent level

^bSignificant at the 5-percent level.

^cSignificant at the 10-percent level.

TABLE XII.—Minnesota: The Models Chosen to Predict the SW/SG Harvested Acreage Ratios for Each CRD

[Based on ref. 1, table 16]

CRD	Ordinary least-squares estimates of the model coefficients								R ² (a)
	Constant	PWO _j ^d	WAD _{ij} ^d	SH _{ij} ^d /SG _{ij} ^d					
10	0.324 (4.15) ^a	1.21E-3 (2.92) ^b	-6.30E-6 (-0.65)	-2.07E-2 (-0.07)					94.1
20	7.81E-2 (2.28) ^c	6.57E-4 (1.96) ^c	1.95E-4 (.23)	.118 (.35)					83.5
30	4.49E-2 (2.49) ^b	3.09E-4 (2.30) ^c	2.83E-2 (1.67)	-.304 (-.67)					^b 61.8
40	.113 (2.88) ^b	8.42E-4 (1.68)	-1.48E-5 (-.68)	.476 (1.59)					97.4
50	-4.37E-2 (-2.38) ^b	1.05E-3 (5.62) ^a	2.20E-4 (3.00) ^b	.426 (2.75) ^b					98.3
60	-8.02E-3 (-2.00) ^c	2.97E-4 (7.81) ^a	1.82E-3 (3.77) ^a	.274 (2.31) ^c					98.6
70	-2.66E-2 (-.81)	-3.09E-4 (-.74)	-1.65E-5 (-.07)	2.86 (3.72) ^a					90.6
80	-7.91E-2 (-1.62)	1.50E-3 (3.52) ^a	2.71E-4 (2.28) ^c	.463 (1.58)					91.5
90	-1.60E-2 (-1.19)	5.58E-4 (3.95) ^a	2.27E-4 (2.02) ^c	.373 (1.55)					95.8

^aSignificant at the 1-percent level

^bSignificant at the 5-percent level.

^cSignificant at the 10-percent level.

on barley allotments were not available at the CRD level (ref. 1).

South Dakota CRD models.— The SW/SG model equation for harvested acreage was

$$SW_{i,j}^t / SG_{i,j}^t = \beta_0 + \beta_1 PWO_j^t + \beta_2 WAD_{i,j}^t + \beta_3 D65-70^t + \beta_4 D71^t$$

where D65-70^t is a dummy variable equal to 1 for the years 1965 to 1970 and 0 otherwise, and D71^t is a dummy variable equal to 1 for 1971 and 0 otherwise. The dummy variable D65-70^t is thought to represent the effects of a change in government programs which allowed increased substitution between wheat and feed-grain plantings (table XIII). The negative sign (in most CRD's on D65-70^t) indicates a shift out of wheat acreage during this period. D71^t was included in part to account for a large 1-year decline in barley acreage in CRD 10. Although increasing the R² value for this CRD, D71^t had a small and inconsistent impact in other CRD's and probably should be dropped from the model. The 3-year moving average of the difference between wheat and oats prices was the most significant variable in the model. The R² values ranged from 0.46 in CRD 40 to 0.93 in CRD 60.

The WW/WG ratio model consisted of the following equation.

$$WW_{i,j}^t / WG_{i,j}^t = \beta_0 + \beta_1 (WW_{i,j}^{t-1} / WG_{i,j}^{t-1}) + \beta_2 PWWR_{i,j}^t + \beta_3 D67^t$$

where WW_{i,j}^t = area in winter wheat (hundreds of acres) for year *t* in CRD *i* of state *j*
 WG_{i,j}^t = area in winter wheat plus rye (hundreds of acres) for year *t* in CRD *i* of state *j*
 PWWR_{i,j}^t = a 3-year weighted moving average of the difference between the state season average price (cents per bushel) received by farmers for winter wheat (PWW^t) and that received for winter rye (PR_j^t); that is,

$$PWWR_j^t = 0.5(PWW_j^{t-1} - PR_j^{t-1}) + 0.3(PWW_j^{t-2} - PR_j^{t-2}) + 0.2(PWW_j^{t-3} - PR_j^{t-3})$$

TABLE XIII.—South Dakota: The Models Chosen to Predict the SW/SG Harvested Acreage Ratios for Each CRD

[Based on ref. 1, table 19]

CRD	Ordinary least-squares estimates of the model coefficients					R ²
	Constant	PWO _j ^t	WAD _{i,j} ^t	D65-70 ^t	D71 ^t	
10	0.471 (11.89) ^a	9.38E-4 (4.89) ^a	7.15E-5 (4.22) ^a	-5.58E-2 (-1.39)	0.127 (2.53) ^b	.474.2
20	.377 (9.94) ^a	1.16E-3 (6.38) ^a	2.56E-5 (4.06) ^a	-5.07E-2 (-1.39)	4.20E-2 (.91)	.83.8
30	.173 (4.69) ^a	1.10E-3 (6.20) ^a	3.00E-5 (1.24)	-3.45E-2 (-.82)	1.96E-3 (.04)	.90.9
40	.190 (3.56) ^a	5.79E-4 (2.26) ^c	6.33E-5 (1.23)	-3.93E-2 (-.73)	5.51E-2 (.84)	46.3
50	.235 (4.11) ^a	1.28E-3 (4.66) ^a	6.25E-5 (2.14) ^c	-3.55E-2 (-.58)	7.10E-2 (.30)	.75.8
60	4.45E-2 (6.07) ^a	2.89E-4 (7.67) ^a	5.09E-5 (1.32)	-1.62E-2 (-1.30)	-1.23E-2 (-1.10)	.92.9
70	6.37E-2 (2.30) ^b	1.05E-3 (6.87) ^a	3.67E-6 (.59)	2.30E-2 (.65)	-7.80E-2 (-1.55)	.88.6
80	6.27E-2 (1.04)	1.12E-3 (3.85) ^a	4.18E-5 (1.10)	-2.07E-2 (-.35)	-1.38E-2 (-.18)	.876.1
90	5.36E-2 (7.10) ^a	1.16E-4 (2.77) ^b	1.82E-6 (1.07)	-2.70E-3 (.28)	-2.14E-2 (-1.56)	.63.9

^aSignificant at the 1-percent level

^bSignificant at the 5-percent level

^cSignificant at the 10-percent level

Prices are assumed to be zero before 1971.

D67' = a dummy variable equal to 1 in 1967 and 0 otherwise

The model coefficients for harvested acreage are shown in table XIV. The signs of the coefficients are all positive as theoretically expected. Most of the winter rye acreage is in CRD's 20 and 30. The high R^2 for these two CRD's indicates that a large amount of the historical variation was explained by the models. The economic variable $PWWR'_{i,j}$ had a highly significant t-value in these CRD's.

A model for the harvested acreage ratio of winter wheat to total grains (WW/GR) was also developed for South Dakota. The equation follows.

$$WW'_{i,j}/GR'_{i,j} = \beta_0 + \beta_1(WW'_{i,j}{}^{-1}/GR'_{i,j}{}^{-1}) + \beta_2PWWR'_{i,j} + \beta_3D67'$$

where $PWWR'_{i,j}$ = a 3-year weighted moving average of the state season average winter wheat price (cents per bushel) received by farmers; that is,

$$PWWR'_{i,j} = 0.5PWWR'_{i,j}{}^{-2} + 0.3PWWR'_{i,j}{}^{-3} + 0.2PWWR'_{i,j}{}^{-4}$$

Prices were assumed to be zero prior to the 1971-72 marketing year.

The model coefficients for harvested acreage ratios are shown in table XV. Although the signs of the model coefficients were all theoretically correct, the statistical results were inconsistent among the CRD's. While the R^2 value for CRD 30 (a CRD with little wheat acreage) showed that the model explained 90 percent of the historical variation in the WG/GR ratio; the R^2 value for CRD 50 (a CRD with one of the largest winter wheat acreages) showed that the model explained only 35 percent of the historical variation in the WW/WG ratio. The reasons for this poor performance were not understood and need further investigation.

Montana CRD models.—The model equations for

predicting the SW/GR ratios were of the following form.

$$SW'_{i,j}/GR'_{i,j} = \beta_0 + \beta_1PSWB'_{i,j}{}^{-1} + \beta_2D69-71' + \beta_3WAD'_{i,j}$$

where

$PSWB'_{i,j}{}^{-1}$ = the spring wheat price (cents per bushel) minus the barley price (cents per bushel) lagged 1 year. Prices are season average prices received by farmers and are assumed to be zero before 1972.

D69-71' = a dummy variable equal to -1 for the year 1969, 1 for 1971, and 0 otherwise

The coefficients for the seven Montana CRD SW/SG models for harvested acreage are shown in table XVI. Of Montana's spring wheat acreage in 1976, 58 percent was grown in CRD 30 and another 31 percent was grown in CRD 20 (ref. 5). Except in CRD 70, the signs of the model coefficients are all theoretically correct and their t-values are mostly significant. The negative sign on $WAD'_{i,j}$ in CRD 70 is possibly related to the fact that spring wheat acreage accounts for about 20 percent of all wheat acreage in Montana. Winter wheat is a major crop in Montana, and, in all CRD's, harvested winter wheat acreages were more than twice the spring wheat acreages harvested in 1976. Perhaps a basic problem with the SW/SG models is a lack of a variable accounting for winter wheat planting decisions. Also, spring wheat plantings may depend partly on the amount of winterkill in winter wheat. Data to estimate winterkill were not available. The R^2 values, which are typically smaller in Montana than in other states, also indicate that an important independent variable may be missing from the models.

Because Montana winter rye acreage was not reported by CRD, the winter-wheat/total-grain model was the only winter wheat model developed. The model equation for WW/GR follows.

$$WW'_{i,j}/GR'_{i,j} = \beta_0 + \beta_1WW'_{i,j}{}^{-2} + \beta_2WAD'_{i,j} + \beta_3D67-68'$$

where

WW_j^{t-2} = the season average winter wheat price (cents per bushel) received by farmers, lagged 2 years. Since t represents the marketing year which begins when the crop is harvested and the crop is planted the previous fall, the price must be lagged two periods.

D67-68' = a dummy variable equal to 1 for 1967 and 1968 and 0 otherwise

The CRD model coefficients for WW/GR ratios for harvested acreage are shown in table XVII. Model coefficients have the correct sign in all CRD's. However, the R^2 value for CRD 20, where about one-half of Montana's winter wheat is grown, was disappointing. The model explained only 57 percent of the historical variation in the WW/GR ratio.

Accuracy test of 1976 CRD predictions.—A paired difference test, similar to the one run on the Canadian CD data and described in the section entitled "Precision and Accuracy Tests," was made between the CRD predictions based on the model for 1976 and those based on last year's ratio. The test results for the 1976 CRD-harvested acreages are shown in table XVIII.

The results show that the ratio values predicted by the model were closer to the actual 1976 ratio values

than were the lagged (1975) ratio values at the 10-percent (or higher) significance level in North and South Dakota, but not in Minnesota and Montana. Nevertheless, in the latter two states, the model predictions were generally closer to the actual value than were the lagged values.

Model predictions for 1977.—The 1977 ratio predictions, the associated variance of estimates, and the associated variance of prediction are shown in table XIX. In contrast to the Canadian models, which predicted that the ratio of SW/SG acreage would increase in 1977 from the 1976 values, the U.S. CRD models generally predicted a decline in the 1977 SW/SG ratios from the 1976 values. A major reason for a decline in the SW/SG ratios was a reduction in the 1976-77 price of wheat from the previous year. It should be noted that the only new information needed to make 1977 predictions was the estimated season average prices received by farmers for small grains.

A large jump in the lagged dependent variable in 1976 caused an unrealistic prediction for 1977 in Minnesota CRD 70 (table XIX). The SW/SG ratio increased from 0.161 in 1975 (and less than 0.12 in earlier years) to 0.406 in 1976. CRD 70 accounted for 5.4 percent of the total Minnesota spring wheat acreage in 1976, up from 2.2 percent in 1975. Consequently, the rather large predicted ratio for CRD 70

TABLE XIV.—South Dakota: The Models Chosen to Predict the WW/GR Harvested Acreage Ratios for Each CRD

[Based on ref. 1, table 23]

CRD	Ordinary least-squares estimates of the model coefficients								R^2
	Constant		WW_{ij}^t / WG_{ij}^t		$D67^t$		PWW_{ij}^t		
10	0.435	(2.51) ^b	0.503	(2.58) ^b	0.124	(3.33) ^a	4.32E-4	(2.54) ^b	^a 68.3
20	5.28E-2	(1.09)	.612	(3.92) ^a	.216	(3.80) ^a	1.57E-3	(5.43) ^a	^a 88.4
30	-2.31E-2	(-1.35)	1.146	(4.90) ^a	6.67E-2	(1.73)	1.25E-3	(3.34) ^a	^a 96.1
40	.587	(2.79) ^b	.406	(1.90) ^c	7.92E-3	(.94)	2.36E-5	(.57)	36.1
50	.380	(1.92) ^c	.477	(1.84) ^c	.146	(2.31) ^b	8.82E-4	(2.70) ^b	^a 70.1
60	1.29E-2	(.51)	.840	(4.31) ^a	.132	(2.78) ^b	1.87E-3	(4.28) ^a	^a 96.0
70	.680	(5.72) ^a	.297	(2.35) ^b	1.85E-2	(1.18)	5.90E-5	(.70)	^c 47.5
80	.512	(1.94) ^c	.460	(1.64)	3.48E-2	(1.27)	1.72E-4	(1.19)	40.9
90	.440	(3.13) ^b	.454	(2.42) ^b	8.63E-2	(1.78)	5.07E-4	(1.61)	^a 71.5

^aSignificant at the 1-percent level

^bSignificant at the 5-percent level

^cSignificant at the 10-percent level

will not have a significant impact on state results. Nevertheless, in this case, the 1976 ratio would undoubtedly provide a better predictor in CRD 70 than the model value.

All other models provided estimates that appear reasonable. The accuracy of all models can be tested better when 1977 CRD crop acreage estimates become available.

CONCLUSIONS

The results for both the United States and Canada show that econometric models can provide estimates of confusion crop ratios that are more accurate than historical ratios. Whether these models can support the LACIE 90/90 accuracy criterion is uncertain. In the United States, experimenting with additional

TABLE XV.—South Dakota: The Models Chosen to Predict the WWGR Harvested Acreage Ratios for Each CRD

(Based on ref. 1, table 25)

CRD	Ordinary least-squares estimates of the model coefficients								R ²
	Constant		$WW_{ij}^{1-1} / AGR_{ij}^{1-1}$		$D67^A$		PWW_{ij}^1		
10	3.94E-2	(1.69)	0.361	(1.84) ^c	6.94E-2	(2.19) ^c	1.96E-4	(3.86) ^a	^b 66.1
20	2.05E-2	(2.65) ^b	-2.17E-2	(-.09)	1.31E-2	(1.02)	6.78E-5	(3.05) ^b	^c 49.8
30	-6.89E-3	(-1.99) ^c	2.690	(3.04) ^b	7.07E-3	(1.14)	-2.10E-6	(-.08)	^a 82.6
40	.286	(1.36)	.509	(1.34)	.117	(1.78)	4.24E-4	(2.41) ^b	^a 78.8
50	.180	(1.47)	-.162	(-.25)	5.31E-2	(.53)	3.12E-4	(2.24) ^b	35.3
60	-2.27E-3	(-2.25) ^b	1.610	(4.98) ^a	5.32E-3	(2.84) ^b	9.93E-6	(2.08) ^c	^a 90.1
70	.342	(3.23) ^a	.493	(2.99) ^b	.114	(2.97) ^b	2.13E-4	(2.94) ^b	^a 74.8
80	.577	(2.11) ^c	-.154	(-.29)	.135	(1.67)	2.25E-4	(1.58)	32.7
90	1.38E-2	(2.13) ^c	.358	(1.48)	1.25E-2	(3.13) ^b	1.32E-5	(1.89) ^c	^c 52.5

^aSignificant at the 1-percent level.

^bSignificant at the 5-percent level.

^cSignificant at the 10-percent level.

TABLE XVI.—Montana: The Models Chosen to Predict the SWISG Harvested Acreage Ratios for Each CRD

(Based on ref. 1, table 27)

CRD	Ordinary least-squares estimates of the model coefficients								R ²
	Constant (a)		PSW_{ij}^{1-1}		$D69-71^A$		WAD_{ij}^1		
10	0.111	(7.20)	4.15E-4	(4.13) ^a	4.62E-2	(2.70) ^b	1.80E-4	(5.06) ^a	^a 74.5
20	.351	(8.33)	6.30E-4	(2.33) ^b	.174	(3.75) ^a	5.41E-6	(1.39)	^b 77.0
30	.672	(21.75)	6.16E-4	(3.11) ^b	.102	(3.01) ^b	5.70E-6	(1.79)	^b 64.4
50	.195	(.45)	2.42E-4	(.86)	.104	(2.16) ^c	2.69E-6	(.21)	43.5
70	.295	(7.27)	-1.45E-5	(-.08)	9.41E-2	(3.05) ^b	-2.16E-5	(-.50)	^b 62.7
80	6.47E-2	(3.79)	4.13E-4	(3.73) ^a	5.04E-2	(2.69) ^b	3.17E-5	(3.70) ^a	^b 65.4
90	.382	(12.64)	7.56E-4	(3.88) ^a	.119	(3.59) ^a	7.72E-5	(4.75) ^a	^a 73.4

^aSignificant at the 1-percent level.

^bSignificant at the 5-percent level.

^cSignificant at the 10-percent level.

model formulations could provide improved models in some CRD's, particularly for winter wheat. Improved models may also be possible for the Canadian CD's.

The more aggregate province/state models outperformed individual CD/CRD models. This result was expected—partly because acreage statistics are based on sampling procedures and the sampling precision

TABLE XVII.—Montana: The Models Chosen to Predict the WWIGR Harvested Acreage Ratios for Each CRD

[Based on ref. 1, table 31]

CRD	Ordinary least-squares estimates of the model coefficients							R ²
	Constant (a)		WW_j^{1-2}	$D67-68^d$		$WAD_{1,j}^f$		
10	0.301	(15.81)	2.36E-5 (0.32)	-3.73E-2	(-1.53)	2.21E-4 (4.47) ^a	^a 85.1	
20	.266	(3.98)	7.66E-4 (2.98) ^b	7.50E-2	(.92)	1.26E-5 (1.89) ^c	^b 57.7	
30	.114	(6.12)	1.83E-4 (2.55) ^b	4.17E-2	(1.88) ^c	2.05E-6 (1.00)	^b 58.8	
50	.415	(9.23)	4.82E-4 (2.79) ^b	5.84E-2	(1.08)	5.12E-5 (3.52) ^a	^a 74.6	
70	.398	(22.72)	1.25E-4 (1.86) ^c	-1.01E-2	(-.47)	1.64E-4 (5.55) ^a	^a 87.8	
80	.544	(15.46)	5.47E-4 (4.05) ^a	9.38E-2	(2.17) ^c	2.72E-5 (1.41)	^a 75.6	
90	.412	(12.65)	1.67E-4 (1.34)	5.78E-2	(1.46)	2.63E-5 (1.38)	^c 51.2	

^aSignificant at the 1-percent level

^bSignificant at the 5-percent level

^cSignificant at the 10-percent level

TABLE XVIII.—Results of the Paired Comparison Tests of the Difference in Accuracy Between the "Best" Model and Last Year's Ratio as Predictors of the 1976 Harvested Acreage Ratios for CRD's by State and Ratio^a

State	Ratio	Number of CRD's	Mean value of weighted paired differences, \bar{X} (b)	Standard error of paired differences, $S_{\bar{X}}$	Calculated t-value
North Dakota	SW/SG	9	26.2	12.0	^c 2.18
Minnesota	SW/SG	9	92.8	53.0	1.75
South Dakota	SW/SG	9	79.6	31.6	^d 2.52
	WW/WG	9	36.9	15.5	^d 2.38
	WW/GR	9	73.6	18.5	^c 3.98
Montana	SW/SG	7	18.0	15.0	1.20
	WW/GR	7	18.5	14.4	1.28

^aDifferences are based on the absolute values of the percentage error in using (1) the ratio model to predict the actual ratio and (2) last year's ratio to predict the actual ratio
$$h_{i,j,k} = (|d_{i,j,k}^1| - |d_{i,j,k}^2|)W_{i,j}^{-76}$$
 where $d_{i,j,k}^1$ is the difference between the model-predicted ratio value and the actual ratio value for CRD k , state i , and ratio j , $d_{i,j,k}^2$ is the difference between last year's (1975) actual ratio value and the actual (1976) ratio value for CRD k , state i , and ratio j , and $W_{i,j}^{-76}$ = proportion of harvested wheat acreage in CRD k of state i in 1976.

^bIndicates model prediction is better than the lagged (last year's) ratio at the 10-percent level of significance

^cIndicates model prediction is better than the lagged (last year's) ratio at the 5-percent level of significance

^dIndicates model prediction is better than the lagged (last year's) ratio at the 1-percent level of significance

declines from the province/state to the CD/CRD level. Also, CD and CRD data were not always available for predictor (independent) variables such as prices. More aggregate province/state-level observations had to be substituted for the desired CD/CRD data which were not available. Declining sampling precision and the need to substitute province/state data for CD/CRD data introduced measurement error into the CD/CRD models. When

the independent variables are subject to measurement errors, ordinarily least-squares techniques give estimates of the model coefficients that can be both biased and inconsistent.

To minimize operational problems, the same variables were used in all CD or CRD models of a province or state. Using the same set of variables in all CD's/CRD's of a province/state may introduce equation errors in at least some of the CD's/CRD's.

TABLE XIX.—Predicted Confusion Crop Harvested Acreage Ratios for 1977

[Based on ref. 1]

State	Geographic unit	Predicted 1977	Variance of estimate	Variance of prediction
<i>SWISG ratios</i>				
Minnesota	CRD 10	0.553	0.00183	0.00382
	CRD 20	.240	.00168	.00226
	CRD 30	.080	.00029	.00037
	CRD 40	.572	.00115	.00554
	CRD 50	.331	.00041	.00169
	CRD 60	.078	.00002	.00006
	CRD 70	1.073	.00182	.06043
	CRD 80	.491	.00343	.01560
	CRD 90	.176	.00026	.00117
Montana	CRD 10	0.140	0.00046	0.00057
	CRD 20	.395	.00311	.00392
	CRD 30	.715	.00164	.00207
	CRD 50	.212	.00337	.00425
	CRD 70	.204	.00141	.00178
	CRD 80	.094	.00054	.00067
	CRD 90	.435	.00165	.00207
	North Dakota	CRD 10	0.822	0.00237
CRD 20		.655	.00063	.00075
CRD 30		.627	.00316	.00379
CRD 40		.741	.00096	.00115
CRD 50		.676	.00170	.00204
CRD 60		.606	.00218	.00261
CRD 70		.739	.00053	.00064
CRD 80		.641	.00116	.00139
CRD 90		.567	.00146	.00175
South Dakota	CRD 10	0.676	0.00166	0.00202
	CRD 20	.631	.00139	.00169
	CRD 30	.414	.00139	.00169
	CRD 40	.317	.00281	.00341
	CRD 50	.516	.00330	.00401
	CRD 60	.108	.00009	.00011
	CRD 70	.294	.00194	.00235
	CRD 80	.308	.00361	.00438
	CRD 90	.079	.00015	.00018

"Page missing from available version"

4. USDA SRS. Minnesota Agricultural Statistics (St. Paul, Minn.) 1963 to 1977.
5. USDA SRS. Montana Agricultural Statistics (Helena, Mont.) 1963 to 1977.
6. USDA SRS. North Dakota Crop and Livestock Statistics (Fargo, N. Dak.) 1963 to 1977.
7. USDA SRS. South Dakota Agricultural Statistics (Sioux Falls, S. Dak.) 1963 to 1977.
8. Lidman, R.; and Bawden, D. L.: The Impact of Government Programs on Wheat Acreage. *Land Econ.*, vol. 1, no. 4, Nov. 1974, pp. 327-335.
9. Garst, G. D.; and Miller, T. A.: Impact of the Set-Aside Program on U.S. Wheat Acreages. *Agr. Econ. Res.*, vol. 27, no. 2, Apr. 1975, pp. 30-37.
10. Weaver, R.; Morzuch, B.; and Helmberger, P.: Acreage Supply Response: Spring Wheat in the Dakotas. Prepared for the Institute of Environmental Studies, University of Wisconsin (Madison), Jan. 1977.
11. Schmitz, A.: Canadian Wheat Acreage Response. *Canadian J. Agr. Econ.*, vol. 16, 1968, pp. 79-86.
12. Capel, R. F.: Predicting Wheat Acreage in the Prairie Provinces. *Canadian J. Agr. Econ.*, vol. 16, 1968, pp. 87-89.
13. Meilke, K. D.: Acreage Response to Policy Variables in the Prairie Provinces. *American J. Agr. Econ.*, Aug. 1976, pp. 572-576.
14. Meilke, K. D.; and Kramar, R. E.: Acreage Response in Ontario. *Canadian J. Agr. Econ.*, vol. 24, no. 1, 1976, pp. 51-66.
15. Canadian Wheat Board Annual Reports. Canadian Wheat Board, Winnipeg, Manitoba, 1949 to 1977.
16. Jolly, R. W.: An Econometric Analysis of the Grain Livestock Economy in Canada With Special Emphasis on Commercial Agricultural Policy. Ph. D. dissertation, University of Minnesota, Mar. 1976.
17. Rasmussen, W. D.; et al.: A Short History of Agricultural Adjustment 1933-75. USDA/SRS Economic Research Service. *Agr. Inf. Bull.* 391 (Washington, D.C.), Mar. 1976.
18. Department of Agriculture. Acreage, Yield and Production of Principal Grain Crops—Saskatchewan. Annual Reports of the Government of the Province of Saskatchewan (Regina, Sask.), 1950 to 1976.
19. Statistics Canada. Field Crop Reporting Series—No. 20. Catalogue 22-002, annual reports.
20. Statistics Canada. Grain Statistics Weekly. Catalogue 22-004, annual reports.
21. Statistics Canada. Grain Trade of Canada. Catalogue 22-201, annual reports.
22. Statistics Canada. The Wheat Review. Catalogue 22-005, annual reports.
23. USDA ASCS. Minnesota ASCS Annual Report (St. Paul, Minn.) 1966 to 1973.
24. USDA ASCS. Montana ASCS Annual Report (Bozeman, Mont.) 1964 to 1973.
25. USDA ASCS. North Dakota ASCS Annual Report (Fargo, N. Dak.) 1963 to 1973.
26. USDA ASCS. South Dakota ASCS Annual Report (Huron, S. Dak.) 1963 to 1973.
27. USDA ESCS. Grain Stocks (Washington, D.C.), 1962 to 1977.
28. Draper, N. R.; and Smith, H.: Applied Regression Analysis. John Wiley and Sons, Inc. (New York), 1966.
29. Johnston, J.: Econometric Methods, 2nd edition. McGraw-Hill Book Co. (New York), 1972.
30. Nerlove, M.: Distributed Lags and Demand Analysis for Agricultural and Other Commodities. USDA Agricultural Handbook 141, 1958.

LACIE Sampling Design

A. H. Feiveson,^a R. S. Chhikara,^b and C. R. Hallum^a

INTRODUCTION

The sampling design used in LACIE consisted of two major components, one for wheat acreage estimation and one for wheat yield prediction. The acreage design was basically a classical survey for which the sampling unit was a 5- by 6-nautical-mile segment; however, there were complications caused by measurement errors and loss of data. Yield was predicted by sampling meteorological data from weather stations within a region and then using those data as input to a previously fitted regression equation. Most of the discussion in this paper refers to the acreage sampling design, since there was considerably more freedom in planning for the collection of Landsat data (used for acreage estimation) than there was for the collection of meteorological data, a situation in which one was forced to make use of what was currently available. Wheat production was not estimated directly; instead, it was computed by multiplying yield and acreage estimates (see the paper by Chhikara and Feiveson entitled "Large-Area Aggregation and Mean-Squared Prediction Error Estimation for LACIE Yield and Production Forecasts").

ACREAGE ESTIMATION SAMPLING DESIGN

Determination of Sampling Units and Frame

All information on current-year wheat acreage was obtained through Landsat imagery of a number of 5- by 6-nautical-mile segments. These segments

were the basic sample units for acreage estimation and were distributed throughout the wheat growing regions of LACIE countries.

Because of various data base engineering constraints, a maximum of 4800 sample segments could be processed within a crop year, regardless of the size of the individual segment. Given the maximum sample size of 4800, the physical dimensions of 5 by 6 nautical miles for sample segments were decided on as large enough for Classification and Mensuration Subsystem (CAMS) analysts to obtain wheat acreage estimates and small enough to not tax computer and manpower resources. Throughout this paper, the term "sample segment" refers to 5- by 6-nautical-mile segments actually in the LACIE sample, whereas "segment" refers to any 5- by 6-nautical-mile area whether or not in the sample.

The LACIE sampling frame was constructed by first covering the wheat growing regions of a country by a large grid of 5- by 6-nautical-mile segments and then excluding those segments which appeared to have less than 5 percent agriculture, as determined by an examination of previous years' Landsat imagery. The remaining segments constituted the frame from which the actual sample segments were chosen.

Allocation of Samples to Countries

In the early years of LACIE, it was decided to allocate the maximum 4800 sample segments to 8 major wheat producing countries in proportion to their most recent wheat production statistics. Two types of sampling strategies were used in LACIE—one for countries with historical wheat data on a detailed level (D) and one for countries with published historical data only for fairly large political subdivisions (N). In table I, the eight LACIE countries, their smallest political subdivision (SPD) for which published historical data exist, and the number of samples in the initial allocation are listed.

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Houston, Texas.

TABLE I.—LACIE Countries

Country	SPD with published wheat data	Level (a)	Number of segments in initial allocation
United States	County	D	637
U.S.S.R.	Oblast	N	1949
China	Province	N	810
Canada	Crop subdistrict	D	283
India	State	N	626
Australia	Shire	D	257
Argentina	Partido	D	165
Brazil	State	N	47
Total			4774

^aD = "detailed" sampling strategy; N = "non-detailed" sampling strategy.

Definition of Strata

In level N countries, sample segments are allocated at random within strata which are approximately the intersection of SPD's with the sampling frame. More precisely, for a given SPD, the corresponding stratum consists of all 5- by 6-nautical-mile

segments which are in the sampling frame and the center points of which lie in the SPD (fig. 1). As a consequence, the agricultural area in each SPD (and hence the whole country) is approximated by the collection of 5- by 6-nautical-mile segments from which samples are drawn. At the country level, the error in the approximation is negligible; however, when SPD's are small, adjustments must be made to the wheat acreage estimate for a stratum to obtain a more precise estimate for the corresponding SPD (see the paper by Chhikara and Feiveson entitled "LACIE Large-Area Acreage Estimation").

In level D countries, each SPD is also approximated by the collection of 5- by 6-nautical-mile segments which lie in the sampling frame and the center points of which lie in the SPD; however, the collection in this case is called a "substratum" rather than a "stratum" because, in some cases, no sample segment is selected from it. To distinguish between an SPD and its approximating collection of segments, the latter will be called a pseudo-SPD (PSPD).

Strata in level D countries are defined to be the union of PSPD's which are contained in the next higher political subdivision of the country. For example, in the United States, where an SPD is a county and the next higher political subdivision is a crop reporting district (CRD) (within a state), the stratum

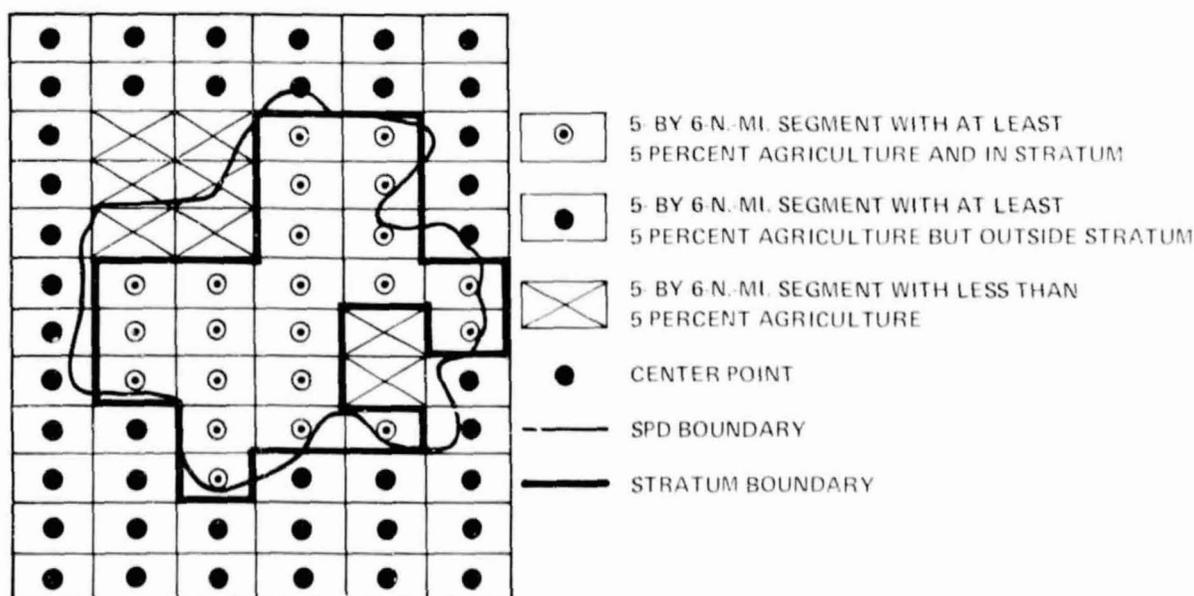


FIGURE 1.—Smallest political subdivision and corresponding stratum in level N countries.

consists of the collection of all "pseudo" counties the corresponding counties of which lie within that CRD.

Allocation and Selection of Segments in Strata/Substrata

In the first 2 years of LACIE, sample sizes for countries were fixed as shown in table I. Since little or nothing was known about the accuracy of yield predictions at that time, it was decided to allocate samples to strata (level N countries) or substrata (level D countries) so as to minimize the best a priori estimate of the variance of the country wheat acreage estimate.

It is well known (cf. Chhikara and Feiveson, "Large-Area Aggregation and Mean-Squared Prediction Error Estimation for LACIE Yield and Production Forecasts") that if one estimates a population total by stratified sampling over L strata with a total sample size of n , the variance (ignoring the finite population correction) of the estimate is minimized if n_k , the sample size for the k th stratum, is proportional to $N_k S_k$, where N_k is the total number of segments in the k th stratum from which n_k samples were selected at random, and S_k is the standard deviation of the segment "characteristics" (in this case, wheat acreages) within the k th stratum. This fact was used in LACIE to obtain allocations to strata in nondetailed countries, where N_k was the number of segments comprising the k th stratum and S_k^2 was assumed proportional to the "binomial" variance $P_k(1 - P_k)$, where P_k was the historical proportion of wheat in the SPD corresponding to the k th stratum. Note that it was not necessary to know the constant of proportionality; i.e., if $S_k \propto P_k^{1/2}(1 - P_k)^{1/2}$, the optimal sample size for the k th stratum would be given by

$$t_k = \frac{n N_k P_k^{1/2} (1 - P_k)^{1/2}}{\sum_{k'} N_{k'} P_{k'}^{1/2} (1 - P_{k'})^{1/2}} \quad (1)$$

except that, in general, t_k would not be an integer. For Phases I and II of LACIE, n_k was taken to be the nearest integer to t_k . In level N countries, this rounding could be done with little effect since the value of

t_k tended to be rather large (between 10 and 50). Once n_k was computed, n_k sample segments were selected at random from the N_k segments comprising the k th stratum.

In level D countries, an attempt to use the preceding technique in substrata would produce many values of t_k of less than 1 or between 1 and 2. As a result, the sample sizes t_k as computed in equation (1) were used to categorize substrata into three groups: Group I, $t_k > 1.0$; Group II, $0.1 \leq t_k \leq 1.0$; and Group III, $t_k < 0.1$.

The substrata in Group I received N_k sample segments, selected at random, where n_k was t_k rounded to the nearest integer. All Group II substrata within a stratum were called a "Group II collection." Each entire collection received an allocation of segments equal to the rounded total of t_k within the collection. For example, in the United States, if there were three Group II pseudocounties (substrata) in a pseudo-CRD (stratum) with respective t_k values of 0.7, 0.6, and 0.5, then the collection of three pseudocounties would receive a total of 2 (rounded value of $0.7 + 0.6 + 0.5$) sample segments. Once the sample size, say m , had been determined for a Group II collection consisting of M substrata, the sample segments were chosen with a two-stage sampling scheme where, in the first stage, m substrata were selected at random with probabilities proportional to their historical wheat acreage, then, in the second stage, one sample segment was selected at random within each of the m chosen substrata (note that $m \leq M$).

The Group III substrata were those that would hypothetically receive less than a tenth of a sample segment in the optimal allocation and thus were not sampled at all. Their wheat acreage was instead estimated by first computing a historical ratio of their wheat acreage to that of neighboring Group I and Group II substrata and then applying that ratio to the current-year estimate for the neighboring Group I and II substrata. (For details, see the paper by Chhikara and Feiveson entitled "Large-Area Aggregation and Mean-Squared Prediction Error Estimation for LACIE Yield and Production Forecasts.")

For Phase III of LACIE, some modifications were made to the allocation procedure. Instead of assuming that within-stratum wheat variances were proportional to the binomial $P(1 - P)$, where P is the historical proportion of wheat in the stratum, it was decided that a better approximation would be to assume that the wheat variance is proportional to the

small-grain¹ variance, which could be directly estimated from a regression model using Landsat imagery from recent years. The advantage of this procedure lies in the ability of analysts to examine a Landsat full-frame color image and to obtain crude estimates of small-grains (but not wheat alone) proportions for all 5- by 6-nautical-mile segments within the area covered by the image.

It is not feasible to use this capability to estimate small-grains variances for all strata/substrata because it is a very time-consuming process and also because appropriate data acquisition dates may not exist in all areas. It is, however, possible to use full-frame imagery to estimate the proportion of agriculture for every segment in the sampling frame and then establish a regression model approximately expressing the small-grains within-stratum variance as a function of the agriculture variance, proportion of agriculture, and historical proportion of small grains in the stratum. Since all the preceding determinations employ data not from the current year but from recent years, there is an implicit assumption that within-stratum small-grains variances are about the same from year to year, at least for the purposes of sample allocation.

The regression model, also known as the "magic formula," was developed as follows. Let a_i and g_i be the respective agriculture and small-grains proportions in the i th segment of a stratum/substratum. Then, if $r_i = g_i/a_i$, one can write

$$g_i = a_i r_i. \quad (2)$$

Within a region where cropping practices are about the same, it is not unreasonable to expect that a_i and r_i are independent; i.e., knowledge about the amount of agriculture in a segment provides almost no information about the ratio of small grains to agriculture in that segment. (Note that for all segments in the sampling frame, $a_i \geq 0.05$; therefore, r_i is always well defined.) Under independence, one can write

$$\text{Var}(g_i) = E(a_i^2) \text{Var}(r_i) + E^2(r_i) \text{Var}(a_i) \quad (3)$$

¹The term "small grains" refers to the combined crops of wheat, barley, oats, and rye.

Note that $E(a_i^2)$ and $\text{Var}(a_i)$ are directly estimable from knowledge of a_i for all segments. The quantity $E(r_i)$, although not known precisely, can be approximated by $r_i = g_i/a_i$, where g_i is the historical small-grains proportion in the stratum and $a_i = E(a_i)$. Although $\text{Var}(r_i)$ is unknown, it is assumed that it can be approximated by $cr_i(1 - r_i)$, where c is a coefficient between 0 and 1. Under all the preceding assumptions, equation (3) can be written

$$\sigma_g^2 = cr(1 - r) [\alpha_2] + r^2 \sigma_a^2, \quad (4)$$

where $\sigma_g^2 = \text{Var}(g_i)$, $\alpha_2 = E(a_i^2)$, and $\sigma_a^2 = \text{Var}(a_i)$.

In the United States, c was estimated by (1) using full-frame imagery to estimate g_i directly and thus compute σ_g^2 for all segments within 40 selected counties (i.e., substrata) and (2) regressing the quantities $(\sigma_g^2 - r^2 \sigma_a^2)/\alpha_2$ against $r(1 - r)$ over the 40 counties to obtain c . It was found that this procedure gave a better fit than that obtained by regressing σ_g^2 against $g(1 - g)$.

Once c was determined, $a_i \alpha_2$ and r were computed for all substrata and σ_g^2 was estimated using equation (4). Finally, the sample sizes t_k were computed using the estimate of σ_g^2 instead of $P_k(1 - P_k)$ in equation (1). In other countries, a similar procedure is used to estimate σ_g^2 ; however, where strata are much larger than U.S. counties, the assumptions which led to equation (4) are more likely to be false.

For level D countries in Phase III, the definition of Group III was changed to be the set S of all substrata such that (1) the total historical wheat acreage for the substrata in S was approximately 2.5 percent of the country's historical wheat acreage and (2) if S_1 and S_2 were substrata such that $S_1 \in S$ and $S_2 \notin S$, then S_2 had (historically) more wheat than S_1 . The values of t_k in equation (1) were then computed only for the substrata remaining after the elimination of those designated as Group III.

Another modification in Phase III was that the values of t_k were "probabilistically" rounded to integers (n_k) in the sense that if $t_k = m + r$, where m is an integer and $0 < r < 1$, then n_k was randomly set equal to m (with probability $1 - r$) or $m + 1$ (with probability r). This revised rounding procedure made the total sample size much closer to n than did the old procedure.

Finally, in Phase III, rather than allocate 4800 sample segments to 8 countries in proportion to their

production, it was suggested that LACIE should estimate for each country the sample size needed to satisfy a given accuracy criterion and use these sample sizes as long as the total was less than 4800. This procedure was followed in the United States and the U.S.S.R. by specifying a desired coefficient of variation (CV) for the production estimate of each country, then calculating the sample size necessary to achieve the CV, given errors due to (1) sampling, (2) classification, (3) yield prediction, and (4) loss of data. Using the best available a priori estimates of the magnitude of the errors, one can approximate the variance of the production estimate as a function of the total sample size n using the optimal allocation strategy. The equation can then be solved for n . The resulting expression is given by

$$n = \frac{\left[\sum_{j=1}^L l_j \sum_{k=1}^{L_j} \sqrt{N_{jk} \Theta_{jk}^2 (Y_j^2 + T_j^2)} \right]^2}{\delta \left[P^2 CV^2(P) + \sum_{j=1}^L \sum_{k=1}^{L_j} N_{jk} \Theta_{jk}^2 (Y_j^2 + T_j^2) - \sum_{j=1}^L A_j^2 T_j^2 \right]} \quad (5)$$

where n = the total number of sample segments allocated to the area of interest
 N_{jk} = the total number of agriculture segments in the k th substratum/stratum in the j th yield stratum²
 Θ_{jk}^2 = estimate of segment-to-segment variance of the estimated small-grains area within the k th substratum/stratum of the j th yield stratum, which is the sum of sampling and classification components. The sampling component comes from equation (4), whereas the classification variance is an estimate obtained from previous testing of classification procedures.
 Y_j = average yield of the j th yield stratum over the most recent 2 to 3 years (Obtained from the LACIE yield models if available; otherwise, obtained from historical information. If neither of these is

²A yield stratum is an area for which wheat yield is assumed to be constant. In current usage, a yield stratum is a union of acreage strata.

available, Y_j may be obtained from soil characteristic maps overlaid on Landsat imagery of the area of interest.)

T_j = estimate of the standard deviation of the yield estimate in the j th yield stratum

L = the total number of yield strata in the area of interest

L_j = the total number of substrata/strata in the j th yield stratum

$CV(P)$ = preassigned value³ of the CV of the production estimate

P = estimate of the total wheat production in the country/area of interest based on historical data

A_j = estimate of wheat area in the j th yield stratum based on historical information

δ = a conservative lower-bound estimate of the expected sample acquisition rate (determined from previous experience with loss of segments due to cloud cover or other factors)

After determining the total number n of segments to be allocated to the area of interest, the optimal sample size n_{jk} to be associated with the k th substratum/stratum in the j th yield stratum is defined by

$$n_{jk} = \frac{[P] n N_{jk} \sqrt{\Theta_{jk}^2 (Y_j^2 + T_j^2)}}{\sum_{j=1}^L \sum_{k=1}^{L_j} N_{jk} \sqrt{\Theta_{jk}^2 (Y_j^2 + T_j^2)}} \quad (6)$$

where $[P]$ denotes probabilistic rounding to the integer either above or below.

³The choice for the preassigned value for the production coefficient of variation is dependent on the desired accuracy of the production estimate for the area of interest. In this case, accuracy is measured with respect to the probability space (Ω, β, P) where P operates on β , the σ -field of Lebesgue measurable subsets of Ω , the set of percentage deviations from true production. For example, if a country estimate is to be made, the goal is to obtain a country production estimate which is within 10 percent of the actual production with a probability of at least 0.9.

Using these revised procedures, the allocation for Phase III was 601 segments for the United States and 1111 for the U.S.S.R. In the United States, wheat area was estimated for a total of 855 counties; 288 were in Group I; 164 were in Group II; and 403 were in Group III.

YIELD SAMPLING DESIGN

In LACIE, wheat yield was not measured or estimated on a point or segment basis. The average yield for a large area called a "yield stratum" was estimated by collecting meteorological information from existing weather stations scattered throughout the stratum and using the average weather as input to a yield prediction regression equation. These yield strata were typically larger than acreage strata, especially in level D countries. For example, in the

United States, yield strata were about the size of states and were delineated such that they were relatively homogeneous with respect to climate, soil type, and other factors which could affect wheat yield. (For details, see the paper by McCrary et al. entitled "Operation of the Yield Estimation Subsystem.")

PRODUCTION

Wheat production was not sampled or estimated directly but was computed by multiplying yield estimates by acreage estimates. (For details, see the paper by Chhikara and Feiveson entitled "Large-Area Aggregation and Mean-Squared Prediction Error Estimation for LACIE Yield and Production Forecasts.")

LACIE Area Sampling Frame and Sample Selection

C. J. Liszcz^a

INTRODUCTION

Early in the design phase of LACIE, it was decided not to use a completely enumerated area for developing the sampling frame because much wasteful (totally nonagricultural) information would be included. The alternative was to delineate all the agricultural and nonagricultural areas and to locate the sample segments in the agricultural areas. An agricultural area was defined on full-frame Landsat color-infrared (CIR) imagery as an area having discernible field patterns. For those sections of a country for which there was no imagery, operational navigation charts (ONC's) were used, and only those areas that were definitely nonagricultural (e.g., mountains, deserts, tundras, and lakes) were eliminated as potential sample segments. In other parts of the LACIE countries, some imagery was unusable because of cloud or snow cover; both these situations required the use of ONC's.

In later LACIE phases, most of the agricultural and nonagricultural delineations were accomplished using full-frame CIR imagery. This procedure necessitated moving some segments from what was potentially agricultural on the ONC's to "real" agricultural as interpreted on the CIR imagery.

This agricultural area, composed of 5- by 6-nautical-mile segments for each country's smallest political unit, was the area's sampling frame. The number of these segments for each country's smallest political unit was the pseudocounty. This terminology was used for all countries even though the smallest political units of some, such as the U.S.S.R. and the People's Republic of China (P.R.C.), were very large and were oblasts, states, etc., instead of counties.

Constraints were imposed on sample segment locations because of the computer hardware available for selecting or stripping the sample segment data from full-frame data. This limited the number

of segments that could be obtained per frame in the east-west and north-south directions.

An additional constraint was imposed, by the gathering, sorting, and stripping agency, to verify the actual acquired segment location and the registration of subsequent acquisitions of that segment. The center point of a segment's first acquisition was specified as being within 2 nautical miles in the north-south direction and within 3 nautical miles in the east-west direction. The location and registration of subsequent acquisitions of the same segment were specified as being within one picture element (pixel) of the first acquisition. To guarantee these specifications, a rectangle of data (approximately 10 by 12.5 nautical miles) enclosing each sample segment had to be stripped from the Landsat full-frame image (100 by 100 nautical miles). To ensure there was no overlap (a hardware requirement), the center points of all sample segments had to be separated by a minimum of 10 nautical miles north-south and 12 nautical miles east-west.

In some instances, primarily in those countries with small political units, this minimum-distance constraint made it impossible to place the allocated number of sample segments in the political area. Fewer segments were thus placed in those countries.

The procedures for the agricultural delineation and sample segment location in a LACIE country are discussed in this paper. These procedures also contain the description of required materials and the definition of terms peculiar to LACIE.

DETERMINATION OF AREA SAMPLING FRAME

Transparent acetate overlays containing agricultural boundaries within each of the LACIE countries were prepared and registered to ONC's. Full-frame Landsat CIR images of the same scale as the ONC's (1:1 million) were used to identify agricultural boundaries based on discernible agricultural field patterns.

^aLockheed Electronics Company, Houston, Texas.

Preliminary Steps

The following preliminary steps were taken in the preparation of the base map overlay.

1. ONC base maps (1:1 million scale) were obtained through the NASA LACIE Physical Data Library (LPDL). All ONC's obtained had to be of the same series and publishing date to simplify the registration of the completed overlays.

2. The availability of ONC base physical feature overlays from the Aeronautical Chart and Information Command (ACIC) was checked. When available, these base overlays should have matched the series and date of the ONC map used in the agricultural and nonagricultural delineations. It was desirable that these physical feature overlays be in some color other than black because of the black lines already on the ONC.

3. The Landsat CIR imagery available from the LPDL and the Classification and Mensuration Subsystem (CAMS) was researched. Landsat CIR imagery of 9 by 9 inches was used to construct the agricultural and nonagricultural overlays. The regional analyst determined what imagery was available to complete the agricultural and nonagricultural tasks. A request was made to the NASA Systems and Facilities Branch to acquire imagery that was not already in the LACIE system. Specific requirements such as location and dates of usable imagery, maximum cloud cover, and image quality accompanied the request.

4. After the in-house CIR imagery had been accumulated, it was reviewed for the following qualities.

- a. Scene number.
- b. Date of imagery (to determine usability with respect to agriculture).
- c. Percentage of cloud cover and areas covered. If the cloud coverage was mostly over a lake or city and not over an agricultural area (even if the percentage was high), the imagery could still be usable.
- d. Seasonal coverage (with respect to agricultural area).
- e. Image quality.
- f. Multidate coverage (to enhance interpretation).

After the review, the imagery was logged and filed on an ONC basis. The review parameters for each CIR image were noted in the log.

Construction of an Overlay of the Base Map Physical Features

The purpose of constructing an overlay of the base map physical features (e.g., streams, rivers, and lakes) was to register the Landsat CIR imagery to the overlay used in constructing the agricultural and nonagricultural delineations. The procedure used was as follows.

1. Control coordinates were marked on an ONC base map overlay. These marks were usually made in each of the four corners and at the upper and lower center points of the ONC. The marks were made with a straightedge, and the appropriate geographical coordinates were printed near each mark.

2. All major physical features (streams, rivers, lakes) were delineated or traced on the overlay with blue ink. Minor features were delineated only if major features were lacking, so that the CIR imagery could be registered to the overlay.

(NOTE: If ONC base physical feature overlays were available from the ACIC, steps 1 and 2 were deleted.)

3. The base map overlay was titled and the date of completion was noted. The title on the overlay included the country, ONC map number, map scale, and map series and date numbers.

Agricultural and Nonagricultural Delineation Overlay

The agricultural and nonagricultural delineation overlay was constructed by registering the agricultural and nonagricultural overlay to the overlay of the base map's physical features. This was accomplished as follows.

1. The Landsat CIR imagery was registered with the overlay by aligning the imagery to the overlay of the base map's physical features.

2. The agricultural and nonagricultural areas were delineated by outlining the field and/or nonfield patterns, enclosing all constructed lines, and marking an appropriate symbol (A for agricultural or N for nonagricultural) within each constructed area.

3. All cloud cover regions, regions in which snow cover prohibited agricultural and nonagricultural delineation, or areas for which imagery was not available were marked.

4. Common coordinate marks were matched to all adjoining ONC overlays.

5. All agricultural and nonagricultural delineation lines were matched between adjoining ONC's.

Rationale for Delineation of Agricultural and Nonagricultural Areas

The following criteria were used in delineating areas as either agricultural or nonagricultural in the construction of the overlays.

1. All recognizable field patterns were classified as agricultural.

2. All other areas were classified as non-agricultural.

3. All contiguous nonagricultural areas greater than or equal to 4 square miles in size were delineated. For India and the P.R.C., where agricultural areas are present but the field sizes are too small to be viewed on the imagery, agricultural was defined as all areas not meeting the nonagricultural criterion. The nonagricultural criterion was defined as including only obvious areas such as mountains, deserts, forests, flood plains, or other physical geomorphic phenomena visible in the imagery. If there was any question about whether the area being viewed was agricultural or nonagricultural, it was called agricultural.

4. All imagery used was recorded.

A log was maintained of each image used to make the agricultural and nonagricultural delineations. The log contained the scene number (orbit and row number); date of image scene; image quality (poor, fair, or good); cloud cover percentage; and other comments, including snow cover, unusual features or field patterns, and rationale for decisions made in interpretation. This log reflected the imagery used on an ONC basis and showed all LACIE imagery used up to the agricultural and nonagricultural delineation completion date. The completed overlay was indexed and distributed.

If a determination had to be made about whether a particular area was agricultural or nonagricultural (other than just described), the delineation was based on the regional geographic knowledge of the area(s) in question. The rationale used in determining the delineation was outlined in the documentation that accompanied each individual agricultural

and nonagricultural product. The CAMS senior country analysts were consulted on this matter.

SAMPLE SEGMENT SELECTION

The equipment or material required for sample segment selection included the following.

1. ONC's of the country.

2. A dot matrix, to the scale of the ONC's, spaced 6 nautical miles in the east-west direction and 5 nautical miles in the north-south direction.

3. An overlay of 100 by 100 nautical miles, to the scale of the ONC's, with a center point surrounded by a rectangle of 5 by 6 nautical miles, surrounded by a rectangle of 10 by 12 nautical miles. The boundaries of this larger rectangle are drawn out to the edges of the Landsat scene of 100 by 100 nautical miles (frame boundaries). Lines were drawn running north-south, 10 nautical miles inside both the east and west boundaries of the square of 100 by 100 nautical miles, producing a rectangle 80 nautical miles in the east-west direction and 100 nautical miles in the north-south direction.

4. Lists of random numbers covering the ranges of 1 to 10, 1 to 20, 1 to 30, etc., up to 1 to 300.

5. A data form indicating zone, stratum, and substratum. If there were more than one segment in the stratum/substratum, the segment that was selected was indicated. Space was provided on the form to enter the number of segments (rectangles of 5 by 6 nautical miles) in the stratum/substratum, the number of segments in nonagricultural areas of the stratum/substratum, the number of segments in the agricultural areas of the stratum/substratum, the number of the Landsat track passing through the stratum/substratum, the latitude and longitude of the center of the selected sample segment, and the spring and winter wheat sample segment.

The selection procedure consisted of the following steps.

1. The dot grid was fastened to an ONC.

2. A stratum/substratum was selected from the data form and located on the ONC. Its boundaries were then drawn on the ONC overlay.

3. The number of segments in the stratum/substratum was counted (a segment was in if its center point was in the stratum/substratum) and entered on the data form. If the center point of a seg-

ment fell exactly on a boundary, a coin was flipped. If it came up heads, the segment was placed in the stratum/substratum being worked; if it came up tails, the segment was placed in the adjacent stratum/substratum.

4. The number of segments that fell entirely in nonagricultural areas was counted and entered on the form.

5. The number in step 4 was subtracted from the number in step 3, and the difference was recorded on the data form.

6. Starting with the most northwestern segment, all segments determined in step 5 were numbered.

7. The lowest random number table that included the last number entered in step 6 was selected.

8. The first unused random number from the table was used to determine which segment was the sample segment.

As mentioned in the introduction, not all sample segments were located because of certain engineering constraints. These constraints and changes in the location of a sample segment were determined as described in the following flow-diagram procedures.

1. Place the center point of the overlay of 100 by 100 nautical miles on the center point of the sample segment.

2. Is there another sample segment within the rectangle of 10 by 12 nautical miles?

a. Yes. Discard newest sample segment and return to step 8 of the selection procedure.

b. No. Proceed.

3. Are there more than four sample segments between the extended boundaries of the rectangle of 10 by 12 nautical miles running in the east-west direction?

a. Yes. Discard newest sample segment and return to step 8 of the selection procedure.

b. No. Proceed.

4. If any other sample segments fall within the east-west boundaries of the rectangle of 10 by 12 nautical miles, move the center point of the overlay to the center points of these segments to determine

whether the new sample segment causes more than four sample segments to be within these boundaries.

a. Yes. Discard newest sample segment and return to step 8 of the selection procedure.

b. No. Proceed.

5. With the center point of the overlay back on the center point of the newest sample segment, are there more than eight sample segments between the extended boundaries of the rectangle of 10 by 12 nautical miles running in the north-south direction?

a. Yes. Discard newest sample segment and return to step 8 of the selection procedure.

b. No. Proceed.

6. If any other sample segments fall within the north-south extended boundaries of the rectangle of 10 by 12 nautical miles, move the center point of the overlay to the center points of these segments to determine whether the new sample segment causes more than eight sample segments to be within these boundaries.

a. Yes. Discard newest sample segment and return to step 8 of the selection procedure.

b. No. Proceed.

7. Determine whether the segment is a spring or winter wheat segment, based on whichever comprises more than 50 percent of the total wheat area in the stratum/substratum.

8. Record data for this sample segment as required on data forms.

(NOTE: Latitude and longitude are recorded to degrees and minutes only.)

9. Return to step 2 of the selection procedure and continue until the number of allocated segments for a country is selected and located.

ACKNOWLEDGMENT

The assistance of W. Daley in documenting the section entitled "Determination of Area Sampling Frame" is gratefully acknowledged.

LACIE Large-Area Acreage Estimation

R. S. Chhikara^a and A. H. Feiveson^b

INTRODUCTION

This paper describes the procedure for estimating wheat acreage for a large area, given estimates for the sample segments. A segment wheat acreage estimate is obtained by multiplying its small-grains acreage estimate as computed by the Classification and Mensuration Subsystem (CAMS) by the best available ratio of wheat to small-grains acreages obtained using historical data. The CAMS approach for estimating segment small-grains acreages is described in the symposium paper by Heydorn et al. entitled "Classification and Mensuration of LACIE Segments," and the econometric models used in predicting the ratio of wheat to small-grains acreages are given in the paper by Umberger et al. entitled "Econometric Models for Predicting Confusion Crop Ratios."

In the United States and in other countries with detailed historical data, sample allocation was made at the substratum level. As a result, the acreage estimation in countries with detailed historical data requires one level of aggregation more than it would in other countries. The estimation procedure described in this paper is for the United States, but it is equally applicable to other LACIE countries with detailed historical data. Also described are the essential features of the estimation procedure for the remaining LACIE countries.

The U.S. counties correspond to substrata and were grouped into three categories for sample allocation. Those in Group I were allocated at least one segment each, two-stage probability-proportional-to-size (PPS) sampling was used in Group II counties, and no sample segments were allocated to counties in Group III. In the United States, a stratum (crop reporting district (CRD)) corresponds to a collection

of counties, a zone corresponds to a state, and a region corresponds to a collection of states (see the paper by Feiveson et al. entitled "LACIE Sampling Design" for details).

Wheat acreage estimates are made for each stratum, zone, and region in a LACIE country. However, no estimate is made for a zone unless it contains at least three segments satisfactorily processed by CAMS. A segment wheat acreage estimate may not be made in the following cases of non-response.

1. The sample segment was obscured by cloud cover.
2. Landsat data quality was insufficient to permit processing.
3. Landsat data acquisition was not properly registered with the reference Landsat image.
4. The acquisition and/or processing procedures failed to provide an acceptable estimate.

ACREAGE ESTIMATION

A CRD (stratum) acreage estimate consists of three components.

1. An acreage estimate is made for the Group I counties (substrata) for which segment data exist. (A Group I county is treated as a Group III county if it does not have at least one segment with an acceptable wheat proportion estimate.)
2. An acreage estimate is made for the entire set of Group II counties in the CRD if there is at least one segment with an acceptable wheat proportion estimate in this set of counties. (Otherwise, the Group II counties are all treated as Group III counties.)
3. An acreage estimate is made for the Group III counties, including the Group I and II counties being treated as Group III counties.

The wheat acreage estimates for these three components are obtained as follows.

^aLockheed Electronics Company, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

Group I Substrata Estimation

Group I counties are treated as strata, and a stratified random sampling estimator is used to estimate their wheat acreages. The estimate for the collection of Group I counties in the j th CRD is given by

$$A_{1j} = \sum_{k=1}^{L_{1j}} A_{1jk} \quad (1)$$

where L_{1j} = number of Group I counties in the j th CRD

A_{1jk} = wheat acreage estimate for the k th Group I county in the j th CRD

The county (substratum) estimate is obtained by

$$A_{1jk} = R_{1jk} N_{1jk} \frac{1}{M_{1jk}} \sum_{i=1}^{M_{1jk}} A_{1jki} \quad (2)$$

where R_{1jk} = ratio of the true k th substratum area to its gross pseudosubstratum (GPC) area (see the paper by Liszcz entitled "LACIE Area Sampling Frame and Sample Selection" for the definition of GPC)

N_{1jk} = number of segments (after exclusion of nonagricultural segments) in the k th substratum of the j th stratum

M_{1jk} = number of sample segments for which estimates are available in the k th substratum of the j th stratum

A_{1jki} = estimated wheat acreage for the i th sample segment in the k th substratum of the j th stratum

Group II Substrata Estimation

A PPS estimator is used to estimate the wheat acreages for the Group II collection of substrata. The wheat acreage estimate for the Group II collection of substrata in the j th stratum is given by

$$A_{2j} = \sum_{k=1}^{M_{2j}} R_{2jk} N_{2jk} \frac{A_{2jk}}{\pi_{2jk}} \quad (3)$$

where M_{2j} = number of sample segments for which acreage estimates are available in the Group II substrata of the j th stratum

R_{2jk} = ratio of the true k th Group II substratum area to its GPC area

N_{2jk} = number of segments (after exclusion of nonagricultural segments) in the k th Group II substratum of the j th stratum

A_{2jk} = wheat acreage estimate of the sample segment belonging to the k th substratum in the j th stratum (There is only one segment allocated in each selected Group II substratum.)

π_{2jk} = probability of selection for the k th Group II substratum of the j th stratum, given by

$$\pi_{2jk} = \frac{M_{2j} W_{2jk}}{W_{2j}} \quad (4)$$

where W_{2jk} = wheat acreage harvested in the primary epoch year in the k th Group II substratum of the j th stratum

$$W_{2j} = \sum_{k=1}^{L_{2j}} W_{2jk}$$

L_{2j} = number of Group II substrata in the j th stratum

Group III Substrata Estimation

The wheat acreage estimate for the collection of Group III substrata is obtained by means of a ratio estimator. Depending on the number of segments in a stratum for which estimates are available, three categories of Group III acreage estimates are possible. Categories 1, 2, and 3 correspond respectively to three or more segments, one or two segments, and no segments having estimates available. The ratio used for the Group III estimator is the historical wheat acreage for the Group III counties divided by the historical wheat acreage for the combined Group I and II counties.

For category 1 estimates (three or more usable segments in the stratum), the ratio is based on historical acreages only within the stratum. The acreage estimate for the Group III substrata in the j th stratum is given by

$$A_{3j} = \left(\frac{A_{1j} + A_{2j}}{W_{1j} + W_{2j}} \right) W_{3j} \quad (5)$$

where A_{1j} and A_{2j} are given by equations (1) and (3) and W_{1j} , W_{2j} , and W_{3j} are the historical wheat acreages for Group I, II, and III substrata in the stratum, respectively.

For category 2 and 3 estimates (less than three usable segments in the stratum), the ratio is based on acreages in the zone containing the stratum for which the estimate is being made, and the acreage estimate for the Group III substrata in the j th stratum is obtained by

$$A_{3j} = \left(\frac{A_{1\cdot} + A_{2\cdot}}{W_{1\cdot} + W_{2\cdot}} \right) W_{3j} \quad (6)$$

where a dot (.) in a subscript denotes the summation over all the Group I or Group II substrata, whichever is the case, in the zone. The reason for differentiating between categories 2 and 3 will become evident in the section dealing with stratum variance estimation.

Stratum Estimation

In the United States and in other LACIE countries with detailed historical data, the wheat acreage estimate of each stratum is computed as the sum of the Group I, II, and III component estimates which comprise the stratum, as follows.

$$A_j = A_{1j} + A_{2j} + A_{3j} \quad (7)$$

In the U.S.S.R. and in other LACIE countries without detailed historical data, the wheat acreage estimate of a stratum is given by

$$A_j = R_j N_j \frac{1}{n_j} \sum_{k=1}^{n_j} A_{jk} \quad (8)$$

where R_j = ratio of the actual area to the pseudo gross area for the j th stratum

N_j = number of segments (excluding the nonagricultural segments) in the j th stratum

n_j = number of sample segments for which estimates are available in the j th stratum

A_{jk} = wheat acreage estimate for the k th segment in the j th stratum

It is required to have $n_j \geq 3$; otherwise, no acreage estimate for the stratum is made.

Higher Level Estimations

The wheat acreage estimate for a zone, a region, or a country is obtained by adding estimates for the strata included in the zone, region, or country. The acreage estimate at the zone level is obtained by

$$A_{rz} = \sum_{j=1}^S A_{rzj} \quad (9)$$

where S = number of strata in the z th zone, r th region of the country

A_{rzj} = acreage estimate of the j th stratum, z th zone, r th region of the country

The acreage estimate at the region level is obtained by

$$A_r = \sum_{z=1}^R A_{rz} \quad (10)$$

where R is the number of zones in the r th region of the country.

The acreage estimate at the country level is

$$A = \sum_{r=1}^K A_r \quad (11)$$

where K is the number of regions in the country.

ACREAGE VARIANCE ESTIMATION

In countries with detailed historical data, the problem of acreage variance estimation involves several complexities resulting from the use of a two-stage PPS sampling scheme for the Group II substrata and the availability of only one sample segment per substratum in most cases. The estimation procedure in such countries consists of a series of steps to be described in the subsections to follow. On the other hand, it is fairly straightforward to estimate the variance for countries that lack historical data; in this case, no direct variance estimate is attempted for a stratum containing less than three processed segments, and all strata belong to the Group I category.

Group I Substrata Variance Estimation

The variance estimate for the Group I substrata acreage estimate is obtained using the variance formula for a stratified random sampling. For the j th stratum, it is computed by

$$V_{1j} = \sum_{k=1}^{L_{1j}} R_{1jk}^2 N_{1jk}^2 \frac{S_{1jk}^2}{M_{1jk}} \quad (12)$$

where L_{1j} , R_{1jk} , N_{1jk} , and M_{1jk} are as defined previously and S_{1jk}^2 is the within-substratum variance estimate to be computed according to the procedure described in a succeeding section. The finite population correction, M_{1jk}/N_{1jk} , is negligible; hence, it is not considered in equation (12).

Group II Substrata Variance Estimation

The variance of the estimate for a collection of Group II substrata consists of a within-substratum variance component and a between-substrata variance component. The first component can be estimated in a manner similar to the Group I case, but estimation of the second component requires additional historical acreages for all Group II substrata in each stratum.¹ Using the Hartley-Rao PPS sampling approach described in reference 1, the follow-

ing variance estimate for the Group II substrata in the j th stratum is obtained.

$$V_{2j} = \sum_{k=1}^{L_{2j}} R_{2jk}^2 N_{2jk}^2 \frac{S_{2jk}^2}{\pi_{2jk}} + \sum_{k=1}^{L_{2j}-1} \sum_{i=k+1}^{L_{2j}} (\pi_{2jk}\pi_{2ji} - \pi_{2jki}) \cdot \left(\frac{Y_{2jk}}{\pi_{2jk}} - \frac{Y_{2ji}}{\pi_{2ji}} \right)^2 \quad (13)$$

where L_{2j} , R_{2jk} , N_{2jk} , and π_{2jk} are as defined previously and S_{2jk}^2 is the within-substratum variance estimate to be computed according to the procedure in the succeeding subsection. The Y_{2jk} 's are the Group II substrata historical wheat acreages during a year other than the primary epoch year; i.e.,

Y_{2jk} = wheat acreage harvested in the secondary epoch year in the k th Group II substratum in the j th stratum

Y_{2ji} = wheat acreage harvested in the secondary epoch year in the i th Group II substratum in the j th stratum

π_{2jki} = probability of having the pair of Group II substrata k and i selected in the sample. For $i \neq k$, π_{2jki} is determined according to the procedure given by Hartley and Rao (ref. 1) and is computed using the following formula.

$$\pi_{2jki} = \frac{M_{2j} - 1}{M_{2j}} \pi_{2jk} \pi_{2ji} + \frac{M_{2j} - 1}{M_{2j}^2} (\pi_{2jk}^2 \pi_{2ji} + \pi_{2jk} \pi_{2ji}^2) - \frac{M_{2j} - 1}{M_{2j}^3} (\pi_{2jk} \pi_{2ji}) \left(\sum_{\alpha=1}^{L_{2j}} \pi_{2j\alpha}^2 \right) + \frac{2(M_{2j} - 1)}{M_{2j}^3} (\pi_{2jk}^3 \pi_{2ji} + \pi_{2ji}^3 \pi_{2jk} + \pi_{2jk}^2 \pi_{2ji}^2)$$

¹H. O. Hartley and D. A. Lamb: Formulas for Variance Estimation in Proposed LACIE Sampling Plan. Technical Report to NASA, 1975.

$$\begin{aligned}
& - \frac{3(M_{2j} - 1)}{M_{2j}^4} (\pi_{2jk}^2 \pi_{2ji} + \pi_{2jk} \pi_{2ji}^2) \left(\sum_{\alpha=1}^{L_{2j}} \pi_{2j\alpha}^2 \right) \\
& - \frac{2(M_{2j} - 1)}{M_{2j}^4} (\pi_{2jk} \pi_{2ji}) \left(\sum_{\alpha=1}^{L_{2j}} \pi_{2j\alpha}^3 \right) \\
& + \frac{3(M_{2j} - 1)}{M_{2j}^5} (\pi_{2jk} \pi_{2ji}) \left(\sum_{\alpha=1}^{L_{2j}} \pi_{2j\alpha}^2 \right)^2 \quad (14)
\end{aligned}$$

where M_{2j} is as defined previously.

Within-Substratum Variance Estimation

Often, there is only one sample segment in a substratum; therefore, no direct estimate of the within-substratum variance is possible. If variances for substrata are assumed to be equal, substrata are collapsed to form a new stratum and its sample variance provides an estimate of the within-substratum variance. But this technique generally leads to an overestimation of the variance and hence provides a biased estimate of the variance. Different methods of collapsing strata have been suggested by Hensen et al. (ref. 2), Cochran (ref. 3), and Seth (ref. 4). Another variance estimate is possible using the method of Hartley et al. (ref. 5), who suggest the regression approach and show that their technique may lead to smaller bias in variance estimation as compared to the collapsed strata technique. But, as is pointed out by Hartley et al., this technique could lead to negative variance estimates, particularly if the concomitant variables are not well correlated with the stratum means. In an earlier study (ref. 6), application of the collapsed strata and Hartley-Rao-Kiefer (ref. 5) techniques for variance estimation did not seem to provide satisfactory results; the former led to overestimation and the latter to negative estimates. However, when the Hartley-Rao-Kiefer approach was combined with the collapsed strata approach, where first substrata were grouped into groups of substrata as homogeneous as possible and then a separate regression was performed for each group of substrata, the empirical results were more satisfactory. Therefore, the method combining the

two approaches was adopted for use in LACIE. It is based on the assumption that the historical county proportions are well correlated with the CAMS estimates of segment proportions. The method consists of (1) forming homogeneous groups of substrata in a zone with respect to a priori estimates of within-substratum variability, (2) performing regression of the CAMS segment wheat proportion estimates on the substratum historical wheat proportions, and (3) taking the residual mean squared error (MSE) as an estimate of the within-substratum variance for each substratum in a group.

Segments within a zone are grouped into collections according to the corresponding a priori within-substratum standard deviations θ_k used in the original allocation (see the paper by Feiveson et al.). These collections of substrata are required to be as homogeneous as possible, and each must have an adequate number of segments to allow a reliable estimate of the variance. The following conditions should be satisfied in forming the collections.

1. No collection should contain less than eight segments unless there are less than eight segments available for the zone. This constraint is to ensure an adequate number of degrees of freedom for obtaining a reliable linear regression equation.

2. All segments in the same substratum shall be in the same collection. It is necessary to use every available segment in a substratum for estimating its variance.

3. The number of collections c shall be given initially by

$$\begin{aligned}
c &= 1 \text{ provided } NS < 16 \\
c &= 2 \text{ provided } 16 \leq NS < 24 \\
c &= 3 \text{ provided } NS \geq 24
\end{aligned}$$

where NS is the number of available segments in the zone. If conditions 1 and 2 cannot be satisfied when NS is greater than or equal to 16, reduce the value of c by 1. This condition is imposed to keep down the number of collections to avoid an unnecessarily fine grouping of substrata.

4. If c is greater than 1, the collections should maximize the ratio of the between-collection

variance to the within-collection variance of the θ_{rj} ; i.e., let

$$F = \frac{\sum_{r=1}^c \sum_{j=1}^{d_r} \frac{(\bar{\theta}_{r\cdot} - \bar{\theta}_{\cdot\cdot})^2}{c-1}}{\sum_{r=1}^c \sum_{j=1}^{d_r} \frac{(\theta_{rj} - \bar{\theta}_{r\cdot})^2}{d_r - c}} \quad (15)$$

where d_r is the number of segments in the r th collection, θ_{rj} is the a priori small-grains standard deviation (see the paper by Feiveson et al. and ref. 7) associated with the substratum containing the j th segment in the r th collection (note that segments in the same substratum have duplicate θ_{rj} values),

$$\bar{\theta}_{r\cdot} = \frac{1}{d_r} \sum_{j=1}^{d_r} \theta_{rj} \text{ and } \bar{\theta}_{\cdot\cdot} = \frac{1}{c} \sum_{r=1}^c \bar{\theta}_{r\cdot}$$

Then, the partitioning of the $\sum_r d_r$ segments into collections should be such that F is maximized, subject to the constraints specified in conditions 1 and 2. If F is less than 1, c is reduced and F is recomputed. This requirement is to form as many homogeneous collections of substrata as possible.

Let \hat{p}_{rij} be the wheat proportion estimate for the j th segment of the r th substratum of the r th collection, and let x_{rj} be the epoch year (historical) wheat proportion for the r th substratum. It is assumed that the segment estimate can be modeled as

$$\hat{p}_{rij} = \alpha_r + \beta_r x_{rj} + \epsilon_{rij} \quad (16)$$

where $E(\epsilon_{rij}) = \Delta r$ and $Var(\epsilon_{rij}) = \sigma_r^2$. Then, for each of the c collections, a regression of \hat{p}_{rij} on the x_{rj} is performed. Let S_r^2 denote the residual mean squared error from the r th regression; i.e.,

$$S_r^2 = (\text{segment acreage})^2 \sum_{j=1}^{d_r} \frac{(\hat{p}_{rij} - p_{rj})^2}{(d_r - 2)} \quad (17)$$

where $p_{rj} = a_r + b_r x_{rj}$ (the predicted value using the regression equation).

$$\left. \begin{aligned} a_r &= \bar{p}_{r\cdot} - b_r \bar{x}_{r\cdot} \\ b_r &= \frac{\sum_j \hat{p}_{rj} (x_{rj} - \bar{x}_{r\cdot})}{\sum_j (x_{rj} - \bar{x}_{r\cdot})^2} \end{aligned} \right\} \quad (18)$$

Since by constraint 2, every substratum is associated with one and only one collection, the variance S_r^2 is assigned to every substratum having segments in the r th collection (whether the substratum is Group I or II); i.e., S_{1jk}^2 (or S_{2jk}^2) = S_r^2 if the segments in the k th substratum of the j th stratum belong to the r th collection.

Stratum Variance Estimation

For countries with detailed historical data, the stratum acreage variance estimation depends on the category of its Group III substrata.

In a stratum having at least three sample segments processed, where ratioing is done only within the stratum, the variance for the j th stratum acreage estimate given in equation (6) is easily seen to be

$$V_j = \left(1 + \frac{W_{3j}}{W_{1j} + W_{2j}} \right)^2 (V_{1j} + V_{2j}) \quad (19)$$

and hence can be estimated by equation (19), where estimates of V_{1j} and V_{2j} replace the actual variances (see eqs. (12) and (13)) and where W_{1r} , W_{2r} , and W_{3j} are as defined previously.

If the historical acreage ratio $W_{3j}/(W_{1j} + W_{2j})$ is different from that of the current year, it will introduce bias into the estimate; hence, V_j in equation (19) provides a biased estimate of the stratum variance. However, because the historical acreage ratios are not expected to show any significant year-to-year variability, equation (19) and the others

below are regarded as providing unbiased variance estimates.

If the j th stratum has at least one but less than three segments processed, the stratum acreage estimate given by equation (6) can be written as

$$A_j = (A_{1j} + A_{2j}) + \frac{A_{1\cdot} + A_{2\cdot}}{W_{1\cdot} + W_{2\cdot}} W_{3j}$$

$$= (A_{1j} + A_{2j}) \left(1 + \frac{W_{3j}}{W_{1\cdot} + W_{2\cdot}} \right) + \frac{A_{1\cdot}^{(j)} + A_{2\cdot}^{(j)}}{W_{1\cdot} + W_{2\cdot}} W_{3j} \quad (20)$$

where $A_{1\cdot}^{(j)} = A_{1\cdot} - A_{1j}$
 $A_{2\cdot}^{(j)} = A_{2\cdot} - A_{2j}$

Since the two terms on the right side of equation (20) have independent acreage estimates, the variance of A_j can be easily obtained. With variance components replaced by their estimates, the stratum acreage variance estimate is given by

$$V_j = \left[1 + \frac{2W_{3j}}{\sum_{i \in S} (W_{1i} + W_{2i})} \right] (V_{1j} + V_{2j}) + \left[\frac{W_{3j}}{\sum_{i \in S} (W_{1i} + W_{2i})} \right]^2 \sum_{i \in S} (V_{1i} + V_{2i}) \quad (21)$$

where S is the set of indices i , such that the j th stratum in the zone has at least one processed segment.

If the j th stratum has no segment processed, the whole stratum is in the Group III category, and the

variance estimate for the stratum acreage estimate is given by

$$V_j = \left[\frac{W_{3j}}{\sum_{i \in S} (W_{1i} + W_{2i})} \right]^2 \sum_{i \in S} (V_{1i} + V_{2i}) \quad (22)$$

If the country lacks detailed historical data, it has large strata, and the within-stratum variance estimate is directly computed by

$$S_j^2 = \frac{1}{n_j - 1} \sum_{k=1}^{n_j} (A_{jk} - \bar{A}_j)^2 \quad (23)$$

where

$$\bar{A}_j = \frac{1}{n_j} \sum_{k=1}^{n_j} A_{jk}$$

Hence, the stratum acreage variance estimate is obtained by

$$V_j = R_j^2 N_j^2 S_j^2 \quad (24)$$

where R_j and N_j are as defined previously.

Variance Aggregation to the Zone, Region, and Country Levels

For a zone in a country lacking detailed historical data, the stratum acreage estimates are independently obtained. Hence, the variance estimate is obtained by aggregating the stratum variance estimates in the zone,

$$V_z = \sum_j V_{zj} \quad (25)$$

where V_{zj} is the variance estimate for the j th stratum in the z th zone.

In countries with detailed historical data, stratum acreage estimates in a zone are correlated unless all Group III ratio estimation is done only within the stratum. In general, the zone acreage estimate given by equation (9) is of the form

$$A_z = \sum_{j \in L} \left[1 + \frac{w_{3j}}{w_{1j} + w_{2j}} + \frac{\sum_{i \in M} w_{3i} + \sum_{i \in N} w_{3i}}{\sum_{i \in S} (w_{1i} + w_{2i})} \right] (A_{1j} + A_{2j})$$

$$+ \left[1 + \frac{\sum_{i \in M} w_{3i} + \sum_{i \in N} w_{3i}}{\sum_{i \in S} (w_{1i} + w_{2i})} \right] \sum_{j \in M} (A_{1j} + A_{2j}) \quad (26)$$

where S is the set of indices associated with strata having at least one processed segment, M is the set of indices associated with strata that have at least one but less than three segments processed, L is the set of indices associated with strata that have at least three segments processed, and N is the set of indices associated with strata that have no processed segments. Since L and M are disjoint sets, making the estimates in the two terms on the right side of equation (26) uncorrelated, an estimate of the variance of the zone acreage estimate is obtained by

$$V_z = \sum_{j \in L} \left[1 + \frac{w_{3j}}{w_{1j} + w_{2j}} + \frac{\sum_{i \in M} w_{3i} + \sum_{i \in N} w_{3i}}{\sum_{i \in S} (w_{1i} + w_{2i})} \right]^2 (V_{1j} + V_{2j})$$

$$+ \left[1 + \frac{\sum_{i \in M} w_{3i} + \sum_{i \in N} w_{3i}}{\sum_{i \in S} (w_{1i} + w_{2i})} \right]^2 \sum_{j \in M} (V_{1j} + V_{2j}) \quad (27)$$

No variance estimation is required for zones that do not contain at least three processed sample segments.

Since zone acreage estimates are obtained independently, the acreage variance estimates at both the regional and country levels are computed by adding the zone acreage variance estimates. This procedure

holds for all LACIE countries. Specifically, the variance estimate at the region level is obtained by

$$V = \sum_{z=1}^R V_{rz} \quad (28)$$

and the variance estimate at the country level is obtained by

$$V = \sum_{r=1}^K V_r \quad (29)$$

MIXED WHEAT AREA AND VARIANCE ESTIMATION

Winter or Spring Wheat Estimation

In a mixed wheat area, separate aggregations are performed for estimating the spring and winter wheat acreage estimates as well as their variance estimates at the stratum and higher levels. In each case, the estimation procedure is the same as that described in the two preceding sections for each aggregation level. Data from sample segments designated as winter or spring wheat segments and the historical substratum winter or spring wheat acreages are used to estimate winter or spring wheat acreages and their associated variance estimates. However, inputs of $\pi_{2/k}$ in estimating the Group II substrata acreages are based on the historical total wheat area and are the same in both cases, because the Group II substrata for sample segment allocation were selected with probabilities that were determined from the historical total wheat area for the collection of Group II substrata.

Total Wheat Estimation

The total wheat area estimate in a mixed wheat area is computed by adding the winter wheat and the spring wheat acreage estimates for the area of interest; i.e., if A_w and A_s denote the winter and spring

wheat acreage estimates, respectively, the total wheat acreage estimate A_t is given by

$$A_t = A_w + A_s \quad (30)$$

This is done at the stratum and higher levels.

The two estimates A_w and A_s are correlated for overlapping winter and spring wheat areas in a zone. Thus, the variance of A_t is given by

$$Var(A_t) = Var(A_w) + Var(A_s) + 2Cov(A_w, A_s) \quad (31)$$

where the covariance (A_w, A_s) can be expected to be negative. If so,

$$Var(A_t) < Var(A_w) + Var(A_s) \quad (32)$$

Accordingly, the variance estimate for the total wheat is biased if obtained by adding the variance estimates for the winter and spring wheat area estimates. Instead, if it is obtained by way of estimating the total wheat area directly, a better variance estimate is expected. The latter procedure is followed in LACIE for the total wheat acreage variance estimation in the mixed wheat areas. The procedure is the same as described in the third section and is applicable to each level of aggregation.

ACKNOWLEDGMENT

H. O. Hartley of Texas A & M University suggested the sampling plan and the acreage estimation approach that form the basis of the methodology described in this paper. Jason Chang, who previously worked for Lockheed Electronics Company, Inc., developed a software system that supported the methodological development.

The authors also benefited from the comments and suggestions made by C. R. Hallum and A. G. Houston of the Earth Observations Division, Johnson Space Center, and by other colleagues.

REFERENCES

1. Hartley, H. O.; and Rao, J. N. K.: Sampling With Unequal Probabilities and Without Replacement. *Ann. Math. Stat.*, vol. 33, 1962, pp. 350-374.
2. Hensen, M. H.; Hurwitz, W. N.; and Madow, W. G.: *Sample Survey Methods and Theory*. Wiley (New York), 1953.
3. Cochran, W. G.: *Sampling Techniques*. Wiley (New York), 1963.
4. Seth, G. R.: On Collapsing of Strata. *J. Indian Soc. Agric. Stat.*, vol. 18, 1966, pp. 1-3.
5. Hartley, H. O.; Rao, J. N. K.; and Kiefer, G.: Variance Estimation With One Unit Per Stratum. *J. American Stat. Assoc.*, vol. 64, 1969, p. 841.
6. Chhikara, R. S.; and Chang, J.: An Empirical Study of Variance Estimation With One Unit Per Stratum. Paper presented at the 136th Annual Meeting of the American Statistical Association (Boston, Mass.), Aug. 1976.
7. Chhikara, R. S.; and Chang, J.: LACIE Area Variance Estimate in the United States. Technical Memorandum LEC-8054, Lockheed Electronics Company, Inc., Houston, Tex., 1976.

Large-Area Aggregation and Mean-Squared Prediction Error Estimation for LACIE Yield and Production Forecasts

R. S. Chhikara^a and A. H. Feiveson^b

INTRODUCTION

In LACIE, large-area wheat acreage estimates are made from the Landsat data acquired according to the sampling design described by Feiveson et al. in the paper entitled "LACIE Sampling Design." The acreage estimation procedure is described for a segment (sampling unit) by Heydorn and Bizzell in the paper entitled "Methods for Segment Wheat Area Estimation" and for large areas by Chhikara and Feiveson in the paper entitled "LACIE Large Area Acreage Estimation." Yield is predicted by establishing a relationship between historical yield and weather data (detailed in the paper by Strommen et al. entitled "Development of LACIE CCEA-1 Weather/Wheat Yield Models"). Though weather also influences crop acreages in an area and harvested wheat acreage and its yield per acre may thus be correlated, these are estimated independently in LACIE.

The terminology to be used in the paper is basically the same as that employed by Chhikara and Feiveson in the paper entitled "LACIE Large Area Acreage Estimation"; hence, no attempt is made here to define again the terms which have appeared in the other paper.

The yield stratification does not necessarily coincide with the stratification used for the acreage estimation. A yield stratum generally consists of several acreage strata and sometimes crosses zone boundaries. For example, the Panhandle yield model covers some crop reporting districts (CRD's) from Oklahoma and some CRD's from Texas. No prediction is attempted for yield and hence for production

below the yield stratum level. It is necessary to define the stratification by which wheat production can be estimated efficiently from the given acreage and yield estimates. Consequently, pseudozones are created in a zone if it is covered by more than one yield stratum. A pseudozone is obtained from the intersection of a yield stratum with a zone (described in fig. 1).

An estimate¹ of the production in a pseudozone is obtained by the product of its area estimate and its

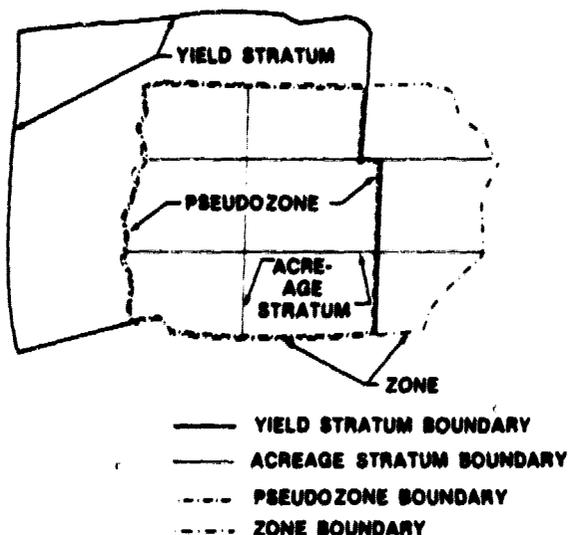


FIGURE 1.—Determination of a pseudozone.

^aLockheed Electronics Company, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

¹Estimates refer to forecasts when these are made prior to crop harvest time.

yield prediction, and these estimates are aggregated to estimate zone and higher level productions. In mixed areas, this is done for each crop type (winter and spring wheat). The total wheat production is estimated by adding the two crop-type production estimates for a zone, a region, and a country.

PRODUCTION ESTIMATION

In this section, the aggregation formulas are listed for production estimation of a crop type.

Zone Estimate

Suppose a zone consists of H pseudozones, G_1, G_2, \dots, G_H , with acreage estimates $A_{z1}, A_{z2}, \dots, A_{zH}$ and yield predictions $Y_{z1}, Y_{z2}, \dots, Y_{zH}$, respectively. Then, the zone production is estimated by

$$P_z = \sum_{i=1}^H A_{zi} Y_{zi} \quad (1)$$

Region Estimate

Suppose that a region consists of R zones with production estimates $P_{r1}, P_{r2}, \dots, P_{rR}$. Then, the regional production estimate is obtained by

$$P_r = \sum_{j=1}^R P_{rj} \quad (2)$$

Country Estimate

If a country consists of K regions with production estimates P_1, P_2, \dots, P_K , then the country production estimate is given by

$$P = \sum_{r=1}^K P_r \quad (3)$$

MEAN-SQUARED PREDICTION ERROR ESTIMATION

Zone Prediction Error Estimate

The yield prediction error is estimated for a yield stratum and is available as a standard output from the yield prediction algorithm described by Strommen et al. Each pseudozone of a yield stratum is assigned the yield prediction error as estimated for the yield stratum. On the other hand, the acreage estimation error (variance) needs to be estimated for a pseudozone. Since stratum acreage estimates in a zone can be correlated, it is necessary to estimate the covariances between different pairs of pseudozone acreage estimates because these estimates can also be correlated.

Each pseudozone acreage estimate is obtained by summing acreage estimates for strata comprising the pseudozone; i.e.,

$$A_{zi} = \sum_{j \in G_i} A_{zij} \quad (4)$$

where A_{zij} is the acreage estimate for the j th stratum in the i th pseudozone. Thus, the variance of A_{zi} is given by

$$V = \sum_{j \in G_i} V_j + \sum_{j \in G_i} \sum_{k \in G_i} V_{jk} \quad (5)$$

where V_j is the acreage variance estimate for the j th stratum (obtained as described in the paper by Chhikara and Feiveson) and V_{jk} is the estimated covariance between acreage estimates for the j th and k th strata in the zone. The covariance between acreage estimates for a pair of strata would depend on how the estimates are obtained for the strata. This, in turn, would depend on how the Group III substrata acreage for each stratum is estimated. When data from at least three segments are available for a stratum, the estimate for its Group III substrata is based on the data from the stratum alone. Other-

wise, all the available data from the corresponding zone are used. Accordingly,

$$V_{jk} = 0 \quad (6)$$

when acreage estimates for the Group III substrata of the j th and k th strata are obtained from data available within each stratum and thus provide independent stratum acreage estimates. In other cases, it is obtained as follows.

When the estimate for the Group III substrata of the j th stratum is based on data only from the stratum, whereas it is obtained for the k th stratum using the data available for the zone in which the stratum lies,

$$V_{jk} = \frac{W_{3k} (W_{1j} + W_{2j} + W_{3j}) (V_{1j} + V_{2j})}{(W_{1j} + W_{2j}) \sum_{i \in \{S\}} (W_{1i} + W_{2i})} \quad (7)$$

When zone data are used to obtain acreage estimates for the Group III substrata of the j th and k th strata, the following three cases arise.

1. When each of the j th and k th strata has less than three but at least one segment available,

$$V_{jk} = \frac{W_{3k} (V_{1j} + V_{2j}) + W_{3j} (V_{1k} + V_{2k})}{\sum_{i \in \{S\}} (W_{1i} + W_{2i})} \quad (8)$$

2. When the j th stratum has less than three but at least one segment processed and the k th stratum has no segment processed,

$$V_{jk} = \frac{W_{3k} (V_{1j} + V_{2j})}{\sum_{i \in \{S\}} (W_{1i} + W_{2i})} + \frac{W_{3j} W_{3k} \sum_{i \in \{S\}} (V_{1i} + V_{2i})}{\left[\sum_{i \in \{S\}} (W_{1i} + W_{2i}) \right]^2} \quad (9)$$

3. When each of the j th and k th strata has no segment processed,

$$V_{jk} = \frac{\frac{W_{3j} W_{3k}}{\left[\sum_{i \in \{S\}} (W_{1i} + W_{2i}) \right]^2} \sum_{i \in \{S\}} (V_{1i} + V_{2i})}{\frac{W_{3j} W_{3k}}{\sum_{i \in \{S\}} (V_{1i} + V_{2i})} + \frac{\left[\sum_{i \in \{S\}} (W_{1i} + W_{2i}) \right]^2}{\sum_{i \in \{S\}} (V_{1i} + V_{2i})}} \quad (10)$$

The quantities V_{1j} , V_{2j} , W_{1j} , W_{2j} , and W_{3j} , and the set $\{S\}$ are defined as follows.

1. V_{1j} and V_{2j} are acreage variance estimates for the Group I and Group II substrata, respectively, in the j th stratum.

2. W_{1j} , W_{2j} , and W_{3j} are the historical wheat acreages for Group I, Group II, and Group III substrata, respectively, in the j th stratum.

3. $\{S\}$ is the set of strata, each of which has at least one segment processed and the wheat proportion estimated.

For computation of V_{1j} and V_{2j} and for an understanding of the categories of strata, refer to the paper by Chhikara and Feiveson on acreage estimation. The formulas given in equations (6) through (10) are fairly straightforward and are easily obtained by considering different possible pairs of strata in a zone.

Acreage estimates and yield predictions are independently made up to the pseudozone level. Although some correlation between crop acreages and their yields per acre within a pseudozone is possible, it is assumed that a pseudozone acreage estimate and its yield prediction are uncorrelated. Also, both the acreage estimate and the yield prediction are assumed to be unbiased. Then, the squared prediction error for a pseudozone production estimate follows from the formula for the variance of product of two random variables (ref. 1, p. 12); and an unbiased estimate of it is obtained by

$$S^2 = VY^2 + UA^2 - VU \quad (11)$$

where V is given by equation (5), A is the acreage estimate, Y is the yield prediction, and U is the yield squared prediction error for the pseudozone.

If wheat production estimates for pseudozones in a zone are uncorrelated, the zone production variance is given by the sum of variances for the pseudozones. However, correlation between acreage estimates for strata in a zone will probably result in some dependence between pseudozone acreage estimates. The following formula for estimating the mean-squared error for a zone production estimate accounts for such dependence; pseudozone yield predictions for a zone are made independently. As such, an estimate of the mean-squared error of the production estimate P for the z th zone is obtained by

$$S_z^2 = \sum_{i=1}^H (V_{zi}Y_{zi}^2 + U_{zi}A_{zi}^2 - V_{zi}U_{zi}) + 2 \sum_{i=2}^H \sum_{l=1}^{i-1} Y_{zi}Y_{zl} \left(\sum_{j \in G_i} \sum_{k \in G_l} V_{jk} \right) \quad (12)$$

where U_{zi} is the estimated mean-squared prediction error of the yield for the i th pseudozone, V_{zi} is the area variance estimate for the i th pseudozone, and other quantities are as defined earlier. Again, it is fairly straightforward to derive equation (12) from equations (11), (5), and (1).

Region and Country Prediction Error Estimates

If yield strata in a region cross zone boundaries, the production estimates for the zones in the region will be correlated. For example, such is the case in the U.S. Great Plains, where the Texas Panhandle winter wheat yield model covers CRD's from both Oklahoma and Texas and where the Red River spring wheat yield model covers two eastern CRD's of North Dakota and three western CRD's of Minnesota. Thus, accounting for both variance and covariance terms for the zones, one derives an esti-

mate of the mean-squared error of P_r , the regional production estimate, which is given by

$$S_r^2 = \sum_{z=1}^R S_{rz}^2 + \sum_{z=1}^R \sum_{z' \neq z} S_{rzz'} \quad (13)$$

where S_{rz}^2 is the estimated mean-squared error of P_{rz} , the production estimate for the z th zone, and $S_{rzz'} = 0$ if the z th and z' th zones have no yield stratum in common. Otherwise,

$$S_{rzz'} = \sum_{k=1}^C A_{rzk} A_{r'zk} U_{rk} \quad (14)$$

where U_{rk} is the mean-squared prediction error for the k th yield stratum estimate commonly applicable to the area estimate A_{rz} for the z th zone and the area estimate $A_{r'z'}$ for the z' th zone, and C is the number of yield strata common to both the z th and z' th zones.

The estimate of the mean-squared prediction error of P , the country production estimate, is given by

$$S^2 = \sum_{r=1}^K S_r^2 \quad (15)$$

This computation of S^2 is made assuming that the regional production estimates are uncorrelated. This assumption certainly holds with regard to the estimation procedure. However, it is possible for the regional productions to be correlated because of weather and economic conditions.

ZONE AND REGIONAL YIELD AND ITS PREDICTION ERROR ESTIMATION

When there is a single yield model in a zone, the yield prediction and its mean-squared prediction er-

ror are obtained as described in the paper by Strommen et al. In case of more than one yield model in a zone, these parameters (i.e., yield and prediction error) are to be estimated for the zone as well as for the higher levels.

The weighted average yield given by production/acreage is used to determine the combined yield per acre for a higher level. Consequently, an estimate of the yield for level $\{S\}$ is obtained by

$$\bar{Y}_{\{S\}} = \frac{P_s}{A_s} \quad (16)$$

where P_s is the production estimate and A_s is the area estimate for the level (zone, region, or country).

The exact formula for the prediction error of \bar{Y}_s is not tractable because both P_s and A_s are random variables. Only an approximation for the variance of this ratio is considered here. However, good approximation can be achieved because of the large sample property of the acreage estimate A_s , its low coefficient of variation, and the fact that $A_s > 0$.

Using the first-order approximation given in reference 1 (Theorem 5.3), an estimate of the mean-squared prediction error of the ratio estimate \bar{Y}_s is obtained by

$$U_s = \left[\bar{Y}_s^2 \frac{S_s^2}{P_s^2} + \frac{V_s}{A_s^2} - \frac{2 \sum Y_i V_i}{P_s A_s} \right] \quad (17)$$

where S_s^2 is the estimated mean-squared prediction error of P_s , the production estimate; V_s is the estimated variance of A_s , the area estimate; Y_i is the yield estimate for the i th pseudozone; and V_i is the estimated variance of the acreage estimate for the i th pseudozone.

PREDICTION ERROR ESTIMATION FOR MIXED WHEAT AREAS

The mean-squared error estimation problem, mainly for the zone and higher levels, is discussed in this section. Three crop-type pseudozones (pure

winter, pure spring, and mixed wheat) are possible in a zone of mixed wheat. The yield predictions and their mean-squared error estimates are available separately for the pure winter and pure spring pseudozones. On the other hand, a weighted average of the two yield predictions, one for spring wheat and another for winter wheat, would provide a combined yield prediction for a mixed wheat pseudozone. The two weights correspond to the two crop-type acreages. However, it is proposed to use the acreage figures different from LACIE estimated acreages so that the assumption of independence between LACIE acreage estimates and yield estimates is not violated; hence, the formulas given in the two preceding sections are applicable. To avoid within-year dependence, the use of historical acreages is suggested. This method, of course, may cause a certain amount of bias. Thus, for a mixed wheat pseudozone, a combined yield prediction and its mean-squared error estimate, Y_{zi} and U_{zi} , respectively, are obtained as follows.

$$Y_{zi} = \frac{A_{wo} Y_w + A_{so} Y_s}{A_{wo} + A_{so}} \quad (18)$$

$$U_{zi} = \frac{\left(A_{wo} U_{zw}^2 + A_{so} U_{zs}^2 \right)}{(A_{wo} + A_{so})^2} \quad (19)$$

where A_{wo} and A_{so} are the historical (primary epoch) year harvested winter wheat and spring wheat acreages, respectively, in the pseudozone; and where U_{zw} and U_{zs} are the mean-squared prediction errors of the winter wheat and spring wheat yield estimates, respectively, for the pseudozone.

By obtaining inputs for pseudozone yields as described in equations (18) and (19), the mean-squared prediction error estimates are computed for production and yield estimates using equations (12) through (15) and (17).

REFERENCE

1. Raj, Desh: Sampling Theory. McGraw-Hill, 1968.

Classification and Mensuration of LACIE Segments

R. P. Heydorn,^a R. M. Bizzell,^a J. A. Quirein,^b K. M. Abotteen,^b and C. A. Sumner^b

INTRODUCTION

A goal in LACIE has been to estimate wheat acreage to a given accuracy using Landsat rather than ground-enumerated data. In this approach, Landsat data are classified into wheat/nonwheat classes in each of a number of randomly allocated areal segments. The acreage classified as wheat is then measured in each segment and these segment estimates are aggregated to obtain the country estimate. For the method to work, a limited amount of manual interpretation is required for each segment. From this interpretation, spectral samples of the crop types of interest are obtained and used to estimate classification parameter values. These parameter values specify the classification rule from a given family of possible rules. The process of selecting and labeling samples for estimating classification parameters is commonly referred to as "training" a classifier.

A procedure for manually training a classifier and a method of machine classification were tested in the first two phases of LACIE. These tests revealed a number of shortcomings; consequently, the approach was redesigned for Phase III.

The theory of the classification methods and the functional steps in the manual training process used in the three phases of LACIE are discussed in this paper. In addition, the major problems that arose in using the earlier approach are discussed to reveal the motivation that led to the design for the third LACIE phase.

A problem with both designs was that wheat could not be separated from the other small grains.

Although studies now in progress suggest that such a separation is possible if Landsat acquisitions at certain critical times in the wheat crop calendar are available, conclusive results are not yet available. Since wheat estimates were obtained by ratioing methods from small-grains estimates, the class of interest in the following sections will be referred to as small grains.

PHASE I AND II CLASSIFICATION AND MENSURATION APPROACH

The basic steps that were used in Phases I and II to estimate the small-grains area in a segment are illustrated in Figure 1. That estimate was the result of both manual and machine-processing operations. The manual operations were required to train a classifier; once trained, it classified each pixel (except for pixels designated as nonagricultural or under cloud cover) as either small grains or non-small-grains. Those pixels classified as small grains were then totaled to obtain a segment acreage estimate.

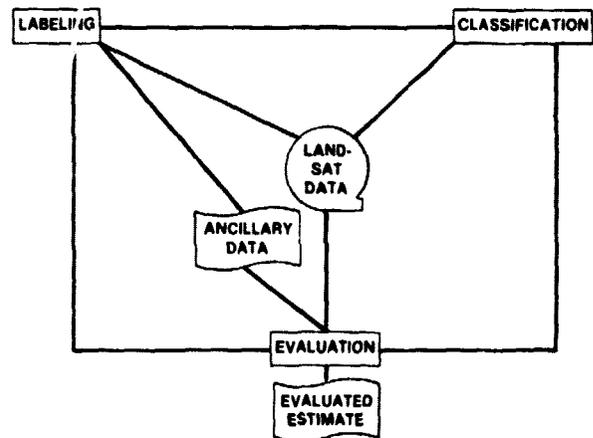


FIGURE 1.—Processing flow in CAMS.

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Houston, Texas. J. A. Quirein is currently with Schlumberger, Houston, Texas; C. A. Sumner is now with Central State University, Edmond, Oklahoma.

Aside from certain screening operations, the starting point of this method was the analyst labeling operation. The analyst was required to choose examples of small-grains and non-small-grains fields that would result in good estimates of the probability density of each of these two major classes.¹ To obtain a good estimate of a class density, two fundamentally different selection processes were required. First, the classification model, as will be discussed later, assumes that a given class can be represented by a sum of normally distributed densities. Thus, the analyst would first specify the number of such distributions to be used in the model and then select example fields from which to estimate the means and covariances of each of these distributions. After the analyst selected and labeled these example fields, the fields were approximated by polygons and the vertices of these polygons inserted into the computer. At this point, the segment was ready to be machine processed.

Certain areas within an agricultural scene that are not agricultural areas can appear spectrally very similar to a small-grains area on a given date. A common example of this is grassland. Since these areas are generally easy for the analyst to spot, they are excluded in the labeling process by again bounding them by polygons and labeling them "designated other" or "DO." These DO areas are not classified but are automatically included in the non-small-grains count.

As mentioned previously, certain screening operations are performed. These operations include checking the imagery to ensure that it is, in the opinion of the analyst, a processable image. Two factors considered in making this decision are image data quality and the crop phenology at the given acquisition time. If possible, it is desirable to process an image taken within the wheat phenology intervals of planting to emergence, emergence to jointing, jointing to heading, and heading to harvest. Another screening operation is performed to exclude clouded

areas from processing.² This is done by enclosing these areas with polygons and labeling them as "designated unidentifiable" or "DU" areas.

Following the labeling operation, the entire segment, less the DO and DU areas, is classified into small-grains and non-small-grains areas. The underlying model on which these classification decisions are based assumes that each major class (i.e., the small-grains and non-small-grains classes) could be described by a mixture of multivariate normal densities. That is, if f_k denotes the class density of the k th class, then

$$f_k(\cdot) = \sum_{i=1}^{n_k} \rho_i^k N(\cdot; \mu_i^k, \Sigma_i^k), \quad k = 1, 2 \quad (1)$$

where $N(\cdot; \mu_i^k, \Sigma_i^k)$ is the i th component normal density function for the k th class. The parameters ρ_i^k , $i = 1, 2, \dots, n_k$, are generally assumed to be $1/n_k$ and the parameters n_k (the number of normal clusters), μ_i^k (the mean vector), and Σ_i^k (the covariance matrix) are estimated from the training field data.

Once the densities f_1 and f_2 are estimated (the estimates are denoted by \hat{f}_1 and \hat{f}_2 , respectively), the classification rule for classifying a given pixel x is as follows: "Decide x is from class 1 if $\pi_1 \hat{f}_1(x) > \pi_2 \hat{f}_2(x)$ and from class 2 if $\pi_1 \hat{f}_1(x) \leq \pi_2 \hat{f}_2(x)$." The weights π_1 and π_2 are nonnegative and add to 1. Their purpose is to estimate the prior probability that a pixel is from class 1 or 2; however, because this prior information is generally not known, the weights are usually taken to be 1/2.

Certain outlier pixels (pixels within an extremely small probability density) are thresholded in the classification process. These thresholded pixels are subsequently counted as non-small-grains. (The details of thresholding are similar to those discussed in the section on classification.)

The percentage of small grains in a segment was then estimated to be

$$\hat{p}_{SG} = \frac{N_{SG}}{22,932 - N_{DU}}$$

where N_{SG} is the number of pixels classified as small grains and N_{DU} is the number of pixels in the DU areas. (Note that there are 22,932 pixels in a LACIE segment.)

¹The estimation of a probability density for a class is done by estimating parameters (means and covariances) of normal density functions. This is accomplished in the classifier training process discussed in the introduction.

²Images are automatically screened as part of the registration process at the NASA Goddard Space Flight Center to ensure that a given segment has less than 10-percent clouded areas. A second screening is done at the NASA Johnson Space Center to doublecheck the automatic process.

Before the estimated proportion of small grains in a segment was passed to the Crop Assessment Sub-system, it was checked by the analyst. Basically, this check was made by comparing the segment imagery to a classification map. The classification map is another piece of imagery on the same scale as the segment imagery with all areas classified as small grains appearing clear and all other areas opaque (see fig. 6). If the analyst believed that areas were not classified as he thought they should be, the segment was reworked. Fundamentally, the rework strategy attempted to correct field-labeling errors and to select new fields that were samples from crop spectral distributions not originally sampled in the training process.

In labeling a given field, the analyst would examine a sequence of imagery ranging over times which included the major crop phenological stages. For the most part, however, machine classifications were done on only single-pass Landsat data. Some multitemporal classification (in which each image pixel is described as a vector of Landsat data from more than one acquisition) was done, but, in general, such attempts proved to be too difficult to execute for all segments. A type of machine processing called "no-significant-change processing" was done; it attempted to select a single acquisition that would provide a good estimate. In this process, the Landsat image from a new acquisition would be visually compared with the class map from the last classification for the given segment. If the analyst decided that there was poor agreement between the two images, then the new image would be classified. On the other hand, if through this visual comparison it was decided that the previous classification was still valid (i.e., no change), no new classification would be made.

Instances arose where very little wheat could be found in a segment. In those instances, the segment was hand-counted and no machine processing was done.

PROBLEMS WITH THE PHASE I AND II APPROACH

The Phase I and II technique used by the analyst is probably better described as an art than as a well-defined procedure. Basically, the analyst was required to interpret color imagery and ancillary numerical data to make decisions related to complex statistical questions, such as "How many normal dis-

tributions will fit the data?" or "How many fields should be sampled to determine an accurate estimate for a given distribution?" Given enough time and the ability to execute a trial-and-error process, a good analyst can obtain a reasonably good answer for a segment. But in a highly automated environment, as LACIE was required to be in order to make estimates at regularly scheduled intervals, such an approach can be inefficient. Moreover, because of the subjective nature of the process and the varying degrees of talent among the numerous analysts who participated in LACIE, the results can contain considerable variance and bias. Some of the specific pitfalls that were inferred from the LACIE results are discussed in the following sections.

Selecting and Labeling Fields

For the machine classification process to work properly, the underlying assumptions of the model given in equation (1) should be satisfied. This model assumes that the data can be statistically described by a sum of normal densities. Thus, in his selection of fields, the analyst must in effect decide how many normal densities are likely to be present. If the decision is too high, a large number of parameters have to be estimated (i.e., several means and covariance matrices); this implies that the number of fields that need to be labeled to obtain good estimates is very large.³ If the decision is too low, it is likely to lead to a poor fit of the model no matter how well the parameters are estimated. Even if a correct decision is made on the number of distributions, a decision must be made as to which fields are samples from a given distribution. This decision is especially difficult when the given distribution has a relatively large variance, because samples must then be drawn from the extremities (tails) of that distribution as well as from the center (around the mean). Samples drawn only around the mean will lead to gross underestimates of distribution variances. It is more natural for an analyst to label a sample far from a given distribution mean as being an observation sampled from a different distribution; i.e., one whose mean is closer to that observation.

³Because the within-field variance is generally much larger than the between-field variance, just choosing large fields is in general not the answer. Hence, to estimate a given parameter, several fields are needed.

Efficiency

The need to reclassify a segment, perhaps several times, to improve the estimate can be a time-consuming process. The amount of rework was not always proportional to the size of the area misclassified. Even a small area classified incorrectly could cause a problem.

Other examples of inefficiency existed in part because of the mechanics of the experiment implementation. A particular example was the requirement to select fields by approximating them with polygons. This procedure led to long delays in preparing computer cards because batch processing rather than interactive processing was used.

Small-Field Versus Large-Field Processing

The Phase I and II procedure was more adaptable to areas that contained large agricultural fields. In areas where the fields were small (e.g., the spring wheat strip-fallow fields in North Dakota), these procedures proved to be difficult to implement. In general, as the size of the fields decreases, the required number of training fields increases. It became increasingly difficult to determine the number of distributions in the model and to choose an appropriate number of samples from each distribution.

Multitemporal Processing

It was well known in the beginning of LACIE that much of the information that distinguishes one crop from another can be obtained from spectral observations over time. In fact, the analysts worked on the premise that, by knowing the crop calendar, it should be possible to relate spectral changes to phenological crop changes and thereby identify that crop. Unfortunately, attempts to machine classify multitemporal data were largely unsuccessful. The underlying reason for this failure is probably related to the problem of selecting and labeling sample fields. To effect a multitemporal classification, an analyst not only must make decisions of the type discussed in the section on selecting and labeling for each image but also must account, again by sampling, for the possible additional classes and covariance terms that arise from the multivariate nature of the problem. Thus, the difficulty in choosing good training fields is greater in such multitemporal applications.

INTRODUCTION TO PROCEDURE 1

Motivated by the problems experienced with the Phase I and II design, a second approach, called Procedure 1, was designed and adopted in Phase III. This design proved to be a significant improvement in terms of both estimation accuracy and efficient use of analyst abilities. More data could be processed with greater accuracy using the same manual resources. A key feature in this improvement was that the analyst was freed to concentrate on the labeling function. Machine processing was used to reduce the variance of an analyst-derived area estimate and to improve labeling accuracy. The classification of a segment was treated as a stratification of that segment into "probably small-grains" and "non-small-grains" strata. Through the use of a poststratified estimation method, the variance of a simple randomly allocated analyst estimate was reduced. Moreover, the ability to cross-check between machine classification and analyst labeling of the same areas and the introduction of analyst labeling aids were elements of the design aimed at improving analyst labeling accuracy.

The analysis of a given segment in Procedure 1 can be described in terms of four interrelated operations, which will be called labeling, classification, area estimation, and evaluation. These operations generally follow the sequence illustrated diagrammatically in figure 2. Labeling refers to all manual functions that result in the assignment of a label to certain specified pixels within the Landsat segment image. The purpose of labeling is threefold: (1) to provide observations from small-grains and non-small-grains classes that are needed to estimate certain classifier parameter values, (2) to provide observations for a stratified area estimate of small grains, and (3) to provide observations for testing the quality of the segment estimates. The classification operation sorts each pixel in a segment into one of two possible classes. The result is a class map, which is subsequently treated as a stratification of the segment area into two (not necessarily connected) regions. Within the limits of classification error, the first region contains pixels primarily of the first class and the second region contains pixels primarily of the second class. Given this stratification, area estimation is performed. This is a stratified area estimate using a second set of labeled dots (independently selected from the set used to estimate classification parameters) allocated within the strata. Finally, the purpose of the evaluation operation is to

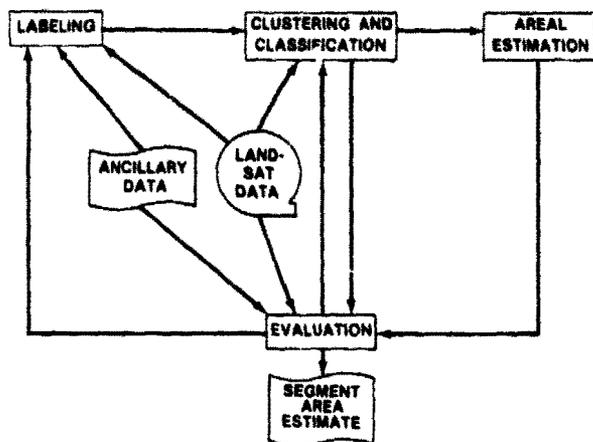


FIGURE 2.—Processing flow in Procedure I.

provide a quality check on the segment estimate and to develop rework strategies if required.

Before an acquisition of a segment is analyzed, it is screened. Two types of screening are used. The first is a manual screening intended mainly to check data quality, including haze distortion and missing pixel data. The second is an automatic screening intended to select from several possible acquisitions the four acquisitions that are likely to give the least classification error. (The details of the acquisition selection process are presented in the section on multitemporal estimation.)

Labeling

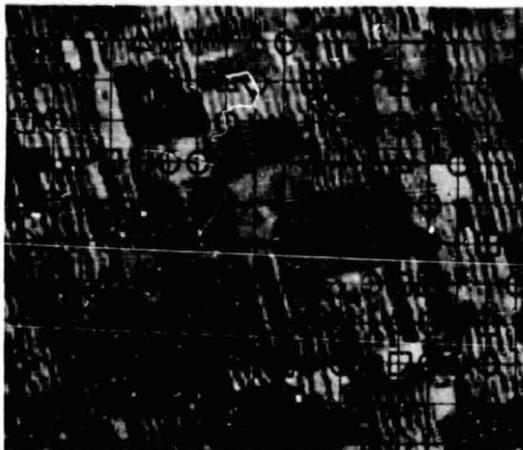
Of the 22 932 pixels in a LACIE segment, 209 (about 1 percent) are selected as a set of candidates to be labeled. These pixels, or dots, coincide with a grid of every tenth column and every tenth row of the segment image. Two randomly and independently chosen subsets of dots (called type 1 and type 2 dots) are selected for labeling. Mechanically, the labeling is done by overlaying a dot template on the scene and labeling each scene pixel directly underneath an indicated dot on the template. An example image and the templates for type 1 and type 2 dot labeling are shown in figure 3. The dot locations appearing on the template have been randomly selected. The same template for type 1 dot labeling is used for each image. Similarly, one template is used for all type 2 dot labeling. Use of the same templates does not violate the intent that the dot selection be random because it is assumed that each segment is a randomly selected

observation from the set of all possible LACIE segments.

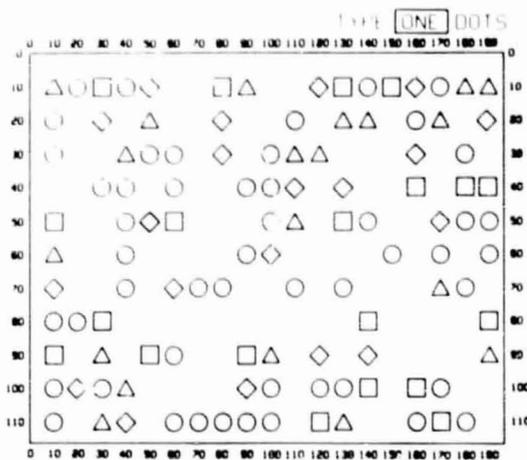
Not all dots in a LACIE segment necessarily fall within an agricultural field. Some can fall on field boundaries, and more will probably fall near field edges so that when a temporal sequence of images is viewed, registration error will cause the dot to appear in different fields on different acquisitions. Since the labeling operation is intended to be a process by which a given pixel is assigned a generic name (e.g., small grains, non-small-grains), it would logically follow that only those pixels that do not fall on or near field boundaries should be labeled. This logic is followed in labeling type 1 dots but not in labeling type 2 dots. Experiments using accurate labeling indicated that skipping boundary and edge dots in type 1 dot labeling produces a classification result not significantly different from the result obtained by including these labels. (For test comparison purposes, boundary and edge dots were assigned labels based on the majority of material represented by the dot.) Moreover, since manual labeling of boundary and edge dots is a highly error-prone process, the decision to skip these dots is appropriate. On the other hand, the type 2 dot labels enter directly into the area estimate; therefore, skipping the boundary and edge dots could bias the estimate. Thus, in type 2 labeling, the analyst must estimate the amount of material in a boundary dot and label that dot according to the majority of material present. In the case of an edge dot, one acquisition is used as a reference and the label of the pixel in that reference image is the one assigned. If the pixel is from an agricultural field, the idea is to select a reference image that clearly contains the dot and assign the field label to the pixel. Note that the field can easily be traced in a multitemporal sequence of images.

The templates in figure 3 are designed to allow for skipping. For type 1 dot labeling, the analyst is required to first label pixels appearing under the circle symbol, skipping boundary or edge dots. If the required number of dots is not labeled, pixels under the square symbol are labeled, again skipping over boundary and edge dots. Finally, if the required number of dots is still not labeled, the triangle symbol is used. Type 2 dot labeling is similar. Skipping in this case, however, is done only if a pixel is in a DO or a DU area (as explained in the following paragraph).

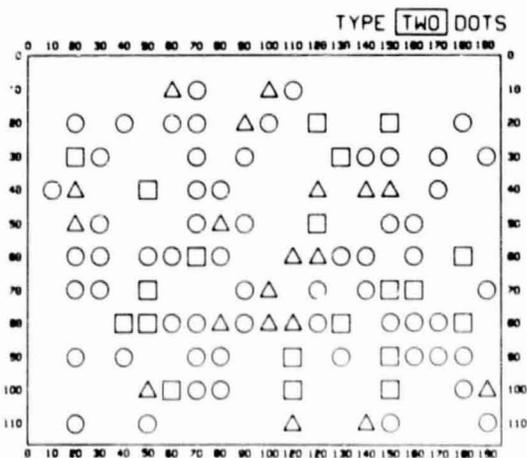
Often in a LACIE segment, there are large areas which are clearly not agricultural. As indicated in the discussion of the Phase I and II approach, these areas



(a)



(b)



(c)

FIGURE 3.—Example scene and type 1 and 2 templates. (a) Landsat scene with a type 2 template overlaid. (b) Type 1 template. (c) Type 2 template.

are called DO areas. Before labeling the dots, the analyst identifies such areas. These DO areas are then skipped in classification and automatically assigned to the non-small-grains portion of the area estimate. Recall that the purpose of assigning an area to the DO category is to eliminate areas which could be spectrally confused by the classifier with a small-grains area. Grasslands are a common example of this category.

The labeling operation uses color-infrared (CIR) imagery, numerical Landsat displays, and ancillary data. Peculiar to Procedure 1 are the trajectory plot and scatter plot displays shown in figure 4. These two displays are intended to present two different types of "information" to the Procedure 1 analyst. The trajectory plot is intended to summarize the "spectral pattern" over time of the crop canopy represented by a dot. A separate trajectory plot is presented for each of the 209 dots. Knowing the nominal pattern of members of the small-grains class for example, an analyst can estimate the likelihood that a given trajectory indicates a small-grains classification. Whereas the trajectory plot can aid an analyst in making decisions about the labeling of a specific dot, the scatter plot is intended to aid the analyst in establishing the consistency of the labeling. The basic idea in the use of a scatter plot is that two dots which are very close are likely to belong to the same class. Both the trajectory plot and the scatter plot are intended to be aids and not infallible indicators of crop type. Indeed, certain classes of grass display trajectories very similar to those in the small-grains class. Also, since two spectral classes can be very close together and in fact have distributions with intersecting supports, the concept of proximity in a scatter plot does not always lead to correct classifications.

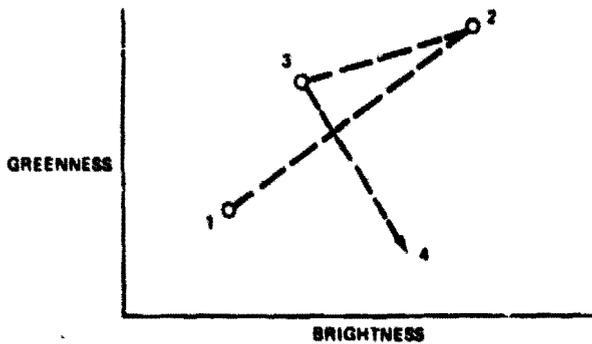
Both the trajectory plot and the scatter plot are plotted against two coordinates known as brightness (abscissa) and green number (ordinate). (For a detailed discussion of these coordinates, see the paper by Kauth and Richardson entitled "Signature Extension Methods in Crop Area Estimation.") The values of these coordinates are obtained from an affine transformation⁴ of the four-dimensional

⁴For Landsat-2 data, the A matrix is

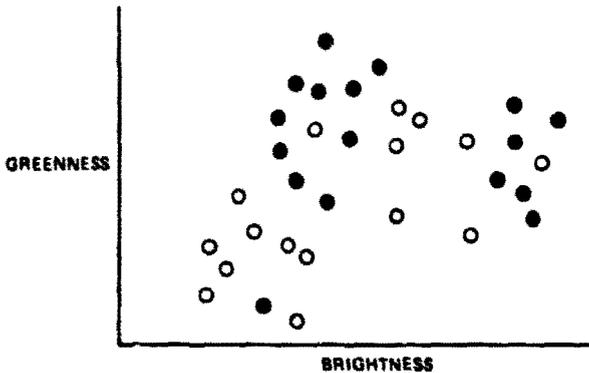
$$\begin{bmatrix} 0.33231 & 0.60316 & 0.67581 & 0.26227 \\ -0.28317 & -0.66006 & -0.57735 & -0.38833 \end{bmatrix}$$

Calibration differences in the satellite data alter this transformation. The vector \mathbf{b} is estimated for each segment and acquisition and is intended to normalize greenness values of bare soil to zero.

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR



(a)



(b)

FIGURE 4.—Example trajectory and scatter plots. (a) Trajectory plot. (b) Scatter plot.

(single-pass) Landsat data; i.e., letting x represent a vector of values associated with a Landsat pixel, the coordinate values y would be $y = Ax - b$, where A is a 2 by 4 matrix and b is a 2 by 1 vector.

Experiments have shown that much of the information in Landsat 1 and 2 data is concentrated in a two-dimensional subspace of the Landsat four-dimensional space (ref. 1). This means that information displays can be constructed to graphically represent important data characteristics on a two-dimensional surface (i.e., a piece of paper). The affine transformation represents an attempt to rotate the coordinate system in that two-dimensional subspace to bring out properties of the data related to soil brightness and canopy growth (or green development). Although there is no claim that this transformation provides a means of measuring these two properties precisely (even if they were precisely defined), experiments have shown that the representation is reasonable (see the paper by Kauth and Richardson). In particular, one finds that the bright-

ness coordinate value for dark soils is less than the value for lighter soils. As a wheat crop grows, its green-number value increases until the wheat reaches the senescent period. After that period, the green number decreases, presumably because the crop is yellowing.

Classification

In Procedure 1, the distributions of the classes of interest (e.g., small grains and non-small-grains) are approximated by a weighted sum of normal distributions. The number of distributions in these sums and the estimates of the parameters related to these distributions are obtained through a clustering process. The algorithm used for clustering is called ISOCLS (ref. 2).

ISOCLS sorts the pixel spectral data in the LACIE segment into a set of clusters, and the elements of a cluster are treated as a sample from one normal distribution. As will be explained in more detail later, these samples are then used to estimate the parameters (i.e., the mean vector and covariance matrix) of that distribution.

The initial iteration of the algorithm operates on a k -means principle (ref. 3) according to which clusters are formed by grouping all points in the segments according to their distance from a given set of points, called "seed" points. Subsequent iterations augment this k -means principle with a split/combine logic in which clusters judged to be too large are split to form smaller ones and clusters judged to be too small are combined with existing larger clusters. The main operations in the algorithm are shown in figure 5. In that figure, x_i , $i = 1, \dots, n$, denotes the i th element of a multispectral pixel vector in a LACIE segment. (In a typical multitemporal clustering, n is 16 and there are 22 932 such vectors.)

There are several parameters that control the algorithm. (These are indicated by quotation marks in figure 5.) The values of these parameters must be specified before the operation of the algorithm. The first parameter is called NMIN. Its purpose is to specify the lower limit of the number of points that can be in a cluster. Small clusters are eliminated because they are likely to represent isolated groupings in spectral regions of extremely small probability densities and therefore are likely to be "outlier" points from some larger cluster. Another parameter is STDMAX. It specifies the maximum sample standard deviation of the elements of the pix-

el vectors in a cluster. In essence, this parameter bounds the volume of a cluster and, depending on the dispersion of data, is a partial control on the number of clusters that can be formed. The parameter DLMIN specifies a minimum distance between two clusters. This parameter is intended to eliminate "false modes" from the data (i.e., two clusters that are more likely to contain samples from a common unimodal distribution than from two separate distributions). The number of clusters is determined by cluster sizes and their volumes (as explained previously) up to a maximum number of clusters specified by "CO" and also by a specified logic sequence called the "split/combine" sequence. An example of a split/combine sequence might be SSSCCS, which would require that four cluster-splitting operations be done, followed by two combining operations, in turn followed by a final split operation. The purpose of including the split/combine sequence is to provide the algorithm with additional information which may be known about the cluster structure of the data other than that information which is automatically obtained from splitting and combining operations based on cluster volume (as measured by the standard deviation) and cluster size (as measured by the number of samples in a cluster). In the final step of the algorithm, small clusters are combined, where small clusters are defined by the parameter $P(N)$ and the number of channels in the data.

The algorithm can be run in a so-called "nearest-neighbor" mode. In this mode, the samples are grouped around the seed vectors, the sample means and covariance matrices are computed for each cluster, and the algorithm is terminated. Experiments have shown that the nearest-neighbor mode of operation is somewhat more predictable than the full-iterative mode (see the paper by Heydorn et al. entitled "An Evaluation of Procedure 1"). This is presumably because the cluster labeling (discussed in the following paragraph) is more predictable since the cluster means are likely to be closer to the seed vectors than when iterative clustering is performed. The nearest-neighbor operation can be specified by appropriate setting of the controlling parameters.

Once the clusters have been formed, the cluster means are compared with each labeled type 1 dot vector. The label of the dot vector that is closest (as measured by the Euclidean metric) to a given cluster mean is assigned to that cluster. In this way, every cluster is automatically labeled from a given set of type 1 dots.

A Bayesian approach is taken to classify every (non-DO and non-DU) pixel in the LACIE segment.⁵ The approach is implemented by first estimating density functions and prior probabilities for the small-grains and non-small-grains classes. Class density functions are modeled as mixtures of normal densities. Thus, the density function for the k th class (where $k = 1$ denotes the small-grains class and $k = 2$ denotes the non-small-grains class) can be expressed as

$$f_k(\cdot) = \sum_{i=1}^{n_k} p_i^k N(\cdot; \mu_i^k, \Sigma_i^k), \quad k = 1, 2 \quad (1)$$

where $N(\cdot; \mu_i^k, \Sigma_i^k)$ is the i th component normal density function for the k th class. The prior probability p_i^k , the mean μ_i^k , and the covariance matrix Σ_i^k are the parameters that therefore specify the exact class density. The parameter set $(p_i^k, \mu_i^k, \Sigma_i^k)$ is estimated from all the points which fall in the i th cluster that has been assigned to the k th class in the above-mentioned labeling process. Generally, the prior probability p_i^k is estimated as

$$\hat{p}_i^k = \frac{N_i^k}{\sum_j N_j^k}$$

where N_i^k is the number of points in the i th cluster assigned to the k th class. However, one can introduce other convex sets of numbers if they are considered to be better estimates. The mean μ_i^k and covariance matrix Σ_i^k are estimated using the usual sample estimators.

Let $\hat{f}_k(\cdot)$, $k = 1, 2$, denote the estimate of the density function, which is obtained by substituting the estimates of the parameters in equation (1). Then, the classification rule for assigning any non-DO or non-DU pixel x in the scene to class 1 or class 2 is as follows: "Decide x is a pixel from class 1 if $\hat{\pi}_1 \hat{f}_1(x) > \hat{\pi}_2 \hat{f}_2(x)$ and a pixel from class 2 if $\hat{\pi}_1 \hat{f}_1(x) \leq \hat{\pi}_2 \hat{f}_2(x)$ ".

⁵This approach is similar to the one discussed earlier in relation to equation (1). The main differences are in the way parameters are estimated.

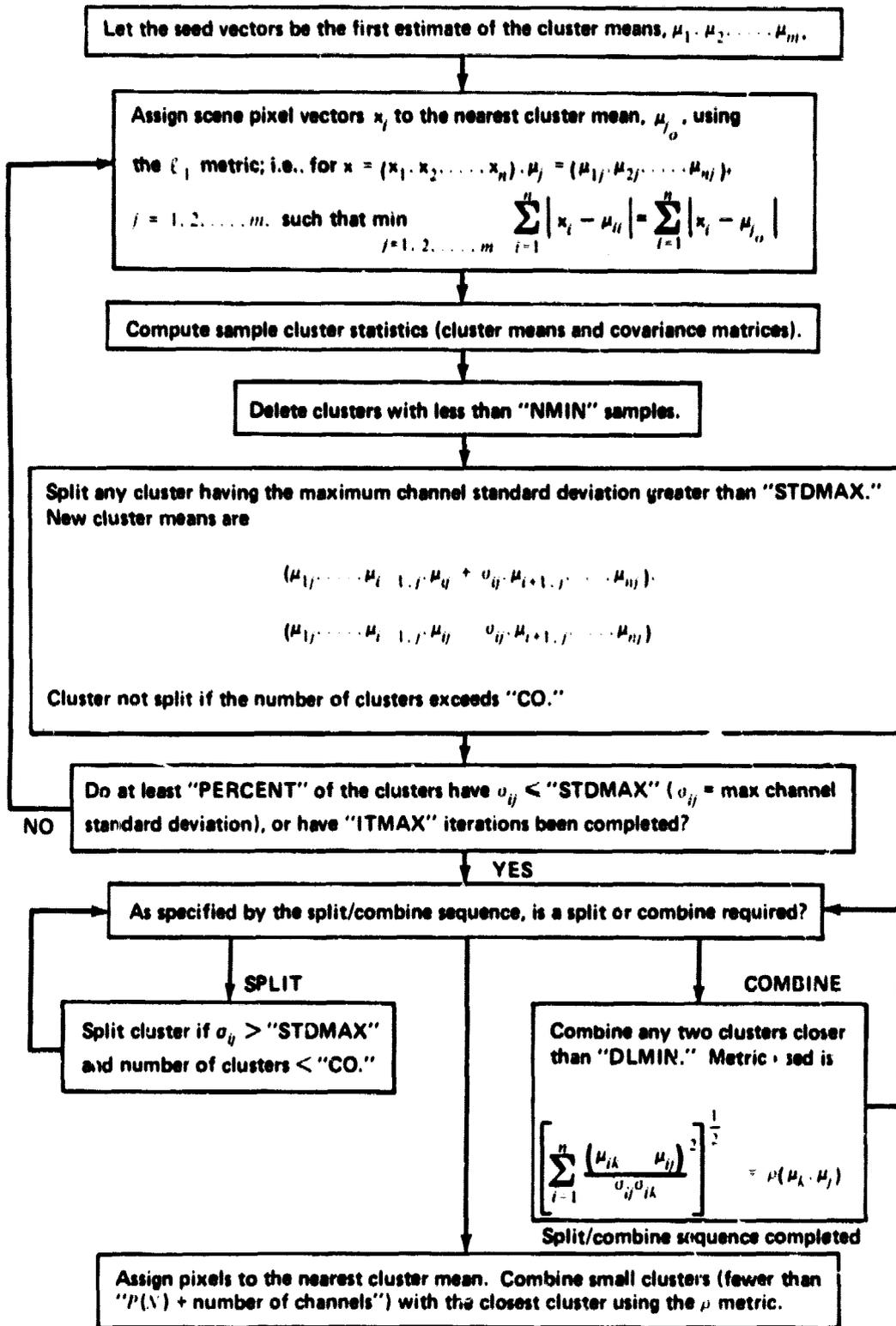


FIGURE 5.—ISOCLS algorithm.

$\hat{\pi}_2 \hat{f}_2(x)$." Here

$$\hat{\pi}_k = \frac{\sum_i N_i^*}{\sum_{i,k} N_i^*}, k = 1, 2$$

As before, any two convex numbers can be used as estimates of π_1 and π_2 if ancillary information indicates that these alternative numbers would be better estimates.

The pixels that are assigned to DO areas in the screening process are automatically tabulated with those classified as non-small-grains. The pixels assigned to DU areas are left unclassified since they represent areas under clouds and therefore are of unknown classification.

Pixels that are a large distance away from any mean of a small-grains cluster are thresholded; i.e., they are not classified as small grains but rather are assigned to the non-small-grains class. To decide whether or not a point x should be thresholded, it is compared with each small-grains cluster sample mean \bar{x}_i using squared metric

$$d^2(x, \bar{x}_i) = (x - \bar{x}_i)' S_i^{-1} (x - \bar{x}_i) \quad (2)$$

where S_i is the sample covariance matrix for the i th cluster. If $d^2(x, \bar{x}_i)$ exceeds a given threshold value δ_i for all indexes i related to the small-grains clusters, then that pixel is thresholded. The threshold is selected so that $Pr[d^2(x, \bar{x}_i) > \delta_i] = 0.01$. Because $d^2(x, \bar{x}_i)$ has an F distribution with p and $N_i - p$ degrees of freedom, where N_i is the sample size (used to compute S_i and \bar{x}_i) and p is the dimensionality of the x vector, the threshold values are tabulated in standard references.

Area Estimation

The small-grains area estimate is obtained by combining an analyst's estimate with the machine estimate. From a statistical sampling point of view, the estimate is a stratified area estimate where the stratification is performed after the type 2 dot allocation (poststratification) or before the allocation

(prestratification). In the latter case, the type 2 dots are allocated in proportion to the stratum sizes.

In LACIE, poststratification was used since the type 2 dots were labeled before a machine classification was obtained. From the analyst's point of view, this type of allocation is simpler to use since the same type 2 dots are labeled at each acquisition. In the other allocation, the location of the type 2 dots would be a function of the classification result and hence would require that, in multitemporal applications, the analyst would label more dots. The prestratification approach has the advantage that the resulting area estimator has a smaller variance than the poststratified estimator. However, as the number of type 2 dots increases, the variances of the two estimators converge. The estimator for both the poststratified and the prestratified estimates is given in the following paragraphs.

For the case where the machine classification produces nonempty strata for both the small-grains and the non-small-grains stratum and type 2 dots fall in both strata,⁶ the estimator is

$$\hat{p}_n = \hat{p}_{11}(n)\lambda + \hat{p}_{10}(n)(1 - \lambda) \quad (3)$$

where $\hat{p}_{11}(n)$ = the number of type 2 dots called class 1 (small grains) by the analyst and classified as class 1 by the machine divided by the number of type 2 dots machine classified as class 1

$\hat{p}_{10}(n)$ = the number of type 2 dots called class 1 by the analyst and classified as class 0 (non-small-grains) by the machine divided by the number of type 2 dots machine classified as class 0

$$\lambda = N_1/N$$

N_1 = number of pixels machine classified as class 1

N = 22 932 minus the number of DO pixels minus the number of DU pixels minus the number of

⁶In the poststratified case, it is possible not to have any type 2 dots fall in a given stratum. In the prestratified case, the allocation is controlled and this will not happen provided enough dots are allocated.

thresholded pixels

n = index denoting the total number of type 2 dots used in the estimate

For the case where the machine classification produces an empty stratum or if no type 2 dots fall in a given stratum, a simple random sample estimate based only on the type 2 dots is used.

Evaluation

Each segment estimate in Procedure 1 is checked in an effort to spot estimates that deviate by large amounts from the true value. If the estimates \hat{P}_{11} and \hat{P}_{10} in equation (3) are a result of an unbiased process (which could be the case if labeling were done by ground sampling), then \hat{P} in equation (3) is an unbiased estimate of the true proportion. Under these conditions, $(N_1/N) - \hat{P}$ is an unbiased estimator of the classifier bias (i.e., an estimator that permits a test of the hypothesis that the proportion estimate N_1/N directly from the classifier is "too far" from the true value P). This would imply that classification errors of omission or commission or both are large and hence that the stratification done by the classifier is resulting in an inefficient process. (See the paper by Heydorn et al. for a discussion of the relation between the omission and commission errors and sampling efficiency.) Unfortunately, because of labeling error, the manual labeling process is not unbiased at the segment level and, consequently, under these conditions, the above statistical test merely accepts or rejects the hypothesis that the classifier estimate is "too far" from the analyst-expected value. When manual labeling is used, the underlying philosophy in the evaluation process is to check the consistency between the machine classification and the manual labeling.

A variety of statistics and visual displays is provided for evaluating a given estimate. As a result of the machine classification process, a classification map, a cluster map, a conditional cluster map, and a classification summary are produced. Examples are shown in figures 6 and 7. The classification map is a transparency the size of a LACIE CIR segment image in which small-grains areas are clear and non-small-grains areas are opaque. If the small-grains class is split into winter small grains and spring small grains, then the winter-small-grains areas are clear and the spring-small-grains areas are gray. A visual evaluation of machine classification performance can be made by overlaying the classification map on

a LACIE CIR image. Evaluation using cluster maps and conditional cluster maps is a similar process. The conditional cluster map is similar to the classification map in that small-grains and non-small-grains areas are clear and opaque, respectively; however, in addition, areas in the image corresponding to spectral clusters whose mean is farther from the closest type 1 dot than a specified threshold value are printed in a color. These colored areas are then "flagged" for the analyst as regions where the spectral values are not representative of the spectral values of the type 1 dots. The analyst can then judge whether the areas corresponding to these conditional clusters have been correctly or incorrectly labeled in the automatic cluster-labeling process. The cluster map is a map in which an area corresponding to a given spectral cluster is assigned a distinct color.

Multitemporal Estimation

By observing the biomass development of a crop canopy at specific phenological stages, it is possible to discriminate that crop from generically different crops with reasonable accuracy. This fact has motivated the use of crop calendar predictions in the labeling process and multitemporal classification in the machine processing. Consequently, Procedure 1 was designed to process Landsat data in a multitemporal fashion. The basic multitemporal strategy is to acquire Landsat data at least once within each of up to four time intervals of the small-grains crop calendar scale (if the Landsat coverage is not obscured by cloud cover), concatenate the Landsat pixel observations into an n -dimensional vector (where n could be 4, 8, 12, or 16, depending on the number of Landsat acquisitions used), and classify those vectors as discussed previously.

If Procedure 1 is to be used to make sequential estimates (as was the case in LACIE where estimates were made in the early, middle, and harvest portions of the wheat-growing season), then, since a maximum of four acquisitions can be used, it is desirable to select the best set of acquisitions to process. The strategy followed in Procedure 1 is based on the assumption that, as the season progresses, later acquisitions will be better for crop discrimination than the worst of the acquisitions that have already been processed. The strategy is implemented by estimating a quantity related to the classification error that would be incurred by classifying the best $M - 1$ of M acquisitions previously processed. The acquisition

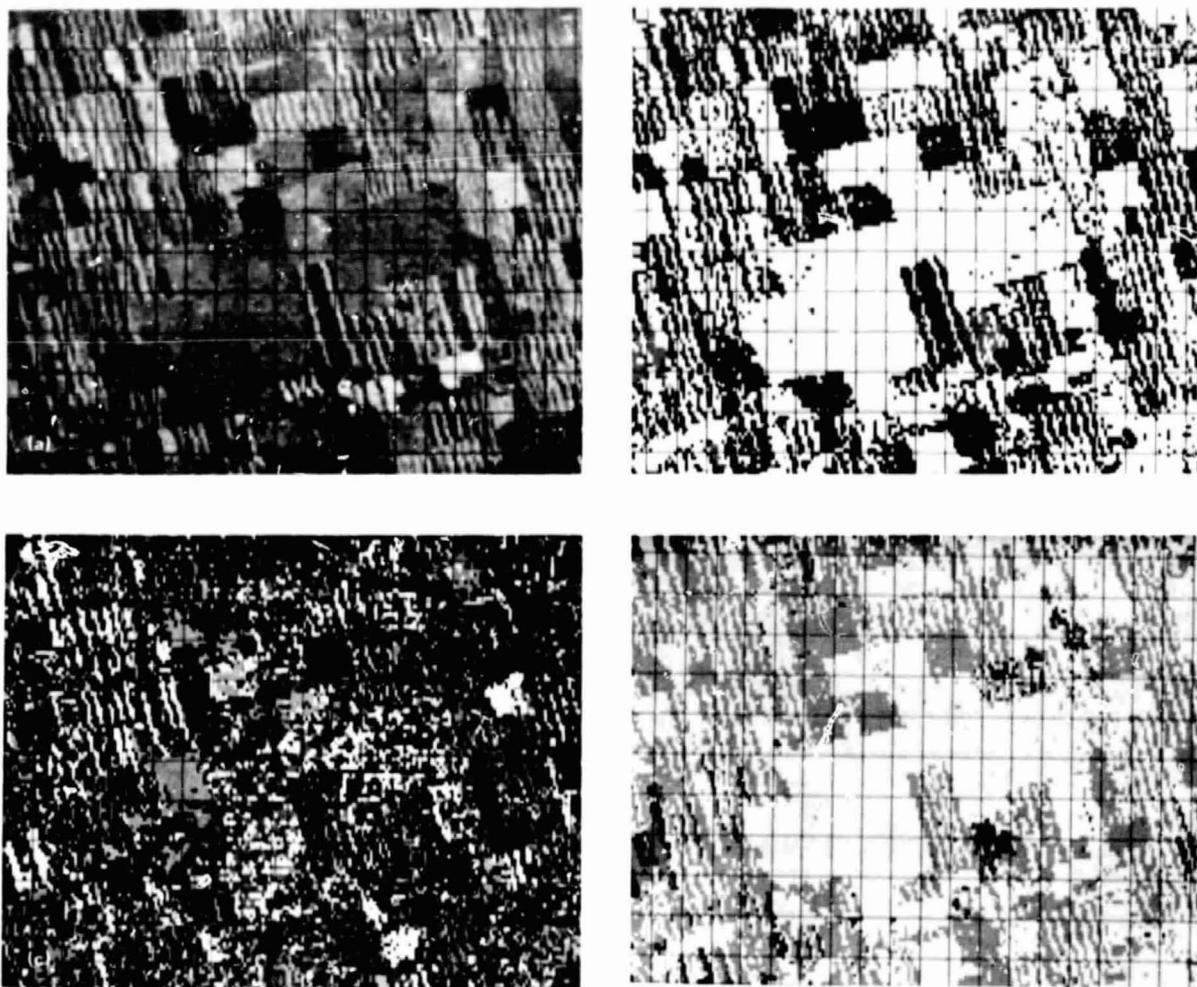


FIGURE 6.—Examples of various maps for a given Landsat CIR image. (a) Landsat CIR image. (b) Classification map. (c) Cluster map. (d) Conditional cluster map.

that is dropped from the M given acquisitions is considered to be the worst, and it is replaced in the multitemporal processing by the newly acquired acquisition. As an example of this strategy, consider the case where four acquisitions have already been processed and another has been acquired. In this case, classification performance for each combination of three acquisitions from the four already in the data base (this does not include the newly acquired acquisition) is estimated. The best combination is retained and combined with the new acquisition to form a new combination of four acquisitions.

Classification performance is estimated using the average Bhattacharyya coefficient between component distributions of the two competing classes. Thus, if $N(\cdot; \mu_i, \Sigma_i)$, $i = 1, 2, \dots, m_1$, are the compo-

nent distributions of the first class (e.g., the small-grains class) and $N(\cdot; \nu_j, K_j)$, $j = 1, 2, \dots, m_2$, are the component distributions of the second class (e.g., the non-small-grains class), then the Bhattacharyya coefficient for the i th and j th distributions of opposite classes is

$$B_{ij} = \frac{1}{\sqrt{|\Sigma_i|}} \frac{1}{\sqrt{|K_j|}} \frac{1}{\sqrt{|\Sigma_i + K_j|}} \exp\left\{-\frac{1}{4}(\mu_i - \nu_j)'(\Sigma_i + K_j)^{-1}(\mu_i - \nu_j)\right\}$$

CLASSIFICATION SUMMARY REPORT
DOT SUMMARY

DOT NUMBER	DOT LOC LIN PXL	DOT TYPE	DOT LABEL	CATEGORY CLASSIFIED	DOT BRIGHTNESS			DOT GREEN NUMBER					
					A1	A2	A3	A4	A1	A2	A3	A4	
3	10	30	1	S	W	79	45	31	60	18	4	2	0
4	10	40	S	W	W	74	94	28	68	20	3	5	4
15	10	150	S	N	N	85	81	37	86	5	0	0	6
17	10	170	S	N	N	86	71	29	75	5	5	2	4
30	20	110	S	W	N	85	61	36	61	7	4	2	4
35	20	160	S	N	N	87	66	31	58	5	5	3	7
39	30	10	S	W	W	77	62	36	72	20	2	0	6
43	30	50	S	N	N	93	88	34	74	4	2	2	2
44	30	80	S	W	W	72	61	30	64	32	2	3	5
48	30	100	S	W	W	80	64	33	71	12	10	0	3
56	30	180	S	W	W	81	88	33	74	17	11	2	6
80	40	30	S	W	W	87	66	44	78	10	8	1	7
86	40	90	S	N	N	74	59	32	59	11	4	1	3
67	40	100	S	N	N	75	66	31	58	9	5	3	6
73	40	160	1	W	W	89	69	32	78	19	3	1	5
75	40	180	S	W	W	83	70	31	69	12	6	3	4
77	50	10	S	N	N	72	66	35	56	5	10	0	4
82	50	60	S	N	N	87	66	29	66	2	3	3	3
86	50	100	S	N	N	77	64	29	53	6	4	3	6
89	50	130	1	N	N	76	66	29	60	6	5	5	5
90	50	140	S	N	N	73	66	31	63	3	8	3	3
94	50	180	S	W	W	75	74	32	67	16	8	1	2
104	60	90	S	N	N	73	65	30	56	4	3	3	7
110	60	150	S	N	N	72	64	29	55	3	11	3	5
111	60	190	S	W	W	77	70	29	76	18	9	3	3
118	70	40	S	N	N	82	68	44	79	3	4	1	5
121	70	70	S	N	N	79	75	33	60	6	8	3	5
122	70	80	S	N	N	76	64	30	57	5	3	2	4
134	80	10	S	S	S	84	53	47	106	16	4	2	6
135	80	20	S	W	W	75	70	29	64	13	10	2	5

FIGURE 7.—Example classification summary (only the first three dots are shown).

The average Bhattacharyya coefficient is then defined as

$$B = \frac{1}{m_1 m_2} \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} B_{ij}$$

CONCLUSIONS

The Procedure 1 design is an attempt to efficiently merge machine classification and analyst interpretation processes. Efficiency is interpreted in terms of both the amount of time spent by the analyst in obtaining an estimate and the variance of that estimate as a function of the number of dots interpreted.

Given that an analyst can label randomly selected dots from an image with reasonable accuracy, a segment estimate can be obtained directly from this labeling process without the intervening machine classification. The benefit of using Procedure 1 is that the classification process will stratify the scene into potential small-grains and non-small-grains areas and thereby can reduce the variance of the analyst's random dot estimate through the stratified area estimation process. The more accurate the classification process, the lower the variance of the estimates.

Classification errors are, of course, dependent on the observations that are to be classified, and one can therefore expect these errors to vary from segment to segment and for different acquisition histories of the Landsat data. Experiments have shown that the variance reduction in using Procedure 1 over a simple random sampling approach differs considerably from segment to segment (see the paper by Heydorn et al.). On the average, however, this reduction appears to be near 0.7 when compared to the variance that can be achieved using only the type 2 dots for the simple random estimate. This suggests that improvement in classification methodology is required, assuming that the Landsat data are more than just marginally effective in discriminating crop types.

When compared to the Phase I and II design, Procedure 1 provides a framework wherein the interaction between analyst and machine is more controlled and therefore more easily studied. Hence, the Procedure 1 approach offers considerable research potential in studying error propagation and a potential means of developing improvements for future designs. As an example, the mere use of a fixed labeling grid has provided a means for more accurately determining the effect of labeling error on classification performance since field selection is controlled and therefore not confounded in the error. As another example, three proportion estimates can be computed for a segment estimate; namely, a direct machine estimate (as given by λ discussed previously), a simple random estimate (as obtained from the labeling of type 2 dots), and the stratified area estimate. Thus, for each segment analysis, it is possible to compare these three estimates and thereby study the advantages of machine and manual processing.

As a final point, it should be noted that Procedure 1 was a design improvement over the Phase I and II design both in terms of increasing the number of segments that could be processed (see the paper by White entitled "LACIE Applications Evaluation System Efficiency Report") and in terms of the accuracy of the estimates (see the paper by Potter et al. entitled "Accuracy and Performance of LACIE Area Estimates").

REFERENCES

1. Wheeler, S. G.; Misra, P. N.; and Holmes, Q. A.: Linear Dimensionality of Landsat Agricultural Data With Implica-

2-2

- tions for Classification. Purdue University, Laboratory for Applications of Remote Sensing (LARS) Symposium, 1976.
2. Kan, E. P. F.: The JSC Clustering Program ISOCLS and Its Applications. LEC-048J, Lockheed Electronics Company, Houston, Texas, July 1973.
 3. MacQueen, J. B.: Some Methods for Classification and Analysis of Multivariate Observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Vol. 1, 1967, pp. 281-297.

LACIE Registration Processing

Gerald J. Grebowsky^a

INTRODUCTION

The basic requirements which have been imposed on the NASA Goddard Space Flight Center (GSFC) LACIE processing system are

1. To extract specified test sites (sample segments) from Landsat multispectral scanner (MSS) data

2. To apply geometric corrections and perform correlations to ensure registration between successive data acquisitions to within 1 pixel (root mean square (rms))

The processing flow necessary to meet these requirements is shown in figure 1. The NASA Johnson Space Center (JSC) defines test sites by specifying the geodetic latitude and longitude of the site center plus the biological window or time window during which MSS data acquisitions are to be extracted. As indicated in figure 1, the output of the GSFC LACIE processing is a digital tape containing sample segments of all four MSS bands. A sample segment consists of 117 lines and 196 pixels representing a rectangular ground area approximately 9.3 by 11.1 km (5 by 6 n. mi.).

The first step in processing is to determine which sample segments are contained in an MSS image frame. The time windows specified for each test site are compared with the MSS data acquisition date to identify the sites desired. The latitude and longitude of each of these desired sites are then compared to the MSS image frame center latitude and longitude. Any site located within $\pm 1^\circ$ of the image center is initially assumed to be contained within the MSS image frame. For each of these sites, the line and pixel location within the MSS image frame is determined. Valid line and pixel numbers provide the final test of whether a sample segment is contained on the MSS image frame.

If available attitude and ephemeris data were precise, the location of the sample segments determined

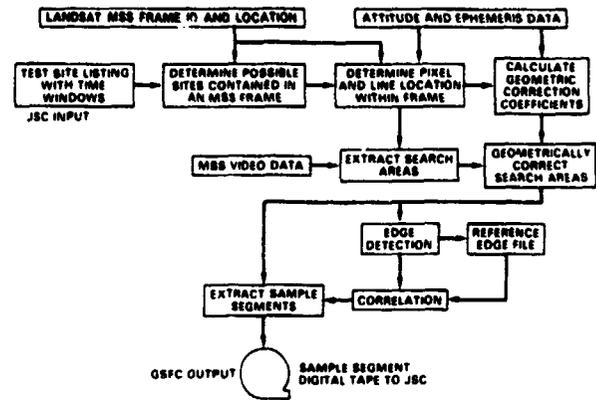


FIGURE 1.—LACIE processing flow.

as described in the preceding paragraph would be sufficient (although a more sophisticated calculation than that actually used would be required). However, errors in attitude and ephemeris data were originally estimated to be as great as ± 4.5 km. These errors required the definition of search areas containing 234 lines and 354 pixels (approximately 18 by 20 km). Provided locational errors do not exceed the ± 4.5 -km limits, the desired sample segment data will be contained within the larger search area. After determination of the estimated line and pixel number location of a test site, geometric correction coefficients are calculated using Landsat attitude and ephemeris data. The line and pixel number location is used to extract search areas from radiometrically corrected (normal Landsat MSS corrections) MSS data. Using these geometric correction coefficients, the search area data for each of the four MSS bands are geometrically corrected using nearest neighbor methods (i.e., corrections rounded to nearest integer pixel).

For the initial acquisition of a test site, the sample segment is assumed to be the center of the search area. Based on previous error estimates, the initial sample segment may be mislocated by ± 4.5 km. However, the initial sample segment becomes the reference, and subsequent acquisitions of a test site

^aNASA Goddard Space Flight Center, Greenbelt, Maryland.

are extracted to coincide with the initial extraction. The location uncertainty of subsequent acquisitions is removed by an edge-dependent correlation scheme. During initial sample segment extraction, an edge detection operation is performed on the sample segment data, and the edge data are stored (edge file) as the correlation reference for later acquisitions. Subsequent search areas are then processed using the same edge detection process, and coincidence of edges relative to the reference edge data determines the location of the correlated sample segment to be extracted. Having reviewed the general processing flow within the GSFC LACIE processing system, the following steps will now be described in detail.

1. Determination of line and pixel location of a search area within an MSS frame
2. Determination of geometric correction coefficients and application of geometric corrections
3. Edge detection
4. Correlation by coincidence of edges

TEST SITE LOCATION CALCULATION

As discussed previously, test sites are specified by the geodetic coordinates (latitude and longitude) of the site center. The calculation of the corresponding search area locations (line and pixel number) within a Landsat MSS frame (computer-compatible tape (CCT)) requires the parameters listed in table I. The diagram in figure 2 illustrates the simple geometry assumed in the location calculations. The X and Y axes are aligned with the scan line direction and the orbit plane, respectively, with the minus- Y axis in the direction of spacecraft heading (i.e., spacecraft velocity). The origin is chosen to be the format center of the Landsat MSS frame (CCT). The Y_1 axis is in the direction of north (i.e., direction of longitudinal meridian through the format center), and the X_1 axis is in the direction of east (i.e., parallel of latitude through format center). The orientation of north (Y_1) and east (X_1) is relative to the format center orbit plane (Y) and scan line (X), respectively. For simplification, it is assumed that parallels of latitude and meridians of longitude are straight lines. It is also assumed that parallels of latitude are orthogonal to the meridian of longitude passing through the format center. Ignoring geometric distortions inherent in raw MSS data, the displacement (Y_1) of a test site relative to the origin is given by the arc length $R(L_a - L_a')$ on a spherical Earth. Similarly, the displacement (X_1) of a test site is given by the arc length $R(L_o - L_o') \cos L_a$ on the circle of a

parallel of latitude. Any estimated errors in format center location can be accounted for by empirical offsets A_i and B_i . The X_1, Y_1 coordinates of a test site are then given by the expressions

$$X_1 = R(L_o - L_o') \cos L_a - A_i \quad (1)$$

$$Y_1 = R(L_a - L_a') - B_i \quad (2)$$

TABLE I.—Input Parameters for Search Area Location

Parameter	Symbol	Value	Unit
<i>Test site</i>			
Latitude ^a	L_a	Variable	rad
Longitude ^b	L_o	Variable	rad
<i>Landsat MSS frame</i>			
Format center			
Latitude ^a	L_a'	Variable	rad
Longitude ^b	L_o'	Variable	rad
Spacecraft heading angle	b	Variable	rad
Earth rotation skew angle	a_f	Variable	rad
Normalized spacecraft altitude change ^c	$\frac{\Delta H}{H}$	Variable	percent
Normalized spacecraft velocity change ^d	$\frac{\Delta V}{V}$	Variable	percent
Lines per MSS frame (CCT)	L_f	2340	lines/frame
Pixels per MSS line (line length)	P_L	3291	pixels/line
<i>Constants</i>			
Earth radius	R	6368	km
Nominal pixel scale	K	0.0565	km/pixel
Nominal line scale	M	0.0799	km/line

^aN = +, S = -
^bE = +, W = -
^c $\Delta H = H - H_N$
^d $\Delta V = V - V_N$

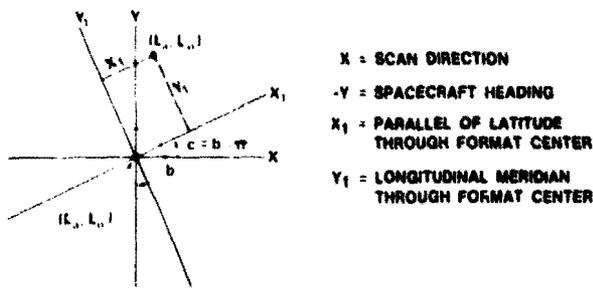


FIGURE 2.—Search area location geometry.

During the early period of operation using Landsat-1 data, errors in attitude values caused format center errors to occasionally exceed the ± 4.5 -km predicted limit. Those errors were found to be seasonally dependent, and measurements of known ground control points in the imagery were used to determine A_i and B_i for each month (i = month). Subsequently, these empirical data were used to adjust the Attitude Measurement System (AMS) model which calculates the Landsat spacecraft attitude. On January 5, 1975, the adjusted AMS model was incorporated into the GSFC operational system for Landsat-1 and Landsat-2 data processing. This change reduced format center errors to ± 2.5 km, and the A_i and B_i parameters have been set equal to zero in the LACIE processing system.

As shown in figure 2, the transformation from X_1, Y_1 coordinates to X, Y coordinates is simply a rotation of angle $c = b - \pi$ as given by

$$X = X_1 \cos c - Y_1 \sin c \quad (3)$$

$$Y = X_1 \sin c + Y_1 \cos c \quad (4)$$

The coordinates used are in units of kilometers. To obtain pixel and line numbers, the pixel and line spacings in kilometers are required. The nominal scales K and M defined in table I are adjusted for frame-dependent scale variations. The significant contributors to scale variations (± 2 percent) are

spacecraft velocity and altitude and the adjusted scales K' and M' are accordingly defined as

$$K' = K \left(1 + \frac{\Delta H}{H_N} \right) \quad (5)$$

$$M' = M \left(1 + \frac{\Delta V}{V_N} \right) \quad (6)$$

Using equations (3) to (6) and recalling that the X, Y origin is at pixel $P_L/2$ and line $L_F/2$, the line (L_c) and pixel (P_c) location of the test site is given by

$$L_c = \frac{L_F}{2} - \frac{Y}{M'} \quad (7)$$

$$P_c = \frac{P_L}{2} + \frac{X}{K'} \quad (8)$$

An adjustment is required in equation (8) to account for Earth rotation. This adjustment requires a shift of $Y \tan a_E$ in the X coordinate to yield the final expressions

$$L_c = \frac{L_F}{2} - \frac{Y}{M'} \quad (7)$$

$$P_c = \frac{P_L}{2} + \frac{X - Y \tan a_E}{K'} \quad (9)$$

Within the GSFC LACIE processing system, the upper left corner of a search is used to extract the

search area data. These data are obtained from L_c and P_c as defined previously by simply subtracting 117 lines and 176 pixels, respectively. A final adjustment in the pixel location, $P = P_c - 176$, is required by the relationship between actual video line length and the parameters P_L and K defined as input parameters. The scale factor, $K = 0.0565$ km, as defined in table I is the pixel spacing at the nadir point. Distances in LACIE algorithms to be discussed later are calculated as true distances on the surface of the Earth. For a nominal MSS total scan angle of 0.201586 rad, the total scan arc on the surface of the Earth is 185.9 km. Thus, at the defined scale, the total of 3291 pixels corresponds to a nominal MSS scan line. However, the pixel location desired for extraction of a search area must be in terms of actual MSS pixels (equally spaced in time, not ground spacing). Since the magnitude of the correction is estimated to be less than 3 percent, a simple scale adjustment by a factor of P_L/LL (LL = actual MSS line length) is considered adequate. The upper left corner of a search area is defined by

$$L = L_c - 117 \quad (10)$$

$$P = \frac{P_L}{LL} (P_c - 176) \quad (11)$$

Although a detailed analysis has not been pursued to define the errors introduced by all the simplifying assumptions, the successful correlation of sample segments within the search areas defined by these algorithms indicates sufficient accuracy for the current GSFC LACIE processing system. The major problems experienced early in the program were errors in format center location, as described previously.

It is appropriate at this point to discuss the use of the all-digital geometrically corrected data (fully processed—HDT-PM tapes) which will be available from GSFC's Information Processing Facility system late in 1978. These data will be resampled to a specified map projection (presently Hotine Oblique Mercator (HOM)) and will include improved location when ground control is available. The orientation of MSS frames relative to the orthogonal coordi-

nates of the map projection will be specified and fixed. The previously described location calculations will then be replaceable by

1. Transformation from latitude and longitude to orthogonal map coordinates
2. Rotation of coordinates based on specified orientation to horizontal and vertical MSS coordinates
3. Division by 0.057 km of both dimensions to obtain pixel and line location

The accuracy of such extractions will be commensurate with the expected accuracies of the geometrically corrected data: ± 1 pixel when adequate ground control is used and ± 44 pixels without ground control. Without ground control, the accuracy of location will continue to be affected by the ± 2.5 -km uncertainty in Landsat attitude values.

GEOMETRIC CORRECTIONS

Four types of geometric corrections, as defined by the diagrams in figure 3, are performed within the GSFC LACIE processing system. As indicated by the vectors in figure 3, these are linear corrections

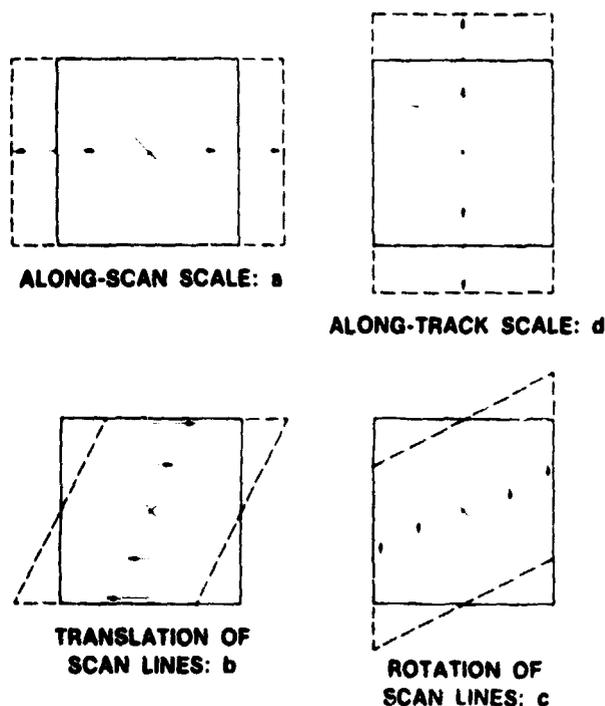


FIGURE 3.—Linear geometric corrections.

with the magnitude of the correction increasing with displacement from the center toward the edges. As indicated in figure 3, each correction is defined by a single correction coefficient a , b , c , or d . The correction algorithms are defined as

$$P' = P + a(P - P_c) + b(L - L_c) \quad (12)$$

$$L' = L + c(P - P_c) + d(L - L_c) \quad (13)$$

where a, b, c, d = correction coefficients defined later

P, L = pixel and line number of an element in corrected data

P', L' = pixel and line number in uncorrected data

P_c, L_c = pixel and line number of center (as determined by eqs. (7) and (9), with P_c value from eq. (9) adjusted by the factor P_L / LL to account for actual line length)

As outlined earlier, the geometric corrections are implemented as nearest neighbor resampling. This simply means that the calculations defined by equations (12) and (13) are rounded to the nearest integer. The resampling or geometric correction process is really a reformatting process—for each location P, L in the desired corrected search area array; equations (12) and (13) are used to calculate the P', L' location in the original array for the pixel value to be used.

Before proceeding with a discussion of the geometric correction coefficients, the order of operations within the LACIE processing flow (fig. 1) should be recalled. The geometric corrections are applied to search areas before edge extraction and correlation. This order was selected to avoid correlation errors or sample segment location errors due to geometric distortions. However, this order does contribute to registration errors. Consider the special case of two data acquisitions with identical geometry (i.e., a, b, c, d equal between acquisitions) but different location errors in extraction of a search area. The calculations of equations (12) and (13) are

referenced to the search area center element, and the location of round-off threshold points (an input pixel either repeated or deleted) will be at the same displacement from search area center for both cases. When the sample segments are extracted, the sample segment centers will not necessarily be the same distance from search area centers for each acquisition. Thus, the round-off threshold points for the two sample segments will not coincide and, in the worst case, 50 percent of the pixels will be misregistered by 1 pixel, although the geometry of the data is identical. These errors can be eliminated by applying the geometric corrections to raw sample segments after correlation is accomplished. Since the present order was selected to ensure correlation success, a test study would be required to determine the effectiveness of correlation before geometric corrections. If corrections are necessary before the correlation step, the sample segment location determined by correlation could be used to extract data from the uncorrected search area at the improved location, and a separate geometric correction could then be applied to the sample segment.

The four geometric corrections defined previously were selected by reviewing the parameters affecting the geometry of Landsat MSS data. In table II, the parameters considered to be significant and estimates of their contributions to registration errors are summarized. The error estimates are worst case—opposite extremes of parameter values and at the edges of a sample segment, where errors are greatest (refer to fig. 3). The maximum absolute values of the correction coefficients using the worst-case parameter values are

$$\begin{aligned} |a| < 0.04 & & |c| < 0.039 \\ |b| < 0.072 & & |d| < 0.024 \end{aligned}$$

The impact of each parameter will become evident with the presentation of the correction coefficient algorithms.

Input parameters required for the calculation of correction coefficients are defined in table III. The along-scan scale coefficient a is defined as

$$a = \frac{K}{K'} - 1 \quad (14)$$

where $K = 0.0565$ km = pixel spacing defined for a corrected sample segment and K' is the pixel spacing in an uncorrected search area. The uncorrected pixel spacing K' is defined by

$$K' = \left(\frac{P_L}{LL} \right) \frac{f(P'_1, P'_2)}{\Delta P'} \quad (15)$$

where

$$f(P'_1, P'_2) = R |\gamma(P'_2) - \gamma(P'_1)|$$

$$\gamma(P') = \sin^{-1} \left\{ \frac{H+R}{R} \sin(\beta + \alpha_R) \right\} - (\beta + \alpha_R)$$

and

$$\beta = A \sin \left(\frac{\omega_s P'}{P_2} + B \right) + C - \beta_c$$

The β calculation accounts for the nonlinear MSS mirror velocity defined by the parameter values listed in table III. Landsat-2 mirror parameters have always been used under the assumption that Landsat-1 and Landsat-3 mirror mechanisms have approximately the same characteristics, and any small differences will be negligible over the size of a LACIE area. The $\gamma(P')$ calculations transform from scan and pointing angle (roll angle, α_R) reference at the spacecraft to subtended arc angle from nadir on the surface of a spherical Earth. The $f(P'_1, P'_2)$ represents the ground arc length on the surface of a spherical Earth. The factor P_L/LL in equation (15) adjusts for the use of P_L in the β calculation (similar to previous line length adjustment discussion). It should be noted here that the use of K in search area location calculations was another approximation. The uncorrected value K' really applies; however, K' is not calculated until after search area location verifies that the sample segment is contained on the Landsat image frame.

The along-track scale coefficient d is defined as

$$d = \frac{M}{M'} - 1 \quad (16)$$

TABLE II.—Maximum Registration Errors Within a Sample Segment

(Uncorrected data)

Parameter	Max. error, percent	
	Pixel	Line
Along-scan scale		
Δ Altitude < 1.6 percent	1.6	—
Mirror velocity < 1 percent	1.0	—
Perspective—Earth		
curvature < 0.5 percent	.5	—
Roll < 1° (=0.1 percent)	.1	—
Root sum square	1.9	—
Translation of scan lines		
Δ Roll ($\pm 0.003^\circ/\text{sec}$) < 0.00017 rad	1.3	—
Adjacent orbit (1.1° rotation)	1.54	—
Root sum square	2.02	—
Along-track scale		
Δ Velocity < 0.2 percent	—	0.12
Δ Pitch ($\pm 0.003^\circ/\text{sec}$) < 0.00017 rad	—	1.0
Root sum square	—	1.01
Rotation of scan lines		
Yaw < 0.6°	—	.75
Adjacent orbit (1.1° rotation)	—	1.31
Root sum square	—	1.51
Total root sum square		
Without adjacent orbits	< 2.7	—
Using adjacent orbits	< 3.3	—

where $M = 0.0799$ km = line spacing defined for a corrected sample segment and M' is the line spacing in an uncorrected search area. The uncorrected line spacing M' is defined by

$$M' = M \left(1 + \frac{\Delta V}{V_N} - \frac{H \Delta \alpha_P}{M \Delta L} \right) \quad (17)$$

The M' calculation accounts for changes in ground distance covered by the fixed number of lines in an uncorrected search area caused by spacecraft velocity variations from nominal and by spacecraft tilting forward or back from first to last line of a search area.

TABLE III.—Definition of Terms

Symbol	Definition	Value	Unit
<i>General</i>			
P, L	Pixel and line number of an element in corrected data	Variable	None
P', L'	Pixel and line number in uncorrected data	Variable	None
P_c, L_c	Pixel and line number of center element of search area	Variable	None
P_1', P_2'	Pixel numbers for first and last elements in an uncorrected search area line	Variable	None
$\Delta P, \Delta L'$	Total number of pixels and lines in uncorrected search area	Variable	Quantity
LL	Number of pixels in a full line of Landsat data (actual line length)	Variable	pixels
P_L	Length of MSS scan line (nominal line length)	3291	pixels
α	Spacecraft yaw angle	Variable	rad (deg)
θ	Rotation angle between adjacent MSS orbit frames	Variable	rad (deg)
$\Delta \alpha_R$	Difference in roll angle from first to last line of search area	Variable	rad (deg)
$\Delta \alpha_p$	Difference in pitch angle from first to last lines of search area	Variable	rad (deg)
$\frac{\Delta V}{V_N}$	Normalized velocity change	Variable	percent
H	Spacecraft altitude	Variable	km
R	Earth radius	6368	km
β	MSS scan angle (0 at nadir)	Variable	rad (deg)
$\gamma(P)$	Arc angle on Earth corresponding to pixel P	Variable	rad (deg)
<i>Landsat-2 MSS mirror velocity parameters</i>			
B_c	—	0.100793	rad
A	—	.36954	rad
B	—	-.26725	rad
C	—	.097588	rad
ω	—	17.0903	rad/sec
t_s	—	.032330	sec

Comments made previously regarding the use of K in search area location calculations also apply to the choice between M and M' .

The scan line translation (horizontal skew) coefficient b is defined as

$$b = \frac{H}{K'} \frac{\Delta \alpha_R}{\Delta L'} + \frac{M}{K'} \sin W \quad (18)$$

The first term accounts for translations between MSS lines caused by the spacecraft tilting to the side

from the first to the last line of a search area. The second term represents one component of the image rotation correction required to register sample segments from adjacent Landsat orbits. Data acquired from noncoincident orbit paths have an inherent image rotation which is corrected in the GSFC LACIE processing system by rotating the vertical image axis through this term and by rotating the scan lines using coefficient c . Normal orbit drifts will introduce this image rotation; however, this correction was implemented to take advantage of the considerable overlap of MSS image data between adjacent orbits at higher latitudes (above 50°N). Use of data from adjacent orbits allows consecutive-day coverage of LACIE sites.

The definition of the rotation angle W will be presented later.

As defined in equation (18), the coefficient b may be oversimplified. Earth rotation skew would normally be accounted for by a correction of this type. Early in the development of the GSFC LACIE processing system, it was decided that an Earth rotation correction was not essential for registration of sample segments. The rationale for this decision was based on the fact that the Earth rotation correction is a constant correction for a specified location as in the case of any LACIE site. In retrospect, it can be seen that this is not entirely correct since the MSS sensor scans six lines in each mirror sweep. Earth rotation introduces translations between mirror sweeps (sets of six lines), not between lines. Since the MSS scan is asynchronous, the location of the Earth rotation between mirror sweeps does not remain fixed within the sample segments extracted for a test site and will contribute to registration errors. This correction could not be made when this effect was recognized since the reference sample segments would have had to be reprocessed and the registration of significant portions of the LACIE data base would have been impacted. A similar (although smaller) six-line mirror sweep effect is introduced by the roll correction term in equation (18) since it applies a separate correction to each line rather than to each six-line sweep.

Within the six lines of an MSS mirror sweep, there is an 0.08-pixel offset between lines due to the finite sampling interval. Although an 0.08-pixel offset is small, the accumulated effect is a 0.4-pixel offset between the first and the sixth lines of a mirror sweep or equivalently a 0.4-pixel offset between the sixth line of a mirror sweep and the first line of the next mirror sweep. This correction is also not accounted for by equation (18). These points should be evaluated and equation (18) modified in any future LACIE processing system. The scan line rotation of coefficient c is defined as

$$c = \frac{K}{M} \tan(\alpha_Y - W) \quad (19)$$

This coefficient accounts for the combined rotation of scan lines due to yaw and adjacent orbit image rotation W .

The image rotation angle W is defined as the rota-

tion from an actual orbit path to a fictitious orbit path passing through the test site center. This fictitious orbit path was devised to minimize the rotation required and to avoid the complication of referencing rotations to the orbit during which an initial sample segment was extracted. This image rotation angle, due to the orbit path not passing through sample segment (or search area) centers, is defined by

$$W = \tan^{-1}(\sin \phi_S \tan \Delta L) \quad (20)$$

where ϕ_S = geodetic latitude of test site center
 ΔL = difference in longitude between orbit path and test site center at geodetic latitude

Sign conventions used are north-positive, south-negative, east-positive, and west-negative. The longitude difference ΔL is defined by

$$\Delta L = L_S - L_n - (l_S - l_n + \delta L_r) \quad (21)$$

where L_S = longitude of test site, L_n = nadir longitude at Landsat frame center, and l_S , l_n , and δL_r are defined by algorithms

$$l_S = \cos^{-1} \left[\frac{(\cos^2 \phi'_S - \cos^2 I)^{1/2}}{\sin I \cos \phi'_S} \right] \quad (22)$$

$$l_n = \cos^{-1} \left[\frac{(\cos^2 \phi'_n - \cos^2 I)^{1/2}}{\sin I \cos \phi'_n} \right] \quad (23)$$

$$\delta L_r = \frac{T}{1440} \left[\sin^{-1} \left(\frac{\sin \phi'_S}{\sin I} \right) - \sin^{-1} \left(\frac{\sin \phi'_n}{\sin I} \right) \right] \quad (24)$$

where I = nominal orbit inclination angle = 80.886°

T = nominal orbit period = 103.267 minutes

ϕ'_n = geocentric latitude of nadir
= $\tan^{-1} [(1 - f)^2 \tan \phi_n]$

ϕ'_S = geocentric latitude of test site
= $\tan^{-1} [(1 - f)^2 \tan \phi_S]$

f = $1/298.3$

ϕ_n = geodetic nadir latitude

ϕ_S = geodetic test site latitude

Table IV is a summary of the estimated maximum registration errors in geometrically corrected sample segment data. These estimates are conservative.

TABLE IV.—Maximum Registration Errors

[Resampled data]

Parameter	Max. error, percent	
	Pixel	Line
Along-scan scale = a		
ΔAltitude < 0.02 percent	0.02	—
Mirror velocity < 0.2 percent	.2	—
Roll < 0.14°	.02	—
Root sum square	.2	—
Translation of scan lines = b		
ΔRoll < 0.00002 rad	.15	—
Image rotation < $0.4'$.58	—
Root sum square	.60	—
Along-track scale = d		
ΔVelocity < 0.02 percent	—	0.02
ΔPitch < 0.00002 rad	—	.12
Root sum square	—	.12
Rotation of scan lines = c		
Yaw < 0.14°	—	.18
Image rotation < $0.4'$	—	.52
Root sum square	—	.55
Total root sum square	< .8	—

especially with regard to the uncertainty in the image rotation angle W . The 0.5-percent round-off error inherent in the nearest neighbor process and another possible ± 0.5 -percent error due to nearest pixel correlation are also present. An accurate combination of these errors with the geometric uncertainties of table IV has not been devised. It is estimated that an overall registration accuracy of ± 1 pixel (rms) is attained through the GSFC LACIE processing system.

EDGE DETECTION

The correlation of temporally separated MSS data, such as LACIE processing requires, is complicated by changes in gray level (even contrast reversals) during the seasons of the year. A fully satisfactory method for treating these changes using the normal correlation process (i.e., gray-level correlation) has not been devised. This problem led to the GSFC selection of an edge detection process. It was assumed that even though the MSS signal levels for fields may change during the year, it would be possible to detect edges (e.g., field boundaries) at any time. A study (ref. 1) was conducted to determine an edge detection process. The edge value E for pixel e is defined as

$$E = |a - b| + |c - g| + |b - h| + |d - f| \quad (25)$$

where the letters a through i represent the MSS pixel values at the locations diagramed in figure 4.

Using equation (25), the edge value E is calculated for each pixel in the search areas of MSS bands 5 and 7. These two bands were selected for edge detection after studies (ref. 1) demonstrated that some features appear best in band 5, others in band 7, and some even appear as edges in band 5 in one season and then in band 7 during another season. This procedure led to the derivation of edge images which are a composite of MSS bands 5 and 7 as described later. After calculation of the edge value for each pixel in a search area the edge values for a band are histogrammed. A threshold edge value is then defined such that 15 percent of the pixels in a search area have edge values greater than the threshold value. All pixels having edge values greater than the threshold are called edges and those with edge values

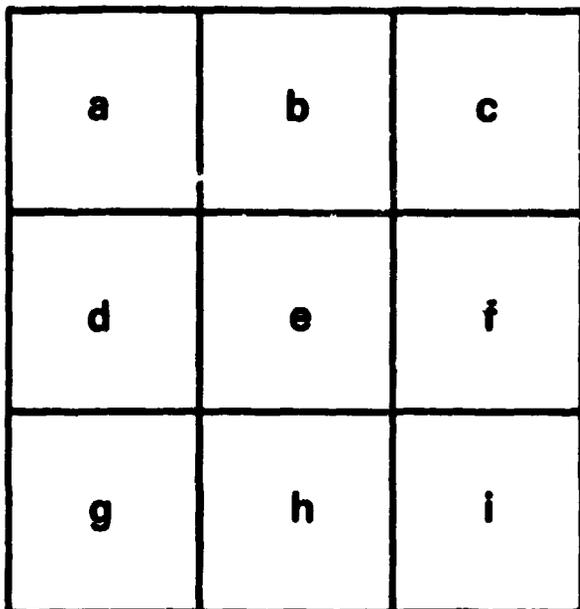


FIGURE 4.—Three by three array used for edge calculations.

below the threshold are not edges. Binary edge images for MSS bands 5 and 7 are then constructed by setting edges equal to one and nonedges equal to zero. The final composite edge images are obtained by a logical OR process; i.e., a pixel which was classified as an edge in either band 5 or band 7 is an edge in the composite edge image. The 15-percent edge density criterion was selected by testing correlation success rate for several edge density values. Using the 15-percent edge density for the individual MSS band 5 and band 7 edge images results in a composite edge image density on the order of 20 percent.

During operational use of the GSFC LACIE processing system, it was found that this edge detection technique was sensitive to small scattered clouds. Since the threshold edge process selects the strongest edges and clouds are usually brighter than any other target, the presence of small scattered clouds results in cloud edges predominantly rather than in the temporally invariant ground edges. A study (ref. 2) was conducted to determine solutions to this problem. A simple threshold test on MSS band 5 pixel values was devised. Before calculating the edge value for a pixel, each of the pixels in the 3 by 3 array (fig. 4) is compared to a defined value. If any one of the 9 pixels is greater than the threshold value, the edge value is set equal to minus 1, which eliminates it from contention as an edge. Although the cloud test is only on

band 5, a resultant cloud edge detection of minus 1 for a band 5 pixel is used to eliminate the corresponding band 7 pixel. Sufficient testing was not performed to optimize the cloud threshold value; however, the impact of cloud edges has been reduced using an operational value of 60. The investigator also attempted to establish a lower threshold level to identify cloud shadows, which would be given an edge value of minus 2. The minus-1 and minus-2 edge values assigned to clouds and cloud shadows would be useful in identifying pixels which are not valid for LACIE analysis. The cloud-shadow threshold definition is complicated by the fact that actual ground data can have values down to pixel level zero, and, thus, some true ground data would be identified as cloud shadows. The cloud and cloud-shadow identification capability has not been implemented. Additional studies utilizing Sun elevation considerations were not completed.

EDGE CORRELATION

The edge detection process is applied to every geometrically corrected search area. In the case of an initial extraction (first acquisition of a test site), the sample segment is defined as the center portion of the corrected search area and is extracted accordingly. The corresponding portion of the composite search area edge image is also extracted and stored in the edge image file for future correlations.

When a sample segment extraction is to be accomplished for a previously acquired test site, the initial edge image of a sample segment is retrieved from the edge image file for correlation with the new corrected search area. The correlation process consists of simply counting coincidence of edges (a logical AND operation) between search area and reference sample segment edge images for every possible overlay position. The overlay positions are restricted to those for which no sample segment edge pixels fall outside the search area. The location with the maximum number of coincident edges defines the center location of the sample segment to be extracted.

To reduce correlation processing time, a rapid correlation rejection technique based on partial correlation estimates was devised (ref. 3). The complete correlation is performed for only one thirty-second of the possible overlay alignments; i.e., every sixteenth column and every other row. For each of

these alinements, the normalized correlation coefficient p is calculated.

$$p = \frac{N_c}{N_E} \quad (26)$$

where N_c = number of coincident edge pixels

N_E = number of edge pixels contained within overlaid portion of search area

The mean correlation coefficient \bar{p} is calculated as well as the standard deviation σ about this mean. A rapid rejection threshold T_{RR} is then defined as

$$T_{RR} = \bar{p} + 3\sigma \quad (27)$$

For all remaining alinements, a correlation estimate is determined, using one-ninth of the edge pixels. The edge pixels in the reference edge image are formatted in 16-bit words (16 edge pixels per word), and the correlation estimate uses every third word on every third line. The correlation estimate for these edge pixels is determined using equation (26). If the value of p obtained is less than T_{RR} , the alinement is

rejected and the next alinement correlation is initiated. The complete correlation is then performed only for alinements that produce correlation estimates greater than T_{RR} . It is estimated that after the threshold T_{RR} is defined, only about 5 percent of remaining alinements require complete determination of the correlation coefficient. Naturally, the alinement producing the maximum correlation coefficient p determines the location for extracting the sample segment from the search area.

Additional studies (ref. 2) were pursued to determine the feasibility of correlating subareas of sample segment edge images in order to improve geometric corrections and registration. Preliminary results indicated such methods are feasible. Further testing and analysis would be required before actual implementation could be initiated.

REFERENCES

1. Nack, M. L.: Interim Report on Temporal Registration of Multispectral Digital Satellite Images Using Their Edge Images. Computer Sciences Corporation, NASA Contract NASS-11999, Apr. 1975.
2. Nack, M. L.: Final Report on Image Registration Research. Computer Sciences Corporation, NASA Contract NASS-11999, Nov. 1976.
3. Nack, M. L.: Temporal Registration of Multispectral Digital Satellite Images Using Their Edge Images. AAS paper 75-104, 1975.

Development of LACIE CCEA-I Weather/Wheat Yield Models

N. D. Strommen,^a C. M. Sakamoto,^b S. K. LeDuc,^b and D. E. Umberger^c

INTRODUCTION

The LACIE required an estimate of wheat yield throughout the wheat growing season. The initial test area of the experiment included the winter wheat and spring wheat regions of the U.S. Great Plains. In subsequent phases, Canada, the U.S.S.R., Argentina, Brazil, Australia, and parts of India were added to the coverage.

A basic premise behind the LACIE project was that there existed considerable technology for yield estimation which could be further developed and tested for its ability to make accurate, real-time yield estimates in a quasi-operational setting. The task of developing the initial operational yield models in LACIE was assigned to NOAA's Center for Climatic and Environmental Assessment (CCEA),¹ which was officially started in November 1974. The U.S. Great Plains yield models were scheduled to be in operation by April 1975.

MODELING APPROACHES

Crop/climate modeling can be approached in many ways. The selection of an approach depends primarily on the objective for which that model is developed and the data and resources that are available for development. For example, if the primary objective is to better reflect the physiological processes of the crop in response to its environment, the

best modeling approach may differ from that taken where the objective is to estimate grain yield over large areas.

The various approaches to modeling for grain yield may be classified as follows: (1) causal (phenological, dynamic, physiological), (2) statistical regression, and (3) analog. Each of these approaches has its advantages and limitations.

Causal Approach

Early discussions at CCEA recognized that conceptually a crop/climate model should have the capability to simulate closely plant development as a response to changing environmental elements. The estimation of model relationships using this approach has been done on research data where the crop being modeled was grown either under controlled conditions or within a homogeneous environment (refs. 1 to 3). This approach tries to model the biological effects of environment (climate, cultural practices, etc.) on crop growth and grain yield. The models attempt to specify the complex processes inherent in plant development and reproduction. Many researchers are attempting to use experimentally derived information to model these relationships. Statistical procedures are required to estimate the coefficients of some of these relationships once sufficient data become available. However, the present lack of data and the uncertain accuracy of much of the available data for wide ranges of the environment make the potential gain from such models questionable. There are two major problems associated with this approach in an operational system. First, development and adaptation of the causal crop/climate model to large areas would require acquisition of data much of which is not routinely available for many areas of the world. For example, soil moisture capacities are not available for all areas. Second, knowledge of causal relationships needed to

^aNOAA Environmental Data and Information Service, Washington, D.C.

^bNOAA Environmental Data and Information Service, Columbia, Missouri.

^cUSDA Economics, Statistics, and Cooperatives Service, Columbia, Missouri.

¹In a NOAA reorganization during 1978, CCEA was made a part of a new Center for Environmental Assessment Services.

quantify the effects of weather events on biological/physical processes and ultimately on yield was incomplete at the beginning of LACIE and in many areas remains incomplete today.

Statistical Regression Approach

The historical regression approach, as it has been applied, attempts to "shortcut" the complex biological processes and to estimate statistically some relationship between observed environmental information and grain yield. This approach usually has been limited to single equation models. Historical data series have been the most commonly used sources of information to estimate the coefficients of the model for large areas. Explanatory variables may be seasonal, monthly, daily, or for some other time frame. Considerations include environmental factors that are hypothesized to significantly affect grain yield. These models are usually limited to a few explanatory variables.

Analog Approach

In the absence of historical data for a region of interest, it is sometimes feasible to seek a like area for which data are available, to develop a model, and then to use that model as a surrogate for the area of interest. Unfortunately, exact agricultural analog areas are rare, if they exist at all. Prime candidates for this type of approach are China and portions of Alaska, where reliable records of weather and crop yield are not available. The concept of this approach is similar to developing a "universal" wheat yield model, with selected agronomic and climatic data for application to areas where such data do not exist, the distinction being that only one specified area is considered in the strict analog model.

CCEA-I MODEL DEVELOPMENT

Model Philosophy

Given LACIE's primary goal, that of estimating wheat production for the large areas of eight major wheat-growing regions, the statistical regression approach of correlating historical yield and climate data

offered the greatest potential return within the constraints of time and data resources. Thus, CCEA's decision was to develop models which utilized the best currently available data. Monthly weather data were readily available for the United States and were available with reasonable effort for other geographic areas in a form compatible with the yield data. The advantages included the availability of sufficient data to build these models in the LACIE test areas and an established flow of current meteorological data capable of supporting a real-time operational program. These data are available from the World Meteorological Organization's Global Telecommunications System (GTS). This is the system through which global weather information is transmitted and made accessible to all participating World Meteorological Organization members. A more detailed account of the GTS is given in the plenary paper entitled "The Impact of LACIE on a National Meteorological Capability" by N. Strommen, M. Reid, and J. Hill.

Model Form

Time and data constraint considerations and the problem shown by the models developed by Thompson (refs. 4 to 6) led CCEA to develop its first-generation wheat yield model, the CCEA-I, using the historical regression approach.

The basic equation for the CCEA-I model is of the form

$$\hat{Y} = \text{constant} + \hat{A}(\text{technology trend}) + \hat{g}(WX)$$

where $\hat{g}(WX)$ includes variables measuring the impact on yield of moisture stress and heat stress based on monthly temperature and precipitation data. The term $\hat{A}(\text{technology trend})$ can include variables indicating the impact on yield from changes in hybrid varieties, fertilization rates, herbicides and pesticides, and other management or cultural practices. In the absence of long-term quantitative data on these technology variables, time (year) is used as a surrogate for technology trend. The basic concept is to have the trend term describe the sustained yield achieved in the region of interest and let the weather term reflect the variation of yearly yield around the trend term. The variation of yield on a global scale due to weather is believed to be close to 10 percent. The variation of yield is often larger than 10 percent in many areas, particularly drier areas. Yield varia-

tion may be less than 10 percent in areas where moisture is less restrictive.

High multicollinearity among the predictor variables causes statistical problems in determining the most appropriate variables and also in estimating their impact. Two different weather variables—temperature and precipitation—are highly correlated, and variables for successive months are also correlated because of persistence. Since complete independence is not practically attainable, one must determine a priori what variables have important causal relationships.

The use of monthly meteorological data also assumes implicitly that particular phenological stages occur at the same time each year. If occurrence of a particular stage varies between two calendar months, the model may not be able to estimate the impact of meteorological conditions during one or both of those months. The necessary assumption of an average crop calendar is one of the limitations of the use of monthly data.

Truncation

One of LACIE's requirements was within-season estimates; i.e., an estimate before harvest. This requirement led to the concept of truncated models for making within-season estimates. Separate regression equations or truncated models were developed to capture information on yields contained in weather data available before harvest. If regression analysis indicated that weather during periods prior to and through a month contributed to the final yield estimate, a model was truncated for that month. This estimate was made without any explicit assumptions about weather during the rest of the crop season. The implicit assumption is either that weather in later months will be similar to that which occurred in the period used to develop the model or that much of the final yield is determined by conditions prior to the time of the truncation. Several alternatives to this approach are possible, but the intent in CCEA-I was to determine how much information on final yield might be available from early-season weather and no predictive knowledge of future weather.

Trend Variable

Bond and Umberger (ref. 7) discuss seven factors that affect changes in the technological trend of

wheat yield. These are irrigation, varietal changes, fertilizer application rates, pesticide usage, cultural practices, soil productivity base, and government programs (e.g., the land bank program). A complete series of historical data for all these nonweather factors does not exist for areas compatible with areas for which meteorological data exist. Consequently, the effects of these nonweather factors cannot be statistically estimated with sufficient confidence. With little or no data on these factors, it is necessary to assume a certain yield based on known practices in a given area and then subjectively specify the trend.

How a trend term and possible changes in direction can be specified is illustrated by analyzing the historical yield data series for North Dakota (fig. 1). The initial CCEA models were derived from observations for the period 1931-74 and were updated each year for Phases II and III. For illustration purposes, however, the period 1879-1976 is plotted in figure 1. The yield series shows that yield trended downward from 1879 until the drought period of the 1930's. This downward trend is partially attributed to two factors: (1) soil fertility deterioration with time and (2) expansion of wheat acreage to the less humid western areas of the state.

After World War II, an upward trend in yield occurred. The period after World War II was a period of gradual increase in fertilizer application, while during the 1950's the major impact was the introduction of new varieties. Other factors that have led to changes in the North Dakota trend include increased use of summer fallowing, changes in the location of wheat growing activities, improved weed control, and governmental actions, such as beginning or ending the land bank program. For North Dakota, at least, the yield trend appears to have leveled in 1972. The reasons for this 1972 break can be traced to increases

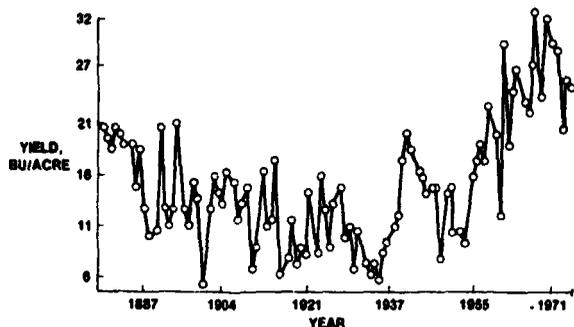


FIGURE 1.—Plot of average state wheat yields versus year for North Dakota spring wheat, 1879-1976.

in acreage planted, which increased the use of land with lower wheat production potential, reduced the percent of wheat grown on fallow, and reduced the fertilizer application rates.

The CCEA-I models were not capable of jointly estimating the weather effects, the technology effects, and the possible weather-technology interactions. An example of this interaction is shown for Oklahoma (fig. 2), where the September through December precipitation, a significant variable in that area, clearly shows a dry period in the 1950's and a more favorable moisture regime in the 1960's. The abrupt shift in the yield series in the mid-1950's, therefore, may not be associated solely with changing technology.

Changes in technology can contribute to rapid changes in yield. In Uttar Pradesh, India, for example, yields have increased substantially since 1965, although monsoonal precipitation has tended to decrease from the early 1960's through the mid-1970's (fig. 3). The combination of high-yielding varieties and the corresponding increase in irrigation and fertilization on acreage planted to high-yielding varieties is responsible for much of the increase in yield. But would the yield response have been higher still had the precipitation trend been on the upswing rather than on the downswing? There is no question that quantitative relationships between trend changes and known agronomic and climatic variability are urgently needed.

Selection of Weather Variables

In choosing the type of weather variables to include in the model for a given area, important factors

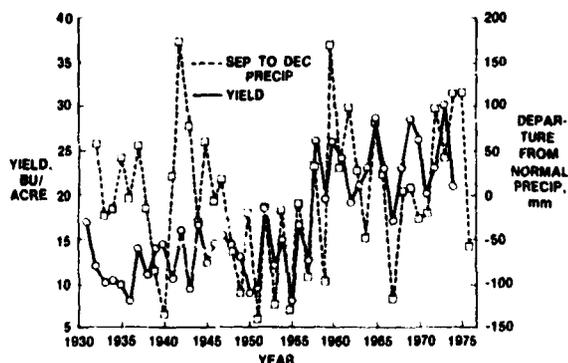


FIGURE 2.—Plots of average Oklahoma wheat yield and early-season precipitation, 1931-76.

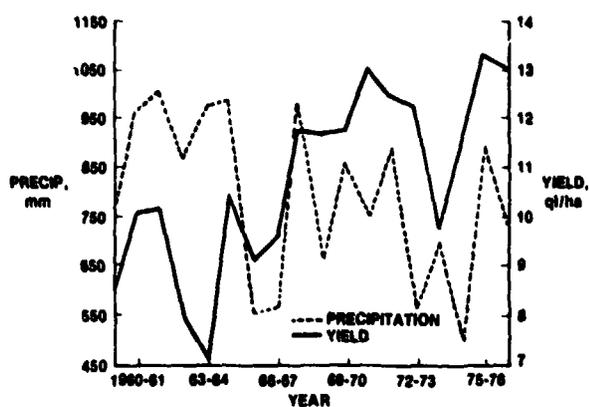


FIGURE 3.—Plots of average wheat yield and crop season rainfall for the State of Uttar Pradesh, India, 1959-77.

that must be considered are (1) data availability, (2) the phenological stage of wheat, and (3) the long-term climate for a given modeled area. For example, climatically, frequent May storms in the form of heavy thundershowers and gusty winds can be detrimental to yield in Oklahoma. Basically, the weather variables of a model estimate the effects of moisture stress and temperature stress on yield during various assumed average growth stages. As a first approximation, precipitation and temperature values were utilized as surrogates for moisture and temperature stress.

A soil moisture variable that included a water balance might have been a better indicator for soil moisture stress. However, no reliable soil moisture estimating procedures using monthly data are known to exist.

Variables were initially selected from a priori indications that they should be important. The final choice of which variables to include in the model depended on two criteria: (1) most importantly, the estimated effect of the variable on yield at the assumed phenological stage had to agree with the agronomic expectations; i.e., the sign of the coefficient had to be correct; and (2) each variable's coefficient had to be statistically significant from zero, but the level of significance was not a set value and at times was higher than the usual 5-percent significance level. The omission of a meaningful predictor was undesirable; thus, variables were accepted that might otherwise have been rejected. In addition to the application of these two criteria, the potential candidate variables were plotted with the detrended yield to permit a visual evaluation of their importance.

In addition to the departure from average of monthly precipitation amounts, two other candidate variables were tested as possible moisture stress indicators in most models. These were (1) the difference between precipitation and potential evapotranspiration (PET) and (2) the ratio of evapotranspiration (ET) to PET. Potential evapotranspiration was estimated by Thornthwaite's procedure (ref. 8). Conceptually, these variables are moisture supply-and-demand indicators, so that when demand (PET) exceeds supply (precipitation plus soil moisture reserves), the crop is under a measure of stress. The ratio of ET/PET has been used as an indicator of stress, although the form $1 - ET/PET$ has been preferred (refs. 9 and 10).

In the models for Australia and Argentina, a monthly water budget procedure (ref. 11) was included to consider, in part, runoff. A soil moisture index called the Z-index is in effect another moisture stress variable that uses monthly climatic data and a knowledge of the local soil water-holding capacity. Sakamoto (ref. 12) used the Z-index in the semiarid zones of Australia and found it to be a reliable indicator of moisture stress and hence a predictor of wheat yield for South Australia. The Z-index was not included in the U.S. Great Plains models, which were developed before the Australian models. The model for Australia with the Z-index responded well to the wide fluctuations in the data. A completely independent 10-year (1963-72) test was also run for the model developed from the data set of 1940-62. The results indicate that the variables selected for the Australian wheat yield models were stable for the two data sets.

In some cases, it was also prudent to include precipitation outside the growing period. The total precipitation prior to planting for spring wheat or the fall-plus-winter precipitation for winter wheat was often used as an indicator of the soil moisture reserve. For example, in the Canadian models, the monthly precipitation for the period 20 months before planting was included in order to consider the beneficial effects of summer fallowing. Unfortunately, an excessive amount of precipitation within a short period of time often leads to runoff, depending on soil type, slope, and preexisting soil moisture levels. Runoff was not considered in the initial models.

To cope with this problem and also the problem of evaluating the predictive capability of the model for climatic events outside the range within which it was developed, a censoring procedure was instituted.

Values of precipitation and/or temperature outside the limits of the data base were adjusted to bring them within the data from which the models were developed. In the case of CCEA-1, monthly precipitation values used in the model for the prediction year could be no greater than the 90th percentile of the historic data, and temperature values were censored to between the 5th and the 95th percentile. This means that if the precipitation for the prediction year exceeded the 90th percentile, the input value reverted to the 90th percentile.

The temperature stress indicator variable was the departure from normal of the mean monthly temperature. Although it would have been desirable to include mean maximum or mean minimum temperatures as candidate variables in all the models, lack of data and resources did not permit this to be done in the allocated time. However, for some cases in Canada, mean minimum and mean maximum temperatures were included as candidate temperature variables, allowing consideration of the potential effect of temperature on the length of the growing season and the reduction in yield that may be produced by untimely freeze.

Excessively high temperature during the flowering to heading stages is detrimental to wheat yield. In the U.S. Great Plains models, a dichotomous variate "degree days above 90°" was developed to consider this effect. The number of degrees that the maximum temperature is above 90° is the daily degree days. This value was accumulated for a month and plotted with detrended yield to estimate the relationship between the two. In all cases, the correlation was negative. From the plots, it was observed that if the number of degree days exceeded a critical value (which differed for different regions), yield was reduced; but, if the threshold value was not exceeded, no decrease occurred.

In the U.S.S.R. winter wheat area, winterkill is a major yield-reducing factor in some years. This seems to be more of a problem in the U.S.S.R. than in the U.S. Great Plains. It was found that in many areas of the U.S.S.R., the December-January or January-February mean temperature was often a significant indicator of winterkill effect.

Episodic Events

An episodic event is defined as an occasion in which yield is affected by a relatively rare occurrence, natural or social. Examples of natural events

include frost, hail, rust outbreak, flood, cattle trampling the crop, etc. (ref. 13). Social events might include revolutions or significant changes in national agricultural policy which cause widespread changes in farming practices. Generally, episodic events are not modeled by the selected set of independent variables. When a particular episodic event was known to greatly depress yield, the datum for that year could be eliminated from the analysis, and in some cases this was done. No known quantitative estimation technique is currently available to handle episodic events. However, an objective adjustment of yield based on historically averaged damage or a very gross adjustment as was done for Argentina (ref. 14) is sometimes possible.

DATA

Strata Selection

For the United States, historic yield estimates from the U.S. Department of Agriculture (USDA) are available at county, crop reporting district (CRD), state, and national levels. These estimates are made using statistical sampling techniques. The coefficient of variation of the wheat estimates is about 2 percent at the national level and from 3 to 8 percent at the state level depending on the statistical sampling in use (ref. 15). In most states, the historic yield estimates at the state and CRD levels are geographically compatible with meteorological data; i.e., meteorological data are summarized by state and crop district. The compatibility of the historic yield and climatic data at these levels minimizes the data handling problems for the United States. Common boundaries exist for historic yield and meteorological data in other regions of the globe, but these data for most other global regions are limited to larger areas. Therefore, state models were developed in the United States, although in some areas smaller strata were selected (fig. 4) because of climatic differences. In other countries, the model included that area for which the wheat yields were reported.

Operational Data Flow

Real-time meteorological data flow to support the operational program of LACIE requires summarization of daily maximum and minimum temperatures

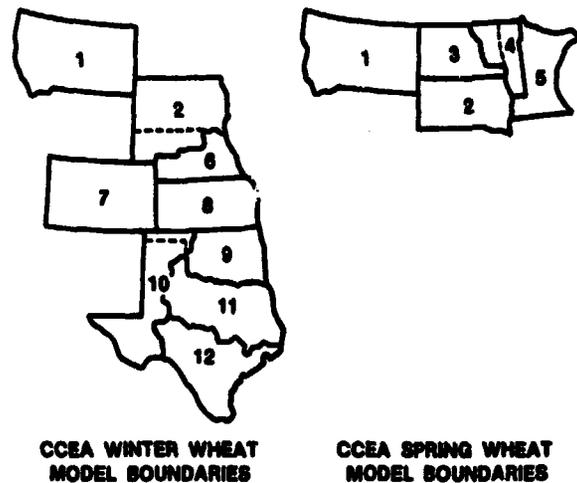


FIGURE 4.—Boundaries of regions for which U.S. wheat yield models were developed.

and precipitation. This information is available from the synoptic scale network supplemented by the monthly climatic data. These are limited data, and individual stations must often cover large areas. Extension of the limited ground data could be accomplished subjectively using meteorological satellite data. The meteorological satellite data are, however, considered most effective in identifying areas of no precipitation, noted by absence of clouds. For example, the meteorological satellite confirmed the extent of the drought in the U.S.S.R. wheat regions in 1975.

Weighting

In the estimation of yield for a selected area such as a state, the weather data for each CRD are weighted to obtain a state-level value. The weight for each CRD is given by that CRD's percentage of the total harvested area in the state, using historical acreage. No weighting of area climatic data was required for foreign areas.

MODEL RESULTS

The results of a 13-year (1965-77) simulated operational run, based on the "bootstrap" test, are shown in figure 5. Comparisons are made for five sample states. The bootstrap test is a simulation where the years prior to the prediction year are used

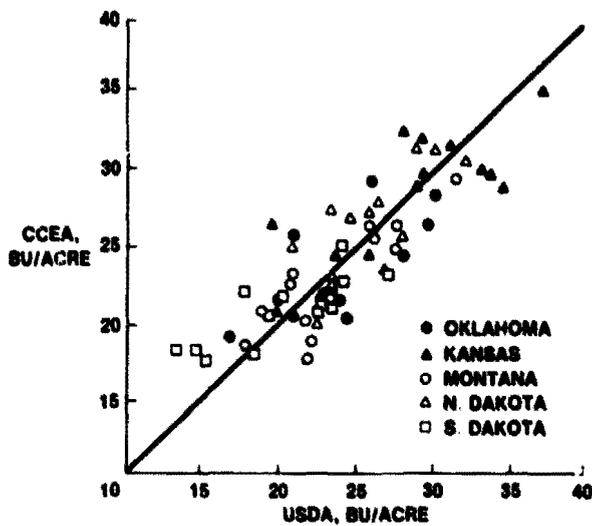


FIGURE 5.—USDA wheat yield estimates versus CCEA model estimates for selected states, 1965-77.

in developing the coefficients for the equations. Each subsequent year adds a year for estimating the coefficient. For example, in 1977, the years 1931 through 1976 were used in recalculating the model coefficients whose variables remained unchanged. Production estimates using the USDA acreages are compared in figure 6. In most test years, CCEA model estimates compare favorably with USDA estimates. In 1971 and 1974, unusual departures of the observed crop calendar from the historical average led to the large yield discrepancies in the estimates. The experience in 1977 suggests a need to reevaluate the trend component of the CCEA-I model.

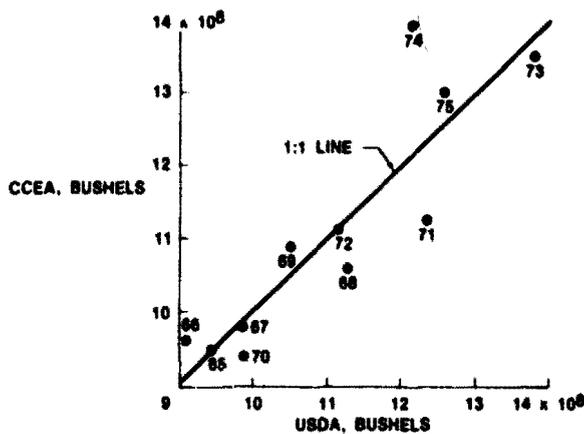


FIGURE 6.—USDA wheat production estimates versus CCEA estimates (using reported USDA acreage), 1965-75.

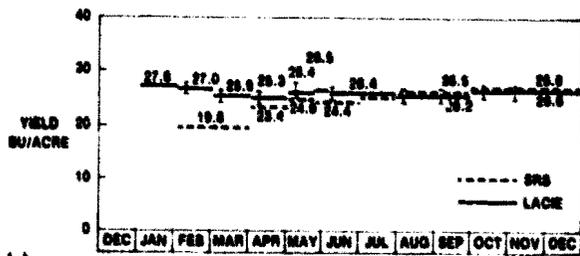
In the development of the model, historical CRD precipitation and temperature data are used. These data come from both the synoptic network (principal weather stations) and the cooperative climatological network. In real-time operations, often data from only the principal (first-order) stations are available. To test if these data (precipitation in particular) can be improved using a denser network of weather stations, data from which are available in a delayed time frame for the United States, these models were rerun and the results compared with the results using only the synoptic network (ref. 16). The differences between the two estimates were very small. Thus, timeliness and cost considerations may indicate that the operational data base should be limited to the synoptic scale.

The CCEA-I models also revealed a potential for making early-season estimations. Figures 7(a), 7(b), and 7(c) show an example of this tracking for the U.S. Great Plains and the U.S.S.R. Although much can happen to alter yields for localized areas, the results of the aggregated yields over large areas indicate that reasonably accurate information is possible in certain areas with early-season truncated models. The performance of these models over the past three (1975-1976-1977) crop seasons is evidence that they can effectively provide useful real-time forecasts of yield for the principal wheat-growing regions of the world.

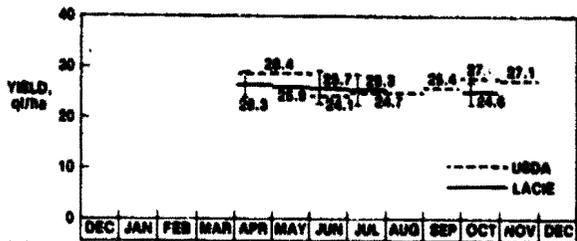
FUTURE CONSIDERATIONS

New efforts are currently being undertaken by CCEA to improve or complement this initial performance. Using the logic shown in figure 8, the CCEA-I models are being reviewed with the intent to include variables that may better explain the year-to-year variability. This review is the first major attempt to reanalyze these initial CCEA-I models.

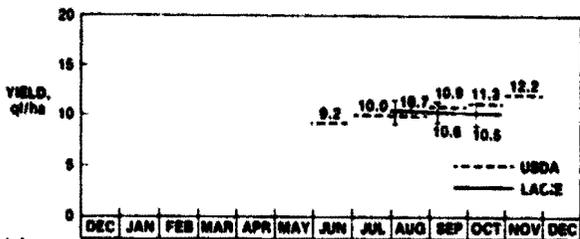
The reader will recall that the CCEA-I truncated yield models were developed and operated on the implicit assumption that the weather after the time of the truncation had only a limited effect on the yield. LeDuc, in an unpublished report (ref. 17), investigated two alternative approaches. She used the CCEA-I truncated yield model for North Dakota spring wheat and made estimates in two different ways: (1) by assuming normal weather after a truncation and (2) by using the current weather reported to the month of truncation, then inputting historical data (1932-74) for the remainder of the season to obtain a distribution of possible yields. The result of (2)



(a)

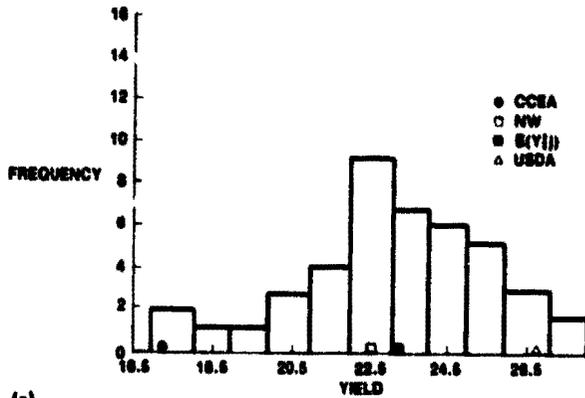


(b)

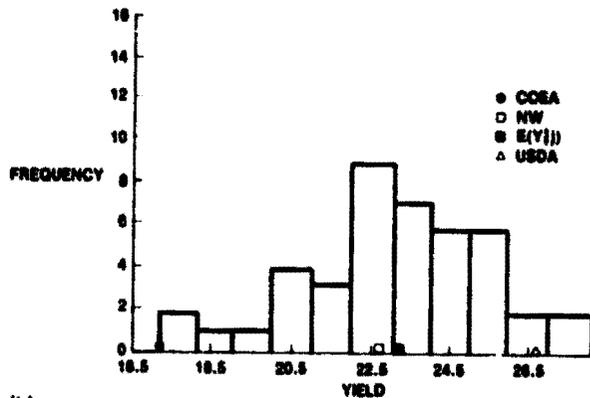


(c)

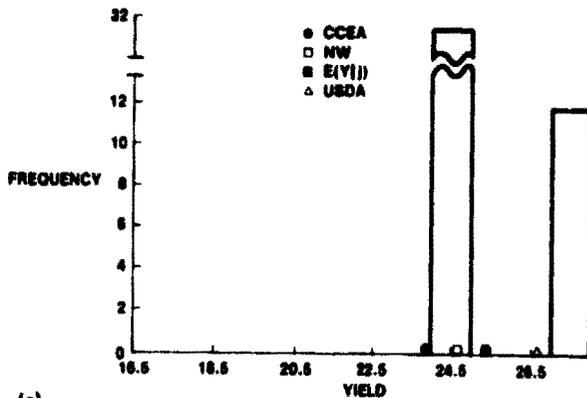
FIGURE 7.—Comparison of LACIE and USDA monthly wheat yield estimates in U.S. and U.S.S.R. for 1976. (a) U.S. Great Plains winter wheat. (b) U.S.S.R. winter wheat. (c) U.S.S.R. spring wheat.



(a)



(b)



(c)

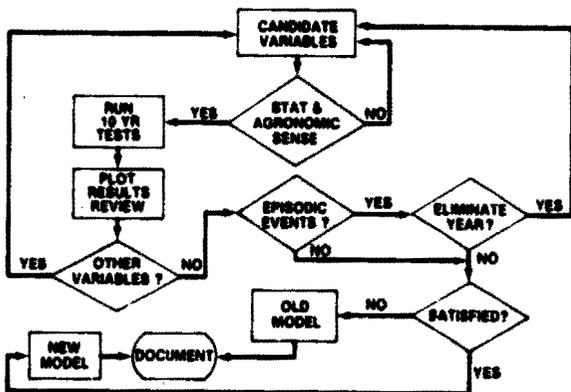


FIGURE 8.—Logic flow for review and revision of CCEA-1 wheat yield models.

FIGURE 9.—Comparison of three methods for early-season estimation of wheat yields. (a) April. (b) May. (c) June.

is the set of histograms shown in figures 9(a), 9(b), and 9(c). Also indicated is the single estimate from the CCEA-I truncated model as used in LACIE and the estimate resulting from (1). From these figures, it is evident that the assumption of normal weather (NW) for the remaining crop season or the mean of the histogram $E(Y_j)$ may provide an estimate closer to the final observed yield (USDA). This approach needs additional investigation for operational use.

LeDuc (ref. 18) has also reported on a statistical phenological spring wheat model for North Dakota, which uses the crop reporting district as a basic unit and considers the crop calendar as well as a soil moisture budget and heat stress term. Steyaert et al. (ref. 19) have also developed a procedure to use atmospheric pressure directly in large-area modeling. Eventually, other approaches may lead to improved models for areas where historical data are not available for model development and to improved accuracy over the time-series regression approach. However, until improved long-term weather forecasts are available, it is unlikely that these models can significantly improve estimates of harvested wheat yields in many of the large wheat-growing regions.

SUMMARY AND CONCLUSIONS

The statistical regression approach to crop modeling, under many conditions, represents an effective way to achieve reasonably accurate estimates of wheat yields for large areas in several important wheat-producing regions of the world. The operational performance during the last three growing seasons has demonstrated that, with the current state of the art, the historical regression approach is a feasible method to convert the flow of meteorological data available into useful wheat yield information.

The yield estimates have been provided in a timely manner at a low cost. Such estimates supplied regularly in an operational system would help provide needed information to government planners, agribusiness decisionmakers, and farmers. The experience of LACIE has provided better insight into the problem areas that need further work.

REFERENCES

1. Runge, E. C. A.: Effects of Rainfall and Temperature Interactions During the Growing Season on Corn Yield. *Agronomy J.*, vol. 60, 1968, pp. 503-507.
2. DeWit, C. T.; et al.: A Dynamic Model of Plant and Crop Growth. In *Potential Crop Production: A Case Study*, P. F. Wareing and J. P. Cooper, eds., Heinemann Educational Books, 1971.
3. Haun, J. R.: Determination of Wheat Growth Environmental Relationships. *Agronomy J.*, vol. 65, 1973, pp. 813-816.
4. Thompson, L. M.: Weather and Technology in the Production of Corn in the United States Corn Belt. *Agronomy J.*, vol. 61, 1969, pp. 453-456.
5. Thompson, L. M.: Weather and Technology in the Production of Wheat in the United States. *J. Soil Water Conservation*, vol. 24, 1969, pp. 219-224.
6. Thompson, L. M.: Weather and Technology in the Production of Soybeans in the Central United States. *Agronomy J.*, vol. 62, 1970, pp. 232-236.
7. Bond, J.; and Umberger, D.: Technical and Economic Causes of Changes in U.S. Wheat Production, 1949-1976. USDA Foreign Agricultural Service, 1978 (to be published).
8. Thornthwaite, C. W.: An Approach Toward a Rational Classification of Climate. *Geog. Rev.*, vol. 38, 1948, pp. 55-94.
9. Shaw, R. H.: Use of Moisture Stress Index for Examining Climate Trends and Corn Yields in Iowa. *Iowa State J. Res.*, vol. 51, no. 3, 1977, pp. 249-254.
10. EarthSat Spring Wheat Yield System Test 1975. First Report. Earth Satellite Corporation, Contract Number NAS 9-14655, NASA Johnson Space Center, 1975.
11. Palmer, W. C.: Meteorological Drought Research Paper No. 45. U.S. Department of Commerce, National Weather Bureau, 1965.
12. Sakamoto, C. M.: An Index for Estimating Wheat Yield in Australia. LACIE-00502, NASA Johnson Space Center, Houston, Texas, 1977.
13. Yield Advisory Group Report LACIE-00466, NASA Johnson Space Center, Houston, Texas, 1978.
14. Sakamoto, C. M.: The Z-Index and Subitaneous Events as Variables for Estimating Wheat Yield in Argentina. LACIE-000503, NASA Johnson Space Center, Houston, Texas, 1977.

15. **Preparing Crop and Livestock Estimates.** USDA Statistical Reporting Service, March 1974.
16. **Sakamoto, C. M.; Strommen, N. D.; and LeDuc, S. K.: Synoptic Vs. Cooperative Climate Data As Inputs for Wheat Yield Model Estimates.** Paper presented at the 13th American Meteorological Society Conference on Agriculture and Forest Meteorology, Purdue University, West Lafayette, Indiana, 1977.
17. **LeDuc, S. K.: Truncated Model Versus Yield Expected (From Climatology) of Final Model Using Weather to Date.** Unpublished CCEA Report, NOAA, Columbia, Missouri, 1977.
18. **LeDuc, S. K.: CCEA Second Generation Model.** Paper presented to the Crop Modeling Workshop, Columbia, Missouri, 1977.
19. **Steyaert, L. T.; LeDuc, S. K.; and McQuigg, J. D.: Atmospheric Pressure and Wheat Yield Modeling.** Agr. Meteorol., vol. 19, 1978, pp. 23-24.

Growth Stage Estimation

V. S. Whitehead,^a D. E. Phinney,^b and W. E. Crea^c

CROP CALENDAR MODELING APPROACH

Identification of wheat by the analyst-interpreters requires that they integrate all the knowledge available to them concerning the appearance of wheat and the farming practices and natural events that can change that appearance.² One of the tools employed is a crop calendar that describes the progression of the crop through detectable and/or agronomically significant events in its life cycle (i.e., planting date, date of emergence, date of heading, etc.). This, of course, can change from place to place and from year to year; this calendar is also a function of variety. Localized mean crop calendars for wheat and some confusion crops were derived from the data available at the beginning of LACIE for U.S. areas. The local year-to-year changes in this crop calendar due to differences in weather and the normal crop calendars for foreign areas were not so well understood, however.

Early in the preparation for LACIE, it became apparent that these year-to-year variations in the seasons made the use of localized normal crop calendars to aid in distinguishing wheat from other crops a questionable procedure. Further, it was recognized that because wheat yields could be drastically affected by unusual events at critical times in its development (e.g., hot temperatures at heading), yield models to be developed would probably require a good estimation of the development stage of the crop throughout the crop year for the year of interest.

A literature search was performed for candidate approaches to adjustment of the crop calendar to account for year-to-year weather differences. Three candidates were identified: the heat unit, a function of temperature alone; the photothermal unit, a function of temperature and day length; and the Robertson triquadratic unit, a nonlinear function of

maximum and minimum temperature and day length (ref. 1). After comparative testing on an independent set of data, the Robertson model was chosen as best describing the rate of phenological development of wheat because (1) there existed empirical and theoretical evidence of nonlinear responses to temperature and day length; (2) the number of phases and related interval lengths in the corresponding scale appeared meaningful and convenient; and (3) application of the Robertson model to both winter and spring wheat had met with preliminary success (accuracies 28 to 14 percent better than the heat unit and photothermal unit models between emergence and heading).

Data required to operate this model are initiation (planting) date, duration of daylight (date and latitude dependent) and daily maximum and minimum air temperatures. Planting date can be taken as normal or modeled; date and latitude are known; and maximum and minimum temperatures can be taken from reported maximum and minimum values or high and low hourly values, or estimated by use of 3- or 6-hourly synoptic reports. Since no equivalent model for winter wheat was available, a contract was let to Kansas State University to modify the spring wheat model so that it would track the development of winter wheat and account for the dormancy characteristics of that crop.

The following error sources and constraints in the initial models were recognized.

1. Coefficients were derived for spring wheat varieties used in Canada. The applicability of these coefficients to U.S. and U.S.S.R. spring wheat varieties and particularly to winter wheat and dwarf wheat was questionable.

2. Use of normal planting dates could lead to significant errors, particularly in early season (before and immediately after dormancy for winter wheat). Further, even these dates were questionable in some foreign areas.

3. The period of vegetative growth before vernalization in winter wheat and the handling of dormancy posed definite problems. The initial model

^aNASA Johnson Space Center, Houston, Texas

^bLockheed Electronics Company, Houston, Texas

^cFormerly with NASA Johnson Space Center, currently with Crea Farms, Fern, Idaho.

did not account for vernalization, and it also showed sporadic development to continue during warm periods in midwinter.

4. Reports of daily maximum and minimum temperatures were not generally available from foreign areas, nor were hourly reports generally available for these areas. Techniques to estimate the daily temperature extremes from 3- and 6-hourly synoptic reports existed, but they were not available initially.

It was believed, however, that even these crude early models would be preferable to the use of the normal values and that the use of these crop calendar models should lead to their refinement. This did prove to be the case.

MODEL FORMULATION

Initial Model Form

The adjustable crop calendar (ACC) developed by Robertson (ref. 2) describes the progress of spring wheat toward maturity as a function of daily maximum and minimum temperatures and day length. The adjustable crop calendar, as implemented for LACIE (ref. 3), is used to calculate the daily increment of development (DID) through six physiological stages of growth. These stages are tabulated as follows.

<i>Development stage for 50 percent of crop</i>	<i>ACC stage</i>
Planting	1.0
Emergence	2.0
Jointing	3.0
Heading	4.0
Soft dough	5.0
Ripe	6.0

A triquadratic equation is used to calculate the DID within each stage. The DID's are accumulated from stage to stage.

The rate equation for each stage may be written as

$$DID = \left[a_1 (DL - a_0) + a_2 (DL - a_0)^2 \right] \left[b_1 (T_X - b_0) + b_2 (T_X - b_0)^2 + c_1 (T_N - b_0) + c_2 (T_N - b_0)^2 \right] \quad (1)$$

This may be written more simply as

$$DID = (V_1)(V_2 + V_3) \quad (2)$$

Each of the increment development terms V_1 , V_2 , and V_3 is examined to see if it is negative; if negative, the value of the term is set to zero.

Since wheat responds differently to the environment during each physiological stage of growth, five separate equations are required. The individual regression coefficients are given in table I. For the techniques used to arrive at the values of these coefficients, the reader is referred to Robertson's original work.

Dormancy Modeling

The Robertson crop calendar was developed for Marquis spring wheat grown in Canada. Systematic bias due to varietal differences in maturation rate and to large variation in the length of dormancy was observed when the initial model was applied to winter wheat.

To apply the ACC to winter wheat, Feyerherm (ref. 4) developed modifications to reflect the effect of dormancy on winter wheat. Each DID from the emergence to the heading stage is multiplied by a factor calculated from the following equation.

$$M = 0.5684 + 0.025081(ADTJ) - 0.006139(AAPR) \quad (3)$$

The use of normal average daily temperature for January and normal annual precipitation was based on an observed systematic bias in the original crop calendar from cold/wet to hot/dry conditions. The $ADTJ$ term was found to be related to the length and severity of the winter dormancy period. The $AAPR$ term was used to compensate for increased development rate under conditions of increasing moisture stress.

This multiplier was derived for winter wheat varieties typically planted in the U.S. Great Plains during the early 1970's. For foreign areas where suffi-

TABLE I.—Characteristic Coefficients Developed by Robertson for the Spring Wheat Crop Calendar

Coefficient	Development stage of crop (a)				
	P-E	E-J	J-H	H-S	S-R
a_0	$V_1 = 1$	8.413	10.93	10.94	24.38
a_1	$V_1 = 1$	1.005	.9256	1.389	-1.140
a_2	$V_1 = 1$	0	-.06025	-.08191	0
b_0	44.37	23.64	42.65	42.18	37.67
b_1	.01086	-.003512	.0002958	.0002458	.00006733
b_2	-.0002230	.00005026	0	0	0
c_1	.009732	.0003666	.0005943	.00003109	.0003442
c_2	-.0002267	-.000004282	0	0	0

*P-E = planting to emergence; E-J = emergence to jointing; J-H = jointing to heading; H-S = heading to soft dough; and S-R = soft dough to ripe.

cient information is available on varietal maturities to relate them to U.S. varietal maturity classes, additional adjustment factors were derived (ref. 4).

$$M(\text{late}) = 0.7243 + (0.009623) ADTJ - (0.003536) A APR \quad (7)$$

$$M(\text{early}) = 0.7037 + (0.023445) ADTJ - (0.006735) A APR \quad (4)$$

$$M(\text{mid-early}) = 0.7613 + (0.018766) ADTJ - (0.007251) A APR \quad (5)$$

$$M(\text{mid-late}) = 0.7905 + (0.012568) ADTJ - (0.005733) A APR \quad (6)$$

The equations may be used for varieties similar to those shown in table II.

Occasionally, seasons arose in which the model showed jointing occurring before dormancy. Since this is physiologically impossible for winter wheat, some adjustment was required. Feyerherm suggested that if the accumulated DID's exceeded 2.85 (3.0 is jointing) on any day before January 1, the accumulated value be reset to 2.80.

TABLE II.—Winter Wheat Varieties Used to Define Maturity Classes

Maturity	ADTJ < 20° F		ADTJ > 20° F	
	Hard wheats		Hard wheats	Soft wheats
Early	Lancer, Warrior, Hume		Triumph class	Monon, Benhur, Knox
Mid-early	Nebred, Winoka, Winalta		Scout class	Arthur
Mid-late	Minier		Comanche, Pawnee	Dual, Fairfield
Late	Kharkof, Yogo, Cheyenne		Turkey, Kharkof	Trumbull, Redcoat

Spring Restart Model

As an alternative to dormancy modeling, the possibility of simply restarting the crop calendar after dormancy was examined (ref. 5). The normal end of dormancy for the U.S. winter wheat regions was determined from climatological data by plotting the mean monthly minimum temperature against midmonth day length at each station. This effort results in climagraphs like the one shown in figure 1. There are two lines in this figure, one that intersects the fall portion of the climagraph and one that intersects the spring portion. These lines represent the beginning and end of dormancy as defined by the following criteria.

1. When the sum of the development rates for the last 15 days becomes less than 0.02 of a unit, that day is said to be the beginning of dormancy.

2. When the sum of the development rates for the last 15 days becomes greater than 0.10 of a unit, that day is said to be the end of dormancy.

The Robertson model was run with the emergence-to-jointing coefficients from actual historical emergence dates to obtain the geographic distribution of beginning- and end-of-dormancy dates.

Climagraphs were prepared for all the synoptic weather stations in the winter wheat areas in the United States, the U.S.S.R., and China. These were used in determining the climatic analogs and to transfer the dormancy criteria to foreign areas in the early Phase II crop calendar adjustments. Later in Phase II, this approach was replaced by the use of mean planting dates with the Feyerherm multipliers, as that method was simpler to use and provided similar accuracies.

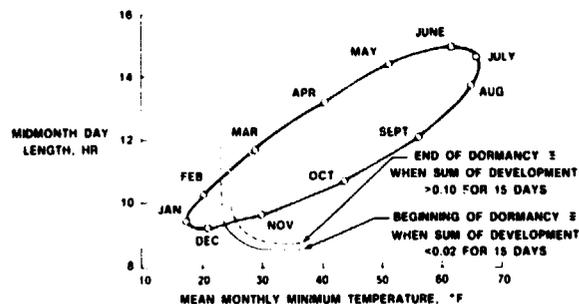


FIGURE 1.—Day length versus minimum temperature climagraph for Dodge City, Kansas (37°45' N 99°58' W, 2594-foot elevation).

Spring Wheat Starter Model

In order to use crop calendar models, knowledge of the planting dates is required. Feyerherm (refs. 4 and 6) considered the effects of temperature and precipitation on accumulated warming/planting (WP) days. The general form of the model was as follows.

$$\begin{aligned} WP &= 0 & TA &\leq 32 \\ &= \alpha (TA - 32) (PRE) & 32 < TA < 32 + 1/\alpha \\ &= 1 & TA &\geq 32 + 1/\alpha \end{aligned} \quad (8)$$

His study found that for spring wheat, $\alpha = 0.1$.

Tests of this spring wheat planting model indicated no statistically significant precipitation effect, and $PRE = 1$ was ultimately used for operations.

The date for 50-percent planting of spring wheat is estimated from a degree-day-type summation beginning on January 19. When the number of accumulated warming/planting days reaches 35.5, it is assumed that 50 percent of the crop has been planted.

Stuff and Phinney (ref. 7) developed an equation for the daily rate of spring wheat planting based on temperature, precipitation, and the normal planting date.

$$R = -0.77 + 0.045(T) - 0.032(P) + 0.053(N) \quad (9)$$

Tests on independent data indicated that this model did as well as but no better than the Feyerherm model already in use, and it was dropped without implementation.

Winter Wheat Starter Model

Feyerherm (ref. 6) confirmed earlier studies conducted at the NASA Johnson Space Center which found no agrometeorological variables that showed improvement over the use of the normal fall planting

date for winter wheat. Consequently, local normal planting dates for winter wheat have been used as the starting date for ACC model operation. It appears that, with regard to weather in the U.S. Great Plains, the farmer has a great deal of leeway in choosing the optimum date for fall seeding and the planting date is driven primarily by other factors.

Inclusion of Moisture Variable

The effect of moisture variation on crop development was studied by Seeley et al. (See paper entitled "Prediction of Wheat Phenological Development—A State-of-the-Art Review.") The moisture was treated indirectly through the use of rain-days. The model is a triquadratic form similar to the Robertson model for spring wheat except that the day-length variable has been replaced by a moisture variable. This new variable is based on the mean frequency of rain-days, which is computed on a daily basis by means of a low-pass-filter function.

$$\overline{RD}_i = \overline{RD}_{i-1} + K(RD_i - \overline{RD}_{i-1}) \quad (10)$$

- where \overline{RD}_i = the weighted running mean value of RD at day i
 \overline{RD}_{i-1} = the weighted running mean value of RD at day $i-1$
 K = an arbitrary constant
 RD_i = 1 for a day with measurable precipitation
 RD_i = 0 for a day with no measurable precipitation

A value of 0.1 was selected for K and used throughout development of this model. This allows RD_i to account for approximately 95 percent of the variation of RD in the past 30 days.

The form of the new crop calendar model is as follows.

$$DID = \left[a_1 \left(\overline{RD} - a_0 \right) + a_2 \left(\overline{RD} - a_0 \right)^2 \right] \left[b_1 \left(T_X - b_0 \right) + b_2 \left(T_X - b_0 \right)^2 + c_1 \left(T_X - b_0 \right) + c_2 \left(T_X - b_0 \right)^2 \right] \quad (11)$$

where \overline{RD} = the daily weighted running mean value of the filter function of the rain-day occurrence

Results of this approach may be found in the paper by Seeley et al. This improvement occurred too late for inclusion in the Phase III adjustable crop calendar operational program.

Display of Crop Stage Estimation Results

It was not until the users (that is, the analyst-interpreters performing labeling) had the opportunity to work with several potential display formats that this part of the system could be designed. During Phase I, the adjustments were made on the crop-reporting-district level. Several experimental formats were considered during this period. The display finally developed by Wilcox (ref. 8) employed a grid system. The crop calendar models would be operated at the grid point nearest the data input (meteorological station) location, and an objective analysis interpretive scheme was employed to extend the estimate from the meteorological stations to the sample segments.

CONCLUSIONS

Assessment of the ACC accuracy over the period of LACIE operation indicates that the adjustable crop calendars used did provide more accurate information than would have been available using historical normals. The models performed best under the conditions from which they were derived (Canadian spring wheat) and most poorly for the dwarf varieties and Southern Hemisphere applications. Refinements introduced into the model during LACIE resulted in some improvement in accuracy, and the supporting research and development ac-

tivities have ready for use other modifications that appear to provide increased accuracy. Any major improvement in accuracy, however, will be dependent on (1) a reliable starter model and (2) a developmental data set collected over a wide range of conditions with the specific goal of supporting development of crop calendar models. Recognition of this improvement, once obtained, would require acquisition of a more reliable test data set.

ACKNOWLEDGMENTS

As indicated by the referenced material, the authors of this paper are summarizing the work of a team of aggressive crop calendar developers. Particular credit should go to S. K. Woolley, J. W. Easley, R. G. Stuff, R. L. Baskett, W. W. Hildreth, and D. D. Wilcox for the work performed during the period they supported this effort through the Lockheed Electronics Company.

REFERENCES

1. Stuff, R. G.: Preliminary Evaluation of Phenological Models for Spring Wheat. Technical Memorandum 6134, Lockheed Electronics Co., Houston, Texas, June 1975.
2. Robertson, G. W.: A Biometeorological Time Scale for a Cereal Crop Involving Day and Night Temperatures and Photo-Period. *Internat. J. Biometeorol.*, vol. 12, no. 3, 1968, pp. 191-223.
3. Wilcox, D. D.; Champagne, G. J.; Baskett, R. L.; and Woolley, S. K.: A FORTRAN Implementation of the Robertson Phenological Model. Technical Memorandum LEC-5974, Lockheed Electronics Co., Houston, Texas, Apr. 1975.
4. Feyerherm, A. M.: Adjusting Robertson's Biometeorological Time Scale to Winter Wheat Environments and Varietal Maturities. Kansas State University, Feb. 23, 1976.
5. Baskett, R. L.; Hildreth, W. W.; Wilcox, D. D.; and Woolley, S. K.: Large Area Crop Inventory Experiment (LACIE) Phase II Winter Wheat End-of-Dormancy Restart Model. Technical Memorandum LEC-8268, Lockheed Electronics Co., Houston, Texas, Apr. 1976.
6. Feyerherm, A. M.: Planting Date and Wheat Yield Models, Final Report. Contract NAS9-14533, NASA Johnson Space Center, Houston, Texas, 1977.
7. Stuff, R. G.; and Phinney, D. E.: Climatic and Statistical Analysis of Spring Wheat Planting in the Dakotas. Proceedings of 68th Annual Meeting of American Society of Agronomy (abstract), 1976.
8. Wilcox, D. D.: System Description of the LACIE/YES Gridded Crop Calendar Report Writer. Technical Memorandum LEC-10004, Lockheed Electronics Co., Houston, Texas, Jan. 1977.

Accuracy Assessment: The Statistical Approach to Performance Evaluation in LACIE

A. G. Houston,^a A. H. Feiveson,^a R. S. Chhikara,^b and E. M. Hsu^b

INTRODUCTION

An important function in the LACIE is the evaluation of results obtained at various stages of the experiment. The objective of LACIE is not only to demonstrate the technological feasibility of estimating large-area wheat production using the LACIE approach but also to produce estimates which satisfy certain accuracy and reliability goals. The accuracy assessment effort is designed to check the accuracy of the products of the experimental operations throughout the crop growing season and to determine whether the procedures used are adequate to accomplish the desired accuracy and reliability goals. These goals are set out in greater detail in the LACIE requirements documents (refs. 1 to 3).

Objectives

The objectives addressed in the development of statistical methodology for assessing LACIE performance are as follows.

1. To determine whether the accuracy goal of the LACIE estimate of wheat production for a region or a country is being met—The LACIE accuracy goal is a "90/90" criterion for at-harvest wheat production, meaning that the at-harvest wheat production estimate for the region or country should be within 10 percent of the true production with a probability of at least 0.9.

2. To determine the accuracy and reliability of early-season estimates and estimates made at regular intervals throughout the crop season before harvest—This objective includes a determination of the degree to which the 90/90 criterion is supported at these intervals during the crop season.

3. To investigate the various sources of error in the LACIE estimates of wheat production, area, and yield; to quantify and relate these error sources to causal elements in the LACIE estimation process; and to recommend procedures for reducing error

Such an effort satisfies the need to provide timely identification of major problem areas that require improvement so that LACIE goals can be met. Once a major problem area is identified, it is the responsibility of the accuracy assessment function to relate the problem to causal elements in the LACIE estimation process and then to make recommendations to improve the technology.

Most of the accuracy assessment investigations are performed in the U.S. Great Plains, which is called the "yardstick" region. This region was selected because reliable independent estimates of wheat production, area, and yield for the state and higher levels are available from the U.S. Department of Agriculture (USDA) Statistical Reporting Service (SRS, recently renamed the Economics, Statistics, and Cooperatives Service) and because ground observations may be obtained at the segment level through the assistance of personnel in the USDA Agricultural Stabilization and Conservation Service (ASCS). However, the studies in the yardstick region are used to promote the development of LACIE procedures for obtaining reliable estimates for other countries.

Background

The LACIE was conducted in three phases.

Phase I.—During LACIE Phase I, wheat acreage in the U.S. Great Plains was estimated. Yield and production feasibility studies also were performed, but the accuracy assessment effort consisted of evaluating only the acreage estimation and aggregation procedures. The specific objectives for LACIE Phase I were (1) to develop consistent estimators of

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Houston, Texas.

the variance of the LACIE acreage estimates, (2) to assess the sampling and classification components in terms of their respective contributions to the variability of a large-area acreage estimate, and (3) to isolate the factors significantly affecting the classification performance and segment wheat proportion estimation by the Classification and Mensuration Subsystem (CAMS).

The bias in the LACIE acreage estimate for a particular region was estimated by the difference between the LACIE estimate and the at-harvest estimate for that region released by the USDA/SRS for the 1974-75 crop year. Separate classification and sampling error components were estimated, the former by comparisons of LACIE proportion estimates with ground observations obtained from 27 intensive test sites (ITS's) in 8 states and from 30 ground-observed segments (blind sites) in 2 states (Montana and North Dakota) and the latter by comparisons with county-level data from the 1969 U.S. Agricultural Census. Several investigations were performed to develop a better estimator of the variance of the stratum acreage estimate and to study factors important to the processing of segment data for subsequent wheat proportion estimation (refs. 4 and 5).

Phase II.—In Phase II, the accuracy assessment group continued to test and evaluate LACIE acreage estimates but expanded its efforts to include evaluation of yield and production estimates as well. Methodology for assessing LACIE performance in terms of the 90/90 criterion and for estimating sampling and classification error components was developed. Detailed error source investigations were made employing LACIE proportion estimates and ground-observed proportions for 150 blind sites and 27 ITS's.

Estimates of the coefficient of variation (CV) and the bias were used to evaluate the LACIE production estimate in terms of the 90/90 criterion at the U.S. Great Plains level, and a sensitivity analysis was performed to determine the effect of various errors on the LACIE production estimate. In the foreign area, 10 ITS's in Canada were studied and evaluated.

Phase III.—During Phase III, the accuracy assessment group continued to enlarge the scope of detailed evaluation of LACIE estimates and procedures over the nine-state yardstick spring and winter wheat region. Evaluations were also performed for the U.S.S.R. and Canada. The investigations made in Phase III were similar to those per-

formed in Phase II. The relative contributions of classification and sampling error components were assessed using 212 blind sites and 24 ITS's in the United States, 30 test sites and 10 ITS's in Canada, and the 1974 U.S. Agricultural Census data. In addition, in support of the development of CAMS Procedure 1, efforts were made to determine analyst labeling and classification omission and commission errors.

STATISTICAL ANALYSIS OF PRODUCTION ESTIMATION

Evaluation Technique

A major part of the accuracy assessment effort is devoted to determining whether the operational procedures produce an estimator that meets the 90/90 accuracy goal of LACIE. This accuracy criterion was specified by the experimenter to derive technology improvements. To describe the criterion in exact terms, it is formulated statistically as follows.

Let \hat{P} be the LACIE estimate of wheat production for the region or country and let P be the true wheat production of the same region or country. The accuracy goal of LACIE is a 90/90 criterion for at-harvest wheat production, which is defined by the following probability statement.

$$\Pr \left[\left| \hat{P} - P \right| \leq 0.1P \right] \geq 0.9 \quad (1)$$

Equation (1) is a statement that the accuracy goal is for the LACIE estimate of wheat production to be within 10 percent of the true wheat production with a probability of at least 90 percent.

In LACIE, estimation of acreage, yield, and production is made for large areas, using data from many sample segments. Thus, it is assumed that the LACIE estimate \hat{P} is normally distributed, with mean $(P + B)$ and variance $\sigma_{\hat{P}}^2$, where B is the bias given by

$$B = E(\hat{P}) - P \quad (2)$$

Accordingly, the probability statement, equation (1), can be expressed

$$Pr \left[\frac{0.1 - 0.9RB(\hat{P})}{CV(\hat{P})} < Z < \frac{0.1 + 1.1RB(\hat{P})}{CV(\hat{P})} \right] > 0.90 \quad (3)$$

where $Z = [\hat{P} - (P + B)]/\sigma_{\hat{P}}$ follows the standard normal distribution $N(0,1)$. The parameter $CV(\hat{P})$ is the coefficient of variation of \hat{P} defined by

$$CV(\hat{P}) = \frac{\sigma_{\hat{P}}}{E(\hat{P})} = \frac{\sigma_{\hat{P}}}{P + B} \quad (4)$$

The term $RB(\hat{P})$ is called the "relative bias" of \hat{P} and is defined by

$$RB(\hat{P}) = \frac{E(\hat{P}) - P}{E(\hat{P})} = \frac{B}{P + B} \quad (5)$$

It follows that the accuracy goal of LACIE is attained if

$$\Phi \left[\frac{0.1 + 1.1RB(\hat{P})}{CV(\hat{P})} \right] - \Phi \left[\frac{0.1 - 0.9RB(\hat{P})}{CV(\hat{P})} \right] > 0.90 \quad (6)$$

where Φ represents the cumulative standard normal distribution. The area under the curve in figure 1 contains combinations of $CV(\hat{P})$ and $RB(\hat{P})$ for which equation (6) is satisfied.

Variance and Bias Estimation for the Wheat Production Estimate

To apply the evaluation technique described in the previous section, knowledge of the variance $\sigma_{\hat{P}}^2$ and the bias B of the LACIE wheat production estimate for a country or a region is required. Since values for these parameters are unknown in LACIE, estimates have to be obtained. The estimation of the production variance at different aggregation levels is described in detail in the paper entitled "Large-Area

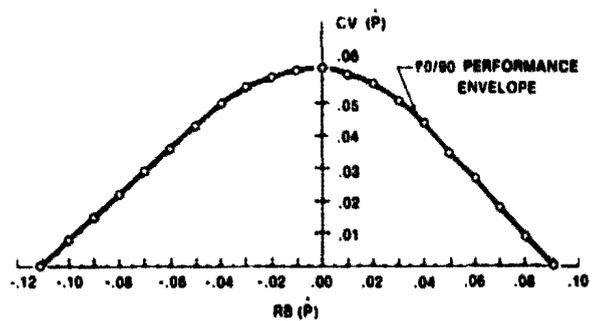


FIGURE 1.—Relative bias compared to coefficient of variation of LACIE production estimator \hat{P} . The shaded area includes combinations of $RB(\hat{P})$ and $CV(\hat{P})$ that will satisfy the 90/90 accuracy criterion.

Aggregation and Mean-Squared Prediction Error Estimation for LACIE Yield and Production Forecasts" by Chhikara and Feiveson.

An estimate of bias can be obtained from the difference between a LACIE production estimate and the corresponding USDA estimate of production, but this estimate is realistic only in the United States and is based only on a single sample. For foreign countries, the USDA Foreign Agricultural Service (FAS) makes periodic forecasts, which are generally for total grain production, using ad hoc methods. Although FAS estimates may be used to indicate a major problem, they cannot be used for a quantitative assessment of bias in a LACIE estimate.

The 90/90 Evaluation

Given the LACIE estimator $\hat{\sigma}_{\hat{P}}$ of the standard deviation $\sigma_{\hat{P}}$ of the LACIE production estimator, an estimate of $CV(\hat{P})$ is

$$\hat{CV}(\hat{P}) = \hat{\sigma}_{\hat{P}}/\hat{P} \quad (7)$$

The computation of $\hat{\sigma}_{\hat{P}}$ is described in detail in the paper by Chhikara and Feiveson and is the result of the Crop Assessment Subsystem (CAS) aggregation software. An estimate of $RB(\hat{P})$ is

$$\hat{RB}(\hat{P}) = \frac{\hat{B}}{\hat{P}} \quad (8)$$

where \hat{B} is the difference between the LACIE production estimate and the corresponding SRS estimate.

The distribution of the estimated value of the left side of equation (6) with $CV(\hat{P})$ and $RB(\hat{P})$ replaced by their estimates, $CV(\hat{P})$ and $RB(\hat{P})$, respectively, has been found to be intractable because of problems in obtaining a joint distribution of $CV(\hat{P})$ and $RB(\hat{P})$. However, if $CV(\hat{P}) > 0.061$, there is a fair indication that the LACIE estimate may not satisfy the 90/90 criterion even if \hat{P} is assumed unbiased. Since $CV(\hat{P})$ has been found to be very stable at the country level (U.S. Great Plains level in the case of the United States) and less than 0.061, one can treat $CV(\hat{P})$ as the parameter $CV(\hat{P})$ and solve equation (6) to determine the tolerable values of $RB(\hat{P})$ that would meet the 90/90 accuracy goal. That is, given $CV(\hat{P})$, there exist real numbers R_0 ($R_0 < 0$) and R_1 ($R_1 > 0$) such that equation (6) is satisfied for

$$R_0 < RB(\hat{P}) < R_1 \quad (9)$$

Equivalently, there exist corresponding tolerable bias limits B_0 and B_1

$$B_0 < B < B_1 \quad (10)$$

where $B_0 = [R_0/(1 - R_0)]P$ and $B_1 = [R_1/(1 - R_1)]P$, where P is the actual production.

Suppose next a null hypothesis H_0 that the LACIE production estimate is from a 90/90 estimator; i.e., suppose $CV(\hat{P}) \equiv CV(\hat{P}) < 0.061$ and $RB(\hat{P}) \in [R_0, R_1]$ and hence $B \in [B_0, B_1]$. To test the hypothesis that H_0 is true, first fix a value of B , say $B^* \in [B_0, B_1]$, then test the subhypothesis $B = B^*$ against the alternative $B \neq B^*$, using the statistic $\hat{B} = \hat{P} - P_{SRS}$ and assuming $\hat{B} \sim N(B, \hat{\sigma}_{\hat{P}}^2)$. A "p-value" for this test is given by

$$\Pi(B^*) = Pr\left\{|\hat{B} - B^*| > |b - B^*|\right\} \quad (11)$$

given $\hat{B} \sim N(B^*, \hat{\sigma}_{\hat{P}}^2)$, where b is the observed difference, $\hat{P} - P_{SRS}$. The overall hypothesis, $H_0: \hat{P}$ is

90/90, is rejected if

$$\max_{B^* \in [B_0, B_1]} \Pi(B^*) < \alpha \quad (12)$$

where α is a predetermined significance level. If the test fails to reject H_0 , it is *not* immediately inferred that the LACIE production estimator is a 90/90 estimator. (The test has low power since only one observation is available to estimate $RB(\hat{P})$.) In this situation, the statement is made that "support of the 90/90 accuracy goal" is indicated; however, results obtained from blind-site analyses and other accuracy assessment tasks are then considered for further assessment of whether or not the 90/90 criterion is achievable.

Comparison of LACIE Estimates With Reference Standards

The reference standards to which the LACIE estimates are compared are the USDA/SRS estimates for the United States and the FAS estimates and/or official country estimates for foreign countries. The statistic used for making these comparisons is the relative difference (RD) in percent defined as follows.

$$RD = \left(\frac{LACIE - STANDARD}{LACIE} \right) \times 100 \quad (13)$$

where LACIE stands for the LACIE estimate of wheat production, area, or yield and STANDARD represents the corresponding reference standard estimate. This definition expresses the difference between the two estimates as a percentage of the LACIE estimate.

Significance tests of no difference are made only at the region or country level for the LACIE production, area, and yield estimates of spring wheat, winter wheat, and total wheat. For a significance test, the LACIE estimate (of wheat production, area, or yield) is assumed to be approximately normally distributed with unknown mean μ and variance σ_{LACIE}^2 . A test of the hypothesis $H_0: \mu = STANDARD$ against the alternative hypothesis $H_1: \mu \neq$

STANDARD is then made using this assumption. The test statistic is given by

$$Z = \frac{\text{LACIE} - \text{STANDARD}}{\hat{\sigma}_{\text{LACIE}}} \quad (14)$$

which, under the null hypothesis, is approximately normally distributed with mean 0 and variance 1. The null hypothesis is rejected in favor of the alternative at the α level of significance if

$$|Z| > z_{\alpha/2} \quad (15)$$

where $z_{\alpha/2}$ is the $(1 - \alpha/2)$ critical point of the standard normal distribution. For $\alpha = 0.10$, $z_{\alpha/2} = 1.645$, and if $|Z| > 1.645$, it is concluded that the mean of the LACIE estimator is significantly different from the reference standard estimate.

ERROR SOURCES IN LACIE

Any uncertainty in a LACIE prediction of wheat production at the country level is directly related to errors in wheat acreage estimates and yield predictions at the zone level. These errors are incurred independently and, hence, estimated accordingly.

The yield prediction error is evaluated on the basis of the residual mean square error obtained by regressing yield on weather data for past years. The error in country-level yield prediction is assessed by taking into account the variability with which LACIE acreage estimates are obtained. (See the paper by Chhikara and Feiveson.) When extreme weather conditions prevail, the yield prediction is likely to be biased. In addition to making a comparison between LACIE yield estimates and the reference standard estimates, another evaluation is made using historical data. (For details of this methodology, see the paper entitled "Accuracy and Performance of LACIE Yield Estimates in Major Wheat Producing Regions of the World" by Phinney et al.)

The acreage estimate is subject to both bias and variability. Sampling and classification are the two major error components of an acreage estimation error. Sampling error contributes primarily to the

variance of the acreage estimate, whereas classification error is the main contributor to the bias in acreage estimate. In general, estimated within-stratum variances are input to the variance estimate of a zone acreage estimate and consist of sampling as well as classification variance components. The bias is incurred at two levels: segment and stratum. The segment-level bias is due to the classification procedure that first determines the small-grains proportion in a segment and then converts it to a wheat proportion by applying the stratum-level historical ratio of wheat to small grains. The stratum-level bias in the United States is due to the segment-level bias and to the ratio estimation of wheat acreage for Group III counties. (For a definition of Group III counties, see reference 1 or the paper entitled "LACIE Large-Area Acreage Estimation" by Chhikara and Feiveson.)

The error sources that contribute to the prediction error (mean squared error) of a LACIE production estimate are outlined in figure 2. Though it is desirable to assess the individual contributions of these error sources, the complexity of their interaction and lack of knowledge of true area and yield make intractable the estimation of all error components and the relation of the components to the overall error.

First-Order Error Source Investigations

First-order errors are those errors contributing to the LACIE production estimate \hat{P} , which can be approximately quantified using LACIE, SRS, historical, and blind-site data. (A "blind site" is a sample segment selected from the set of LACIE-allocated segments for the purpose of acquiring ground observations of the true distribution of crops.) The error in \hat{P} depends on its sources in a complex way; thus, it is unrealistic to assume the total error can be written as a sum of uncorrelated random components. Instead, the effect of each component is assessed by estimating the reduction in the prediction error in \hat{P} achieved by removing that error. A major accuracy assessment effort is devoted to the development of statistical methodology for estimating the acreage bias and its error components.

Effect of sampling, classification, and yield variability on the variance of the production estimate.—The effect of a particular error source is assessed by determining the reduction in the production variance estimate when the error is eliminated from the computation of the variance estimate. Suppose $\sigma_{\hat{P}}^2$ is the

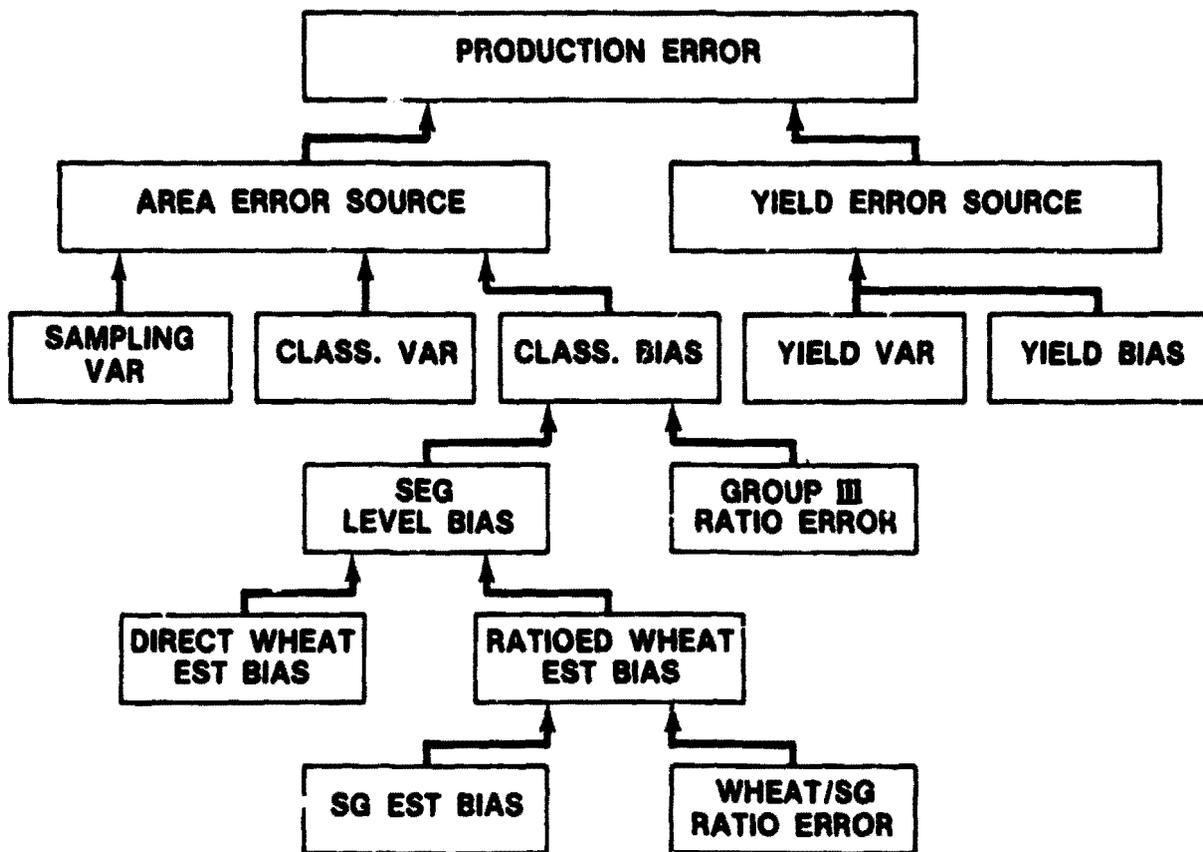


FIGURE 2.—LACIE first-order error components.

variance of the country production estimate. Then, as described in the Chhikara and Feiveson paper, $\sigma_{\hat{p}}^2$ can be expressed as

$$\sigma_{\hat{p}}^2 = \sum_{i=1}^n \sigma_i^2 + 2 \sum_{i > j} \sigma_{ij} \quad (16)$$

where σ_i^2 is the variance of the i th and j th pseudozone production estimates. The variances $\{\sigma_i^2\}$ and covariances $\{\sigma_{ij}\}$ can further be expressed in terms of acreage and yield error components, and both the acreage error and the covariance term can be further subdivided into sampling and classification error components if the latter can be estimated by means of the following procedure.

To assess the effect of an error component on the production variance, σ_i^2 and σ_{ij} need to be estimated with that error component omitted. Suppose $S_{\hat{p}}^2$, $S_{\hat{p}_1}^2$, $S_{\hat{p}_2}^2$, $S_{\hat{p}_3}^2$, and $S_{\hat{p}_4}^2$ are the estimates of $\sigma_{\hat{p}}^2$ when the error component omitted is none, yield,

acreage, sampling, and classification, respectively. Then the ratios $S_{\hat{p}_1}^2/S_{\hat{p}}^2$, $S_{\hat{p}_2}^2/S_{\hat{p}}^2$, $S_{\hat{p}_3}^2/S_{\hat{p}}^2$, and $S_{\hat{p}_4}^2/S_{\hat{p}}^2$ are determined to evaluate the sensitivity of the production variability to the yield, acreage, sampling, and classification error components, respectively.

Area error source investigations.—Area error source investigations consist of estimating bias at the regional and segment levels and determining contributions of sampling and classification error to segment and regional acreage estimation variability.

Estimating bias at the regional level: The method for estimating bias described in this section is valid for any area having a sufficient number of blind sites. In the accuracy assessment of LACIE area estimates, it is applied at the state and higher levels.

The LACIE estimate of wheat acreage \hat{A} for a given area can be written

$$\hat{A} = \sum_{i=1}^n w_i \hat{A}_i \quad (17)$$

where \hat{x}_i is the wheat proportion estimate in the i th LACIE segment, n is the number of processed LACIE segments, and w_i ($i = 1, \dots, n$) is a known weight based on the size of the substratum in which the i th segment is located, the number of segments in this substratum, and the historical data of any Group III substrata the wheat acreages of which are estimated by means of the Group III ratio involving this substratum.

Corresponding to the estimate \hat{A} is the true acreage A , which may be expressed as

$$A = \sum_{i=1}^n w_i^* c_i \quad (18)$$

where c_i is the true wheat acreage for the substratum containing the i th segment and w_i^* is the value of the weight which would give perfect Group III estimates of wheat acreage for unsampled areas using these n acquired segments.

The wheat proportion estimate for the i th segment can be expressed by the identity

$$\begin{aligned} \hat{x}_i &= c_i + (x_i - c_i) + (\hat{x}_i - x_i) \\ &= c_i + \epsilon_i + \delta_i \end{aligned} \quad (19)$$

where x_i is the true wheat proportion of the i th segment, ϵ_i is the sampling error, and δ_i is the classification error. Since segments are located randomly in substrata, the sampling is unbiased and $E(\epsilon_i) = 0$. However, unbiased classification is not assumed and

$$E(\delta_i) = \theta_i \quad (20)$$

where θ_i is unknown.

The bias in \hat{A} , defined by $E(\hat{A} - A)$, is given by

$$\begin{aligned} B &= E\left(\sum_{i=1}^n w_i \hat{x}_i - \sum_{i=1}^n w_i^* c_i\right) \\ &= \sum_{i=1}^n w_i E(c_i + \epsilon_i + \delta_i) - \sum_{i=1}^n w_i^* c_i \end{aligned}$$

$$\begin{aligned} &= \sum_{i=1}^n (w_i - w_i^*) c_i + \sum_{i=1}^n w_i \theta_i \\ &= B_1 + B_2 \end{aligned} \quad (21)$$

Note that the first term, B_1 , represents a bias caused by the lack of exactness of the Group III ratios (i.e., $w_i \neq w_i^*$), whereas the second term, B_2 , is the bias due to classification.

The classification bias component B_2 is estimated by

$$\hat{B}_2 = \frac{n}{m} \sum_{j=1}^m w_j \hat{\theta}_j \quad (22)$$

where m is the number of blind sites in the area containing n processed segments.

$$\hat{\theta}_j = \hat{x}_j - x_j \quad (23)$$

for the j th blind site, where x_j is the ground-observed proportion of wheat for that segment. Since the blind sites are a random subsample, \hat{B}_2 is an unbiased estimator of B_2 ; i.e.,

$$E(\hat{B}_2) = B_2 \quad (24)$$

The variance of \hat{B}_2 is

$$\text{Var}(\hat{B}_2) = \frac{n^2}{m} \left(1 - \frac{m}{n}\right) S^2 \quad (25)$$

where

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n \left(w_i \theta_i - \frac{1}{n} \sum_{j=1}^n w_j \theta_j\right)^2 \quad (26)$$

This variance is estimated by replacing S^2 with its estimate

$$\hat{\sigma}^2 = \frac{1}{m-1} \sum_{j=1}^m \left(w_j \hat{\theta}_j - \frac{1}{m} \sum_{j=1}^m w_j \hat{\theta}_j \right)^2 \quad (27)$$

An approximate 90-percent confidence interval for B_2 is constructed by $(\hat{B}_2 - 1.645\hat{\sigma}, \hat{B}_2 + 1.645\hat{\sigma})$, where $\hat{\sigma}^2$ is the estimate of $\text{Var}(\hat{B}_2)$.

Reliable county-level data are not often available for estimating B_1 , the bias due to Group III ratio estimation. Agricultural census data at the county level are available only every 4 to 5 years, the latest in 1974. These most recent census data are used to obtain the Group III ratio estimates in the LACIE aggregation scheme and hence cannot be used for estimating B_1 . Therefore, county-level SRS estimates, which are the only independent estimates available, are used for estimating B_1 . It is known that the SRS estimates are not very reliable at the county level; therefore, the following estimate of B_1 is obtained only at the U.S. Great Plains level and is used with caution.

Because current SRS county-level estimates are not available during the crop year, previous-year county-level SRS estimates are used to obtain the c_i in the equation

$$B_1 = \sum_{i=1}^n c_i (w_i - w_i^0) \quad (28)$$

for each of the processed LACIE segments in the U.S. Great Plains. Then, B_1 is estimated by

$$\hat{B}_1 = \sum_{i=1}^n w_i c_i^{\text{SRS}} - A_{\text{SRS}} \quad (29)$$

where c_i^{SRS} is the SRS wheat proportion for the county containing the i th segment and A_{SRS} is the SRS wheat area estimate for the U.S. Great Plains. A reliable estimate of the variance of B_1 is not available; thus, for practical purposes, the bias due to Group III ratio estimation is considered negligible if B_1 is less than 2 percent of A_{SRS} .

Estimating bias at the segment level: In this section, the statistical methodology for estimating the wheat proportion estimation error expected for the CAMS-processed segments is described. Let N be the number of segments acquired in a region (state or higher level) and let n be the number of blind sites selected randomly from these N segments. For a region, let \hat{x}_i represent the wheat proportion estimate in the i th segment and let x_i represent the ground-observed proportion of wheat in the i th segment, where $i = 1, \dots, N$. Then, the average error μ_D is given by

$$\mu_D = \frac{1}{N} \sum_{i=1}^N (\hat{x}_i - x_i) \quad (30)$$

The estimate of μ_D is given by

$$\bar{D} = \frac{1}{n} \sum_{i=1}^n (\hat{x}_i - x_i) \quad (31)$$

Letting $D_i = \hat{x}_i - x_i$, $i = 1, 2, \dots, n$, the variance of \bar{D} is estimated by

$$s_D^2 = \left(\frac{1}{n} - \frac{1}{N} \right) \frac{\sum_{i=1}^n (D_i - \bar{D})^2}{n-1} \quad (32)$$

Lower and upper confidence limits, respectively, for the population average difference μ_D are given by

$$\mu_{D_L} = \bar{D} - t_{1-\alpha/2} s_D \quad (33a)$$

$$\mu_{D_U} = \bar{D} + t_{1-\alpha/2} s_D \quad (33b)$$

where $t_{1-\alpha/2}$ is the value of the $1 - \alpha/2$ percentage point, from the Student's t distribution with $(n-1)$ degrees of freedom, corresponding to the desired confidence level of $1 - \alpha$. If μ_D is inferred to be significantly different from zero, contributions to the

bias and mean squared error (MSE) due to small-grains classification error and wheat-to-small-grains ratio error are estimated (unless a direct wheat classification procedure was used).

Let \hat{r}_i and \hat{x}_i ($i = 1, 2, \dots, n$) be the estimates of r_i and x_i , respectively, for the i th blind site, where r_i is the ground-observed ratio of wheat to small grains, x_i is the ground-observed small-grains proportion, and n is the number of blind sites. In LACIE, \hat{r}_i is the forecast ratio of wheat to small grains, and \hat{x}_i is the CAMS estimate of the small-grains proportion.

The bias B and the MSE of the wheat proportion estimate obtained after ratioing may be estimated by

$$\hat{B} = \frac{1}{n} \sum_{i=1}^n (\hat{r}_i \hat{x}_i - r_i x_i) \quad (34)$$

and

$$\hat{MSE} = \frac{1}{n} \sum_{i=1}^n (\hat{r}_i \hat{x}_i - r_i x_i)^2 \quad (35)$$

It is clear that both these errors are caused by two factors: the CAMS classification of small grains and the estimated ratio of wheat to small grains. The contribution of a particular error factor may be assessed by the reduction in the bias or the MSE which would be achieved if that error factor were omitted. Specifically, the following formulas are used in this study.

1. Bias estimate with no ratioing error:

$$\hat{B}' = \frac{1}{n} \sum_{i=1}^n (r_i \hat{x}_i - r_i x_i) \quad (36)$$

2. Bias estimate with no classification error:

$$\hat{B}'' = \frac{1}{n} \sum_{i=1}^n (\hat{r}_i x_i - r_i x_i) \quad (37)$$

3. MSE estimate with no ratioing error:

$$\hat{MSE}' = \frac{1}{n} \sum_{i=1}^n (r_i \hat{x}_i - r_i x_i)^2 \quad (38)$$

4. MSE estimate with no classification error:

$$\hat{MSE}'' = \frac{1}{n} \sum_{i=1}^n (\hat{r}_i x_i - r_i x_i)^2 \quad (39)$$

These quantities are calculated at the state and U.S. Great Plains levels, and a sensitivity analysis is conducted to measure the effect of classification and ratio error on the bias and the MSE for ratioed wheat proportion.

Contributions of sampling and classification error to segment and regional acreage estimation variability: The variance of the LACIE acreage estimate \hat{A} for a large area (e.g., a zone) can be written

$$\text{Var}(\hat{A}) = V^2 = \sum_i w_i^2 \sigma_i^2 \quad (40)$$

where σ_i^2 is the variance of the acreage estimate for the i th substratum (county) and w_i is a weight which depends on the size of the substratum, the number of segments in the substratum, etc. (For a description of the estimation of V^2 , see the paper entitled "LACIE Sampling Design" by Feiveson et al.)

The variance σ_i^2 represents a mean squared deviation between the LACIE estimate for the county wheat proportion and the true county wheat proportion. This variance is caused mainly by two factors: sampling error and classification error.

It follows from the assumptions in equation (41) that the i th substratum acreage error variance σ_i^2 can be written $\sigma_i^2 = \sigma_c^2 + \lambda^2 \sigma_s^2$, where σ_c^2 is a contribution resulting from classification, and $\lambda^2 \sigma_s^2$ is a contribution caused by sampling. To determine the effect of no classification error, the variance of the LACIE acreage estimate will be calculated using $\rho \sigma_i^2$ instead of σ_i^2 where ρ is the ratio $(\lambda^2 \sigma_s^2) / (\sigma_c^2 + \lambda^2 \sigma_s^2)$. Similarly, the effect of no sampling error is estimated by replacing σ_i^2 by $(1 - \rho) \sigma_i^2$. In the following discussion, the method employed for estimating sampling and classification variances and the function ρ is described.

It will be assumed that, for some reasonably large area (e.g., a zone), the sampling and classification errors ϵ_i and δ_i have the following properties.

$$\left. \begin{aligned} \epsilon_i \text{ and } \delta_i \text{ are uncorrelated} \\ E(\epsilon_i) = 0 \\ E(\delta_i | X_i) = \lambda^* X_i + \theta \\ V(\epsilon_i) = \sigma_s^2 \\ V(\delta_i | X_i) = \sigma_c^2 \end{aligned} \right\} \quad (41)$$

It is also assumed that there is a linear model relating the current-year county proportion C_i to the historical proportion, which will be denoted by z_i ; i.e.,

$$C_i = \alpha + \beta z_i + \zeta_i \quad (42)$$

where $E(\zeta_i) = 0$, $V(\zeta_i) = \sigma_h^2$, $\text{Cov}(\zeta_i, \epsilon_i) = \text{Cov}(\zeta_i, \delta_i) = 0$, and α and β are regression coefficients.

From the preceding assumptions and definitions, three basic regression models are obtained.

1. True segment proportion compared to historical county proportion—From the definition of ϵ_i ,

$$\begin{aligned} X_i &= C_i + \epsilon_i \\ &= \alpha + \beta z_i + \zeta_i + \epsilon_i \end{aligned} \quad (43)$$

It follows that

$$E(X_i) = \alpha + \beta z_i \quad (44)$$

$$V(X_i) = \sigma_h^2 + \sigma_s^2 \quad (45)$$

2. LACIE segment proportion compared to

ground-truth segment proportion—From the definition of δ_i ,

$$\hat{X}_i = X_i + \delta_i \quad (46)$$

It follows that

$$E(\hat{X}_i | X_i) = X_i + \lambda^* X_i + \theta \quad (47)$$

$$V(\hat{X}_i | X_i) = \sigma_c^2 \quad (48)$$

Writing $\lambda = 1 + \lambda^*$, one obtains

$$E(\hat{X}_i | X_i) = \lambda X_i + \theta \quad (49)$$

$$V(\hat{X}_i | X_i) = \sigma_c^2 \quad (48)$$

3. LACIE segment proportion compared to historical county proportion—From equations (44) to (49),

$$\begin{aligned} E(\hat{X}_i) &= E_{X_i} [E(\hat{X}_i | X_i)] \\ &= E_{X_i} (\lambda X_i + \theta) \\ &= \lambda(\alpha + \beta z_i) + \theta \end{aligned} \quad (50)$$

$$\begin{aligned} V(\hat{X}_i) &= E_{X_i} [V(\hat{X}_i | X_i)] + V_{X_i} [E(\hat{X}_i | X_i)] \\ &= \sigma_c^2 + \lambda^2 (\sigma_h^2 + \sigma_s^2) \end{aligned} \quad (51)$$

As stated previously, it is desirable to estimate $\rho = (\lambda^2\sigma_s^2)/(\sigma_c^2 + \lambda^2\sigma_s^2)$. None of the three regression models enables an estimate of σ_s^2 separately from σ_h^2 ; i.e., one can only estimate $\sigma_s^2 + \sigma_h^2$, not σ_s^2 alone. If current-year county proportions C_i were available, σ_h^2 could be estimated; but, since this is not the case, $\rho^* = [\lambda^2(\sigma_s^2 + \sigma_h^2)]/[\sigma_c^2 + \lambda^2(\sigma_s^2 + \sigma_h^2)]$ will be estimated instead of ρ . If $\sigma_h^2 \ll \sigma_s^2$ (a reasonable assumption), then $\rho^* \approx \rho$.

The procedure for estimating ρ^* is described as follows. Suppose a given zone has m blind-site segments and n ordinary (i.e., not blind site) segments, and let the blind-site segments be numbered 1 to m . It is assumed that ground-observed wheat proportions X_i , $i = 1, \dots, m$ are available for the blind sites and LACIE estimates \hat{X}_i , $i = 1, \dots, m+n$ are available for all segments. It is also assumed that historical wheat proportions Z_i , $i = 1, \dots, m+n$ are available for the counties containing the segments. If $\sigma_h^2 \ll \sigma_s^2$ so that $\rho \approx \rho^*$, regression models 1 to 3 are applicable.

$$E(X_i) = \alpha + \beta Z_i; V(X_i) = \sigma_s^2; i = 1, \dots, m \quad (52)$$

$$E(\hat{X}_i | X_i) = \lambda X_i + \nu; V(\hat{X}_i | X_i) = \sigma_c^2; i = 1, \dots, m \quad (53)$$

$$\left. \begin{aligned} E(\hat{X}_i) &= \theta + \lambda\alpha + \lambda\beta z_i; \\ V(\hat{X}_i) &= \lambda^2\sigma_s^2 + \sigma_c^2; i = m+1, \dots, m+n \end{aligned} \right\} \quad (54)$$

If there is one segment per county, then the errors ϵ_i and δ_i are independent for different values of i ; hence, the likelihood function of the sample can be written

$$L = \prod_{i=1}^m f(X_i, \hat{X}_i) \prod_{i=m+1}^{m+n} h(\hat{X}_i) \quad (55)$$

where $f(X_i, \hat{X}_i)$ is the joint density of X_i and \hat{X}_i for $i = 1, \dots, m$ and $h(\hat{X}_i)$ is the density of \hat{X}_i for $i = m+1, \dots, m+n$.

The function

$$\prod_{i=1}^m f(X_i, \hat{X}_i)$$

can be expressed

$$\prod_{i=1}^m f(\hat{X}_i | X_i) g(X_i)$$

where $f(\hat{X}_i | X_i)$ is the conditional density of \hat{X}_i given X_i , and $g(X_i)$ is the density function of X_i . Assuming errors to be normally distributed, the likelihood function L can be specified since

$$\begin{aligned} \prod_{i=1}^m f(X_i, \hat{X}_i) &= \prod_{i=1}^m \frac{1}{\sigma_c \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma_c^2} \sum_{i=1}^m (\hat{X}_i - \lambda X_i - \nu)^2 \right\} \\ &\times \frac{1}{\sigma_s \sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma_s^2} \sum_{i=1}^m (X_i - \alpha - \beta z_i)^2 \right\} \end{aligned} \quad (56)$$

and

$$\begin{aligned} \prod_{i=m+1}^{m+n} h(\hat{X}_i) &= \frac{1}{(\lambda^2\sigma_s^2 + \sigma_c^2)^{\frac{1}{2}} \sqrt{2\pi}} \\ &\times \exp \left\{ -\frac{1}{2(\lambda^2\sigma_s^2 + \sigma_c^2)} \sum_{i=m+1}^{m+n} (\hat{X}_i - \lambda\alpha - \lambda\beta z_i)^2 \right\} \end{aligned} \quad (57)$$

Letting $Q = -2 \log L$ and equating to zero the partial differentials of Q with respect to all unknown parameters, the maximum likelihood equations are obtained for $\alpha, \beta, \theta, \lambda, \sigma_c^2$, and σ_s^2 . If $\hat{\alpha}, \hat{\beta}, \hat{\theta}, \hat{\lambda}, \hat{\sigma}_c^2$, and $\hat{\sigma}_s^2$ represent the solutions of these equa-

tions, then the invariance theorem for maximum likelihood estimation can be used to obtain

$$\hat{\rho} = \frac{(\hat{\lambda})^2 \sigma_s^2}{\hat{\sigma}_c^2 + (\hat{\lambda})^2 \hat{\sigma}_s^2} \quad (58)$$

as the maximum likelihood estimate of ρ . The maximum likelihood equations are nonlinear but can be solved using numerical techniques; e.g., Newton's method.

Since $\hat{\rho}$ is a complicated function of the data, it is impossible to write down the variance of $\hat{\rho}$ for finite sample sizes m and n . However, the asymptotic variance of $\hat{\rho}$ can be estimated using the information matrix, i.e., if

$$V = (V_{ij}) \\ = E \left\{ \frac{-\partial^2 \log L}{\partial u_i \partial u_j} \right\} \quad (59)$$

and $g(u)$ is a differentiable function of the parameter vector $u = (\hat{\alpha}, \hat{\beta}, \hat{\theta}, \hat{\lambda}, \hat{\sigma}_c^2, \hat{\sigma}_s^2)$, then the variance of $g(u)$ is asymptotic to

$$[g'(u)]^T V^{-1} g'(u) \quad (60)$$

where $g'(u) = (\partial g / \partial u_1, \dots, \partial g / \partial u_6)^T$ and T stands for the transpose of a vector or matrix. Thus, in this case, $g(u) = (\lambda^2 \sigma_s^2) / (\lambda^2 \sigma_s^2 + \sigma_c^2)$ and

$$g'(u) = \left[\begin{array}{l} 0, 0, 0, 2\lambda^2 \sigma_c^2 (\lambda^2 \sigma_s^2 + \sigma_c^2)^{-2}, \\ -\lambda^2 \sigma_s^2 (\lambda^2 \sigma_s^2 + \sigma_c^2)^{-2}, -\frac{\lambda^2 \sigma_c^2}{(\sigma_c^2 + \lambda^2 \sigma_s^2)^2} \end{array} \right] \quad (61)$$

To estimate V , the values $\{x_i\}$, $\{\hat{x}_i\}$, and $\{z_i\}$ and the estimated parameters $(\hat{\alpha}, \hat{\beta}, \hat{\theta}, \hat{\lambda}, \hat{\sigma}_c^2$ and $\hat{\sigma}_s^2)$ are substituted into the matrix $H = (h_{ij}) = (\partial^2 \log L) / (\partial u_i \partial u_j)$. Then, equation (60) is used to obtain an approximate variance for $\hat{\rho}$. Assuming that $\hat{\rho}$, which is the ratio of the within-county sampling variance estimate to the total within-county area variance estimate, also applies to a large area, the estimated variances of the regional area estimate due to classification ($\hat{\eta}^2$) and sampling ($\hat{\nu}^2$) are given by

$$\hat{\eta}^2 = (1 - \hat{\rho}) \hat{\nu}^2 \quad (62)$$

$$\hat{\nu}^2 = \hat{\rho} \hat{\nu}^2 \quad (63)$$

where $\hat{\nu}^2$ denotes the estimated acreage variance for the large-area estimate. Consequently, the estimated CV of a large-area estimate \hat{A} due to classification is given by

$$\hat{CV}(\hat{A}|C) = \frac{\hat{\eta}}{\hat{A}} \quad (64)$$

and that due to sampling is given by

$$\hat{CV}(\hat{A}|S) = \frac{\hat{\nu}}{\hat{A}} \quad (65)$$

where $\hat{CV}(\hat{A}|C)$ and $\hat{CV}(\hat{A}|S)$ are often casually referred to in LACIE as the classification CV and the sampling CV, respectively.

Second-Order Error Source Investigations

A major effort is made in LACIE to study the sources of the errors that influence the LACIE production estimation to ascertain the accuracy of the procedures being used and to devise ways of improving these procedures.

Yield estimation investigations.—The purpose of yield estimation investigations is to determine factors introducing errors into the Center for Climatic

and Environmental Assessment (CCEA) yield model predictions at the pseudozone levels. For diagnostic purposes, the following plots are developed for each pseudozone for each yield truncation.

1. Precipitation as a function of time of year for the current year and for the 3 maximum and the 3 minimum yield years, as determined from the historical data

2. Temperature as a function of time of year for the current year and for the 3 maximum and the 3 minimum yield years

3. Means and standard deviations of temperature and precipitation as a function of time of year—The monthly temperature and precipitation for the current year are plotted on these charts for diagnostic purposes.

The following diagnostic checks are also made.

1. Calculate sampling error by using meteorological data from the cooperative station network.

2. Check the data base at the pseudozone level for clerical errors.

3. Check for episodic weather conditions and resultant impact on yield estimates.

4. Evaluate models:

- a. Reevaluate variable selection by adding current-year meteorological data.

- b. Perform latent root regression on pseudozone data to calculate the most stable variables for predicting yield (without allowing for a trend term).

- c. Investigate trend term by performing latent root regression without allowing for trend and calculating trend on the residuals from the most stable fit.

During LACIE Phases I, II, and III, several investigations were performed to evaluate and improve the classification program for estimating segment wheat or small-grains proportion. Analyses of variance models were employed in several of these studies. (See references 6 and 7 and the paper by Chhikara and Feiveson entitled "LACIE Large-Acreage Estimation.") Biostage, analyst-interpreter (AI), segment location, and ground-observed wheat or small-grains proportion were the factors evaluated for their effect on the variability of segment wheat or small-grains proportion estimation by CAMS. Studies on omission and commission errors in labeling of classes by AI's, as well as those resulting from classification algorithms, have also been conducted. Evaluations were often investigative in nature and the methodology used was generally restricted to plotting and tabulating data, fitting data by regression to examine relationships,

and performing tests of significance for comparative analysis.

The possible sources of error in the classification of a segment for estimating its wheat or small-grains proportion are outlined in figure 3. Most of these factors are causative and are called second-order error sources. Some of these sources contribute mainly to the variation in the segment proportion estimate, some sources introduce bias, and others are influential in both respects. Brief descriptions of a few useful investigations are presented in the following paragraphs. For the actual studies made and the scope of second-order error source evaluations, see reference 6 and Chhikara and Feiveson's paper.

Segment-level accuracy investigations.—Accuracy of ground-observed proportions obtained by dot counting, of CAMS proportion estimation, and of crop calendars comprise the segment-level accuracy investigations.

Accuracy of ground-observed proportions obtained by dot counting: Two methods are used to determine the true wheat and small-grains proportions for blind sites; namely, dot counting and computer digitization. The first method gives wheat and small-grains proportions by evaluating the ground-observed labels of a subsample of 400 (or more) random dots from the 9- by 11-kilometer (5- by 6-nautical mile) sample segment. This method produces only approximate results. In the second method, the Bendix 100 system and computer programs SPATL and MLTCRP are used to generate the wall-to-wall digitized ground-observed proportions for wheat and small grains. In this task, the dot count proportions are compared to the wall-to-wall digitized proportions to determine the accuracy of proportions obtained from the first method. The purpose of this task is to validate conclusions about bias due to classification that may have been made before the

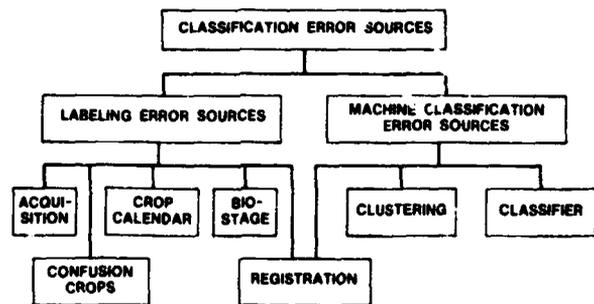


FIGURE 3.—Sources of error/variation in the classification error for estimating segment wheat and small-grains proportions.

wall-to-wall ground-observed proportions obtained using the second method were available.

Accuracy of CAMS proportion estimation: For the blind sites, the estimation error ($\hat{P} - P$), where P is the ground-observed wheat or small-grains proportion and \hat{P} is the estimate of P made by CAMS, is determined. Two types of studies are performed on these errors: (1) analysis of variance/covariance is done on the absolute errors to investigate the effect of various factors likely to influence the classification performance (e.g., AI, wheat biostage, segment location, wheat proportion) and (2) a linear regression is performed for the estimation errors on the ground-observed proportions, biostage, field size, crop type, etc., to explain the variability in segment wheat proportion estimation errors.

Accuracy of crop calendar: A major reference used by analyst-interpreters in their classification procedures is the nominal (mean historical) crop calendar and the adjustable crop calendar (ACC). Since the ACC provides the latest reference information on the stage of development of wheat in an area being classified and estimated, it is necessary to determine the accuracy of this reference information.

The basic data set for these evaluations is the growth-stage data acquired by USDA/ASCS personnel from ITS's in the United States. These growth-stage data are acquired in periodic ground observations of the ITS's over the crop reporting districts (CRD's).

Plots are made of the ACC outputs (for the ITS's), the mean of the ground observations of wheat growth stages, and the nominal crop calendar. Confidence interval estimates are made on the basis of the distribution of the ITS ground-truth observations, and whether the ACC results fall within these limits is determined. The relationship of the crop calendar information to known episodic events of the current year, such as drought, is also investigated by the accuracy assessment group.

Pixel-level comparison investigations.—In the pixel-level investigations, the ground-truth data are compared with the AI-labeled pixel data and with the cluster and classification maps produced by CAMS. This procedure also enables a determination of the actual composition (in terms of ground-truth classes) of each pixel, of each cluster on the cluster map, and of each class on the class map.

Blind-site data: Blind-site ground observations and CAMS data are compared at the pixel level to

evaluate the omission and commission errors and then to develop a method of assessing the labeling, clustering, and classification performance in a quantitative manner. A ground-truth data processing procedure is used to produce a tape on which the ground observations are presented as an image, similar to the Landsat imagery and to the cluster and classification maps generated by CAMS. Details are given in the paper entitled "Accuracy Assessment System and Operation" by Pitts et al.

Each subclass in the ground-truth data has its own assigned gray-scale level on the ground-truth tape. The subclasses used are shown in table I. The image on the ground-truth tape is registered to the corresponding Landsat image. However, the data on the ground-truth tape are at a finer resolution. There are six subpixels on the ground-truth image for each pixel on the Landsat image.

Analyst-interpreter dot (pixel) labeling accuracy: To investigate dot (pixel) labeling, the composition of each dot is obtained first. This procedure consists of determining the representation of the various ground-truth classes (table I) among the six subpixels for each dot on the ground-truth tape. Each dot is then given the label of the subclass having the largest representation among the six subpixels corresponding to that dot on the ground-truth tape.

Each dot is also given a class name (as distinguished from its subclass name). The classes are those used by the analyst to label the dots: spring grains (SG), spring wheat (SW), winter grains (WG), winter wheat (WW), grains (G), wheat (W), other (O), and a class denoted "X" which consists of dots that fell on clouds or cloud shadows and therefore were unidentifiable.

Dot labeling accuracy is studied by estimating two confusion matrices—one for classes and one for subclasses. The class confusion matrix consists of errors of omission and commission by the AI and indicates the degree of accuracy of the AI labeling with respect to the eight classes mentioned. The subclass confusion matrix of AI omission and commission errors describes AI skill in labeling pixels with respect to the subclasses listed in table I.

Labeling accuracy depends on several factors (fig. 3). The effect of these factors is evaluated whenever feasible and/or critical. Finally, a study is made to determine whether the probability of a dot being correctly labeled is higher if the analyst label agrees with the classifier label for that dot.

Clustering accuracy: Three aspects of clustering

TABLE 1.—Subclasses Used in Accuracy Assessment and the Corresponding Gray-Scale Levels on the Ground-Truth Tape

Subclass	Gray-scale level
Fields 1 to 80	1 to 80
Alfalfa	90
Beans	91
Corn	92
Safflower	93
Sunflower	94
Sudan grass	95
Sorghum	96
Soybeans	97
Sugar beets	98
Winter wheat	99
Spring wheat	100
Barley	101
Rye	102
Flax	103
Oats	104
Grass	105
Hay	106
Pasture	107
Trees	108
Same as 90 to 108 except:	
Harvested	115 to 133
Abandoned	140 to 158
Strip fallow	165 to 183
Strip fallow, harvested	190 to 208
Strip fallow, abandoned	215 to 233
Water	240
Homestead	250
Idle cropland, stubble	251
Idle cropland, cover crop	252
Idle cropland, residue	253
Idle cropland, fallow	254

are studied: cluster composition and purity, cluster labeling accuracy, and the cluster confusion matrix. The data used are blind-site cluster maps.

Cluster composition is the set of percentages of subpixels in a given cluster that belong to each of the major classes. Class is determined by comparing the cluster map with the image on the ground-truth tape. The major classes are SG, WG, G, O, and Y, where Y consists of both designated other (DO) and designated unidentifiable (DU) areas. The "purity" of a cluster is the percentage of the total number of subpixels in the cluster that belongs to the class with the largest representation. The composition and purity of clusters are of interest since they indicate the capability of the clustering algorithm to separate the classes into relatively "pure" clusters. These quantities are studied as a function of segment, stage, and acquisition history.

Cluster labeling accuracy is studied first by assigning each cluster the name of the class having the largest representation of subpixels. The cluster is assumed correctly labeled if the label given by the labeling logic corresponds to this name. In the case of nearest neighbor labeling logic, an incorrect label may result from AI mislabeling of the dot used to label the cluster or from poor performance by the labeling logic. If the identity of the dots that were used to label each cluster can be determined, these two sources of error are studied separately. Cluster labeling accuracy is studied as a function of cluster purity, segment, state, and acquisition history.

Two confusion matrices are estimated for clusters—a class confusion matrix and a subclass confusion matrix. The clustering confusion matrices are evaluated as a function of segment, region, and acquisition history.

Classification accuracy: Classification performance is studied by estimating the classification confusion matrices for both classes and subclasses. The classes are SG, WG, G, O, X, and T, where T indicates pixels which have been thresholded by the classifier; subclasses are the same as for dot labeling and clustering.

An important investigation is made to determine the effect of crop height and ground cover on classification accuracy. In this study, crop height and ground cover data acquired every 18 days for 15 selected wheat fields in each blind site are used. The probability of correct classification is computed for each of these fields and is plotted as a function of crop height, Robertson biostage, ground cover, and "green number." Means and other relevant statistics are calculated at the segment, state, and regional levels.

Intensive test site data: The purpose of ITS data is to determine the causes of labeling and classification errors that cannot be determined from blind-site data. For example, one use for these data is to examine the relationship, if any, between the labeling and classification accuracy and errors in the adjustable crop calendar. Studies of accuracy are made for both wheat and small grains if CAMS estimates of both are available.

To evaluate the dot labeling accuracy, CAMS personnel analyze the imagery and attempt to determine the dot labels by photointerpretation. The labeling accuracy (omission and commission) is determined by comparison with ground-truth labels for each classification. The "18-day observation" fields in the ITS's are used to determine the crop growth stage

and other agromet activities and hence the cause of mislabeling. Whenever a dot falls on an 18-day observation field, CAMS investigates the ancillary data from the 18-day observations such as crop height, ground cover, stand quality, planting date, irrigated or dry land, farming practice, and growth stage for correlation with the Landsat data. Each acquisition is processed so that the accuracies, as a function of growth stage, may be determined; CAMS performance analysis then determines the following for the ITS labeling.

1. Number of ground-truth wheat and other dots
2. Number and percentage of incorrect labels, for wheat, other, and total
3. Cause of error for each dot
 - a. Necessary acquisitions missing
 - b. Poor stands
 - c. Late planting, emergence, or development
 - d. Strip fields
 - e. Analyst error
 - f. Confusion crops
 - g. Border/edge pixels
 - h. Unknown cause

SUMMARY

The methodology described in this paper for assessing the accuracy of LACIE estimates illustrates the detailed and extensive evaluations performed during the experiment. This methodology was necessary for validation of the implemented wheat production forecasting technology. As intended, it has allowed the identification and isolation of key problems in wheat area and yield estimation, some of which have been corrected and some of which remain to be resolved.

The major unresolved problem in accuracy assessment is that of precisely estimating the bias of the LACIE production estimator. This problem will continue to be an issue in the United States and more so in foreign countries. In the United States, reliable ground observations, like those obtained over blind sites and intensive test sites during LACIE, can be obtained for further assessment of the bias in the crop area estimation technology. In the future, if reliable yield information at the field level can also be obtained together with the crop acreage information, an improved assessment of the bias in the crop production forecasting technology can be achieved.

REFERENCES

1. LACIE Crop Assessment Subsystem (CAS) Requirements. Vol. IV, LACIE-C00200, Rev. B, NASA Johnson Space Center, July 1976.
2. LACIE Yield Estimation Subsystem (YES) Requirements. Vol. III, LACIE-C00200, Rev. A, NASA Johnson Space Center, Sept. 1976.
3. LACIE Crop Assessment Subsystem Software Requirements Document. NASA Johnson Space Center, Nov. 25, 1975.
4. LACIE Accuracy Assessment Report - Phase IA, November/December 1974. NASA Johnson Space Center, June 1976.
5. LACIE Phase I CAMS Rework Experiment. Technical Memorandum, LEC-8655, Lockheed Electronics Co., Houston, Texas, June 1976.
6. LACIE Phase I and Phase II Accuracy Assessment Final Report. LACIE-00450, NASA Johnson Space Center, Feb. 1978.
7. LACIE Third Interim Phase III Accuracy Assessment Report. LACIE-00467, NASA Johnson Space Center, Jan. 1978.

Manual Interpretation of Landsat Data

C. M. Hay^a

INTRODUCTION

The LACIE analyst is required to estimate the proportion of small grains in a given sampling unit. These sampling units are 5- by 6-nautical-mile areas located in accordance with a statistical sampling design. The estimation process requires that the analyst interpret a 1-percent sample of the segment area using both Landsat multispectral scanner (MSS) imagery data and ancillary data. Ancillary data include crop calendar summaries, cropping practice reports, meteorological data, and other pertinent regional data.

This paper discusses the manual interpretation process that has been developed within LACIE. Details regarding the role of interpretation in the machine processing approach are discussed in the paper by Heydorn et al. entitled "Classification and Mensuration of LACIE Segments," and the implementation of this approach is discussed in the paper by Abotteen and Bizzell entitled "The Classification and Mensuration Subsystem."

As will be pointed out subsequently, the methodology for the interpretation of Landsat and ancillary data for inventory purposes is in a state of heuristic development that has continued through the 3 years of LACIE. With any heuristic development, concepts are formed, methods are developed, resulting problem areas are analyzed, and new methods are proposed. Such cycles have certainly been experienced within LACIE. This paper, however, will not attempt to document the progression of thought that occurred throughout these 3 years but rather will discuss the fundamental concepts that evolved. Still, no claim is made that these concepts are entirely satisfactory, and, in fact, problems with the methods will be addressed at the end of the paper.

HISTORY OF MANUAL INTERPRETATION IN LACIE

LACIE Phases I and II

Throughout LACIE Phases I and II (1975 and 1976), the analyst performed two main tasks. The first task was to outline representative areas (fields) for all spectral classes within a segment on the basis of their appearance on the Landsat image product. The spectral statistics generated from these areas were used as training for maximum likelihood classification. The second task was to label the crop type (wheat/nonwheat) within the selected training areas. This process of first selecting representative training areas and then labeling the crop type within the areas comprised what is called the "Fields Procedure." An analyst took approximately 12 hours to process a segment by the Fields Procedure and evaluate and possibly rework the results. Half of this time was spent selecting and recording training areas; only one-eighth of the time was spent actually labeling the areas as to crop type.

LACIE Phase III

By contrast with the Phase I and II procedure, in LACIE Phase III (1977), a procedure was developed and implemented which incorporated clustering for spectral class definition and training statistics generation. This procedure is called Procedure 1. The analyst was freed from the time-consuming task of spectral class definition and could concentrate solely on crop-type labeling. A new within-segment sampling strategy involving randomly selected sample dots (pixels) was another innovation of Procedure 1. Because the analyst now has only to label sample dots as to crop type, his segment processing time is reduced to approximately 3 to 4 hours. In Phase III, therefore, the analyst had only one main analysis task—crop-type identification.

^aUniversity of California at Berkeley.

MANUAL INTERPRETATION PROCESS

Perspective on Landsat Data

It must be remembered that Landsat data is a new type of data which had not been used extensively before LACIE. Landsat's uniqueness, aside from its low resolution, small scale, and synoptic coverage, is more significantly due to the increased spectral resolution (compared to conventional photographic systems), its regular periodic coverage, and the digital format of the data. It is important to recognize the differences between Landsat data and conventional photographic data. Although manual image interpretation procedures that were developed with and for conventional photographic imagery still have relevance to image-formatted Landsat data, they need to be modified and restated within the Landsat context. Furthermore, Landsat data provide additional information not obtainable from photographic data. This additional information is a function of the *digital* temporal-spectral response data from relatively narrow-band (compared to photographic data) multispectral sensors.

The Analysis Process: The "Art of Probabilities"

Crop (feature) identification from Landsat (or any other remotely sensed) data and ancillary data is basically the "art of dealing in probabilities" (ref. 1). An analyst must (1) collect information from the Landsat data about the characteristics of a feature, (2) factor in additional evidence from a priori knowledge and ancillary data, (3) judge the importance and relevance of all the evidence, (4) formulate several reasonable working hypotheses, (5) test these hypotheses against the evidence, (6) select the most probable conclusion, and (7) judge the degree of probable correctness of this conclusion. In LACIE, the most probable conclusion is recorded as a crop-type label for a given pixel.

In simple terms, the interpretation process (called labeling in LACIE) consists of two main components: (1) feature detection and physical characteristics determination and (2) feature evaluation including identification and condition assessment. Although these processes may occur simultaneously and iteratively, they can be treated separately for the purpose of explanation. *Feature detection* is the action of discriminating a unique feature (a field in the LACIE case) based on spectral, spatial, and temporal

characteristics observable within Landsat multitemporal-spectral data. *Feature characteristics evaluation* is the process of assessing the available data by analytical means and then synthesizing the pertinent data for the purpose of concluding the feature's identity and condition state. *Feature identification* is the action of assigning a name (e.g., wheat, nonwheat) to the detected feature. *Feature condition assessment* is the more refined identification of a feature such that some quality or state (e.g., late developing, poor stand, harvested) is ascribed to the feature. Correct feature identification cannot proceed properly unless feature detection has occurred. Feature detection, however, does not ensure feature identification. Thus, errors in labeling may result from either (1) failure to detect a feature of interest or (2) failure to identify correctly a detected feature.

Feature Detection and Characteristics Determination

Within LACIE, the *features* that an analyst wishes to detect are cropped fields. The *feature characteristics* that an analyst must determine are (1) the size and shape, the type of boundary elements, and the spatial relationships of similarly and dissimilarly appearing features; (2) the temporal-spectral patterns throughout the growing season for a given region; and (3) the magnitudes of the actual spectral values within specific time periods corresponding to given crop-type biostages (spectral-temporal characteristics). Of these three characteristics, the second is the most important to the analyst for the detection and identification of wheat or any other crop. The other two characteristics are necessary when significant overlap exists between wheat and confusion crops, when key acquisitions are missing or of poor quality, and/or when any ambiguity exists in the data. Obviously, the probability of correctly identifying a crop within a given field will be low if a spectral response indicating vegetation canopy cannot be detected during any one portion of the growing season or during particularly significant vegetation biophases specific to given crop types.

Feature Characteristics Evaluation for Crop Identification

Although site-specific Landsat data enable an analyst to detect a feature and determine its physical

characteristics, ancillary data and a priori knowledge from outside the Landsat data are necessary for the analyst to identify and label a detected feature. That is, nowhere will one find the word "wheatfield" written across a field as observed on Landsat data. A priori knowledge and ancillary data supply information about what crops are grown within a region, the rate and timing of canopy development for specific crops, cropping and cultivation practices employed within a region or specific to a given crop type, the characteristic appearances of given features on Landsat data, etc.

A priori knowledge.—An analyst gains a priori knowledge from past experiences, educational background, specialized training, and specially prepared interpretation keys. For example, many LACIE analysts were geographers who had studied as part of their formal education (1) interrelationships between land use and cultural patterns and (2) physical environmental parameters such as climate and soils. Since some analysts grew up on a farm, they had a firsthand understanding of agricultural cropping practices. Still others had many years of photointerpretation experience upon which to draw. With such varied backgrounds, there was a need to bring all the analysts to some minimum level of common experience. Thus, before the start of LACIE, the analysts attended an intensive specialized training course in Landsat data analysis, wheat physiological development, and regional cropping practices relative to small grains.

Ancillary data.—In addition to the a priori base information, additional information is required for crop-type identification. This additional information is by convention called ancillary data in that it consists of data different from and outside the site- and date-specific Landsat spectral data.

Types of ancillary data which have been recognized as being necessary for crop-type identification are (1) crop calendar information including average-normal and year-specific data; (2) historical crop proportions for several recent years; (3) regional cropping practice information such as crop rotation, cultivation practices, and irrigation practices; and (4) occurrence of meteorological events affecting crop development and/or crop spectral response. Each of these data types should contain quantitative descriptions of mean normal conditions, as well as the variability that can be expected about the normal. The variability data should include temporal variability (year to year) as well as spatial variability.

Crop calendar data.—Crop calendar data compose

the one *most important* type of ancillary data for crop identification from Landsat data and are important to all aspects of the crop identification task. First, before any imagery is acquired, analysis of crop calendar data will determine the time periods during which data should be collected for the particular crop or crops of interest. Second, crop calendar data in conjunction with historical crop proportion information enable an analyst to predict possible confusion crops and thus assess the need for additional confusion crop separability information. Third, crop calendar data serve to set initial expectations of temporal-spectral response patterns for the major crops. Crop calendar data contain information about the time periods during an average year when significant stages in the cultivation and development of a given crop can be expected to occur. Planting and harvesting dates are often given, and information about the timing of other intermediate phases such as seedbed preparation, crop emergence, heading, or flowering is frequently presented.

Crop calendar data available to LACIE analysts include average-normal year data for all major crops within an area and a year-specific adjusted wheat crop calendar. Figure 1 is an example of an average-normal year crop calendar. In this example, the percentage of the given area undergoing an indicated development stage on specified dates is indicated for each major crop. Figure 2 is an example of the year-specific adjusted wheat crop calendar information.

Relying on past experience and a limited number of spectral-response-to-ground-data correlations, the analyst must translate ground crop calendar data into expected spectral responses and image characteristics. For example, from field experience, the analyst knows that a healthy small-grains crop in the headed biostage has a 90- to 100-percent canopy cover. The major part of the spectral response from such a field will be from the vegetation canopy; there will be very little response from underlying soil or surface litter. Therefore, the analyst expects high infrared reflectance and low red reflectance (because of chlorophyll absorption) from this field. On a color-infrared (CIR) image, such a field should display an intense red color. On the other hand, for earlier stages of development, such as emergence, the soil background reflectance would contribute more to the overall spectral response from the field. Thus, a less intense red color would appear on the CIR image.

Historical agricultural statistics.—A significant input for crop identification is the most recent year's crop acreage statistics for the specific region. From

**CROP CALENDARS PLOTTED JAN 1, 1976;
PERCENT OF AREA IN DEVELOPMENT STAGE BY SPECIFIED DATE FOR
NORTH DAKOTA, 1952 - 1984; AVERAGE CROP CALENDARS**

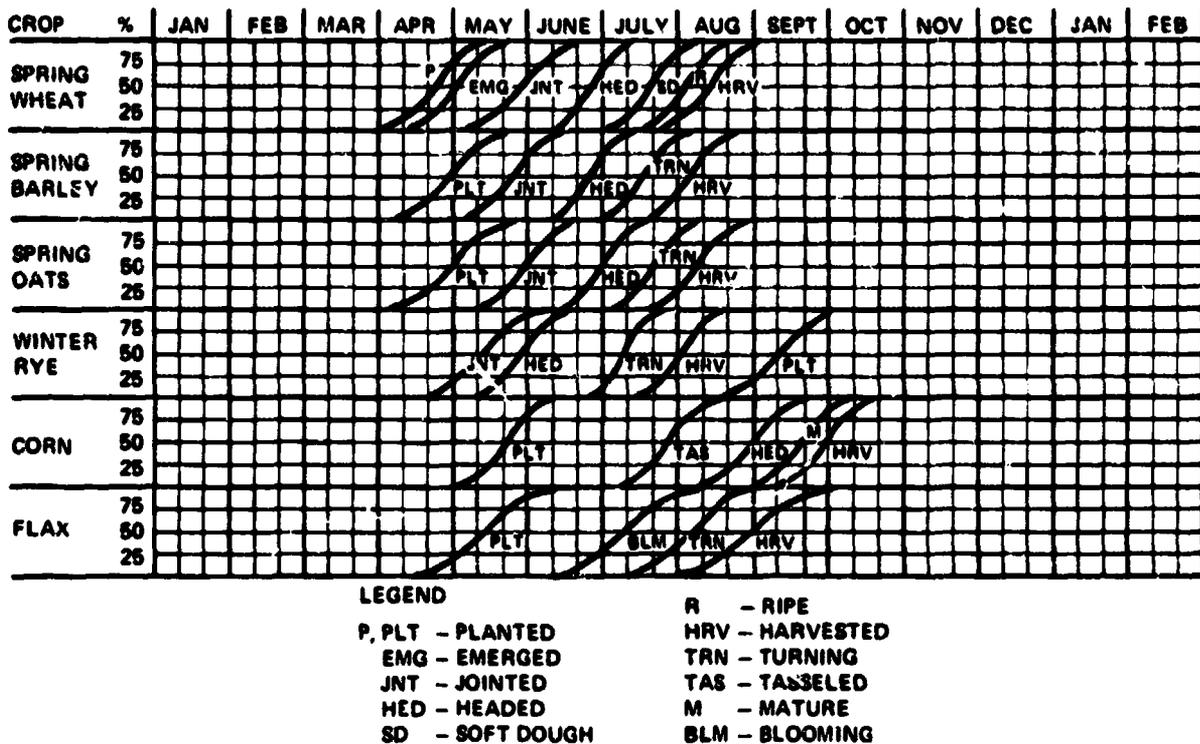


FIGURE 1.—Example of average-normal crop calendar for North Dakota.

these data, initial probabilities of occurrence of the principal crop types and possible confusion crops can be set.

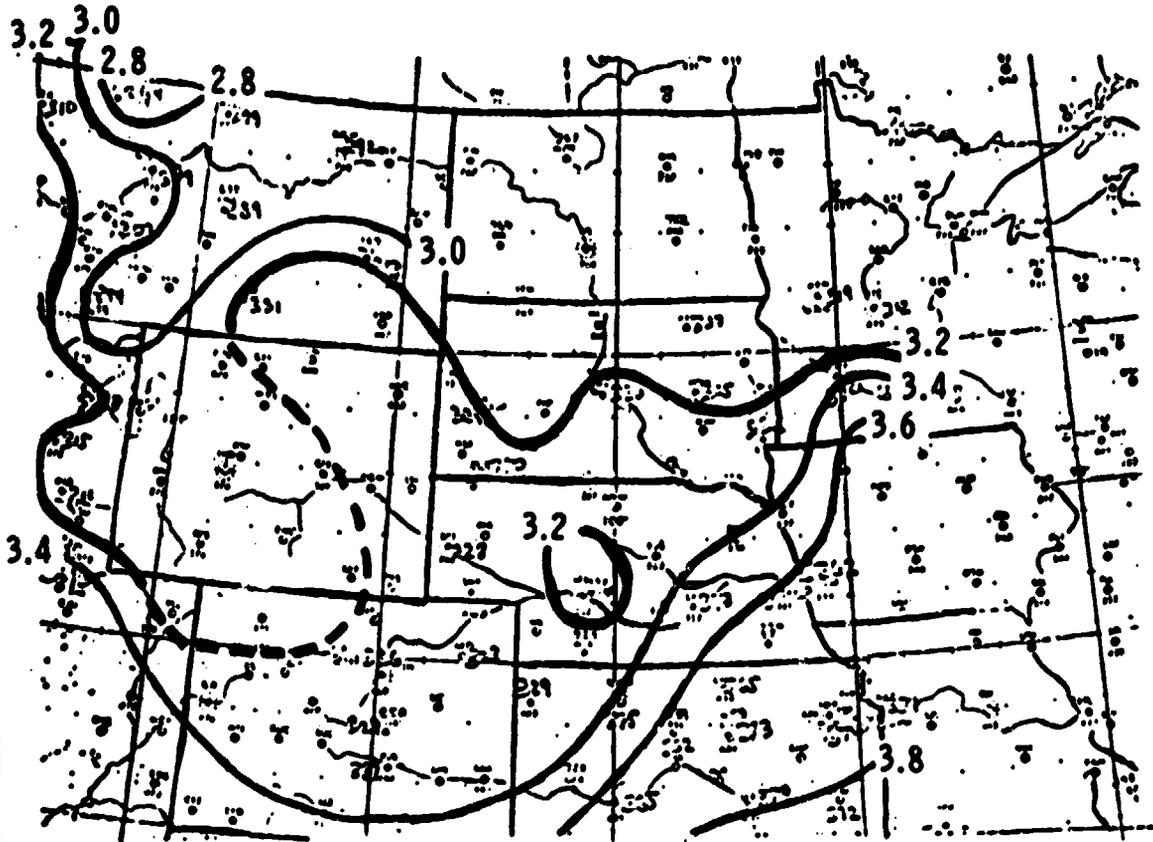
The historical agricultural statistics are an indication of the expected percentage of possible confusion crops (confusion crops predicted from calendar data) present within an area relative to the primary crop or crops of interest. In western Kansas, for example, other small grains (namely, barley, rye, and oats) are always possible confusion crops for wheat. Reference to published historical agricultural statistics, however, indicates that the combined percentage of cropland occupied by these three crops was approximately 1 percent. By comparing these percentages with the approximately 45 percent of cropland devoted to wheat, the analyst can have confidence that his wheat identifications will, in general, contain a very low (less than 2 percent) commission error (identifying these other small-grains confusion crops as wheat). Figure 3 is an example of the

agricultural historical statistics provided to LACIE analysts.

Cropping practice and environmental relationships.— Relationships between the physical environment and the presence of certain crop types can often be used effectively in crop identification. Physical parameters that exert significant influence on what crops will be planted within a region are climate, soil type, and availability of irrigation water. For example, on the dry-farmed sandy-loam and loamy-fine-sand soils of western Kansas, sorghum is planted more often than wheat. Sorghum, which is more tolerant of water stress than wheat, grows well on these highly permeable, low-water-holding-capacity soils. However, if irrigation is available and if there is supporting evidence that wheat is irrigated within this region, the possible proportion of wheat within these soil types must be expected to be equivalent to the proportions found on loam soils of the area. Irrigation also leads to more crop diversity within an

CROP CALENDAR ADJUSTMENT

May 16, 1976



These data are based on the Robertson Triquadric model to compute biometeorological time. Data output is by CCEA (NOAA). The biological stages of wheat are divided into seven development stages:

1. Planting
2. Emergence
3. Jointing
4. Heading
5. Soft Dough (turning greenish-yellow to yellow)
6. Hard Dough (ripe)
7. Harvest

The model is driven by daily maximum and minimum temperatures and run for selected met stations in the LACIE countries. Isolines are drawn to connect geographical points that show equal development stages. These isolines are drawn to the nearest .2 of a development stage.

FIGURE 2.—Example of year-specific wheat crop calendar adjustment data.

area than will be found in dry-farmed areas. Thus, the possibility of confusion crops is greater than that in an adjacent, less diversified area. An example of one type of cropping practice information available

to LACIE analysts is shown in figure 4.

Full-frame Landsat imagery.—The usefulness of the ancillary data described previously can be significantly enhanced when the data are analyzed in con-

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

County/State CASS CO., N. DAKOTA

Land Use	Acreage (Most Recent Year)	Percent of Total County Area
A. Total County Area	1,119,296	
B. Total Cropland	947,051	84.6
1. Cropland Harvested	803,104	71.8
2. Cropland Pastured	34,123	3.0
3. All Other Cropland	109,824	9.8
C. Woodlands, Woodland Pasture	1,184	0.1
D. All Other Land	82,157	7.3
E. Average Field Size (Acres)	_____	
F. Wheat (% Total of County Area)	_____	

Conservation Practices (Most Recent Year)

A. Irrigated Land (% County Area)	_____
B. Contour (Acres)	_____ ; Strip (Acres) _____ ; Terrace (Acres) _____

Crops and Agricultural Lands (Absolute Acreage and % of Total County Area)

Crops (Include All)	Most Recent Year 75			Second Most Recent Year 74		
	Acreage	Percent	Percent Irrigation	Acreage	Percent	Percent Irrigation
Wheat, All	460,700	41.2		447,800	40.0	
Durum	65,100	5.8		45,300	4.0	
O.S. Wheat	394,500	35.2		401,000	35.8	
W. Wheat	1,100	0.1		1,500	0.1	
Barley	178,600	16.0		129,100	11.5	
Rye	3,200	0.3		2,900	0.3	
Oats	19,000	1.7		20,200	1.8	
Flaxseed	18,100	1.6		20,200	1.8	
Soybeans	87,500	7.8		89,300	8.0	
Sugarbeets	17,000	1.5		18,900	1.7	
Sunflowers	71,800	6.4		83,500	7.5	
All Hay	33,100	3.0		33,600	3.0	
Alfalfa Hay	22,100	2.0		21,600	1.9	

FIGURE 3.—Example of historical agricultural statistics.

cert with full-frame Landsat imagery. One factor that limits the usefulness of historical agricultural statistical data and environmental relationship data is the lack of spatial variability information. Two questions arise: how do the crop proportions reported vary throughout the reporting unit area, and what signifi-

cance does this variability have for crop-type identification within the area? Full-frame Landsat imagery (fig. 5) clearly displays the distribution of major land use types (cropland, rangeland, etc.), as well as minor land use differences that may affect local (within-segment) crop proportion mixes.

Identification # US-20

Segment # 1518
Strata Overlay _____
File Location _____

UNIVERSAL STRATA DESCRIPTORS

Country: United States

State, CRD: North Dakota..CRD's 3,6,9
Minnesota. CRD's 1,4

Full Frame Landsat Imagery Numbers:

ND: 1398, 1399, 1462, 1464, 1473, 1482, 1565, 1584
1586, 1618, 1619, 1621, 1624, 1641, 1642, 1645
MN: 1514, 1515, 1518, 1519, 1521, 18 1841

AGRICULTURAL LAND USE - General Description

Important cash crops are spring wheat, potatoes, sugar beets, and soybeans. Legume seed are widely important in the northeast part of the stratum. Fallowing is practiced mainly for weed control and for accumulating nitrogen. Sweet clover (green manure) is grown widely for soil improvement.

Ag/Non-Ag Overlay _____ File Location _____
DSF _____ Texture (High/Low) _____
Overlay _____ File Location _____

Field Size:

CLIMATE - General Description

The climate of this stratum is continental. The summer temperatures are generally comfortable with very few days of hot and humid weather. Nights, with few exceptions, are comfortably cool. The winter months are cold and dry, with maximum temperatures rising above freezing only on an average of 6 days each month, and nighttime lows dropping below zero approximately half of the time. (See representative stations.)

Precipitation is the most important climatic factor in the area. The Red River Valley lies in an area where lighter amounts fall to the west and heavier amounts to the east. Seventy-five percent (75%) of the precipitation occurs during the growing season (April to September) and is often accompanied by electrical storms and heavy falls in a short time. Winter precipitation is light and indicates that heavy snowfall is the exception rather than the rule. The first light snow in the fall occasionally falls in September, but usually very little, if any, occurs until October or November. The latest fall is generally in April. (See representative stations.)

With the flat terrain, surface friction has little effect on the wind in the area, and this fact has led to the legendary Dakota blizzards. Strong winds with even light snowfall cause much drifting and blowing snow, reducing visibilities to near zero. Fortunately, these conditions occur only several times during the winter months.

FIGURE 4.—Example of cropping practice information.

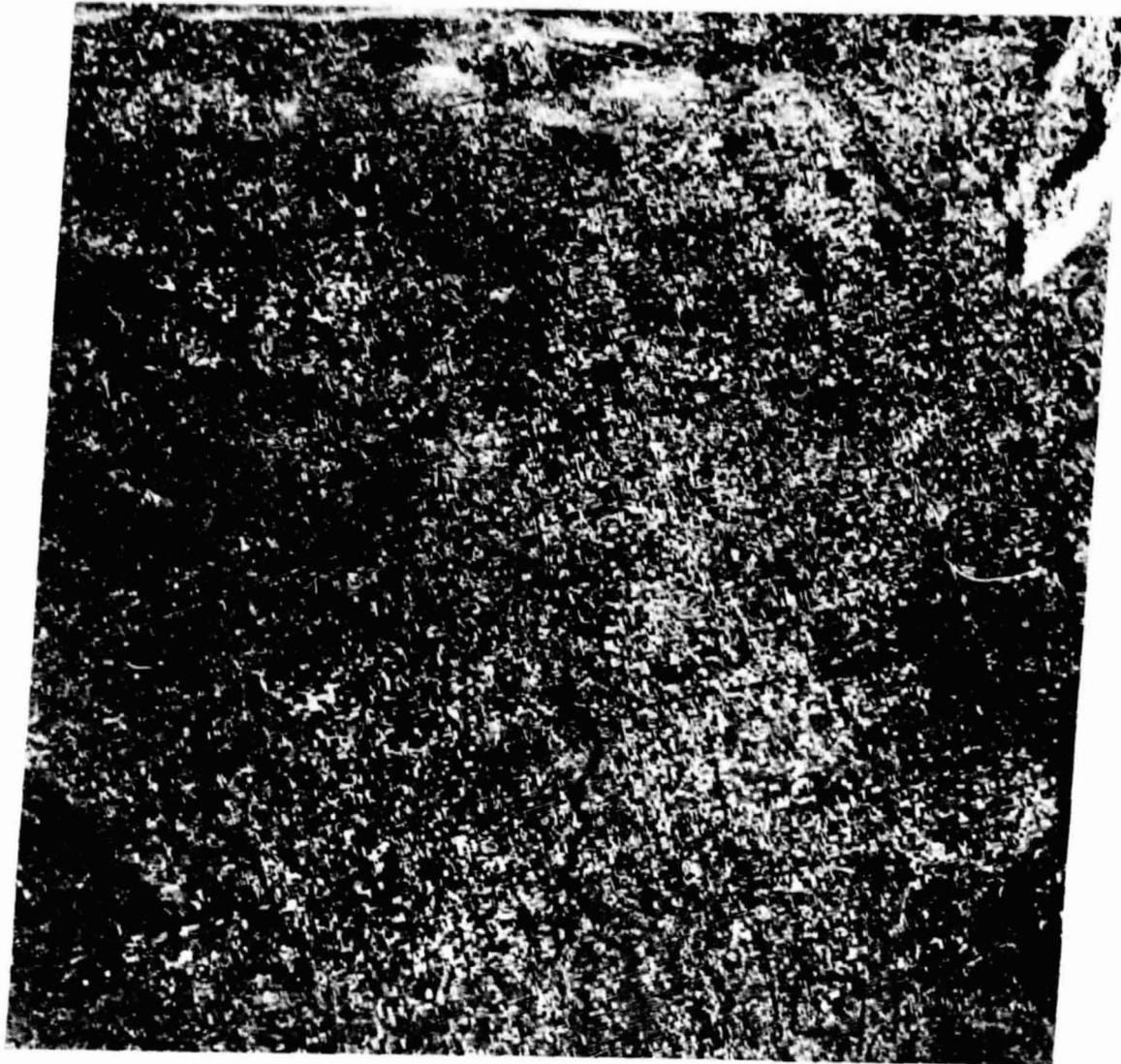


FIGURE 5.—Full-frame Landsat data for part of eastern North Dakota. The small outlined area is the segment discussed in the multitemporal data analysis example (Aug. 11, 1975; Landsat frame E-2201-6421).

Another situation in which full-frame Landsat imagery is useful is where cultivated land is thinly interspersed among rangeland and other wildland areas. During certain wheat biophases, it is difficult to distinguish some wheatfields from native grassland range. When only a small area (such as a sample segment) is available for viewing and only less-than-optimum temporal-spectral data are available, a number of misidentifications may occur. Upon reference to full-frame imagery for the area, however, the analyst can obtain a better appreciation for the distribution of grassland range within the sample segment. In addition, the evaluation of at-

mospheric effects in segment-sized areas is difficult, and full-frame data have been found helpful here, also.

Multitemporal Data Analysis for Crop Identification

As stated previously in the section entitled "Feature Detection and Characteristics Determination," the feature's temporal-spectral pattern throughout the growing season is the most important characteristic for detection and identification of crop

type. The procedure used to evaluate this particular feature characteristic is called multitemporal data analysis. Multitemporal data analysis is based on the assumption that, within a specific region, a given crop type or group of crops has a temporal-spectral development pattern that is relatively unique. Therefore, by monitoring the spectral changes within a field throughout the growing season, an analyst can identify with a high degree of accuracy the crop or crop group grown within a field.

To demonstrate the principle of multitemporal data analysis for crop-type identification, a "walk-through" of the analysis of a segment in southeastern North Dakota will be presented. The previously presented ancillary data examples will be referred to since they are applicable to this segment.

First, from a priori knowledge and ancillary data, a conceptual *spectral* crop calendar for the area must be developed. The spectral crop calendar describes the expected Landsat temporal-spectral response pattern for various crop types found within the area of interest. Figure 6 is a graphic illustration of one way of portraying such a spectral crop calendar for North Dakota. In this particular spectral crop calendar, the ratio of 2 times MSS band 7 divided by MSS band 5 was used as a green vegetation indicator to portray the Landsat spectral-temporal patterns corresponding to crop canopy and phenological development through time. Spectral crop calendars are not currently available to LACIE analysts directly. However, it must be realized that every analyst consciously or unconsciously carries the concept of a spectral crop calendar within him. Therefore, to facilitate the discussion of multitemporal data analysis, part of the spectral crop calendar for North Dakota has been presented in concrete form.

From figure 6, one can see that the maximum and minimum Landsat vegetation indicator values occur at different times for each of the four crop types shown. It is these temporal-spectral differences that will enable the identification of crop types within given fields. Notice that the temporal-spectral differences between some crop groups such as small grains (wheat and barley) and large grains (corn) are quite pronounced. Therefore, there is little risk of confusing these two crop groups within this region, given moderately adequate timing of Landsat acquisitions. Thus, small grains (wheat, barley, oats, rye) as a group are expected to be fairly consistently identifiable within North Dakota.

The temporal-spectral differences between closely related crops, however, such as between small-grains

crops (wheat versus barley), are seen to be more subtle. Precise timing of Landsat acquisitions will be critical to the separation of these two crops.

Some Landsat imagery will now be examined to see how this information can be used for crop-type identification. Figure 7 is a multitemporal sequence of Landsat imagery. The image product displayed here is known within LACIE as Product 1. Relative feature temporal-spectral response characteristics are determinable from this image product. High near-infrared response coupled with low visual-red response and relatively low to medium visual-green response, as is typical of green or actively metabolizing vegetation, appears red on the CIR composite. Relatively equal response in all three bands, as may be given by bare soil fields, appears as various shades of gray, depending on overall total reflectance.

From past experience and the ancillary data, the image characteristics for wheat are expected to be as follows. From the emergence to the jointing biostage, the wheat crop canopy cover increases from 0 to 100 percent. The Landsat vegetation detection threshold appears to be approximately 20-percent canopy cover. Thus, after a 20-percent canopy cover has been achieved on through jointing (May 19 to June 10, 1976), the wheat is expected to appear on Product 1 as various shades of pink or red, depending on the amount of canopy cover. From jointing through heading (June 10 to 30), the crop is actively metabolizing and should appear bright red. From the heading through the soft-dough biostage (July 1 to 26), prior to turning golden, the crop is increasing its dry matter percentage, and wheatfields tend to decrease their near-infrared reflectance slightly. This slight decrease in infrared reflectance makes wheat in this stage appear as a darker red (brick or blood red) than in the previous stage. As turning proceeds (July 26), infrared reflectance drops drastically and red reflectance increases slightly. Wheatfields in the turning stage appear dull yellow, brown, or brownish green on the CIR image product. After the wheatfield has been harvested (August 7), it will appear whitish or yellowish white if stubble remains in the field. The harvested field may appear in various shades of gray if disking or plowing has occurred. The more the field has been plowed, the less stubble there will be on the surface of the soil and the darker gray the field will appear. All small grains go through roughly the same color sequence on the CIR image product.

Before specific field labeling is attempted, the expectations for this segment must first be established

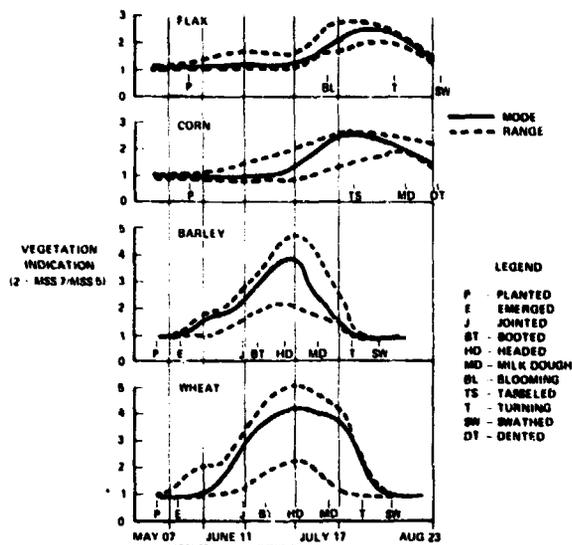


FIGURE 6.—1976 spectral crop calendar for flax, corn, spring barley, and spring wheat in Cass County, North Dakota.

by referring to the available ancillary data. From the nominal North Dakota crop calendar (fig. 1), spring wheat, spring barley, and spring oats have very closely timed coincident biostages. Significant confusion among these small grains can be expected if they all occur within the same area. Remembering that there is an indication that barley matures and turns golden a little sooner than wheat, one can look for an acquisition around the critical time period and, if it is available, attempt to separate wheat and barley on the basis of the expected subtle temporal-spectral differences. Also from the crop calendar data, it is observed that another crop, flax, has some overlap with the small grains. Although the calendar indicates later planting and turning for flax as compared to the small grains, there may be confusion between small grains and flax if acquisitions are missing. No other summer season crops (corn, potatoes, sunflowers, beans, etc.), however, are expected to be significantly confused with wheat and the other small grains in this area.

The historical agricultural statistics (fig. 3) indicate that, of the possible wheat confusion crops identified from crop calendar analysis, barley is the only crop type that occurs in significant proportions along with the spring wheat in this area. Analysis of the segment position on full-frame imagery (fig. 5) does not indicate any conditions that would lead to adjustment of expectations from the countywide statistics.

In figure 7, several fields of various crop types have been interpreted and outlined. Crop calendar

information for the first acquisition (May 7, 1976) indicates that the spring small grains, wheat and barley, have just been planted. Spring small-grains fields, therefore, should appear as bare soil (dark gray to black). Winter wheat is in the tillering (pre-jointing) stage according to the crop calendar, and significant vegetation canopy is expected (definite red color on image). Corn, flax, and other summer crops are not yet planted and therefore are still seen as bare soil.

On the second acquisition (May 25), according to crop calendar information, spring small grains are emerged to tillering. They should begin to show some indication of red (dark purple to pinkish to red) on the Product 1 image. Winter wheat should be jointed and will appear bright red. Corn and flax are just being planted so still appear as bare soil.

On acquisition 3 (June 11), spring small grains are jointed to booting and should have significant vegetation canopy and a definite red color on the image. Winter wheat is headed and should be bright red or beginning to darken. Flax and corn are emerging but may not have sufficient canopy cover to allow detection of vegetation within these fields as yet.

On acquisition 4 (June 30), spring small grains should be headed and bright red or beginning to darken slightly. Winter wheat is in the soft-dough stage and should be showing definite signs of darkening. Corn and flax may show signs of canopy development but may still not have adequate canopy cover present to indicate vegetation on the image.

On acquisition 5 (July 17), spring barley is starting to turn (light greenish or yellowish on image). Spring wheat, however, is not yet to the turning stage but is still in the soft-dough stage and should show darkening on the image. Winter wheat has turned and harvest has started. These fields will appear bright white on the image if stubble is still present or darker gray if the field has been plowed after harvest. Corn and flax are tasseling and blooming, respectively, and should appear as some shade of red on the image.

On the last acquisition (August 23), all small grains (winter and spring) have been harvested (whitish for stubble, gray for plowed field), flax is starting to turn (darkening on image), and corn is dented (still appears red on the image).

This example has been presented in simplified form to demonstrate the principles involved in multitemporal analysis for crop-type identification. It can be seen that small grains have a unique pattern compared to other crop types. The slight difference in rate of maturation of spring barley and spring

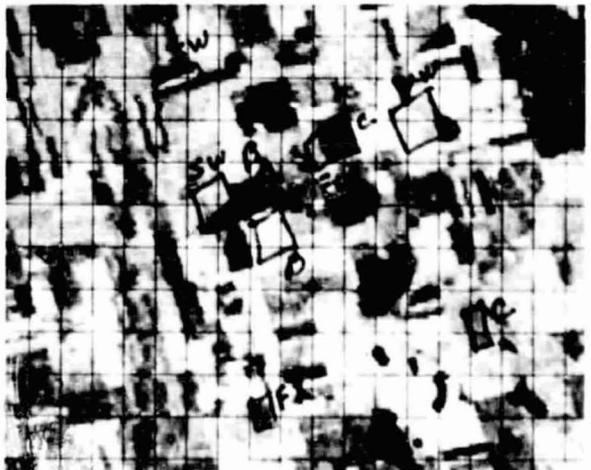
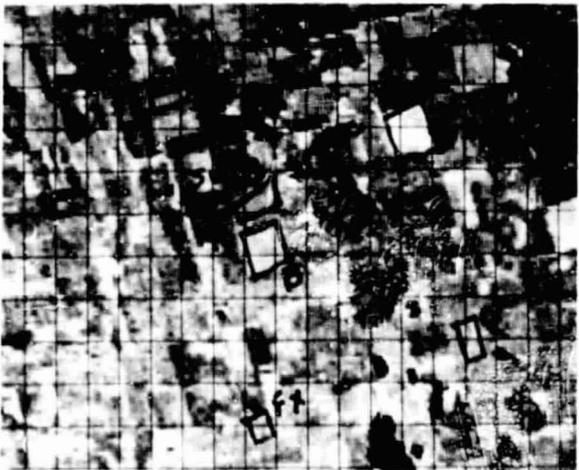
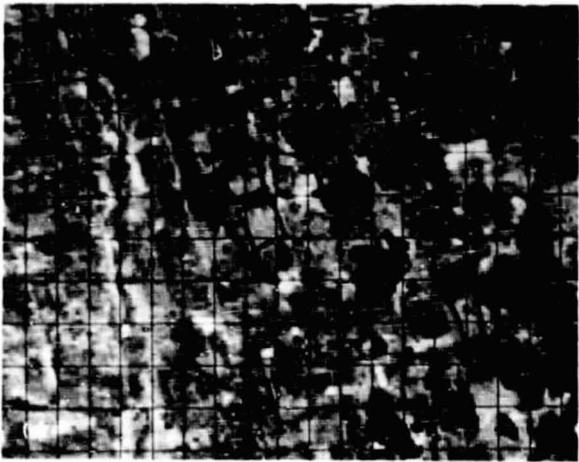
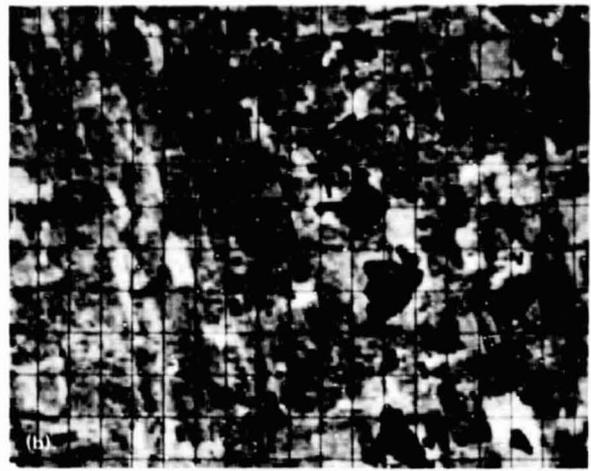
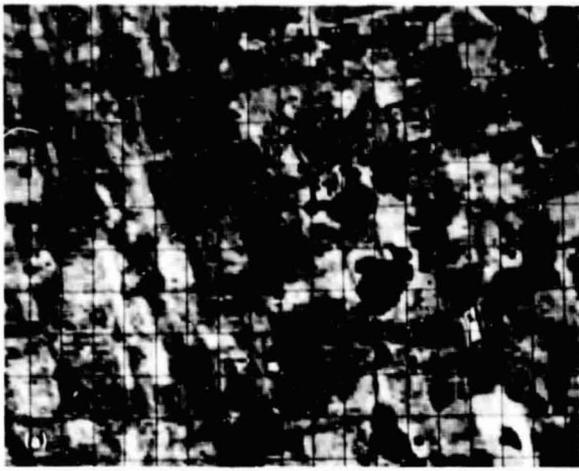


FIGURE 7.—1976 crop year multitemporal set of acquisitions for a segment in Cass County, North Dakota; SW = spring wheat, B = spring barley, WW = winter wheat, FX = flax, and C = corn. (a) May 7. (b) May 25. (c) June 11. (d) June 30. (e) July 17. (f) August 23.

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

wheat is also demonstrated. Precise optimum timing of acquisitions at wheat-soft-dough and barley-turning biostage is critical. An acquisition missing altogether at this period or less than optimally timed can make wheat and barley separation almost impossible. Also, the heavy dependence on crop calendar data should be noted. As stated earlier, crop calendar data is the single most important piece of ancillary data in crop-type identification.

MANUAL CROP IDENTIFICATION PROBLEMS

In Phase I and Phase II of LACIE, it was found that in some regions, the analyst's interpretation error was beyond the tolerance limits. Several problem areas associated with each of the two main interpretation components, feature detection and feature identification, were identified, and the solution of these problems was addressed by LACIE through cooperation between the research community and LACIE operations personnel. A detailed description of the manual interpretation problems encountered and the supporting research efforts on these problems is presented in the paper by Hay entitled "Manual Landsat Data Analysis for Crop Type Identification." A briefer discussion of manual interpretation problems will be presented here for the purpose of completeness within this paper.

Problems in Feature Detection and Characteristics Determination

As stated earlier, Landsat data are used for feature detection and physical characteristics determination. Thus, the capability of Landsat data products to clearly and accurately represent spatial and spectral data to the analyst is of great concern. During LACIE Phases I and II, the only Landsat data products available to analysts were the CIR image Product 1 and positive-negative image Product 2. These image products are good, effective data display formats for the extraction of spatial information, such as feature size, shape, relationship to neighboring features, and distribution throughout an area. However, image Product 1 can provide only gross, relative spectral information about a feature. Although gross, relative spectral information is often sufficient for crop-type identification using multitemporal analysis procedures, numerous situa-

tions were encountered in LACIE Phase I and Phase II where the image Product 1 did not represent the Landsat spectral data sufficiently for correct crop-type labeling.

Another crop identification problem related to feature characteristics determination is insufficient temporal sampling. As stated previously, the temporal-spectral pattern throughout the growing season is the most significant feature characteristic for crop-type identification. If this characteristic pattern is not adequately determined, there is a greater probability of confusion between crop types. Two causes of insufficient temporal sampling which lead to inadequate temporal-spectral pattern determination are (1) missing Landsat acquisitions because of cloud cover or other reasons and (2) periodicity of Landsat overpasses.

Features below the resolution limit of the Landsat sensors (approximately 1 acre) cannot be detected. Thus, correct crop identification with Landsat-1 and Landsat-2 data is impossible for fields less than 1 acre and improbable for fields from 5 to 10 acres. The improbability of correctly identifying fields from 5 to 10 acres is a function of misregistration between acquisitions and boundary (mixed) pixel problems. It is necessary to determine the spectral changes of a field over time fairly accurately. If data points representing a given ground location cannot be overlaid from one acquisition to another to a fairly precise degree, an accurate temporal-spectral pattern and crop type cannot be determined.

Problems in Feature Evaluation

Most of the remaining sources of error in manual crop-type labeling in LACIE were a function of insufficient a priori knowledge or ancillary data or of nonoptimum labeling procedures.

Insufficient a priori knowledge and ancillary data.— One deficiency in a priori knowledge which had an effect on analyst labeling accuracy, particularly in the early phases of LACIE, was the lack of adequate information concerning the *variability* in the temporal-spectral patterns of wheat, small grains, and other crop types. A related deficiency was the lack of adequate crop-type temporal-spectral separability information. No specific information about the temporal-spectral patterns of crop types other than wheat was available to the analysts. These deficiencies resulted in omission errors for wheat and small grains. Analysts incorrectly assumed less variability in

wheat temporal-spectral patterns than was actually present. Thus, interpreted labels were conservative with respect to wheat. As the analysts' experience increased through LACIE, they developed a better feeling for the true wheat temporal-spectral pattern variability. However, additional variability information was definitely needed in abnormal situations, such as the occurrence of drought, winterkill, or other episodal events. Similarly, without specific information about the temporal-spectral patterns of crops other than wheat, analysts could not "doublecheck" their identifications by working the problem in reverse. That is, in addition to responding to the question, "Is this pixel wheat?" the analyst could have posed and responded to the questions (1) "What crop type is this pixel?" and (2) "What crop types are definitely not represented by this pixel?" Being able to eliminate certain crop types often forces the analyst to go back, reconsider, and change his initial answer to the limited question first posed. However, since the analyst did not have the other specific crop-type data and the needed temporal-spectral variability information, he could not doublecheck his initial answer. The result was that some wheat was mislabeled or omitted.

Nonoptimum labeling procedure.—A large number of labeling errors traced to the analyst consisted of the labels affixed to misregistered and boundary (mixed) pixels. Misregistered pixels are those that jump back and forth between one field and an adjacent one on successive acquisitions. Boundary pixels are mixtures of the signatures from two adjacent fields. In LACIE, the analyst affixed a definite crop-type label to a boundary or misregistered pixel. To do this, he specified a reference acquisition on which he labeled the pixel. He "guaranteed" the pixel label for that reference acquisition only, and not for any other acquisitions. This led to analyst-credited "mislabeled" when the pixel label was not appropriate for the majority of the segment's acquisitions that were machine processed and subsequently checked in accuracy assessment.

SUMMARY AND CONCLUSIONS

This paper constitutes an attempt to analyze the manual interpretation process. Although manual crop identification in LACIE did not achieve 100-percent accuracy, the error in crop identification was more a consequence of insufficiencies in the Landsat, a priori, and ancillary data than totally the result of inaccurate or nonoptimum procedures. Situations

will occur where the "correct" interpretation cannot be reached on the basis of the data available for interpretation. No procedure, whether it be manual or automatic, can consistently reach the correct conclusion if the set of necessary and sufficient data is not available. The manual interpretation procedures used in LACIE were adequate to support the accuracy goal for winter wheat and U.S.S.R. spring wheat (see the paper by Marquis entitled "Lacie Area, Yield, and Production Estimate Characteristics: U.S. Great Plains," the paper by Hickman entitled "LACIE Area, Yield, and Production Estimate Characteristics: U.S.S.R.," and the paper by Potter et al. entitled "Accuracy and Performance of LACIE Area Estimates"). In areas where the accuracy goal was not met (namely, U.S. and Canadian spring wheat areas; see the papers by Marquis and Potter et al. and the paper by Conte et al. entitled "LACIE Area, Yield, and Production Estimate Characteristics: Canada"), the failure to provide adequate acreage estimates was at least as equally, if not more directly, due to Landsat sensor limitations (resolution limit, temporal sampling rate, etc.) as to manual interpretation deficiencies. This does not say that better manual interpretation and measurement procedures are not possible. Indeed, much research is in progress toward such improvements. It does say, however, that current procedures are adequate to support a technology capable of supplying valuable agricultural resource information.

This paper has stressed only the logical processes involved in interpretation rather than giving a highly detailed step-by-step description of analyst procedures. Such a step-by-step description of LACIE analyst procedures is available elsewhere (ref. 2). In conclusion, the LACIE experience has clarified the perception of the manual interpretation process. This clearer perception will significantly aid current research in improving manual interpretation procedures and developing automated or semiautomated crop-type labeling procedures.

REFERENCES

1. Rabben, E. L.; et al.: Fundamentals of Photointerpretation. Manual of Photographic Interpretation, R. N. Colwell, ed., American Society of Photogrammetry (Washington, D.C.), 1960, p. 109.
2. LACIE Phase III CAMS Detailed Analysis Procedures. LACIE-00720, NASA Johnson Space Center, Houston, Texas, Aug. 1977.

System Implementation and Operations

FOREWORD

The LACIE Applications Evaluation System (AES) was composed of the personnel, procedures, and systems which over a 3-year period operated and evaluated the LACIE technology. This was a very diverse activity with three participating government agencies, utilizing numerous data systems and facilities located across the United States. The varied system products were integrated to produce the final LACIE output at the NASA Johnson Space Center (JSC).

The AES was functionally separated into data acquisition and management, area estimation, yield estimation, production estimation, system operations and control, and efficiency and accuracy assessment. This chapter will be directed toward the phased implementation of the Large Area Crop Inventory Experiment and the manner in which the AES was operated and evaluated.

The acquisition, preprocessing, and storage of data for LACIE was probably the most diverse of the functions. The collection and processing of Landsat data was the responsibility of the NASA Goddard Space Flight Center (GSFC). The existing Landsat ground processing system was used to obtain the Landsat imagery, and the LACIE processing system implemented at GSFC provided custom processing to produce the 5- by 6-nautical-mile segments required by the project. This activity is described in "Acquisition and Preprocessing of Landsat Data" by Horn et al. The other major source of real-time data was the existing worldwide weather station networks. This weather information was assembled and formatted for LACIE use by NOAA and is described in "Operation of the Yield Estimation Subsystem" by McCrary et al.

The reformatting, storage, and retrieval of the data required in the LACIE project was a major activity. A vast amount of digital Landsat data was acquired and processed on a daily basis and maintained in electronic data bases as described in "LACIE Data-Handling Techniques" by Waits. Ground inventories were obtained for about one-third of the U.S. seg-

ments each year. The collection and handling of this "ground truth," performed for the accuracy assessment program, is addressed in "Ancillary Data Acquisition for LACIE" by Spiers and Patterson. The majority of the nonelectronic data used in the LACIE project (e.g., maps, periodicals, photographic products, reports) was stored in an extensive data library as detailed in "The Acquisition, Storage, and Dissemination of Landsat and Other LACIE Support Data" by Abbotts and Nelson.

The element of the AES that had the operational responsibility for Landsat data analysis was the Classification and Mensuration Subsystem (CAMS). The implementation and operation of CAMS was the responsibility of the JSC Earth Observations Division (EOD). A discussion of the phased implementation and operation of CAMS is presented in "The Classification and Mensuration Subsystem" by Abotteen and Bizzell. The major portion of the CAMS operation utilized a batch-run capability on a large computer in the JSC Mission Control Center. During the latter part of Phase II, an interactive capability was implemented on a small computer in the JSC-EOD as described in "Concepts Leading to the IMAGE-100 Hybrid Interactive System" by Mackin and Sulester. This system was used during Phase III to develop procedures and operationally process 50 segments in the U.S.S.R. This operational processing was performed by USDA analysts and is addressed in "USDA Analyst Review of the LACIE IMAGE-100/Hybrid System Test" by Ashburn et al.

The NOAA played a major role in the project by supplying real-time and historical meteorological data and developing and operating yield models and crop calendars. These activities are explained in the paper by McCrary et al. The yield model results were aggregated with the area estimates to produce the production estimates. The crop calendar outputs were provided to the CAMS analysts to aid in relating expected crop growth stage to the signatures observed in the Landsat imagery.

The area and yield estimates were input to the Crop Assessment Subsystem (CAS). CAS was implemented by the JSC-EOD during Phase I and oper-

ated by USDA personnel during the remainder of LACIE. CAS periodically, on a predefined schedule, aggregated the area and yield inputs, produced production estimates, and generated detailed reports to document the LACIE estimates. The development and operation of the CAS system is described in "The Crop Assessment Subsystem: System Implementation and Approaches Used for the Generation of Crop Production Reports" by McAllum et al.

The various subsystems of the AES were each controlled by a subsystem manager primarily concerned with the accomplishment of the objectives of that particular subsystem. The coordination of the activities of the major functional subsystems and the collection of the information to provide project management with insight into the actual operation of the AES as a system was provided by a number of coordinators, including the LACIE operations manager, quality assurance manager, facilities manager, the

manager of the Data Acquisition, Preprocessing and Transmission Subsystem, and the manager of the Information Storage, Retrieval, and Reformatting Subsystem. Descriptions of these coordination activities are presented in "LACIE Status and Tracking" by Dauphin et al., "LACIE Quality Assurance" by Gutschewski, "Operations Reporting" by Musgrove and Dale Marquis, and "EOD Facilities Configuration Management Office" by Dauphin and Palmer.

The results of the experiment operations were evaluated by the Accuracy Assessment System, which concentrated on the accuracy of the area, yield, and production estimates. The approach and the system developed for this assessment is delineated in "Accuracy Assessment System and Operation" by Pitts et al. The efficiency of the experiment operation was monitored during the three phases of LACIE and is discussed in the "LACIE Applications Evaluation System Efficiency Report" by White.

Acquisition and Preprocessing of Landsat Data

T. N. Horn,^a L. E. Brown,^b and W. H. Anonsen^b

INTRODUCTION

Early in 1974, a development effort was undertaken at the NASA Goddard Space Flight Center (GSFC) to establish a data acquisition and processing system to support the LACIE. Designated the Data Acquisition, Preprocessing, and Transmission Subsystem at GSFC (DAPTS/GSFC), this system was to provide Landsat data inputs to the LACIE system at the NASA Johnson Space Center (JSC), where a joint NASA/U.S. Department of Agriculture (USDA)/National Oceanic and Atmospheric Administration (NOAA) team would perform analyses and evaluation in pursuit of LACIE objectives. Requirements imposed on the GSFC system included the following.

1. Temporal registration of selected Landsat data to within 1 pixel root mean square
2. Data acquisition, processing, and transmittal to JSC within 7 days
3. Capacity to handle data for 4800 site locations, with multiple coverage required for 960 sites and four-time coverage required for the remaining 3840 sites

In response to these requirements, DAPTS/GSFC was configured to operate as an integral part of the established Landsat ground system at GSFC. This system was designed to use existing equipment and processing capabilities as much as possible in order to maximize hardware compatibility and minimize software development. An operation team with considerable Landsat experience was assembled to support the start of production processing in January 1975. The development of that initial system and the evolutionary changes which followed have consistently been aimed at providing in a timely fashion the data required by the LACIE organization. A review of system performance since early 1975 substantiates the success this system has achieved.

^aGeneral Electric Space Division, Beltsville, Maryland.

^bNASA Goddard Space Flight Center, Greenbelt, Maryland.

LANDSAT OVERVIEW

The first Earth resources technology satellite, Landsat-1, was launched in July 1972. Its mission was to orbit the Earth and return images of the Earth's surface. A second Landsat satellite was launched in January 1975, and a third was launched in March 1978. Each satellite contained a multispectral scanner (MSS) and a return beam vidicon (RBV) camera system. The MSS (fig. 1) consists of an oscillating mirror that scans the Earth's surface horizontally, while the forward motion of the spacecraft provides vertical displacement. The Landsat-1 and Landsat-2 scanners are four-channel systems, while the Landsat-3 MSS is a five-channel scanner. The RBV units (fig. 2) are shuttered camera systems. On Landsat-1 and Landsat-2, three cameras were arranged to provide coincident images in three spectral bands. The Landsat-3 system substitutes two panchromatic cameras that are aligned to provide adjoining images which have improved resolution. The spectral characteristics of both sensors are listed in table I.

The operation of all Landsat satellites in collecting image data is directed and controlled by the Landsat Operations Control Center, which is located at

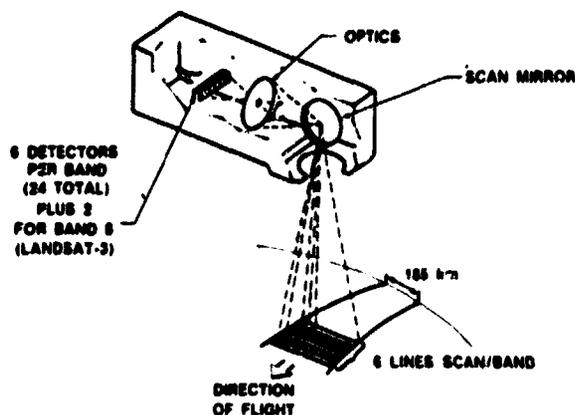


FIGURE 1.—MSS scanning arrangement.

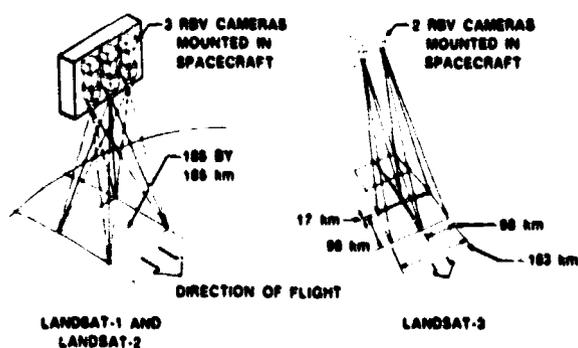


FIGURE 2.—RBV scanning patterns.

TABLE I.—Sensor Spectral Characteristics

Channel	Spectral band, μm
<i>RBV system</i>	
1	0.5 to 0.6
2	^a 0.6 to 0.7
3	0.7 to 0.8
Panchromatic	^b 0.5 to 0.8
<i>MSS</i>	
4	0.5 to 0.6
5	0.6 to 0.7
6	0.7 to 0.8
7	0.8 to 1.1
8	^b 10.4 to 12.6

^aLandsat-1 and Landsat-2 only

^bLandsat-3 only

GSFC. Once acquired, image data are either relayed directly to a Landsat ground station or recorded on the satellite for later transmission. The three primary Landsat ground stations are located in Alaska, in California, and at GSFC; a portable ground station was also deployed in Pakistan between October 1976 and September 1977 to assist in collecting Landsat image data. The ground station records the data on wideband video tapes, and these tapes are then sent to GSFC for processing. The overall Landsat system is illustrated in figure 3.

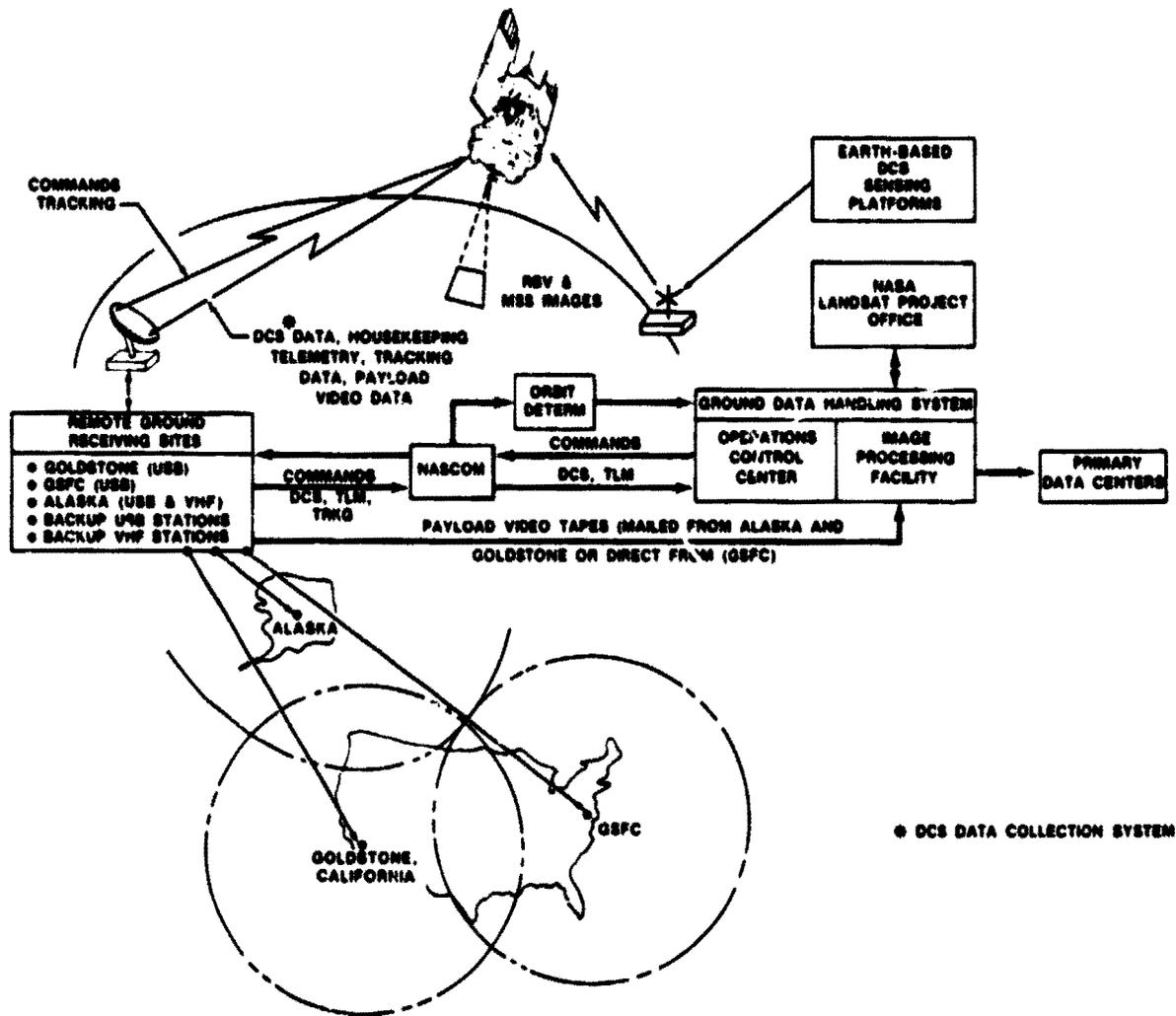
At GSFC, MSS and RBV data are processed to produce both photographic and digital data products. The processing involved includes reformatting, annotation, radiometric calibration, and geometric correction of various types, depending on data and product type. The Landsat data processing flow is shown in figure 4. In its current hybrid configuration, the Landsat processing system produces a film archive in 70-millimeter format. This archive is then used to generate photographs for Landsat users. Upon request, the original video tape is used to produce digital products. Copies of the film archive are also provided to several other data centers for use in generating and distributing Landsat data products of various types. Future plans at GSFC call for conversion to a digital archive, with accompanying improvements in image processing capability. More thorough descriptions of various Landsat system elements are provided in the Landsat Data Users Handbook (ref. 1).

DAPTS/GSFC OVERVIEW

In order to meet LACIE requirements for Landsat MSS data, a dedicated acquisition and processing flow path was established at GSFC within the Landsat ground system. Although many of the processing functions are similar, throughput requirements and time-line constraints dictated that a separate end-to-end flow be established. Several elements included in this path also perform other Landsat functions on a shared basis, while other elements are totally dedicated to LACIE support. The DAPTS/GSFC system is shown in figure 5. Comparison of this figure to figure 4 illustrates the similarity between the two systems.

Each block in figure 5 represents a separate processing subsystem at GSFC. As shown, GSFC receives LACIE requirements on computer tapes (or cards). These inputs are processed within the general-purpose image processor subsystem and a test site tape is generated. This tape is provided to the Control Center in order to schedule Landsat MSS coverage of LACIE test sites. Unlike most other Landsat coverage requirements, LACIE coverage scheduling does not include consideration of predicted cloud cover (although the capability to do so is available). Data are recorded on wideband video tapes and shipped to GSFC as an integral part of the Landsat data flow.

The Control Center generates two computer tapes



• DCS DATA COLLECTION SYSTEM

FIGURE 3.—Overall Landsat system.

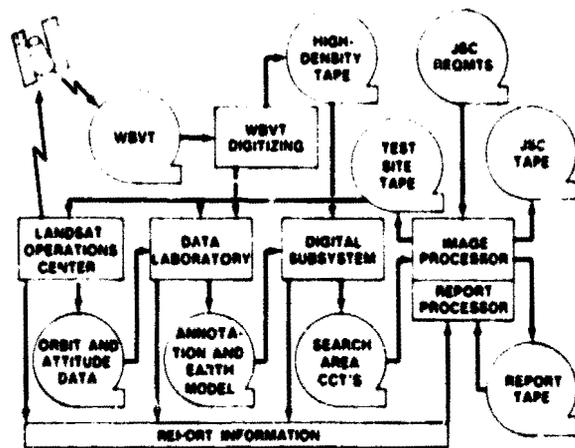


FIGURE 4.—Landsat data acquisition and processing flow at GSFC.

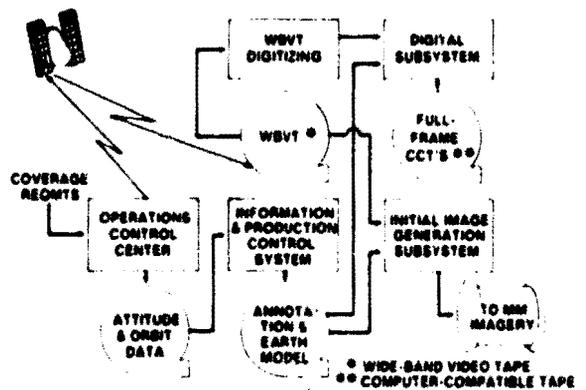


FIGURE 5.—DAPTS/LACIE data acquisition and processing flow at GSFC.

to enable processing and tracking of MSS data. After data acquisition has been confirmed, a spacecraft location and attitude tape is delivered to the Data Services Laboratory (DSL) for processing. This tape defines the acquisition dates for which processing is to be performed and contains the ephemeris and telemetry data required for processing. A status tape for inclusion in the LACIE master file is also generated to report on the scheduling and data acquisition activities performed by the Control Center. The Control Center uses XDS-Honeywell Sigma 3 and Sigma 5 computers in performing its functions.

The DSL locates LACIE data within the MSS data stream, calculates geometric correction coefficients, and performs the image annotation processing required for LACIE data. A copy of the test site tape provided to the Control Center is used by the DSL to identify LACIE test site data. Detailed descriptions of the algorithms involved in data location and correction coefficient development are provided in reference 2. The resulting annotation data are recorded on the LACIE image annotation tape, which is then provided to the digital subsystem. This annotation tape contains the information required by the digital subsystem to extract 10- by 11-nautical-mile "search areas" from the 100-nautical-mile-wide swaths of MSS data. In addition to the annotation tape, the DSL also produces a status report tape, which provides processing activity inputs for the LACIE master file. Processing for the Data Services Lab is performed on an XDS-Honeywell Sigma 5 computer.

When the video tapes have been received at GSFC, the MSS data for LACIE are digitized by the MSS preprocessor. Data from the video tapes are transferred to a high-density digital tape, which can be further processed by the digital subsystem. During the transfer process, data quality checks are performed, and poor-quality data are replaced by adjacent data on a line-by-line basis. (Replacements of this kind result in data being flagged as marginal when transmitted to JSC.) Information on MSS preprocessing is manually transferred to the DSL for inclusion in the LACIE status report tape.

The digital subsystem processes data from the high-density tape, using the LACIE image annotation tape as a control. Data for each LACIE site are extracted from the high-density tape and transferred to nine-track computer-compatible tapes (CCT's). Radiometric calibration is performed during the extraction process. Following extraction, each data set is reprocessed to permit reformatting, development

of a radiometric histogram table, and cloud-cover screening. Automatic cloud-cover detection is performed within the digital subsystem using a simple level-slicing technique. The criterion for rejection was set at 10 percent of the pixels in the search area. The radiance threshold was initially set at an absolute count of 90 in the 0.5- to 0.6-micrometer channel (band 4). However, after several months of operation, this level was adjusted to 60 in order to more closely match the cloud sensitivity level of the temporal registration process. If more than 10 percent of the pixels have a count greater than 60, the data set is discarded; otherwise, it is transferred to the search area computer-compatible tape, which is provided to the general-purpose image processing subsystem. In addition to the search area tape, the digital subsystem produces a report tape containing inputs to the LACIE master file.

The general-purpose image processor performs the geometric correction and temporal registration functions required of DAPTS/GSFC. The data within each search area are first geometrically corrected through resampling and then registered with previous data for the same LACIE site. An edge detection and correlation technique is used to establish temporal registration between data sets. This technique involves using radiance gradients within each data set to identify feature edges, such as field boundaries. (Feature recognition of this type should not be affected by seasonal changes in radiance levels.) The edge patterns developed for two data sets are then compared statistically to determine a registered alignment. This technique is illustrated in figure 6, and the algorithms involved are described elsewhere

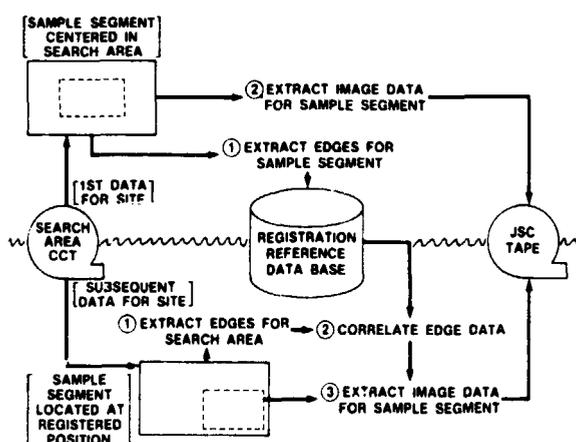


FIGURE 6.—Temporal registration process.

in this volume (G. Grebowsky, "LACIE Registration Processing System") and in reference 2. The results of the correlation process are used to identify and extract from within the search area a 5- by 6-nautical-mile "sample segment" that precisely matches previously processed data. If no prior data has been processed for a site, the 5- by 6-nautical-mile sample segment to be sent to JSC is extracted from the center of the search area. The histogram data provided on the search area tape are used to calculate film recorder parameters used in JSC processing, and the extracted sample segment is transferred to an output tape for transmission to JSC. As in other subsystems, a status report tape is produced for use in updating the LACIE master file.

In order to support the temporal registration process, the controlling data base for DAPTS/GSFC is maintained within the image processor. Each requirement update received from JSC is processed within the image processor, and the registration reference file is also maintained in this subsystem.

As a separate function, the image processor updates the LACIE master file, processing status report tapes from other subsystems and from its own registration processing. In addition, requirement updates, final data transmission reports, and postprocessing quality and inspection reports are also recorded in the master file, resulting in an integrated end-to-end record of DAPTS/GSFC activities. The image processing subsystem also produces master file reports in data list form, primarily for analysis purposes. General-purpose image processing involves the application of an enhanced XDS-Honeywell Sigma 3 computer.

As an addition to the LACIE system shown in figure 2, a master file retrieval system was developed to provide an additional capability to produce status, summary, and analytical reports on DAPTS/GSFC operation. Master file tapes provide input data to this system, which operates on the UNIVAC 1108 computer at GSFC (not a part of the Landsat ground system).

INITIAL CONFIGURATION

In January 1975, DAPTS/GSFC was configured to support LACIE through the use of Landsat-1. Newly acquired data were to be processed and transmitted to JSC within 7 days. At that time, video tape data were transferred to high-density tape by the initial image generation subsystem, as the MSS

preprocessor was not yet operational. Telemetry and ephemeris data were provided by the Control Center on separate tapes, and only a data list could be used in producing master file reports.

In this initial configuration, constraints were imposed on LACIE site location in order to preclude search area extraction overloads in the digital subsystem, where only one pass through the Landsat data stream was permitted. The image processing Sigma 3 computer system, which was originally a part of the Landsat scene correction (i.e., precision processing) subsystem, did not initially include a disk storage capability; as a result, a tape-oriented controlling data base system was implemented. With only research/development study results rely on in implementing the temporal registration process, an interactive registration system was readied to supplement the automatic correlator (or to replace it if necessary). Arrangements for daily data transmissions to JSC involved courier service to a nearby airport, airfreight transportation to Houston, and courier service pickup and delivery to JSC.

When production operations began on January 13, 1975, an evolutionary sequence also began, eventually leading to the system currently in operation at GSFC. Some enhancements were actually Landsat system improvements, from which the LACIE system also benefited, while others were improvements specifically intended to upgrade the GSFC LACIE support capability or to meet a newly established LACIE requirement.

SYSTEM ENHANCEMENTS

Landsat-2 was successfully launched on January 22, 1975, and, after being declared operational in early February, was designated to replace Landsat-1 as the prime source of data for LACIE use. Both satellites, however, remained in operation to serve the Landsat user community, and processing loads at GSFC increased accordingly. In March, the MSS preprocessor became available to support LACIE processing and Landsat digital data product generation for other users. The introduction of this subsystem allowed the initial image generation subsystem to be dedicated to film production. This modification enhanced the GSFC capability to support two Landsat satellites and also eliminated the possibility of conflicts between LACIE processing and Landsat film archive generation.

In November 1976, the LACIE system at GSFC

again benefited from a Landsat system enhancement. Spacecraft telemetry and ephemeris data, which had previously been transmitted across the Control Center/processing system interface independently, were merged with video tape information into a spacecraft location and attitude tape. This operational simplification resulted in time-line improvements for both Landsat and LACIE processing.

Several enhancements unique to LACIE have also been implemented since early 1975 and have resulted in both increased throughput and improved performance. Early analysis results indicated that the GSFC system processed data that were "too cloudy" to be useful at JSC. (Excessive snow cover was also considered unacceptable.) Accordingly, an interactive screening function was established within the image processor to permit operator rejection of data not meeting JSC criteria. In order to expedite processing, an overlapping configuration was established, in which two data sets would undergo processing in a time-shared fashion. In this configuration, the data from one set are displayed for operator examination while the other set is involved in computer processing. As illustrated in figure 7, the two data sets alternately undergo display and processing in staggered fashion through the several steps necessary to complete registration processing. This design allows early rejection of data judged to be too cloudy, provides for a later "second look" to handle marginal cases, and maximizes the efficiency of both the operator and the computer system throughout the processing sequence.

OPERATOR/DISPLAY FUNCTIONS	GENERAL-PURPOSE IMAGE-PROCESSOR COMPUTER FUNCTIONS
SCREEN 1st INPUT (SEARCH AREA)	GEOMETRIC CORRECTION OF 1st INPUT
	TEMPORAL REGISTRATION OF 1st INPUT
SCREEN 2nd INPUT (SEARCH AREA)	GEOMETRIC CORRECTION OF 2nd INPUT
SCREEN 1st OUTPUT (SAMPLE SEGMENT)	TEMPORAL REGISTRATION OF 2nd INPUT
SCREEN 3rd INPUT (SEARCH AREA)	GEOMETRIC CORRECTION OF 3rd INPUT
SCREEN 2nd OUTPUT (SAMPLE SEGMENT)	TEMPORAL REGISTRATION OF 3rd INPUT
SCREEN 4th INPUT (SEARCH AREA)	GEOMETRIC CORRECTION OF 4th INPUT
SCREEN 3rd OUTPUT (SAMPLE SEGMENT)	TEMPORAL REGISTRATION OF 4th INPUT
SCREEN 4th OUTPUT (SAMPLE SEGMENT)	

FIGURE 7.—Interactive screening/ processing sequence.

After 2 years of operation, it became evident that cloud/snow-cover rejection rates were seasonally dependent. During the spring and summer months, lower data rejection rates often could not justify the operator/display time involved in the screening process. Accordingly, the image processor configuration was again modified to provide cloud-cover screening as a selectable option. In this present configuration, the on-line screening option is selected when rejection rates are high and when the data flow rate is low (i.e., primarily late fall through early spring). When screening is not performed on-line, data not meeting JSC criteria are rejected as a part of the postprocessing quality inspection function.

In a separate enhancement relating to cloud cover, a revised correlation technique was implemented in the summer of 1976 to minimize the adverse effects of clouds and cloud shadows. This technique involved recognition of cloud/shadow areas within the data being processed and avoidance of these areas in correlation processing. Details of this enhancement are provided in the paper by Grebowsky. In October 1976, GSFC and JSC participated jointly in establishing a data link interface for use in transmitting Landsat data to JSC. The use of this link has improved data delivery time considerably and has eliminated the inefficiency and complexity of the courier service/airfreight interface.

As the image processor configuration was refined and improved, initial emphasis on tape storage of data gave way to a disk-storage design. Late in the fall of 1976, the data base that contained all registration reference data was established in permanent disk-resident form. The reduction in reference data access time which resulted represents a significant improvement in image processing efficiency.

NEW CAPABILITIES

New requirements in supporting LACIE have been imposed on GSFC at various times over the 3 years LACIE has been underway. Responses to these requirements have resulted in provision of several new acquisition and processing capabilities at GSFC. These new capabilities combine with previously described system enhancements to account for the evolutionary development of the current DAPTS/GSFC configuration.

When production processing began in January 1975, the data being processed had actually been acquired some 3 months earlier, during the fall of 1974

Although initial DAPTS requirements excluded such "retrospective" processing, a capability to perform retrospective processing was added to the DAPTS/GSFC configuration at JSC request. In order to minimize the impact of this processing mode, the system was modified to accept bulk image annotation tapes generated during Landsat processing and then to perform only the additional calculations required to produce the LACIE image annotation tapes. Although this approach helped, retrospective processing remains an expensive capability, particularly in terms of throughput and time. Accordingly, GSFC has recommended minimal use of this capability.

Several of the constraints (ref. 3) which were imposed on LACIE site location by the initial DAPTS/GSFC configuration were eliminated when the DSL and the digital subsystem were modified to allow more than one pass through the MSS data stream. As a result of these changes, a search area that overlaps another or that violates other location constraints can now be deferred and extracted in a second pass through the data stream. This modification has permitted JSC to obtain critical data that were previously unavailable. Although a second pass involves rewind and reprocessing delays, it does not significantly reduce DAPTS/GSFC throughput.

An accompanying enhancement recently installed in the DSL eliminates the redundant processing of data located within Landsat frame overlap regions. Such data, although of no value to JSC, were processed and transmitted to JSC before this change was incorporated. The time saved in not processing these data is now available to process other data that are useful, thereby increasing the throughput of the system.

The initial LACIE concept called for a complement of 4800 sites, with data for most sites to be acquired on a selective basis. When JSC analysis planning revised the requirement to involve full-time coverage, data overloads both at GSFC and JSC were immediately projected. Accordingly, a "pseudocoverage" system was established to permit the acquisition of all data but the processing of only a selected subset, with the remainder archived for possible later use. This pseudocoverage capability was placed in operation in August 1975 and has been used periodically since then to facilitate LACIE data requirements adjustments. The pseudocoverage capability does not involve any ground system elements except the Control Center, for acquisition scheduling purposes.

After several months of LACIE production processing, the need to provide additional acquisition and processing status information to JSC became apparent. In response to this need, a status report interface was established to report on all data which were rejected from processing at GSFC. These reports were produced as a byproduct of the LACIE master file update sequence and were transmitted to JSC on a regular basis. However, the incremental nature of these reports made them difficult to summarize, and in the fall of 1976, a cumulative report interface was established to supplement, and then replace, the incremental reports. This cumulative report is actually a historical record of the postprocessing quality inspection activity at GSFC and has been provided to JSC since 1976 on a monthly update basis.

As a part of the registration process, the first data set processed for each site becomes the reference data for use in subsequent registration processing. Data of poor quality occasionally appeared as reference data, leading to correspondingly poor correlation results. In May 1976, a new capability was added to the image processing system which permits replacement of such reference data without losing registration. When directed by JSC requirement inputs, the reference data for specified sites are now replaced by other data for which registration has been successfully accomplished, thereby maintaining registration continuity. This capability has since been used to improve correlation results for a number of sites suffering from the effects of poor-quality reference data. An additional general-purpose image processor enhancement requested by JSC at the same time involved manipulation of an annotation flag to indicate the first data set of each "biological window" (i.e., acquisition time), for use in generating appropriate film data products at JSC.

In the summer of 1976, analysts at JSC established a requirement for regular Landsat imagery to complement the sample segment data provided via DAPTS/GSFC. Arrangements were made with the USDA facility in Salt Lake City to provide the required imagery as a part of their Landsat data dissemination function. In order to support this arrangement, the Landsat processing system at GSFC was modified to produce a work order that identified each Landsat scene from which one or more LACIE sample segments had been extracted. Production of this work order continued until late 1977, by which time USDA capabilities had been upgraded to eliminate the need for this information.

Landsat-3 was successfully launched on March 5,

1978, making five-channel MSS data available to Landsat users for the first time. However, extensive LACIE system modifications would have been necessary both at GSFC and at JSC in order to handle the fifth-band data. As a result, the LACIE system at GSFC was modified to exclude the fifth channel and to extract and process Landsat-3 data only from channels 4 through 7. Further modifications were made to permit two-satellite acquisition and processing activities, thereby allowing more frequent coverage over selected LACIE sites. These modifications were completed in the summer of 1978.

FUTURE ENHANCEMENTS

Throughout 1978, the Landsat ground system at GSFC has been undergoing conversion from the familiar hybrid processing system to a new all-digital configuration. A major element in this conversion involves the new master data processor, which will significantly increase the GSFC capability to geometrically correct and register image data. When fully operational, this system is expected to displace the current DAPTS/GSFC configuration as the source of Landsat data for LACIE, with standard full-frame data replacing the subframe sample segments currently transmitted to JSC. Toward that objective, efforts have been underway since mid-1977 to establish the ground-control-point data base needed to support master data processor registration in regions of LACIE interest. Activities are also underway to establish the high-density tape interface through which master data processor output will be transmitted to LACIE users and to accept and preprocess Landsat data in this form at various user facilities. Several longer range planning exercises have also been undertaken jointly by GSFC, JSC, and other LACIE participants to anticipate and prepare for LACIE or follow-on activities in the early 1980's.

PRODUCTION SUMMARY

In the final analysis, a summary of LACIE processing accomplishments at GSFC best indicates the success of the DAPTS/GSFC endeavor. Between January 1975 and May 1978, some 130 000 data sets have been acquired to support LACIE, with more than 45 000 sample segments successfully extracted

and transmitted to JSC. The system responsible for this performance was effectively implemented in a 1-year period during 1974, at a cost of roughly \$600 000. This effort included the research and development of a heretofore untried registration technique which has since performed beyond all expectations. Registration performance throughout the 3-year period of LACIE operations has satisfied the 1-pixel root-mean-square requirement established in 1974, with more than two of every three attempts at data registration proving successful, notwithstanding the data cosmetic faults or content inadequacies to which the process is inherently susceptible. The cloud/snow rejection rate experienced throughout the last 3 years has approached 50 percent, as expected in most Landsat data use situations. A detailed summary of production processing performance in each year of LACIE operation is provided in table II.

TABLE II.— DAPTS/GSFC Performance Summary

Function	Jan. to Sept. 1975 LACIE Phase I	Oct. 1975 to Sept. 1976 LACIE Phase II	Oct. 1976 to Sept. 1977 LACIE Phase III
	Number of acquisitions	7380	31 296
Number of search areas extracted	3979	15 676	30 438
Number of sample segments transmitted to JSC	2602	10 645	22 718
Rate of cloud snow rejections, percent	30	52	49
Rate of correlation rejections, percent	14	9	4
Rate of miscellaneous rejections, percent	12	5	11

REFERENCES

1. Landsat Data Users Handbook. Report 76 SDS 4258, NASA Goddard Space Flight Center, Greenbelt, Md., Sept. 2, 1976.

2. Thomas, V. L.: Techniques Used To Register Temporal Data Which Have Agricultural Features. Paper presented at NTA Manpower Symposium, Cleveland, Ohio, Oct. 1974.
3. Project Plan for the Data Acquisition, Preprocessing, and Transmission Subsystems of the Large Area Crop Inventory Experiment (DAPTS/LACIE). Appendix A, NASA Goddard Space Flight Center, Greenbelt, Md., Mar. 1978.

Ancillary Data Acquisition for LACIE

B. E. Splers^a and R. L. Patterson^b

INTRODUCTION

The design, implementation, and operational functions of the three phases of LACIE required several types of data in addition to Landsat multispectral digital data. This paper will summarize the types of data required, cover the various collection processes, and describe the procedures for obtaining the data for the user.

The data required by the users in the project fell into four main categories: ancillary data packets, full-frame Landsat imagery, intensive test site (ITS) ground observations, and blind site data.

To aid in the computerized classification process, a packet including the following ancillary data was needed:

1. Statistical data for all crops in each of the countries for a period of at least 15 years at the lowest political subdivision
2. Agronomic data describing farming and crop rotation practices in the wheat growing areas for each of the LACIE countries
3. Soil, topographic, political subdivision, and crop density maps for areas of interest in each country
4. Yearly phenological crop development data for each crop in the area of interest for each of the eight countries for at least 10 years
5. Current-year phenological crop development reports plus periodicals and annual statistical reports for all crops

To develop the sampling strategy and to support crop assessment, full-frame color-infrared (CIR) Landsat imagery needed to be taken throughout the crop season.

To improve LACIE procedures, certain intensive test sites in the wheat growing regions of the United

States and Canada were selected for which the following data were collected:

1. Land use inventories
2. Periodic crop observations
3. Solar radiometer measurements
4. Rainfall
5. Wheat yield for selected fields

To assess the accuracy of the LACIE results, some operational segments in the U.S. Great Plains were designated blind sites for which ground truth was collected; this ground truth consisted of land use inventories and wheat development estimates.

ANCILLARY DATA PACKETS

During the design stage of the project, it was decided that ground-observed data would not be used by the analyst in the computer training classification process. Instead, other supporting data would be used in the identification of crops used to train the classifier. These ancillary data had to be provided for each segment of each country being worked.

The requirements for ancillary data were established using U.S. data sources as a guide. The ancillary data package consisted of 2 years of recent statistics on all crops grown in the county where the segment was located, a summary of farming practices and crop rotations for the general area, a description of the general soil type and productivity, a nominal phenological crop calendar for all crops grown in the segment, and various large- to medium-scale topographic maps.

The ancillary data packets for the United States were developed in-house from data that had been obtained through various contacts with local, state, and federal agricultural agencies. All of the data used in the packet preparation had to be extracted from the reference sources and reformatted or summarized to meet format requirements. This task was time consuming, and the work required close coordination between the preparers and the data analyst.

^aUSDA Agricultural Stabilization and Conservation Service, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

During Phase I of LACIE, U.S. segments and a few scattered segments throughout the other seven countries were processed. The ancillary data for the foreign segments were more difficult. Statistical data sets could not be found to meet all project needs—they were either nonexistent as in China, incomplete as in Russia, or very difficult to obtain as in India. Very little had been published about farming practices in any of the foreign countries, and no phenological data were available to develop segment crop calendars. Small-scale maps were available but medium- or large-scale maps were impossible to get for most of the areas of interest. Many hours were spent during Phase I on developing ancillary data for the few foreign segments that were worked.

As LACIE progressed from phase to phase, compromises were made on many of the data requirements or substitutes were developed to replace important items.

The collection of long-term detailed historical data of the type required for a sampling approach was a paradox. If the data had been readily available, the need for the new technology would not have been apparent; without the data, an optimum sampling strategy to produce an accurate production report could not be designed. Therefore, any and all types of historical crop data at any political level for the countries involved were obtained, hoping that from this mixed assemblage of statistics of varying degrees of detail and varying degrees of accuracy a decent sampling strategy could be devised.

FULL-FRAME LANDSAT IMAGERY

Full-frame Landsat data had been collected for almost 2 years prior to LACIE Phase I. Landsat data had been collected at least once over most of the area defined in the LACIE countries. The U.S. Air Force (USAF) operational navigations chart (ONC) maps (1:1 000 000 scale) that were available and useful did not show agricultural areas, field sizes, or patterns, so a requirement was generated to use black and white 9- by 9-inch single-band full-frame Landsat imagery prints to help delineate the areas of interest. Data searches were made and the imagery was produced by the Aerial Photograph Field Office (APFO) of the U.S. Department of Agriculture (USDA) at Salt Lake City, Utah. The Cartographic Section of the Earth Observations Division (EOD) of the NASA Johnson Space Center (JSC) prepared sectional mosaics using

the ONC's as a base. This product was used in defining the agricultural areas, delineating the sample frame, and developing data to be used in the sampling strategy.

When it was discovered that the sampling strategy had to be refined and the ancillary data could not be provided to meet specifications, full-frame CIR transparencies were considered and used by the project. A set of CIR imagery was produced to cover each of the wheat-growing areas. This imagery was used to redefine the agricultural and nonagricultural areas. This determination was then factored into an improved stratification and sampling frame.

Since the ancillary data were not complete or satisfactory, the CIR imagery was also incorporated into the analysis procedures and crop assessment activities of the experiment. A requirement was defined to obtain coverage for at least one full-frame image with less than 20-percent cloud cover for each LACIE biowindow. The APFO did not have access to a statusing system for full-frame coverage that would meet the timing requirements—data within 14 to 18 days of acquisition. The NASA Goddard Space Flight Center (GSFC) implemented a work order system whereby they provided APFO with the identity of each frame containing LACIE segments. This work order was shipped to APFO with the archival rolls of 70-millimeter film used by APFO to generate the LACIE product. This system provided a method to identify in a timely manner the base product, but the turnaround was still 21 to 30 days from date of acquisition to receipt of CIR imagery. The system still required many manhours of manual labor at the APFO to select the frames to be produced. In the middle of Phase III, APFO implemented a system to generate LACIE full-frame work orders directly from an in-house data base updated with a GSFC update tape received with the roll of film. By the end of Phase III, LACIE was receiving imagery within 18 to 25 days from the date of acquisition.

In addition to the CIR data received to support the acreage analysis, additional data were also ordered from APFO and GSFC to support crop assessment situations, such as the drought in the U.S. Great Plains that occurred in 1976. These data orders were included in the work order system by changing the criterion from less than 20-percent cloud cover to less than 50-percent cloud cover and getting the data for each overpass of both Landsat 1 and Landsat 2. During the three phases of the experiment, over 9000 frames of imagery were generated by APFO and GSFC.

INTENSIVE TEST SITES

Intensive test sites (ITS's) were selected and data collection requirements were established to provide ground-observed data for procedural development and accuracy evaluation for LACIE. Landsat experiments conducted by EOD prior to LACIE had shown the importance of acquiring ground-observed data from areas where Landsat data were being taken for processing. The experience gained in these projects was incorporated into the development of the LACIE requirements and collection procedures. The size, number, and location of the sites were established by the LACIE Intensive Study Area Task Group. In addition to the U.S. sites, 10 sites in Canada were made a part of the LACIE ITS program by using an existing agreement between the Canadian Centre for Remote Sensing (CCRS) and the USDA.

Procedures were developed to handle the ITS data in such a manner that they could be readily reproduced and made available to users within 1 week after receipt. The data collected from these sites were shared by USDA, NASA, and the CCPS. USDA entered the data in a master data file and provided the project a tape file and printout of the complete data base at the end of each crop year.

There were 42 ITS's in Phase I, 37 in Phase II, and 34 in Phase III. Changes were made between phases because of changes in workload. Data from 30 of the sites were received in each of the three phases.

Methods of data handling for the ITS's gradually changed throughout the development of LACIE primarily because of changing participants and scope of involvement by each group. During the 1974-75 crop year, the USDA LACIE Project Office was the organization responsible for collecting all field observation data; however, most of the management of this function was transferred by contract to the Earth Satellite Corporation. This function included preparing data forms, training USDA and Canadian field personnel, checking data for inconsistency and errors, and preparing a compilation of the field data at the end of the crop year. JSC was responsible for obtaining high-altitude aerial photography and subsequently preparing all rectified prints and field boundary maps for each ITS. Copies of all field observation data were sent to JSC to be logged and put into a data library for use by LACIE personnel.

During the next two crop years (1975-76 and 1976-77), the USDA elected to manage directly the field observations program, including preparation of

data forms, production of instruction manuals, and compilation of computerized field data. The NASA functions continued to be a JSC responsibility, with the addition of building and calibrating solar radiometer instruments for use at each ITS. The types of data reported for each ITS were land use inventories, periodic crop observations, solar radiometer, rainfall, and wheat yield for selected fields. The site inventories and periodic crop observations were the most important data obtained, and these will be described in more detail in the following paragraphs.

All ITS's received a complete "wall-to-wall" inventory once every crop year. This task was carried out by USDA or Canadian personnel, usually in May or June, depending on site location. An additional fall inventory was taken at the sites containing winter wheat. The fall inventory of the winter wheat sites identified the content and field boundaries for all fall-planted crops such as winter wheat, barley, or rye along with the following information for each planted field: acreage, crop and variety being grown, irrigation method (if applicable), fertilizer used, and planting date. The annotated photographs or field maps and tabular data were forwarded to JSC for duplication and distribution. The spring inventories were scheduled to begin after spring planting was complete and before winter wheat harvest began. The inventory included the same type information as in the fall inventory; however, it was for all crop types and current land use status within the site. During the growing season, high-altitude aerial photography was acquired over each site, and, after it was processed and screened, the imagery was rectified and scaled. These data were then combined with the annotated photographs or field maps and tabular data forwarded to JSC by the field personnel, and an updated photographic overlay containing current field boundaries and identification was produced. These new photographs, overlays, and field maps constituted the data base for the next site inventory as well as a reference for current-year data processing.

The periodic (18-day cycle) observations within an ITS were scheduled on the days of the Landsat overpass for that site. The observations of approximately 50 fields within the site provided a record of the crop changes for these specific fields throughout the growing season. They began with the planting of fall crops or with spring crop planting where there were no fall crops, and they continued through September or spring wheat harvest. The field personnel

at each site selected approximately 50 fields of which about half contained wheat and the remainder a representative sample of the other major crops grown in that site. The periodic record of crop development throughout the growing season included information such as plant growth stage, percent ground cover, plant height, surface moisture conditions, weed growth, field operations (farming activities in progress), disease or insect problems, and estimated crop quality rating.

During the first two project years, 35-millimeter color photographs and solar radiometer measurements were taken. The photography was taken at each selected field on the observation dates. The solar radiometer measurements were made at the scheduled Landsat overpass time using equipment provided by JSC. This information would permit evaluation of atmospheric interference effects on the Landsat data being acquired. Rainfall was recorded from a network of gauges spaced throughout the test site and reported with the next periodic observation. All periodic observation data were forwarded to JSC, where they were recorded and duplicated for distribution to the appropriate users and to the LACIE data library.

Wheat yield data were reported at the end of the crop year for a minimum of 10 of the observed wheat fields. The fields selected were representative of the yield values within the particular site. The reported data were yield estimates by the farmers or the USDA rather than actual production figures. The lack of specific values was due to the fact that the harvested production was not isolated and available on a per-field basis. The estimated yields were therefore subjective and of varying accuracy.

BLIND SITE DATA

The principal objective of the LACIE was to assemble, operate, and evaluate the remote-sensing technology for providing country-level wheat production estimates. A necessary part of this experiment was to evaluate results and assess errors in order to improve subsequent operational systems. A phased approach was chosen which involved expansion in two directions: in the technical complexity of the functions performed and in the geographic size and difficulty of the area being surveyed. The technical evaluation tasks were performed primarily by an Accuracy Assessment Team.

The LACIE Phase I Accuracy Assessment ac-

tivity in the United States concentrated on the analysis of the ITS data. In order to better assess the LACIE operations, some regular LACIE segments, so-called blind sites, were "ground truthed." The expression "blind site" was merely a designation applied to selected aggregatable segments for which, unknown to the analyst, ground-observed data were acquired for subsequent evaluation purposes. The implementation of this approach occurred late in the growing season of LACIE Phase I. Thus, all of the selected sites fell in the northern spring wheat regions.

High-resolution CIR aerial photography was acquired over 29 LACIE segments in North Dakota and Montana in mid-August 1975. Simultaneously, field teams were collecting ground information for a substantial portion of these segments. These data were combined to obtain both field and total segment ground-observed data.

The Phase I procedure for obtaining these blind site ground data was relatively costly and cumbersome because it involved using personnel from JSC, which necessitated travel and field operation expenses. Because of the usefulness of the data, however, a much larger number of blind sites, with wider distribution, was proposed for Phase II. This required that a new and more cost-effective procedure be developed for obtaining a field data set of this larger magnitude.

The new procedure used USDA personnel in the counties to obtain the actual field data. This provided field observers who were more experienced and readily available. The participating agency was the Agricultural Stabilization and Conservation Service (ASCS) of the USDA. In addition to providing a larger data set, this method allowed some of the winter wheat sites to be surveyed twice during the growing season: to evaluate the LACIE technology for early-season as well as at-harvest wheat assessment.

Color-infrared aerial photography was taken for the new procedure by NASA aircraft at an altitude of 20 000 to 25 000 feet along two flight lines over each proposed site. The developed film was then screened and specific frames (usually about four per site) were selected to use for print enlargements of the site. After the prints were made, a frosted plastic overlay was attached to each print for use in recording field identification data. The overlay material was sufficiently transparent to show the field patterns on the aerial photographic prints and had a suitable texture to permit writing on the overlay.

The site boundaries were drawn on the overlays by JSC personnel and these prints, along with appropriate instructions and examples, were sent to the ASCS county offices in the areas where blind sites had been selected. The type of data recorded was a crop or land use code for each field or area within the site. A simple but uniform set of crop codes was used for all sites to simplify data interpretation when the prints and overlays were returned to JSC for analysis. In addition to identifying the crops within the site, each ASCS participant was asked to complete a two-page questionnaire containing a few comments about weather, insect, or disease conditions affecting the wheat crop within the site. Other items included were an estimate of the stage of wheat development at that date compared to "average" years and a section to identify any special crop or land use codes used on the overlay.

In Phase II (1975-76 crop year), there were 40 early-season blind site inventories throughout the Southern Great Plains states and 168 blind site inventories prior to or at wheat harvest. Thirty-seven of these late-season inventories were actually revisits to sites previously surveyed. This provided an indication of the final disposition of winter wheat fields identified in the early-season inventories. Some of these fields were plowed under, grazed by livestock, replanted to other crops, or allowed to mature to harvest depending on many factors such as stand quality of the field, farming practices, weather growing conditions, and economic conditions affecting wheat production.

For Phase III (1976-77 crop year), there were 67 early-season site inventories and 202 site inventories near harvest. Fifty of the late-season inventories were repeats of the earlier set of sites surveyed. Two changes or improvements were incorporated into the procedures for the Phase III blind sites. The first was to use the NASA RB57 aircraft to photograph the sites from an altitude of 50 000 to 60 000 feet. This permitted complete photographic coverage of a site on a single frame of film. This procedure decreased the number of print enlargements required and simplified the field survey by having only one print to annotate. As in Phase II, frosted plastic overlays were attached to the prints to be used by the field observers to note crop codes. The second change incorporated in Phase III was the selection of 15 wheat fields within each site and periodic reporting of their development status. These field reports were scheduled to correspond with the periodic (every 18 days) Landsat passes for the given site. Parameters

such as plant height, percent ground cover, and drill (or row) spacing were helpful in correlating and interpreting the Landsat imagery acquired throughout the growing season.

After receipt at JSC of the annotated print overlays and the questionnaires, these items were transmitted to the Accuracy Assessment group for preparation for analysis. This included verification of crop codes and outlining of the fields containing wheat or other crops of interest. The next major task was to planimeter the photographic overlays to determine the relative areas of different crop types within the site. The change in Phase III to the use of high-altitude photography and the resulting single print per site greatly reduced the manhours required for preparation and also resulted in a reduction of computational errors in the utilization of blind site data.

SUMMARY

The functions performed by the data acquisition subsystem during the three phases of LACIE supported the data needs of all other elements of the project. The nonelectronic data base consisted of statistical data, printed reports, periodicals, ground-observed data received from intensive test sites and operational segments, and full-frame multispectral scanner CIR photographs. Requirements were received as part of an overall project requirements document; they were written into an implementation plan; and finally they were satisfied within the subsystem or farmed out to be implemented by other units of the project.

Collecting, statusing, and providing the data required by LACIE was an enormous task. Requirements were constantly changing because of the dynamic nature of the project. The volume of data handled far exceeded initial estimates. The data provided were used to develop procedures to operate a system and subsequently test the results to determine whether it was possible to use remotely sensed data to inventory a crop (wheat).

ACKNOWLEDGMENTS

Data for LACIE came from many sources. We would like to express a special thanks to the following groups.

Intensive Test Site Data Collectors

Idaho
Bannock County ASCS Office
Franklin County ASCS Office
Oneida County ASCS Office

Indiana
Boone County ASCS Office
Madison County ASCS Office
Shelby County ASCS Office

Kansas
Ellis County ASCS Office
Finney County ASCS Office
Morton County ASCS Office
Rice County ASCS Office
Saline County ASCS Office

Minnesota
Polk County ASCS Office

Montana
Glacier County ASCS Office
Hill County ASCS Office
Liberty County ASCS Office
Toole County ASCS Office

North Dakota
Burke County ASCS Office
Divide County ASCS Office
Williams County ASCS Office

South Dakota
Hand County ASCS Office

Texas
Deaf Smith County ASCS Office
Oldham County ASCS Office
Randall County ASCS Office

Washington
Whitman County ASCS Office

**Operational Blind Site Data Collectors in the
ASCS county offices in the states of**

Colorado
Kansas
Minnesota
Montana
Nebraska
North Dakota
Oklahoma
South Dakota
Texas

**Cooperative Statistical Reporting Services from
the Foreign Agricultural Service and from the states
of**

Arizona
Arkansas
California
Colorado
Idaho
Illinois
Indiana
Iowa
Kansas
Minnesota
Missouri
Montana
Nebraska
North Dakota
Ohio
Oklahoma
South Dakota
Texas
Washington

LACIE Data-Handling Techniques

G. H. Waits^a

ABSTRACT

The data-handling techniques that were implemented to facilitate processing of Landsat multispectral data between 1975 and 1978 are described in this paper. The data that were handled during the LACIE and the storage mechanisms used for the various types of data are defined. The overall data flow, from placing the Landsat data orders through the actual analysis of the data set, is discussed. An overview of the status and tracking system that was developed and of the data base maintenance and operational task is provided. Finally, the archiving of the LACIE data is explained. The perspective gained by the adoption of this data-handling framework will be helpful in addressing these specific areas of system design in future applications.

INTRODUCTION

Until recently, far more Earth resources application data had been collected than could be practically managed and utilized in a cost-effective manner. During 1977, 17 000 Landsat acquisitions were arrayed in mass storage for LACIE. It is always possible to store information randomly or in arrival sequence and to retrieve it by an exhaustive search; however, the disadvantages are obvious. It is also possible to file all information orderly and search for it sequentially. Indeed, there was little other choice before direct-access memory was introduced. To fully exploit the potential value of the Landsat data collected every 18 days, the most rapid, cost-effective data-handling methods available must be used. Therefore, the LACIE data-handling system evolved from existing individual data-processing component systems used in varying remote-sensing disciplines.

^aLockheed Electronics Company, Inc., Systems and Services Division, Houston, Texas.

These various information constituents were modified, transformed, and integrated into the LACIE data-handling methodology.

The design of the LACIE data-handling system was predicated on the concept of man-machine interaction. The objective of the design was to provide the LACIE analyst a complete array of analysis and interpretation tools to interact with and operate on the available data. To fulfill this task, a subsystem was created to be responsible for the active collection, organization, storage, statusing, retrieval, and dissemination of remotely sensed data. Although the primary use of these data was in direct support of LACIE, the preservation of such data for future use by various secondary user groups has been ensured. This paper briefly outlines the nature of both the electronic and physical aids that were utilized to process Landsat data, the data-processing system developed to process these data, the interfaces in the use of the information, and the integrated information system to support the use of the data.

DATA DEFINITION

The LACIE data-handling system was developed to manage three basic forms of remotely sensed data: electronic data, physical data, and derived data.

The electronic data entered the system in Landsat multispectral digital format on nine-track computer-compatible tapes (CCT's). These data are input directly into disk storage to provide interactive display capability and mass storage of an entire crop year's Landsat acquisition history. The electronic data were primarily handled and statused automatically.

The physical data consisted of spacecraft imagery, aircraft photography, field observation data, crop calendars, topographic maps at several scales, and ancillary summary data. These data were manually handled and automatically statused. Much of the physical data was placed into LACIE segment

packets for Classification and Mensuration Subsystem (CAMS) analyst utilization. The field observation data and aircraft photography were provided to the Accuracy Assessment Subsystem for use in evaluating analyst estimations.

The derived data appeared as several types of computer printouts, microfiche, spectral aids, classification and cluster map film products, and CCT's. The final derived data were the wheat proportion estimates that were forwarded to the Crop Assessment Subsystem (CAS) for the generation of crop production reports. The Accuracy Assessment Subsystem was provided the classification and cluster tapes, batch run decks, and statistical results output tapes.

DATA STORAGE

The principal function of the LACIE data-handling system was to provide a contingent of information as required in a timely fashion, and to extract that information from the electronic and physical data repositories in an orderly and consistent manner. The implementation of this objective resulted in the establishment of an on-line mass disk data storage system to accommodate the electronic Landsat imagery and the creation of a 4000-square-foot LACIE Physical Data Library (LPDL) to manage and store the comprehensive physical data set.

The electronic storage capability is centered around an IBM System 360 Model 75J computer. The system is located in the Mission Control Center (MCC) and consists of equipment which was originally used for the Apollo lunar landing project. The core memory of the 360-75 is supplemented with 42 packs/drives of 7330 disk storage, providing direct-access storage for more than 4200 megabytes of data. The companion disk packs are removable and interchangeable between the 7330 disks. Each pack contains 11 disks with 20 recording surfaces, giving more than 100 megabytes of data storage.

Initial requirements for data storage were slightly shortsighted in that a clean and abrupt transition was expected for data acquisition and analysis from one crop-growing season to the next. Additionally, no provisions were made for maintaining data acquired from previous years on-line for research and development purposes. These data were available only from stored data tapes, thus increasing the com-

plexity and time involved in working with previous years' data sets.

This situation was rectified by expanding the disk storage space available and building a separate data base for each LACIE crop season. A research data base was also created so that data sets not maintained for operations would be available for analysis. These added flexibilities greatly enhanced ease of access to data, minimizing operational problems involved with processing such large quantities of data.

A specialized technical library, augmented by an automated status and tracking system, was established to store the LACIE physical data. The complexity of the job to be done, together with the huge volume of data to be handled and processed, required the adoption of a total systems approach and the automation of the LPDL. Approximately 3000 operational segment packets were maintained during the third phase of LACIE. Each segment packet contained Landsat segment film, crop calendars, topographic maps at several scales, and ancillary summary data. All packets were stored sequentially in filing cabinets with controlled access. The implementation of this facility is addressed in detail in another LACIE symposium paper.

FLOW OF DATA AND INFORMATION

It was within this framework that the LACIE data-handling system was planned and developed into its current integrated information and data-processing system. Perhaps the best way to examine the composition of this system is to follow incoming electronic and physical data through their processing cycle and then describe the tracking mechanism designed to identify each significant event. A visual depiction of the LACIE data flow is contained in figure 1.

Initial Data Order

Following the project selection and definition of the LACIE sample segment location, electronic data orders are placed with the NASA Goddard Space Flight Center (GSFC) via the data transmission line. The segment locations are defined by geographical coordinates at the center of the sample area. The Landsat acquisition date range—the beginning and

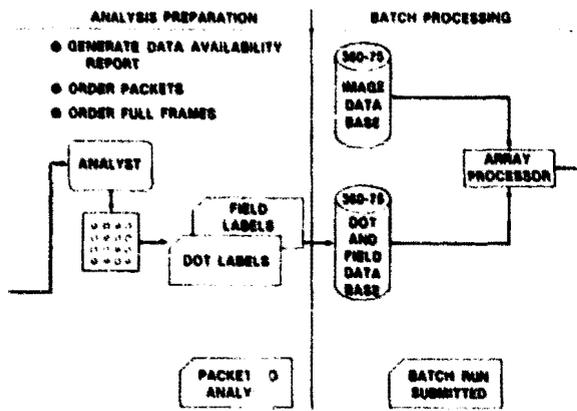


FIGURE 1.—LACIE data flow.

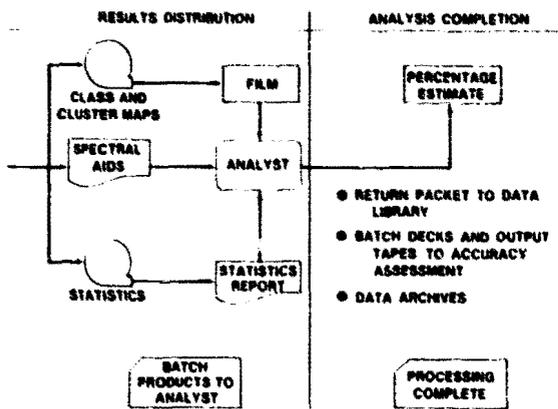


FIGURE 1.—Continued.

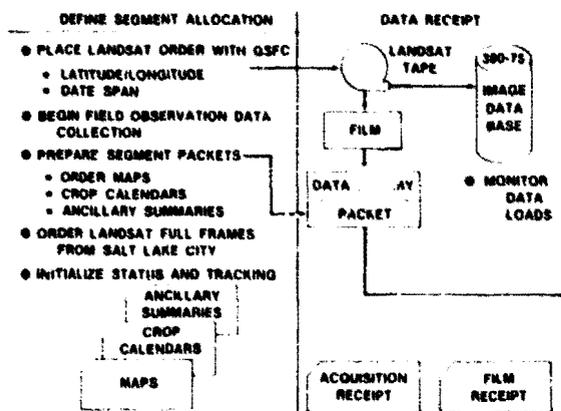


FIGURE 1.—Concluded.

the ending of collection periods—is also specified in the GSFC data order.

Concurrently, the segment number set is randomly distributed over the disk storage packs in anticipation of storing up to 16 Landsat acquisitions for each segment. Experience has shown that an average of 5 to 6 acquisitions can be expected for each Landsat data order placed with GSFC during one crop-growth year.

The supportive products such as topographic maps, crop calendars, and ancillary data are defined and ordered as soon as the sample segment location is specified. These supportive products and the Landsat acquisition film are placed in the LACIE analyst packets, which are stored in the LPDL until required for analysis.

Currently, the NASA Johnson Space Center (JSC) does not have the capability to generate full-frame 9- by 9-inch imagery from Landsat tapes. Since the synoptic view of the area surrounding the LACIE segment is important in making an accurate analysis, arrangements were made with GSFC to send the full-frame negatives to Salt Lake City for further processing. Therefore, the initial segment list was transmitted to Salt Lake City to support this film-processing effort for development of three-channel color infrared products.

Finally, the field observation data collection is initiated as soon as the "ground-truth" segments are identified within the total segment allocation. This task is accomplished by Agricultural Stabilization and Conservation Service (ASCS) and Canadian personnel. Crop inventory data are collected for these segments and forwarded to JSC.

Data Receipt

The Landsat data from GSFC are transmitted to JSC via a communication/image transmission line and recorded on nine-track magnetic tapes. All magnetic tapes received by JSC are first entered into the storage records and assigned a unique accession number. One to five tapes were received at JSC each day, with the Landsat data arranged in files, each file containing one sample segment acquisition. Each of these files consisted of header identification data; parameters used in film product generation; and the image data, which consisted of 117 scan lines with 196 pixels of data each in four spectral bands.

All acquisitions are entered into the image data base on the IBM 360-75. Subsequent acquisitions for the sample segments are stored on the same physical disk device but not necessarily in sequential order. Established indexes allow retrieval and composition of data consisting of up to four acquisitions of the same sample segment as required for application processing. A report of the stored data is automatically generated at the time of update, and queries concerning stored data may be generated at any time.

The Landsat data, on magnetic tape, are filmed on the Production Film Converter (a digital tape-to-film conversion device), which produces three different three-channel color transparencies for development by the Photographic Technology Laboratory at JSC. After development, the film roll is cut into individual products and packaged. All film processed in this manner is forwarded to the LPDL for inclusion in the LACIE analyst packets.

Preparation for Analysis

A status and tracking system (discussed later in this paper) provides a report of all electronic and physical data available to support an analysis. When a decision is made to analyze a particular region, a packet order list is generated requesting the LPDL to transfer analyst packets from storage to the analyst. After the analyst has received the packet, fields and dots are labeled and batch run decks are prepared by incorporating the labels into a sequential input processing deck. These cards are forwarded to the IBM 360-75 for interaction with the stored imagery and statistical analysis using a Staran array processor which is linked to the data base. Details of the classification and clustering processes are contained in other symposium papers.

Batch Processing and Results Distribution

Data classification runs (from batch or interactive processing) for area determination on the IBM 360-75 result in output tapes that are used to produce color transparencies of cluster and classification images. The images are generated on the Production Film Converter, as were the incoming Landsat image products in the "Data Receipt" phase. These batch or interactive jobs on the IBM 360-75 system also result in statistical report tapes (CAMS/CAS Interface Tape, CCIT) and microfiche. The CCIT is further

processed on a PDP 11-45 to provide the analyst with Type 1 and Type 2 dot label classifications, bias correction classification reports, Type 1 and Type 2 dot label cluster assignment, bias correction cluster reports, and separability reports. All these result products are forwarded to the analyst for evaluation. The Accuracy Assessment Subsystem receives all batch input decks and all output tapes after operations is through with them.

Analysis Completion

It should be clear by now that the LACIE analyst spends a great amount of time studying and working with the electronic, physical, and derived data to gather all the information necessary to produce an area estimate for the CAS to utilize. On completion of the analysis cycle, area proportion estimates are given to CAS and the analyst packet is returned to the LPDL.

INFORMATION STATUS AND TRACKING

A description of the information flow is not complete without a mention of the Automatic Status and Tracking System (ASATS). The ASATS, as the centralized source of information, is the hub around which the LACIE data revolves. It is built on the concept of a management tool to trace the flow of LACIE materials from collection and data storage through the various imagery interpretation/mensuration stages and finally to the compilation of a crop area and production estimate.

As the LACIE data collection and data bases increased in size, the simple sequential ordering of units was not adequate to effectively organize and report on the information within the system. A custom-built software package was designed to solve the problems unique to LACIE. The resulting system, ASATS, was maintained on a PDP 11-45 computer. This system is discussed in detail in another paper. The ASATS tracks Landsat data, from arrival at JSC to completion of the processing cycle, at various designated stations in the LACIE data flow previously discussed. Figure 1 indicates the status points within ASATS that were used to generate management reports, plan data-processing work, and track work assignments. In the earlier development of the system, the ASATS was a valuable aid in determining whether data were delinquent or lost; it

highlighted problem points in the data flow that were subsequently improved.

DATA BASE MAINTENANCE AND OPERATION

The receipt of all incoming Landsat data is monitored, all data updates are verified, and data storage levels on the imagery data bases are observed to maintain sufficient space for additional data. This monitoring is required since the storage system is configured to allow an average of 5 acquisitions per segment (up to a maximum of 16) to be stored. Overflow is captured on an overflow pack; however, the system monitoring is intended to prevent this occurrence. Data base restructuring or data purges are sometimes necessary as corrective measures. Data are periodically deleted from the system when quality is questionable. The entire data bases are checkpointed weekly to minimize recovery pro-

cedures in the event of data base failures. Thus, data base integrity is maintained at all times.

DATA ARCHIVES

At the completion of a project crop year, all electronic data resident in the operational data bases are unloaded on computer-compatible magnetic tapes. Copies of these tapes are sent to the Federal Archives and Records Center in Washington, D.C. Working copies are retained at JSC for further research and evaluation. A directory of all archived data is maintained for each LACIE phase or crop year.

The physical data remains in the LACIE segment packet as a historical reference tool for the analyst to use in subsequent crop years. If a particular segment is not used in the next crop year, the segment packet is removed from the operational data library and placed in an inactive data repository.

The Acquisition, Storage, and Dissemination of Landsat and Other LACIE Support Data

L. F. Abbotts^a and R. M. Nelson^a

INTRODUCTION

Background

The Data Research and Control (DR&C) Section of the Earth Observations Division (EOD) had been in existence for several years prior to the LACIE program but was configured to provide remote-sensing data support to projects of much lesser magnitude. Data available at the onset included aerial photography (mainly domestic, but serving as a good asset throughout the project to support research efforts); a full-frame Landsat and tape file started in 1972 with the launch of Landsat-1; a visual aid file; remote-sensing/Earth sciences reference collections; a map/chart acquisition and storage facility; and data to support projects such as Cornblight and the Crop Identification Technology Assessment for Remote Sensing (CITARS) compilations.

The LACIE support requirements, initially outlined in a number of baseline requirement documents, were eventually consolidated as part of a plan adjusted throughout the experiment. Because DR&C also supported other EOD projects, the LACIE Physical Data Library (LPDL) was established to separate LACIE-type tasks from other division support requirements.

Premise

With 4800 sample segment study sites located throughout the eight LACIE countries, it became apparent that the total volume of data for Landsat and correlated Intensive Test Site (ITS) areas would be greater than it had been for any other remote-sensing project undertaken by EOD. With Landsat data

being collected a number of times per site throughout the wheat growth year, it was apparent that such imagery could be overwhelming and could easily overtax the existing manual data-handling systems.

However, since the overall LACIE effort was in three phases, ranging from about 600 active segments in Phase I to nearly 3000 in Phase III, it was believed there would be ample time to adequately train personnel and develop efficient support systems. Adaptations were made without an excessive expenditure of manpower, time, or facilities as the experiment progressed.

In addition, field measurements data—consisting of aircraft remote-sensing and ground data collected during Landsat overpasses of ITS areas for research, test, and evaluation (RT&E) purposes—would be variable and intermittent. It was also expected that the rate of data input to RT&E (and subsequently to LPDL) would vary with the needs of their programs, including material for their supporting research and technology (SR&T) contracts.

LACIE DATA SUPPORT REQUIREMENTS

The overall LACIE requirement for physical data handling performed by LPDL is presented in figure 1. Many of these data are sample-segment dependent and require extensive handling and storage. Since the data were produced by other organizations, a basic function of the LPDL was to interface with these organizations, to serve as a central repository of data, and to transmit the data to LACIE users as needed. The general LPDL function was to research, acquire, index, maintain, distribute, track, and control LACIE operational data and documents. To execute this function, the LPDL was required to interface with LACIE users and NASA support organizations. It was represented during the daily LACIE Operations Coordination Center (OCC) meetings, where operations problems could be resolved in real time.

^aLockheed Electronics Company, Inc., Houston, Texas.

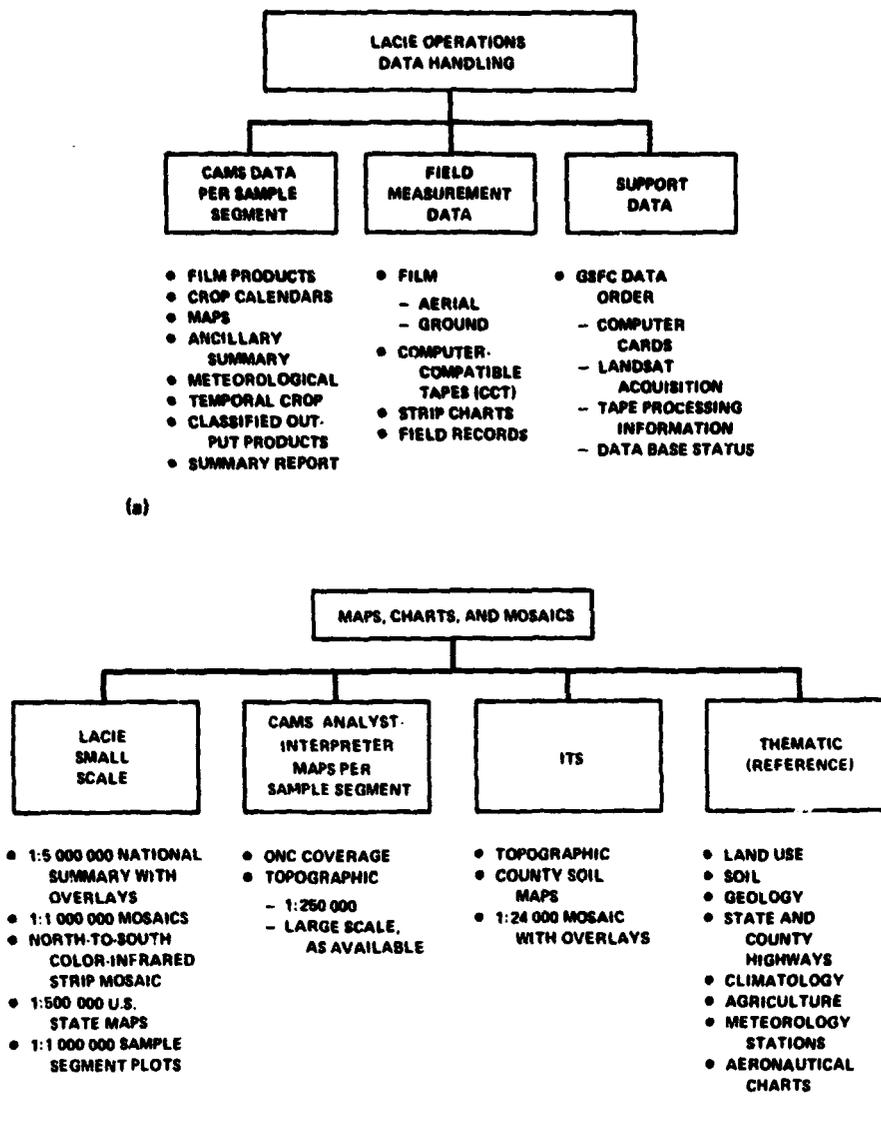


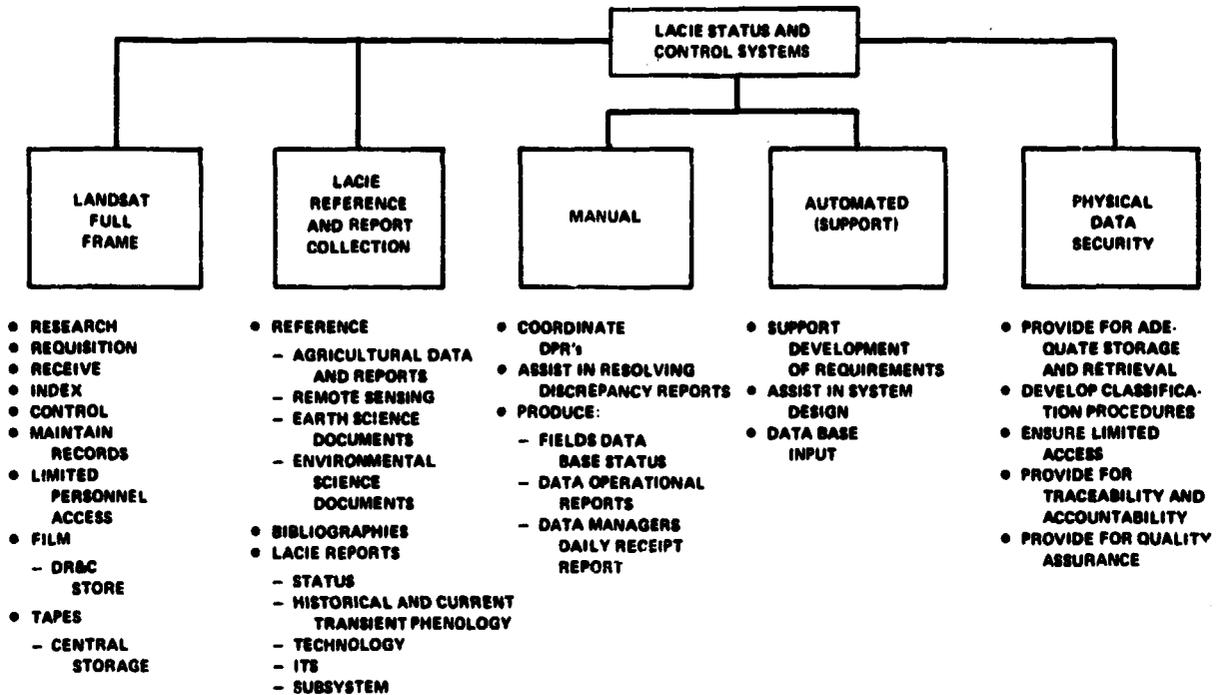
FIGURE 1.—LPDL requirements tasks. (a) Operations data handling. (b) Maps, charts, and mosaics. (c) Status and control systems.

SYSTEMS AND OPERATIONS

Small-size remote-sensing experiments often entail relatively small amounts of physical data that can be stored or filed by the experimenter in a couple of filing cabinet drawers. Large-size remote-sensing experiments such as LACIE, however, with its repeated coverage of many parts of the world extending over several years, produce a huge amount of data. Such systems need a formal management method and centralized data repositories with con-

trolled input/output mechanisms to make the material valuable to many users without excessive data duplication or loss of time. The data support systems developed for LACIE were designed to satisfy this need, in addition to meeting the other requirements.

During the Phase III peak level when approximately 3000 actual segments were being acquired by Landsat, a staff of 20 people was required to maintain the functions in support of the data depicted in figure 1. Housing that many people and providing



(c)

FIGURE 1.— Concluded.

facilities for staging and storing the data required more than 4000 square feet of office/warehouse space.

ANALYSIS PACKET PREPARATION AND STATUSING

Packet Development

The data required for the area mensuration analysis of a LACIE sample segment needed to be assembled, coordinated, stored, tracked, and retrieved conveniently. The method developed to solve the requirement was the Sample Segment Packet, one of which was prepared for each of the 4800 LACIE sample segment sites.

Data inserted in each packet by LPDL included LACIE sample segment film from Landsat, crop calendars, topographic maps at several scales, and ancillary summary data. The LPDL was also responsible for direct acquisition of available topographic maps that covered the LACIE sites.

Whenever there is a large volume of data—such as that associated with the 4800 sample segment packets (including the receiving, sorting, and organizing of thousands of pieces of paper and film)—there are bound to be problems. Therefore, documented procedures were established to maintain order, foster longevity, and handle problem areas. The packet materials were placed in sturdy, large envelope folders that contained sample segment numbers and indexing cards. The packets were stored sequentially in filing cabinets with controlled access and only checked out and in by authorized personnel.

All incoming production film converter imagery was screened on receipt and checked for problems. Problem photography was referred back to the production supervisors for refilming. Other data problems were referenced back to the sources for correction. The LPDL handled any map problems, and the Data Acquisition, Preprocessing, and Transmission Subsystem (DAPTS) was responsible for correcting problems with crop calendars and ancillary summary data.

Sample Segment Packet Status and Tracking System

The LPDL provided status and tracking data to the Automatic Status and Tracking System (ASATS). During packet development, a coded key-punched status card for each type of data (maps, crop calendars, and ancillary data) was submitted to ASATS immediately after such material was inserted in each packet. When a Landsat acquisition had been received and entered into the IBM 360-75 imagery data bases for storage, a status card indicating data availability at the NASA Johnson Space Center (JSC) was prepared and forwarded to ASATS.

After Landsat film had been received and inserted in the packet, a film-ready status card was entered in ASATS. When the packet was released from ancillary hold and film was received, LPDL prepared another coded status card indicating that all necessary data had been placed within the packet and the sample segment was ready for analysis. These status and tracking steps were repeated throughout the LACIE program.

Data Coordination and Reporting

The LPDL coordinated the preparation and tracking of Data Product Requests (DPR's) for LACIE. DPR's prepared by the analysts were checked for completeness, recorded, approved, and forwarded. The returned data was checked for completeness, recorded, and forwarded to the requester. DPR's ranged from requests for Landsat film products to data base queries, keypunch requests, and special requests for data search. During Phase III, these requests averaged over 30 per week.

Packet Operations and Statusing

Prior to the start of Phase I analysis, sample segment lists were used to create the segment packets and to generate the process of ordering maps. As maps, crop calendars, ancillary summary data, and film were received and filed, the sample segment information was statused through manual reports and inputs to the Interim Status and Tracking System (ISATS).

During Phase I, 1033 segments were analyzed, while 2649 Landsat film sample segment acquisition sets were received. Each initial film data set con-

sisted of two color composite film transparencies and four black and white film transparencies, each of which had to be filed in the packets and statused ready for analysis. Numerous problems emerged during Phase I. While most were solved, several persisted due to time pressures and the need for rapid responses. These problems are listed below.

1. Sample segment lists (based on the sampling strategy output) were often not received in sufficient time to search, order, status, and file maps in the packets without a "crash" program. For example, it required 1 to 3 months after placing an order to receive maps from Canada; 1 to 2 months to get maps from the Defense Mapping Agency Topographic Center; 2 months from Australia; and 1 month for large map quantities from the U.S. Geological Survey. These time intervals include first-class/airmail delivery in the United States and airfreight shipments from foreign countries. This situation persisted and was aggravated each time there was a major relocation of segments.

2. Initial packet handling required that packets ready for analysis be removed from the sequential files and placed in separate cabinets in order to satisfy user needs to examine packet contents and to plan daily workloads prior to checkout. In addition, packets with data problems were placed in another separate cabinet. This method was operationally satisfactory to users but was disruptive to the regular data handling process, causing a high expenditure of manpower time.

3. All types and formats of data (e.g., computer printouts, detailed processing results, etc.) were saved, causing bulky packets and file crowding. Direction regarding retention time of such data was difficult to obtain as there was often a lack of consensus concerning the longevity value of the data. This problem persisted and in fact may be the unfortunate byproduct of any large-scale experimental/quasi-operational system that handles large and varied quantities of data.

4. At the start of Phase I, input data cards on the required data and packet status were created through manual coding and punching. Batch card inputs to overnight ASATS updates on status had operational problems and breakdowns. Therefore, critical reports on packet data availability and statistics on data flow and time lines often were prepared through a time-consuming manual process. Since LPDL provided all these data and essentially performed most of the tasks for ASATS, as well as maintaining the manual systems and records, manpower requirements for re-

puting were exceptionally high. This situation was eventually solved as the ASATS reporting capabilities became reliable.

In Phase II, the major emphasis was to increase and diversify the number of segments and Landsat data acquisitions to be analyzed. Segments increased to 1437 and Landsat data acquisitions to 9211. The color composite transparencies and black and white products remained the same.

By Phase II, computer-generated packet-ready lists were provided through ASATS, thus eliminating the need to physically shift packets from cabinet to cabinet. Film was sorted and placed in annotated envelopes, making it easier for Classification and Measurement Subsystem (CAMS) analysts to organize these products sequentially. Early in Phase II, data and packet availability cards were still prepared manually. Initially, the batch card input to perform overnight update, statusing, and daily output had low dependability, and manually recorded data were utilized as necessary. However, as problems with system dependability were solved, the ASATS became more stable.

In Phase III, there was a significant increase in segments to be analyzed. The number of color composites per site increased to three but the black and white transparencies were dropped. In addition, some sample segments were dropped; some were relocated; and new segments were added in keeping with a new sampling strategy. Analyzed segments increased to 3006 and Landsat sample segment acquisition sets to 19 000. The changes and the additional volume began to test the full data handling capability of the entire LACIE system, including the LPDL.

Every effort was made to obtain the sample segment relocations lists as soon as possible, since each relocated or added segment had to be plotted on a map index to determine the new map coverage (for ordering purposes). This effort required the full resources of LPDL to prepare the requisitions and provide expeditious handling on arrival so that maps could be available to assist in the analysis effort.

The CAMS operational procedures changed, incorporating Procedure I. Bulky computer printout materials retained in the packets caused packet storage problems, so that file drawers normally storing 25 packets now could only accommodate 10 to 12. This problem was solved by removing an operational packet from the active staging area as it was completed for the season and storing it with other cyclical records.

Machine-generated film labels received via

ASATS for the film product envelopes were also placed on the plastic sleeve used to protect each piece of film. The use of the computer-generated stickyback labels saved considerable time in the process of labeling and statusing incoming film products.

Prepunched packet data-statusing cards were now prepared in advance for all new or relocated segments. As the segment listings were entered into ASATS, these cards were prepunched except for date. The date was entered automatically as the cards were forwarded by LPDL to indicate that the data had been received. Thus, while the data rates nearly doubled from Phase II to Phase III, the ever-increasing automation of the status and tracking system often offset to some degree the impact of the increased workload.

LANDSAT FULL-FRAME SYSTEM

In 1972, a Landsat full-frame data acquisition and storage system was initiated by EOD to support remote-sensing projects. The film was retained for operational use, while the tapes (after initial analysis) were placed within a centralized tape library for operational storage and retrieval.

Early in the program, LACIE management stated the need for a coordinated Landsat full-frame film file with adequate status and tracking procedures. This file was established as part of LPDL, which developed a method for coordinating full-frame data acquisition and management.

Corrected full-frame 9.5-inch-format film imagery was stored from both Landsat-1 and Landsat-2. The imagery from the multispectral scanner (MSS) often comprised all four bands in black and white transparencies and/or prints and in various color combinations. Since a limited quantity of return beam vidicon (RBV) data was acquired, there were only a few selected frames in the file. The digital full-frame Landsat data was received in the form of nine-track computer-compatible tapes.

A Landsat path and row (footprint) indexing system was adopted, with data being filed in standard filing cabinets and grouped sequentially by acquisition date. An index card was prepared for each image, and film could be checked out only by authorized personnel. Manual reporting procedures were developed to provide the LACIE OCC information on the status of full-frame film.

When LACIE I became operational, the file was

composed of about 15 000 pieces of film, mostly black and white transparencies of the four MSS bands. During LACIE, approximately 9200 color composite acquisitions were received, producing a total of 25 000 pieces of film in the full-frame inventory.

MAP SEARCH AND ACQUISITION

Maps and charts depicting most of the basic land-associated themes that covered LACIE areas in the eight countries were researched and acquired, where possible; see figure 1 for a summary list. These maps were obtained from U.S. government agencies and commercial publishing organizations. In addition,

LPDL was the repository for specialized LACIE maps, mosaics, and overlays created by the EOD Cartographic Laboratory.

However, a major effort was necessary to search and acquire topographic maps to cover the 4800 LACIE sites in eight countries. The map management process is illustrated in figure 2. Map orders were sent to a U.S. Geological Survey map distribution center for large- and medium-scale U.S. maps; the Defense Mapping Agency Topographic Center for other domestic and foreign maps; and a foreign cartographic agency for nationally printed maps.

Each segment packet was provided a 1:1 000 000 Operational Navigation Chart (ONC), a 1:250 000 topographic map, and large-scale maps. U.S. large-scale maps were 1:62 500 (15-foot quadrangle) and

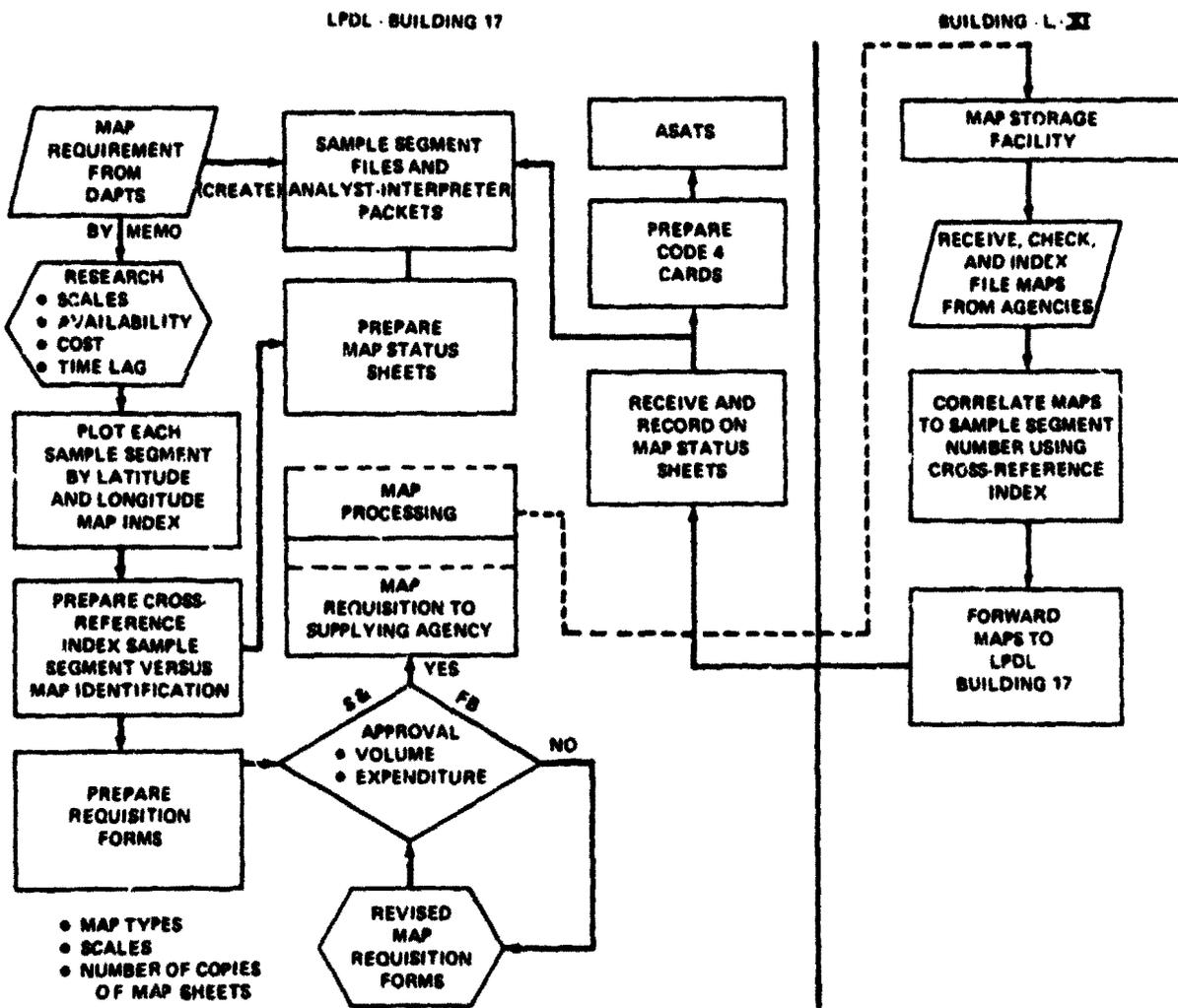


FIGURE 2.— DR&C map management for LACIE.

1:24 000 (7.5-foot quadrangle) scale maps, as available. In foreign areas, they were 1:100 000 and 1:50 000 scale maps, as available. LPDL created a map statusing sheet as a device to ensure that when the maps were received the correct combination would be assembled, numbered, folded, and inserted in the appropriate packet.

As described in the Landsat systems design paper, the best the Goddard Space Flight Center (GSFC)/Landsat acquisition system could guarantee was that the sample segments were located within a nominal 10- by 11-mile area. Therefore, to ensure adequate map coverage, each segment was provided maps that covered a 10-mile-diameter circle centered at the sample segment coordinate point, as plotted on a map index. Consequently, if the coordinates of a U.S. segment were located near map boundaries, the packet would require up to four 1:1 000 000 scale maps, four 1:250 000, four 1:62 500, and nine 1:24 000. After the segments were sited, the unused maps were returned and stored. Later in the program, the 1:1 000 000 scale ONC maps were removed from the packets and used in a separate file for sample segment plots.

It was necessary to acquire a sufficient number of these maps so that not only each packet had one copy of the appropriate map sheet but that there were also a few spares in the map storage facility. In time, maps were cut up or lost and replacements were needed; because a few spares were available, a great deal of time and labor were saved.

The Map Acquisition and Storage Facility prepared an index card for each map, mosaic, and overlay received; card data included the number of copies and where the map was used. During the LACIE period, the facility increased from approximately an 11 000 index card, 30 000 map file to 34 000 index cards and 90 000 maps, charts, mosaics, and overlays. A semiautomated information retrieval system was utilized as a manual assist, thus permitting one person to manage the indexing process. The system reduced the time required to record data and retrieve information, but it did not have the capability of producing map listings.

LACIE REFERENCE/PROJECT DATA AND REPORTS

The LACIE Reference/Project Data and Reports was subdivided into four areas: (a) LACIE reference and report collections, (b) LACIE ITS data, (c) con-

tingency data, and (d) records storage. These areas are shown in greater detail in figure 3.

The LPDL interfaced with the users and received data and documents on a daily basis. It also provided data, reference materials, bibliographical information, and reports for project use. Over 60 000 items (excluding maps) were received, filed, and managed in this part of the LPDL operation during the LACIE program. Information requests for this data varied from 50 to 100 per week.

Reference and Report Collections for LACIE

Reference information was composed of agricultural and other Earth/environmental science texts, statistics, reports, remote-sensing documents, and data. Most agricultural reference material, apart from basic information sources, was received either from the statistical service offices of the federal government or directly from the states. Specific LACIE project-required data was received via DAPTS transmittal reports.

The report collection consisted principally of reports generated by LACIE management and the subsystems. In addition, data and information was supplied to offsite technical investigation centers under NASA SR&T contracts. Reports from these investigations also became part of the LACIE report collection.

The DR&C/LPDL regularly used a terminal connected to a Remote Console (recon) that provided a capability to research bibliographical data from the several library systems at GSFC. Additional bibliographical data were obtained from other library sources across the country.

Special LACIE Data

Data covering LACIE ITS and three "supersites" were received throughout the LACIE program; the data input reached a peak during Phase III. These data were composed of two basic types: field measurements data collected over the test sites (comprising only basic preprocessing, such as film development), and data produced from secondary or tertiary processing or development (such as a color land use map overlay). See figure 3. Suitable storage and retrieval systems were established for each type of data. One of the larger files was created for the 35-mm slides (10 325 images) depicting crop phenology in the various test fields during all the LACIE phases.

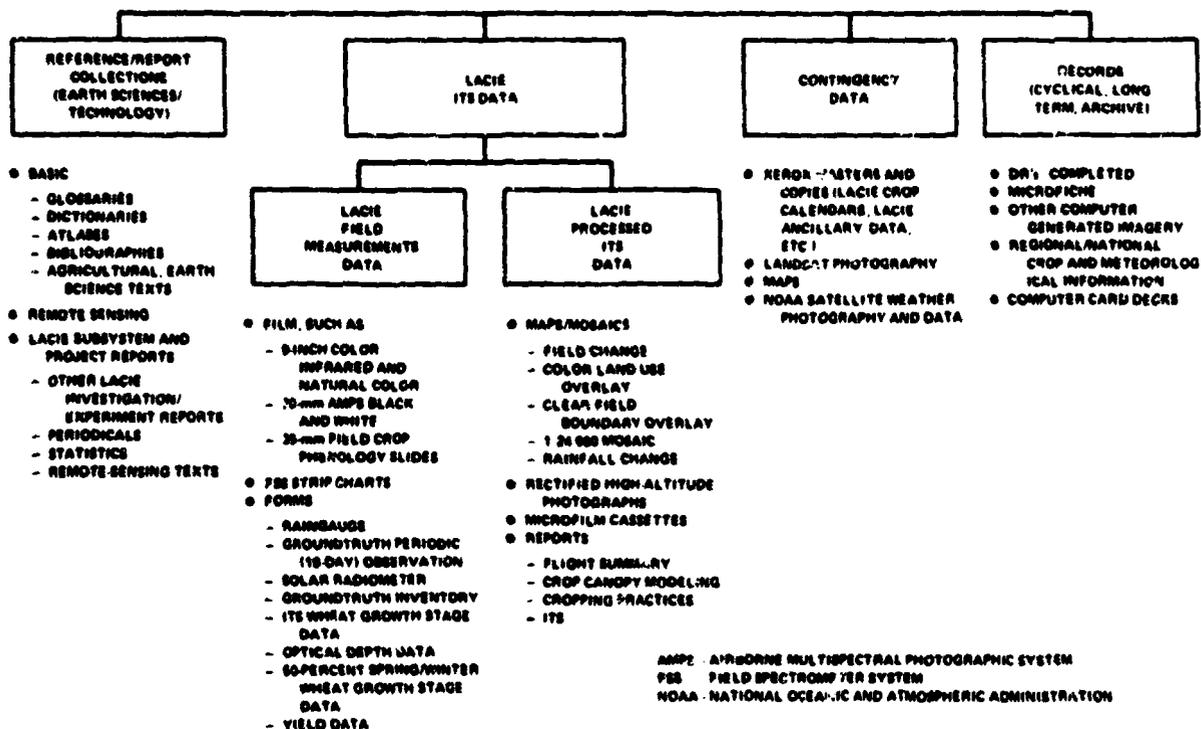


FIGURE 3.— LACIE Reference/Project Data and Reports.

Contingency Data

Contingency data comprised all classes of material for which there was a rapid response requirement. Such data was located, organized (often copied), and filed so that short-time retrieval action could be instituted. This concept was based on providing rapid support to the basic data flow of the LACIE subsystems, such as CAMS, the Crop Assessment Subsystem (CAS), and the Yield Estimation Subsystem (YES), and to the support elements, such as ASATS. The specific methodology varied with data base size and allowable response delay. For example, copies of crop calendars and ancillary data were located in a room next to the packet operations so that a replacement was almost instantly available if a crop calendar or ancillary summary were lost. However, the map storage facility, because of its size and space requirements, was located offsite. To replace a map (assuming a duplicate was available) required less than half a day. Most data, such as LACIE reference data, was stored so that retrieval could be completed within 24 hours.

Records

Records storage and management was created separately for data having a cyclical or long-term need and for data that had completed their usefulness and were archived. Cyclical or long-term data could be made readily available based on the advanced planning and timing need per project; e.g., for the start of each LACIE phase, for a particular time period within a phase, or to support an SR&T contract. Depending on priorities, available personnel, and data volume, most records-type data could be obtained from the files in 1 to 2 days.

Most ITS data were categorized as records and stored offsite due to lack of onsite space. Retrieval action for some of the data, such as the 35-mm slides stored in LPDL, was almost instantaneous. However, for data that required processing, such as copies of aerial photography, the timing varied with the processing cycle.

Data placed in an archive were basically "used" data, such as a card deck for a no-longer-used computer program. Such data were resurrected only on

special request and had a low retrieval priority. The data, however, were stored for the life of the project.

OTHER SUPPORT SERVICES

The DR&C performed other standard support services throughout LACIE. Of particular value was a complete worldwide microfilm file of Landsat data utilized intermittently in conjunction with changing priorities or countries. The expertise gained in day-to-day use was employed as required throughout the experiment.

CONCLUSIONS AND RECOMMENDATIONS

The management of data is an important element in any remote-sensing experiment or operational program. It requires appropriate considerations in system planning and development, and suitable decisions should be made early in a program so that data will be readily available when needed.

During the course of the LACIE program, several persistent data handling problems were evident. These problems were as follows.

1. Delays in receiving requirements, which lead to a lack of sufficient implementation time to supply data when needed (e.g., delays in getting segment lists so that maps could be ordered and received in time for analysis without a crash program)

2. Inadequate user interface plans and requirements, so that when data were received they could be indexed, stored, and retrieved properly to satisfy specific users

3. Difficulty in obtaining system management plans and requirements for the retention and distribution of data. LACIE follow-on planning, with the definition of supporting data successive systems requirements, has demonstrated the need for initial and updated planning for data retention. For example, the blind site data for Phase II operations, which was utilized later in the program, was essentially lost because of inadequate plans. However, the interface between requirements and what can be afforded in the way of physical storage will continue to present problems to program planners. Data generated by a project the size of LACIE is tremendous, and storage space and associated personnel requirements exceeded the capability of providing complete storage for all data.

Such problems can be traced back in part to the lack of initial input into LACIE baseline requirements documents or into a LACIE integrated implementation plan, where not only qualitative tasks but also data volume, rate of input, diversification, and timing are significant considerations. Therefore, it is recommended that for future programs, provisions for physical data management be made an early integral function of system development.

THE FUTURE

The data collected and stored as a result of the LACIE program, coupled with the data initially available to support LACIE, constitute a valuable data collection structured to support operational and experimental remote-sensing programs. Much of the data can be incorporated into an Earth resources data base. Elements of the data collection that could support future remote-sensing programs include the following.

1. The Landsat full-frame image files
2. The microfilm file of aerial and space photographic and multispectral scanner data that encompasses a large portion of the Earth's surface
3. The map/chart collection that includes various scale maps and charts for a good portion of the United States and the LACIE area in foreign countries
4. Computer-compatible tapes of good quality Landsat scenes particularly adaptable to agricultural application site research
5. A collection of basic remote-sensing data, project data, reference material, and associated publications
6. Visual aids that can be used in part to support presentations on remote-sensing projects
7. Research acquisition and handling procedures for managing data for a high-density remote-sensing program that will be applicable to future programs

The Classification and Mensuration Subsystem

K. M. Abotteen^a and R. M. Bizzell^b

INTRODUCTION

The Classification and Mensuration Subsystem (CAMS) was responsible for the acreage component of the wheat production estimates produced by LACIE. The wheat acreage for a region or a country was produced from the individual wheat proportion estimates of 5- by 6-nautical-mile sample segments using Landsat imagery and supporting historical data. To accomplish this task, CAMS implemented a processing system to respond to both the accuracy and throughput requirements of LACIE.

From an operational standpoint, the most significant item CAMS had to overcome was the scope; i.e., segment volume processing requirements. The obvious conclusion from a review of the requirements, listed in the following chart, shows that a significant increase in data handling and processing was necessary.

	Phase I	Phase II	Phase III
Total segments	700	1700	3 000
Acquisitions received	2000	9000	18 000
Peak processing requirements per day	16 to 20	35 to 40	75 to 80

In Phase I, the state-of-the-art classification technology was assembled into a machine-processing system capable of handling the large volume of data required to evaluate and improve the technology. The design of the initial system was simplified to allow for subsequent modifications with minimal impact. A significant portion of the operational elements (e.g., data handling, computer card deck generation, etc.) was accomplished manually. The classification technology implemented consisted of classifying Gaussian maximum likelihood per pic-

ture element (pixel) from defined training fields and their associated statistics. However, one major exception existed: the definition, identification, and labeling of training fields were accomplished without the benefit of ground observations. Never before had such a task been attempted on the basis of analyst-labeled satellite imagery. Thus, a key element of CAMS was the development of consistent and accurate labeling and analysis procedures that used Landsat and supporting data in a high-volume, high-throughput environment.

By exercising the first-generation technology in Phase I, CAMS personnel identified several key issues. During Phase II, answers to many of these technical questions evolved. The interrelationships between man and machine, technology and operations, and accuracy and throughput started to become clearer. Thus, a significant design effort was initiated in parallel with the Phase II operations to define an improved technology. The result of these efforts was the design of an analysis approach called Procedure 1. (Procedure 1 is the subject of the paper by Heydorn entitled "Classification and Mensuration Approach of LACIE Segments.")

An experimental design to test and evaluate Procedure 1 was conducted during the latter stages of Phase II. When these tests showed positive results, the tasks necessary to implement Procedure 1 (e.g., software modification, procedures development, analyst training, etc.) were initiated and continued through the initial Phase III processing period for winter wheat (fall-winter, 1976-77).

The implementation of Procedure 1 into Phase III operations was accomplished in two stages. First, a concept implemented through analyst procedures with minimal software changes was utilized operationally in the processing of spring 1977 winter wheat. This period was used to accelerate analyst training, final system debugging, and on-line testing. Finally, beginning with Phase III spring wheat processing (June 1977), the "full-up" Procedure 1 was implemented operationally.

^aLockheed Electronics Company, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

CAMS OPERATIONS

The training backgrounds of the analysts, the available data, the labeling logic, the analysis procedures, and the overall integration of these factors into the large-scale LACIE environment are described in the following sections.

Analyst Training Background

Varied backgrounds among the analysts were required to ensure that all aspects of LACIE were adequately covered—photograph interpretation, geography, agronomy, mathematics, statistics, and computer science. The analysts were extensively trained in image interpretation techniques, photographic film production, pattern recognition theory, applied statistical techniques, and available data analysis systems. Figure 1 depicts the concept of the operational system utilized by the analyst and the interactions of the various functions.

Available Reference Data

Reference data (e.g., data on what crops to expect in a given area and what growth stages to expect for the particular acquisition(s) being analyzed) were available to the analyst. Imagery (film products) provided the analyst with spatial and spectral information. A machine-processing system took the analyst's input, classified the total segment, and generated output products (classification maps, cluster maps, classification summaries, etc.) for evaluating the processing results.

For every sample segment, the analyst had a packet that contained imagery (film products), maps, ancillary data, and previous machine classification data. Available reference materials not included in the packet were weekly meteorological summaries, full-frame imagery, and analyst interpretation keys.

Film products.—Imagery for the segment includes product 1 of all acquisitions during the past crop year, if collected; and products 1, 2, and 3 for each acquisition in the current crop year. LACIE product 1

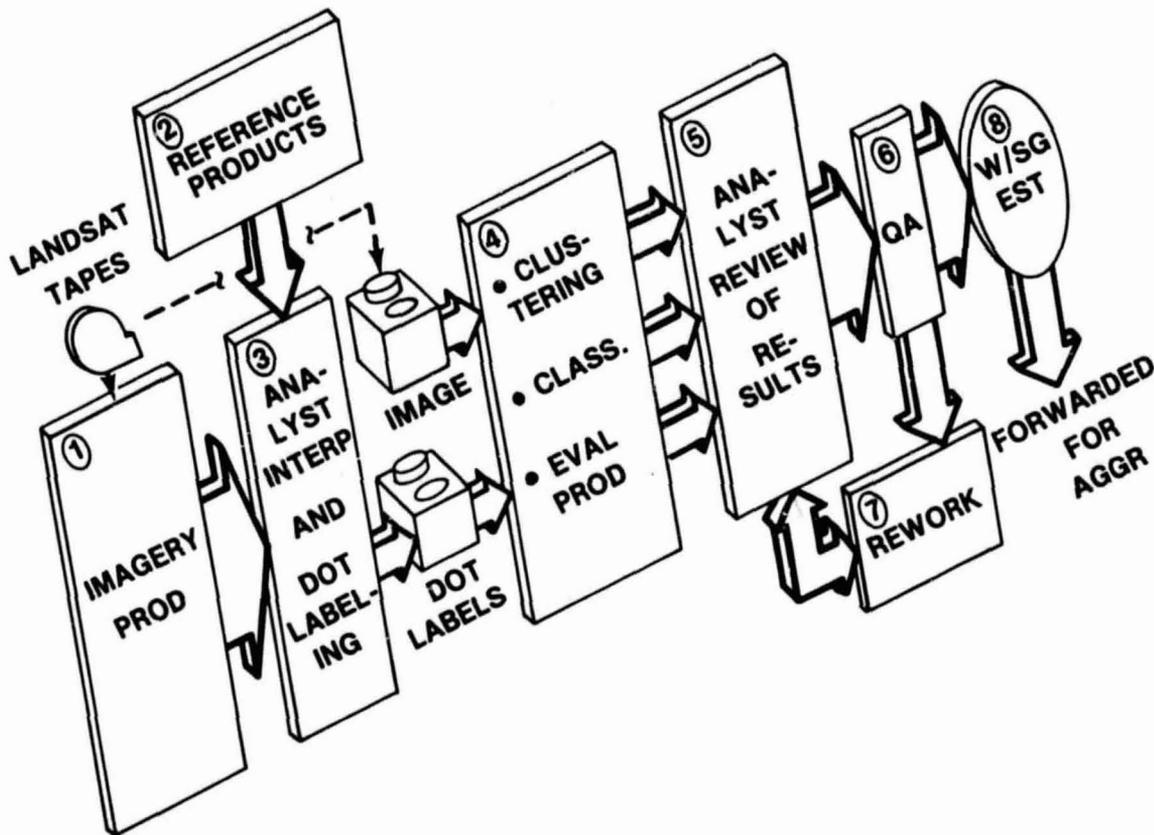


FIGURE 1.—LACIE Procedure 1 — areal estimation procedure.

(or simulated false-color imagery), shown in figure 2(a), is created from digital values in channels 1, 2, and 4; color assignments are blue, green, and red, respectively. LACIE product 2 (or positive-negative imagery), shown in figure 2(b), is created from values in channels 2, 3, and 4 to make information in channel 3 available to the analyst. Color assignments in this product are red for channel 2, blue for channel 3, and green for channel 4; polarities are reversed for channels 3 and 4. Products 1 and 2, generated to emphasize contrast, are excellent for field delineation and enhanced spatial features. However, depending on the data in the scene, contrast is sometimes achieved at the expense of consistent color depiction of spectral values. LACIE product 3, shown in figure 2(c), is imagery developed for Phase III specifically for a more consistent color display of spectral signatures. (The magnitude of color distortion varies from scene to scene; the range of possible distortion is displayed in figure 3.) Product 4 (black-and-white images of each channel) is used for detecting data dropout problems etc.

Maps.—Maps at scales of 1:24 000 and 1:250 000 were useful in identifying topographic features, unusual signatures, and natural vegetation; they also ensured the proper location of segments.

Ancillary data.—Ancillary data include information on cropping practices and soils for the crop-reporting district (CRD), historical crop percentages for the political subdivision for the preceding 4 or 5 years, and a nominal crop calendar for the CRD based on 10-year averages. Sample nominal and adjusted nominal crop calendars for CRD 20 (Montana) are given in figures 4(a) and 4(b), respectively. Average length of crop development stages, the nominal dates when these stages occur, and the relative growth stages of other crops in the area are presented.

Machine classification data.—Machine classification data include previous wheat acreage estimates on the segment and classification products for the most recent estimate.

Available reference materials not included in the analyst's packet are the following.

Weekly meteorological summaries.—Weekly meteorological summaries¹ provide current data on tem-

¹A weekly meteorological summary of the United States is published as an internal memorandum by the National Oceanic and Atmospheric Administration (NOAA), the U.S. Department of Agriculture (USDA), and the National Aeronautics and Space Administration (NASA). LACIE personnel had limited access to these summaries.

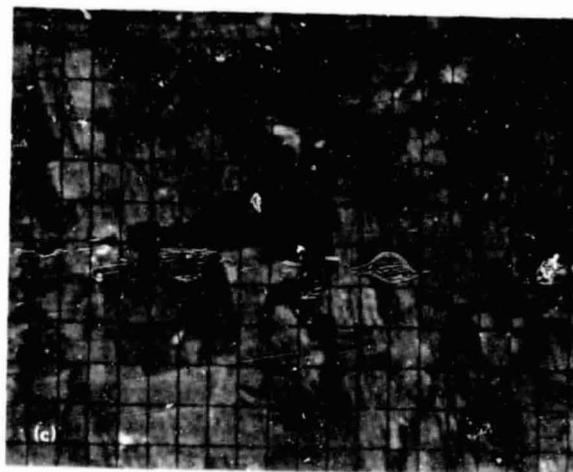
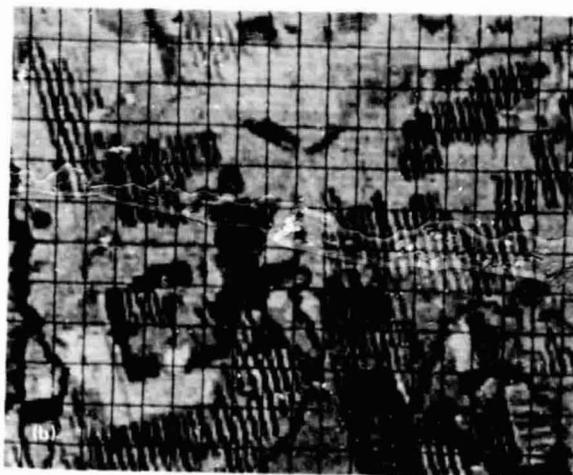


FIGURE 2.—Landsat film products for Blaine County, Montana, acquired on July 3, 1977 (from ref. 1). (a) Product 1 created from digital values in channels 1, 2, and 4. (b) Product 2 created from digital values in channels 2, 3, and 4. (c) Product 3 created from digital values in channels 1, 2, and 4.

ORIGINAL



MODIFIED

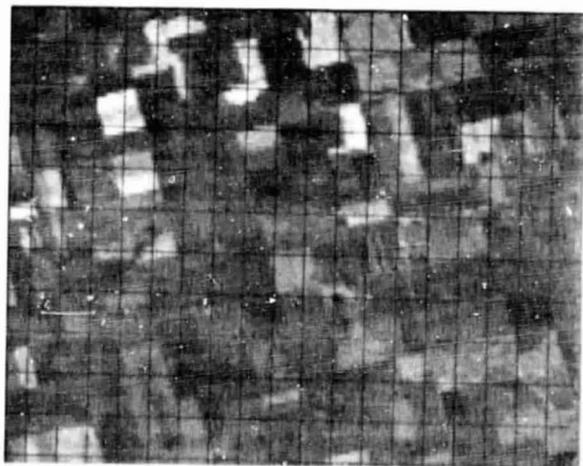
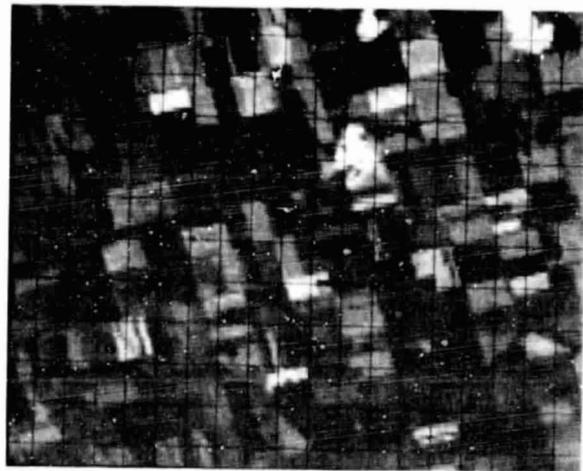
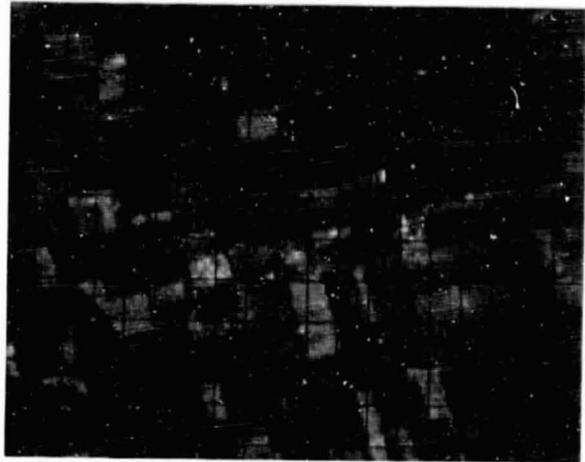


FIGURE 3.—Comparison of Products 1 and 3 (original versus modified version). (a) June 7, 1977. (b) July 31, 1977. (c) August 18, 1977.

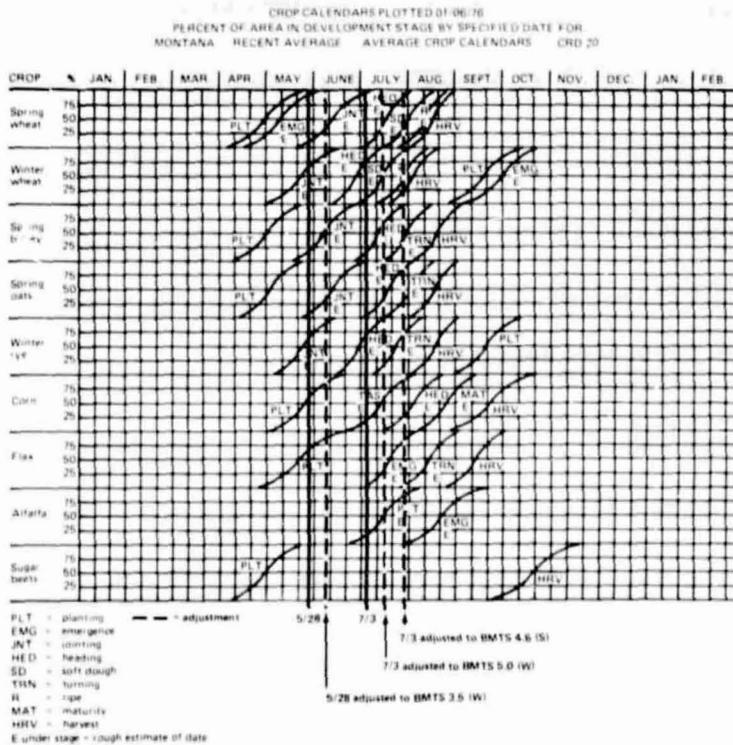
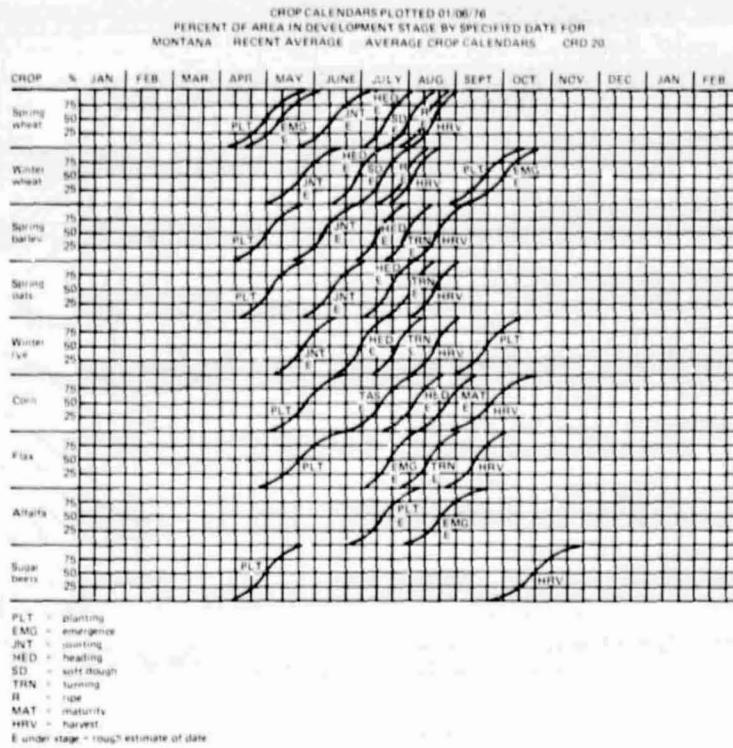


FIGURE 4.—Nominal and adjusted nominal crop calendars for CRD 20 (Montana) (from ref. 1). (a) Nominal (historical) crop calendar. (b) Nominal crop calendar adjusted for crop year 1976-77 (segment 1528).

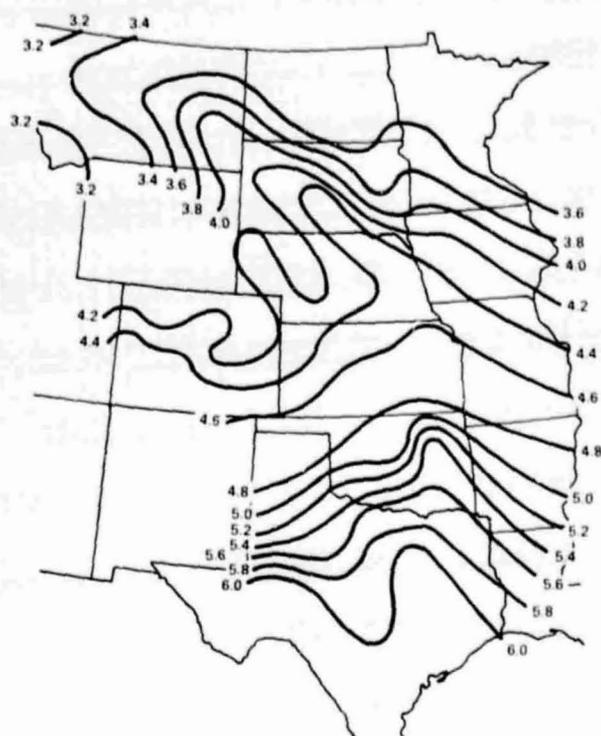
perature and precipitation, a crop calendar adjustment reflecting the current Robertson biometeorological time scale (BMTS) discussed in reference 2, the growth stage of wheat, and a summary of statewide crop and weather assessments.

By correlating the current year's specific growth stages found on the crop calendar adjustment (fig. 5) from the weekly meteorological summary, the analyst adjusts the nominal crop calendar to be more specific for the current year as shown in figure 4(b). This information can be very useful because episodic events, such as recent rainfall, alter crop signatures. Soil reflectivity contributes an indeterminate component to the average reflectance value for an acre recorded by Landsat, and the reflectivity of wet soil differs from that of dry soil. Long-term events, such as drought, affect the expected Robertson BMTS for wheat and the expected spectral signatures.

Landsat full-frame imagery.—Full-frame imagery of all areas in which LACIE segments exist are provided to the NASA Johnson Space Center (JSC) four times during the growing season for analyst use. To fulfill this requirement, cloud cover must be less than 20 percent when Landsat passes over an area in order to acquire usable imagery. Figure 6 is a full-frame image of sample segment 1528 in Blaine County, Montana, acquired on July 3, 1977.

The coverage of the full frame can be used for better distinction between agricultural and non-agricultural patterns and signatures within the segment area and in the area surrounding the segment. Drainage patterns, streams, and areas of natural vegetation, such as rangeland in the U.S. Great Plains, are frequently easier to identify on full-frame imagery. Knowing the relationship of the segment to the county through analysis of the full-frame image and agricultural statistics can be very important to the analyst. The agricultural statistics in the ancillary summary can be understood better when the political subdivisions to which they apply (e.g., counties in the United States) and the segment are viewed together after plotting on the full-frame image. In this manner, the analyst can observe how the segment compares to the remainder of the county with respect to the proportion of agricultural land in the county and in the segment. Certain crops may be grown in some areas of a county and not in others. For example, different crops may be grown along rivers and streams rather than on drier hillsides with poorer soil.

Analyst Interpretation Keys.—Two volumes of Analyst Interpretation Keys (ref. 3) were available to



EACH GROWTH STAGE BEGINS WHEN APPROXIMATELY 50 PERCENT OF THE CROP HAS REACHED ONE OF THE FOLLOWING NUMBERED STAGES.

GROWTH STAGE	DESCRIPTION	GROWTH STAGE	DESCRIPTION
1.0	PLANTING	4.0	HEADING
2.0	EMERGENCE	5.0	SOFT DOUGH
	DORMANCY (WINTER WHEAT)	6.0	RIPE
	SPRING GROWTH (WINTER WHEAT)	7.0	HARVEST
3.0	JOINTING (GREENING UP PERIOD)		

FIGURE 5.—Example of crop calendar adjustment from the weekly meteorological summary (from ref. 1). The numbers refer to the Robertson BMTS growth stages for wheat.

the analysts in Phases II and III. Volume I is an operational overview of wheat and nonwheat signatures. Examples are relatively general, and nominal photophenology for wheat is illustrated. In addition, volume I gives examples of, and causes for, some of the variations in wheat signatures seen on Landsat imagery. It is used as a general training and information aid. Volume II is a regional key of Canada and the U.S. Great Plains and is used by production analysts as a guide in operations. It is designed to lead to the correct identification of wheat and small-grains areas.

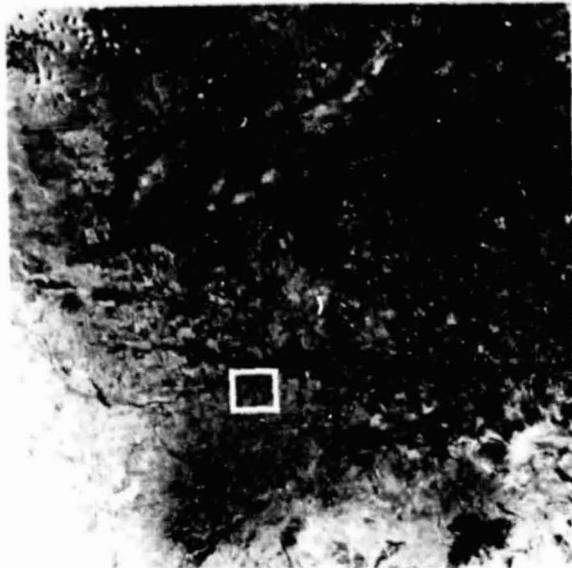


FIGURE 6.—Full-frame image of sample segment 1528 in Blaine County, Montana, acquired on July 3, 1977.

Labeling Logic

Plant phenology is indicated by a temporal change in the infrared reflectance relative to the reflectance in the visible range. Figure 7 shows a typical curve for green growing vegetation plotted in wavelength versus percentage of reflectance. Since channel 4 (0.8 to 1.1 micrometers) is assigned red in LACIE product 1 and 3 imagery, a red spectral signature on this imagery reflects healthy green vegetation. The analyst reviews and interprets the available imagery and reference data to determine the potential small-grains areas in a segment by correlating the temporal growth stages of small grains and the expected sequence of changes in spectral signatures to the sequence of spectral signatures evident on the imagery. Table I shows the expected sequence of small-grains signatures for the acquisition dates shown in figure 8. The field labeled "W" on the sequence of imagery in figure 8 is clearly a winter grain field. The field labeled "S" follows a spring grain temporal growth pattern. The field labeled "N" does not follow the temporal growth pattern of wheat. A detailed discussion of the interpretation process is contained in the paper by Hay entitled "Manual Interpretation of Landsat Data."

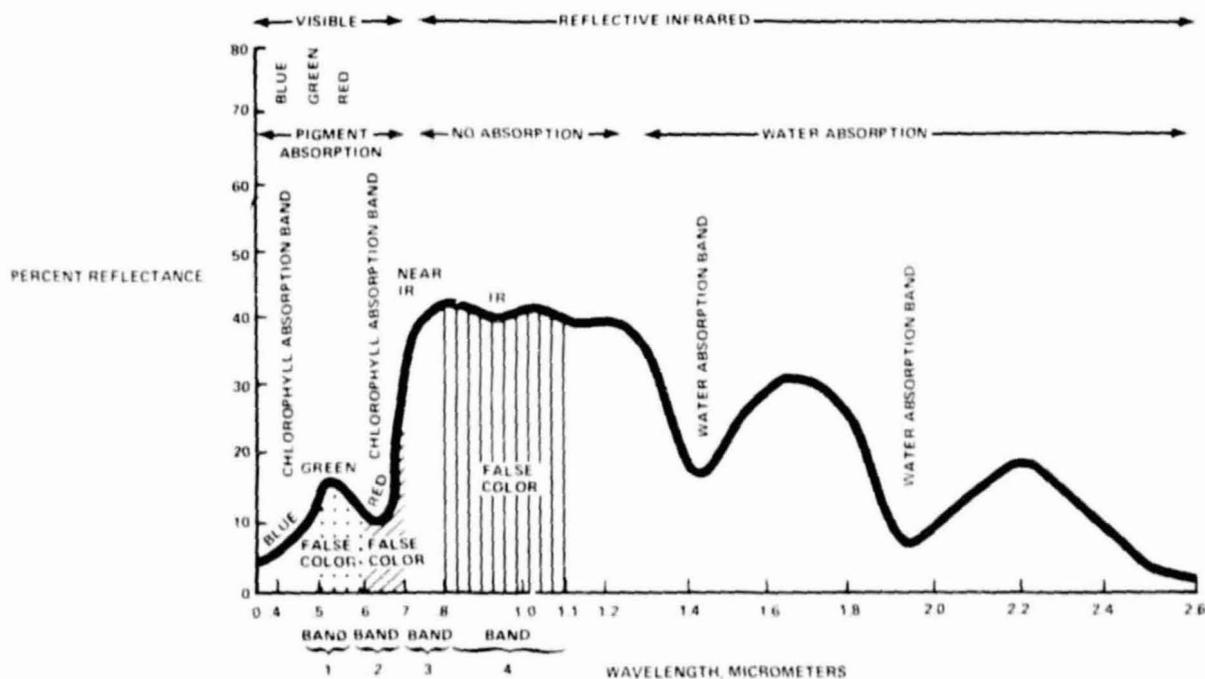


FIGURE 7.—Reflectance curve of healthy green vegetation (from ref. 1).

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

TABLE I.—Expected Spectral Signatures for Small Grains, Segment 1528

[From ref. 1]

Date	Winter grain	Spring wheat
Nov. 11 to Mar. 30	Planting/emergent stage—no red to red	Preplanting—no red
Apr. 22	Vigorous growth stage—red	Planting—no red
May 28	Vigorous growth stage perhaps some change preparatory to turning—red or brick red	Emergent—red or pink, depending on plant canopy
July 3	Ripe stage—orange, yellow, brown	Vigorous growth stage—red
Aug. 8	Harvest stage—yellow, white, tan	Harvest or ready for harvest—yellow, brown, olive green, white

A problem detected early in LACIE was that the labeling logic should be more sophisticated. The within-segment variability was very large, sometimes as great as the within-state variability. Thus, some knowledge about the parameters that affected the spectral signatures and labeling logic had to be developed and implemented operationally.

The analyst must consider interpretation variables every time he works a segment.

1. Knowledge of regional effects—such as soils, cropping practices, and climatic conditions—is an important aspect of the interpretation process.

2. Field size and registration in small- and/or strip-field segments need to be considered because a signature becomes more difficult to label with the inherent increase in the number of field boundaries and, subsequently, on the resolution of Landsat, with an increase in the number of mixture pixels. Figure 9 shows examples of large, typical, and small fields in the U.S.S.R. and the United States.

3. Acquisition history is important in preventing the occurrence of confusion crops and ensuring the accurate identification of spectral signatures. Figure 10 is an example of an acquisition history for a segment in Richland County, North Dakota. The dashed lines on the partial crop calendar at the bot-

tom of the figure indicate the adjusted growth stages for the available acquisitions for spring wheat and sunflowers, the latter a potential confusion crop. Acquisitions 1 and 4 are essential for the separation of trees and natural vegetation. Because the spring wheat has already begun to emerge on acquisition 1, crop confusion can occur; however, used with acquisition 4, the separation would be complete. With just acquisitions 1 and 3, spring wheat would be confused with sunflowers. Note the similarity of the fields on both dates.

4. Episodic events are any phenomena that cause the spectral signatures to deviate significantly from the nominal. The most pronounced episodic event monitored during LACIE was severe drought (fig. 11(a)).

The use of ancillary data, such as meteorological data and full-frame imagery, was incorporated into the CAMS labeling logic several times during LACIE. Other parameters which varied regionally to affect signatures were irrigation (fig. 11(b)); fertilization; plant varieties; planting dates, (fig. 11(c)); planting densities; etc. In the presence of any or all of these factors, the labeling logic had to be flexible enough to account for them when they were significant, yet general enough not to bog down the labeling process.

Analysis Procedures

To ensure that consistent output for subsequent evaluations came from CAMS, a controlled set of analysis procedures had to be developed, implemented, and maintained (refs. 4 to 6). These procedures had to be adaptable to new technology whenever deemed necessary, yet provide consistent results to meet the LACIE accuracy and throughput goals.

Phase I was primarily a learning experience. No one knew how well the state-of-the-art technology, current at the time, would work in an operational environment in which all segments were processed alike and high throughput was important. Mechanical procedures were available for using all the technology, but no documented decision logic and analysis procedures existed.

Two types of analysts were thought necessary in Phase I. The analyst-interpreter (AI), expert in image interpretation techniques, was responsible for identifying potential small-grains areas on Landsat film products, delineating "training" fields for all

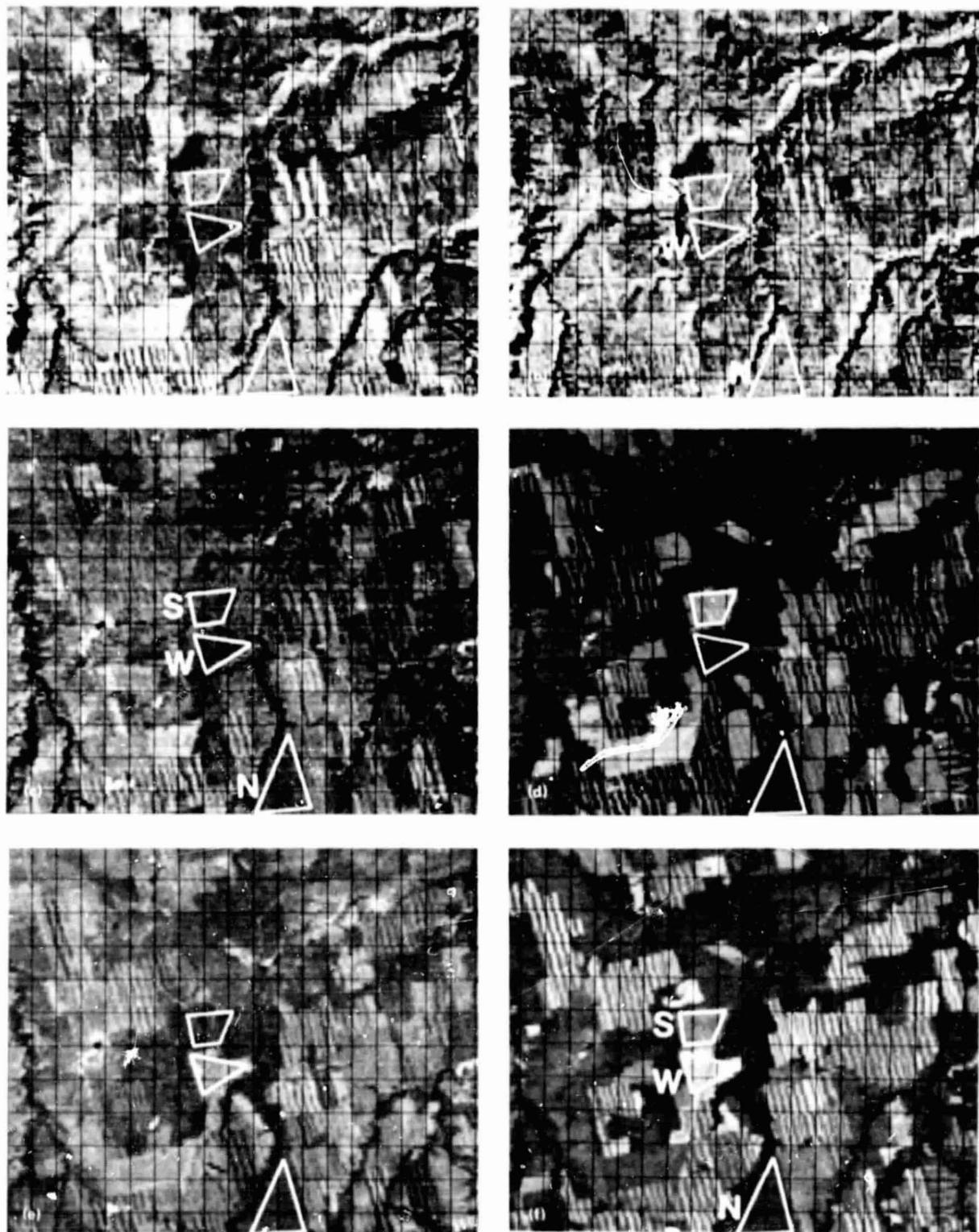


FIGURE 8.—Interpretation process exemplified. (a) November 11, 1976. (b) December 17, 1976. (c) April 22, 1977. (d) May 28, 1977. (e) July 3, 1977. (f) August 8, 1977.

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

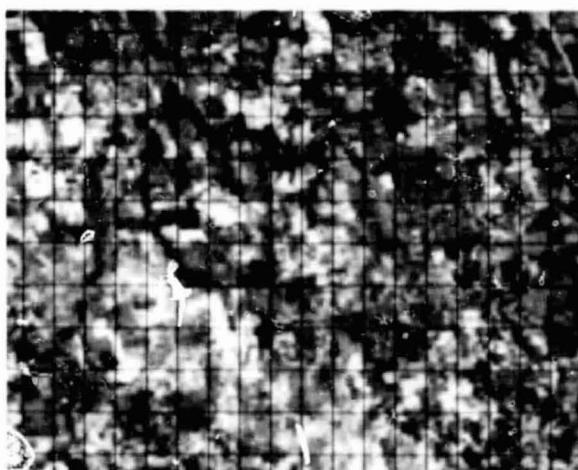
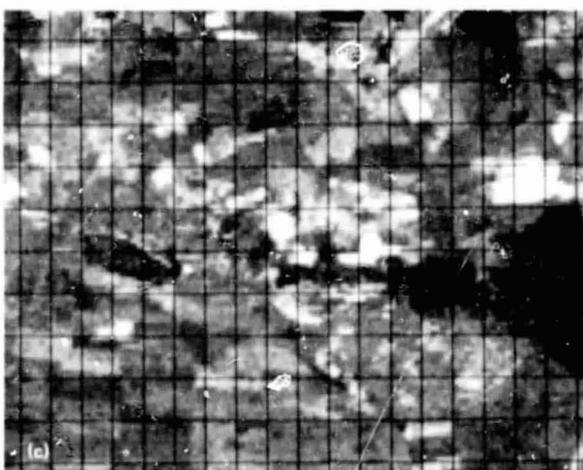
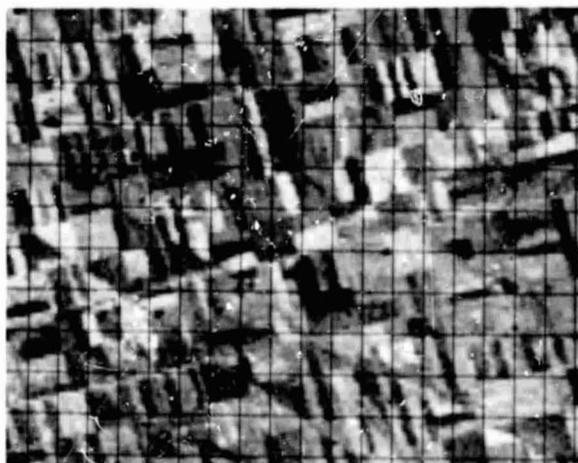
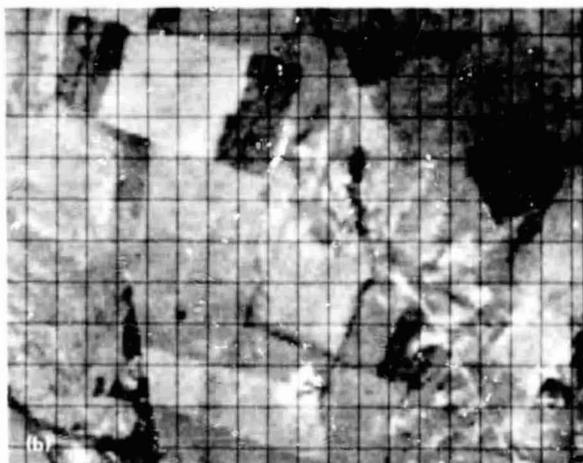
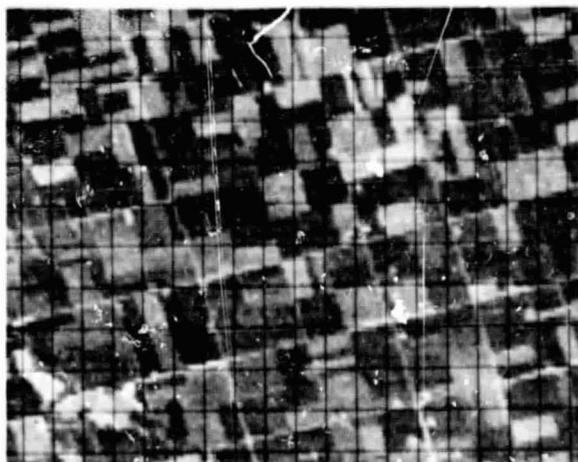
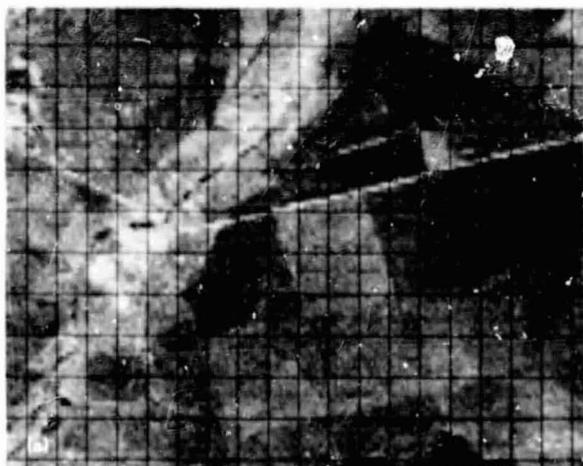


FIGURE 9.—Comparison of field sizes in the U.S.S.R. (left) and in the United States (right). (a) Large fields. (b) Typical fields. (c) Small fields.

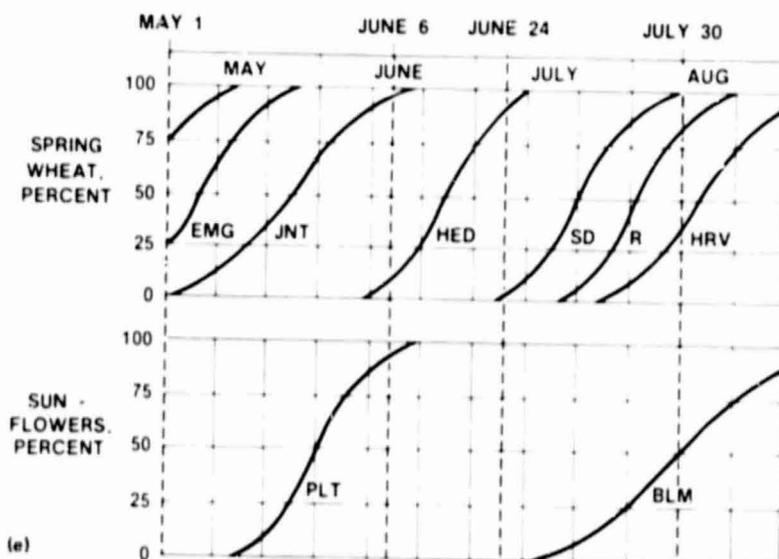
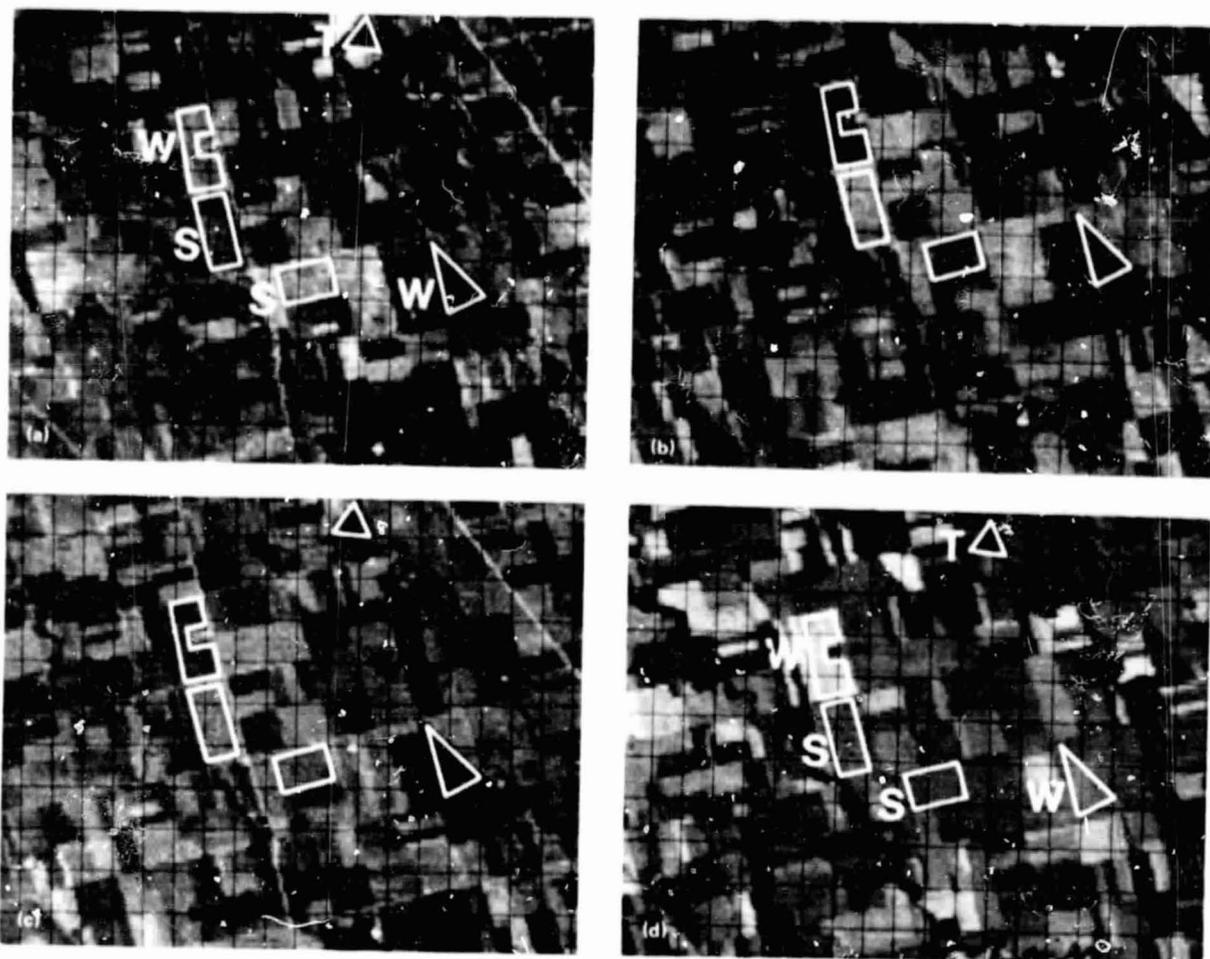


FIGURE 10.—Acquisition history for Richland County, North Dakota. (a) Acquisition on May 1, 1977. (b) Acquisition on June 6, 1977. (c) Acquisition on June 24, 1977. (d) Acquisition on July 30, 1977. (e) Partial crop calendar for the available acquisitions for spring wheat and sunflowers.

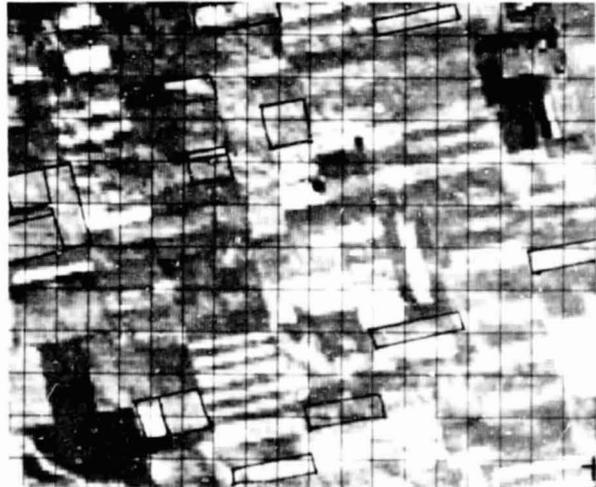
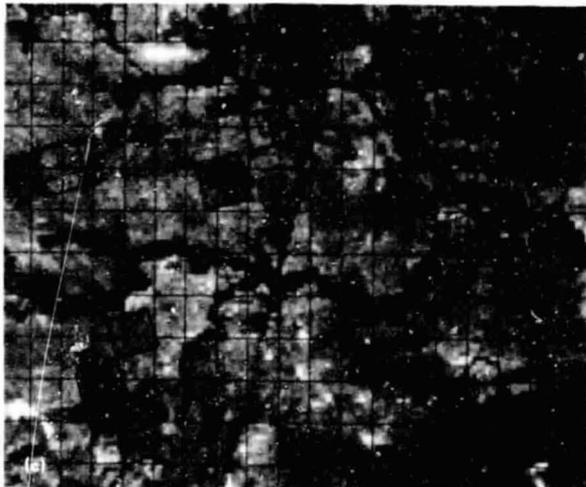
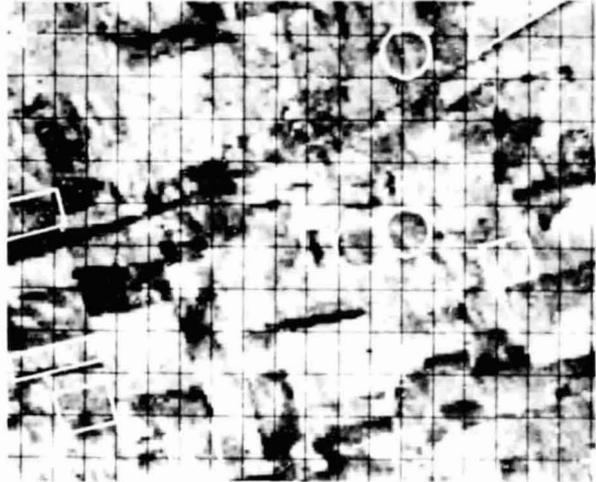
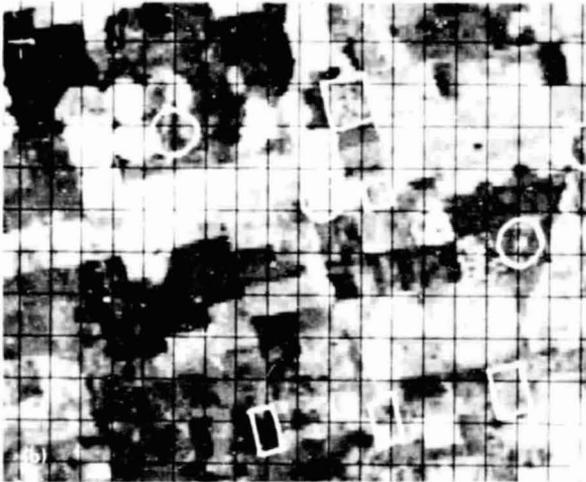
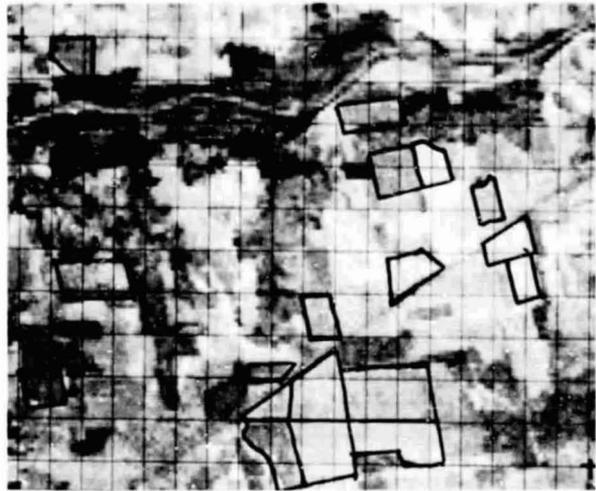
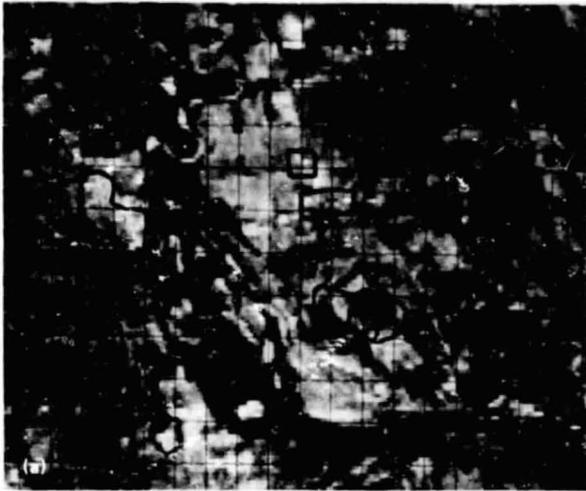


FIGURE 11.—Conditions that contribute to the variability of wheat signatures. (a) Drought stress. (b) Irrigated versus dryland cropping. (c) Planting date.

spectral signatures, and inputting the training information for computer processing. The data processing analyst (DPA), expert in pattern recognition theory and available data analysis systems, was responsible for processing the data and evaluating the results.

The AI had many obstacles to overcome. The quality of the film products was poor, and color representation was inconsistent. The acquisition coverage was poor mainly because of resource limitations at the NASA Goddard Space Flight Center (GSFC). More acquisition coverage would have been beneficial at this stage of LACIE because no documented wheat key was available. No current-year growth stage adjustments were available during this phase either; thus, the signature variability problem was compounded.

The inadequacy of the information available to the AI made his labeling task very difficult. In addition, the AI had to assume the manual drudgery of many of the data handling tasks (fig. 12). The analyst's major functions and his interaction with the operational system are further discussed in the following paragraphs. This discussion concentrates on the basic initial characteristics of the system in Phase I. Improvements made in Phases II and III are discussed subsequently.

Screening.—Not all the acquisitions received at JSC were acceptable for further processing. The initial system implemented at GSFC to extract LACIE segments, though improved in later phases of LACIE, allowed a substantial number of segments with excessive cloud cover, data dropouts (especially with Landsat-1), and misregistration. The analyst had to prescreen all the imagery and select for further processing only those acquisitions that satisfied a prespecified criterion.

Definition of training areas.—The initial system used the typical classification approaches; i.e., it required that the analyst identify, select, and label training fields for the classifier and test fields for evaluation. Each spectral signature had to be represented in the training fields selected. Figure 13 exemplifies a typical selection of fields. Sufficient samples of each signature also had to be defined so that sufficient statistics could be calculated for the classifier. These steps had to be repeated for the test fields. Then, the analyst labeled all the selected fields as wheat or nonwheat using the labeling logic.

The problems with this approach were numerous. It was very difficult for the analyst to identify all the spectral signatures. If sufficient samples were not allocated in the correct proportion to the classes

identified, then subsequent classifications were found to be affected. These tasks were very tedious and time-consuming and detracted the analyst from his most important task — labeling the signatures. These problems had the additional undesirable effect of indirectly making multitemporal processing operationally difficult, if not impossible.

Data base updating.—Data base updating the CAMS functional flow was a series of manual functions to go from signature labeling to segment classification. The analyst manually had to delineate on the imagery the vertices of all the selected fields (from 40 to more than 100). In a separate operation, the vertices were read with a grid-encoded magnifier, recorded on keypunch forms, and subsequently entered into the processing data base. About halfway through Phase I, a technique that used a semiautomated digitizer for field delineation was developed. The output was then reformatted for compatibility with the data base. This task was less tedious but still somewhat time-consuming. The next task was to complete all the forms, card decks, etc., necessary to submit the job for machine processing.

Machine processing.—Once the fields were successfully loaded into the data base, the DPA was notified. All machine processing was interactive. Many techniques were attempted: class statistics and cluster statistics (obtained by clustering the fields of each category) for maximum-likelihood classification, multitemporal classifications, and signature extension runs. Halfway through the phase, batch capabilities and statistical manipulations for signature extension were added to the system. The rework rate in Phase I was 250 percent; signature extension attempts were unsuccessful and therefore abandoned in Phase II; and the majority of the multitemporal classifications was unsatisfactory because of misregistration and AI procedural labeling deficiencies.

Evaluation of results.—Before the submission of the resultant segment wheat proportion estimate from the machine classification, a final review of the analysis output was conducted. This was accomplished by examining the probabilities of correct classification (PCC's) of the analyst-selected training and test fields. If the PCC's were above preset values, the proportion estimate was deemed satisfactory. However, in addition, the classification map was manually correlated with the color-infrared imagery to ascertain whether the wheat areas and nonwheat areas appeared to be in good agreement with the classification map.

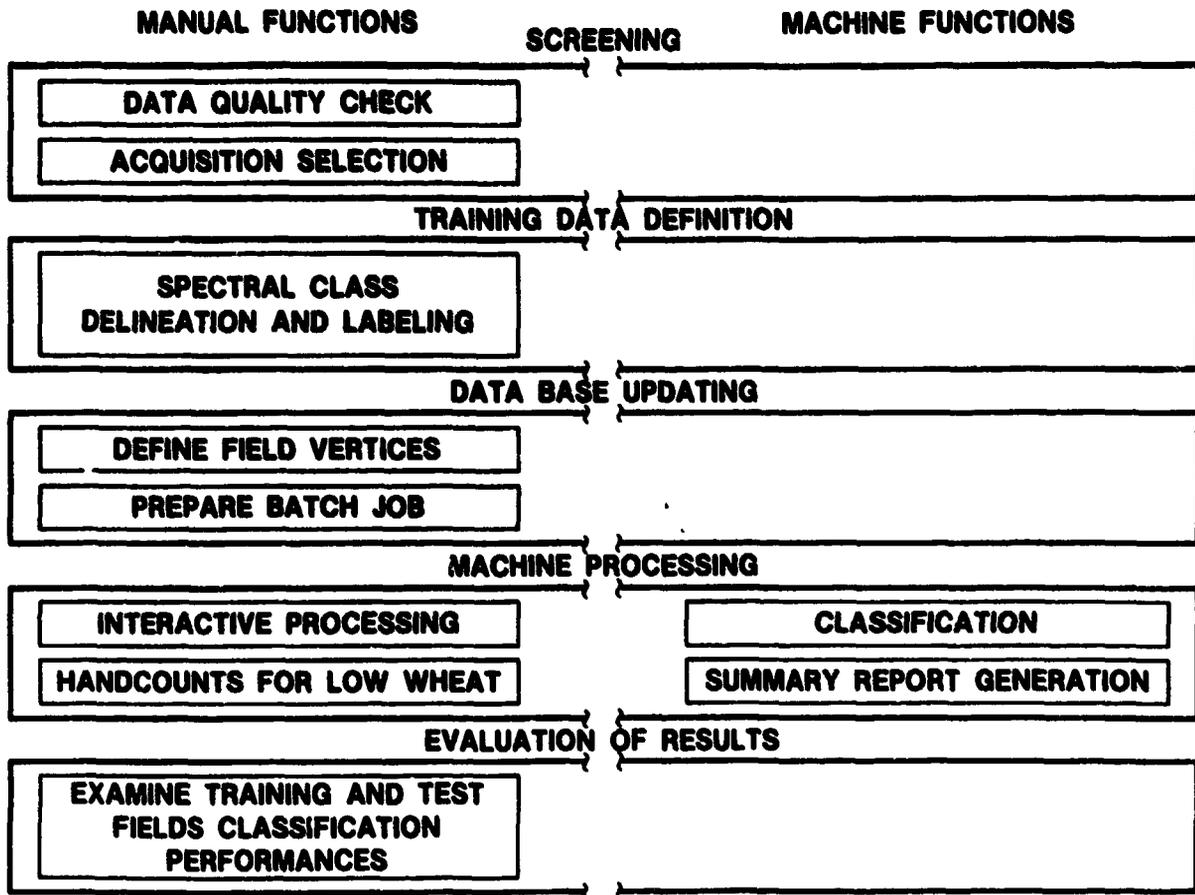


FIGURE 12.—Initial data processing tasks.

Overall Integration of CAMS Operations into LACIE

Phase I.—At the conclusion of Phase I, several problems became obvious. Most of these were due to the operational problems inherent with the large throughput requirements. Technology that was straightforward when performing remote-sensing applications in a low-key environment became very cumbersome when applied to a large number of segments. Naturally, these problems were compounded by the constraint of there being no ground data for the generation of training statistics. More importantly, no ground data were available to evaluate the performance of the system during operations. This mode was new but understandably necessary for the overall goals and objectives of LACIE.

Phase II.— Because of the problems identified in Phase I, it became apparent that new approaches to the LACIE environment would have to be developed during Phase II. One of the obvious problems in Phase I was the segregation of the AI and the DPA. It was very difficult for each to perform his job satisfactorily without a clear understanding of the other half of the analysis. A successful attempt to resolve this problem was made in Phase II when two-man teams were formed in a cross-training effort to produce end-to-end analysts. Increased acquisition coverage (all clear acquisitions) and improved film products were also aids to the analyst. Documented decision logic procedures came into existence, and Robertson growth stage adjustments were available to refine the crop calendars for the current crop year.

Training fields were still used to train the

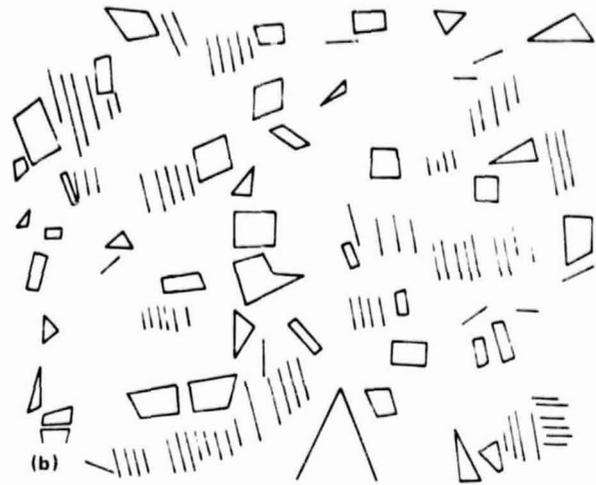
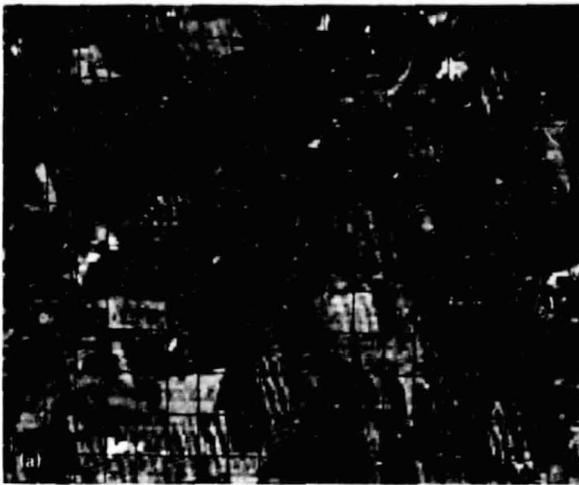


FIGURE 13.—Phase II training field delineation. (a) Imagery. (b) Field overlay.

classifier, but more emphasis was placed on multitemporal signature selection when processing was multitemporal. Because this task was very tedious and time-consuming, multitemporal processing was avoided as much as possible. Another deterrent for multitemporal processing was a machine requirement that when more channels were used for processing, the training sample had to be larger to prevent excessive thresholding. One of the major problems with Phase II was the increased processing load (from 700 to 1700 segments). Some operational efficiencies had to be developed and implemented to allow the analyst to increase his throughput.

A "no significant change" procedure was introduced in Phase II to expedite processing. If the analyst thought that processing a new acquisition would not significantly change a previously satisfactory estimate for the segment, the code was assigned to the new acquisition, no machine processing occurred, and the least satisfactory estimate was used for the aggregation. In addition, because of a lack of sufficient samples, segments subjectively thought to contain less than 5 percent wheat were processed by handcounting the individually labeled wheat pixels. Although these processes helped reduce the backlog, the subjectivity involved made them technically undesirable.

During Phase II, a data base management system was implemented to relieve the analyst of a significant portion of the data base updating task. This system performed all of the reformatting routines,

quality checks, and status and tracking job submissions. A batch processing capability was also implemented; it eliminated all the interactive processing except for reworks (approximately 25 percent).

In addition to the analyst's functions, an independent group of personnel performed a final quality assurance review of each segment analysis before submittal for subsequent aggregations. The group's primary function was to ensure that established procedures were followed during the analysis so that subsequent evaluations and problem-solving tasks could be conducted on a consistent and controlled data set. In addition, the group reviewed the segments in a regional framework to ascertain whether specific spectral signatures were consistently labeled and whether the signatures interfaced with the aggregation and sample allocation elements of the project. The findings were reported to aid in problems particular to those areas.

Many problems were resolved in Phase II, but some needed more attention. Labeling was a major problem in Phase I; however, with a year's experience, the analyst gained confidence in labeling by having a mental wheat key. The decision logic was quite general, however, and did not cover the variability of wheat signatures under different growing conditions (i.e., drought, winterkill, dryland, cropping, irrigation).

Another problem area was analyst bias. The continued tendency to underestimate caused further questioning of the assumption that the analyst could accurately sample and label the segment with only

imagery for spectral information. The classifier depended on training sample sizes in proportion to the spectral signature across the entire segment. This was a difficult task for the analyst to achieve for all the possible combinations of multitemporal classes from imagery alone. Consistency between analysts was impossible to maintain with this type of procedure. If two analysts were given the same segment to process, they would use different quantity and population training fields, and their proportion estimates could vary considerably. To begin to develop solutions to the underestimation problem, it was obvious that an opportunity to assess the Landsat data at a more detailed level was necessary.

Phase III.— Solutions to the labeling and bias problems were the major goals of Phase III. Interpretation teams were formed so that the interpretation process could be a consensus of a group of three to five people with diverse backgrounds and levels of experience. A supplemental film product, product 3 (fig. 3) and Analyst Interpretation Keys, as discussed previously, were available to the analyst during this phase as labeling aids. However, a more serious signature masking effect due to the film generation process was corrected for Phase III. Note the fields outlined on the color-infrared product in figure 14(a). In film space, little or no signature difference is apparent. Note the position of these same fields in figure 14(b), which is a spectral scatter plot of the actual digital data from Landsat channels 2 and 3. Thus, to improve the analyst's labeling capability, spectral aids, scatter and trajectory plots of the Landsat digital data, were made available to the analyst for the first time, but they were only initially used as an evaluation tool to check dot label consistency. The spectral data were transformed into two variables, green number and brightness (refs. 7 and 8). Figure 15 shows scatter plots of analyst-labeled dots for segment 1528. Figure 16 shows sample trajectory plots for winter wheat, spring wheat, and nonwheat pixels.

The implementation of Procedure 1 during Phase III originated from a cluster-based procedure designed for small-fields areas. This procedure was refined and improved by considering analyst bias, machine bias, and efficiency. A systematic grid (fig. 17) was developed for dot labeling so that sampling among analysts would be more consistent and signature sampling more statistically accurate across the whole segment. Thirty to fifty dots are identified to start and label the clusters, and another 40 to 60 dots are labeled and used as a stratified areal sample after

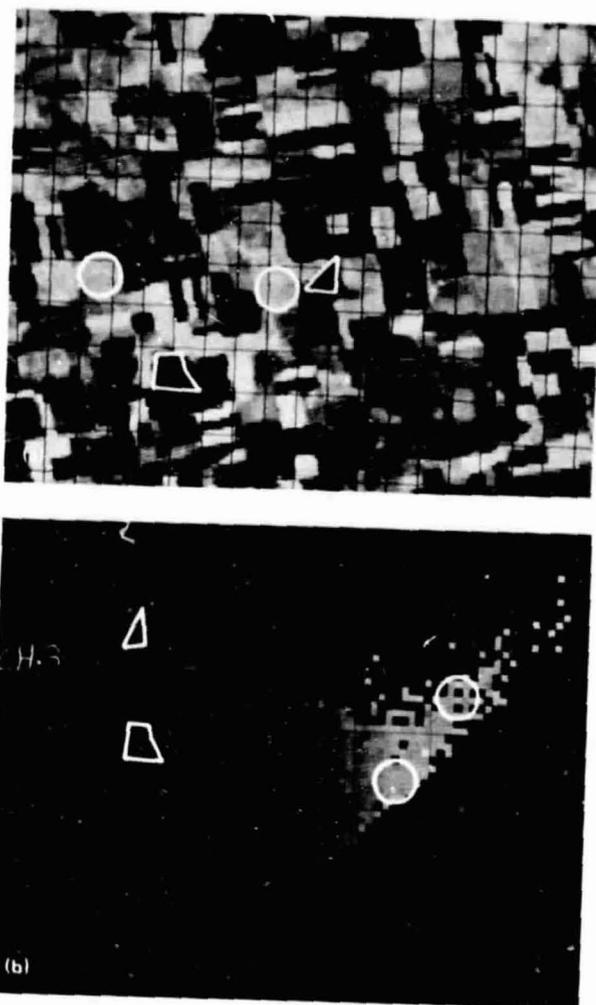


FIGURE 14.—Relationship of Landsat imagery to the spectral scatter plot. (a) Product 1. (b) Scatter plot.

classification to correct for machine bias and to reduce reworks.

Clustering is the process of grouping pixels according to some distance measure (ref. 9). The pixel vector ($X = x_1, x_2, \dots, x_n$, n = number of channels) of each of the 22 932 pixels in the segment was compared with the pixel vector of each of the 20 starting dots ($Y = y_1, y_2, \dots, y_n$). Each pixel was assigned to the closest starting dot.

$$\text{Distance} = \sum_{i=1}^m |y_i - x_i|$$

where m is the number of channels used in clustering.

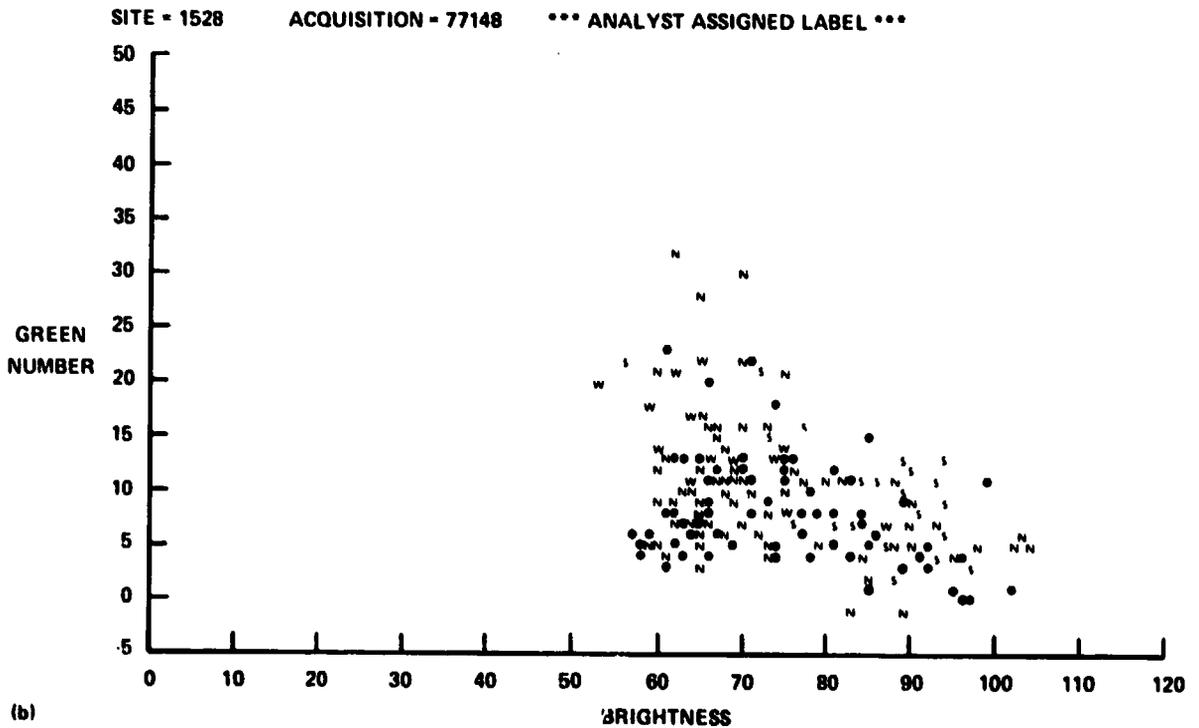
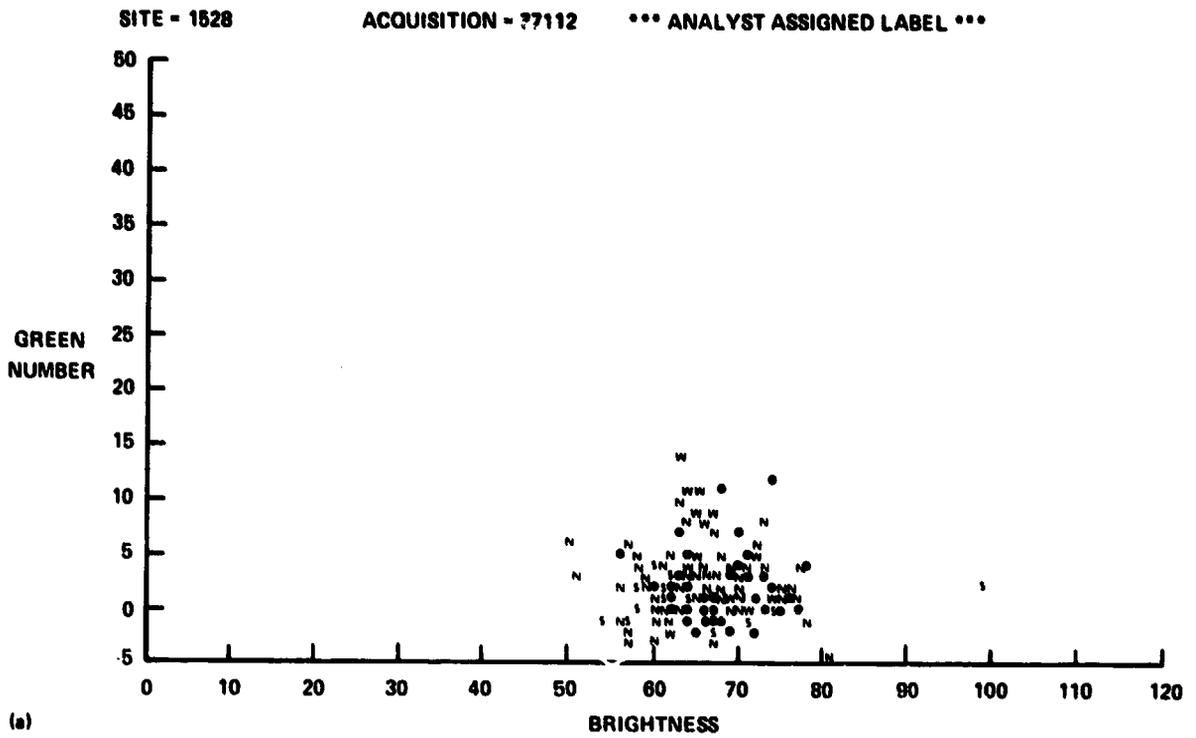
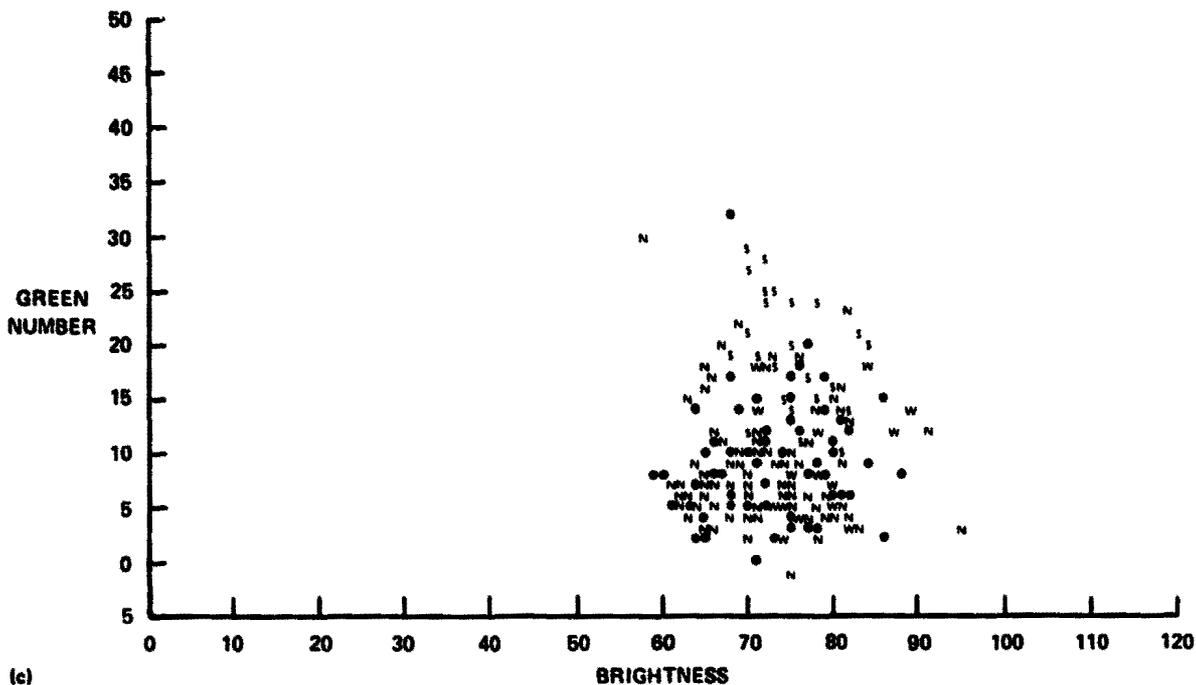


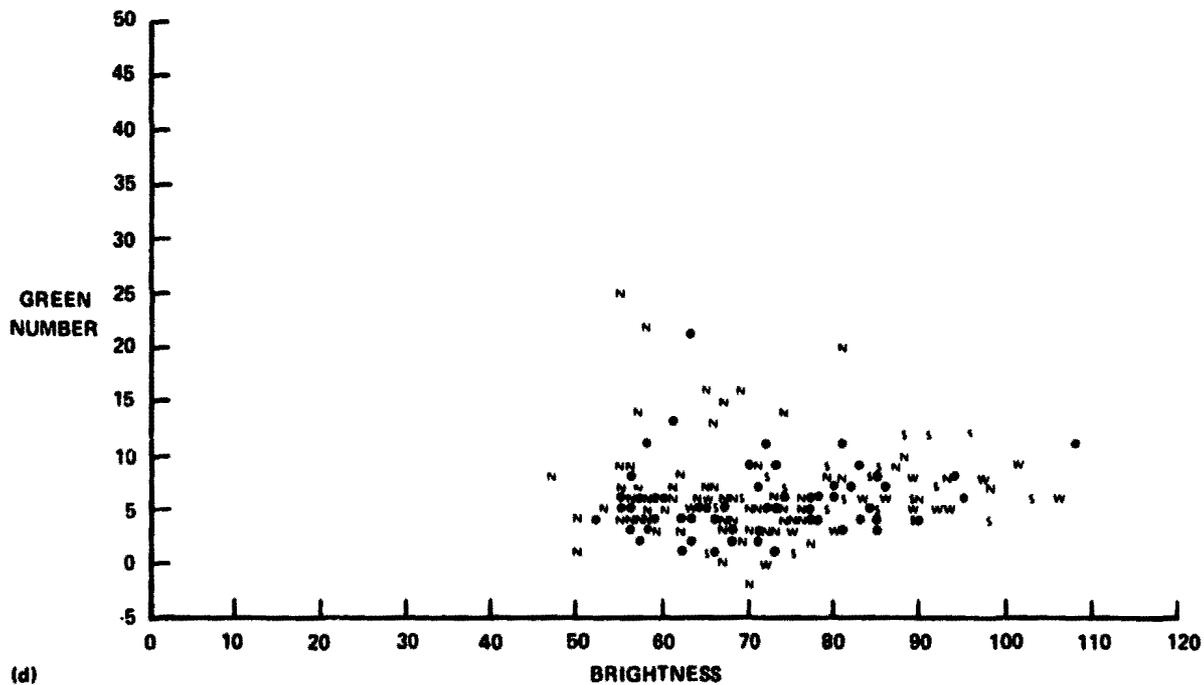
FIGURE 15.—Scatter plots of analyst-labeled dots for segment 1528 (from ref. 1). (a) Acquisition on April 22, 1977. (b) Acquisition on May 28, 1977. (c) Acquisition on July 3, 1977. (d) Acquisition on August 8, 1977.

SITE = 1528 ACQUISITION = 77184 *** ANALYST ASSIGNED LABEL ***



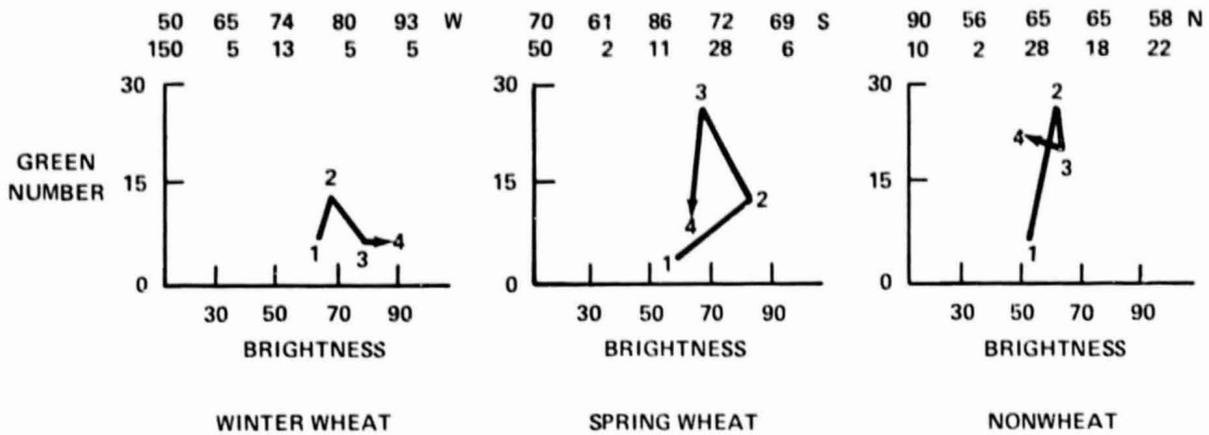
(c)

SITE = 1528 ACQUISITION = 77220 *** ANALYST ASSIGNED LABEL ***



(d)

FIGURE 15.—(concluded).



ACQUISITION 1: APR. 22, 1977; BMTS OF 2.5 FOR WINTER WHEAT AND 2.0 FOR SPRING WHEAT
 ACQUISITION 2: MAY 28, 1977; BMTS OF 3.5 FOR WINTER WHEAT AND 3.1 FOR SPRING WHEAT
 ACQUISITION 3: JULY 3, 1977; BMTS OF 5.0 FOR WINTER WHEAT AND 4.6 FOR SPRING WHEAT
 ACQUISITION 4: AUG. 8, 1977; BMTS OF 7.0 FOR WINTER WHEAT AND 6.0 FOR SPRING WHEAT

DOT NUMBER	GRID INTER-SECTIONS		ANALYST LABEL	CLASSIFIED	ACQUISITION 1		ACQUISITION 2		ACQUISITION 3		ACQUISITION 4	
					G	B	G	B	G	B	G	B
91	50	150	W	W	5	65	13	74	5	80	5	93
119	70	50	S	S	2	61	11	86	28	72	6	69
153	90	90	N	N	2	56	28	65	18	65	22	58

FIGURE 16.—Trajectory plots of typical winter wheat, spring wheat, and nonwheat pixels (from ref. 1).

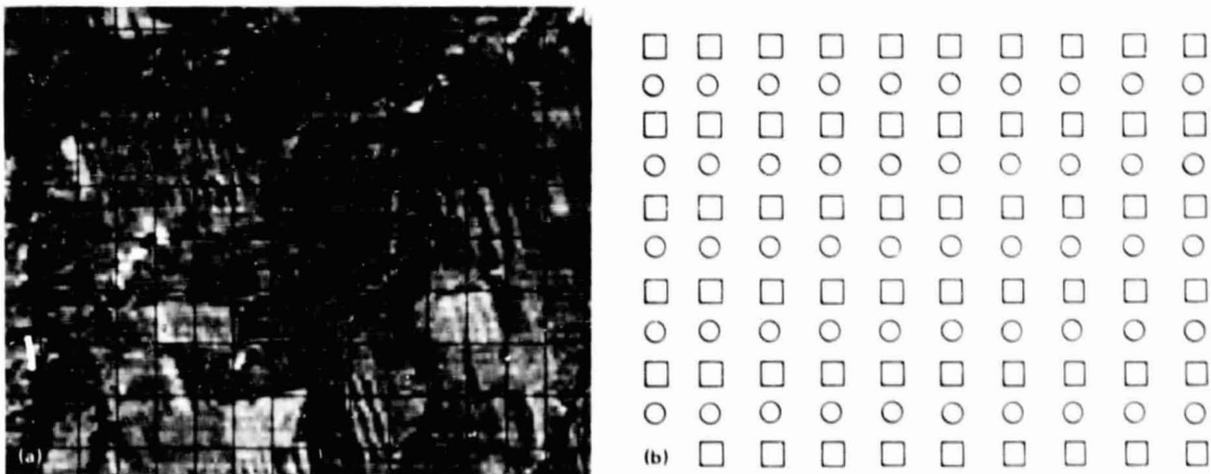


FIGURE 17.—Example of Phase III dot labeling. (a) Imagery. (b) Dot overlay.

After all pixels were assigned, the mean and the standard deviation of each cluster were computed for each channel. Figure 18(a) is a color-coded cluster map of segment 1528. Each cluster is assigned a unique color. Figure 18(b) is a conditional cluster map. If the distance between the cluster and the dot that labeled it was within a specified threshold of a dot labeled "S," "W," or "N" (spring wheat, winter wheat, or nonwheat), the cluster was color-coded green, cyan, or yellow, respectively. If it was not within the threshold, the cluster was assigned a unique color. The cluster was then automatically labeled by the closest dot, and the cluster statistics were used to classify the segment. In classification, the probability of each pixel in the segment belonging to each cluster is computed. Values are summed to the category level. Pixel assignment is made to the most likely category and then to the most likely cluster or subclass within that category. After classification, thresholding is applied to remove from the final classification results pixels with a low probability of belonging to the assigned subclass (ref. 10). Figure 18(c) is a classification map of segment 1528.

The analyst time line decreased by 75 percent (fig. 19). He no longer had to delineate and digitize fields. He simply labeled the prescribed number of dots; filled out a process request form; and, after the classification was complete, evaluated the results (fig. 20). The evaluation process was also designed to be more objective. The percent of correctly classified labeled dots and the variance between the machine estimate and bias-corrected estimate were used to determine whether the results were satisfactory or unsatisfactory. Less than 10 percent of the acquisitions processed required rework. Acquisitions with low PCC's were machine-reworked by relabeling the clusters and reclassifying. When the PCC's were marginal and the variance was large, a dot rework was performed by manually labeling more bias correction dots and recomputing the bias correction estimate and variance.

The implementation of Procedure I for Phase III and significant improvements in the data management system were major breakthroughs and account for the biggest success stories in LACIE. This vast improvement in the allocation of tasks to man and machine allowed the accomplishment of the increased scope of Phase III with no reduction in performance. In fact, because more tasks were transferred to the machine (fig. 21), the analyst could now concentrate on the labeling function and, with the improved products, do a better job.

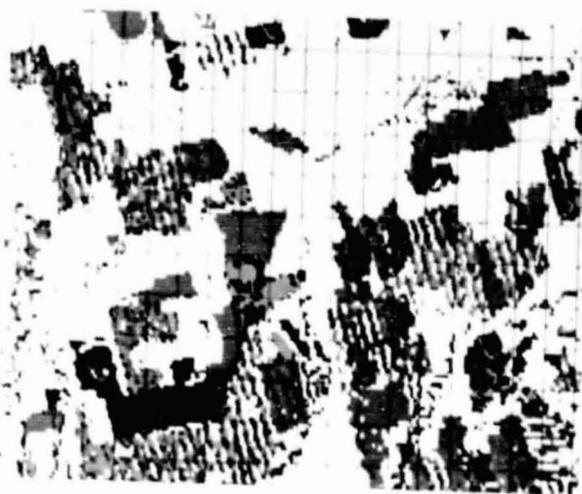


FIGURE 18.—Evaluation products. (a) Plain cluster map. (b) Conditional cluster map (yellow = nonwheat, green = spring wheat, and cyan = winter wheat). (c) Classification map (white = nonwheat, light gray = winter wheat, dark gray = spring wheat).

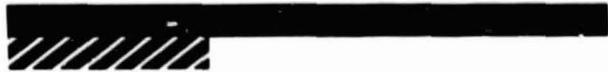
SCREENING



DATA PREPARATION



DEFINITION OF TRAINING AREAS



DATA BASE UPDATING



MACHINE PROCESSING



EVALUATION OF RESULTS



REWORK



OVERHEAD



PHASE I 

12-14 HR

PHASE III 

3-4 HR

FIGURE 19.—Analyst time line.

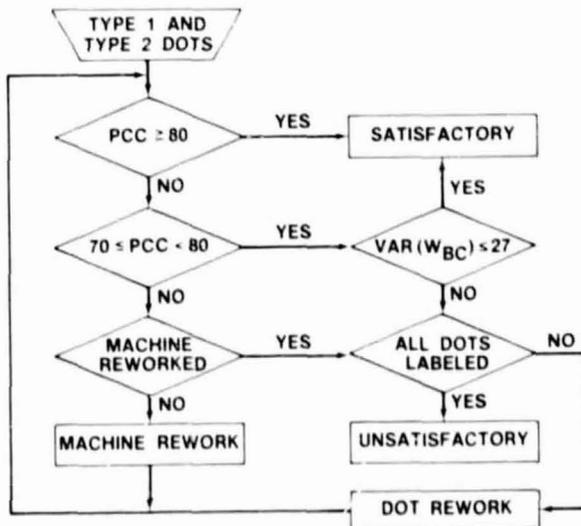


FIGURE 20.—Evaluation flow.

SUMMARY

The key accomplishments of the CAMS operations are related to the accuracy and throughput goals of the basic LACIE output requirements.

Accuracy Goals

The accuracy goals achieved are best described in the paper by Potter entitled "Accuracy and Performance Characteristics of LACIE Area Estimates." However, it should be emphasized that the per-segment errors in bias and variance measured against ground truth decreased significantly from Phase I to Phase III. The decrease in variance, noted during Phase II and continued into Phase III, is attributed to the stabilization of repeatable analyst procedures. The bias component of error which was manifested

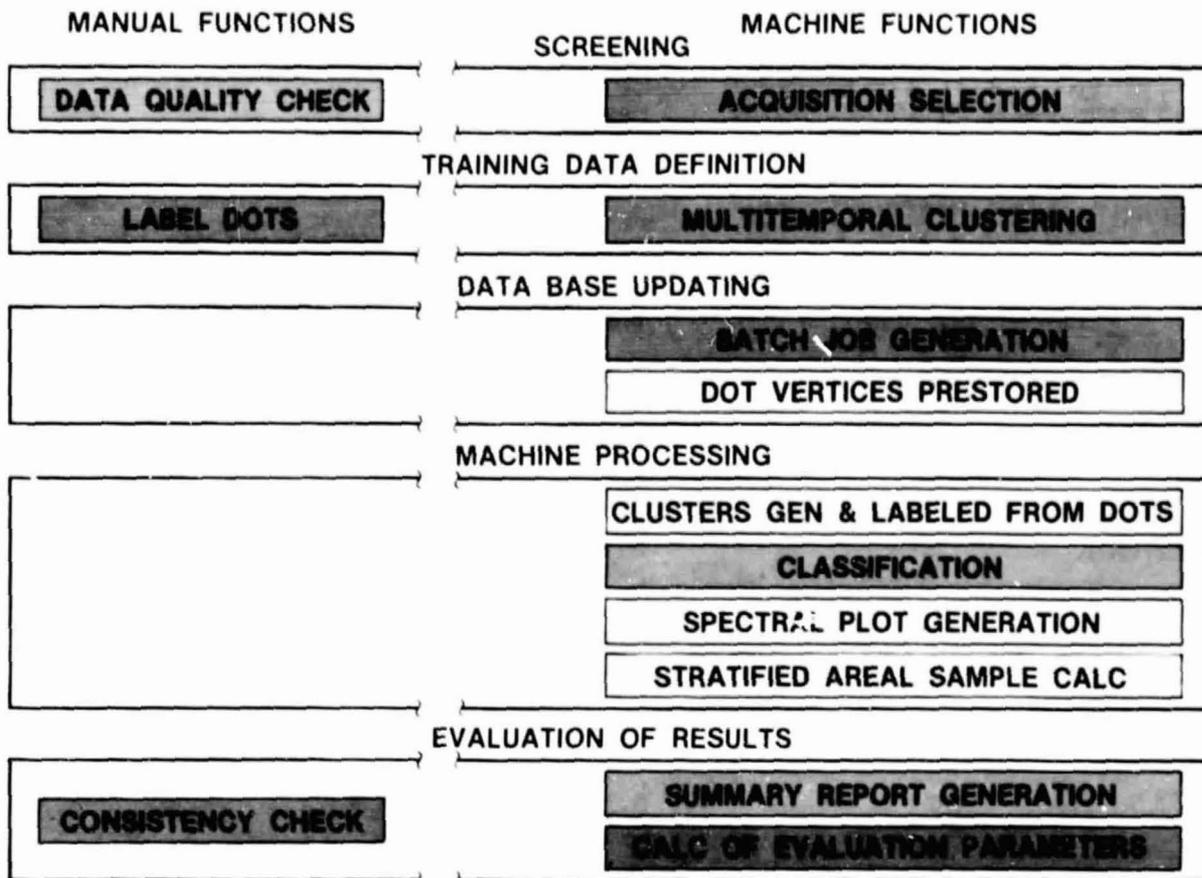


FIGURE 21.—Procedure 1 data processing tasks.

in Phase I and somewhat in Phase II by underestimates of wheat proportion, especially in the spring wheat regions, is attributed to weaknesses in the labeling decision logic. In Phase I, the decision logic was fairly simple; i.e., there was very little allowance for the total within-segment signature variability. This logic tended to result in errors of omission (off-nominal signature of wheat called nonwheat) and thus in the underestimation of the true wheat proportion. However, the process matured during the final phases of LACIE and the bias was reduced to insignificant levels, except for the highly concentrated regions of spring wheat and spring barley. The primary factors that helped support this decrease in bias were overall improvements in the decision logic, the compilation of the Analyst Interpretation Keys, the incorporation of digital spectral aids, and the implementation of Procedure 1 into the CAMS operation.

Throughput Goals

The significance of Procedure 1 to the CAMS accuracy goals can best be appreciated when one recalls that the initial motivation leading to its design and subsequent implementation was for increasing the CAMS throughput. In Phase I, the scope for CAMS was 700 segments and 2000 acquisitions. In Phase II, it increased to 1700 segments and 9000 acquisitions. To accommodate this increased load, questionable schemes, such as handcounts and "no significant change" procedures, had to be used to keep the CAMS operation from bogging down. It became apparent very quickly that to meet the Phase III scope of 3000 segments and 18 000 segments, a more technically palatable technique was necessary; thus, Procedure 1 was developed.

During this period, other elements of CAMS, concerned with the systems aspects of the processing

problem, noted the substantial number of functions being performed by the analyst which dealt more with data handling than with data information extraction. In fact, the majority of his segment time line was spent on data-handling functions instead of data analysis. The implementation of a data management system to relieve the analyst of the data handling, statusing, and tracking was the significant contributor to the accomplishment of the segment loading and throughput goals.

The development and implementation of Procedure I represents one of the first really significant approaches to solving some of the major man-machine interactions involved in a large-scale classification application. Comparison of figure 12 to figure 21 readily shows the improvement in the man-machine distribution of tasks as realized by the Procedure I concept. In addition, as shown in figure 19, the decrease in the analyst's time spent per segment between Phase I and Phase III using Procedure I is significant. The payoff comes not only in more throughput realized per analyst but also in probable improvements in accuracy because of a better allocation of analyst tasks. For example, although the analyst's total time was reduced from 12 to 14 hours to 3 to 4 hours per segment, his actual interpretation time was increased from approximately 1 hour to 2 hours per segment. Thus, the analyst spent less time on clerical tasks and more time on interpretation and labeling.

These accomplishments indicate that LACIE has indeed been successful. The mistakes made within CAMS were many; however, weighed against the stage CAMS personnel were in 3 or 4 years ago, without state-of-the-art technology because of LACIE, they now seem trivial or at least worthwhile as experience gained.

This paper has attempted to capture the significant highlights of the total CAMS experience during LACIE. It has been very difficult for CAMS personnel intensely involved in the daily workings of CAMS with its broad spectrum of technological and operational activities to "broadbrush" a happening that dealt its full measure of emotion on their professional and personal lives for 3.5 years. Most hopefully, the contributions of the CAMS analysts to the success of CAMS in LACIE have been suffi-

ciently evident, for these contributions represent the backbone of LACIE and the builders of the programs to follow.

REFERENCES

1. Austin, W. W.: Detailed Description of the Wheat Acreage Estimation Procedure Used in the Large Area Crop Inventory Experiment. LEC-11497, Lockheed Electronics Co., Houston, Tex., Feb. 1978.
2. Robertson, G. W.: A Biometeorological Time Scale for a Cereal Crop Involving Day and Night Temperatures and Photoperiods. *Int. J. Biometeor.*, vol. 12, 1968, pp. 191-223.
3. Analyst Interpretation Keys. Vol. 1, Wheat/Small Grains Inventories; Vol. 2, United States Great Plains Regional Keys. NASA Johnson Space Center, Houston, Tex., May 1977.
4. Chang, C. Y.; Finley, D. R.; Melvin, M. A.; and Sommers, J. P.: CAMS Detailed Analysis Procedures for the LACIE Operations. LEC-5252, Lockheed Electronics Co., Houston, Tex., Jan. 1975.
5. CAMS Detailed Analysis Procedures. LEC-7673, Lockheed Electronics Co., Houston, Tex., Mar. 1976.
6. LACIE Phase III CAMS Detailed Analysis Procedures. LACIE-00720, NASA Johnson Space Center, Houston, Tex., Aug. 1977.
7. Kauth, R. J., and Thomas, G. S.: The Tasselled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. *Proceedings of the 10th Annual Symposium on Remote Sensing of Environment (Ann Arbor, Mich.) Oct. 1-2, 1975.*
8. Thompson, D. R.; and Wehmanen, O. A.: The Use of Landsat Digital Data To Detect and Monitor Vegetation Water Deficiencies. *Proceedings of the 11th International Symposium on Remote Sensing of Environment (Ann Arbor, Mich.) Apr. 25-29, 1977.*
9. Kan, E. P. F.: The JSC Clustering Program ISOCIS and Its Applications. LEC-0483, Lockheed Electronics Co., Houston, Tex., July 1973.
10. Abotteen, K. M., and Austin, W. W.: Comparison Results of Chi-Squared and Fisher's F Thresholding in LACIE. LEC-7555, Lockheed Electronics Co., Houston, Tex., Dec 1975.

Concepts Leading to the IMAGE-100 Hybrid Interactive System

T. F. Mackin^a and J. M. Sulester^b

SUMMARY

As LACIE Procedure 1 evolved from the Classification and Mensuration Subsystem small-fields procedures, it became evident that two computational systems would have merit—the LACIE/Earth Resources Interactive Processing System based on a large IBM-360 computer oriented for operational use with high computational throughput, and a smaller, highly interactive system based on a PDP 11-45 minicomputer and its display system, the IMAGE-100. The latter had advantages for certain phases; notably, interactive spectral aids could be implemented quite rapidly. This would allow testing and development of Procedure 1 before its implementation on the LACIE/Earth Resources Interactive Processing System. The resulting minicomputer system, called the Classification and Mensuration Subsystem IMAGE-100 Hybrid System, allowed Procedure-1 operations to be performed interactively, except for clustering, classification, and automatic selection of "best" acquisitions, which were offloaded to the LACIE/Earth Resources Interactive Processing System. This article comments on the development and use of this new system.

BACKGROUND

As Procedure 1 (P-1) evolved from the Classification and Mensuration Subsystem (CAMS) small-fields procedures, it became evident that most of its major elements could be easily incorporated into the LACIE/Earth Resources Interactive Processing System (LACIE/ERIPS). These included the use of

dots, bias corrections, selection of best acquisitions, clustering from starting vectors, automatic cluster labeling, and classification with labeled cluster statistics.

It was also apparent that some functions could not be implemented as quickly as desired. These included the complete set of spectral aids to assist in labeling dots in real time. In addition, an easily used rework capability would not be implemented on a timely basis on the LACIE/ERIPS, where segments were required to be reclassified by reworking.

The overall throughput time line of the system was also an issue. It was thought that a more directly interactive system would allow an analyst to complete real-time analysis on the system in a convenient way with a shorter data turnaround time. It was also apparent that a fast, more direct system would allow more experimentation with P-1. Much testing and development would be necessary to refine the concepts of P-1; to choose appropriate parameters; and, in particular, to define the specific spectral aids and procedures to be used in analysis.

IMAGE-100

At the time P-1 was being developed, the Interactive Multispectral Image Analysis System model 100 (IMAGE-100), coupled with a Programmed Data Processor model 11-45 (PDP 11-45) minicomputer, was used for supporting research in several activities in the Earth Observations Division (EOD) at the NASA Johnson Space Center (JSC). The IMAGE-100, built by the General Electric Corporation, was widely considered to be an advanced system when the EOD had acquired it only 2 years earlier, and it had been continuously upgraded since that time.

Several unique features of the IMAGE-100 made it ideal for implementing P-1. Its display system was remarkably flexible and efficient; it provided com-

^aLockheed Electronics Company, Inc., Systems and Services Division, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

plete flexibility in color presentation of images; it furnished a movable cursor whose shape could be defined in several ways; and it allowed eight themes (single-bit images), all of which could be used to store and display results of image operations such as classifications and cluster maps. These bit images could be combined in Boolean operations to make still others, and the results could also be stored and displayed as bit images. All bit images could be displayed alone, in combination with each other, or as overlays to a regular image.

In addition to these capabilities, which probably did not exist on any other commercial machine at the time, there was a "pixel-alarmed" capability. This unique feature allowed a pixel (picture element) or a group of pixels to be identified in one of many ways and then all other identical pixels to be automatically identified. Pixels could be chosen by the cursor or identified manually on a histogram displayed on its graphics screen. Once identified, all identical pixels in the scene were alarmed or "flashed" in an operation that was apparently instantaneous because it depended on hard-wired programs. This capability was especially promising for the spectral aids visualized for P-1.

The IMAGE-100 had other advantages. It already existed in the Data Techniques Laboratory of EOD and hence required no lengthy procurement. Although used for other projects, it could be dedicated to the needs of LACIE; and, quite importantly, in the work force were applications programmers with the depth and experience to design changes and prepare the programs in a very short time.

Designing the Hybrid System

The decision was made to implement P-1 on both systems—the ERIPS for production and the IMAGE-100 mainly for accuracy—and a large number of personnel began working out details for both systems. Some of the personnel were involved in the finalization of the concepts of P-1 without regard to implementation on any one specific hardware/software system. Others worked at the functional design of P-1 as it could be implemented on the IMAGE-100 (concurrently, the same was being done for the LACIE/ERIPS).

Consideration of total run time, admittedly based only on educated guesses, suggested that the clustering algorithms in particular would take an inordinate amount of time on the PDP 11-45 computer.

Although this was never directly confirmed, an early decision was made to offload to the LACIE/ERIPS the three most lengthy computations—clustering, classifying, and selecting the best acquisitions. That is, these three functions would be performed on the LACIE/ERIPS in batch runs, whereas all other operations in P-1 were to be run on the IMAGE-100 system. In this sense, the IMAGE-100 form would be known as the CAMS IMAGE-100 Hybrid System.

The most notable feature of the final design was the spectral aids to be made available to the analyst, including both trajectory plots and spectral plots. Both would be displayed on the IMAGE-100 screen and be available in real time. They would allow complete flexibility of display colors, placement on screen, etc.

The spectral plots, including the Kauth coordinates, were to be made available for any two channels. As an example, an analyst might choose to display all pixels in clusters labeled as wheat, with green number and brightness as coordinates. Full use was made also of alarming—the analyst could encircle a dot on the spectral plot with the cursor and it would be alarmed on the color-infrared (CIR) image of the acquisition, along with all other identical pixels. In addition, dots could be displayed in arbitrary colors; however, the most useful and most generally used display of the dot was in the same colors as the actual image being displayed.

DEVELOPMENT OF THE SYSTEM

The baseline document on requirements for the system became available in June 1977. For the next month, an implementation design team worked to define the configuration of the hybrid system and to develop a specification document, which was published prior to the design review held on February 1, 1977. The system was completed and delivered on June 2, 1977.

Although multitemporal analysis had existed previously, it was not more convenient. In the new system, the labels and types of all dots were retained from a given analysis; therefore, analysis of subsequent acquisitions could be made with the same set, perhaps with minor modifications. This further reduced analyst contact time on the IMAGE-100 system. Because the computer required to do the large computation loads was no longer actively involved in interactive displays, higher volume could be realized, if required.

CAMS/Crop Assessment Subsystem (CAS) statistics file build, and CAMS/CAS data base update.

Once the data are available for work, software modules are required in the following sequence.

1. Initiate segment analysis
2. IMAGE-100 control
3. Image display
4. Field definition
5. Dot crosshair
6. Dot scatter plot
7. Theme logical
8. Window erase
9. Dot selection
10. Single-dot labeling
11. Automatic cluster labeling
12. Classification map display
13. Recompute proportions
14. Cluster map main
15. Unconditional cluster map display
16. Conditional cluster map display
17. Mixed cluster map display
18. Display report generator
19. Reports

These software modules provide for the type of display on the screen of the IMAGE-100 that is shown in figure 2. It should be added that, in the spectral plot area, the axes normally were Kaouth green number versus brightness, but greenness versus time or individual bands could also be displayed.

The system disk and data disk configurations are defined in figure 3. It should be noted that the CAMS

IMAGE-100 Hybrid System software is a complete system operating with its own "driver" or "control" program and interfaces directly with the hardware operating system software, the Resource Sharing Executive model 11D (RSX-11D). The control program can access the image library or other system software modules outside the CAMS software, but all communication or transfer of data or information from within the CAMS software to "outside" modules must go through the CAMS "driver" or "control" program.

The data disk shows functionally the arrangement of the data files retained in the operating data base. A maximum of six acquisitions per segment is retained. To load an additional acquisition, one of the existing acquisitions must be deleted.

The hardware configuration for which the software was developed is shown in figure 4. The interactive software modules provide the analyst with all of the tools necessary to select, one by one, all of the grid overlay dots or pixels; to alarm the pixel in the spectral domain and the CIR image; and to show the trajectory of the vegetation through time represented spectrally. Knowing the spectral signature of wheat and the normal crop planting and growth cycle of wheat, the analyst can identify those pixels or dots that spectrally represent wheat and label them accordingly with the labeling module.

When labeling with the IMAGE-100 is complete, the output in a card-image format is sent to the IBM 360-75 for the clustering and classification processing. The labeled dots allow the analyst to check his identification of wheat against the computer's ability to discriminate between spectral classes and provide a ratio of performance. He may elect to relabel the dots or pixels and start over; or he may relabel whole clusters of pixels, change the classification, and change the statistics which reflect the percent of wheat in the segment.

OPERATIONAL IMPLEMENTATION

The operational data flow in figure 1 and the process prior to that operation in figure 5 identify the functions performed for analysis of a LACIE segment.

1. As tapes with segment imagery data are received from the NASA Goddard Space Flight Center (GSFC), they are copied. One tape goes to the production film converter (PFC) and one goes to the computer for loading into the data base.

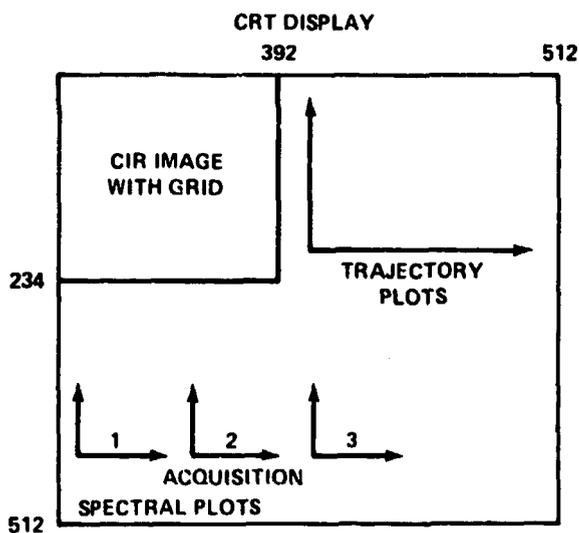


FIGURE 2.—Diagram of display provided by the IMAGE-100.

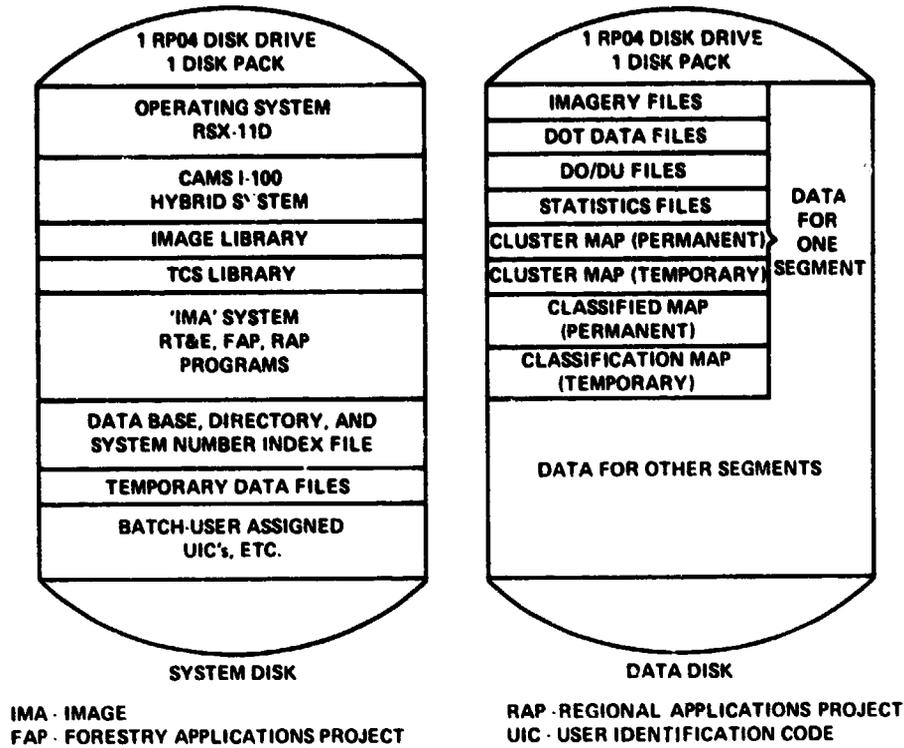


FIGURE 3.—System disk and data configuration.

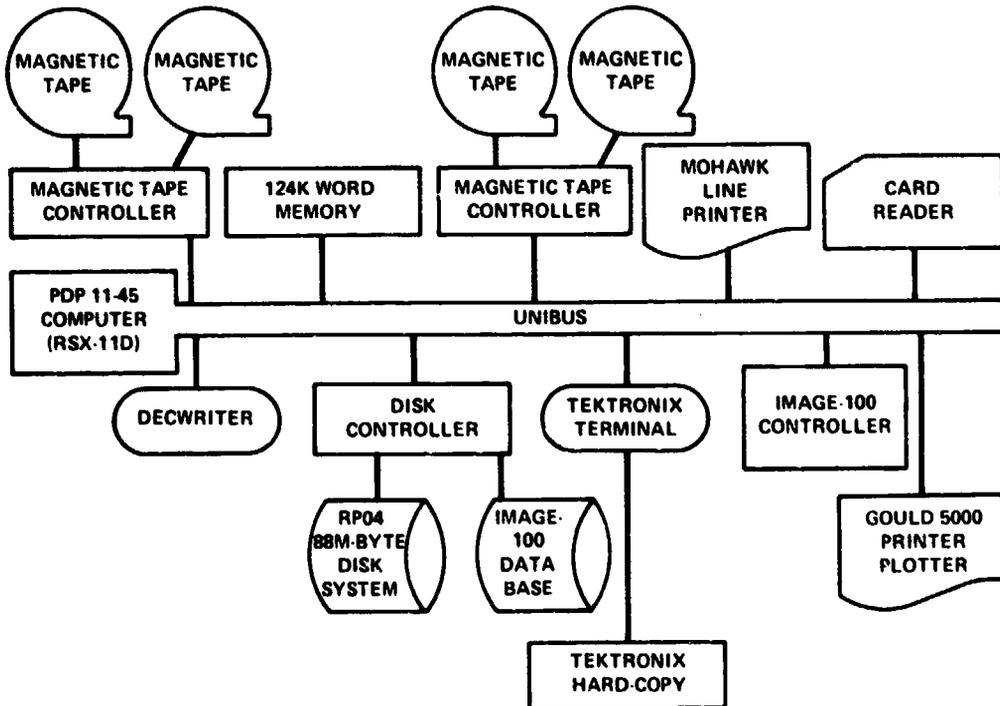


FIGURE 4.—CAMS/IMAGE-100 Hybrid System hardware configuration.

and standard job orders, as described in the paper on Landsat data acquisition and storage.

Because the CAMS IMAGE-100 Hybrid System was developed for use by analysts with a wide variety of backgrounds, a thorough training program in basic analysis, IMAGE-100 software, and operations was required. The training included interpretation methods, interpretation of photographs, regional analysis, multispectral sensing, and signature analysis of agriculture data.

Generally, analysts who are trained in photo-interpretation found the use of a cathode-ray-tube (CRT) display in false color somewhat less desirable

than photographic imagery; however, the added availability of analysis aids, such as spectral plots, trajectory plots, and color maps of clusters and classes, made the data analysis task easier and, hopefully, decreased labeling errors.

It was found, further, that the increase in automation could speed up the analysis and provide much more flexibility in performing one. The total speed of the overall system was found to be dependent on logistics control and data handling; and, although the analyst's "hands-on" time decreased in this prototype system, total elapsed time was governed more dramatically by the system data flow.

USDA Analyst Review of the LACIE IMAGE-100/Hybrid System Test

P. Ashburn,^a K. Bülow,^a H. L. Hansen,^b and G. A. May^c

INTRODUCTION

Late in LACIE Phase II, a proposal was submitted to implement and test an interactive imaging system during Phase III (1976-77). This proposal was approved, and the initial test of the system was initiated in February 1977. The major purpose of the test was to provide and evaluate a pseudointeractive classification capability during LACIE that would provide the project some experience with a type of system that the user (U.S. Department of Agriculture (USDA)) would be implementing. The test was also to evaluate a timeliness factor in segment processing.

Initially, the test was to utilize the classification capabilities of the General Electric (GE) Interactive Multispectral Image Analysis System Model 100 (IMAGE-100 or I-100). This was later changed to programing and implementing a procedure called Procedure 1 (P-1) on the GE I-100 Hybrid System. This procedure called for the labeling of single-pixel training fields (dots) using the I-100 Hybrid System; the clustering and classification was to be done on the Earth Resources Interactive Processing System (ERIPS). Evaluation and minor rework could then be done on the I-100. This total system was called the I-100/Hybrid System.

NASA, USDA, and the supporting contractor, Lockheed Electronics Company (LEC), cooperated in the planning of the test. NASA and LEC were the principals in programing and implementing the test. USDA supported the test by providing the primary operational analysts. Analysts were also provided by LEC for instructing the USDA analysts on how to use the system and for identifying and correcting

procedural and system utilization problems that were encountered.

The I-100 test included operational segments from the U.S.S.R. and test segments from Canada and the United States. These segments provided the USDA analyst with a wide range of geographic conditions from which a broadened range of learning experiences was achieved.

In addition to training, the USDA analysts had the opportunity to test and evaluate additional capabilities of the interactive imaging system and to suggest how the system and procedures could be improved. A number of problems with hardware, software, and procedures were identified and corrected during this process.

Finally, the USDA analysts learned that procedural options were needed to work with the widely varying climatic, geographic, and cultural conditions that exist in the major countries of the world.

THE TEST

The USDA I-100 analysts used essentially the same procedure (P-1) as the Classification and Measurement Subsystem (CAMS) analysts; details for these procedures are documented in the CAMS Image 100/Hybrid System Procedures/Requirements (LACIE-C00202, JSC-11669, Jan. 1977). The major exception was that USDA analysts had at their disposal an interactive system that provided a wide range of on-line spectral aids that could be used in the initial dot-labeling process. They also had an on-line capability for relabeling dots and clusters.

The LACIE sample segments to be processed by the USDA I-100 team were selected from the U.S. Intensive Test Sites (ITS's) (24), Canada ITS's (10), Canada Blind Sites (30), and Kokchetav, U.S.S.R. (50). Each of the four analysts had a list of segments for which he was responsible.

The segments were loaded onto the I-100 disk and

^aUSDA Foreign Agricultural Service, Houston, Texas.

^bUSDA Soil Conservation Service, Houston, Texas.

^cUSDA Economics, Statistics, and Cooperatives Service, Houston, Texas.

the analyst was advised which segments were ready for processing. A total of six different acquisition dates for the same segment could be loaded and retained on the system at any one time. The analyst could call up a segment number and view the Julian date for each acquisition, then select the date to be used for processing.

A color-infrared (CIR) image (both a film product and a cathode-ray-tube (CRT) image) provided the analyst with a visual product of the segment. Space was provided under the CRT image for displaying up to six different scatter plots. When appropriate, the analyst used this space to build a scatter plot for each acquisition date.

The analyst viewed the image to locate and outline areas that should be labeled "DO" (Designated Other than wheat) areas. He also outlined areas that could not be identified because of clouds or cloud shadows; these areas were labeled "DU" (Designated Unidentifiable). Each pixel falling within these outlined areas, including any of 209 single-pixel fields, would no longer be counted in the subsequent processing of the segment. The DO and/or DU areas could be viewed as a color overlay on the CRT image. After the DO/DU definition was completed, the analyst proceeded to a single-pixel field selection procedure where he requested a selection of 30 or more unlabeled single-pixel fields (dots). These did not include any of the pixels contained within DO/DU areas already labeled DO/DU pixels. This selection was stored in one of the eight theme tracks and viewed as an overlay on the CRT image. These 30 or more dots were cross-referenced in the scatter plots to determine if the entire spectrum of the scene was represented. The analyst then began labeling the single-pixel fields to establish labels to be used as starting vectors for the clustering and classification algorithms.

The analyst used two methods for accessing single-pixel fields or dots. One was by typing in the number of the dot and the other was a cursor selection that required putting a cursor-formed box around the dot that represented the field. Either of these two methods prompted a display of information about the dot, including the green number for each acquisition date, the raw digital channel values for each date, the dot number, the location by coordinates of the dot in the scene, and any previous analyst or classifier label for the dot. The dot was alarmed in both the CRT image and the scatter plots at this time. In addition, all other dots in the image with the same radiance values were also alarmed. If

the analyst was unsure of a label for a dot, he could enlarge the area surrounding the dot for better viewing. He then labeled the dot as wheat, small grain, or nonwheat. If the dot was on the edge of an agricultural field, he could skip labeling that dot and proceed to the next dot. This process was repeated for each of the 30 or more selected dots in the scene. These dots were then designated "Type 1" and provided the starting vectors for the clustering and classification algorithms.

After the first selection of dots was labeled, another selection of 40 or more dots was made from the unlabeled dots. The same procedure used for labeling the Type 1 dots was used for labeling these dots, which were designated Type 2 dots, with one exception; no dots could be skipped when labeling. The labels for the Type 2 dots were later used to provide a bias correction of the machine classification.

Computer cards for all labeled dots were generated and sent to ERIPS for processing, after which the analyst was notified that the classification results had been returned and loaded for the continuation of processing on I-100.

The analyst reinitiated the segment processing and requested to see the cluster results and/or classification map. He also checked the percent-correct classification (PCC) for both Type 1 and Type 2 dots. If a PCC of 80 percent or higher was achieved and if the classification map was satisfactory, the results were deemed suitable for aggregation and the results were sent to the Crop Assessment Subsystem (CAS). If not, a rework of the data was required.

The rework could take several forms. One was to check the analyst label and change it to agree with the classifier label, which would increase the PCC. Another form was to change cluster labels by viewing the cluster results and map overlay and changing the labels of mislabeled or misclassified clusters. However, this only improved the PCC if there were analyst-labeled dots associated with the cluster. The last resort was to add more dots and send the segment back to ERIPS for reclassification.

This same procedure was used for all subsequent acquisitions. Generally, the same labels were used on the new acquisitions prior to their classification.

I-100/HYBRID SYSTEM RESULTS

The USDA analyst team analyzed segments from three countries. Twenty-four intensive test sites were

chosen in the United States. These sites were located throughout the Great Plains but the majority were in the spring wheat area. Forty intensive test and blind sites in Canada were selected for processing. These segments fell in Alberta, Saskatchewan, and Manitoba. The team also analyzed 50 Russian segments located in Kokchetav.

Since the USDA analysts were working with an on-line interactive imaging system, they had at their disposal a number of capabilities not provided to the CAMS analysts as an on-line system. These on-line capabilities—such as interactive dot labeling, class or cluster map overlay flicker, and flashing of all dots of equal spectral value—were very useful.

In working with and labeling dots, the multidot alarm capability of the I-100 was a very useful tool, allowing the analyst to select a single pixel on the imagery and alarm within the image all pixels with the same spectral value. This tool also helped the analyst in determining the spectral confusion that may exist within the scene. For example, on a given acquisition, a selected wheat pixel would alarm a pixel that occurred in a hayfield.

Another useful tool of the I-100 was the eight theme tracks. This allowed the analyst to store, at one time, different cluster and classification results that could then be used to overlay the Landsat image. By flashing the different themes on and off, the analyst could view the different overlays for comparison and analysis. However, even with these capabilities, it was sometimes very difficult to identify the wheat.

It was very difficult to interpret and analyze the Russian segments because of a low soil moisture condition that existed at seeding. The native vegetation was showing moisture stress problems on the first acquisitions obtained in early spring; these conditions resulted in poor wheat emergence and the analyst had difficulty identifying fields of emerged wheat. The USDA team flagged to the project the poor wheat condition that existed in Kokchetav.

Very little precipitation occurred in Kokchetav during wheat tillering and jointing. Therefore, much of the wheat that emerged produced little vegetative growth. As a result, the red signature from healthy growing wheat did not appear on the imagery. This made it difficult to obtain a good area estimate of the wheat; in some cases, the estimates for these segments were only 25 percent of the previous year's estimates.

For the Canadian segments, the USDA team passed direct wheat estimates or spring wheat esti-

mates only. Significant amounts of flax, rapeseed, and barley are grown in the spring wheat areas of Canada. The analyst found the green-number scatter plots of the 209 dots to be extremely valuable in separating wheat from flax and rapeseed. Those dots representing flax and rapeseed group together and separate from the small grains and other green vegetation. The flax and rapeseed tend to appear pinkish on the CIR image, which is distinctly different from the appearance of wheat.

No repeatable technique was established for separating wheat from barley, however. At one point in the processing, it appeared that these two crops could be separated by a pattern that existed in a plot of the raw Landsat values for channels 5 and 6. However, after applying this phenomenon against ground truth, it was determined that no apparent correlation existed between the wheat and barley in these two channels.

ANALYST PROBLEMS AND RECOMMENDATIONS

During the process of utilizing the interactive I-100/Hybrid System, the USDA analysts identified several major and many minor improvements that could be made in the procedures and software in the system. Since LEC was the primary contractor in the development of the software, interaction with the LEC staff and the USDA analysts became imperative.

After all users of the system were thoroughly familiar with the P-1 software, several meetings were held to collect inputs on how the system should be upgraded. Inputs were provided by LEC system analysts and instructors and by USDA analysts. Often the same problem was identified by more than one of the three users.

As a result of the meetings on how to improve the CAMS I-100/Hybrid System, a long list of recommendations for improvements was made, ranging from correcting spelling in menus and prompts to major changes in the different processors. All problems that were identified could be divided into five categories: (1) wrong capabilities were stressed; (2) unnecessary data were provided to the analysts; (3) methods were needed to ease the man-machine interface; (4) additional capabilities were needed; and (5) designer performance needed improvement.

There were many areas where systems analysts could improve the crop analysts' activities. However,

in order to do so, it was necessary for the systems analysts to understand the background of the crop analysts. When error messages occur, they can be worded for the system analyst or the crop analyst. It is, however, the crop analyst who will use and respond to them most often. Therefore, these messages, as well as all others, should be oriented to the primary user.

There was a problem in looking at film images and CRT images because of an inherent incompatibility between the two images. This problem was most apparent when the analyst was trying to adjust the CRT colors to match those of the film images. More than 5 minutes are required to display the entire image of a segment on the CRT; this time has to be minimized because image manipulation can occur frequently.

During processing, numerous messages, menus, and prompts were shown to the analysts. Although some of the messages were critical, they were displayed the same as all other messages. Because of this great similarity, the analyst could very easily overlook a message that would render invalid everything he had done.

Outlining a field and designating it as DO/DU was considerably more complex than originally anticipated. A prime problem was the limitation in the number of vertices which a single area could have. The analysts often had to break the areas into arbitrary subareas which contained no more than the limit of 10 vertices. With a large number of fields, it became difficult to keep track of the names of the fields; it would have been simpler to have been able to place the cursor in the field to identify it.

Since an I-100/Hybrid System procedure was specified, the software could have been designed so that one automatically proceeded from one step to another in the correct sequence. This step-by-step sequence could have been accomplished through the use of global (total) defaults. Similarly, the computer software should be streamlined. It should be developed to parrot recognized procedures. It would be advantageous to make the nonstandard processing easy to use.

Because of the nature of the Hybrid System, it was necessary to perform part of the work on the I-100 and the remainder on the LACIE ERIPS. The interfacing problems could have been streamlined considerably. There was also a problem of delay between processing and a timely receipt of results. It was orig-

inally hoped that the timelag would be only a day or two, but experience indicated lapses of more than a week. Because of this situation, the analysts had to work with eight or nine segments at a time, which was inconvenient and complex.

The experience gained from the exercise on the I-100 proved to be invaluable to USDA analysts. When the Applications Test System (ATS) was being designed, many of the problem areas were avoided, and a more efficient and user-oriented system was delivered.

CONCLUSIONS AND RECOMMENDATIONS

The USDA analysts obtained valuable experience from the I-100 effort. The varied wheat conditions throughout the three-country study area enabled analysts to study and become familiar with different cultural practices, weather conditions, and farming methods and with how these factors affect wheat-growing conditions and analysis approaches.

It soon became apparent that a single analyst processing procedure was insufficient in analyzing Landsat data for the purpose of obtaining a wheat area estimate. Procedure 1 worked fairly well when used in areas having small fields and heterogeneous signatures from the land cover types within those fields. However, since it took an average of 3.5 hours per sample segment, it was extremely cumbersome and time-consuming in agriculture areas having relatively large fields and/or homogeneous spectral signatures. A processing procedure should be developed for a type or specific set of agricultural conditions; for example, when working in an area having large fields, the analyst should use a procedure that takes advantage of such conditions.

The I-100 is mainly being used as a display device for images, classification maps, spectral aids, etc. Implementation of P-1 on the machine allows for some manipulation of the display. All clustering and classification was done on the ERIPS with the results being evaluated and minor rework done using the I-100. Utilizing the two different systems to process a segment resulted in innumerable problems, as mentioned previously. Many of these problems concerned logistics, but others involved such factors as an analyst needing two or more processing sessions just to initiate and complete the analysis of a segment. In analyzing remotely sensed data, a totally in-

teractive system is needed to accomplish the task most efficiently. The system should have the capability to process a segment from start to finish, which includes both classification and evaluation.

The ATS is built around such an interactive system concept, and the experience gained on the I-100 has impacted the design of the system.

Operation of the Yield Estimation Subsystem

D. G. McCrary,^a J. L. Rogers,^b and J. D. Hill^a

INTRODUCTION

The Yield Estimation Subsystem (YES) was, as its name implies, the operational element of LACIE responsible for developing yield estimation techniques and using these techniques to produce the yield estimates necessary for each of the project's production reports. However, its work was considerably broader since YES had the responsibility to develop all the project's uses of meteorological data. This paper not only describes the operations necessary to produce yield estimates but also covers the other facets of the subsystem's work.

THE SUBSYSTEM

Overview of Products

The yield estimates were the most important products of YES since they were essential for aggregation with acreage estimates to prepare the project's production estimates. The technical approach to yield estimation is described elsewhere in this collection (see the paper by Strommen et al. entitled "Development of LACIE CCEA-I Weather/Wheat Yield Models") and will not be discussed in detail here. After the technical approach was established, it was necessary for the yield subsystem to apply it to the many areas of the world where LACIE would be performing its investigations. Ultimately, this was to produce a total of 452 separate yield models covering the principal wheat-growing areas of the U.S. Great Plains, Canada, the U.S.S.R., India, Australia, Brazil, and Argentina. Routine weather data were collected

in a timely manner to produce the input data necessary to operate on a monthly basis the 308 models for the United States, Canada, and the U.S.S.R. during the three phases of LACIE. The remaining models were evaluated as part of the project's exploratory work.

The second application YES made of the meteorological data consisted of operating the crop calendar models which produced estimates of wheat development stage for any particular day of the year. These models, which are also described in detail in this collection (see the paper by Whitehead and Phinney entitled "Growth Stage Estimation"), required daily maximum and minimum temperature reports as input data to monitor the crop's development through the growing season. Estimates derived from observations at 2500 separate weather reporting stations were provided on a biweekly basis to the analysts who were interpreting the Landsat imagery. These estimates of growth stage were intended to assist the analysts in separating wheat from other crops in the scene.

In addition to the yield and crop calendar estimates which were objectively derived from the models, subjective evaluations of crop growing conditions were prepared by YES. For each state in the U.S. Great Plains and each major wheat-growing area in the foreign countries, narrative summaries were written to describe the weather in particular and how the growing conditions may have been departing from normal. These narrative summaries were published weekly to aid the analysts in identifying regions of abnormal crop development (and consequently abnormal crop appearance) and to document the conditions under which the project was gathering its results.

Organization of YES

The Yield Estimation Subsystem blended the abilities of specialists who were skilled in

^aNOAA Environmental Data and Information Service, Houston, Texas.

^bUSDA Federal Crop Insurance Corporation, Houston, Texas.

meteorology with those who were skilled in agronomy. Personnel were assigned from all three of the participating agencies, but, because of the reliance on meteorological data, the National Oceanic and Atmospheric Administration (NOAA) had the lead responsibility for YES. The yield subsystem manager and his staff were located at the NASA Johnson Space Center (JSC) in Houston, Texas, and were supported by personnel at other NOAA working locations in Columbia, Missouri, and Washington, D.C.

The Houston staff was primarily responsible for overall management by defining requirements of the yield subsystem, evaluating its products, and integrating these products into the project output. It also prepared the weekly meteorological summaries sent to the analysts and other project reports detailing crop growing conditions.

At Columbia, the Modeling Division of NOAA's Center for Climatic and Environmental Assessment (CCEA)¹ was responsible for extending its yield modeling methodology to all areas of interest to LACIE and conducting preliminary evaluations of model accuracy. Once the yield models were developed, this division programmed them for routine operation on the main NOAA computer located at the National Meteorological Center (NMC) in Suitland, Maryland. A remote terminal at Columbia was connected to the main computer by a high-speed data transmission line. The crop calendar model, developed by contractor personnel supporting YES, was also implemented on the NMC computer and was operated by commands from the staff at Columbia.

The NOAA Center in Missouri was supplemented in its LACIE work by personnel from other project components. The U.S. Department of Agriculture (USDA) assigned to that location an agricultural economist and an agronomist to assist in the development of the yield models needed by YES. Contractor support was also available from Lockheed Electronics Company, principally in the gathering and formatting of historical weather and yield data needed by the scientists who were developing the models.

A third NOAA group, located in Washington, prepared the real-time data needed for operation of the

models and wrote the general weather/crop assessments for each country. This Assessment Division of CCEA complemented the Modeling Division by providing the processed meteorological data that division required. The USDA supported the Assessment Division by detailing a foreign commodity analyst to work 1 day each week helping to prepare the foreign crop assessments, which formed the basis of the narrative material provided to the analysts reviewing the Landsat data in Houston.

Note should be taken of an ad hoc team established to advise the YES manager on key technical matters, particularly in the area of evaluating and developing alternate approaches to yield modeling. This Yield Advisory Group was composed of two representatives from each of the three participating agencies and was chaired by a seventh member, Dr. E. C. A. Runge, Chairman of the Agronomy Department at the University of Missouri at Columbia. During the project, this group was very effective in the preliminary screening of candidate procedures.

OPERATIONS

Meteorological Data Acquisition

The Yield Estimation Subsystem was dependent on timely and comprehensive meteorological data from all areas, both domestic and foreign, for which the project was preparing wheat production estimates. To support this project requirement, NOAA provided a data base of global weather observations which is described in a separate paper within this collection (see the plenary paper by Strommen et al. entitled "The Impact of LACIE on a National Meteorological Capability"). The files of daily weather observations included data from about 8000 locations globally with about 2500 of these located in the countries of LACIE interest. The reports were collected at NOAA's NMC from foreign sources, which consisted predominantly of the Global Telecommunications System of the World Meteorological Organization and the U.S. Air Force military weather collection network. Data from the U.S. Great Plains were available from the domestic weather observation collection facilities of the National Weather Service. The data elements of principal interest were daily rainfall totals and temperature extremes, which were necessary input for operation of the yield and crop calendar models.

¹In a 1978 NOAA reorganization, CCEA became part of a new Center for Environmental Assessment Services.

The data available at NMC were placed on a file which could be accessed by the remote terminals at CCEA's Assessment Division in Washington and the Modeling Division in Columbia. In addition to the basic observational data available from NMC, a limited amount of meteorological data preprocessed by the U.S. Air Force was received by the CCEA Assessment Division. These data included average temperatures and rainfall and estimates of soil moisture conditions for the principal crop regions in the U.S.S.R. and the People's Republic of China (P.R.C.) at 10-day intervals. These data were forwarded to Washington in table and map form in time to be used routinely by LACIE.

Yield Estimation

The wheat yield models operated by YES made their estimates from inputs of average monthly weather elements, such as departure from normal monthly precipitation or departure from average monthly temperatures. This input required that the models be operated at the end of each month as soon as the data could be assimilated and the values derived. For the U.S. Great Plains, 4 working days were set aside after the end of each month for the CCEA Assessment Division to process the domestic weather data and to make the data available to the Modeling Division. The models were normally operated on the same day the data became available, thus providing the project with yield estimates for the United States on the fifth working day of the month.

Longer periods were needed to process the foreign data that had been collected on the NMC data base. For the U.S.S.R., 9 working days were used to prepare the input data needed for the 33 different regions for which models had been developed.

One might expect that, since daily weather data had been routinely collected throughout the month, it would be simple to operate the yield models on the first day of the next month. This was not the case, however, since one additional step was necessary to prepare the data for input. The reports placed on the daily data base were individual station observations representative of point weather. The yield models used average weather over the state or zone for which the estimate was made and thus required that the individual station reports be analyzed to obtain average areal values. The analysis was accomplished by plotting the individual station data on a map and drawing the appropriate isolines. The approximate

density of these stations in the U.S. Great Plains is shown in figure 1. The analysis was augmented by meteorological satellite imagery of cloud patterns. Once the data were analyzed as shown in figure 2, representative values were assigned to each zone and these became the input data for model operations. The analysis was made by CCEA's Assessment Division, and the model input data were provided to the Modeling Division. That division operated the models, reviewed the output, and placed the estimates on a file to be accessed by the project elements in Houston. The yields were copied on a portable computer terminal at JSC and placed in a secure area from which they were available for use in the preparation of the routine LACIE production estimates.

The number of yield models operated varied from month to month, depending on the country and on whether spring wheat, winter wheat, or both were being estimated. During June of Phase III, when both U.S. and U.S.S.R. spring and winter wheat were being estimated, a maximum of 56 models was operated.

Crop Calendars

The crop calendars used by LACIE to estimate wheat development stages also utilized the data base of daily weather observations available at NMC. A selected subset of individual stations was used to provide the daily maximum and minimum temperature values necessary for operation of the models. For each of those stations, a representative crop development stage was calculated for each day of the growing season. The density of these crop calendar stations in the U.S. Great Plains is shown in figure 3.

The models were operated for the United States, the U.S.S.R., Canada, and the People's Republic of China during LACIE Phase II. Operations in the P.R.C. were on a limited exploratory basis. During Phase III, the spring wheat crop calendar was also run for the exploratory investigation in Brazil, Argentina, and Australia. The estimates were updated every 2 weeks during the growing season, with about half the countries being updated each week.

Since the crop calendar model operated directly on the observed temperature reports, it was not necessary for the CCEA Assessment Division to preprocess the data. The Modeling Division directly accessed the observation files at NMC, updated the crop stage estimates for each location, and placed the new estimates on file. This file was then accessed from JSC to acquire a printed listing of the new up-

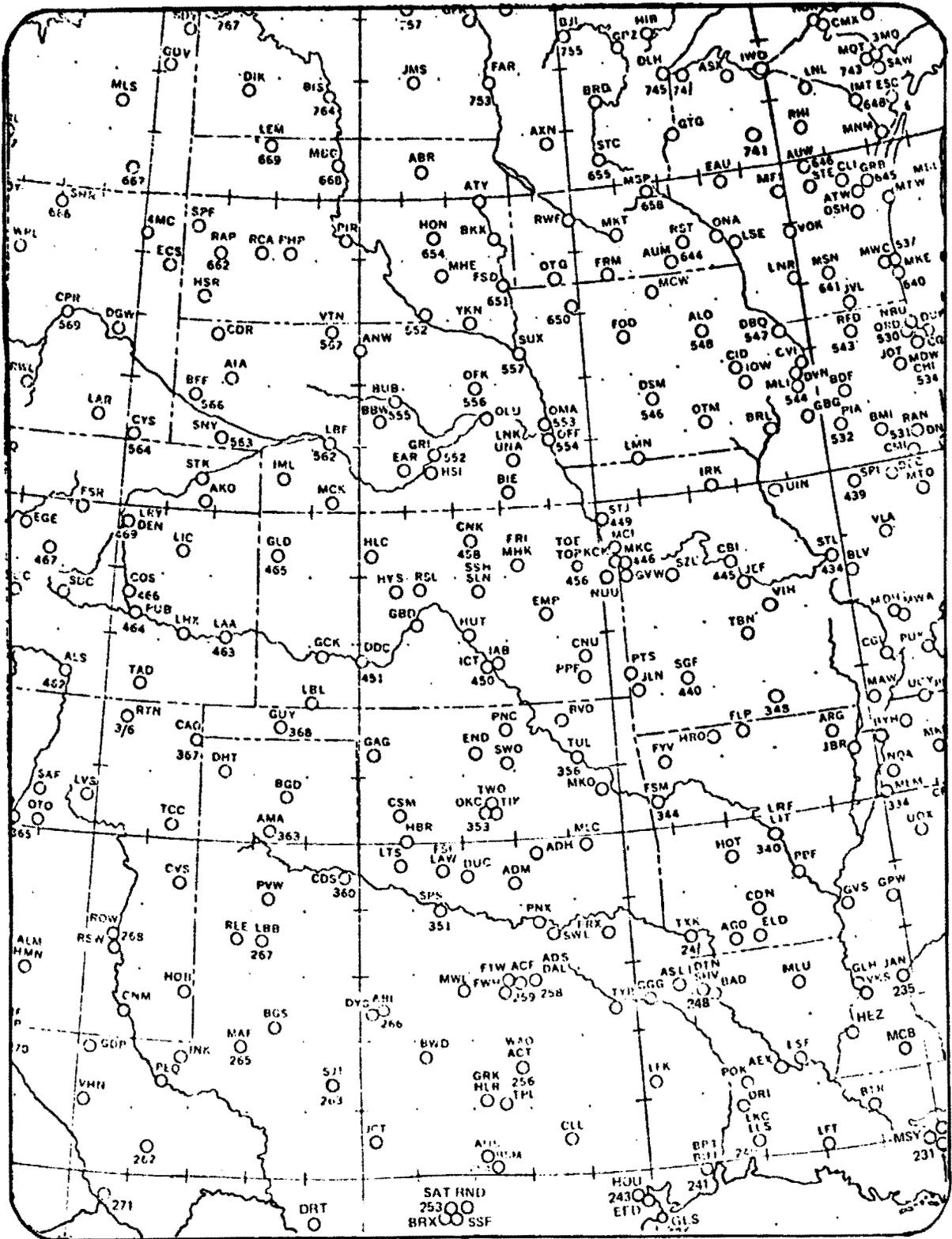


FIGURE 1.—Weather observing network in the U.S. Great Plains.

dates. In addition, a magnetic tape was prepared by the computer in Washington and mailed to Houston.

The individual station crop stage estimates at the end of the 2-week period were plotted, analyzed, and presented to the analysts in the form shown in figure 4. During Phase III, a computer program was developed which used the individual station estimates to prepare crop stage estimates interpolated in both space and time. A crop stage estimate was made for each specific Landsat segment location corresponding to the date of the satellite overpass. This prevented the necessity of subjectively interpolating between locations and times on two consecutive maps.

As with any series of estimating equations, a crop development stage estimation model or its subroutine to estimate planting date may occasionally provide troublesome, inaccurate values. A procedure was developed in Phase II to allow the analysts who were reviewing the Landsat imagery to advise YES of obvious inaccuracies in the crop calendar. This analyst feedback was reviewed by a panel of agronomists, plant scientists, and agrometeorologists. The panel made its recommendation to the YES manager as to what corrections should be made to the crop calendar model for a station, a group of stations, or a region. These changes in esti-

mate of growth stage for each required station for a given day were made and documented, and the crop calendar model was rerun from that date to predict later growth stage estimates. This revised output was again reviewed for accuracy and further corrections.

A specific example of such modifications occurred in the 1977 crop year for spring wheat in the New Lands area of the U.S.S.R. The spring wheat starter model indicated planting in this area from April 25 to about May 5. Emergence was indicated by the analysts to have occurred from May 3 to May 15. Mid-June Landsat acquisitions in the area indicated very little visible small grains at that time. A growth stage of 2.4 was entered as the datum for each station where the analyst was able to detect significant areas of emerged small grains. These operational corrections were made and the crop calendar estimates remained reasonably accurate during the remainder of the crop year.

Weekly Weather Summaries

Each week, the YES meteorologists at JSC prepared narrative weather summaries describing growing conditions and likely crop response. These were based on the domestic and foreign assessments written in Washington by CCEA's Assessment Division, but they provided additional detail needed to aid the analysts in their interpretation of the satellite imagery. These summaries also helped the project management to understand the results being obtained, to evaluate problem areas, and to develop alternate techniques which would be more appropriate for application to these problem situations.

The summaries consisted of maps and charts illustrating current crop stages and depicting the distribution of precipitation, temperature, soil moisture, or other pertinent weather factors. The narrative material then described and discussed the growing conditions and their likely effect on the crop's development and appearance. In some cases, the narrative material included ancillary information such as agricultural attache reports or other onsite observations; however, such data were not widely available in a timely manner nor were these data sufficiently comprehensive, as they usually related only to local problem areas. The narrative material was an attempt to provide timely updates of crop conditions rather than to summarize the entire season for each region. A typical weekly update would appear as follows.

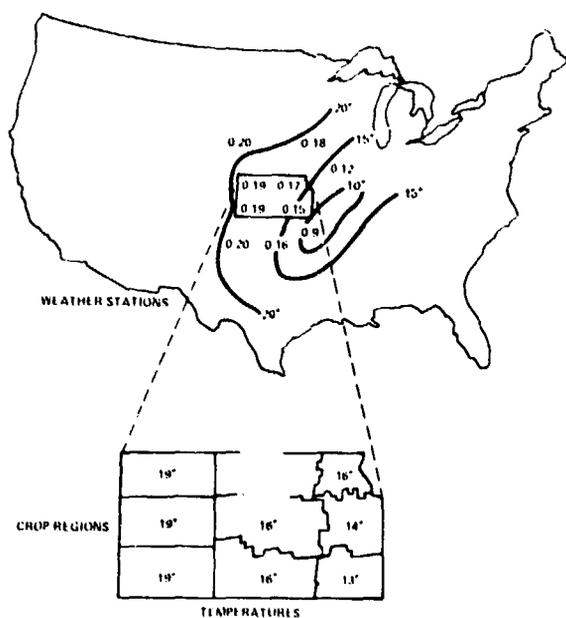


FIGURE 2.—Typical temperature analysis for input to yield model.

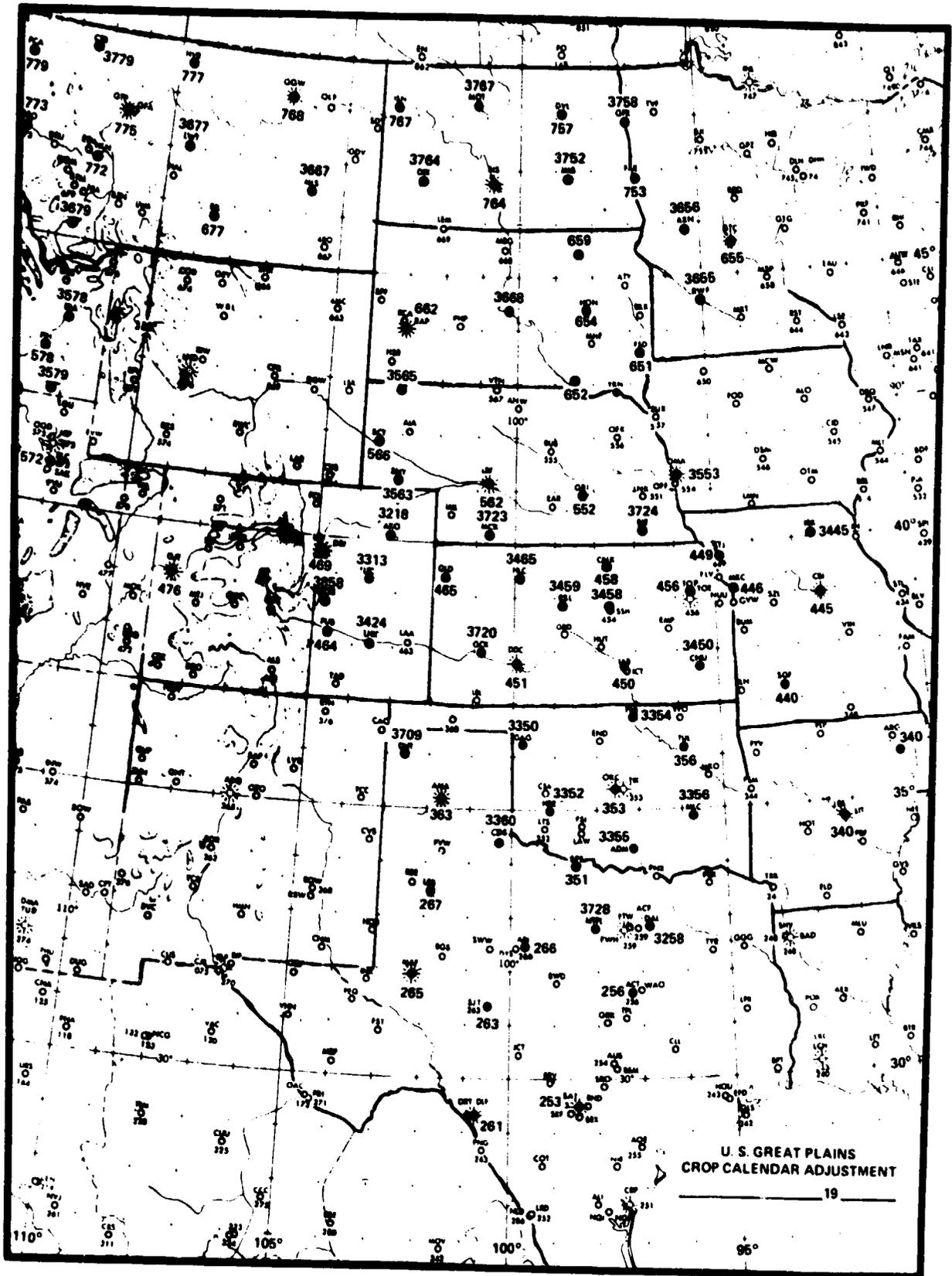


FIGURE 3.—Weather observing stations used for crop calendars in the U.S. Great Plains.

NEBRASKA

For the week ending May 29, 1977

WEATHER: Daytime temperatures dropped about 10° after midweek from around the mid 80's to the mid 70's.

Minimum temperatures remained in the high 50's and low 60's, dropping a few degrees at the end of the week.

Average temperatures ran 4° to 6° above normal.

Adequate rainfall occurred during midweek with some excessive amounts causing flooding in the northeast part of the state.

WHEAT: The growing conditions in the state for wheat continued to be very good. The above-normal temperature has caused the crop to mature ahead of the normal, but the wheat is expected to fill well with the abundant moisture. The wheat is all in the heading stage with some beginning to turn in the southeast and south central.

The printed summaries were distributed regularly to the analysts. Also, oral briefings were presented to enable specific areas and potential problems to be discussed in detail.

Typical charts prepared for the weather/crop assessments are shown in figures 5 and 6. In addition to temperature and precipitation, a soil moisture budget was available for the U.S. Great Plains and the U.S.S.R. The weekly soil moisture status for the United States was depicted by the crop moisture index shown in figure 7. The crop moisture index relates the available water to the normal requirements of the crops grown in each region. This analysis was extremely useful to project scientists who were monitoring drought and its effect on the appearance of wheat in the Landsat imagery.

Other specialized analyses were used to describe local problems and weather episodes. For instance, climographs, such as the one in figure 8, were prepared to determine the nature of weather trends over time at a particular location. In this instance, the onset of unusually cold temperatures near the normal wheat planting time in the northeastern Caucasus region of the U.S.S.R. limited fall establishment of the crop, and analysts were alerted to expect poor wheat signatures in that area at that time.

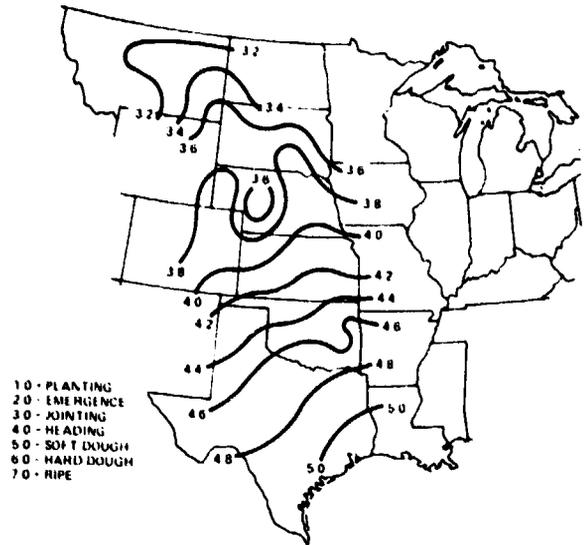


FIGURE 4.—Winter wheat crop stages predicted for May 1, 1977.

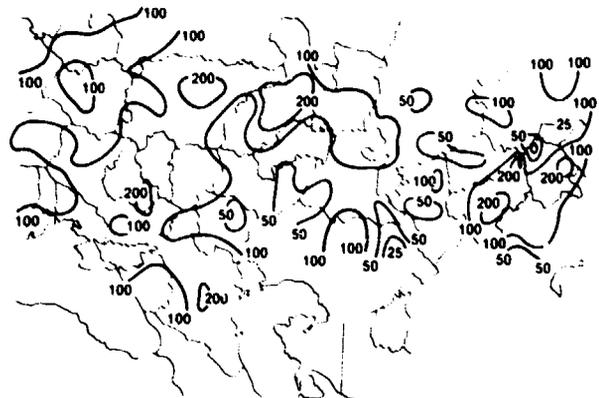


FIGURE 5.—Percent of normal precipitation for May 1977 in the U.S.S.R.

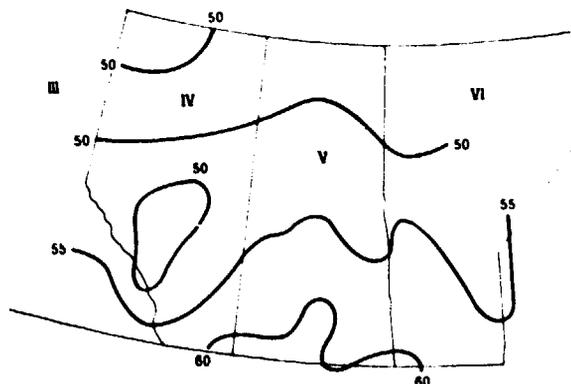


FIGURE 6.—Observed weekly average temperature in the Canadian prairie provinces, August 23 to 29, 1977.

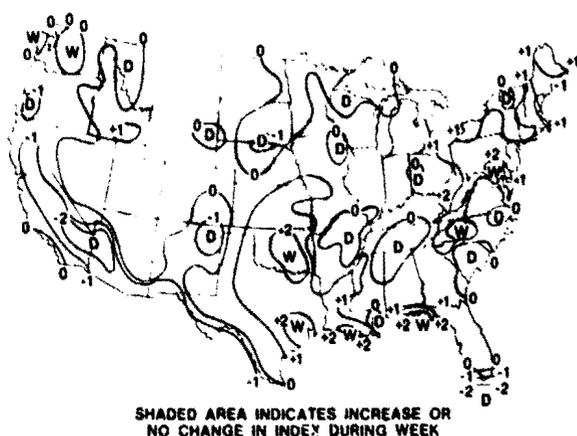


FIGURE 7.—U.S. crop moisture index: June 7, 1975.

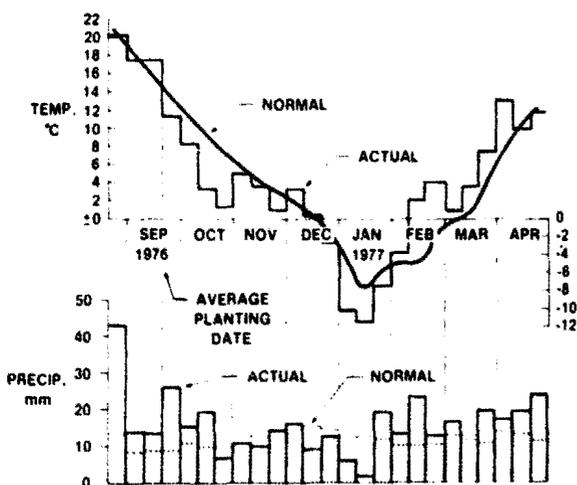


FIGURE 8.—Climograph for Krasnodar Kray, U.S.S.R.: 1976-77.

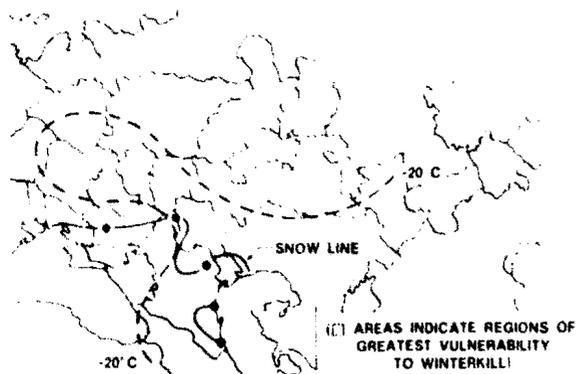


FIGURE 9.—Winterkill analysis for U.S.S.R.; January 5, 1977.

Another type of analysis was used that combined the temperature and snow cover data to determine areas where critically cold temperatures on a single night, without the benefit of an insulating blanket of snow, could have caused injury to the dormant crop. In figure 9, the analysis for the U.S.S.R. indicates a region in the southern Ukraine where conditions may have led to cold injury. In that area, analysts might expect a weak or mottled wheat signature in the spring because the damaged fields would not respond vigorously to warmer weather.

The weather assessments and analyses of particular episodes not only supported the analysts but also aided YES in assessing its own yield estimates. In most instances, the models did not account for reduced yields caused by events such as winterkill which occurred on a single day. The models were designed to be responsive to moisture or temperature stresses which were apparent in the monthly weather data. The episodal analysis made it possible to identify areas where additional influences on yield may have occurred; however, these factors could only be described qualitatively since there were no methods to adjust the yield estimates quantitatively for the shorter episodes.

Project Reports

In addition to the weekly weather assessments, other routine reports were prepared by YES at JSC. An overview of the growing-season weather was written and included in each of the Crop Assessment Subsystem monthly reports which released the project's estimates. A second section entitled "Yield Tracking" was also included to describe the response of each individual yield model to the growing conditions. The reports provided an opportunity to describe the various weather episodes or other problems which could affect yield but which might not be specifically accounted for in the models.

SUMMARY

The Yield Estimation Subsystem demonstrated during the three phases of LACIE that it is possible to use the flow of global meteorological data and provide valuable information regarding global wheat production. First, it was able to establish a capability to collect, in a timely manner, detailed weather data

from all regions of the world. Second, it was able to develop methods for evaluating the data and converting it into information appropriate to the project's needs.

Although the various elements involved in generating the products of YES were widely dispersed geographically, it was possible to coordinate their efforts and to provide the needed information for integration with other project data. Most notable was the utility of the information, particularly the objectively derived yield estimates, which were demonstrated to be capable of isolating problem areas, such as the shortfall in the U.S.S.R. spring wheat during Phase III, and to do so early in the crop season. This information alone has significance to foreign commodity analysts, but it takes on additional mean-

ing when combined with the LACIE estimates of wheat area available for harvest.

The techniques developed and demonstrated by YES to monitor and qualitatively assess the significant growing-season weather factors have added a dimension to global crop assessment capabilities. The demonstrated timeliness and available detail can provide early warning of significant weather conditions and alert analysts to likely effects. Even though these effects are not quantified, they can be very useful by simply pointing the direction in which production may depart from normal. The further development of these analysis techniques and the refinement of the yield models will be major activities for the group which succeeds the LACIE Yield Estimation Subsystem.

The Crop Assessment Subsystem: System Implementation and Approaches Used for the Generation of Crop Production Reports

W. E. McAllum,^a R. E. Hatch,^b S. M. Boatright,^c C. J. Liszcz,^d and S. M. Evans^b

INTRODUCTION

The primary responsibility of the Crop Assessment Subsystem (CAS) during the three phases of the LACIE was to produce crop reports that included estimates of wheat area, yield, and production, as well as a specified set of associated statistical descriptors. Report preparation and transmission were based on a documented reporting schedule which was reviewed and approved by project management at the beginning of each phase of the experiment. Generally, monthly reports were submitted; however, procedures existed to allow for the transmittal of an unscheduled report if circumstances warranted. Annual reports were used not only as a means of documenting the end-of-season wheat estimates but also to document results obtained by re-creating (simulating) estimates for each of the scheduled reports using the end-of-phase capabilities (i.e., latest aggregation software and/or approved procedural changes).

Successful performance of assigned CAS functions was heavily dependent on input data provided by other elements of the Applications Evaluation System (AES) and on the aggregation/report generation capabilities (status of system implementation in terms of hardware and software) available for CAS analyst use.

The purpose of this paper is to provide insight regarding CAS operations during the three phases of LACIE in terms of sampling strategy, CAS input/output data, evolution of aggregation/reporting system capabilities, and CAS aggregation procedures.

^aNASA Lyndon B. Johnson Space Center, Houston, Texas.

^bU.S. Department of Agriculture, Houston, Texas.

^cFord Aerospace & Communications Corporation, Houston, Texas.

^dLockheed Electronics Company, Inc., Systems and Services Division, Houston, Texas.

SAMPLING STRATEGY

The term "sampling strategy" is used to encompass the entire realm of methodologies involved in the definition of the basic sampling unit, the allocation of sample segments to specific political or administrative units, and the actual geographic location of the sample segments. A Sampling Strategy Team (SST) was established as the control agent for developing and/or modifying sampling allocation and location procedures. In addition, the SST was responsible for specifying the basic aggregation/expansion framework and the appropriate formulations for a specified set of statistical descriptors. This total group of functions is extremely critical since (1) the details inherent in the selected sampling strategy are primary determinants of expansion methods to be used; (2) a determination is made regarding the appropriate level at which essentially independent estimates of area and yield should be combined to estimate production; and (3) the techniques or formulations of statistical descriptors to be used in evaluating the accuracy and reliability of results are specified. Basically, the culmination of these functions determines many of the primary characteristics that a software package must possess to adequately support an aggregation/reporting function.

One of the major objectives of the initial sampling strategy was to allocate sample segments at the lowest (or smallest) political subdivision for which historical data were available. For example, segments were allocated at the county level in the United States and at the oblast level in the U.S.S.R. The initial sample segment allocations were based on the total area devoted to wheat production in a selected year, hereafter referred to as the epoch year. In foreign areas, the most recent available data were used for allocation purposes; the 1969 U.S. Census of Agriculture was used as the data source to support

the allocation process in the United States. The actual location of sample segments was confined to agricultural areas (i.e., areas having discernible field patterns), as defined by interpretations of available Landsat imagery. In addition, segment locations could not violate a set of prespecified constraints established by the NASA Goddard Space Flight Center (GSFC). These constraints included (1) a minimum distance between locations and (2) a restriction on the number of segments that are contained in a full frame of Landsat data. Supportive data were provided by the Data Acquisition, Preprocessing, and Transmission Subsystem (DAPTS) of the AES. More intricate details of the initial sampling strategy and modifications during the three phases of LACIE are available in the published requirements document (ref. 1) and are also addressed by other papers prepared for this symposium (presentations and supporting papers in the Experiment Design Session, which addresses sampling and aggregation). Country-specific issues are referred to in several papers that report results for individual countries.

BASIC INPUT DATA TO THE CROP ASSESSMENT SUBSYSTEM

As was previously mentioned, CAS utilized data elements provided by other elements of the AES to generate estimates of wheat area, yield, and production for specified geographic areas. The general categories of data required were (1) proportion estimates of wheat for each sample segment represented by usable Landsat data; (2) yield estimates and associated estimates of yield variance for predefined geographic areas; and (3) historical statistics, including wheat area, yield and production, as well as area for major competing crops grown during the wheat-producing season. More detailed discussions regarding these input categories follow.

Landsat Data

The CAS aggregation software was designed to utilize an estimate of the percentage of wheat for all sample segments allocated within a country as the basis for area expansion. Historical wheat area data, specifically that designated as epoch year data, were used in conjunction with available Landsat-based estimates to produce a wheat area estimate for

geographic units that were not assigned sample segments in the original allocation and/or that did not have Landsat data acquired for allocated sample segments (because of cloud cover, haze, correlation problems, etc.). Classification results (proportion estimates of wheat or small grains for a sample segment) were produced by the Classification and Measurement Subsystem (CAMS) and transmitted to CAS for use as input data in the aggregation process. Each segment-level classification result was identified by segment number and associated with a Landsat acquisition date, a date transmitted to CAS, an evaluation code, and several classification-oriented factors (e.g., unitemporal or multitemporal classification, bias corrections, analyst remarks, etc.). Key characteristics were identified and eventually implemented as control parameters during the evolution of the basic aggregation system.

Methods of transmitting and handling segment-level data changed significantly during the course of LACIE. The primary forces that encouraged changes were (1) the necessity for accommodating increased data loads; (2) a desire to minimize the chances for transcription errors; and (3) the development of specific input formats to support aggregation software. In Phase I, segment-level results were transmitted to CAS via a worksheet prepared by the CAMS analyst. It was then necessary for CAS personnel to code needed data in the appropriate format, to have the data set keypunched, and to perform sufficient checks to ensure that an accurate data set was ready for entry into the aggregation data base. Significantly increased data volumes were anticipated for Phases II and III, and it became apparent that data-handling tasks were likely to be a bottleneck in report preparation. As input formats for the CAS system stabilized, a procedure was established for CAMS to provide segment-level results via punched cards utilizing a prespecified format. This procedure provided the CAS analyst the capability of handling large quantities of data in a relatively short period of time. However, data quality checks in terms of accuracy and completeness of data transmitted were still necessary.

Confusion Crop Ratios

Increasingly large overestimates of wheat area during Phase I alerted LACIE personnel to potential problems in segment-level analyses. Subsequent investigations identified the major problem as the con-

fusion of wheat with other crops being grown in the sample segment. The principal source of confusion was other small grains crops that had a growing season similar to that of wheat. Since this crop separability was not an easily resolved issue, project management required CAMS analysts to identify small grains (spring, winter, or total) for each segment, especially in spring and mixed wheat areas, until reliable techniques and procedures could be developed to identify spring and/or winter wheat. Since sampling and aggregation procedures and all supporting data bases were designed to estimate wheat area, the implementation of small grains estimation at the segment level necessitated the development of confusion crop ratios that could be used to derive the required spring and/or winter wheat proportion estimates for each segment.

Initially, confusion crop ratios were applied only in spring and mixed wheat areas (e.g., Minnesota, Montana, North Dakota, and South Dakota). This procedure was based on the assumption that winter wheat classifications and winter small grains classifications in the pure winter wheat areas of the U.S. Great Plains (USGP) were essentially synonymous. During Phase III, however, the use of confusion crop ratios was extended into the pure winter wheat areas. It is important to recognize that the relative accuracy of this ratioing procedure is heavily influenced by two factors: (1) the degree to which the ratios being used reflect the true distribution of crops in the current year, and (2) the accuracy of the classification process in terms of including all confusion crops (e.g., small grains) in the segment-level proportion estimate.

At the end of Phase I, confusion crop ratios were constructed using state-level historical data for the previous crop year. These ratios were then applied to segment-level small grains estimates prior to aggregation.

The basic approach used during Phase II for the U.S. Great Plains, the Canadian prairie provinces, and the U.S.S.R. was to derive needed ratios from the most recent data available reported for the lowest political subdivision identified in the allocation hierarchy. The four necessary ratios were (1) winter wheat to winter small grains, (2) winter wheat to total small grains, (3) spring wheat to spring small grains, and (4) spring wheat to total small grains. Small grains crops considered in deriving ratios were rye, barley, wheat, oats, and flax. During the classification process, the CAMS analyst identified the proportion estimate for a segment as winter

wheat, spring wheat, winter small grains, spring small grains, or total small grains. In the CAS data base update process, the proper ratio was applied to obtain appropriate spring or winter wheat proportions, thus creating two additional classes (i.e., ratioed winter wheat and ratioed spring wheat).

A task was initiated in December 1976 to use econometric modeling techniques to estimate confusion crop ratios for the four U.S. Northern Great Plains states and for Canada. Ratios estimated by the developed models were used in Phase III analyses. (For further details, see "Econometric Models for Predicting Confusion Crop Ratios" by D. E. Umberger et al., which is included in the symposium proceedings as a supporting paper in the Experiment Design Session.) Because of existing time constraints, limited resources, and an anticipated lack of necessary supportive data, similar modeling efforts were not attempted for countries having planned economies. If confusion crop ratios are deemed a necessary element of future Landsat-based crop estimation systems, further testing and evaluation are needed to ensure that optimal techniques are used to estimate any required ratios.

Data Editing Procedures

During Phases I and II, all data received from CAMS were resident in the aggregation data base and were used to support aggregations throughout the season; i.e., a segment classification became inactive only if replaced by a later classification. During Phase III, it was suspected that early-season acquisitions failed to detect all planted wheat because of poor and/or later emergence. The following country-specific methods were used to delete the "questionable" acquisitions from the data base.

1. U.S.S.R. newspapers were used to determine wheat tillering dates for oblasts, and segment data acquired prior to the established date were eliminated from the aggregation process.

2. In the U.S. Great Plains, rates of change in segment wheat proportion estimates were monitored for segments that had multiple acquisitions. At the average date when the rates of change became small, the crop-growth stage was estimated, and all segment estimates based on Landsat data acquired prior to the derived growth-stage date were excluded from the aggregation process.

A screening procedure, which identified outlier wheat proportion estimates via comparison of seg-

ments with similar historical county statistics, was used in the U.S. Great Plains as a tool for excluding questionable segment estimates from the aggregation procedure.

Yield Estimates

Estimates of wheat yield and the associated estimates of yield variance were provided by the Yield Estimation Subsystem (YES). Phase I was a testing period for yield models; thus, a schedule for generating and transmitting yield estimates and variances was not established. Beginning in Phase II, yield estimates for active countries were generally provided on a monthly basis during the growing season. Yield variances were not available for the U.S.S.R. and Canada until Phase III.

The CAS aggregation software was designed to utilize yield estimates at the stratum level; e.g., the Crop Reporting District (CRD) level in the United States. In Phase II, the boundaries of the yield strata (area represented by a specific model) were defined as CRD's in the United States, as Crop Districts (CD's) in Canada, and as crop regions in the U.S.S.R. However, the historical data used to develop U.S. and Canadian models represented an area larger than a CRD/CD (e.g., several CRD's/CD's or an entire state/province). Therefore, it was necessary to adjust the computation of yield and production statistical descriptors to account for correlations resulting from the definition of model development boundaries which did not match the boundaries used to delineate the area stratum. Input of yield estimates and variances was retained at the stratum level in the CAS software to avoid extensive (and expensive) software modifications.

Basic variables in the yield models are monthly averages of specified weather parameters such as temperature and precipitation. Thus, yield estimates for a particular month would include weather data through the end of the previous month. The schedule for delivery of yield estimates and variances was established to allow time (1) to include the previous month's weather in yield model updates; (2) to operate the models; and (3) to support an established reporting schedule for wheat area, yield, and production. The yield delivery schedule adhered to during Phases II and III was (1) estimates for the United States delivered to CAS on the fourth working day of each month and (2) estimates for foreign areas delivered to CAS on the ninth working

day of each month. Yield data were transmitted via telephone, telefax, and magnetic tape.

CROP ASSESSMENT SUBSYSTEM OUTPUTS

The primary product of the CAS has been crop reports containing estimates of wheat area, yield, and production for each country that was actively being worked in a particular phase of the LACIE. The general format and content of the reports were controlled by an Interface Control Document (ICD) between CAS and the Information Evaluation (IE) group (USDA/LACIE, Washington, D.C.) (ref. 2). Reports included formatted computer outputs of area, yield, and production estimates and a set of statistical descriptors (standard errors, coefficients of variation, probability of less than 10 percent error, and 90-percent confidence intervals—upper and lower) associated with each of the included estimates; output tables were provided at the country, region, zone, and stratum levels. In addition, summary tables of supportive data (segment-level data and stratum-level yield estimates) were provided. A narrative section of the report was utilized to summarize results and to present pertinent analysis of input data with special emphasis on major factors responsible for changes between report dates. Modifications were made in the narrative section to accommodate special needs, such as an assessment of drought conditions or an unusual country-specific situation. During Phase III, a crop condition assessment section was added to all reports in order to more adequately highlight unusual circumstances that could impact the potential production level of a country (e.g., drought and winterkill).

Report Schedules

As mentioned previously, only estimates of wheat area were produced in Phase I. Thus, the established report schedule was based primarily on projected LACIE processing capabilities and a stated requirement for monthly estimates. Phase I reports were generally prepared on the last working day of the month. Only an end-of-season report was transmitted to IE for evaluation purposes.

In addition to the requirement to generate monthly estimates, reports in Phases II and III were to include estimates of yield and production, were to be mailed prior to comparable official releases by

USDA, and were to demonstrate an operational capability to prepare and release crop reports in a timely manner. This total set of requirements, in combination with established input data availability constraints, resulted in the following reporting scenarios.

1. U.S. Great Plains

a. Reports should be mailed no later than the day preceding an official release of domestic estimates by USDA.

b. Yield estimates should be received on the fourth working day of the month.

c. The flow of segment classification results would continue up to and include the scheduled aggregation date.

2. U.S.S.R.

a. Yield estimates should be received on the ninth working day of the month.

b. Phase II reports were scheduled for 5 to 7 working days following receipt of yield estimates.

c. Phase III report dates were adjusted to support scheduled meetings of the FAS/USDA/U.S.S.R. Grain Estimation Task Force.

In summary, tight scheduling in terms of report release dates and availability of required input data necessitated efficient, accurate handling of large data volumes as well as a quick analysis and documentation of results.

The following summarized the countries covered by crop reports.

1. Phase I—U.S. Great Plains: spring and winter wheat, area only

2. Phase II—U.S. Great Plains: area, yield, and production for spring and winter wheat; U.S.S.R.: area, yield, and production for a winter wheat indicator region and for a spring wheat indicator region; Canada: area, yield, and production of spring wheat for the three prairie provinces

3. Phase III—U.S. Great Plains: area, yield, and production for spring and winter wheat; U.S.S.R.: full country estimates of area, yield, and production for spring and winter wheat

Security for Commodity Estimates

Although LACIE was formally designated an experiment, it was recognized that the ultimate product (i.e., crop estimates) should not be widely distributed until appropriate evaluations were completed. In addition, it was particularly important to avoid having

experimental results confused with or mistakenly interpreted as official releases by the USDA. Therefore, a commodity data control plan (ref. 3) was implemented to provide necessary control guidelines. Some of the salient features of the security program are highlighted here; if more detail is desired, interested parties are referred to the referenced document.

A controlled access area was established to house required hardware (remote terminals, hardcopier, and printer), to provide workspace in which reports could be assembled and prepared for mailing, and to provide a storage area from which material of a sensitive nature could be distributed in accordance with the established commodity security guidelines.

Two basic protection periods were established to cover materials (all reports and briefings) which contained aggregated area, yield, or production estimates. A maximum protection period was defined as extending from the date of issuance until the next working day following the official release of an estimate by USDA. For all practical purposes, access to data during this period was limited to CAS, YES, Accuracy Assessment, and project management personnel. During the restricted access period, all controlled data were available to the LACIE staff in order to support evaluation efforts, the preparation of technical reports, program modifications, and other assigned project-related tasks. The duration of the restricted access period was 4 months from the end of the maximum protection period. The cover and each page of all controlled documents were to carry ending dates for both protection periods.

EVOLUTION OF AGGREGATION/REPORTING SYSTEM CAPABILITIES

Currently existing capabilities in terms of hardware and software available to support aggregation and reporting functions were not available at the beginning of LACIE. The CAS grew from a relatively simple, individual dependent system to a fairly complex system that is usable by several trained analysts that have diverse academic and job experience backgrounds. This evolutionary process was strongly influenced by the rate at which requirements were defined and documented, by the time required to develop and receive concurrence on statistical formulations, and by the degree to which project personnel understood the characteristics and poten-

tial uses of available input data. The following sections trace the development of the CAS from Phase I through Phase III. It is important to recognize that existing system capabilities at any particular point during the LACIE experience are necessarily reflected in the quality and the quantity of results obtained.

Phase I

The CAS development system was designed, developed, implemented, and operated on a UNIVAC 1110 computer located in Building 12 at the NASA Johnson Space Center (JSC). All interfacing was through the demand terminals located on the second floor of JSC Building 17.

The aggregation equations implemented were those specified in the requirements for the LACIE Phase I CAS (ref. 4). These equations were basically unchanged through LACIE Phase III, except for a logic change requiring a minimum number of aggregatable segments before a direct area estimate would be made at the stratum level. Otherwise, the stratum would be estimated by applying a zone-level ratio of the current estimate to historical data.

The variance estimation equations, the so-called standard statistics, were implemented in a like manner. However, as data were collected and the algorithm was run, the resulting estimates indicated some erroneous assumptions concerning the manner in which wheat was distributed at the substratum level. Further studies were conducted (ref. 5), and modifications were made to the variance estimation equations. These modifications were based on real-time data and reflected new assumptions concerning the within-substratum wheat distribution.

These new standard statistics became available late in LACIE Phase I and were incorporated into the development system software.

Near the end of LACIE Phase I, further studies of the wheat distribution were conducted to evaluate the standard statistics model with respect to the latest findings. The assumption that the wheat was relatively homogeneously distributed throughout the sampled area of a zone was found to be invalid. However, usually within a zone, two or more substratum groups could be found that were fairly homogeneous. Based on these findings, a new model was developed, tested, and implemented and then was used to recalculate the LACIE Phase I standard

statistics (ref. 5). This model was then used throughout LACIE Phases II and III.

The development software and data bases were converted from the UNIVAC 1110 to the Programmed Data Processor, model 11-45 (PDP 11-45) computer in JSC Building 17 late in Phase I in preparation for controlled access operations during Phase II. The system was further modified in Phase II to support the operational requirements of CAS.

Phase II Batch Aggregation Software

The batch aggregation software was developed to bridge a gap between the UNIVAC 1110 system and the interactive system. The CAS analysts defined data-handling and report requirements for Phase II which could not be met by the Phase I software, even though it had been converted to run on the PDP 11-45. These requirements were included in the interactive software; but, because of the complex nature of an interactive system, they required a 9-month lead time from definition to implementation. The decision was made to design and implement an interim system to be executed in a batch mode that would support generation of CAS reports beginning in March 1976. This system also would have the capability of providing Phase II report formats and data handling.

Design.—The batch system was designed to provide the capability of aggregation qualified by date of acquisition, date passed to CAS, biostage, evaluation code, and level of allocation hierarchy. Requirements also indicated a need to construct a yield estimate data base with predefined edit specifications and to manipulate that data by making additions, changes, and deletions. A similar requirement for maintaining a CAMS data base (containing the date of acquisition, date passed to CAS, biostage, evaluation code, and proportion of wheat in the segment) existed with the addition of requiring a limited query capability. The Phase II CAS requirements called for 10 basic report formats with variations of each according to the desired level of hierarchy. Of these 10, 4 were levied against the interim batch system. These were the area estimate, yield estimate, area-yield-production summary, and production estimate reports.

Summary.—The interim batch system was a useful means of meeting early Phase II CAS software requirements. One of the major factors in the comple-

tion of the design was the cooperation of the CAS analysts in freezing new requirements until after the system had been declared operational and clarifying quickly any vague or inconsistent areas in the established requirements. This afforded the designers and programmers an opportunity to complete their work without major revisions.

Phase II Consolidated System

The interactive system software was not delivered until May 1976 and then only with partial capability. The system was more advanced in aggregation and statistical estimates, but the data bases and output reports were behind the batch system because of the longer lead time required to incorporate design modifications. To obtain the best of both systems, the batch and interactive systems were combined into a consolidated CAS system. The batch data bases were made compatible with the interactive aggregation software through an extract program. The results of the aggregation were extracted from the interactive system and run through the batch report generator to obtain the IE reports. This type of operation was continued through Phase II.

CAS Interactive Software—Phases II and III

The software discussed in this section provided CAS analysts an interactive estimation and reporting capability in a single operational system. The first system delivery occurred in May 1976. Several modifications have been made since. The final version (delivered in July 1977) is described below.

Basic design decision.—The estimation and reporting software was implemented on the PDP 11-45 computer in an interactive, multiuser environment. This necessitated the choice of the Resource Sharing Executive, model 11D (RSX-11D) operating system. The Fortran language was chosen for software implementation because it was the only language available on the PDP 11-45 with enough versatility for a complex system.

Requirements were for a software system which would be operated by the CAS country analysts. Since the analysts' training and experience varied widely, the interactive interface between the CAS analyst and the CAS software included a relatively complete set of tutorial prompts to guide the analyst

through various phases of system operations.

The CAS software performs the following functions.

1. Processing analyst inputs in response to tutorial prompts.

2. Processing and storage of data from external sources; i.e., CAMS, yield estimates, historical data, and confusion crop ratios.

3. Crop area, yield, and production estimation. Area estimates are computed from CAMS and historical data; production is computed from area estimates and input yield estimates; and average yield is computed at the higher hierarchical levels by dividing production by area.

4. Computation of standard statistics associated with the crop estimates to assess the reliability of the estimates.

5. Aggregation of crop estimates to the various hierarchical elements of the active LACIE countries.

6. Generation of aggregation reports.

The software is subdivided into functional categories which include batch programs to initialize the data base; data base management software; data base change software; area, yield, and production estimation; and report generation programs.

Data base design.—The CAS data base must include data such as classification results for each sample segment (CAMS data); yield estimate data; historical area, yield, and production values; data describing the hierarchical structure of the country; and small grains historical statistics used to define the confusion crop ratios. In addition to the data files and allocation files, index files are used to access the data files.

Dual software systems, both batch and interactive, are used to maintain the data base. The batch system initializes or makes large-scale updates to the data base. The interactive system is used for small updates or diagnostic changes to the data base. Both systems produce a line printer record of the data base transaction in the form of a file dump or a data base change report.

Data base security.—The operating system secures the data base through User Identification Codes (UIC's) and passwords. The analyst does not work directly with the master data base, but instead works with his own copy. A copy can be made only by individuals with access to the system password or by analysts with passwords to their own production UIC. The analyst's copy of the master data base is secured by the analyst's password. Thus, several

analysts can be working simultaneously, each with a secured copy of the data base for the LACIE country with which he is working.

Analysts can add files to the master data base, but only individuals with access to the system passwords can delete files or change files in the master data base.

Application program design.—Three types of application programs are in the CAS software system. One type includes all programs to perform aggregations; i.e., the area, production, and yield estimations. Another type consists of the report generators. The last type is the data base change program.

The aggregation programs are split into three tasks—area, production, and yield. Yield estimates are input for each stratum, and CAMS data are input for each sample segment, along with historical area data and a set of analyst-specified parameters. Outputs are area, production, and/or yield estimates and supporting statistics for the requested hierarchical element and all of its hierarchical subelements. These outputs are linked to the report generators through the application report files. All of the aggregation tasks have separate report files for spring, winter, and total wheat.

Two report generators—interactive terminal and line printer—use the application report files to produce reports in two different formats. The interactive reports are displayed on the cathode-ray tube (CRT) with optional line printer output; information is obtained directly from the report files and displayed with a minimum of reformatting. The second report generator produces line printer output in a fixed format that is not compatible with the CRT display; these are intended for use by the IE in Washington, D.C. Since this generator gathers information from many different files for each single report, it requires significantly more time to complete than the interactive report generator.

A third report generator obtains needed information from the data base rather than from the application report files. It outputs yield estimate data using analyst-specified parameters. This generator enables the analyst to see only the data which satisfy the specified input criteria and which are qualified for use by the aggregation programs.

The data base change program is the interactive counterpart of the batch programs for updating the data base. This data base change program, together with the interactive report generator, can be used as a diagnostic tool. Temporary data base changes can be entered and aggregation results viewed immediately

on the CRT to assist the CAS analyst in evaluating test aggregations and in complying with requests from Accuracy Assessment.

OVERVIEW OF CAS OPERATIONS

This section provides an overview of the operations within CAS that were necessary to support established reporting schedules. CAS analysts used documented procedures that covered the CAS from computer terminal operation to report distribution. In general, the procedures guided the analyst through data base initialization and maintenance, selection of aggregation parameters, computer report generation, and preparation of the total CAS report. A new learning situation was encountered whenever technical modifications or procedural changes resulted in changes to the existing software.

The CAS operations were initiated in December 1974 with a test aggregation of 28 segments collected in Kansas for crop year 1973-74. This manual aggregation was used to verify expansion algorithms that were implemented to support Phase I area estimation tasks. As might have been expected, considerable time was devoted to development of procedures and documentation of requirements early in the project. Because of the experimental nature of the LACIE, continual changes in procedures were necessary, both to improve the quality of the output and to solve unexpected problems. Figures 1, 2, and 3 illustrate the technical modifications that were made during the three phases of the LACIE and correlate them in terms of time with CAS system deliveries and reports that were issued.

The following discussion focuses on data bases used, aggregation methods, time lines required for report generation, and analyst interaction with other project elements.

Data Bases

Separate data bases were defined for each country. These data bases were updated with segment proportion estimates and yield estimates as required to obtain timely area, yield, and production estimates for each level in the aggregation hierarchy (i.e., the stratum to the country).

The primary data base required by the CAS system is the allocation data base, which defines the country's area aggregation hierarchy, the yield strata,

and the location of the sample segments. The allocation data base is initiated when a country is activated on the CAS system and requires no updates unless the aggregation parameters are modified (e.g., the number of segments, agricultural area, total political subdivision or hierarchy). The data base contains the same data that were used to define the sample for a country. The system will not accept other types of input data and will not aggregate unless the allocation data base is populated and the hierarchy is correctly defined.

The historical data base contains space for 15 years of historical area, yield, and production data. Three of these years are dedicated to primary, secondary, and allocation epoch years, which leaves 12 years for other historical data. The primary epoch

year is used for ratio estimating and area statistical calculations. These data are used for determining the maximum, minimum, average, and variance of the crop of interest in the hierarchy for reporting and analysis purposes. The data base can be updated as desired. The primary epoch year is required for the system to aggregate.

A confusion crop ratio for each segment is maintained in a data base to ratio winter and/or spring wheat from the CAMS small grains estimates. These ratios are used to obtain the ratioed wheat stored in the CAMS data base, as previously defined in the section entitled "Basic Input Data to the Crop Assessment Subsystem," subsection "Landsat Data." Ratios can be calculated external to the software and input for each segment, or the small grains

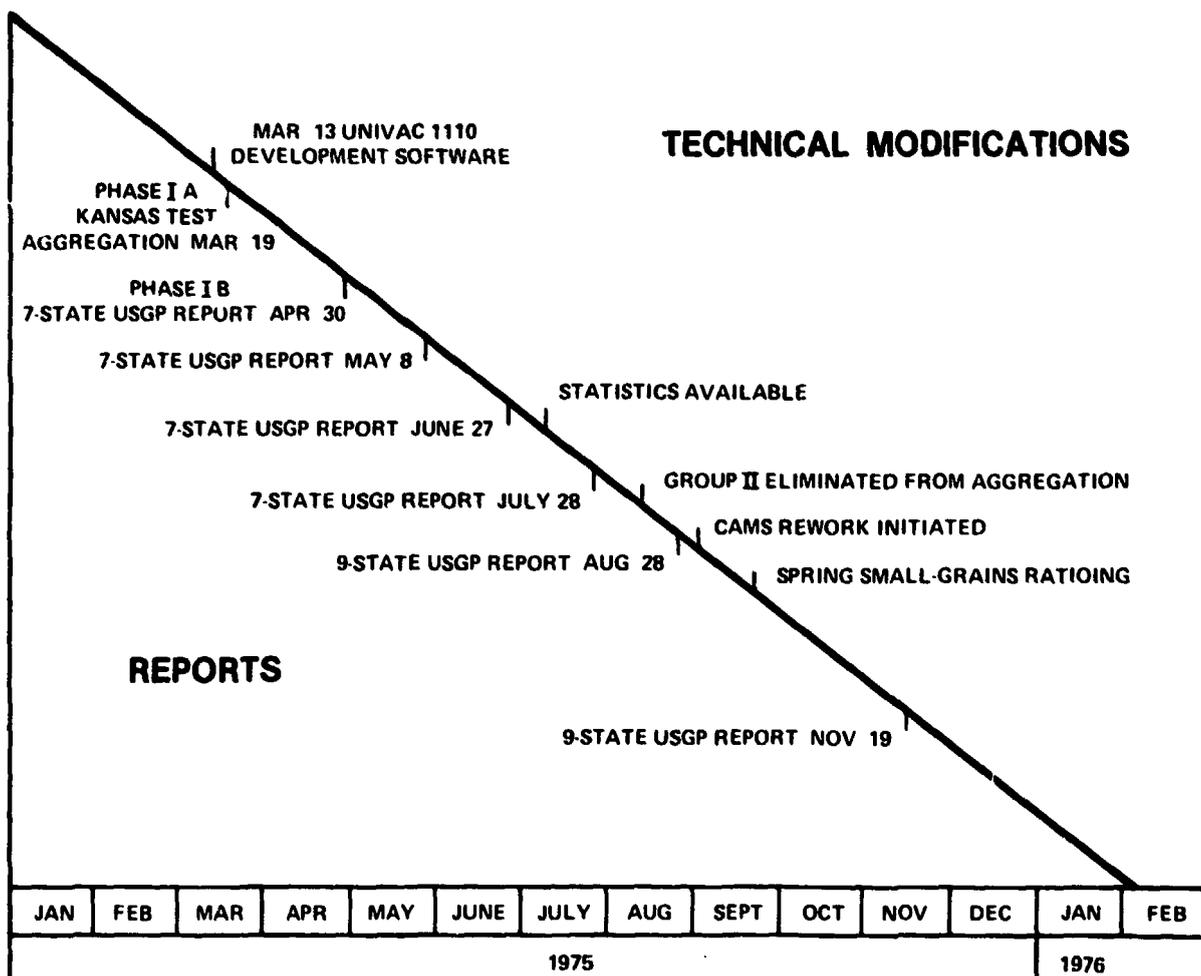


FIGURE 1.—Phase I CAS implementation/operations scenario.

historical data may be input and the software will calculate the ratios. The system will operate without confusion crop ratios in the data base, but the ratioed wheat estimates will be identical to nonratioed wheat estimates because the blanks are interpreted as a ratio of one.

The CAMS data base will accommodate all the sample segment classification results received with a classification code of 10 or greater for all segments in a country. (Table I lists the CAS classification codes.) Five classes are stored—winter wheat, spring wheat, winter small grains, spring small grains, and total small grains. From these classes, two additional classes (ratioed winter wheat and ratioed spring wheat) are obtained by applying the confusion crop ratios to the small grains estimates. In addition, the following parameters also are stored for each

classified segment; i.e., date passed to CAS, date of Landsat acquisition, and biostage of wheat development. The system will not aggregate a zone unless at least three aggregatable segments are available within a lower level of the hierarchy.

The allocation and, consequently, the aggregation hierarchy were based on political subdivisions to utilize the historical and census data that existed for these areas. The base level in a country was the lowest political subdivision for which detailed historical crop statistics were available. The hierarchy for each country became (1) United States—U.S. Great Plains—state, CRD, county; (2) U.S.S.R.—crop region—economic region oblast; and (3) Canada—prairie provinces—province, CRD, county. This resulted in the United States and Canada being substratum-level countries (because of

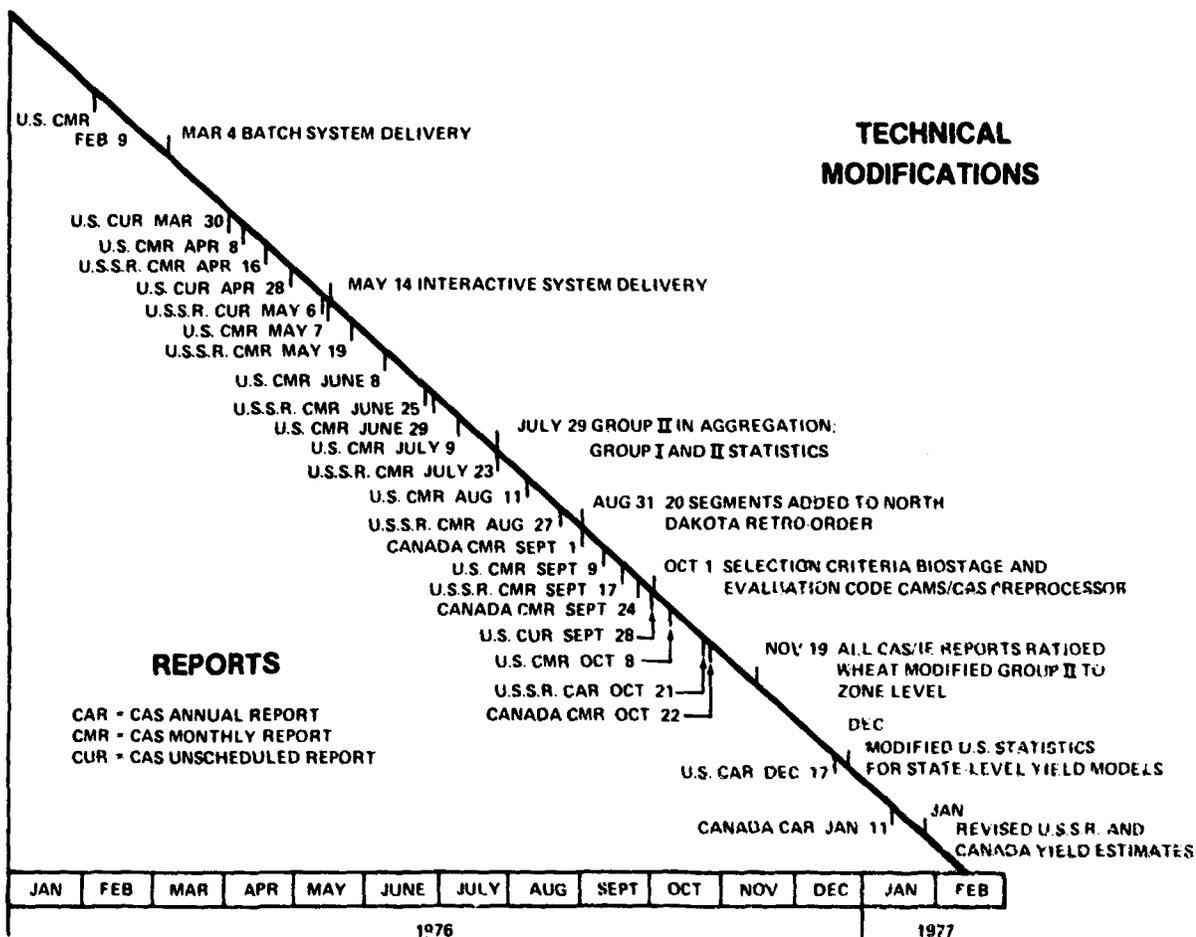


FIGURE 2.—Phase II CAS implementation/operations scenario.

the availability of historical data) and the U.S.S.R. being a stratum-level country (because of the non-availability of historical data).

The area was estimated for the base-level area by direct expansion from the segments and was summed to higher levels in the hierarchy. For base-level areas with no segments, the area was ratio-estimated using areas with segments available for aggregation and historical data.

The selection criteria for the CAMS data to be aggregated assumed that the latest data collected would result in the most accurate classification; therefore, the latest acquisition data and biophase were the primary selection parameters. The classifications were rated satisfactory, marginal, and unsatisfactory by CAMS evaluation procedures, and the satisfactory and marginal ratings were considered acceptable for aggregation. Although the unsatisfactory segments were not aggregated, they were carried in the CAMS

data base for information. The segment classifications were added to the data base as they became available to CAS or as required to produce scheduled reports.

As the experiment progressed in Phases II and III, it became necessary to add the CAMS classification date to properly handle segments that were reworked in CAMS. The four biowindows were expanded to seven crop development biostages (table II) to conform to LACIE crop calendar outputs. This parameter was added to the CAS criteria for selection of segments as a single biostage or as a set in order of predetermined priority. The selection criteria of CAMS segment data were classification date, acquisition date, biostage, and classification code.

The early-season winter wheat area estimates in Phases II and III were low because the LACIE system identified detectable wheat: i.e., wheat area with sufficient ground cover to be detected by

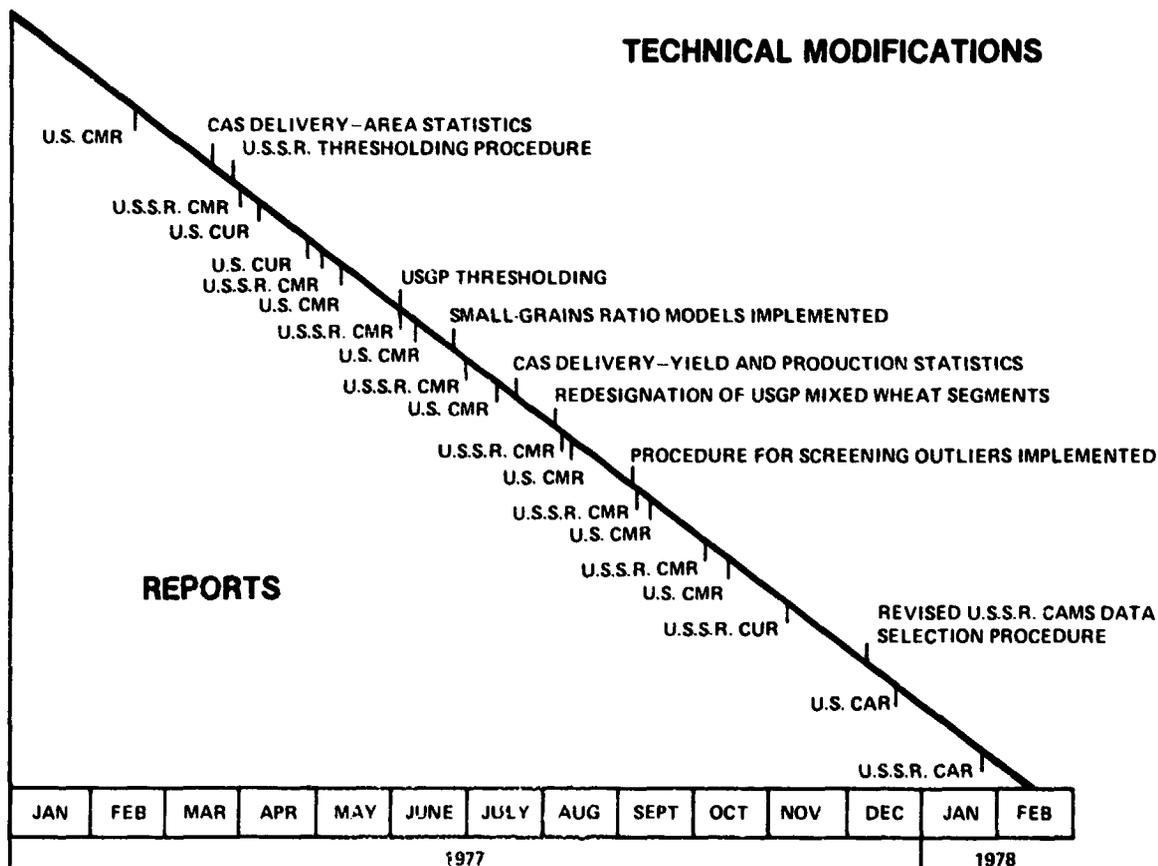


FIGURE 3.—Phase III CAS implementation/operations scenario.

TABLE I.—CAMS Evaluation Codes^a

Code	Description	Code	Description
01	Not machine processed—clouds, haze, snow, etc. This means that the segment cannot be processed through the system because clouds, haze, etc., make interpretation and analysis impossible.	10	Unsatisfactory—unsatisfactory results for segment This code is to be used for any acquisition that has been processed through the system and, based on CAMS evaluation procedures, designated unsatisfactory.
02	Not machine processed—confusion crops or other interpretation problems This code should be used when a segment cannot be processed because of interpretation difficulties, especially when confusion between wheat/small grains and other crops is such that a wheat/small grains decision cannot be determined. The rules for such a decision are to be negotiated between CAMS and CAS and included in the CAMS Detailed Procedures.	12	Unsatisfactory—no significant change This code is used when the new acquisition is evaluated to have no significant change from the previous unsatisfactory evaluation for the segment.
03	Bad data—due to technical problems, not reordered In this case, the segment cannot be processed due to technical problems arising from an unsatisfactory histogram, excessive scan-line drop, etc. If the segment is not reordered, this code is used. NOTE: If the segment is reordered, no code should be listed and no data sheet should be passed to CAS since the segment will again be sent through the CAMS for evaluation. Any segment that cannot be processed due to clouds, etc., or technical problems should be passed as 01 or 03.	14	Unsatisfactory—rework, reevaluated segment This code is used when a segment that was previously passed to CAS is reworked.
05	Not machine processed—dormancy In cases where recognition is a problem because the crop is in a state of dormancy, this code is used.	18	Unsatisfactory—machine-processed multitemporal analysis This code is used when more than one acquisition date is used to produce an unsatisfactory proportion estimate. NOTE: All acquisition dates used in processing should be listed on the CAMS Evaluation Form.
07	Not machine processed—preemergence In cases where the acquisition is prior to the criteria established for fall wheat recognition (to be determined), this code is used.	20	Marginal—marginal results for segment This code is to be used for any acquisition that has been processed through the system and, based on CAMS evaluation procedures, designated marginal.
09	Not machine processed—multiple acquisitions Code 09 was developed to take care of the problem of processing multiple acquisitions of a segment at the same time. When multiple acquisitions are available at one time, only one segment—which is determined to be the better acquisition for interpretation and analysis by the CAMS analyst—is selected for complete processing. Other acquisitions are listed as Code 09 to indicate that they have been reviewed but were not processed through the system. This code enables CAS to review the segment results and account for each acquisition and how it was evaluated by CAMS.	22	Marginal—no significant change This code is to be used when the new acquisition is evaluated to have no significant change from the previous marginal evaluation for the segment.
		24	Marginal—rework, reevaluated segment This code is used when a segment that was previously passed to CAS is reworked.
		28	Marginal—machine-processed multitemporal analysis This code is used when more than one acquisition date is used to produce a marginal proportion estimate. NOTE: All acquisition dates used in processing should be listed on the CAMS Evaluation Form.

TABLE I.—Concluded

<i>Code</i>	<i>Description</i>
30	Satisfactory—satisfactory results for segment This code is to be used for any acquisition that has been processed through the system and, based on CAMS evaluation procedures, designated satisfactory.
32	Satisfactory—no significant change This code is to be used when the new acquisition is evaluated to have no significant change from the previous satisfactory evaluation for the segment.
34	Satisfactory—rework, reevaluated segment This code is used when a segment that was previously passed to CAS is reworked.
36	Satisfactory—less than 5 percent manually (hand) counted This code is to be used for any segment in which the proportion estimate is manually counted rather than machine processed. NOTE: This category would not be used if the segment was a rework segment. Code 34 should be used.
38	Satisfactory—machine-processed multitemporal analysis This code is used when more than one acquisition date is used to produce a satisfactory proportion estimate.
40	Segment is totally nonagricultural This code is used when the segment is evaluated as having no agriculture at all; i.e., no discernible field patterns. NOTE: Segments in an agricultural area that have a 0-percent proportion estimate are to be designated Code 30.

LACIE classification methods. If all wheat is not detected early in the season, then the estimates will be biased low and will continue to be low until the segment data are replaced by later classifications. Thresholding procedures to eliminate the early-season biased data from the aggregation were implemented in Phase III in the United States and in the U.S.S.R.

TABLE II.—LACIE Biowindows and Corresponding Crop Development Biostages

<i>Biowindow</i>	<i>Biostage</i>
1. Crop establishment	1. Planting
2. Greening	2. Emergence
3. Heading	3. Jointing
	4. Heading
	5. Soft dough
4. Maturity	6. Ripening
	7. Harvest

The CAMS data base is updated as required to maintain the most current data available for aggregation. Data are selected from this data base for area estimates depending upon the type of aggregation desired and the priority of the data selection parameters. A data base used for aggregation may be saved for future evaluation and/or reaggregation if required by procedure modifications.

The yield data base contains a yield estimate and associated yield variance for each area stratum. The yield data base is updated as the yields are received from YES, generally on a monthly basis during the growing season. The data base can be updated interactively or in the batch mode using cards or tape. The allocation data base defines the aggregation hierarchy for a country. The area strata and yield strata that are combined to obtain production are identified in the allocation data base.

Aggregation

The LACIE used the same sampling and aggregation techniques through the three phases of operation with some modifications in the allocation parameters. The statistics calculations were modified to accommodate yield models covering areas larger than strata and to adequately consider mixed wheat (spring and winter) areas. The historical ratios for estimating areas with no segments available (Group III) were changed to use the zone-level history rather than the stratum if less than three segments were available for aggregation. The CAS procedures were in a development mode throughout the three phases

and were modified to accommodate new techniques in sampling and classification as the project progressed. Figures 1, 2, and 3 show the progression of LACIE through the three phases in relation to the reports generated.

During the course of CAS operations, statistical calculations were modified to accommodate changes in the system such as yield model boundaries, variances for substrata that are estimated with probability proportional to size (Group II), and mixed wheat areas. (See the presentations and supporting papers in the Experiment Design Session which address sampling and aggregation issues.) These algorithms were implemented in an off-line development system as a test prior to implementation on the CAS system. During the implementation period, aggregations were performed on the interactive system, and statistics were produced on the development system.

During Phase II, algorithms and procedure modifications occurred faster than they could be implemented into the CAS operational system. For this reason, operations were carried out on two aggregation systems with the extract software to manipulate data between them. This situation actually existed at the end of Phase III for thresholding and screening of the CAMS data base prior to aggregation, because these techniques were not incorporated into the interactive system.

In addition to thresholding CAMS data, a procedure called screening also was used in Phase III to identify segment wheat proportions that were statistical outliers. The test statistic was the ratio of the CAMS estimated wheat proportion to the historical proportion of wheat in that county. If this ratio fell outside the 3-standard-deviation limit calculated for its group, the segment was an outlier and was eliminated from the aggregation. This screening procedure was used in the United States in Phase III but was not applied to the U.S.S.R. because the lowest political subdivision for which historical data are available is the oblast (stratum), and the procedure was not applicable to the larger geographic area.

The yield was input at the stratum level and combined with the stratum area to obtain production. The production was then summed to obtain totals for higher levels in the hierarchy. The derived yields were obtained by dividing the production by the area at any desired level in the hierarchy.

The CAS analyst reviewed the CAMS segment data inputs and updated the CAMS data base as re-

quired to keep the data base current. A preliminary area estimate was generated and evaluated about 7 working days prior to a scheduled report. The yield estimates were reviewed by the YES and the CAS prior to updating the yield data base. The area, yield, and production were estimated and reviewed for accuracy and reasonableness. If data base errors existed, they were corrected and a final set of estimates was generated.

Time Line for Report Preparation

The CAS operations time line required 7 working days before a report date to update the CAMS data base and to prepare an area aggregation. A backup aggregation was prepared to submit as a monthly report to cover any computer failures that might occur during the critical period. One backup aggregation was submitted as a CAS report during Phase II because of computer failure. Under special circumstances during Phases II and III, the lead time was shortened to as little as 3 days prior to a report to accommodate segment processing and to include the latest data in an aggregation.

Analyst Interaction With Other Project Elements

The CAS analysts provided feedback to other elements of the AES concerning operational requirements for segment processing or problem areas identified during the report preparation. Sample segment results which produced aggregate area estimates that deviated from the expected values (based on historical data combined with current weather information) were referred to the CAMS for review.

Yields that did not follow expected trends were referred to the YES for verification. If an evaluation of the data resulted in a modification or deletion, the appropriate CAS data bases were updated.

The CAS analysts also met with the Crop Condition Assessment Team and the YES personnel to review their inputs to the CAS report. Information required to complete a report, such as operations processing and evaluation of segment data, were obtained from other AES elements. The completed report was delivered to the Commodity Control Office for reproduction and distribution.

CONCLUSIONS

The system utilized by the CAS during the LACIE evolved to accommodate the AES requirements and to fulfill the objectives of the experiment. The basic elements of the system remained constant, and these elements were necessary to make crop estimations of the LACIE type, regardless of the new environment in which the system was required to operate.

The CAS demonstrated that crop reports could be produced monthly in accordance with a schedule. Additional system improvements in the AES are required to decrease the report generation time to follow rapidly changing crop conditions that affect production.

The CAS software system was under configuration control and required as long as 6 months to make changes; as a result, the CAS output did not maintain pace with the technology. Since remote-sensing techniques are in a state of rapid development, systems should be designed to accept modifications more quickly to ensure that the output products are up to date.

REFERENCES

1. LACIE Crop Assessment Subsystem (CAS) Requirements. LACIE-C00200, Vol. IV, Rev. C, NASA JSC (Houston), Oct. 1977.
2. LACIE Crop Assessment Subsystem/Information Evaluation (CAS/IE) Interface Control Document. LACIE-00709, NASA JSC (Houston), Dec. 1975.
3. LACIE Commodity Data Control Plan. LACIE-C00609, Rev. B, NASA JSC (Houston), Nov. 1977.
4. LACIE Crop Assessment Subsystem (CAS) Requirements. LACIE-C0020, Vol. IV, Baseline issue, NASA JSC (Houston), Dec. 1974, and Rev. A, Dec. 1975.
5. Chhikara, R. S.; and Chang, J.: Large Area Crop Inventory Experiment Area Variance Estimate in the United States. LEC Tech. Memo. 642-1869, NASA JSC (Houston), Apr. 1978.
6. LACIE CAS Detailed Analysis Procedures. LACIE-00718, NASA JSC (Houston), June 1977.

LACIE Status and Tracking

V. M. Dauphin,^a C. H. Jeffress,^b and J. M. Everette^b

INTRODUCTION

The LACIE production processing system at the NASA Johnson Space Center (JSC) called for the flow of electronic and physical data products in a timely and efficient manner. The LACIE data bases at JSC included electronic imagery on the IBM 360-75 computer in the Data Systems and Analysis Directorate (DSAD, Building 30), LACIE crop assessment data on the Earth Observations Division (EOD) Programmed Data Processor (PDP) 11-45 support processor (Building 17), and physical products in the LACIE Physical Data Library (LPDL, Building 17). In addition, imagery tapes were obtained from the NASA Goddard Space Flight Center (GSFC); film products were provided by the JSC Production Film Converter (PFC), and crop calendars were acquired from the U.S. Department of Agriculture (USDA). The need for data and data products to be available at given stations simultaneously dictated that accurate data status be available. Further, there was a requirement to measure throughput rates and perform efficiency analysis. The data management method employed to meet these requirements is known as the Automated Status and Tracking System (ASATS).

The ASATS went through a number of evolutions as the LACIE program matured over four phases in the last 3.5 years. Data flows changed as a result of modifications to basic data analysis; man/machine interfaces became more complex with the advent of more machine terminal processing; and the ASATS data base grew in size from 4000 blocks (256 16-bit words per block) to more than 20 000 blocks of data.

This paper will discuss the operational requirements, the evolution of the status and tracking system in meeting these requirements, and a definition of the final ASATS developed to meet the

LACIE needs. Additionally, the lessons learned during this evolutionary process will be discussed.

LACIE STATUS AND TRACKING OPERATIONAL REQUIREMENTS

The LACIE operational elements were charged with the measurement and analysis of system throughput in both numbers of sample segment acquisitions and times required to process these acquisitions. Further, there were baseline goals against which progress was required to be measured at or between predetermined stations in the LACIE processing system. Operations also had to identify those conditions that deviated from specifications for both the numbers of throughput acquisitions and the processing times at predetermined status points and then flag problem areas and assist in implementing solutions.

In order to accomplish these objectives, a status and tracking system was required with inputs provided by various system elements. Evaluation of the level at which tracking should be accomplished and the information to be maintained dictated that a system be developed so that a large part of the testing, correlating, and reporting could be done in an automated manner.

The data bases were planned to contain several types of information. There was a basic accounting set that was unique to a sample segment or set of segments and subject to very limited change or none. The second data type contained production status parameters, which were subject to change as a function of data being moved in the production system. It was possible that up to 16 sets of production status parameters could exist for each accounting set. Considering the probability of both standard and variation reporting and changes in the stations being tracked due to changes in operating procedures, it was necessary to allow for dynamic manipulation of data base definitions.

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Inc., Systems and Services Division, Houston, Texas.

Additionally, capabilities were required, such as the comparison of elements and the use of the arithmetic tools of addition, subtraction, multiplication, division, and statistical calculation.

There was a need to allow for batch update and reporting, as well as the capability to perform interactive queries. Outputs were required in report formats as defined by the user at run time or as previously stored for batch operation. There was also the need to allow for preprogrammed queries that could be called by the user.

EVOLUTION OF THE LACIE DATA MANAGEMENT/STATUS AND TRACKING SYSTEM

October to December 1974

Discussions were initiated regarding the need for a status and tracking system dedicated to determining the progress of LACIE data as they flowed through the analysis process within the EOD. At this time, it was generally believed that the machine processing status to be provided by the 360-75 computing system might prove adequate for the purpose.

January and February 1975

Further investigations showed that because of data preparation, data quantity, and the analysis steps involved, the status and tracking of LACIE sample segments from the machine processing of GSFC imagery data tapes to the storage of data on the Earth Resources Interactive Processing System (ERIPS) imagery data bases and the PIC-generated products was not sufficient to determine status during the analysis process.

Based on the roughly 600 LACIE sample segments ordered for Phase I and the plan to analyze one acquisition per biowindow, a 2400-file data base with nine status stations (preselected points in the data flow where products availability was the key to the process continuing), some averaging 40 to 60 transactions a day, was too large to maintain by manual processes.

March 1975

A decision was made to implement an automated status and tracking system to support the LACIE

program by June 1, 1975. It was obvious that a fully operational status and tracking system to support the requirements that were at the time beginning to be identified was difficult—if not impossible—to develop, debug, and acceptance test in slightly less than the 4 months remaining prior to Phase I production. An alternative, proposed by the EOD support contractor, was to utilize the TRAC-8 series of status and tracking systems developed for the Data Reduction Complex (DRC) in the Institutional Data Systems Division (IDSD), on an interim basis, until a system that adequately met all the requirements of LACIE could be found and/or developed. It was at this point, March 14, 1975, that the LACIE Interim Status and Tracking System (ISATS) was born.

April to June 1975

Because of the limitations on flexibility within the TRAC-8 software, ISATS was implemented with three directories. One directory cross-referenced the Data Product Requests (DPR's) and the LACIE sample segment numbers and tracked the processing of a sample segment from receipt of data at JSC until the Crop Assessment Subsystem (CAS) procedure was completed.

The second directory tracked each data product request, whether batch, interactive, or update, until all electronic data processing products were received by EOD.

A third directory, which was later dropped from the requirements, tracked Discrepancy Reports until the discrepancy was cleared. (This function was assumed by the Facilities Configuration Control Office.)

Initially, inputs to the ISATS were planned to be made on four terminals. However, the only terminal put into use was located in the LPDL, the control point for all LACIE data products. ISATS went "on-line" for LACIE use on June 1, 1975.

June to October 1975

The inadequacies of ISATS became apparent very early in the operation. Part of the problem was the inflexibility of the TRAC-8 software with regard to minor changes in the LACIE data flow and to any new reports required. This was compounded by the fact that there was no input verification and no audit capability for the relatively new users of the system. The rest of the problem was physical access to the

computer via the one-demand terminal within EOD. At times during this period, the ISATS reports were running 3 to 4 days after the fact; and, while ISATS was building up and saving a valuable data base, it was not helpful on a real-time basis in accounting for LACIE sample segment data products as they flowed through the EOD analysis process. Before the end of July, it became apparent that ISATS was not a satisfactory interim system and another search had begun to find a system more compatible and responsive to the LACIE needs.

The support contractor proposed, as a result of a capability study, that the LACIE Phase II status and tracking be accomplished on a commercially leased data base management system known as "COMSHARE/COMPOSIT 77." The proposal, as finally accepted, called for the status and tracking of all acquisitions from roughly 1800 sample segments—the LACIE Phase II scope—from the time they were ordered from GSFC, through JSC processing, and until sample segment summaries were provided CAS. This activity used card input for overnight batch to update the data base and to generate reports by the next morning containing data less than 12 hours old. The requirements called for nine daily reports (seven statistical summaries and two tabular activity listings) plus a weekly and a monthly summary report. Software optimization was required to provide 21 status stations for each sample segment acquisition identified in the LACIE Phase II requirements.

Cost of operating COMSHARE was originally estimated at \$4000 per month. This estimate was based on projected data base size and terminal activity (connect, line, and central processing unit (CPU) costs).

Even though cost control was recognized as a problem (special query reports for operations were estimated at \$50 per report), the software flexibility and overnight update of each day's activity were desirable in meeting the LACIE operational requirements, and the use of the system was approved by management.

October 1975 to March 1976

During the first 2 weeks of October, the ISATS data base was verified and transferred to the COMSHARE/ASATS. Concurrently, optimization of the software to meet the requirements of LACIE Phase II began. ISATS and ASATS ran simultaneously for the last 2 weeks of October to ensure a correct data

base prior to full-scale operation on ASATS alone.

Additional requirements for LACIE Phase II were generated during the period from the cross reference of Phase I and Phase II data bases, extended reporting capabilities, and additional query capabilities to support LACIE performance analysis. These modifications were completed, documented, and acceptance tested by the middle of March, and the system was turned over to LACIE operations as of April 1, 1976.

April to September 1976

All of LACIE Phase II was supported on the COMSHARE/ASATS, but the software required considerable modification to meet the changes in requirements. The largest modification to the system was to support a change in the Classification and Mensuration Subsystem (CAMS) analysis process. This required the addition of one new card format, autopunch of five cards, two new reports (daily packet order lists and operations throughput summary), the automatic update of historical data bases, and a more comprehensive input verification subroutine.

None of the above changes came cheaply. In fact, the increased operational activity in addition to the software modification costs slowly escalated ASATS costs far beyond what had been anticipated from the information originally available. In July, upon receipt of the LACIE Phase II/III status and tracking requirements, it became obvious that a cheaper method had to be found to support this function.

Prior to the COMSHARE/ASATS original implementation, an EOD machine-loading study was done in the hope that the status and tracking function could be done in-house on its PDP 11-45 computer. Unfortunately, the resources were not available, but late June brought EOD the possibility of obtaining a second PDP 11-45, which would solve the resource problem. By August, EOD management had approved a feasibility study on the duplication of the COMSHARE/ASATS function on the second PDP 11-45.

October 1976 to January 1977

The LACIE Phase II/III requirements were acceptance tested on the COMSHARE/ASATS system and turned over to LACIE operations by October 15, 1976. During the same month, a proposal was made

to duplicate the function on the PDP 11-45 and move the entire operation in-house upon acceptance testing. The plan, as presented, would provide all the required features at a significant cost reduction via utilization of the newly acquired EOD computer (the second PDP 11-45). The plan was approved and work began immediately to modify the PDP-compatible data management software, the Regional Information Management System (RIMS), to provide an in-house RIMS/ASATS as a replacement for COMSHARE/ASATS.

The RIMS/ASATS replacement ran simultaneously with COMSHARE/ASATS for the last half of December to verify the new data management software and was acceptance tested in January 1977. At that time, formal release of the use of COMSHARE was made. Average operating costs dropped from \$20 000 per month to \$2150 per month and have not exceeded \$4000 per month since the conversion was made.

February 1977 to February 1978

The RIMS/ASATS has remained the basic status and tracking function during this period. Only three modifications have been made as a result of changed requirements. The first change was made in June 1977 to accommodate a new data flow that involved the Image-100 ("Procedure I") method of CAMS data analysis. The second change was to accommodate Transition Year requirements for a different batch input stream. The last modification was made in June 1977 to the in-house RIMS to add data base protection and to provide a better arithmetic operator in the comparison of fields.

As of this writing, no further modifications to what has become a very satisfactory Data Management/Status and Tracking System are anticipated.

DESIGN CONSIDERATIONS

From the initial analysis, it was determined that most of the data desired in the accounting data set was already available in card format from the Data Acquisition, Preprocessing, and Transmission Subsystem (DAPTS) Landsat data orders. This set of data became the DAPTS data base. There were additional LACIE operational requirements for this data base and an additional card input was devised to add to the DAPTS set. A record in this data base was generated each time the DAPTS ordered data from

GSFC by utilizing the order cards. Also, records were updated or deleted by utilizing the cards produced by DAPTS operations whenever a segment was changed or deleted.

Additional information pertaining to the segment was input manually. This information pertained to product availability in the analysis packet. The products tracked in this manner were considered critical to the analysis, and their absence caused data to be held or backlogged until these support products became available.

The data base that contained segment acquisition status and tracking information was generated by electronic imagery from an acquisition being received at JSC and entered into the ERIPS imagery data bases. Originally, an update card was prepared manually by entering the segment number and acquisition date into an update card. This data base was called the FLOCON data base after a CAMS subelement responsible for CAMS internal operational flow control.

Once segment acquisitions were generated in the FLOCON data base by inputting the segment number, acquisition date, and date received at JSC, there were several operational requirements imposed on ASATS.

First, the segment was checked against DAPTS to assure that the data received were valid and that they were within the biological window dates established for analysis. In the initial ISATS, these requirements were met by listing the two data sets and comparing them visually. In later systems, this became an automated operation. Also in the later systems, a special report was formatted for printing information on gummed labels. These labels were used by LPDL for film and analysis packet identification.

It was previously stated that inputs were made into the DAPTS data base regarding the availability of analysis products. Some of these analysis products were tracked on an acquisition basis. These were the initial film products and selected computer products. Once the film products were received in EOD, an update was made to the FLOCON data base containing the date of receipt and a comparison was made to information in the DAPTS data base to determine if the key ancillary products were available. If all the products were available, the segment acquisition was reported to CAMS as available for analysis. Once more, in ISATS this was done manually; it was automated in later systems.

The report listing segments available for processing became known as the CAMS Order Form and

was used by CAMS to request data packets from LPDL.

Once the packet was delivered from LPDL to CAMS, a card was input to the system denoting the date analysis began. At various stages within the analysis cycle, date cards were input to ASATS from different stations to update the status of the segment acquisition. When CAMS completed processing of an acquisition, input consisted of a card containing the date the results were sent to the CAS, a relative biostage indicator, a code relating to the process used, and an estimate of data quality. This essentially closed the tracking record for a specific acquisition.

Various reports were required by different LACIE elements. Because of changing procedures, these report requirements were periodically evaluated for content and frequency of need. After each evaluation, reports were often dropped, the reporting frequency changed, or two or more reports were integrated to make a single report.

The data bases were updated each night in the batch mode and the standard reports were run at this time. Run streams were set up for batch operations for daily, weekly, and monthly reports. The proper run stream was selected by LACIE operations at the close of business each day.

Procedures were developed so that LACIE operational elements provided operations with input cards each afternoon and received their output reports at the next morning's operations coordination meeting.

Special queries were performed by the Operations Section on request from LACIE management. Some queries were "canned" (i.e., preprogrammed and maintained on the disk file), with arguments for initialization and execution being entered interactively just prior to batch run time.

RIMS/ASATS IMPLEMENTATION

The ASATS design is based on achieving maximum use of RIMS to perform ASATS functions. The design provides a standard batch update capability, standard reports (both periodic and aperiodic), an ad hoc report capability, and an ad hoc update capability. All data base transactions reflecting LACIE activity are entered as standard batch updates. Ad hoc updates are used normally for correcting such problems as when cards are erroneously entered into the system. About 20 standard reports currently exist for the system (e.g., packet order report, biowindow open/close report). Ad hoc reports are requested frequently to meet special needs; they

usually originate as one-time reports but sometimes become standard system reports.

The ASATS software is composed of special processors built for ASATS to facilitate the auditing of ASATS data base updates, RIMS commands augmented to support specific ASATS requirements, command files (sequences of RIMS commands) that will cause the generation of specified reports, data base definitions describing the ASATS data base to RIMS, and format descriptions describing input files and report formats.

All reports and ad hoc updates are made using the Data Manipulation Language (DML) of the RIMS data base management system. Standard batch updates were implemented using a data preprocessor program and a special data base update program (via FORTRAN interface with RIMS).

The RIMS, which may be used either interactively or in batch, provides both a Data Definition Language (DDL) and DML. DDL provides for defining the data base structure, input formats, and output formats. DML provides commands which support data base update and data base queries.

Data base record formats are defined in terms of field names, field lengths, data type, and whether or not the field is a key. The input format is defined in terms of field name, field start location on input record, field length, and an input verification type for a special update processor. The output format is defined in terms of field name, field start location on output record, and field length.

The data base update capabilities include adding a new record or modifying existing records from records on an external file or changing the contents of a field or fields of a record or records within the data base.

Reporting capabilities include output of field values, specified text, and arithmetic computations for records in the data base. Formats of output may be determined by either a predefined format or by the syntax of the command. The ability to specify text allows for the annotation of reports. Arithmetic and statistical computations include arithmetic expressions involving field values and/or numerical literal values, standard deviations, mean values, maximum, minimum, summation, and count of occurrences for a group of records.

Groups of records for update or reporting may be selected by explicit identification of records, key field value, range of key field values, arithmetic relationships between fields of a record or hierarchical-related records, logical relationships between

records, and/or hierarchical relationship between records.

The RIMS provides for device independence by a command that allows for reassigning system files, including the command file, the data input file, the message file, and the report file. This feature allows flexibility in how the system is used.

Standard reports are implemented as a file of RIMS commands. These files, when assigned as the RIMS command file, produce the designed report. Execution of individual reports during the nightly batch run is caused by entering the file name to be executed in a particular file.

Ad hoc reports are either generated from command files or produced interactively from the terminal. Ad hoc updates are generally performed interactively from the terminal.

The standard batch update program provides audit reports, tape labels, and punch cards that are reentered into the system in addition to updating the data base. The preprocessor program sorts cards in a particular order for the update processor, rejects all cards that are exact duplicates, rejects invalid card types, and separates the sorted card images into separate files by LACIE phase. It also provides audit reports for invalid card types, cards as input, and cards as sorted.

LESSONS LEARNED

Each LACIE subsystem had the responsibility for definition of its own input, data flow, software, and output requirements. Only after these requirements were documented and approved was the need for extensive status and tracking realized. In future systems, more emphasis should be placed on early analysis of preliminary subsystem requirements for the definition of data flow and status points. Further, a data management system should be selected very early in the effort in order to help ensure that necessary modifications can be accomplished before the targeted start-up time.

By the time the LACIE status and tracking requirements were identified, the acquisition and throughput of data were imminent. This led to the selection of an interim system with practically no data management capability. The next step taken was to a commercial system which gave limited, but increased, flexibility at great expense. Only after a long experience and development effort was a system realized that provided the required flexibility as well as a reasonable operational cost.

LACIE Quality Assurance

G. L. Gutschewski^a

INTRODUCTION

This paper describes and explains the LACIE Quality Assurance (QA) Program. It addresses the beginnings of QA, the objective and concept of LACIE QA, its responsibilities, and its accomplishments.

What did QA do for LACIE? What were the methods and rationale of the QA group? This paper will provide answers to these questions but will not delve into detail on how the QA tasks were performed. The reader is referred to the LACIE Quality Assurance Program Plan (ref. 1, section 5.0) and the LACIE Quality Assurance Procedures for more detail on specific tasks.

The LACIE Quality Assurance Program Plan delineates the QA system to include all of the organizational elements within LACIE (listed in ref. 1, section 2.0) and has the goal of assisting all of these organizations in attaining the highest level of performance possible. This paper will describe the extent to which this was done.

ORIGIN OF LACIE QUALITY ASSURANCE

At the beginning of LACIE, there were references in such documents as the LACIE Project Plan (ref. 2) on the need for quality control, but no definitive statements were made on the direction or methodology of this function. This situation prompted two Review Item Dispositions (RID's) in December of 1974. One of the RID's defined the need for a QA plan, and the other defined the need for a data quality plan. These two RID's were major factors in the decision to establish a LACIE QA program.

When the decision was made to have a formal LACIE QA program, it became necessary to develop a concept as to what type of QA program would work

best in LACIE. A LACIE QA group was established to develop a LACIE Quality Assurance Program Plan encompassing all of the LACIE organizations and to assure that this plan was implemented. The size of this QA group (3.5 man-year equivalents) and the number and diversification of the LACIE organizations necessitated that each organization perform its own QA tasks with the QA group checking periodically to assure that the tasks were being performed satisfactorily. Keeping these factors in mind, the next step was the actual development of the LACIE QA program.

DEVELOPMENT OF THE LACIE QUALITY ASSURANCE PROGRAM

In the development of the LACIE QA program, the first tasks undertaken by the LACIE QA group were (1) the preparation of a QA program plan, (2) the definition of QA checkpoints which each organization should include in its operations, (3) the formulation of operational procedures by each of the LACIE organizations, and (4) the compilation of a document entitled "LACIE Quality Assurance Procedures," which included all the QA checkpoints for each of the LACIE organizations. All these tasks were interrelated and were necessary for the development of a useful QA function which could help the LACIE organizations attain the highest level of performance possible. Following are comments on the four tasks.

Preparation of the Plan

The LACIE Quality Assurance Program Plan (ref. 1) was required to define the QA functions. Whenever a function was defined and agreed upon informally by management, the implementation of that function was begun immediately. For example, very early in LACIE, it was known that procedures and products would be audited regularly; so a set of

^aNASA Johnson Space Center, Houston, Texas.

procedures for auditing was prepared as quickly as practical. Therefore, when a given set of operational procedures was available, QA immediately began auditing those procedures. Thus, by the time a formal plan was approved, a large portion of the QA effort was already implemented, and the effects of good quality control had been present for a major portion of the project.

Definition of Checkpoints

When the QA effort started, quality control in the project was almost nonexistent; the quality control that did exist was not formalized or consistent. The definition of the LACIE QA checkpoints within each organization's operations was a necessary first step in establishing formal quality control in the project. This was a cooperative effort between the QA organization and the other LACIE elements.

Formulation of Operational Procedures

While the QA plan was being developed and the QA checkpoints were being established, the concurrent effort of writing formal procedures for each of the LACIE organizations was begun by the organizations. There was reluctance by some organizations to write procedures because they did not understand the value of procedures in their operations and because this task impacted their resources. This delayed some of the procedures as long as a year into the project. However, most of the organizations were cooperative and proceeded to write their procedures as quickly as they could, considering their operational constraints. This task did not end with the first writing but continued throughout LACIE as new techniques, hardware, and software were developed and incorporated into the experiment.

Compilation of Procedures Document

As the QA checkpoints for each organization were established, the QA organization compiled them into a general document which included the QA procedures for every LACIE organization functioning at the time of publication. These procedures were audited, and the results of that audit were included in the document, making it a representative document at the time it was compiled.

The compilation of a QA procedures document was a necessary first step in establishing a system of quality control for LACIE. This document established guidelines for writing the operational procedures and for checking the output of the various organizations and provided the mental discipline whereby the LACIE organization could achieve a high level of performance. The document was eventually absorbed into the operational procedures as they were written.

While this document was being compiled, the LACIE Quality Assurance Program was being implemented. Therefore, when the LACIE Quality Assurance Procedures document was completed in July of 1975, the major portions of the QA program were close to full implementation. Some areas of responsibility were not fully implemented as quickly as others; but, basically, from August of 1975 until December of 1977, most of the QA activities were being implemented.

LACIE QUALITY ASSURANCE RESPONSIBILITIES

This section will cover the QA responsibilities of all of the organizational elements of LACIE, including the LACIE QA group.

Audits

The QA group conducted audits of both procedures and products. The audits on procedures were conducted to ensure that the procedures were being followed or that, if necessary, the procedure was updated in a timely manner. The audits on products were conducted to ensure adherence to specifications. More simply, QA checked to see if the user was satisfied with the product. If the user was not satisfied, either a work-around technique was devised, the product was upgraded, or both. Also, the question, "Is this product really necessary?" was asked.

Following are several examples of problem areas encountered during QA auditing.

1. Frequently, there was a slow turnaround on updating procedures. This was caused by a lack of resources and a formal procedure for updating procedures. Repeated attempts to get a formal procedure for updating procedures approved failed.

2. At the beginning of the experiment and at

various other times during the experiment, a lack of procedures caused unnecessary operational problems. These unnecessary problems were often caused by the personnel not having a clear definition of their duties, and the solution to the problems was a good set of procedures.

3. In some situations, there was an absence of a required product or the product did not meet specifications. This lack of products or inferiority of a product caused some groups to find alternate solutions, thus wasting precious resources.

4. In the early days of LACIE, at one time or another, some of the organizational elements were in doubt as to the products they should be providing or receiving or where to deliver or receive their products.

As a solution to problems (3) and (4), formal product lists and their regular verification were established.

The answer to these and similar problems is a good quality control system that includes definitive procedures, QA checkpoints in the procedures, product checklists, and audits of the procedures and products. The quality control of the procedures and products was considered the primary duty of QA, and this is the area where the most benefit was provided to LACIE. QA accomplished its objectives in this area.

Discrepancy Reporting

The QA group was responsible for maintaining and coordinating a LACIE Discrepancy Reporting System. This system was a means of monitoring the problems in LACIE; of determining problem areas that needed extra attention; and of statusing, reporting, categorizing, and documenting the various problems in LACIE. The LACIE Discrepancy Reporting System was not all-inclusive of LACIE problems, since some problems were never written on Discrepancy Reports (DR's) but were handled via memorandum or personal contact. However, a high percentage of operational problems have been documented by the Discrepancy Reporting System.

Examples of major problem areas in discrepancy reporting are as follows.

1. Division of authority. The many political boundaries in the LACIE system proved a tremendous obstacle to the smooth and efficient operation of the LACIE Discrepancy Reporting System. Constant interaction by QA with the various groups finally resulted in a relatively efficient system which pro-

vided the representative status of LACIE.

2. Untimely response to DR's. Some of the organizations in LACIE did not respond in a timely manner to DR's, even though they might have solved the problem in question. The degree of seriousness of this problem varied between organizations. The QA group at times had to check the internal records of some organizations and urge them to respond to the DR's. This slow response to DR's did not help to impress on personnel the usefulness of the Discrepancy Reporting System.

3. Reluctance to write DR's. At the beginning and at the end of LACIE, some of the organizations were reluctant to write DR's, and QA constantly had to urge personnel to perform this task.

Test Certification

The LACIE Quality Assurance Program Plan defines the acceptance testing function as being applicable to any of the LACIE organizations. However, the test certification effort of the LACIE QA group was concentrated in the Data Techniques Laboratory (DTL) of the Earth Observations Division (EOD) in Building 17 of the NASA Johnson Space Center (JSC). However, the LACIE QA group was available to investigate other support areas if requested by management.

There were two major reasons why the LACIE QA group usually participated primarily in the EOD acceptance testing:

1. Resources. The available manpower in the LACIE QA effort and the time necessary to run an acceptance test prohibited the QA group from monitoring the acceptance tests of organizations outside EOD. For example, hundreds of acceptance tests were run in EOD; some took a few hours, and others took several days or more. Generally, support organizations external to EOD would test their software and hardware systems prior to releasing them to EOD.

2. External organizations. Many organizations outside EOD were performing acceptance testing but had their own QA monitors. These included the JSC Ground Data Systems Division (GDSD) and the Goddard Space Flight Center (GSFC).

Three problems related to test certification existed at the beginning of LACIE QA: (1) inadequate documentation of test plans and procedures, (2) lack of internal testing of software before test scheduling, and (3) lack of adherence to test plans and pro-

cedures. These problems were quickly solved.

Acceptance tests were required to assure that a new hardware, software, or procedural system met specifications and/or performed as represented by its manufacturer or originator. To install a new system or add to an existing working system without testing the new component is to invite total system failure or an immense load of useless data. Thus, to verify that the new system or component was ready to be part of the operation, the LACIE organizations were required to write a test plan and a test procedure, to review and approve them formally, and then to adhere to their own test specifications. This procedure assured that a reliable test was performed and verified that the new component was ready to be part of the operational system. Unfortunately, this procedure was usually opposed; however, to avoid needless operational problems and sometimes total shutdown of an entire operational system, the QA group had to rigidly enforce test plans and procedures.

Procedures Reviews

Whenever a new or revised procedure was issued, LACIE QA would review it for its adequacy as a procedure. In addition, the adequacy of the QA checkpoints contained in the procedures would be reviewed. Initially, many of the groups writing their procedures did not understand that they should write how to perform each task. As a result, some of the procedures were inadequate. As these groups came to understand the purpose of the procedures, this problem disappeared.

Other Tasks

In addition to the responsibilities just described, the LACIE QA function included defining QA requirements, giving status reports of QA activities, providing audit reports, assisting the LACIE organizational elements on QA policies and procedures, identifying necessary configuration changes, and making recommendations to management.

The QA responsibilities of the LACIE organizational elements exclusive of the LACIE QA group can be stated briefly as follows and are included in the LACIE Quality Assurance Program Plan.

1. Procedures preparation. Each organizational element was responsible for writing and updating its

operational procedures, including the QA in-process checkpoints.

2. Discrepancy reporting. Each organizational element was responsible for originating and replying to DR's relative to its area of interest.

3. Test controls. Each organizational element was responsible for testing software, hardware, data flow, and techniques within their respective areas of responsibility. Test plans and procedures were subject to review. The actual acceptance test was subject to monitoring by QA.

INTERNAL QUALITY ASSURANCE SUPPORT

Four of the organizational elements within LACIE—the DTL, the GDSD, the Mission Planning and Analysis Division (MPAD), and GSFC—retained their own formal QA group.

Data Techniques Laboratory

The DTL of the EOD had a one-man QA effort which consisted of monitoring the internal DTL Discrepancy Reporting System, approving the closure of a DR, and approving acceptance tests before they could be closed.

Ground Data Systems Division

The GDSD supported LACIE through the LACIE/Earth Resources Interactive Processing System (LACIE/ERIPS). This group had several aspects of QA performed internally.

1. The IBM Corporation provided some system design, wrote the software, tested the software before it was made operational, and responded to DR's on the software.

2. The GDSD QA personnel monitored the acceptance tests and the internal Discrepancy Reporting System in JSC Building 30.

3. The MPAD provided internal QA support (more fully described in the next section).

Mission Planning and Analysis Division

The MPAD provided QA support to LACIE in two major modes.

1. Primary mode. In support of GDSD, the

MPAD provided LACIE with independent testing of the LACIE/ERIPS software/hardware system. The MPAD goal was to assure that the system met the requirements, that the quality of the system output products was consistent with the quality of the input data, and that the system performance remained stable.

2. Secondary mode. In support of EOD, the MPAD provided independent technical evaluations in problematic technical areas and in the technical performance of operational output.

Detailed information on the MPAD activities is available (ref. 3), but the following is a brief delineation of the tasks performed by MPAD.

Input imagery evaluation

1. Imagery screening with film. The film coming out of the LACIE/ERIPS system was screened daily for both LACIE/ERIPS problems and GSFC problems.

2. Imagery screening on the LACIE/ERIPS. The imagery was screened on the LACIE/ERIPS primarily for GSFC problems but also to compare LACIE/ERIPS film with the original imagery. This task gradually phased out but was replaced by such tasks as the development of an automated cloud screening capability.

3. Imagery registration evaluation with film. The object of this task was to assess visually, using film, the accuracy of registration between sample segments for a given site. This task continued throughout LACIE.

4. Imagery registration validation with the Sequential Similarity Detection Algorithm (SSDA). A computer algorithm called SSDA was used to evaluate selected segments to determine the magnitude of misregistration and to detect subtle registration errors. Written reports were provided on these segments.

Fields definition evaluation

1. Field definition screening with film. In this task, the production film converter (PFC) product 12 (field boundary overlay) was used to check the field definitions for such errors as overlapping fields and misplaced vertices. PFC product 12 was a computer plot on film of the analyst's fields in a classification. This task was phased out when LACIE Procedure 1 became operational.

2. Field definition screening on the LACIE/ERIPS. A more detailed analysis of fields definitions was performed on those selected segments being processed under the Classification and Mensuration Subsystem (CAMS) product evaluation procedures

where anomalies were observed. Using the LACIE/ERIPS for this task provided statistical and measurement capabilities not available when PFC product 12 was used alone.

Software confidence testing.—This testing was conducted on a regular basis and consisted of a representative sequence of production-oriented operations using known input data. This was done to determine that the same input data processed in the same manner will produce the same results on each of the LACIE/ERIPS software systems (versions). It established the consistency, reliability, and accuracy of the relative systems.

CAMS product evaluation.—As an independent check of the CAMS analysts, selected segments were processed on the LACIE/ERIPS using the field definitions of the CAMS analysts. Any errors, such as overlapping fields, were delineated in a detailed report provided to the EOD. This was an independent check on the analysts and the system. This task was gradually phased out.

Problem isolation and error analysis.—This task consists of the attention to special problems and the attendant efforts at possible solutions. Special studies may be placed in this category, as they usually were performed with a specific problem in mind, on request and sometimes at the initiative of MPAD.

Quality assurance data base.—The MPAD kept a computer record of its findings on the data studied and could recall portions of this information from its computer.

Goddard Space Flight Center

To complete the overall view of the QA being performed in LACIE, the QA tasks being performed at GSFC are listed under their general headings—inspection of LACIE sample segments from imagery generated on the color film recorder or black and white film recorder; inspection and analysis of output from the General Purpose Image Preprocessor (GPIP) line printer and teletype; data retrieval from the GPIP line printer; and reporting the LACIE QA assessment of these to the Special Projects Group in the production control section for the Landsat project.

The actual number of QA data inspections is so great that it is impractical to include them in this document. However, the following are given as examples to show the thoroughness of QA at GSFC: Landsat identification, correlation checks, film flag

checks, cloud pixel checks, edge threshold and edge density checks, alinement checks, Sun elevation and Sun azimuth checks, correlation parameters report, image data geometry, and pixel dropouts. The reader is referred to the GSFC detailed quality assurance procedures (ref. 4) for more detail.

ACCOMPLISHMENTS

The foregoing sections demonstrate the breadth and thoroughness of LACIE QA. Some of the results or benefits of the LACIE QA program are discussed in the following paragraphs.

Quality Control

As the QA procedures came into effect, the number of DR's diminished; the reprocessing of computer tasks diminished; the technical errors decreased; and, generally, the overall efficiency of the organizational elements increased. The reliability of the LACIE product increased, and the products became more measurable and more consistent as a result of quality control.

Procedures

Probably the single most important result that the LACIE QA group accomplished was to pressure all the organizational elements to write procedures and keep those procedures updated. In addition to establishing and maintaining quality control, the documented procedures became a tremendous source of information about LACIE. The documentation of technical and operational changes, technical or operational mistakes, and historical information is an invaluable aid in writing LACIE symposium or follow-on papers and as a reference for future planners.

The procedures also saved time and resources. For example, in the beginning stages of LACIE, some individuals were spending much time and resources trying to determine how to perform their tasks or even what their task was. An evidence of this situation was the fact that, at the beginning of LACIE, the DR's pertaining to procedural errors comprised more than 50 percent of the total number of DR's. In contrast, when all the organizations had written pro-

cedures at the end of Phase III, the number of procedural DR's usually was less than 3 percent (weekly basis). As a matter of fact, the number of procedural errors in a given organizational area would drop drastically when good operational procedures were provided to those performing the tasks.

Product Definition

For each of the three phases of LACIE and during the Transition Year, each functional element was required to review its requirements and input to LACIE QA a list of all output products and the products required to complete the assigned tasks. QA would then compile a complete list for LACIE and verify the list through an audit. If a deficient product was found, a solution was agreed upon by the receiving and the providing organizations. If an excess product was found, it was eliminated. Compiling this product list was a very difficult task in Phase I, but in the subsequent phases it became much easier.

What was the value of this exercise? First, the compilation of these product lists on a regular basis forced the LACIE organizations to review their requirements on a regular basis; in fact, this was the only review of requirements performed in a formal, systematic method. Secondly, the elimination of excess products and the attention to problem areas resulted in much more efficient LACIE operations. And, lastly, the actual operational requirements of LACIE are documented as products in the product lists.

Discrepancy Reporting System

Any system, program, or project needs a method of documenting, statusing and tracking, and monitoring its problems and directing the solutions to these problems. The method used in LACIE is called the Discrepancy Reporting System.

The LACIE Discrepancy Reporting System documented the reported problems, providing a summary reference. Documenting the problems helped to avoid repeating the same mistakes. The statusing and tracking of discrepancies helped to avoid system shutdown by pointing out problem areas that needed immediate attention. Monitoring and directing the solutions to the various problems assured LACIE of adequate solutions to the problems, thereby improving the system's performance.

Product Definition

For each of the three phases of LACIE and during the Transition Year, each functional element was required to review its requirements and input to LACIE QA a list of all output products and the products required to complete the assigned tasks. QA would then compile a complete list for LACIE and verify the list through an audit. If a deficient product was found, a solution was agreed upon by the receiving and the providing organizations. If an excess product was found, it was eliminated. Compiling this product list was a very difficult task in Phase I, but in the subsequent phases it became much easier.

What was the value of this exercise? First, the compilation of these product lists on a regular basis forced the LACIE organizations to review their requirements on a regular basis; in fact, this was the only review of requirements performed in a formal, systematic method. Secondly, the elimination of excess products and the attention to problem areas resulted in much more efficient LACIE operations. And, lastly, the actual operational requirements of LACIE are documented as products in the product lists.

Discrepancy Reporting System

Any system, program, or project needs a method of documenting, statusing and tracking, and monitoring its problems and directing the solutions to these problems. The method used in LACIE is called the Discrepancy Reporting System.

The LACIE Discrepancy Reporting System documented the reported problems, providing a summary reference. Documenting the problems helped to avoid repeating the same mistakes. The statusing and tracking of discrepancies helped to avoid system shutdown by pointing out problem areas that needed immediate attention. Monitoring and directing the solutions to the various problems assured LACIE of adequate solutions to the problems, thereby improving the system's performance.

Acceptance Testing

An acceptance test of a new system or component prior to operation is necessary to prevent the absolute operational shutdown caused by the new system or to prevent the operational disturbances

caused by a defective component. This new system or component could be software, hardware, or even procedures in certain cases. With a reasonable amount of testing, most of the major impacts of new systems or components can be avoided. This is what the LACIE QA acceptance testing accomplished.

CONCLUSION

The foregoing has stated simply what the LACIE QA program did and why, so that both the critics and the defenders of QA can appreciate the magnitude of the task. This paper does not delve deeply into the details of QA. Rather, it was designed to give the reader a better understanding of the whys and wherefores of QA and the contributions of the LACIE QA effort.

REFERENCES

1. LACIE Quality Assurance Program Plan. LACIE-C00626, NASA Johnson Space Center, Aug. 1977.
2. LACIE Project Plan. LACIE-C00605, NASA Johnson Space Center, Aug. 1975.
3. MPAD LACIE Product Quality Assessment Software and Procedures — Program Description and Computer Operations. Rept. T-00722, NASA Johnson Space Center, Nov. 1977.
4. DAPTS-GSFC Detailed Quality Assurance Inspection Procedures. Part I, LACIE-00721, NASA Johnson Space Center, Nov. 1977.

00117

Operations Reporting

R. G. Musgrove^a and Dale R. Marquis^b

The need for operations reporting had become obvious to the personnel involved in establishing, coordinating, and monitoring data flow in the infancy of the LACIE operations in the fall of 1974. Even though all subsystem elements were sincerely interested in and concerned with ensuring the success of LACIE operations, an effective coordination and integration function was required to provide the cohesion necessary for a smooth operational system. In fact, the operations coordination, integration, and management function required effective operations reporting for it to succeed.

Anytime a complex data flow is constructed, individual components will break down, bottlenecks will occur, and backlogs will build. It quickly became apparent during LACIE Phase I that the very simple accounting originally envisioned was not providing sufficient information regarding the status of the data system. While accountability was kept in terms of raw numbers for data received from the NASA Goddard Space Flight Center (GSFC) or for data in work, nothing was really known about the constituency of these numbers. For example, if equipment broke down and a bottleneck occurred, it was of course known. What was not known, however, was whether the data involved were U.S. segments, U.S.S.R. segments, spring or winter wheat data, etc. The inability to provide appropriate information or relevant statistics, let alone a coherent status, prompted the need for a comprehensive status, tracking, and reporting system. Thus, operations reporting began with an effort to satisfy a need for improved communication among personnel not only at the working level but at the project management level as well.

One of the first goals of operations reporting was to ensure that management understood the basic operational data flow and its attendant data handling

operations and, just as importantly, the constraints on the operational system. LACIE management also had to be apprised of accomplishments, problem areas, etc. This requirement influenced the evolution of operations reporting as strongly as the need for basic operations information at the working level.

As these various requirements became understood, specific data and information requirements were in turn placed on the appropriate parts of the operational system. These requirements were the driving force for the development of the Automated Status and Tracking System, which is the subject of a separate paper.

The development of the information requirements began by identifying the critical stages in the operational data flow. Then the parameters for reporting the progress of LACIE operations and the operations status were identified. Specific information requirements for input into the operations reporting system were levied on the operational elements. The operational elements responded to these requirements by supplying the inputs derived from manual systems or, in some instances, from automatic systems devised by the operational groups themselves.

The preceding papers show how involved the flow of data through the LACIE system is. The process starts at the NASA Johnson Space Center (JSC) with the ordering of data to be collected by the satellite. Once the data is received at GSFC from Landsat, it must undergo a number of screenings before it is shipped to JSC.

One of the first reporting and statusing systems instituted in LACIE was at GSFC. Originally, the hope had been to have a "full-up" status and tracking system for data in the GSFC system much like the one established at JSC. For a number of reasons (primarily resource), an all-encompassing system that could be interrogated was never implemented. Fortunately, however, it was possible to prioritize the reporting needs from GSFC so that information about the key areas was available. For example, statistics

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Inc., Systems and Services Division, Houston, Texas.

have been meticulously kept since the first data order in Phase I on the number of acquisitions (spacecraft hits), rejections for excessive cloud cover, correlation failures (as discussed previously), and quality rejects. By the end of Phase I, a definite pattern had emerged indicating that about 50 percent of the acquisitions would be rejected because of clouds, 10 percent because of correlation failures, and about 5 percent because of miscellaneous quality reasons. Naturally, these statistics—particularly those for cloud cover—would vary by season and by country; on the whole, they provided a reasonable yardstick by which to gauge performance. If significant short-term deviations were experienced, queries would be made to GSFC asking that they assess their operations to determine whether problem areas existed.

In a preceding paper, an account has been given of both the evolution and the design of the status and tracking system. The evolutionary process was arduous, but by the time spring windows had opened for Phase II, a viable system for data statusing and tracking had been developed. In what follows, the system will be addressed with more or less its current configuration and capabilities and its contribution to the management of daily operational activities.

To those charged with managing the operations data flow, there were several basic pieces of information desired from the system.

1. What is at JSC (by country and crop type)?
2. Is it in work? If so, where?
3. How long is each phase of processing?
4. What segments have been in work excessively long?
5. How old is the data supporting a production estimate?

These questions comprised the essence of the requirements for the status and tracking system. They comprised the basic information required to manage the system as well as that needed to status higher management levels concerning the health of the system.

Potentially one of the most valuable reports generated was one that showed the length of time each segment was in work at each processing station. For example, how long was it from the date of acquisition until GSFC shipped the data to JSC? How long did it take to make the film products and prepare the analyst packet ready for work? How long did the analyst have the segment in work? These were compared against a set of predetermined nominal processing times. Segments that had been in work at a

given station longer than the reference time allocated were automatically printed out.

From its inception, the "Delinquency Report," as it was called, proved to be a valuable management tool. In Phase III, it was common to have 500 segments in analysis alone plus another 500 to 600 coming in weekly. In all, 2000 to 3000 acquisitions might be on the move through the data flow at any given time. With this much active data in the system, it was easy for some of it to get sidetracked. The Delinquency Report caught these and flagged them by station and segment number. After a few iterations of these reports where the individual responsible for a given station had to prepare a response to the Delinquency Report, there was a noticeable tightening of the data flow and a significant reduction in the size of the report.

As a footnote to this activity, however, it was found that the Delinquency Report was a useful tool in managing the data flow only as long as the system resources were sufficient to meet the processing requirements. When the system became overloaded, backlogs began building, the system became saturated with data, and the report was of little value as data was often set aside deliberately to allow for processing of higher priority segments. Thus, it was often possible for large blocks of data to be "delinquent."

The ability to sort data by country, crop type, and acquisition date proved the greatest asset of the status and tracking system. Inevitably, there would be breakdowns in the systems that processed and/or manipulated the data. This was compounded by the fact that, during Phase III, the incoming flow rate of data exceeded the analysis capability. The data was prioritized by country and fed into the status and tracking system. In areas such as the U.S.S.R., where the flow rate exceeded analysis resources, determinations could be made on how to prioritize the data to provide the best possible data set for analysis and aggregation.

As the growing season would progress, there would be some segments that the Crop Assessment Subsystem (CAS) believed had satisfactory estimates which could be carried forth from one report to another and other segments for which they believed either new estimates should be generated or old estimates improved. It was possible to compare the CAS needs with what was available in the status and tracking system and to flag for the analysts those segments that were to be processed on a priority basis.

As understanding of the flexibility of the status and tracking system grew, many other reports, updates, etc., were generated, often for specific users or purposes. As a way of integrating the various user inputs from other status, tracking, and problems, an Operations Coordination Center (OCC) was established. The purpose of the OCC was to provide a focal point for the exchanged information regarding the status, problems, etc., of each of the individual components that comprised the data flow. Here much of the detailed data that came out of status and tracking was summarized into concise displays to provide, at a glance, data flow tracking of each major component in the system.

As a vehicle for obtaining the necessary information with which to manage the data flow, a debriefing was held each morning in the OCC to discuss the previous day's activities. Information exchanged during these sessions consisted of throughput statistics, problem areas, and suggested solutions to problems. Personnel from each major functional area of the project attended these sessions. Where appropriate, action items were given, and tracking of these was instituted to assure their completion.

The primary processing status displays in OCC integrated the many diverse inputs and reported the status and progress of the processing activities from receipt of the data from GSFC through completion of the Classification and Mensuration Subsystem (CAMS) analysis of the data. Specific items reported were receipt of data from GSFC and update of the JSC data base, arrival of the film products at the LACIE Physical Data Library (LPDL), availability of the data to CAMS for analysis, CAMS interpretation of the data and preparation of computer runs, arrival of computer processing data products, and evaluation of the computer run and delivery of estimates to CAS. The quantity of data that could not be satisfactorily classified and the reason were also reported. In addition to providing these daily operations reports, the OCC summarized and assimilated them into the weekly production reports that were provided to LACIE management. Presided over by an operations manager, the daily meetings were the key to the operation of LACIE in that they provided a direct exchange of information among the working level personnel directly responsible for the management and operation of the data flow. Problems could be quickly identified, tracked, and brought to project management's attention if necessary. Often they could be worked by direct assignment. Further, they provided a mechanism for keeping all the functional

areas apprised of the latest operational developments.

Other weekly reports were generated from daily inputs or obtained from the status and tracking system. These included reports on the throughput time required for processing data by the key elements of the operating system. These reports were soon limited to the GSFC and CAMS processing because the other subsystems exhibited a reasonably constant time within that nominally expected. This throughput time report enabled operations management to monitor the processing time line, identify bottlenecks, and initiate corrective measures as required.

In addition to compiling summary statistics based on the status and tracking system, the OCC also tracked Discrepancy Reports (DR's). As discussed earlier in a paper on quality assurance (QA), these DR's represented documented summaries of major failures in either hardware, software, or procedures and were categorized as critical or noncritical. Critical DR's meant that the potential existed for a major failure in the LACIE system, possibly resulting in a work stoppage. These were rigorously tracked and statused until satisfactory closeout had been accomplished.

Another major source of information coming from QA was the procedure audit. Although this type of QA departed significantly from what is thought of in the standard context of product QA, it made a substantial contribution to the overall management of the system. Because of the complexity of the LACIE operating system, it was imperative that standard operating procedures be developed for each of the major functional areas. Generally, these procedures were built around an operating system so that adherence to them was necessary for proper operation of the system. The QA audits of personnel compliance to procedures often proved enlightening; pressures generated by these audits eventually resulted in a much higher quality of documentation than would have been the case otherwise.

Other special reports were generated internally to operations management for its own use. Many of these reports were those generated at the beginning or near the end of a processing cycle for a specific set of data. They typically detailed the status of data acquisitions and processing such that specific operational "miniplans" were developed to process the data. This assured an orderly beginning and ending to processing data sets.

The goals of providing LACIE management with

the appropriate level of information and ensuring the understanding of the data flow and the operation's accomplishments on a weekly basis were realized because of an effective operations reporting system. It enabled the progress of data processing to be measured against the established product goals and objectives. It also provided management with a view of the problems and pitfalls associated with the operations system, permitting it to provide additional direction to the operations system.

The status and tracking system, dialog from the daily OCC debriefings, and QA reporting armed those responsible for managing the data flow with a significant amount of information. The problem then existed of how to distill this information into a weekly summary briefing that would keep management adequately informed of progress, problems, and overall adherence to production reporting schedules.

Early attempts to brief management by letting the numbers speak for themselves produced elaborate matrices showing what had been received, what was in work, average time in work, etc. To management's credit, it withstood this barrage of statistics with a measure of restraint. Over a period of time, it became apparent that while this procedure was historically

documenting what had occurred, it was not providing a means of comparing actual versus expected performance. In trying to do this, it was discovered that while the status and tracking system did contain all the relevant statistics, methodical extraction of summarized reports from it was not very refined.

As the understanding of the capabilities and the limitations of the status and tracking system evolved, so did the reporting. Confidence was gained in the ability to correctly interpret and summarize the reports, and the number of elaborate charts and matrices constructed decreased. By the middle of Phase III, the reporting had come full cycle: data flow information was presented essentially in a tightly summarized form, consisting of an assessment of (1) the goal, (2) progress to date, (3) schedule variations, (4) forecasting regarding adherence to milestones, and (5) problem areas.

In retrospect, this portrayal of the status of the LACIE data system is so obvious that one must wonder how any other course for depicting the information could have been considered. It must be pointed out, however, that synopsis was more of an acquired skill than an exact science and was in fact subjected to an evolutionary process throughout LACIE.

EOD Facilities Configuration Management Office

V. M. Dauphin^a and R. F. Palmer^b

INTRODUCTION

During its early stages, the Earth Observation Division (EOD) Data Techniques Laboratory (DTL) at the NASA Johnson Space Center (JSC) was a test bed for experimental software and hardware techniques supporting scientific data processing. A central control point for management of configuration updates and/or modifications did not exist. Individual DTL users updated or modified the facility at will, within their own projects and assignments. This naturally resulted in nondocumented, uncontrolled configurations evolving through software updates and modifications by programmers. Attendant problems included scheduling conflicts, lack of defined procedures, and software processing via untested software.

The requirements of the LACIE project on the DTL changed its role from a testing facility to a production facility. It was paramount that resource scheduling and control of laboratory hardware and software be implemented.

PLANS FOR CONTROL

In May 1975, EOD management approved the formation of the Facilities Configuration Management Office (FCMO). FCMO was to provide process control for LACIE support within the division. The primary task was to establish standards and procedures for central DTL configuration control for hardware changes, software changes (both system and application), future modification to established baselines, temporary changes, and anomaly resolutions. The goal of these standards and procedures was to assure compatibility of requirements, plans, and applica-

tions as interrelated with the multitude of users within the framework of well-defined support facility resources. This task was later expanded to include DTL system operating software, machine-user identification access control, and documentation of pertinent information required by in-house processor users. Since the DTL was a functioning user facility, implementation of classic configuration control procedures could not be implemented without interruption of service. Implementation had to proceed firmly, but smoothly, along established user modes.

The first step in establishing configuration management for the FCMO was the development of the DTL Software Users Guide. This guide outlined DTL processor configuration, defined the system capabilities available to users developing application software, and established system resources rules and limitations.

Then, while DTL users were absorbing and being tutored in the DTL Software Users Guide, FCMO personnel developed a Configuration Management Plan (CMP)—the first step in a classic approach—that was reviewed by JSC management at periodic stages to assure adherence to established agency standards.

IMPLEMENTING CONTROLS

The purpose of the CMP was to maintain the integrity of all DTL production software. Because implementation of the plan could not be accomplished in one motion because of the continuing multiuser nature of the DTL, the plan had to be instituted in an orderly, nondisruptive manner to support the LACIE project.

First, system baselines were established. This was achieved by accepting all production software as de facto acceptance tested. With this baseline established, all further changes were required to undergo acceptance testing as defined by CMP procedures.

^aNASA Johnson Space Center, Houston, Texas.

^bFord Aerospace and Communications Corporation, Houston, Texas.

The Controlling Organization

As the DTL was evolving into a real-time, multiuser facility, it became apparent that a single responsible body was required to control modifications, access, and utilization of the operating system.

This requirement was satisfied with the establishment of the Systems Management (SM) function within the FCMO. As the technical arm of FCMO, SM was given the responsibility to maintain and control the modifications to all the baselined software utilized for production in the DTL.

In addition to user control, SM continues to provide the following functions.

1. Install and maintain the operating system software.
2. Investigate system failures and either report the failure to the vendor for response or initiate a local system correction or workaround.
3. Install all pertinent operating system corrections received from the vendor.
4. Inform the general user community of all new features, modifications, problems, and workarounds in regard to the DTL operating systems.
5. Assign user identification codes and protection codes that control access and level of access to system capabilities.
6. Provide analysis support to users experiencing problems with either the system features during software development or acceptance-tested software during production runs.
7. Provide system backups and backup procedures to allow regeneration of the baseline system in the event of total system failure.

Enforcement of Controls

To ensure the effectiveness of the SM function, FCMO has the power to enforce the rules and procedures as defined in the CAPP. For instance, if a user chooses not to follow established rules or procedures, FCMO can refuse the user access to the system. This system denial may be in the form of refusal to schedule an acceptance test, or it may be through computer lockout of the user via a computer console entry. These rules and procedures are necessary to avoid one user making modifications to the system that would, in effect, put other users of the system out of business. Review, coordination, and impact analysis of all planned modifications in relation to other users is the responsibility of FCMO.

Included in the configuration management plans are acceptance test procedures for quality assurance (QA) signoff for both hardware and software modifications.

User Allocations and Identification

Because of the limited disk data storage in the multiuser environment, the amount of block storage utilized by each user must be defined and controlled. To accomplish this, each system user must file a formal request with FCMO via a "User Identification Code (UIC) Action Form." This form specifies the amount of storage required to support each user task. The form is submitted to the Data Base Manager (FCMO NASA Task Monitor), who approves or disapproves the request. If the task is disapproved, the Data Base Manager notifies the requester. If the request is approved, the requester is assigned a unique UIC and the allotted storage block. The Data Base Manager directs FCMO SM to install the new UIC in the system. When this is accomplished, the UIC requester is notified and the request for system access has been approved and the new unique UIC has been installed. Each system user must sign on and enter the specific unique UIC to gain system access.

Because of personnel attrition, transfers, etc., all UIC's must be kept current. This is accomplished by requesting all system users to renew their UIC and computer storage request on a quarterly basis. All UIC's not renewed within 10 days after the specified renewal time are purged to tape and retained for 60 days. UIC renewal dates are communicated through an FCMO bulletin. The FCMO bulletin is also used to convey information concerning newly installed hardware and modifications to systems and production software.

SUMMARY

All DTL configuration management procedures and guidelines were implemented by the FCMO within an 18-month period. This time frame allowed easy transition from an unmanaged to a managed condition. It further allowed users ample time to become acclimated to the newly established procedures, and it allowed the procedures themselves to be "fine tuned" for more effective control.

Configuration management has improved the efficiency of the operational system in the DTL. System

scheduling and operational problems have diminished to no more than an occasional operator error. Because of the formal acceptance test procedures, software reliability has improved and users have gained respect for system integrity. Finally,

FCMO tracking of all deliverable documentation assures users the documents associated with new and/or updated software are available at the time of software installation.

Accuracy Assessment System and Operation

D. E. Pitts,^a A. G. Houston,^a G. Badhwar,^a M. J. Bender,^a M. L. Rader,^b W. G. Eppler,^c
 C. W. Ahlers,^b W. P. White,^b R. R. Vela,^b E. M. Hsu,^b J. F. Potter,^b and N. J. Clinton^b

INTRODUCTION

The LACIE crop estimation system is composed of several operational subsystems: data collection, classification, yield estimation, crop aggregation and reporting, data storage and retrieval, and accuracy assessment. The Accuracy Assessment Subsystem is responsible for determining the accuracy and reliability of LACIE estimates of wheat production, area, and yield made at regular intervals throughout the crop season and for investigating the various LACIE error sources, quantifying these errors, and relating them to their causes. Timely feedback of these error evaluations to the LACIE project was the only mechanism by which improvements in the crop estimation system could be made during the short 3-year experiment.

Figure 1 illustrates the accuracy assessment data flow. Estimates of wheat production, area, and yield are compared with accurate reference data. For example, in the yardstick region of the nine states in the U.S. Great Plains (USGP), the U.S. Department of Agriculture (USDA) Economics, Statistics, and Cooperatives Service (ESCS) estimate is used as the reference. In areas outside the United States, the USDA Foreign Agricultural Service (FAS) and official country estimates are used as the reference standard. In most cases, the LACIE reports are published a few days before the release of the corresponding ESCS or FAS report. The LACIE estimates and standard error estimates are compared month by month to the corresponding reference data as well as to the end-of-crop-year reference data to determine whether the project accuracy goal of 90/90 (i.e., 90 percent of the time, the estimate of wheat production

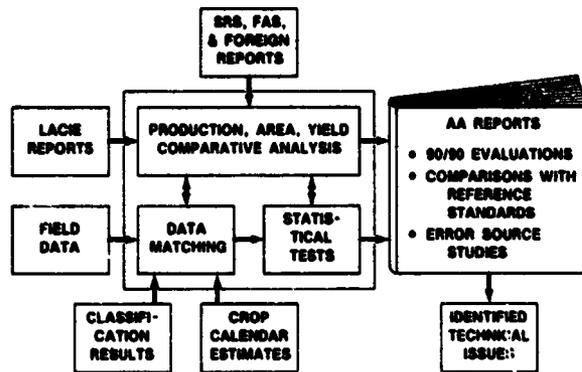


FIGURE 1.—Accuracy assessment analysis steps and flow of data.

should be within ± 10 percent of the reference) is being met (see the paper by Houston et al. entitled "Accuracy Assessment: The Statistical Approach to Performance Evaluation in LACIE").

To produce timely reports of these results, a brief accuracy assessment report called a "quick-look report" is published a few days following each LACIE crop estimation report. Four times each year, the Accuracy Assessment Subsystem produces comprehensive reports of the error source studies, which separate the wheat production error into its component parts of wheat area error and wheat yield error. These errors are further divided into component parts based on field observations (fig. 2).

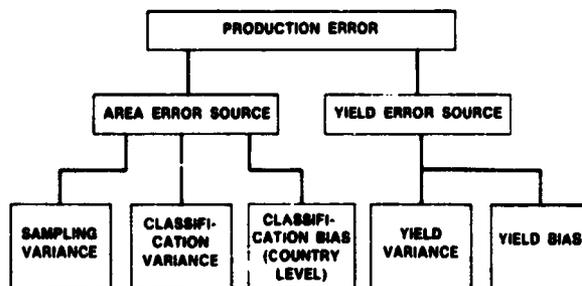


FIGURE 2.—Production error components.

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Houston, Texas.

^cLockheed Missiles and Space Company, Palo Alto, California.

Figure 3 shows the field data used in accuracy assessment. The 29 intensive test sites are special nonoperational sites on which very detailed data are collected each 18 days. These sites are used for Classification and Mensuration Subsystem (CAMS) procedure verification and are also used in the quick-look reports to illustrate particular situations encountered during the crop year. Because of the widely varying crop conditions from one county to the next, ground truth on large numbers of operational sample segments must be obtained to properly separate potential error sources such as classification and sampling. The 166 blind sites with their "wall to wall" inventories meet this requirement. Throughout the crop year, all elements of the project receive information resulting from the blind site studies. However, to protect the integrity of the blind sites as a testing tool, the ground-truth information from the blind sites is not released to elements of LACIE other than the Accuracy Assessment Subsystem until after the year is over.

BACKGROUND

During the mid-1960's, agricultural remote-sensing experiments were performed using aircraft platforms. Photographs from these missions were used to construct photomosaics, which were used as base maps for recording the crop identity observed

on the ground. An example is the familiar C-1 flight-line multispectral scanner (MSS) experiment conducted in Indiana by the Purdue University Laboratory for Applications of Remote Sensing (LARS) (refs. 1 and 2), which included the recording of ground-truth and ancillary data for about 400 fields. Field vertices were manually registered to the MSS gray-scale printouts, which were then used to construct test fields in the classification data base for evaluation of accuracy. This procedure was not viable for LACIE, since the use of test fields in the operational data base would have potentially compromised the results.

With the advent of the Landsat spacecraft, ground truth could be registered to the Landsat MSS and used several times during the crop year so long as crop rotations, abandonment, etc., did not occur and the pass-to-pass registration system (see the paper by Grebowsky entitled "LACIE Registration Processing") was accurate to ± 1 pixel. This method was used in Crop Identification Technology Assessment for Remote Sensing (CITARS),¹ in which six counties in the Corn Belt were sampled, with one 5-by-20-statute-mile sample located in each county. To train the CITARS classifier to distinguish corn and soybeans from other crops and to assess classification accuracy, ground-truth inventories were performed in 20 randomly chosen quarter sections in each of the 6 segments. These data were transferred by image interpretation to gray-scale printouts or cluster maps of Landsat data. Thus, for these sample segments comprising 600 square statute miles, 30 square statute miles were inventoried by the USDA Agricultural Stabilization and Conservation Service (ASCS) each 18 days—six times during the growing season.

The CITARS effectively illustrated the need for ground truth to determine error sources in classification, but the project did not involve large-region production estimation and therefore did not include area estimation sample error, yield estimation error, or the wide range of classification error that can occur over large areas containing hundreds of counties with widely varying climatic conditions and cropping practices. Early in LACIE, the only tool for identifying sources of classification error was the intensive test site. However, with only 29 intensive test sites, none of which were located in Colorado, Oklahoma, or Nebraska, this sample did not adequately represent the 9 states of the U.S. Great Plains, containing

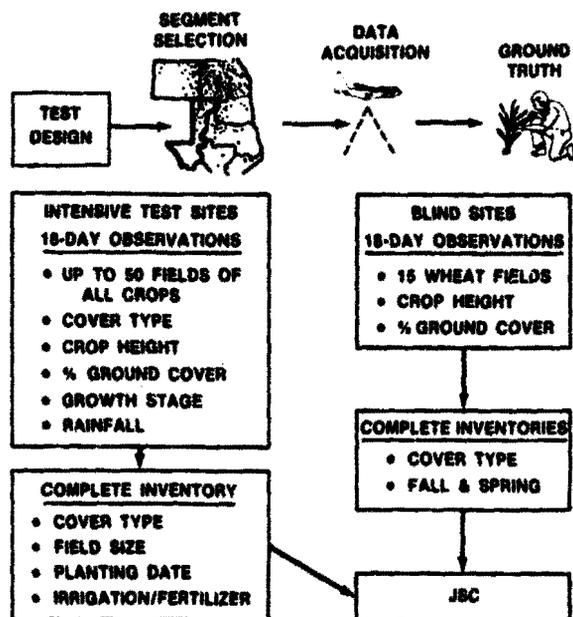


FIGURE 3.—Field data acquisition.

¹Task Design Plan, Vol. 1, CITARS. Rep. JSC-09384, Feb. 1975.

several hundred segments. Furthermore, the intensive test sites were chosen to include high percentages of wheat and small grains and therefore were not necessarily representative of the sites chosen by the LACIE sampling strategy.

To overcome these shortcomings, a pilot effort was undertaken in Phase I in which 29 LACIE segments in North Dakota and Montana were inventoried in August 1975. The term "blind site" was chosen for these sites because they were part of a blind test in which none of the LACIE analysts knew the sites chosen or had any information about the identity of the fields as determined in the inventory. In Phase I, the 29 blind sites were selected randomly from the set of about 55 sites in Montana and North Dakota with biwindow 1 and 2 acquisitions. Acquisition histories as of late July 1975 were used as the basis of selection. Proportions of small grains were determined by planimetry of the ground-truth annotated photography. In the Phase I blind sites, wall-to-wall ground truth was prepared by a combination of ground surveys, aerial inspection from light aircraft, and interpretation of current aerial photographs. Without visits to all the fields, separation of oats, rye, and barley from wheat was not always possible. In Phases II and III, the inventories were conducted using current aerial photographs as the base map and ground surveys to determine each field crop type.

In Phase II, 136 blind sites were selected randomly from the LACIE crop estimation reports. The segments for the Southern Great Plains were selected randomly from the segments used in the February 1976 Crop Assessment Subsystem monthly report (CMR). The June CMR was used for Montana and South Dakota winter wheat, and the July CMR was used for the spring wheat states. These segments were chosen so as to represent equal numbers of segments with high as well as low estimates of small grains. Wall-to-wall ground truth was then collected for all fields by the ASCS personnel using current aerial photographs as the base map.

In Phase III, 166 blind sites were selected randomly from the USGP segments so that each crop reporting district had approximately one-third of the segments chosen as blind sites. This selection was made in October 1976 before the crop year commenced. Wall-to-wall ground truth was then collected for all fields by the ASCS personnel using current aerial photographs as the base map. Proportions of all crops were determined by (1) digitizing the field vertices from the annotated aircraft photo-

graph, (2) registering the data to the Landsat MSS image, and (3) numerically integrating the area of each field. To ensure timely assessment of proportion accuracy, the proportions of wheat and small grains were also estimated for both the planted and the at-harvest inventories, using a dot grid overlaid on the ground-truth annotated photograph.

CLASSIFICATION ACCURACY

A crop estimation system such as LACIE encounters a wide variety of phenomena that can contribute to the classification error (fig. 4). The Accuracy Assessment Subsystem has the task of determining the contribution of each of the components to the overall classification error as well as to the area estimation error on the state and national levels. Estimates of the percentage of small grains in the LACIE 5- by 6-nautical-mile segments have been found to contain errors caused by various sources, of which the most important were (1) abnormal signature development due to a variety of causes including late planting, drought, cattle grazing, crop rotation, disease, and soil variability; (2) inadequacy of the Landsat scanner in resolving small fields; and (3) mislabeling of small grains, grasses, pasture, and idle fallow when key acquisitions are missing.

In each of the last 2 years of LACIE, ground truth was collected for one-third of the operational sample segments. Approximately 12 000 square statute miles of ground-truth data have been produced (about twice the area of the state of New Jersey)—easily the largest amount ever attempted in agricultural remote sensing. The ground truth was used to evaluate the error sources and magnitudes for 64 000 analyst dot

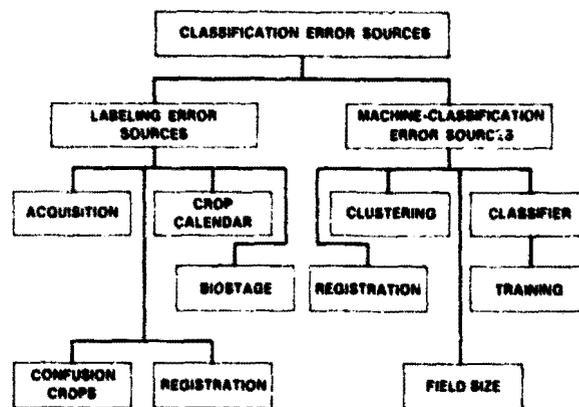


FIGURE 4.—Classification error components.

labels per crop year and the accuracy of clustering and classification of about 14 000 000 pixels per year. In order to process these large amounts of data, the ground-truth wall-to-wall inventories were digitized and registered to the Landsat image digital data. This procedure not only enables objective repeatable experiments to be performed by all users but also enables traceability of the ground-truth data from the fieldwork to the pixel label and facilitates rapid, efficient correction of errors in the digitized ground-truth data. Since a ground-truth label is obtained for each Landsat MSS pixel, it is very easy to compare other LACIE data produced at the pixel level including the following.

1. Landsat data and transformations, such as green number
2. Analyst labels
3. Classification maps
4. Cluster maps

In addition, data produced at the field level can be evaluated in terms of their effect on the pixel-level accuracies. Examples include the following.

1. Crop stage development
2. Yield
3. Fertilizer effects
4. Irrigation effects
5. Cropping practices
6. Rainfall
7. Soil type
8. Atmospheric optical depth

Thus, many other LACIE groups besides Accuracy Assessment find these data important; they include classification technique development groups, spectromet yield model development groups, and spectromet crop development model groups.

The processing of the ground truth into digital forms and its use in Accuracy Assessment are illustrated in figures 5 and 6. A Phase III blind site (number 1523 in Minnesota) is discussed in the following paragraphs as an example of these processing steps.

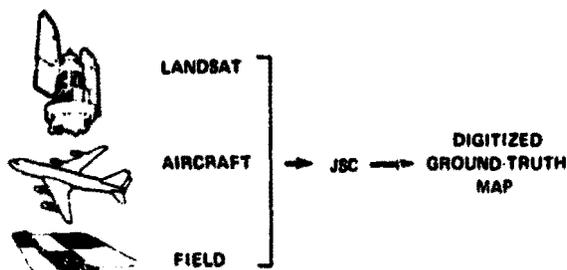


FIGURE 5.—Collection of data for blind sites and intensive test sites.

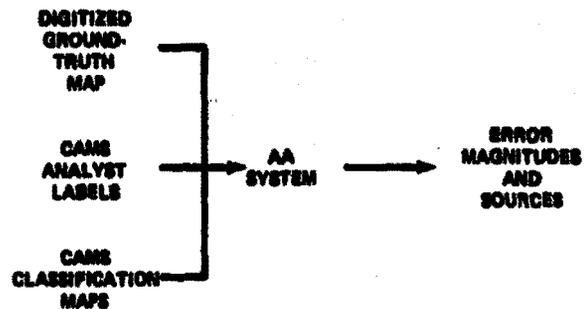


FIGURE 6.—Use of field data to evaluate classification accuracy.

This segment was randomly chosen from the available blind site segments with good acquisition histories.

Aircraft Maps

After selection of the blind sites, Landsat imagery is used to determine the true position of each site, which is plotted on a 1:24 000-scale or 1:12 500-scale map for use by aircrews in acquiring the aerial imagery.

Aircraft Photographs

Aerial photography using color-infrared film is collected for each blind site by using predesignated flight lines. If possible, flight altitudes are greater than 14 kilometers to enable single-photograph coverage of the entire site. These photographs are enlarged to a scale of 1:24 000 for field use (fig. 7).

Field Overlays and Field Segment Kits

If the imagery is of satisfactory quality, transparent overlays are prepared (ref. 3) and forwarded to ASCS personnel in the appropriate county (see the paper by Spiers and Patterson entitled "Ancillary Data Acquisition for LACIE").

Blind Site Field Data Acquisition

The USDA ASCS personnel provide complete inventory data based on ground observations. The overlay is annotated with the standard crop symbols for each field (ref. 3). These inventory packages are completed by the ASCS personnel and forwarded to the NASA Johnson Space Center (JSC) to be logged and tracked by the Data Acquisition, Preprocessing, and Transmission Subsystem (DAPTS).



FIGURE 7.—Aircraft photograph of blind site segment 1523, Wilkin County, Minnesota, June 1, 1977.

Near the time of wheat emergence, 15 fields in each USGP blind site are chosen and annotated on the overlay by ASCS personnel in accordance with the following: 5 below-average fields, 5 average fields, and 5 above-average fields. The ASCS personnel identify the plant height and ground cover of each of these fields at this time. Beginning in October for winter wheat and in May for spring wheat, they revisit these 15 fields in concert with the Landsat overpasses so that classification performance can be related to wheat field stands. Observations are continued until all 15 fields are harvested or abandoned. Inventories for the winter wheat sites are attempted twice (planted and harvested), but only an at-harvest inventory is obtainable for spring wheat because of the shorter growing cycle.

Delineating Photographs

The JSC Cartographic Laboratory overlays the photographs with a second sheet of acetate and outlines each of the homogeneous areas as irregular polygons with 39 or fewer vertices. A homogeneous area has a uniform cover type; examples are a wheat field, a pond, pasture, timber. Curved field boundaries are approximated by a series of straight lines.

Assigning Crop Code and Field Number

As the polygons are being delineated, each one is assigned a digital code (ref. 3) which indicates a par-

ticular crop type determined by the field personnel. A field number is assigned for use in quality assurance and to enable efficient correction of errors on the interactive disk data file.

Digitizing Field Vertices

After the fields have been annotated and delineated on acetate overlays, the polygon vertices are measured and stored on the interactive drafting system. The vertices are measured sequentially and encoded together with the field crop code and number. The digitized results are plotted with a line plotter (fig. 8) for quality checking and prepared for final output and registration. The digitization of the 200 to 1200 fields in a LACIE sample segment takes 6 to 14 hours and is the major throughput problem in the Accuracy Assessment Subsystem.

Registration

To define the geometric relationship of the aerial photograph to the Landsat segment, registration coefficients for each photograph must be obtained (see the paper by Rader and Vela entitled "Cartography: LACIE's Spatial Processor"). This procedure entailed selecting 8 to 12 points per photograph and independently solving for the coefficients of each photograph using least squares techniques. When



FIGURE 8.—An expanded portion of blind site segment 1523 showing a Gerber plot of digitized field delineations for quality control.

more than one photograph is used in the inventory, tie points are necessary to make points common to both photographs occur on the same Landsat line and sample position.

Conversion of Field Vertices to Universal Format Tape

Because the LACIE imagery and the LACIE output products (classification and cluster maps) are in universal format, which can be read by the software of many institutions, the universal format was used for the ground-truth information. To perform registration with accuracy greater than ± 1 pixel and to assess the effect of boundaries and mixed pixels on the classification process, the decision was made to digitize the ground truth with six subpixels comprising each Landsat pixel (fig. 9).

The accuracy assessment software (ref. 4) reads the tape containing the field vertices and determines all the subpixels falling within each field. Each subpixel is assigned a digital code that identifies the crop found in the field observation. Gray-scale printouts of the universal tape are produced at the pixel and subpixel levels. These printouts are checked for accuracy by Cartographic Laboratory, Accuracy

Assessment, and CAMS personnel against the quality assurance plot and the original photograph. If a discrepancy occurs for any field, the cause is determined, the disk data file is updated, and a new tape is produced and checked for accuracy.

Registration Accuracy

To date, only a limited study has been done on the registration accuracy of this process. A segment with large fields in Oklahoma (1048, Cimarron County) and a segment with small fields in North Dakota (1602, Mountrail County) were misregistered by known amounts to determine the number of ground-truth pixel label assignments that would change. The results for the segment with large fields were as follows.

Error	Change
0.5 pixel	4 percent of labels of 209 dots
1.0 pixel	8 percent of labels of 209 dots
1.5 pixels	12 percent of labels of 209 dots

The results for the segment with small fields were as follows.

Error	Change
0.5 pixel	8 percent of labels of 209 dots
1.0 pixel	14 percent of labels of 209 dots
1.5 pixels	19 percent of labels of 209 dots

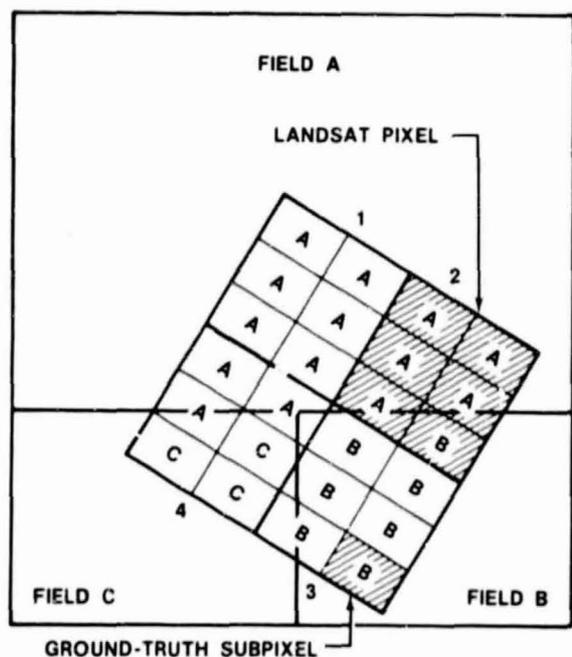


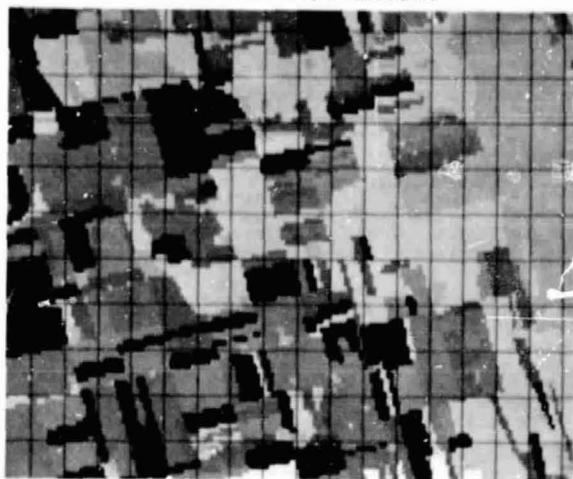
FIGURE 9.—Example of digital ground truth. Pixels 1 and 3 are estimated as pure, while pixels 2 and 4 are mixed.

To verify the registration accuracy, detailed photo-interpretation was done for all dots in 11 segments in Oklahoma and 18 segments in North Dakota, taking into account the NASA Goddard Space Flight Center (GSFC) misregistration from pass to pass. Comparison with the ground-truth digital tape showed 4-percent disagreement in Oklahoma and 10-percent disagreement in North Dakota. In accordance with the preceding tables, these values indicate a registration accuracy of about ± 0.5 pixel for the digital ground-truth product.

Ground-Truth Tape Image

Figure 10 shows a ground-truth tape for blind site 1523 as imaged on the Image-100 (I-100) system. The I-100 can read only every third line of ground-truth tape; therefore, the field boundaries appear more uneven than they are on the digital tape. For clarity,

WILKIN COUNTY, MINNESOTA



INVENTORY DATE, AUGUST 1, 1977

- SPRING WHEAT - SALMON PINK
- HARVESTED SPRING WHEAT - LIGHT PINK
- OTHER SPRING GRAINS - PINK
- HARVESTED OTHER SPRING GRAINS - GOLD
- OTHER CROPS (CORN, SUNFLOWER, SOYBEANS, SUGAR BEETS) - RED
- PASTURE, GRASS, HAY, ALFALFA - BLUE
- IDLE FALLOW, IDLE COVER CROPS, IDLE RESIDUE - GRAY
- NON AGRICULTURE (TREES, HOMESTEAD) - GREEN

FIGURE 10.—Digitized ground truth registered to Landsat imagery for LACIE segment 1523, Phase III blind site.

many of the numerous ground-truth classes are lumped together to form a superclass; e.g., pasture, grass, hay, and alfalfa are all displayed by the color blue. For comparison, the corresponding Landsat data are shown in figure 11.

The preparation and use of LACIE accuracy assessment ground-truth tapes does not require that the entire scene have contiguous inventory. Before digitization of all field vertices, the entire scene is made into one large field containing the crop code representing areas without ground observations and, thereby, a crop code for each Landsat pixel is ensured.

Evaluation of Labeling Errors and Wheat Proportion Error

Using these digital ground-truth data, many types of labeling error analyses are performed routinely for all LACIE blind sites.

1. All crop proportions are correlated with wheat proportion estimation accuracy.

2. Analyst labeling accuracy is determined for all labeled dots.

3. Sampling accuracy is determined for all 209 dots and for the subset labeled by the analyst.

4. Histograms of Landsat data for wheat are compared with histograms of labeled dots to evaluate signatures omitted.

An example for segment 1523 is given in table I, which shows that labeling accuracy and the accuracy of estimation of small grains and wheat improved as the season progressed.

Moreover, it is not sufficient just to know the accuracies. The Accuracy Assessment group must investigate the causes of each mislabeling of a dot. In this error characterization, a special analyst uses the ground truth and the information in the CAMS packet to attempt to deduce the mislabeling cause. These causes can be grouped into three categories.

1. Those causes the analyst can do very little to correct, such as insufficient acquisitions, border/edge locations, and narrow fields near the limit of sensor resolution

2. Those causes representing abnormal signatures in production film converter (PFC) Product 1 (e.g., fig. 11), which are inconsistent with the CAMS procedure, such as wheat fields with temporal color sequences that do not follow the wheat growth cycle, nonwheat fields that do follow the wheat growth temporal color sequence, and temporal sequence signatures of wheat fields that are behind or ahead of the majority of the wheat fields in the same acquisition

3. Those causes that are merely interpretation or clerical errors

Evaluation of Classification and Cluster Maps

The clustering and subsequent classification results from each CAMS run (fig. 12) are transmitted to Accuracy Assessment on digital tape. The Accuracy Assessment Subsystem compares all these products for each classification, pixel by pixel, with the ground truth to determine the accuracy with which they are produced. For example, the classification shown in figure 12 was compared to the ground truth in figure 10. Of the 54 437 small-grains subpixels in the scene, 46 042 were correctly classified as small grains; likewise, of the 83 155 non-small-grains subpixels, 67 571 subpixels were correctly classified as non-small-grains. In fact, each ground-truth class is investigated to see how much is called small grains

BIOWINDOW 1, BIOSTAGE 1.5



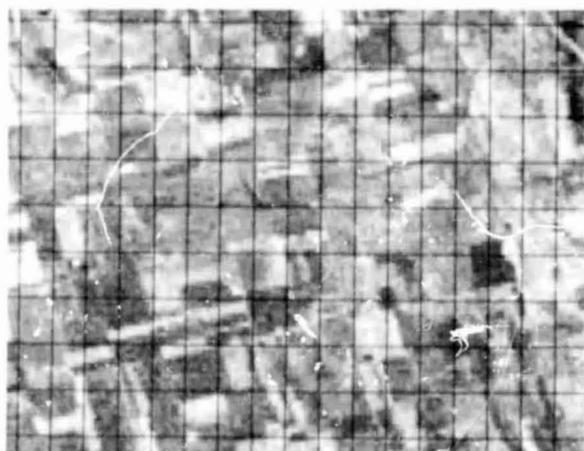
PLANTED, APRIL 30, 1977

BIOWINDOW 1, BIOSTAGE 2.5



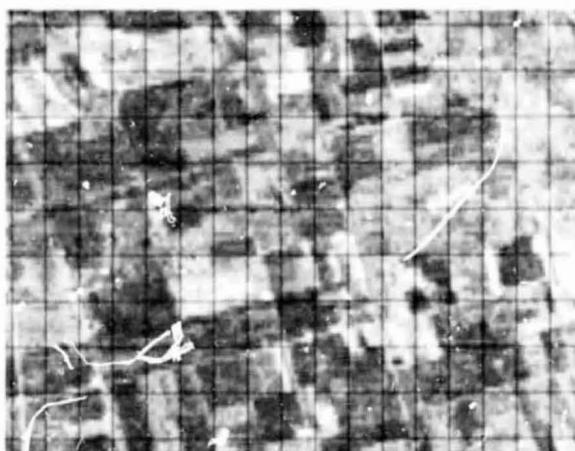
**2.5- TO 4.5- IN. WHEAT,
MAY 18, 1977**

BIOWINDOW 2, BIOSTAGE 3.5



**6.5- TO 13- IN. WHEAT,
JUNE 5, 1977**

BIOWINDOW 2, BIOSTAGE 4.3



**16- TO 30- IN. WHEAT,
JUNE 23, 1977**

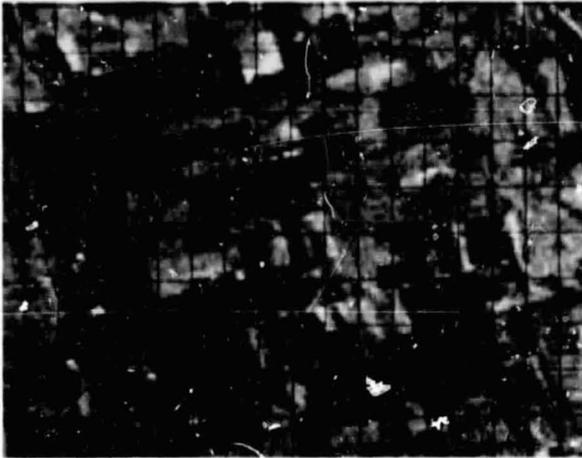
FIGURE 11.—LACIE Phase III blind site 1523, Wilkin County, Minnesota.

and how much is called non-small-grains. In this example, most oats, spring wheat, and barley were correctly called small grains; however, some spring wheat and oats were called non-small-grains. Some corn, sunflower, grass, hay, and pasture were erroneously included in the small-grains category.

Figure 13 shows the location of the 15 wheat fields in blind site 1523. These fields are used to identify the causes of labeling and classifying small grains as non-small-grains. One way in which the 15 fields and their crop height and ground cover information are

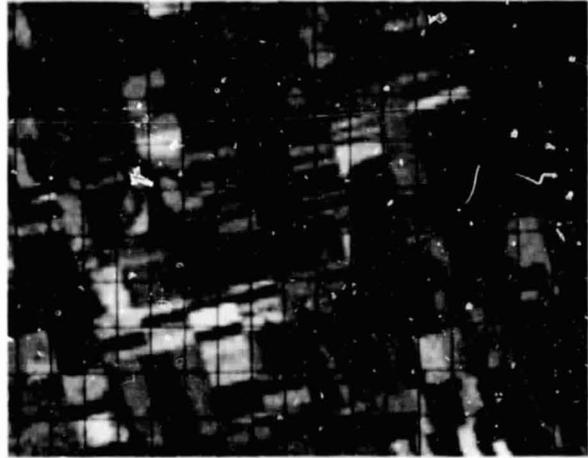
used is to determine the labeling accuracy of dots that fall within these fields. Another use is to investigate the classification accuracy of each of the 15 fields. In the example segment, all the fields had excellent accuracy (better than 95 percent) except two fields which were found to occur in a region of poorly drained soil. The signatures of the wheat fields in this soil were not identified by the analyst because no dots fell within any wheat fields in this area of the scene. However, these signatures had enough commonality with the wheat field signatures

BIOWINDOW 2, BIOSTAGE 4.3



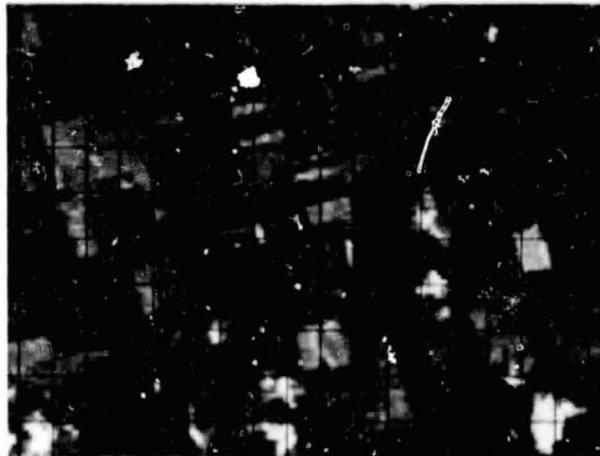
16- TO 30- IN. WHEAT,
JUNE 24, 1977

BIOWINDOW 4, BIOSTAGE 6.0



50-PERCENT HARVESTED,
JULY 29, 1977

BIOWINDOW 4, BIOSTAGE 6.0



POSTHARVEST, AUGUST 17, 1977

PRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

FIGURE 11.—Concluded.

in the remainder of the segment to enable correct identification of about 65 percent of these two fields. This example illustrates the diverse types of labeling and classification problems that can be investigated using the 15-field information together with the accuracy assessment software.

Ground-Truth Spectral Plots

Other diagnostic tools are available to aid in the determination of sources of classification and label-

ing errors. One of these is the scatterplot—a two-dimensional histogram plot for two Landsat MSS channels or for rotations of these data (e.g., greenness and brightness). Figure 14 shows such a plot for spring wheat in the poorly drained soil in segment 1523. Plots of this type are useful in determining spectral signatures of wheat and other crops in the scene, as well as within-segment variability of signatures due to soil type, planting date, irrigation, fertilizer applications, drought, atmospheric effects, crop rotation, variety, and disease.

TABLE I.—Dot-Labeling and Estimated Proportion Accuracy for Four Classifications for Segment 1523 as Compared to Ground Truth

Acquisition date/biostage (1977)	Estimated spring wheat, percent	Estimated spring grains, percent	Dot type	Percent labeled	Percent correct small- grains dots	Percent correct other dots
April 30 1.5, May 18 2.5	11.0	19.3	2 1	60 32	25 80	97 76
April 30 1.5, June 5 3.5, June 24 4.3	24.0	42.0	2 1	54 32	82 80	85 76
April 30 1.5, June 24 4.3, July 29 6.0	24.3	42.5	2 1	51 32	79 93	83 100
Ground truth	20.3	40.0				

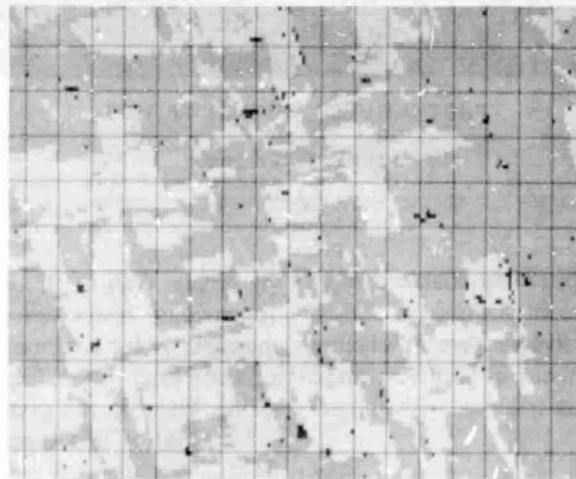
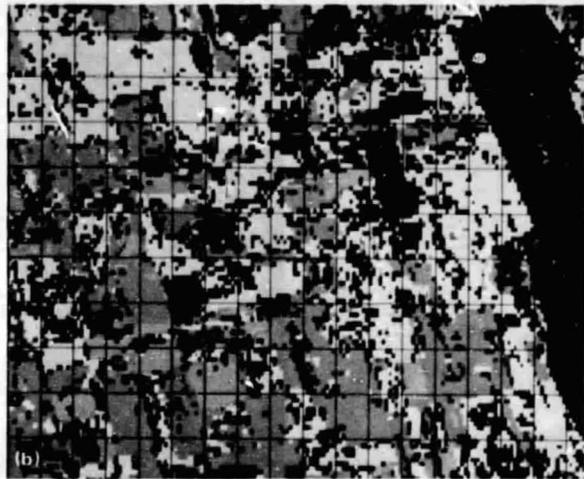
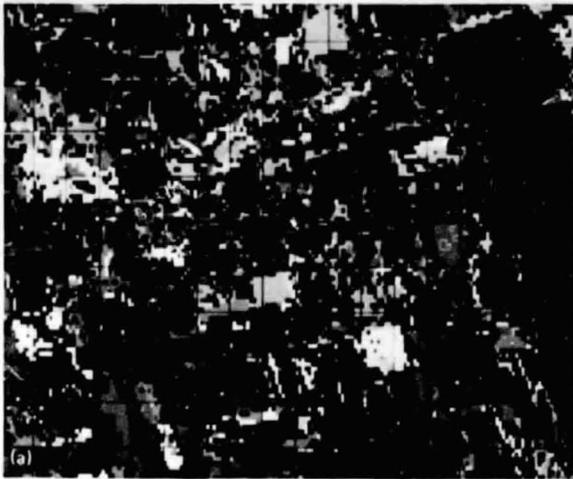


FIGURE 12.—Computer-generated cluster and classification maps for blind site segment 1523, Wilkin County, Minnesota. (a) Unconditional cluster map before assignment of clusters to classes; August 17, 1977. (b) Conditional cluster map; black = threshold, DO, DU; yellow = nonspring small grains; green = spring small grains; other = conditional clusters; August 17, 1977. (c) Classification map; black = threshold; green = spring small grains; orange = nonspring small grains; July 29, 1977.

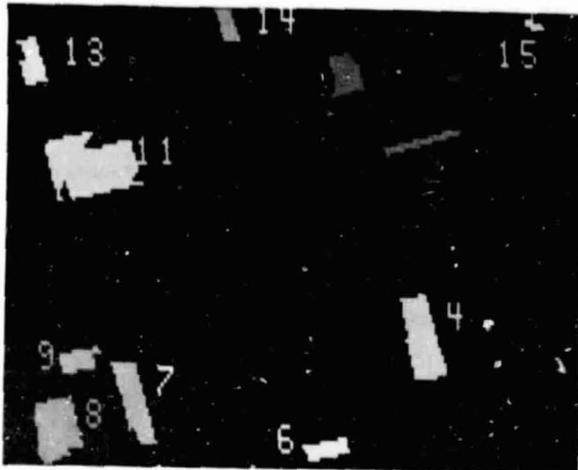


FIGURE 13.—Digitized ground truth registered to Landsat imagery for selected spring wheat fields in LACIE segment 1523, Phase III blind site, Wilkin County, Minnesota, 1977.

YIELD ACCURACY

To support the evaluation of the production estimates in meeting the 90/90 criterion (see Houston's paper), the yield estimates must be tested for bias, and the accuracy of the estimated variances must be determined. Accuracy Assessment (see the paper by Phinney et al. entitled "Accuracy and Performance of LACIE Yield Estimates in Major Wheat-Producing Regions of the World") uses 10 or more years of independent temperature and precipitation data to test the yield model for each zone (state) for all monthly truncations (ref. 5). During the crop year, the Accuracy Assessment group attempts to determine the error sources by investigating the modeling error sources—trend, variable selection, and stability of coefficients—together with the measurement error sources—temperature and precipitation (fig. 15).

To evaluate the contribution of measurement error sources, temperature and precipitation sampling errors are determined for states for which large deviations from the ESCS estimate are observed. Sampling error is determined by comparing the objective analysis using the dense network of meteorological stations (cooperative network) with that using a manual synoptic analysis of the sparse network of meteorological stations (climatic stations) used operationally by LACIE for each climatic district. In the example case (figs. 16 and 17), the sparse network for Oklahoma gave consistently less estimated precipitation than did the dense network,

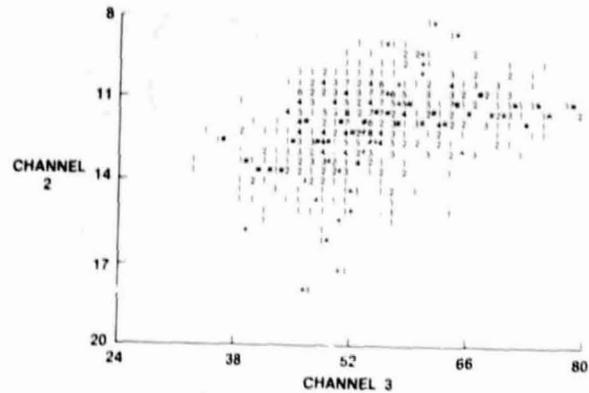


FIGURE 14.—Scatter plot of channels 2 and 3 for segment 1523, June 24, 1977, for spring wheat in poorly drained soil (numbers indicate occurrences).

which agreed closely with the synoptic LACIE estimate (table II).

To search out other measurement error sources, the mean and standard deviation of temperature and precipitation are plotted as a function of time of year. The current-year monthly temperature and precipitation are also plotted on these charts for ease of comparison.

Several tasks are undertaken to check for modeling error sources.

1. Temperature and precipitation are plotted versus month for the 3 highest yield years and the 3 lowest yield years from the historical record.

2. The trend term is evaluated by performing latent root regression without allowing for trend and calculating trend on the residuals from the most stable fit. This procedure has the advantage of removing the long-term changes of the climate from the trend, but it has the disadvantage of not allowing for the interactions between climate and agricultural technology.

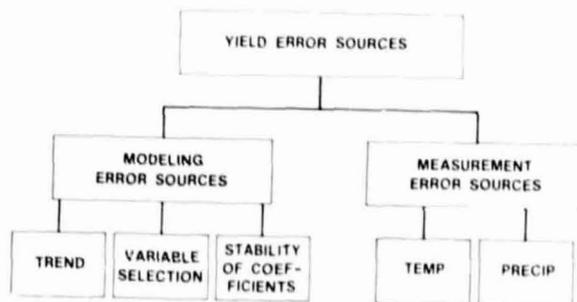


FIGURE 15.—Yield error components.

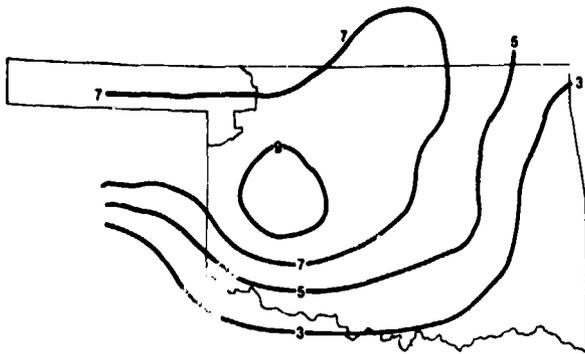


FIGURE 16.—Objective analysis of May 1977 Oklahoma pseudozone precipitation (in inches) using 14 stations.

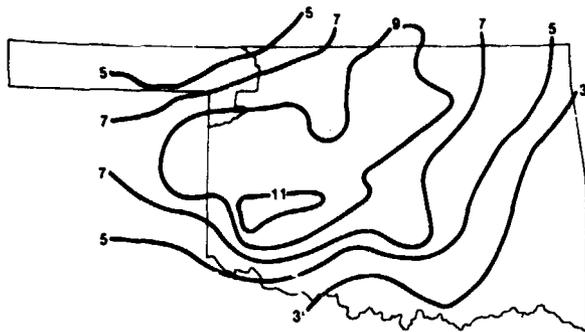


FIGURE 17.—Objective analysis of May 1977 Oklahoma pseudozone precipitation (in inches) using 147 stations.

3. The variable selection is verified by adding the current year to the data set and reselecting the variables. Latent root regression on the pseudozone yield and meteorological data enables determination of the most stable set of variables. An example of this result for Oklahoma (table III) shows little difference between the variables for the LACIE Phase III models and the optimum variables for statistical stability.

RESULTS

The following is a brief description of some of the results of using the Accuracy Assessment Subsystem during the 3 years of LACIE. More detailed discussion of these results can be found in symposium papers by Houston et al., Phinney et al., Marquis, Hickman, and Conte.

In Phase I, operational estimates of wheat area only were made for the USGP. A comparison of the USDA ESCS and the LACIE estimates indicated

TABLE II.—Analysis of Sample Error for May 1977 Precipitation in Oklahoma

Climatic district	Precipitation, ^a in.		LACIE synoptic analysis
	Sparse	Dense	
South central	4.68	5.08	—
Southeast	2.31	3.10	—
North central	8.05	9.08	—
Central	7.34	8.13	—
West central	9.01	10.17	—
Southwest	7.33	8.64	—
Northeast	5.13	6.02	—
State	7.71	8.76	8.90

^aObjective analysis.

support of the 90/90 criterion for winter wheat. Significant underestimates were found in the spring wheat region, the largest in North Dakota. To better understand these differences, the blind site program was initiated. A statistical comparison of the LACIE estimate with the ground-truth data and with the ESCS county estimates for 20 blind sites in North Dakota indicated that the classification accuracy was good and that the source of the problem was sampling error (fig. 18). Because of timely feedback from Accuracy Assessment, approximately 20 additional sample segments were added in Phase II to alleviate this problem.

In Phase II, estimates for all three components—area, yield, and production—were made for the first time. The LACIE estimates of wheat production were encouraging. An overall accuracy of 90/75 was achieved in the U.S. Great Plains (i.e., 90 percent of the time, the estimate was within ± 25 percent of the reference). For the winter wheat in the U.S. Southern Great Plains, the data indicated that the LACIE and

TABLE III.—Oklahoma Yield Model

Latent root variables selected	LACIE Phase III variables selected
October precipitation	August-February precipitation
March precipitation	March precipitation/evapotranspiration
May precipitation	May precipitation
	May precipitation squared
June precipitation	June precipitation
March temperature	—
June temperature	—
	May degree days above 90° F

ESCS estimates of wheat area were not significantly different. However, in some states, the LACIE estimate was higher than the ESCS estimate, and in other states, the LACIE estimate was lower than the ESCS estimate. The largest difference was in Oklahoma, where the LACIE estimate of wheat area was found to be less than the ESCS estimate. The investigation of 20 Oklahoma blind sites indicated that this difference was due to the mislabeling of wheat signatures as nonwheat because of early drought (fig. 19) and the grazing of wheat by cattle. Without this ground-truth inventory of about 700 square statute miles, the error source would never have been isolated, since no intensive test sites were located in Oklahoma and the LACIE crop estimation system had not previously estimated wheat acreage under drought conditions.

For spring wheat, the problem encountered in North Dakota in Phase I was solved. However, in Minnesota and Montana, the estimates of wheat area and production were low, which caused the LACIE estimates for the U.S. Northern Great Plains and the Great Plains as a whole to be significantly lower than the ESCS estimates. The principal problem in Montana was the misclassification of strip-fallow fields that were too narrow for Landsat resolution (fig. 20).

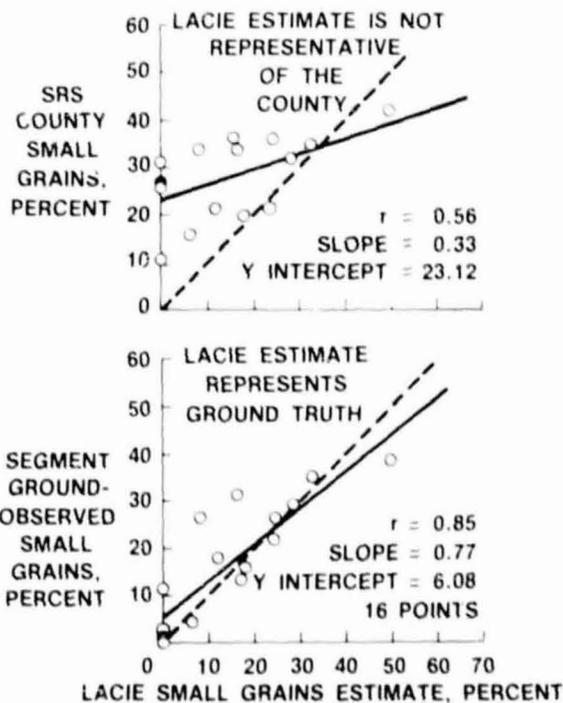
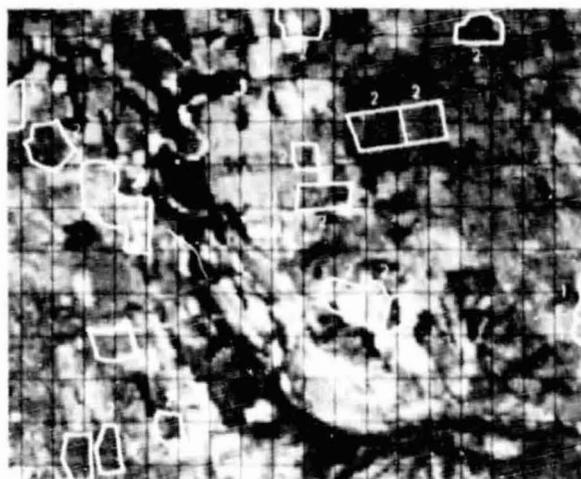


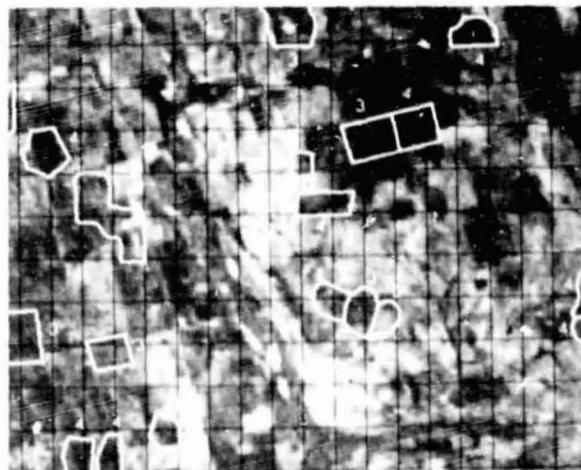
FIGURE 18.—Phase I example of sampling error found in North Dakota.

The use of the blind sites indicated that the underestimate of spring wheat area in Minnesota was the result of sampling error, caused by the use of 1969 data for the sampling, whereas an increase in wheat acreage occurred from 1969 to 1976. However, during this same crop year, this tendency to underestimate spring wheat area was not observed in the U.S.S.R., because the large number of sample segments (2000) placed in the U.S.S.R. gave a low sample error and the large fields in the U.S.S.R. were considerably easier to interpret than the small U.S. and Canadian spring wheat fields.

In Phase III, several steps were taken to solve the problems encountered in Phase II. The number of



NOVEMBER 24, 1975



APRIL 16, 1976

FIGURE 19.—Wheat signature variability due to drought, segment 1232, Kiowa, Oklahoma, 1976 crop year (number indicates biowindow).

sample segments in the U.S. Great Plains was increased to satisfy the required sampling accuracy, and a new multitemporal machine classification procedure was introduced.

As in Phases I and II, the final LACIE Phase III winter wheat production estimate for the USGP supported the LACIE accuracy goal. The LACIE estimates of USGP spring wheat production, however, were significantly different from the corresponding reference estimates. The LACIE 90/90 accuracy goal was not supported by the spring wheat production estimate primarily because of the underestimation of yield, although area also was significantly underestimated. As a result, the final LACIE total wheat USGP production estimate supported a 90/85 criterion, marginally missing the LACIE accuracy goal of 90/90.

As in Phase I and Phase II, the final LACIE winter wheat area estimate for the USGP was not significantly different (at the 10-percent level) from the corresponding USDA ESCS estimate. The final LACIE spring wheat area estimate for the USGP was significantly smaller than the corresponding ESCS estimate, but there was great improvement in the relative difference of this estimate over the corresponding Phase I and Phase II estimates. This improvement is attributed to the implementation of the new classification procedure, Procedure 1.

Based on the blind site ground-truth investigations (for 166 sample segments), the primary source of errors in classification (in both spring wheat and winter wheat) was found to be the mislabeling of wheat signatures as nonwheat because of (1) abnormal signature development caused by late planting, drought, grazing, crop rotation, plant variety, disease, and/or soil type; (2) inability to resolve small fields using Landsat imagery; and (3) lack of Landsat acquisition for both the postemergence stage and the tillering-to-heading stage. In addition to providing a good understanding of U.S. wheat labeling accuracies, the extensive blind site analysis effort added to the confidence in the U.S.S.R. classification accuracy since the small-grains fields in the U.S.S.R. are much larger and the field signatures appear more normal and homogeneous (figs. 21 and 22) than in the USGP (fig. 23).

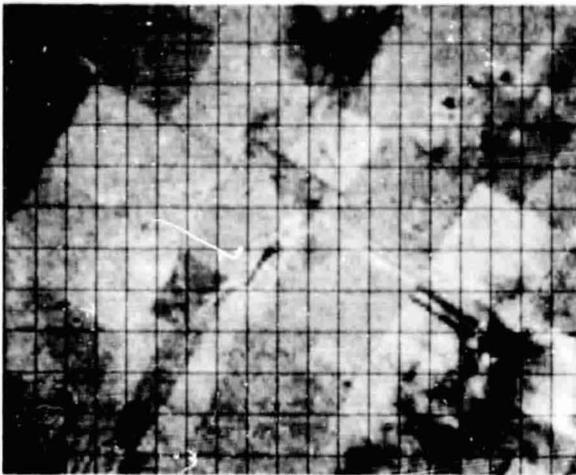
Unlike Phases I and II, the Phase III LACIE total wheat yield estimate was significantly different from the corresponding ESCS estimate in every month because of underestimates for both spring and winter wheat. The largest differences occurred in Oklahoma



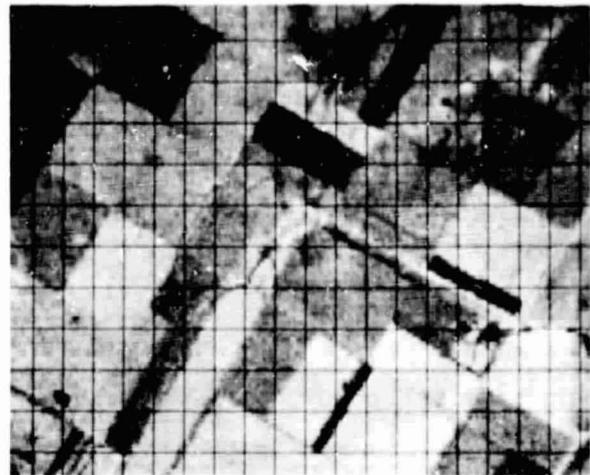
FIGURE 20.—Strip-fallow fields in Montana, August 8, 1977.

and Texas winter wheat yields and in Minnesota and Montana spring wheat yields. The spring wheat yield errors were due primarily to trend terms which failed to account for new varieties of wheat in Minnesota and for increased fertilizer usage in Montana during the past 5 years. The winter wheat yield errors were also due to trend terms which failed to account for more wheat acreage being fertilized in the last two decades in Texas and Oklahoma.

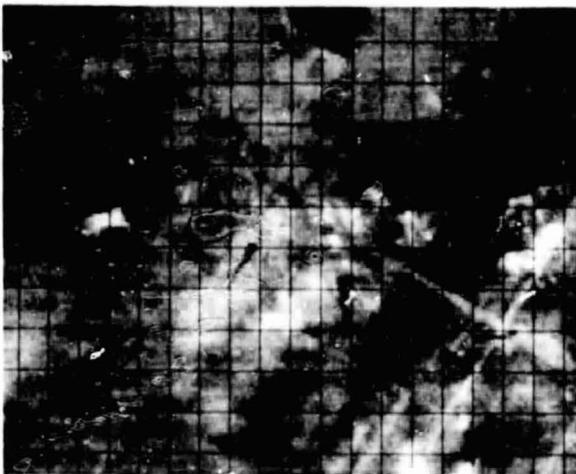
The results of LACIE Phase III production estimation indicated that the accuracy goal of 90/90 was achieved in the U.S.S.R., where the technology was able to identify the shortfall in the spring wheat crop 3 months before completion of harvest and achieved similar results in the winter wheat regions. The initial LACIE estimate in August was within 6 percent of the U.S.S.R. January 28 figure of 92 million metric tons, and the LACIE final modified estimate released on January 23 was within 1 percent. A detailed examination of the conditions which led to the U.S.S.R. shortfall in spring wheat production and the response observed in the LACIE models provided conclusive evidence that the LACIE forecast technology did indeed respond for good reason and in a timely fashion. Over most of the U.S.S.R. spring wheat regions, warmer than average temperatures predominated during the growing season. An investigation of the Landsat data and the yield model response at subregional levels indicated that the drought conditions were clearly observable in the Landsat data and that the yield models accurately responded by reducing yield estimates in the affected regions.



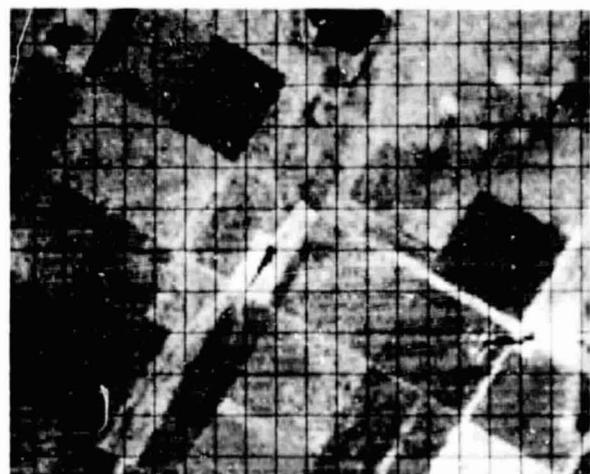
3 MAY 1976 - PLANTING



4 MAY 1976 - PLANTING



14 JULY 1976 - JOINTING TO HEADING



1 AUG 1976 - HEADING TO TURNING

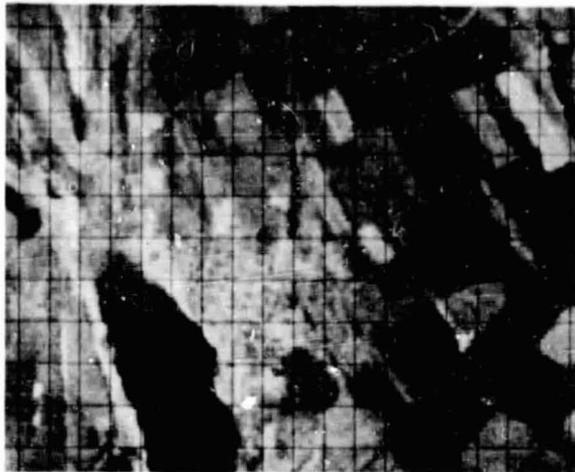
FIGURE 21.—Kustanay, U.S.S.R. segment 8224, spring wheat growth stages.

CONCLUSION

The Accuracy Assessment Subsystem has matured through the 3 years of LACIE. Early in LACIE, the small amount of ground truth that was available precluded accurate statistical estimates of sampling and classification accuracy. As the program culminated, Accuracy Assessment was not only able to meet this goal but was even able to evaluate component labeling errors, such as boundary effects, abnormal signatures, and lack of key Landsat acqui-

tions. Furthermore, the ground-truth data processing matured through LACIE from collecting data for one crop (small grains) to collecting, quality checking, and archiving data for all crops in a LACIE sample segment. This data collection not only assisted LACIE in determining causal factors but has provided an invaluable data set for new system development for other crops, such as corn and soybeans, and for the testing of new classification systems for Landsat data.

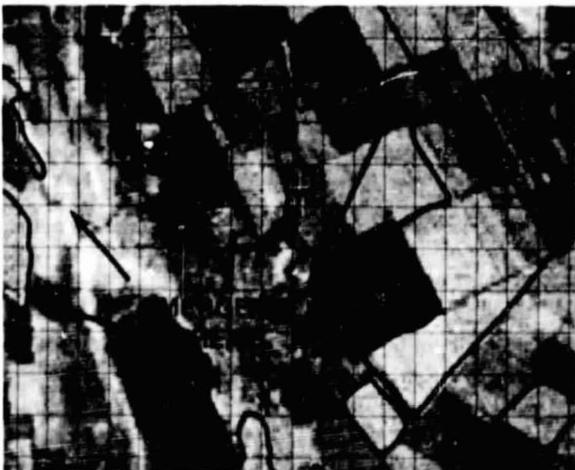
REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR



11 SEPT 1976 - PRE EMERGENCE



29 SEPT 1976 - PRE EMERGENCE



29 DEC 1976 - EMERGENT

FIGURE 22.—Odessa, U.S.S.R. segment 7265, winter wheat emergence.



FIGURE 23.—Very small winter wheat fields in LACIE segment 1503, Stanton, Nebraska, October 26, 1977.

REFERENCES

1. Remote Multispectral Sensing in Agriculture. Annual Report, Vol. 3, LARS, Purdue University, Sept. 1968.
2. Remote Multispectral Sensing in Agriculture. Annual Report, Vol. 4, LARS, Purdue University, Dec. 1970.
3. LACIE Phase III Interim Accuracy Assessment Plan. LACIE-00628, NASA Johnson Space Center, Houston, Texas, Mar. 1978.
4. Loe, D. L.; Hayenga, W. B.; and Ahlers, C. W.: "As-Built" Design Specification for PDP 11/45 Accuracy Assessment System Using Disk Data File. LEC-11881, Lockheed Electronics Company, Houston, Texas, 1978.
5. YES Phase I Yield Feasibility Report. LACIE-00439, NASA Johnson Space Center, Houston, Texas, Jan. 1977.

LACIE Applications Evaluation System Efficiency Report

Timothy T. White^a

SUMMARY

The LACIE Applications Evaluation System (AES) encountered significant increases in scope over the three phases of the project. The increases in scope and data load, as well as increased complexity and sophistication, were accommodated without significant increases in operational resources by developing and implementing a number of system efficiencies. In general, the operation of the AES which was expected to be a major implementation problem turned out to be a manageable one with data timeliness being the only major problem encountered. The timeliness of the data never met the expected goals because of (1) a highly fragmented system, in which the data accumulated large amounts of queue time waiting on delivery to the next function, and (2) the staffing for average loads, which resulted in backlogs during peak loads. With these exceptions, the very complicated LACIE system functioned in an increasingly efficient and productive manner in a state-of-the-art environment. The experience of operating a global remote-sensing inventory experiment produced valuable insight into the operation of a future system.

INTRODUCTION

The operation of the LACIE system required constant monitoring and management to keep all the disparate functions producing in such a way as to meet the operational goals of the project. This required the implementation of a number of monitoring and efficiency analysis tools. The efficiency monitoring function often afforded enough insight into the operation of the system to avert the disruption of the data flow caused by a situation that might arise in the

operation of a given element. This paper discusses the scope of the three phases of LACIE and the system efficiencies which had to be implemented to cope with the resulting Landsat data load. The methodologies used in system analysis, some of the specific data collected, and the inferences of these data and their implication on future systems are also discussed.

LACIE PHASE I

Phase I of LACIE was the most inefficient phase operationally because a major portion of the available resources was being utilized for system development. The scope of Phase I was very modest—692 segments, of which 411 were aggregable U.S. Great Plains segments and the remainder were exploratory sites distributed over the other LACIE countries and intensive study sites. The Landsat acquisitions utilized in Phase I were not as complete as those used in the later phases. The Landsat data were not collected in real time until after April 1, 1975. The data from the first of the crop year, September 1974 through March 1975, were selected manually from the available Landsat-1 acquisitions; and, typically, only one or two images were selected for each segment. Real-time data were obtained after April 1, 1975, from Landsat-2. This scheme resulted in about 2400 acquisitions reaching the NASA Johnson Space Center (JSC) for processing after the Landsat data had been screened for cloud cover and registered. The NASA Goddard Space Flight Center (GSFC), operating five shifts per week, produced 50 to 75 acquisitions per week and fell well behind the incoming data load; turnaround times were running typically 15 days from satellite acquisition to completion of GSFC processing. In June 1975, GSFC implemented a 10-shift-per-week operation; backlogs were eliminated, and turnaround times averaged 6 days.

^aNASA Johnson Space Center, Houston, Texas.

The significant increase in throughput from GSFC, averaging 170 acquisitions per week during the peak months of June and July 1975, inundated the rest of the AES, especially the analyst capabilities, for much of the remainder of Phase I.

The performance of the preprocessing functions at JSC improved steadily during Phase I as procedures were refined, reaching a nominal turnaround time of 5 days from GSFC completion until receipt by the LACIE Physical Data Library (LPDL) at JSC (2 of the 5 days were allocated for air transportation). The major bottleneck at JSC was in the classification of the data. The Classification and Mensuration Subsystem (CAMS) analysts were required to classify the first acquisition for each of four predefined biowindows. This impacted processing considerably in that the acquisition required for processing might not have had an adequate signature or might have contained confusing information which resulted in much rework before satisfactory results could be obtained. The interpretation of the imagery and the definition of training fields was done by 14 image analysts. The segments were subsequently classified by 29 data processing analysts, leading to very inefficient communications. In addition, unproved signature extension techniques were being attempted, along with the implementation of a newly developed batch processing system, which further inhibited the data flow.

These situations caused a prolonged analyst contact time of about 12 hours per classification and a significant amount of rework of 200 to 300 percent (typically two attempts at batch processing and one at interactive rework). Overall, the Phase I CAMS processing produced only about 1100 classifications (approximately 1.6 estimates per segment; see table I). The total man-hours required during Phase I to process the acquisitions classified for each segment ordered was a rather lengthy 25 analyst hours, which includes all rework and the processing of subsequent acquisitions.

LACIE PHASE II

The scope of Phase II was increased significantly by the addition of 680 U.S.S.R. segments and 280 Canadian segments to the Phase I scope, thus raising the total to almost 1700 segments, 2.5 times that of Phase I. The Phase II data collection period was extended, and data were gathered over the entire crop season in real time. This produced more than 9000

TABLE I.—Phase I Landsat Processing Summary

Item	Number	Analyst time
Segments	691	
Acquisitions at JSC	2400	
Acquisitions not processed	1300	
Machine-processed acquisitions	1100	12 hr
Batch-reworked acquisitions	540	4 hr
Interactive-reworked acquisitions	740	3 hr
Time per segment		25 hr
Average time per acquisition		16 hr
Estimates per segment	1.6	
Average throughput time		40 days

acquisitions, almost 4 times as many as in Phase I. The number of Phase II acquisitions per segment increased dramatically from the Phase I value of 3.5 to 5.4, producing more data for the system to contend with but also more information for the analyst to utilize in decisionmaking. GSFC implemented a nominal 10-shift/week operating schedule in Phase II. The peak processing load in the summer brought backlogs of up to a week to GSFC and required the system to operate at maximum capacity in June, July, and August 1976 (17 shifts per week were required to handle the workload). This increased the average weekly output for the months of June and July to 370 acquisitions, which was twice as great as the Phase I peak output.

In Phase I, it was noticed that although initial acquisitions were received for some segments, subsequent acquisitions were not obtained. This problem was diagnosed as being caused by bad reference scenes. A data management system was implemented at GSFC to identify these segments primarily by noting the ones for which fewer than the average acquisitions were received. The bad reference scenes were usually updated toward the end of a phase, allowing the acquisition efficiency on a per-segment basis to improve eventually if the segment was kept for a subsequent phase.

The data preprocessing at JSC remained unchanged in Phase II except for two hardware augmentations. A direct data link between GSFC and JSC was implemented, which eliminated the occasional "misdirection" applied to the data when

shipped via commercial airlines. A significant improvement in the classification time on the LACIE/Earth Resources Interactive Processing System (LACIE/ERIPS) resulted from the implementation of a parallel processor which reduced the four-channel classification time from 6 to 3 minutes. This reduced the required computer operations time in Phase II and allowed a significant increase in computational complexity to be added to the system in Phase III.

In Phase II, about one-half of the acquired Landsat data were unprocessable. As shown in table II, 10 percent of the Phase II acquisitions were of poor quality (caused by the presence of haze, snow, clouds, etc.), 16 percent had preemergence or dormancy (no wheat signatures), 19 percent were next-day passes acquired in overlap requires, and 2 percent were acquired for segments located in non-agricultural areas. These nonagricultural segments were used in the aggregation but were moved to agricultural areas for Phase III.

A number of changes were made in the CAMS analyst operations to improve operating efficiency. The image analysts and data processing analysts were integrated into teams to improve communication and feedback. Thirty-six team equivalents resulted, including about a dozen analysts who performed both functions. The procedure of processing only the first acquisition in a biowindow (one of the four LACIE data collection windows) was replaced with one that required the analyst to analyze and process every acquisition. However, the analysts were not required to machine process every acquisition. If a small percentage of wheat existed in the scene (less than 500 resolution or picture elements), the picture elements (pixels) would be counted by hand. If the analyst could determine that the current acquisition under examination had not changed significantly from a previous estimate for the segment, a "no change" would be submitted and the previous result would be used in the aggregation. Of the acquisitions obtained, 17 percent were machine classified, 9 percent were hand counted, and 27 percent were determined to have no significant change from a previous estimate.

A significant improvement in the total amount of time an analyst spent in processing a segment—from 25 hours in Phase I to 11 hours in Phase II—was the result of a number of factors. A tenfold reduction in rework resulted from the improved communication between the interpretation and the processing of an acquisition, an improved operating system, and a

TABLE II.—Phase II Landsat Processing Summary

<i>Item</i>	<i>Number</i>	<i>Analyst time</i>
Segments	1683	
Acquisitions at JSC	9150	
Acquisitions screened but not processed due to—		
Poor image quality	900	
Preemergence or dormancy	1454	
Next-day pass	1738	
Nonagricultural	183	
Proportion estimates		
Hand-counted acquisitions	823	2.5 hr
Machine-processed acquisitions	1555	7 hr
Acquisitions reworked by machine processing	389	2 hr
No significant change	2470	2 hr
Total time per segment		11 hr
Average time for machine processing		6 hr
Average time per estimate		4 hr
Estimates per segment	2.9	
Average throughput time		33 days

good quality assurance program instituted during the phase. The analyst contact time for machine classification of a segment was reduced to 6 hours by the implementation of the team approach, improved systems and procedures, and experience. The use of the no-change criterion allowed the analyst to provide an estimate without spending the time to machine process the segment. This reduced the average analyst time involved in producing an estimate to 4 hours, with an average of three estimates being produced for each segment in Phase II.

LACIE PHASE III

The Phase III scope was much larger than Phase II with about 17 000 acquisitions for 3000 segments—almost 6 acquisitions per segment compared to 5.4 in

Phase II and 3.5 in Phase I. The biowindows were lengthened again in Phase III, opening earlier to obtain the acquisitions during seedbed preparation and remaining open longer during the winter dormancy period. The GSFC improved its throughput by about 30 percent by eliminating the visual screening step in its procedure, by reducing the scene extraction constraints, and by relaxing some quality assurance constraints. With this increased efficiency and operating 17 shifts per week, the GSFC peak output in June and July 1977 averaged 610 acquisitions per week—a 65-percent increase over Phase II. Even with this increased capacity, a 2-week backlog (21-day turn-around time) was encountered at GSFC during the peak processing period.

The preprocessing steps at JSC remained the same, although some support resources were increased. This was required because of the increased data load and because a number of new products were added to the analyst repertoire, such as cluster maps, dot overlays, bias-correction information, green numbers, spectral plots, feature selection, and trajectory plots. The new products increased the computer time for a four-channel classification back to the Phase I level of 5 to 6 minutes. Late in Phase III, 8- and 16-channel multitemporal classifications were taking as long as 10 to 12 minutes. To a large extent, this was caused by the feature selection portion of the processing logic.

The classification function again underwent a number of changes in Phase III. The team concept was replaced with about 30 individual analysts who were regionalized into two areas. One-half of the analysts were involved in processing about 600 U.S. segments, and the other half processed 2000 U.S.S.R. segments. The processing strategy also was changed. All proportion estimates were to be the result of multitemporal machine classifications, requiring at least two acquisitions, with initial processing being deferred until emergence was detected. A priority system was employed to process the segments which were needed most for an upcoming aggregation. This Phase III processing strategy produced 5000 machine processings. This represented about 1.7 classifications for each segment ordered in Phase III, which was similar to the 1.6 estimates per segment in Phase I. Although it was somewhat less than the 2.9 estimates encountered in Phase II including the no-change estimates, it was slightly more than the 1.4 estimates per segment in Phase II excluding the no-change estimates.

The Phase III processing results are shown in table III. Some differences between Phase II and Phase III are noticeable. The percent of the total acquisitions "screened and not processed" increased significantly in Phase III. This occurred for two reasons: (1) some Phase III end-of-season acquisitions, which were not processed due to resource limitations, were included in the "not processed" category; and (2) backlogs and priority processing in Phase III often caused several unprocessed acquisitions to be encountered in a packet. Usually only the one expected to produce the best estimate was classified and the remaining acquisitions were stashed as a multiple acquisition in addition to any next-day passes encountered. This increase in unprocessed acquisitions corresponds to a decrease in the percentage of total acquisitions for which estimates were passed in Phase III (30 percent) compared to Phase II (53 percent). It can also be noted that the problem encountered in Phase II of having segments located in non-agricultural areas was essentially eliminated in Phase III.

TABLE III.—Phase III Landsat Processing Summary

Item	Number	Analyst time
Segments	2 953	
Acquisitions at JSC	16 640	
Acquisitions screened but not processed due to—		
Not processed	4 041	
Preemergence or dormancy	3 213	
Multiple acquisitions	4 382	
Nonagricultural	16	
Proportion estimates		
Machine classified	4 988	3 hr
Rework	< 10 percent	
Total time per segment		5 hr
Average time per acquisition		3 hr
Estimates per segment	1.7	
Average throughput time		55 days

The implementation of Procedure 1 (P-1) in Phase III made the analyst's job easier by eliminating some mechanical tasks previously required and by providing a number of new products to aid in the analysis. The average contact time was reduced significantly in Phase III to an average of 3 hours per segment, although other portions of the system were burdened somewhat with the increased number of products. Another efficiency of P-1 was the use of the stratified areal estimation procedure, which allowed the analyst to correct for misclassification without reworking the segment. The rework rate in Phase III was reduced to a negligible level, with the rework encountered consisting primarily of batch job format errors.

DEMONSTRATION OF OPERATIONAL TECHNOLOGY

Because of the scope of LACIE and the magnitude of the data, the project encountered a number of new situations related to the development of a fully operational remote-sensing system. The following sections describe some of these situations, the efficiency monitoring associated with them, the importance of these kinds of data, and their relationship to future systems.

Data Acquisition

The ordering, acquiring, quality screening, extracting, and registration of Landsat data results in an unpredictable quantity of data because of the difficulty in forecasting weather conditions and the problems associated with registration. However, it is necessary to estimate an approximate data load as a function of time in order to schedule resources.

A computer prediction model was developed to estimate the number of acquisitions to be expected over a given area for a given time, considering the satellite parameters and historical weather information. The Landsat acquisition frequency for each defined segment is calculated using simple spherical-vector equations and data such as orbit inclination, repeat cycle in days and revolutions, and image frame size. The latitude and longitude of the target of interest and the period of time over which the data are to be calculated are input. The attrition of the satellite acquisitions, because of cloud cover, was

determined using the cloud cover data obtained from the U.S. Air Force Environmental Technical Applications Center (USAF-ETAC). The climatological average number of days per month with less than 25 to 30 percent (depending on available data) cloud cover, closest to the Landsat pass time, was used.

Although relatively simple in structure and operation, the model has proven quite satisfactory for estimating the data to be encountered in LACIE—with a couple of caveats. First, the area of coverage must be rather large (e.g., encompassing several continents). This provides an averaging effect so that the acquired data over the short period of a month or two will be accurately estimated. If the area for which predictions are to be made is a country or a small portion of a continent, the predictions tend to be reasonable only over a minimum period of 6 to 12 months. The model does not predict with any accuracy the coverage to be expected over a small area (e.g., a state) for a short period of time (e.g., a month) because of the variability of local weather.

The cloud cover/acquisition prediction model output is reduced to compensate for attrition in the system because of data quality and because of acquisitions failing registration. This affects only about 15 percent of the data acquired by the satellite. The GSFC processing routinely resulted in the following output: 50 percent lost to clouds and snow, 10 percent lost to registration, 5 percent lost to quality control, and 30 percent being sent to JSC for LACIE processing.

The 10 percent of the Landsat acquisitions lost because of failure to register to the reference scene might be considered insignificant. In some cases, it was, but many of the misregistrations were caused by bad reference scenes. When this occurred, a segment could not be used for aggregation, since one or two acquisitions over a growing period seldom allowed the analyst to produce a reliable estimate. The loss of aggregable segments will cause the sampling bias to increase or require additional samples to be ordered in subsequent years to compensate for these problems.

An analysis of the reference scene problem was performed to determine what did or did not make a good reference scene, and no definitive answer was obtained. Better reference scenes seemed to be obtained early in the growing season, but no criterion was established for LACIE. This problem was handled operationally by monitoring the acquisition history of each segment and flagging the abnormal

ones. This generally occurred after much of the growing season had been completed so that the changing of a reference scene generally helped only in the next crop year, assuming the scene was not moved for the next year. Hopefully, the new full-frame registration system being implemented at GSFC, which is based on ground-control points instead of reference scenes, will eliminate this kind of problem from future large-scale remote-sensing inventory systems.

One additional issue is related to the acquisition of Landsat data; i.e., what are the right data to order and/or extract? Initial LACIE plans called for only the first acquisition in a biowindow to be analyzed. It was quickly realized that crop growth stages could not be predicted with any accuracy in real time, and the set of acquisitions needed by the analyst spanned the entire growing season plus a month or so before and after. For areas where winter crops are grown, data were collected year-round; for spring crops, data were collected for most of the year. It was shown in Phase II and Phase III that the early-season or pre-season acquisitions were the most important in observing crop rotation, seedbed preparation, and initial crop emergence. In future programs, some changes to data collection might be considered. First, the data collection might be reoriented to earlier acquisition dates to focus more attention on the transition period between crops. This would provide insight into planting dates, changes from the last to the current year, and initial growth stages of the current year's crop. This kind of strategy might entail considerable overlap between two crops, whereas the current LACIE approach does not allow much overlap due to data base storage constraints. Secondly, since the early-season data are so important—almost to the point that, if they are not obtained for a given segment, the segment is not usable—a more dense early-season segment population might be planned, and those segments which do not have an adequate early-season acquisition history could be eliminated from the system. This approach would be implemented easily for winter crops; however, because of the shortness of the growing season, it is questionable whether it would be practical for spring crops.

Analyst Contact Time

The need for collecting detailed analyst time as a function of task was established early in LACIE. Midway through Phase II and continued through

Phase III, a system was implemented whereby each analyst would record the start and stop time for each major step in the analysis procedure. These were accumulated and analyzed monthly. A summary of a portion of the data collected is shown in table IV. The improvement through the implementation of P-1 is clearly shown. The reduction in interpretation time resulted in the ease of labeling prelocated dots compared to selecting and identifying fields and in the elimination of the need to extract, reformat, and verify field coordinates. These improvements saved the analyst almost 4 hours of time. However, P-1 postclassification results included cluster maps, spectral plots, etc., which required slightly more evaluation time and time to perform the stratified areal estimate. The net decrease in analyst contact time was 3.3 hours. This saving may have been offset slightly by the time used by the additional clerical help needed to assemble, status, and distribute the additional products. However, since analyst skills are critical in such a project, any improvement in analyst contact time is a significant asset to the project.

Two other important observations can be made from the analyst contact time and should be kept in mind when future project planning is considered. The amount of analyst contact time required to initially process a segment depends on the conditions of that processing; i.e., how many and which acquisitions are available for analysis and whether the segment had been processed in a previous year. In some situations, the analyst may take twice the average time to initially process a segment. It may be necessary for the analyst to become familiar with historical data for the area, to verify the location of the segment, to establish the planting date for crop calendar

TABLE IV.—Analyst Contact Time^a

<i>Phase II Training Field Procedure</i>		<i>Phase III Procedure I</i>	
<i>Task</i>	<i>Time, hr</i>	<i>Task</i>	<i>Time, hr</i>
Interpretation	4.3	Interpretation	1.9
Del Foster	.6		
Reformatting	.6		
Verification	.3		
Evaluation	.7	Evaluation	1.3
Total	6.5	Total	3.2

usage, or to perform a variety of housekeeping activities such as scribing and taping imagery. However, after the initial processing, a general reduction in contact time of 20 percent over a 3- or 4-month period would not be unusual because the analyst would have to make only minor changes to previously labeled signatures, as opposed to labeling the entire scene.

Secondly, analyst contact time (for P-1) is related to the field size and/or signature complexity of the scene. For example, the analyst contact time in Phase III for U.S.S.R. spring wheat segments, which typically consisted of very large fields, required only 2 hours as opposed to spring wheat segments in the U.S. Great Plains (typically smaller fields), where the analyst contact time averaged 4 hours. This variation from country to country will have a bearing on how analyst resources are allocated. It is also important to

future system evaluations for which testing is performed in one country, inasmuch as the results may not be directly applicable to other areas. Furthermore, since most of the areas to be encountered in a global inventory program tend to be as complex as U.S. agriculture or more complex, such as India and China, there may be a need for additional modifications to P-1 to reduce contact time.

System Throughput Time

The throughput goal of LACIE was 14 days from Landsat acquisition to having a proportion estimate available for aggregation. This was predicated on an around-the-clock operation (24 hours per day, 7 days per week). This goal was met only in isolated cases when data were processed through the system in the

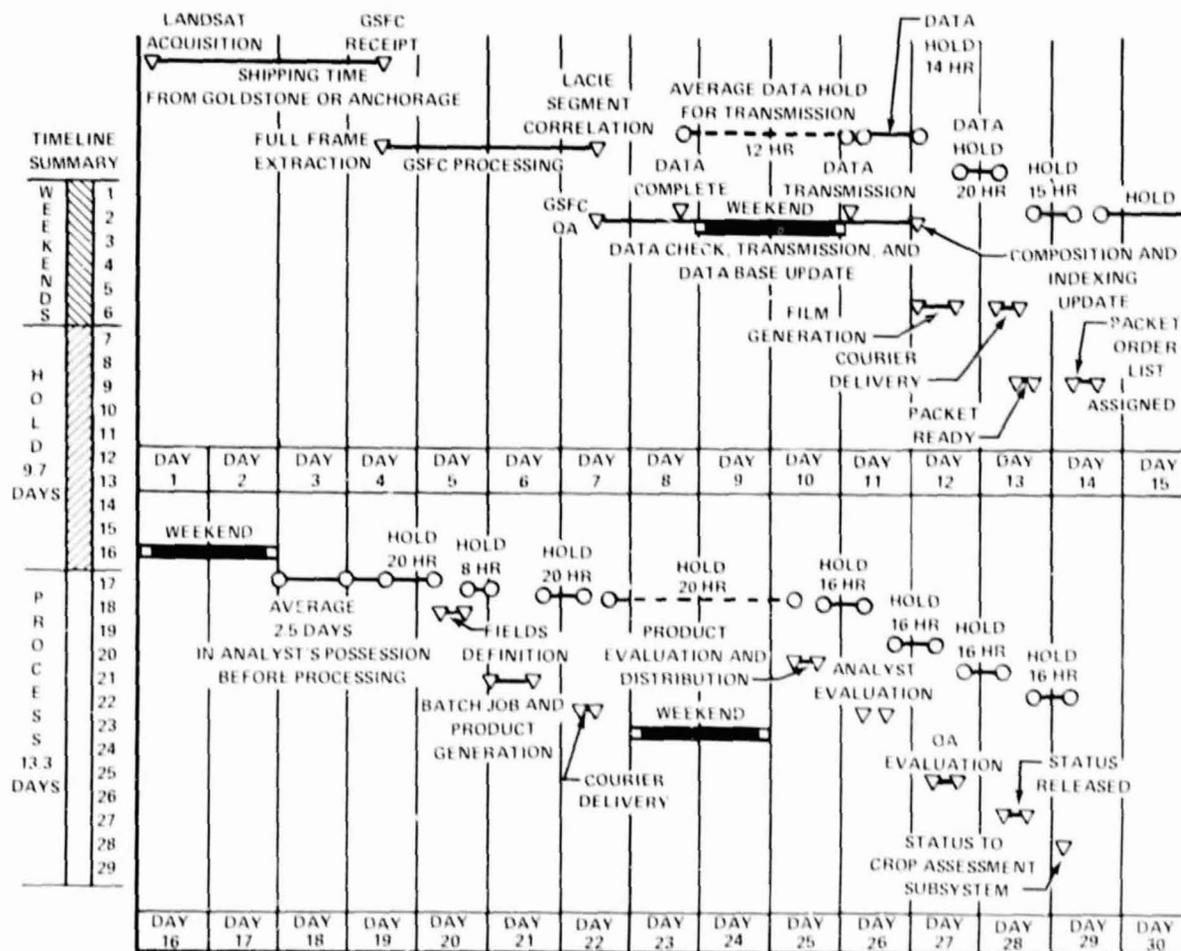


FIGURE 1.—LACIE data flow/nominal throughput time.

optimum manner. The vast majority of the data took about 29 days to flow through the system. This was due to several causes. First, the LACIE system was fragmented, with most major components being located in geographically separated areas. In addition, manual methods were used to transfer data from one component to another (courier, air transportation, mail). A second problem in attaining the throughput time was that a number of the processing components operated only 8 hours per day, 5 days per week. Figure 1 shows the typical flow of data in the LACIE system. As indicated, 6 of the 29 days were lost due to weekends and an additional 10 days resulted from overnight holds. The resulting 13-day in-process time indicates that, if a 3-shift-per-day 7-day-per-week operation could be implemented, the 14-day turnaround goal could be attained.

The nominal 29-day turnaround time was exceeded during the peak processing months (May

through August) of each LACIE phase. As shown in figure 2, the data acquisition rate for the LACIE project was not uniform, and a large influx of data was encountered each summer. Since most of the LACIE components were staffed for the average loads, the peak loads caused major backlogs at GSFC and in the classification subsystem. These backlogs caused the average turnaround time to reach 40 days in Phase I, 33 days in Phase II, and 50 days in Phase III (because of the enormous peak encountered in Phase III). One major reason that the Phase II turnaround time was short compared to Phases I and III was because of the implementation of the no-change and hand-count procedures, which produced almost 70 percent of all the aggregable estimates in Phase II. These procedures did not require the additional 7 to 10 days necessary to submit a batch job and receive and evaluate the results, thereby reducing the average turnaround time considerably.

Several situations encountered in LACIE will likely be encountered in future systems. First of all, the reliability of each of the various system components was reasonably good. However, with so many components in the system, the chances were that at any given time one of them would be inoperable, impeding the flow of data. Secondly, if the data processing cycle is greater than the satellite data acquisition cycle, incoming imagery will very likely have to be placed on hold while previous acquisitions are being completed, thereby impacting the turnaround time of the incoming acquisitions. Finally, unless resources are dedicated around the clock and in sufficient quantity to meet peak load demands in a fully operational remote-sensing data processing system, a project having sizeable scope will most likely not maintain a real-time operation during peak acquisition periods.

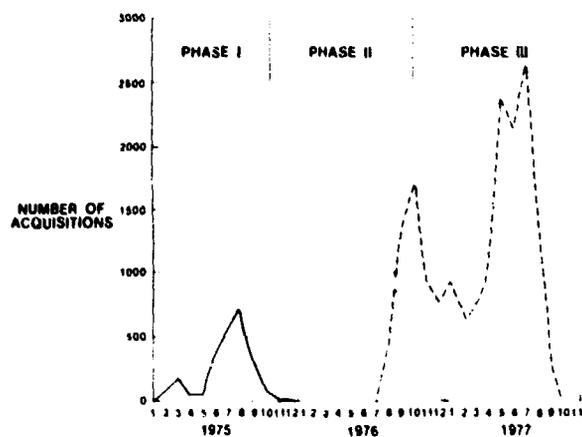


FIGURE 2.—LACIE acquisitions expected at JSC.

Data Processing Systems Design

FOREWORD

In general, the remote-sensing data processing facilities and software available prior to LACIE were primarily suited to the support of essentially "small" research users in laboratory environments. In distinction, the LACIE requirements for a quasi-operational high-throughput and high-volume data system represented a considerable departure from the capabilities of these predecessor installations, both in capacity and in organization. The objectives of this session are to review those key conceptual and design issues identified and addressed in the establishment of an effective and economical LACIE data processing system, and to extrapolate from this experience toward the computational support of analogous future programs. This is accomplished in the following papers through focused discussions of specific elements of several critical subsystems rather than an attempt to be exhaustive or complete in the presentation of all system elements. The evolution of the addressed elements from the predecessor installations (where they existed) to their ultimate form in the Applications Evaluation System (AES) is traced. The final papers in the session draw inferences for future study and development. The reader is referred to two papers in the "Proceedings of the Plenary Session" for the context in which the work described in the current session has been performed; viz, "The LACIE Applications Evaluation System: A Design Overview" and "Data Processing Systems in Support of LACIE and Future Agricultural Research Programs." Additional descriptive papers pertinent to the operations of a number of these subsystems can be found in the System Implementation and Operations Session elsewhere in this document.

Several of the papers in this session pertain to the development of the Earth Resources Interactive

Processing System (ERIPS), which was used as the principal computational vehicle for the Classification and Mensuration Subsystem (CAMS) of the AES. The ERIPS was a representative pre-LACIE capability designated as a facility to be upgraded for LACIE support; many elements of the original design of this system were subject to revision for satisfactory LACIE operations. The introductory paper "LACIE/ERIPS Software System Summary" provides the necessary background for the other papers by outlining the historical development of the total system. The paper entitled "The LACIE Data Bases: Design Considerations" discusses the foundations and behavior of the ERIPS-related mass disk data base on which all crop-year imagery and much related ancillary data were maintained. The paper entitled "Man-Machine Interfaces in LACIE/ERIPS" treats the sometimes difficult conversion from the research-oriented interactive environment to the batch production requirements of LACIE. Finally, the paper entitled "Very High Speed Processing: Applicability of Peripheral Devices to Pixel-Dependent Tasks" outlines the solution to the critical problem of processing speed characteristic of image analysis as implemented in the ERIPS.

The two papers entitled "Cartography: LACIE's Spatial Processor" and "Considerations for Design of Future Research and Development Interactive Image Analysis Systems" specify other significant areas that require future work for the establishment of a satisfactory crop inventory system.

The final papers in the session, "A Look at Computer System Selection Criteria" and "Cost and Performance Characteristics of Data System Configurations for Processing Remotely Sensed Data," assess requirements for subsequent inventory systems and techniques for evaluating and satisfying these requirements in computing organizations.

Cartography: LACIE's Spatial Processor

M. L. Rader^a and R. R. Vela^a

ABSTRACT

The Cartographic Laboratory of the NASA Johnson Space Center Earth Observations Division is responsible for satisfying the spatial processing needs of LACIE. These needs include locating agricultural test sites and registering ground-truth data to Landsat imagery. This paper discusses the technical aspects of the LACIE cartographic support, the unique need for cartography in satellite crop surveys, and proposed improvements which would enhance the cartographic support of future programs.

graphic imagery. Because electronic images such as those from Landsat are optimally processed in the computer, the need for some of the elaborate cartographic electro-optical equipment was eliminated; however, the burden on the NASA computer facilities was increased.

Notwithstanding these problems, the Cartographic Laboratory made significant contributions to LACIE, including test site location and the measurement of LACIE performance. This paper will outline these contributions and explore potential areas for future contributions.

TRANSITION FROM PHOTOGRAPHIC TO DIGITAL IMAGE PROCESSING

From the outset, LACIE was constrained to use Apollo resources that had application to remote sensing. The cartographic capability was one resource which did have significant commonality with remote sensing; it was therefore merged into the Earth Observations Division. This transition into remote sensing created several problems for the Cartographic Laboratory. Many skills needed for traditional photographic mapping were unnecessary in digital image processing. The concept of an electronic resolution element (a pixel) was foreign to the cartographer who was accustomed to continuous-tone photographs. Some cartographers did have experience in automated data processing; however, computer processing of digital images was significantly different from Apollo computer cartography. Image processing was a problem of processing large data sets, whereas mapping was primarily a problem of programming equations.

Another problem was the obsolescence of the electro-optical equipment used in processing photo-

THE CARTOGRAPHIC ROLE IN LACIE

The LACIE process flow (fig. 1) involves three primary tasks: (1) test site selection, (2) classification and yield computation, and (3) performance measurement (accuracy assessment) of the classification and yield computations. The Cartographic Laboratory has primarily supported LACIE operational tasks, including test site selection and performance measurement; it has also contributed to supporting research, including that in yield estimation.

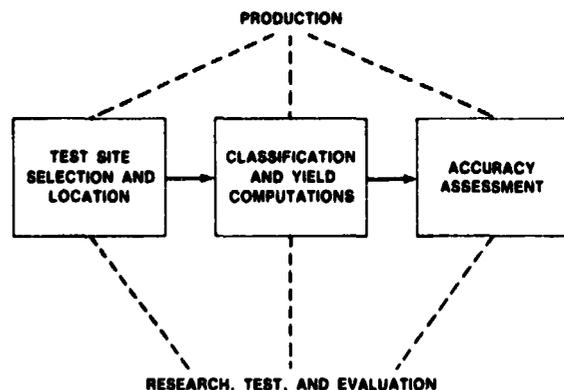


FIGURE 1.—LACIE process flow.

^aLockheed Electronics Company, Houston, Texas.

Test Site Selection

To measure world wheat production, it was necessary to select a set of statistically significant test sites for the world's wheat-growing regions. The Cartographic Laboratory performed a significant part of this task. The first step was to delineate farming regions on 1:1 000 000-scale Operational Navigation Charts (ONC's) over major wheat-growing regions such as the U.S.S.R., Canada, and the United States. These farm regions were delineated from mosaics of Landsat full-frame film products which were at the ONC 1:1 000 000 scale.

Crop Reporting District (CRD) boundaries were also transferred to the ONC map base. CRD's are the units for which agricultural statistics are compiled, such as the county system in the United States where a county agent gathers and reports crop data to the U.S. Department of Agriculture (USDA). Foreign countries have similar systems, the boundaries of which were also transferred to the ONC map base.

These data, along with certain meteorological and soils data, were then used to determine the strategic location of test sites in the wheat-growing regions. The Cartographic Laboratory located the center of each test site, marked it on the ONC, and interpolated the latitude and longitude of the site center. The latitude and longitude center was then used by the NASA Goddard Space Flight Center (GSFC) to strip out a 5- by 6-nautical-mile Landsat image for the test site.

Measuring LACIE Performance

To determine the accuracy of the LACIE wheat production computations, it was necessary to select a limited number of the 5- by 6-nautical-mile LACIE test sites where ground truth could be gathered for comparison with the LACIE results. These special ground-truth sites were designated "blind" sites and were located throughout the United States and Canada, where ground truth could be obtained without diplomatic problems. The ground truth is actually a set of aerial photographs annotated in the field by USDA Agricultural Stabilization and Conservation Service (ASCS) agents as to the crop or cover types for each agricultural field in the site. An example of an annotated aerial photograph is shown in figure 2.

The LACIE was divided into three crop years—Phase I, Phase II, and Phase III. The Phase I and II blind sites were processed differently from the Phase

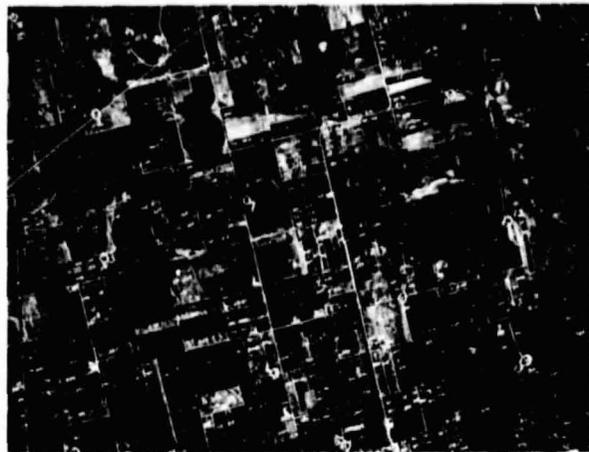


FIGURE 2.—Example of an annotated aerial photograph.

III sites. The wheat area was measured in square inches on an X and Y measuring table on the ground-truthed photographs, which were unrectified but printed at an approximate scale of 1:24 000. The wheat area was divided by the total blind site area, thereby giving a percentage of wheat for the blind site. This percentage of wheat was then used to check the percentage of wheat computed by the LACIE system. However, the error sources and magnitudes were unknown because the photographs were unrectified. This lack of rectification could significantly affect the wheat percentages. A rectified photograph is one in which the geometric distortions caused by the aircraft pitch and roll have been removed. It is printed at a known scale. Figure 3 is an example of

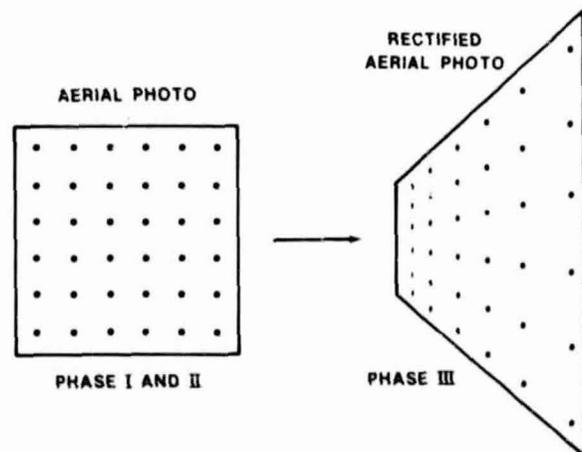


FIGURE 3.—Example comparison of rectified and unrectified photographs.

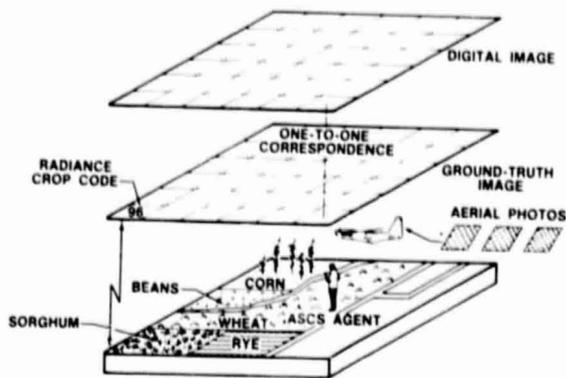


FIGURE 4.—Conversion of ground truth to an artificial digital image.

how an unrectified photograph might appear with respect to a rectified photograph.

The Phase III blind sites have been processed in a much more rigorous manner than those in Phases I and II. The Phase III ground truth is converted to pixel-level ground truth; i.e., each pixel has a crop code assigned to it based on its ground-truth crop-cover type. Figure 4 illustrates this process, which begins by obtaining aerial photography of the site. The aerial photographs are enlarged to a 20- by 20-inch format (approximate scale of 1:24 000). The enlargements are carried to the test site by a county agent, who annotates the crop type of each agricultural field on the photographs. These data are sent to the Cartographic Laboratory, where the fields are delineated as polygons, the polygon vertices measured, and the polygons registered and converted to Landsat-type pixels. The radiance levels assigned

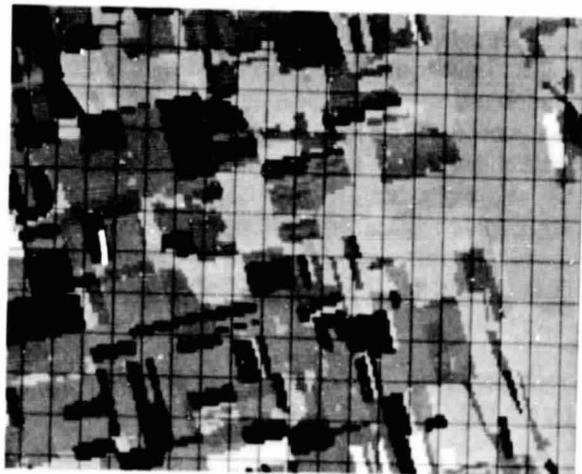


FIGURE 5.—Film image of digital ground truth data.

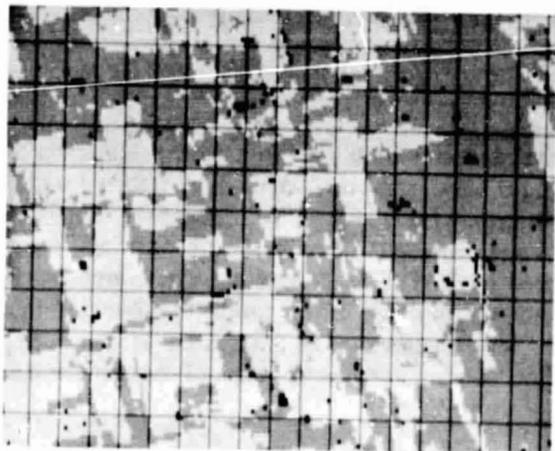
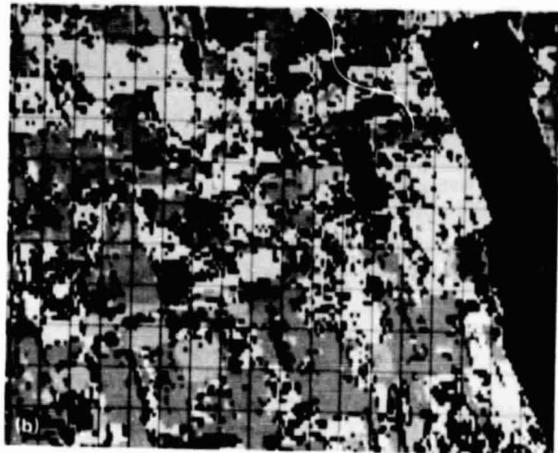
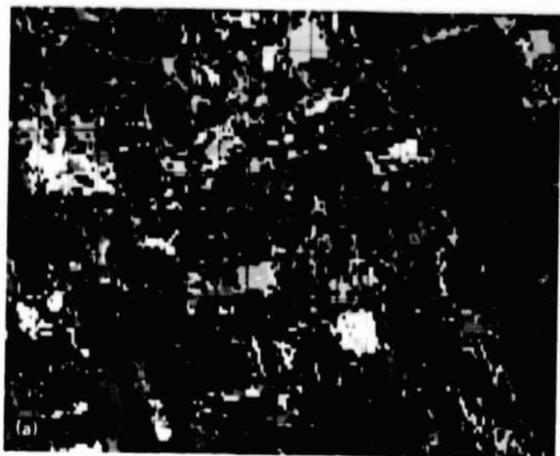


FIGURE 6.—LACIE classification and cluster maps. (a) Unconditional cluster map before assignment of clusters to classes; August 17, 1977. (b) Conditional cluster map: black = threshold, DO, DU; yellow = nonspring small grains; green = spring small grains; other = conditional clusters; August 17, 1977. (c) Classification map: black = threshold; green = spring small grains; orange = nonspring small grains; July 29, 1977.

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

TABLE I.—Approved Symbol List

<i>Symbol</i>	<i>Description</i>	<i>Gray-scale level</i>	<i>Harvested</i>	<i>Abandoned</i>	<i>Strip fallow</i>	<i>Strip fallow, abandoned</i>	<i>Strip fallow, harvested</i>
A	Alfalfa	90	115	140	165	190	215
B	Barley	101	126	151	176	201	226
BN	Beans	91	116	141	166	191	216
C	Corn	92	117	142	167	192	217
CN	Cotton	111	136	161	186	211	236
FX	Flax	103	128	153	178	203	228
G	Grass	105	130	155	180	205	230
H	Hay	106	131	156	181	206	231
I/CC	Idle cover crop	252	—	—	—	—	—
I/CS	Idle cropland stubble	251	—	—	—	—	—
I/F	Idle cropland fallow	254	—	—	—	—	—
I/RE	Idle cropland residue	253	—	—	—	—	—
M	Millet	112	137	167	187	212	237
MT	Mountains	241	—	—	—	—	—
NA	Non-Ag	242	—	—	—	—	—
O	Oats	104	129	154	179	204	229
P	Pasture	107	132	157	183	207	232
PF	Problem field	80	—	—	—	—	—
R	Rye	102	127	152	177	202	227
SB	Sugar beets	98	123	148	173	198	223
SF	Safflower	93	118	143	168	193	218
SG	Sudan grass	95	120	145	170	195	220
SR	Sorghum	96	121	146	170	196	221
SU	Sunflower	94	119	144	169	194	219
SW	Spring wheat	100	125	150	175	200	225
SY	Soybeans	97	122	147	172	197	222
T	Trees	108	133	158	183	208	233
TR	Triticale	109	134	159	184	209	234
VW	Voluntary wheat	110	135	160	185	210	235
W	Winter wheat	99	124	149	174	199	224
.	Water	240	—	—	—	—	—
X	Homestead	250	—	—	—	—	—

to the pixels within a polygon are the numerical crop codes assigned to each cover type (tables I and II). Thus, one has a ground-truth image which has one-to-one correspondence with the Landsat imagery used by LACIE in computing wheat area for that site. A film image of this product is shown in figure 5. This image can be compared to the LACIE classification and cluster maps of the same test site shown in figure 6.

Other Support

The Cartographic Laboratory has also provided many other support items to LACIE. The LACIE Research, Test, and Evaluation (RT&E) group has a set of intensive research sites. Designated Intensive Test Sites (ITS's), these sites have essentially the

same characteristics as the blind sites except that they are formatted in different sizes (2 by 10 nautical miles, 5 by 6 nautical miles, etc.) and have more ground measurements. The Cartographic Laboratory has processed the ITS data and produced 1:24 000-scale maps of the field boundaries. The laboratory has also supported the LACIE yield team by supplying meteorological and agrophysical data in map form. Graphic aids for photointerpretation have been constructed for the LACIE analyst interpreters.

THE UNIQUE NEED FOR CARTOGRAPHY IN SATELLITE CROP SURVEYS

Many of the technical problems encountered in LACIE are the same problems which occur in map-

TABLE II.—Approved Special Crop Codes

Gray-scale level	Description of scene (a)	Approximate relative area proportions
61	Wheat + small grains	1:1
62	Wheat + small grains (2 or more)	1:2
3	Wheat + other annual crop (OAC)	1:1
64	Wheat + OAC	1:2
65	Wheat + OAC	2:1
66	Wheat + small grains + OAC	1:1:1
67	Wheat + small grains + OAC	1:2:1
68	Wheat + small grains + OAC	1:1:2
69	Wheat + small grains + fallow	1:1:1
70	Wheat + small grains + fallow	1:2:1
71	Wheat + small grains + fallow	1:1:2
72	Wheat + OAC + fallow	1:1:1
73	Wheat + OAC + fallow	1:2:1
74	Wheat + OAC + fallow	1:1:2
75	Small grains + OAC	1:1
76	Small grains + OAC	1:2
77	Small grains + OAC	2:1
78	Small grains + OAC + fallow	1:1:1
79	Small grains + OAC + fallow	1:2:1
81	Small grains + OAC + fallow	1:1:2

^aClass designations

- Wheat—winter or spring
- Small grains—barley, rye, triticale, oats, millet
- Other annual crops—beans, sunflower, safflower, sudan grass, corn, soybeans, sorghum, flax, potatoes, peas, mustard, etc
- Fallow—idle/fallow or idle residue

ping applications. An example is the problem where there were a number of aerial photographs (as many as six) for a single blind site. For Phase III, this created a problem because some photographs did not have sufficient correlatable points to register the photographic data to the Landsat image. The Cartographic Laboratory applied a photogrammetric solution (ref. 1) in which the photographic overlap was used to adjust all photographs simultaneously to the Landsat image without requiring control points on every photograph. The mathematical analysis involved is the same as in mapmaking and uses a simultaneous weighted least squares adjustment.

Cartographers have studied the shape of the Earth to improve their mapping product quality. The shape of the Earth may significantly influence the data set where a "flat Earth" assumption has been made. GSFC has begun using map projections in resampling data to improve the geometric quality. GSFC also analyzes the geometry of the scanner to reduce nonlinear systematic error. Figure 7 graphically il-

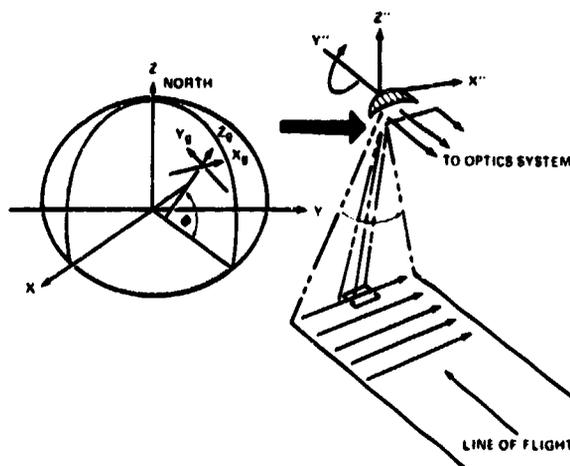


FIGURE 7.—Earth and scanner coordinate systems.

lustrates the scanner and Earth geometry, and figure 8 illustrates a mapping transformation that projects the spheroidal Earth onto a flat plane. The geometric analysis of scanners is very similar to the analysis of the Apollo 17 panoramic camera. Both sensors have rotating optics—the only difference is the recording medium (film or detector), which is significantly different in physical processing but similar in mathematical geometric analysis.

Another area in which cartographic technology may apply to remote sensing is in modeling geometric error and determining its quantitative effect on classification. This technique may be applied to the satellite crop surveys when they begin to approach classification accuracies of 95 percent or better. The effect of geometric errors on computed crop production may become significant at this level.

THE FUTURE CARTOGRAPHIC LABORATORY

Hardware Improvements

The Cartographic Laboratory has greatly enhanced its services for the Earth Observations Division, but it is expected that much more can be done. In particular, the machine interface problems created by physical data products such as maps, photographs, and ground-truth annotations can be reduced by improving the cartographic hardware system. Using the cartographic subsystem that was

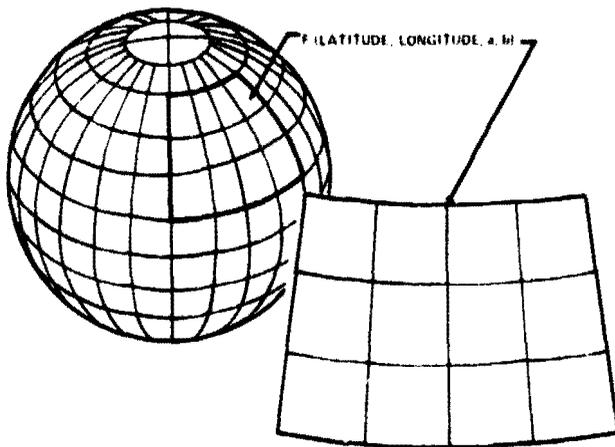


FIGURE 8.—Mapping transformation projecting the spheroidal Earth onto a flat plane.

designed for a proposed NASA Earth Resources Data System (ERDS) would alleviate many of these problems. This design (fig. 9) includes high-speed raster scanners which provide a means of rapid computer inputs of graphics, maps, and photographs. High-resolution black-and-white cathode-ray tubes (CRT's) also provide a means of accurate interactive measurement on digital images created from physical data sources.

Digital Data Base

Another area in which the Cartographic Laboratory can make a significant contribution is in developing and implementing a geographic computer data base which could provide digital imagery, soil types, meteorological history or condition, political membership, and all the other data needed to make satellite crop survey decisions. Because these data have spatial association with the geoid (Earth), they can be organized by spatial location (latitude and longitude). Pointers can be computed to locate data in the computerized data base as a function of spatial position.

Automation of Test Site Location

One of the most expensive tasks performed by the Cartographic Laboratory for LACIE was test site location. For future programs, this process could be automated, resulting in substantial savings. The digital data base would be loaded with the data used

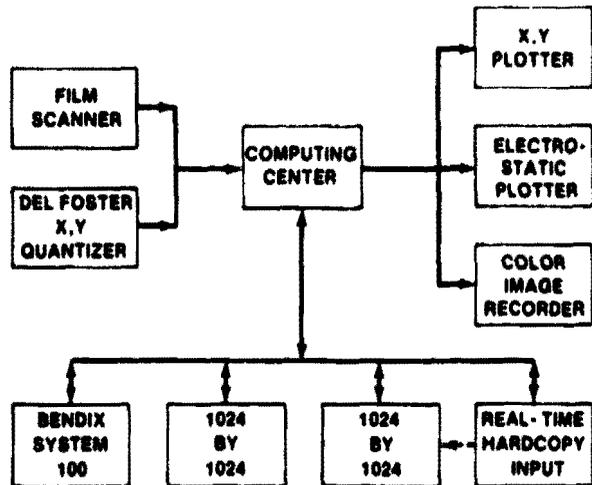


FIGURE 9.—Hardware configuration for cartographic subsystem.

in the selection strategy and the selection process could be machine programmed. This feature would produce a superior product and allow for changes in selection strategy without significant manpower expenditures. It would also provide a means of testing different sampling algorithms.

Automation of Blind Site Registration

Another area to be improved is the registration of ground truth to Landsat imagery for the blind sites. The current process requires the selection of control points that can be measured on the aerial photographs and the Landsat imagery. Because the ground-truth data are compiled and digitized as agricultural field boundaries (polygons), it would probably be better to register the data using an edge-detection correlator such as GSFC uses. Because GSFC has Landsat boundary maps for the LACIE test sites (of which the blind sites are a subset), correlation of the ground-truth boundary maps to the GSFC boundary maps derived from the imagery may be possible. This automatic correlation would then provide the information necessary to generate the registration coefficients.

REFERENCE

1. Rader, M. L.: Registration of "Blind" Site Aerial Photographs to the LACIE Landsat Image. LEC Technical Memorandum 10625, Lockheed Electronics Co., Houston, Tex., 1977.

The LACIE Data Bases: Design Considerations

L. E. Westberry⁴

INTRODUCTION

A primary purpose of the LACIE was to test the concept that large amounts of Landsat data could be analyzed in a real-time environment. Previous Earth resources computer systems relied on tape media for data storage and retrieval, and the volume of tapes necessary to support a LACIE-type system seemed to offer insurmountable physical and administrative problems. Thus, it was decided that LACIE would implement direct-access devices as the data storage media. In this case, all the data would be available immediately without prior staging, and the physical and administrative problems should be minimal.

This paper presents some of the design considerations involved in implementing direct-access storage devices for LACIE. The concentration is on the storage and retrieval of image data because this presented the most significant challenge. The discussion will include a definition of the problem, the solution methodology (or design decisions), the initial operational structure, the modifications which have been incorporated, some conclusions, and projections of future problems to be solved.

THE PROBLEM

The LACIE was initially set up to assess the worldwide production of one agricultural crop, wheat. The site, or sample segment, is the smallest unit of land involved in the assessment process. The analyst or interpreter determines a percentage of wheat contained in a sample segment by using the pattern recognition algorithms available in the LACIE software.

To begin the LACIE process, image data must be requested and received from the NASA Goddard

Space Flight Center (GSFC). The data are input to the LACIE system from the GSFC tape. In addition, field and dot update cards are input to the LACIE system to define or update certain "fields" and "dots" for each image to be analyzed.

The LACIE process allows the data to be statistically analyzed by one of two subsystems, the interactive subsystem or the batch subsystem. The batch subsystem is primarily a card-mode emulation of the interactive subsystem and is more restrictive than the interactive mode.

At the outset, it was planned to have approximately 1200 strata containing a total of 4800 sample segments. For each of the 4800 sample segments, there could be up to 16 acquisitions of data.

Data Base Size

An acquisition is a subscene extracted from a Landsat scene on a given day and composed of four spectral bands. Each acquisition consists of 117 scan lines, and each scan line contains 196 pixels, with four values for each pixel. To be represented in one band, each pixel requires a byte (eight bits). Thus, for each acquisition of a site, there are 91 728 bytes of data.

$$\begin{array}{r}
 117 \text{ lines} \\
 \times 196 \text{ pixels} \\
 \hline
 22\,932 \text{ bytes per band} \\
 \times 4 \text{ bands} \\
 \hline
 91\,728 \text{ bytes of data per acquisition}
 \end{array}$$

In addition to the actual image data, there is information describing the Landsat scene from which the image was extracted. This information must be retained with the image for the data to be useful to the analysts. This information, the header, requires an additional 3062 bytes of storage. The total storage required to retain a single image is at least 94 790 bytes.

⁴IBM Federal Systems Division, Houston, Texas

CA

If the desire is to store and access 16 acquisitions for each of 4800 sample segments, the total storage requirement is

94 790 bytes per acquisition
x 16 acquisitions per site (maximum)

1 516 640 bytes per site (maximum)
x 4 800 sites (maximum)

7 279 872 000 in storage (maximum)

In broad terms, the problem can now be stated: design an integrated data base which can hold up to 7.3 billion bytes of image data along with the associated field and/or dot definitions and allow access in pieces of about 95 000 bytes. However, this simple statement does not fully describe the problem because there were additional constraints and targets.

Since the image data would flow into the data base over a growing season and since not every site would have the entire 16 acquisitions (because of cloud cover and other problems relating to data quality), it was not practical to design a data base which would hold the entire 7.3 billion bytes. Additionally, there was no way of knowing beforehand the exact distribution of acquisitions to sites. Some sites, at year end, might have no data associated with them, while other sites might have the full 16 acquisitions. Thus, it was necessary to design a data base which would allow great flexibility in terms of the distribution of acquisitions to sites.

Cataloging and Cross-Referencing

The time dependency of the image data acquisition also implies a need to catalog the data so that analysts can determine data availability.

As image data is received, the analysts must determine whether the data will allow the pattern recognition process to be performed. Once this determination is made, the analysts must prepare the additional inputs for the pattern recognition algorithms. These inputs, consisting primarily of field and/or dot definitions, are directly related to the specific site under investigation. Thus, when the field and dot definitions are stored, they must be correlated with the imagery for which they were developed. This implies that the cataloging scheme must allow for correlating the image data with the associated field data.

Data Security

Another consideration was the protection of a large cross-referencing data base. The data contained on the LACIE data base had to be reasonably secured from the inadvertent deletion of data by the analysts themselves or the many other users within the NASA Real-Time Computer Complex (RTCC). It was deemed critical that procedures be designed into the LACIE data base accesses such that human or machine errors would not cause complete chaos. The protection had to ensure either that errors could not occur or that, if they did occur, the data base could be restored to its original status.

Processing Constraints

The requirements that the computer analysis operations performed on the RTCC computers be interactive operations or simulate interactive operations to give a batch capability implied additional constraints on the data base design. It was necessary to design sufficient random access capability into the data base such that requested images and fields could be retrieved from the data base in a reasonable amount of time. The processing throughput targets were 30 to 40 segments in a 16-hour period in the initial configuration; the target increased to 120 segments per 16 hours in subsequent configurations. These targets meant that the entire analysis operation had to flow through the RTCC computers on an average of one every 21.5 minutes initially, and finally, one every 8 minutes. Clearly, retrieving the images and fields from the data base could not consume an inordinate amount of time.

Cost Objectives

The budget allocated to the accomplishment of the design and implementation of the data base to support LACIE was an additional constraint. The actual budget levels are not important to this discussion, but it should be stated that cost effectiveness was an important consideration.

The initial problem can now be stated thusly: design and implement, as inexpensively as possible, an errorproof data base structure to support LACIE in such a manner as to allow any given segment to be processed in no more than 8 minutes.

SOLUTION METHODOLOGY

There were two basic designs initially presented to solve the LACIE data base problem. One design proposed a large multivolume image data base to store the image data, a history data base to hold the catalog data, and a field data base to hold the field definitions. The second design proposed multiple image data bases, each a single volume, supported by the same history and field data bases. The IMS-360 was proposed as the data base manager, the system software support for the data bases. The IMS-360 was an off-the-shelf product; thus, the cost of developing a specialized data base support package was eliminated. The IMS-360 also offered sufficient data base management services to support either of the two initial designs completely.

The designers realized that the image portion of the LACIE data base was the most critical because of the potential size. As a result, efforts were begun to determine realistic size limits.

Sizing the Data Base

The first major result of the sizing effort was the understanding that the 3000-byte header was largely duplicated for each acquisition of a given sample segment. Each new acquisition after the first one for a segment really required only 24 bytes of storage to record the differences. This understanding led to the first step in a logical design, which is illustrated in figure 1.



FIGURE 1.—Simple view of image data base logical structure.

Initial Logical Design

The logical design at this point would have an image data base with the key portion being a single 3000-byte header. Each header could have up to 16 acquisition headers associated with it, and each acquisition header would have one data segment associated with it. The resulting size would be

$$\begin{array}{r} 4\ 800\ \text{sites} \\ \times 3\ 062\ \text{bytes} \\ \hline 14\ 697\ 600\ \text{bytes of site headers} \end{array}$$

plus

$$\begin{array}{r} 16\ \text{acquisitions} \\ \times 4\ 800\ \text{sites} \\ \hline 76\ 800\ \text{sites per acquisition} \\ \times 24\ \text{bytes} \\ \hline 1\ 843\ 200\ \text{bytes of acquisition headers} \end{array}$$

plus

$$\begin{array}{r} 16\ \text{acquisitions} \\ \times 4\ 800\ \text{sites} \\ \hline 76\ 800\ \text{sites per acquisition} \\ \times 91\ 728\ \text{bytes} \\ \hline 7\ 044\ 710\ 400\ \text{bytes of data} \end{array}$$

equals

$$7\ 061\ 251\ 200\ \text{bytes}$$

Reducing the number of site header records would save almost 220 million bytes of storage space. The ITTEL disks being considered as the storage devices for LACIE would each hold approximately 100 million bytes of data; thus, the reduction in the number of site header records stored would cut the requirement by two packs.

The number of disk volumes required to retain the image data had been reduced from 73 volumes to 71 volumes, a 3-percent reduction. This savings was not considered significant. Obviously, a better estimation of the maximum number of acquisitions to be retained would yield a more significant reduction in the maximum data base size. The LACIE planning staff determined that the LACIE image data base should be of sufficient size to store up to 4

acquisitions of data for 3840 sites and up to 16 acquisitions for 960 sites.

The new requirement reduced the maximum number of acquisitions from 76 800 to 30 720. The resulting maximum image data base size would be

$$\begin{array}{r} 4\ 800\ \text{sites} \\ \times 3\ 062\ \text{bytes} \\ \hline 14\ 697\ 600\ \text{bytes of site headers} \\ \\ \text{plus} \\ \\ 30\ 720\ \text{acquisitions} \\ \times 24\ \text{bytes} \\ \hline 737\ 280\ \text{bytes of acquisition headers} \\ \\ \text{plus} \\ \\ 30\ 720\ \text{acquisitions} \\ \times 91\ 728\ \text{bytes} \\ \hline 2\ 817\ 884\ 100\ \text{bytes of data} \\ \\ \text{equals} \\ \\ 2\ 833\ 318\ 980\ \text{bytes} \end{array}$$

This reduced the number of disk volumes required to store the image data to 29.

Security Considerations

The next step in the process was to determine the implications of supporting a 29-volume data base using IMS-360 as the data base manager. The designers were aware that IMS-360 provided for data base protection via extensive recovery utilities. There were IMS-360 utilities available for checkpointing data bases by copying to tape either the physical or the logical structure of the data base. The recommended IMS-360 procedure for data base recovery was to copy the most recent checkpoint tapes to the set of disks comprising the data base and then to read in the IMS log tapes that were created between the time of the checkpoint tape creation and the error occurrence.

The error recovery procedures implicit in IMS-360 had significant consequences relative to the two initial data base designs. If a failure occurred on a single disk volume of the multivolume data base, the time required simply to restore the checkpoint tapes to the

29 volumes would exceed 4 hours, given the maximum possible transfer rate from the tape drives. And there was no hope of transferring 3 billion bytes of data at the maximum transfer rate. The more probable transfer rate of 89 600 bytes/sec suggested a recovery time of at least 8 hours. The same single-disk failure in the design involving multiple image data bases would necessitate the recovery of only the disk volume that failed and not the other 28 volumes. In this situation, the recovery time for such a failure would be reduced to a little over 9 minutes at the maximum transfer rate and 18 minutes using the realistic rate of 89 600 bytes/sec.

Upon consideration of the time requirements for checkpointing and restoring the image data bases using the IMS-360 techniques, the design proposing a single multivolume image data base was rejected. The time required to checkpoint (copy from disk to tape) was about equal for both proposals, but the recovery from error situation clearly favored the multiple image data base proposal.

Expanding the Multiple Data Base Design

At this point, the designers began to refine the image data base structure. They had to define a process which would divide the image data into several IMS data bases, one data base per disk volume. However, it was also necessary to make the structure look like a single image data base to the application programs or at least provide a method by which the application programs could easily get to the proper image data bases.

Distribution of the data over the packs.—A master catalog or index appeared to be the solution allowing the application programs to refer to the desired data bases. Knowing that the application programs wanted to access image data simply by supplying the site number and that the IMS required the address of a program control block (PCB), which contained the "key" and the data base, the designers were able to build a catalog. The application program would simply call a subroutine passing the site number, and the subroutine would return the address of the required PCB. The application program would then issue the proper IMS request for the desired site.

The catalog is structured such that allocation of sites to data bases is controlled by a 10 000-byte table, in which sites are represented by table position and the position content specifies the data base to which the site is assigned. Next came the question of how

to assign the sites to each image data base. It seemed simple merely to divide the 4800 sites by 29 packs to get the number of sites to assign to each pack, then to assign the sites consecutively to a pack. Thus, image data base 1 would contain sites 1 through 166; image data base 2 would contain sites 167 through 331; etc.

However, this very simple assignment technique would run into difficulty when "intensive test sites" were introduced. Intensive test sites are acquired year-round with no break in the acquisition cycle, as opposed to normal production segments whose acquisition windows conform to the crop growing season. The intensive study sites are likely to have more acquisitions than the regular sites, and, if a large number of intensive test sites are allocated to a single pack, there is a high risk that that pack will be unable to contain all the required acquisitions. Knowing that site numbers would be assigned consecutively within a country and that some countries could contain many intensive test sites, data base designers were led to a standard IMS randomizing routine that would randomly assign the 4800 sites to the various image data bases. The use of this routine would make the proportion of intensive test sites and the proportion of sites for each country about the same on each pack (fig. 2).

Distribution of the data on a pack.—The same concerns that existed for the distribution of the data over the set of packs also applied to distributing the data on each pack. For the sites assigned to a given pack, the loading of data acquisitions would be uneven and unpredictable. Further, the intensive test sites would require more space than the normal production sites. The same logic that necessitated randomly distributing the sites over the entire set of image data bases applied to randomly distributing the sites over the blocks on an individual image data base.

Blocking the image data.—At this point, the problem of allocating the sites to the multiple data bases and making the structure independent of the application programs had been solved. The problems of efficiently blocking the data for IMS and the physical devices to be used were addressed next. The simple logical view, shown earlier in figure 1, indicated data records of almost 92 000 bytes. Physical records of this size are simply impractical, if only because of the buffer size required to transfer the record from disk to computer storage. The track size of the ITTEL disk allows for slightly more than 13 000 bytes of data, and, since the device uses track addressing, it seemed logical to choose a record size which was either a multiple or a divisor of that track size.

<u>DISPLACEMENT</u>	<u>INDEX</u>	<u>SITE NUMBER</u>
0	DATA BASE 1	0
1	DATA BASE 22	1
2	DATA BASE 3	2
3	DATA BASE 11	3
4	DATA BASE 2	4
5		5
.		.
.		.
9995	DATA BASE 29	9995
9996	DATA BASE 4	9996
9997	DATA BASE 11	9997
9998	DATA BASE 10	9998
9999	DATA BASE 15	9999

FIGURE 2.—Master index structure.

INITIAL OPERATIONAL STRUCTURE

The designers consulted the programmers to determine how the statistical routines would be accessing the image data and learned that the programmers intended to code routines that would compute statistics for up to four acquisitions at a time, line by line across all the acquisitions. A logical data base structure that would match this approach is shown in figure 3.

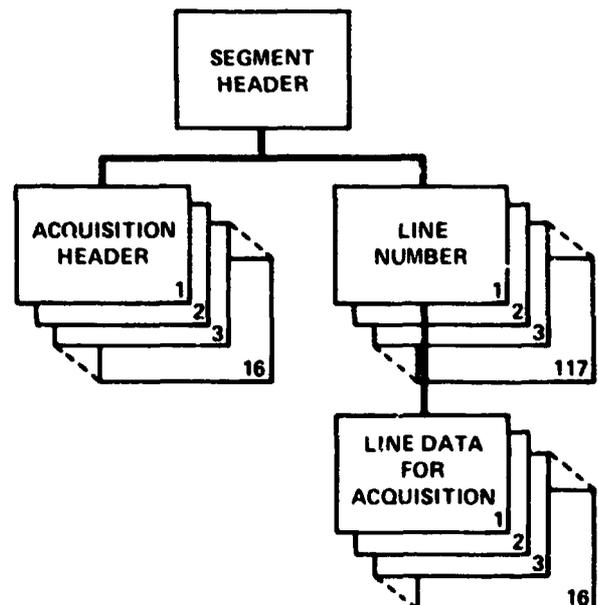


FIGURE 3.—The logical structure to match processing logic.

This logical structure and consideration of buffer size combined with the knowledge that 80 percent of the sites would have only four acquisitions allowed the designers to develop a physical structure which would look like this:

Acquisitions 1-4

Line 1	Line 2	Line 3	...	Line 117	Pad
--------	--------	--------	-----	----------	-----

1 block (6442 bytes)

Approximately 30 tracks of 7330 disk space

However, for those sites which have more than four acquisitions associated with them, the physical structure would be

Acquisitions 1-4

Line 1	Line 2	Line 3	...	Line 117
--------	--------	--------	-----	----------

1 block (6442 bytes)

59 blocks (380 078 bytes)

Acquisition 5

Lines 1-4	Lines 5-8	Lines 9-12	Lines 13-16	...	Lines 11-117
-----------	-----------	------------	-------------	-----	--------------

Acquisition 16 . . . Each additional acquisition requires 7.5 tracks of 7330 disk space.

The block size of 6442 bytes was determined to be the optimum block size for the actual image data based on the buffering within the central processor, the characteristics of the 7330's, the characteristics of the IMS access method, and the data characteristics.

However, the blocking factor that was optimal for the image data was not optimal for storing and retrieving the ancillary data (the site and acquisition headers). Combining all the ancillary data associated with a given site would require approximately 3500 bytes of storage. When the storage requirement, data base manager overhead, and data pointers were considered, the optimum block factor for the ancillary data was determined to be 4248 bytes per block.

The means of obtaining the optimum blocking factor for each type of data was to define two data set groups for each image data base. One data set would contain only image data, and the other data set would contain the ancillary information. Defining two data sets for each image data base also offered an additional advantage. If the data sets were placed on

different packs (volumes), the contention for the arm on the device could be reduced, thus eliminating some wait time required when the device arm is moved from one location to another. This was done, and the resulting image data bases were organized as shown in figure 4.

While the structure, both logical and physical, matched the application logic, it turned out to be highly inefficient for storage and retrieval. The line data for each of the first four acquisitions was scattered over 59 blocks of storage, and, while user requests for all four acquisitions could be handled as efficiently as requests for only a single acquisition, in a worst-case situation (where there were 16 acquisitions available for a site and the analyst wanted the last four in inverse order), 5850 accesses were required. The operational average for image retrieval requests for sites having 16 acquisitions turned out to be 760 accesses. Because there is a significant amount of central processor overhead associated with each access, the designers wanted to eliminate any unnecessary accesses. They proceeded to look for improvements in the data base design to accomplish this. The current image data base structure is the result of the overhead reduction study.

CURRENT OPERATIONAL STRUCTURE

The new design differs from the previous one only in treatment of the first four acquisitions. The data for each acquisition are placed on the data base by line. This procedure puts the lines for each acquisition in a set of 15 contiguous blocks.

Acquisition 1

Lines 1-8	Lines 9-16	...	Lines 113-117
-----------	------------	-----	---------------

15 blocks (96 630 bytes)

Acquisition 16

Lines 1-8	Lines 9-16	...	Lines 113-117
-----------	------------	-----	---------------

15 blocks (96 630 bytes)

As a result, the number of accesses required to retrieve the first acquisition of the site is only 16, and it drops to 15 for every additional acquisition of the same site. The logical structure which maps the new physical structure is shown in figure 5.

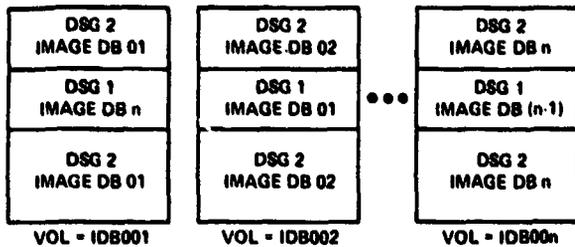


FIGURE 4.—Data set allocation for LACIE image data base.

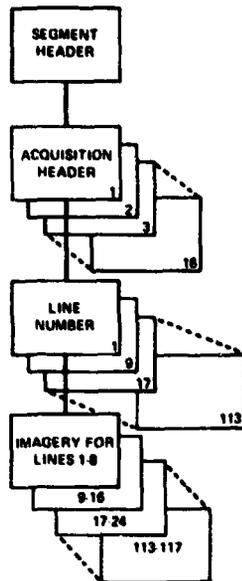


FIGURE 5.—Logical structure implemented.

The first structure incorporated programmer plans into the data base design. Then, in order to increase the efficiency of accessing the data base, the structure was modified. The modified structure did not fit with the programmer's plans. This problem of incompatibility between the retrieval logic and the application logic was resolved by setting up four large-core storage buffers for IMS output and application input and by modifying the call sequence to IMS. Instead of retrieving each acquisition in full before going to the next, the application requests each acquisition in increments of eight lines, which is a full block of data. This allows the application program to transfer the image data to a working storage area on a disk in the order that fits the processing logic. Only the first call in each group of eight is fully qualified with cylinder, track, and block. The following seven are sequential calls, which are most efficient for IMS-360.

Cross-Referencing

The necessary cross-referencing is accomplished by using the site number in all the data bases. The site number is the key for the application program to access the image data bases as well as the history data base and the field data base. The analysts work with one site at a time; the software design takes advantage of this fact. By knowing which site and acquisitions are being worked on, the programs can access all required records for processing.

Security

The data bases are also recoverable within a reasonable time frame. With a multiple data base set for the image data, a pack failure requires the recovery of only that pack and not the entire image data base. Because the data on each image data base are physically organized to minimize the number of accesses required to retrieve each record, the time required to store or retrieve all records is minimal. As a result, a full image data base, one volume, can be completely dumped to tape or restored from tape in approximately 20 minutes.

Data base security is provided by periodic checkpoints, where all the data bases are dumped to tape. The IMS log tapes created between checkpoints are retained. If an error occurs, the recovery procedure begins by identifying the data base that is in error. Then the checkpoint tape for that data base is restored to a disk via an IMS utility. Another utility reads all the log tapes, except the one which contained the error, for the data base identified as being in error. In this way the data base is recovered up to the point of the error. The updates that were done after the error must be redone.

Expansion

The multiple image data bases also offer an additional benefit. The GSFC/JSC Interface Control Document specifies a maximum of 4800 sites with 3840 sites having up to 4 acquisitions and 960 sites having up to 16 acquisitions. These dimensions imply a maximum data base size of 31 volumes. In practice, data arrive at JSC at the rate of 0 to 120 acquisitions per day. Thus, throughout a crop year, a considerably smaller data base can be used to contain the image data. However, with such a compact image

data base, provisions must be made for a data set overflowing its volume. The provisions are twofold. First, all image data set groups are cataloged to an overflow volume (IDBOVF) in addition to their primary volumes (fig. 6). Second, a procedure has been constructed to add volumes and redistribute sites.

Conceptually, the expansion process proceeds in the following manner. The LACIE image data bases consist of a collection of sites and associated data whose distribution over n volumes is specified in the master index. An expansion of the image data bases is accomplished by adding k new volumes to the existing n volumes and redistributing the sites over the new $n+k$ volume configuration. A new master index which reflects site locations in the new data base configuration is constructed by considering number of sites, number of acquisitions, and balance. Those sites which have new data base assignments are unloaded to tape. The new index is used to control the reloading of the unloaded sites at their new data base locations. Finally, the reloaded sites are deleted from their old locations.

Daily Use of LACIE Data Bases

Now that the development of the LACIE data bases has been examined, the day-to-day use of the data bases will be discussed. The LACIE data bases are in use approximately 8 hours per day, 5 days per week. A typical day of support will be divided among several user groups. Usually, the first portion of a production period will be used for data base updating. After the updating is complete, the remaining time will be spent in either batch production or interactive use.

The production activities are fairly consistent and tend to follow a daily pattern, at least as far as the uses of the data bases are concerned.

Image data are requested from GSFC to start the LACIE process. This is accomplished via the history update job, which uses card inputs to update the history data base and produce a JSC interface tape, a history data base query report, and a listing of history data base updates. The JSC interface tape is sent to GSFC to order image data.

Receiving the data from GSFC is the second step in the process. The data are input to the LACIE system from the GSFC tape and are processed by the composition and indexing (C&I) job. The C&I job

may contain card inputs which define the sites to be processed or excluded from the tape. The C&I job updates the history and image data bases and generates the daily report. The C&I job also computes a "green number" for the image being processed. The green number is stored on the image data base in the variable header for the acquisition.

In addition, the LACIE process requires that "fields" and "dots" defined for each image be analyzed. The field update job uses card inputs to update the field data base with the definitions and produce a transaction report, a field report, and a listing of field data base updates. A field overlay tape is also produced by this job.

The dot data base utility is used to define and label 209 pixel locations per segment, called "dots." Dots are used as starting vectors for several of the analysis routines. The category (blank if not specified) and function for each dot are defined. The dot data base utility, which has both complete replacement and partial update capabilities, maintains the dots on the dot data base. The dot data base resides in segments of the field data base. Figure 7 charts the steps in the preparation for an analysis run.

Now the data is ready to be statistically analyzed. There are two subsystems in place for processing the data, the interactive subsystem and the batch subsystem. The batch subsystem is primarily a card-mode emulation of the interactive subsystem and is more restrictive than the interactive mode. Figure 8 charts the analytical process in the batch mode.

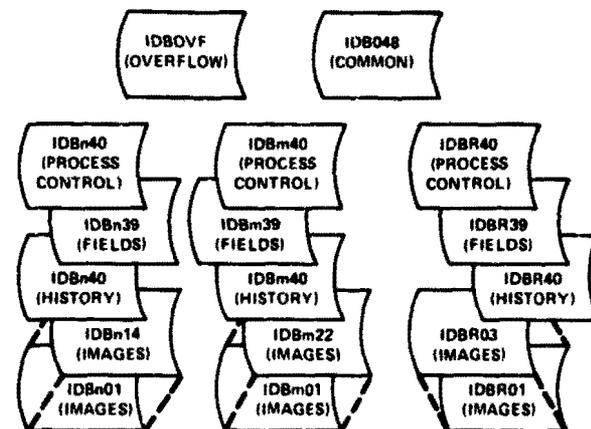


FIGURE 6.—Operational data bases for crop year n , crop year m , R , T , and E .

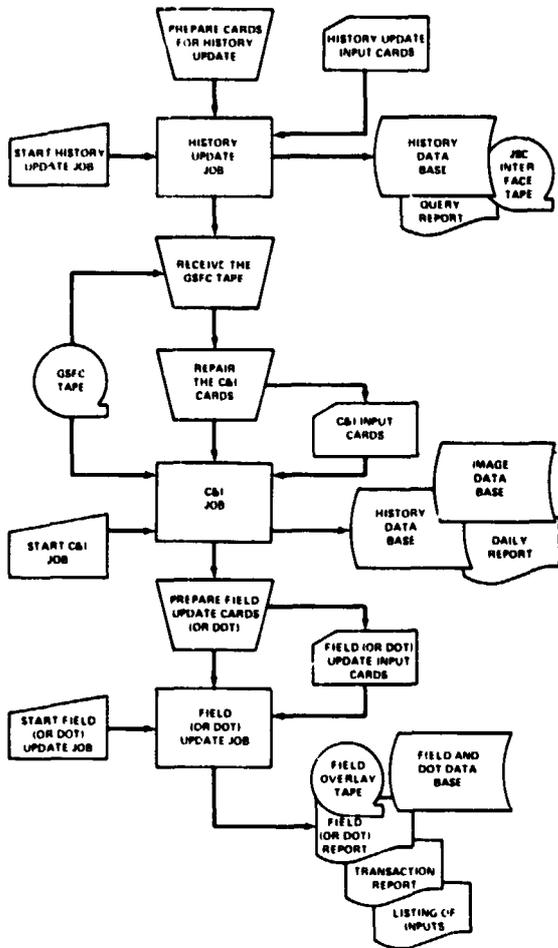


FIGURE 7.—Preparation for an analysis run.

CONCLUSIONS

The LACIE data bases have been in use for almost 4 years. Data for the crop years 1975, 1976, and 1977 have been received and processed; currently, the 1978 image data are being received at JSC. Throughout this support, the image and ancillary data have been efficiently stored and retrieved for processing. The processing throughput has been consistent with design objectives.

The IMS utilities for data security using the checkpoint/restore methodology have performed as expected. Data losses have been negligible.

By taking advantage of the expandability allowed by the multiple image data bases, the most effective use of the 7330's has been made.

FUTURE DATA BASES

The LACIE data bases met the design objectives for the limited environment for which they were intended. However, there are some limitations to the LACIE solution that must be overcome if a data base of worldwide and multicrop coverage is to be developed.

The future data bases to support the new environments must contain more data. The future Landsat will have more sensors with higher resolution to produce more information per unit area of land. Additionally, new applications such as air quality, water quality, and land use will require new access to the data. There will also be new satellites such as Seasat and the Soil Moisture Satellite, and analysts may desire to combine data from several satellites to address a particular problem. All these concepts in combination imply a global data storage and distribution problem that the LACIE methodology has not begun to address.

The most apparent shortcomings of the LACIE data base design in considering the future applications are the floorspace requirements for the direct-access storage devices, the overhead required by the checkpoint system for data base protection, and the overhead associated with reorganizing the data base

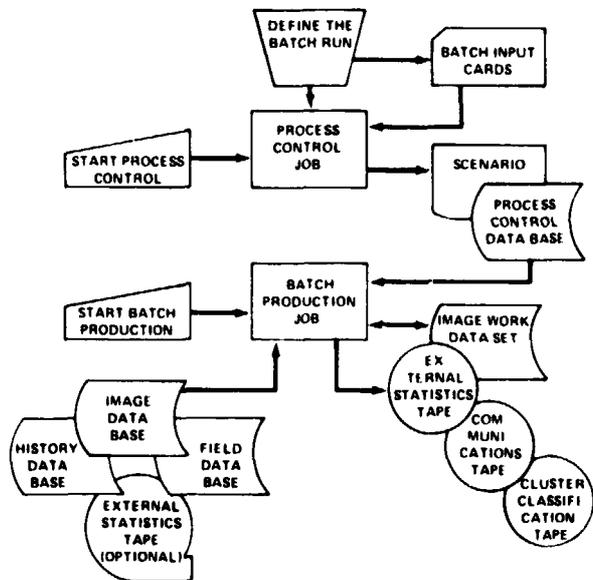


FIGURE 8.—Processing the image in batch mode.

if the balance of data to be retained is greater than anticipated.

The solution to the floorspace problem is available with current technology. There are mass storage devices with storage densities much higher than those of the currently used direct-access devices. The use of these mass storage devices would require the trade-off of some retrieval time. Possibly, a combination of direct-access devices and mass storage devices would allow rapid retrieval and the on-line storage of large amounts of data. Such a combination might use the mass storage devices for permanent storage of the data and the direct-access devices for immediate access to the data.

This same combination would probably also yield some benefits in the areas of error protection and data recovery. If a staging device were the primary source of data available to the analysts, the mass storage devices would not be accessible to them. Any errors that resulted during the analysis process would be confined to the staging device and only it would need to be recovered.

The mass storage devices may offer only a partial solution to the storage problems of the future. To

augment the capacity of the mass storage devices, it may be necessary to apply data compression techniques to the data before storing it. The total solution may even involve distributed data bases with a network capable of transporting data from one geographic location to another.

Assuming the interest in determining worldwide multiple crop production continues, the next data base challenge will almost certainly involve solving a trillion-byte data base problem.

BIBLIOGRAPHY

Project Implementation Plan—Large Area Crop Inventory Project. NASA Ground Data Systems Division, June 28, 1974.

LACIE—Large Area Crop Inventory Experiment. Technical Directions, vol. 2, no. 3, IBM Federal Systems Division, Autumn 1976.

Data Base Administration Manual, IBM Federal Systems Division.

Man-Machine Interfaces in LACIE/ERIPS

Barbara B. Duprey^a

INTRODUCTION

The Earth Resources Interactive Processing System (ERIPS) supports the LACIE. There are three major man-machine interfaces in this system: the use of "menus" for communication between the software and an interactive user; the checkpoint/restart facility to recreate in one job the internal environment achieved in an earlier one; and the error recovery capability, which greatly reduces the impact of errors which would normally cause job termination. This interactive system has also been adapted for use in noninteractive (batch) mode.

The LACIE/ERIPS software system is a large computer program developed by IBM Federal Systems Division, Houston, Texas, in support of NASA's Earth resources data analysis activities. ERIPS executes on an IBM 360/75 mainframe in the Real-Time Computer Complex (RTCC) in Building 30 of the NASA Johnson Space Center (JSC), Houston, Texas. A general description of the development and capabilities of this system is presented elsewhere in these proceedings (C. L. Johnson, "LACIE/ERIPS Software System Summary"). One of the most important aspects of the interactive portion of the system is the way in which the analysis and decisionmaking capabilities of a human being are integrated with the speed and accuracy of a computer to produce a powerful analysis system. A basic goal of the design of the system was for it to be "man-rated"—easy to use, amenable to human direction, and forgiving of human error (by both users and programmers). This paper discusses the techniques used in ERIPS to reach this objective.

The broad objective of "man-rating" the system led to the development within ERIPS of three capabilities: the menu-style user control interface,

the checkpoint/restart facility, and the error recovery facility. The menus are graphic displays which are the user's principal interface with the system. Each menu presents information concerning the status of the system, and it either requests necessary information or shows current options so that the user can make selections as he would from a restaurant menu. Checkpoint/restart allows for recovery from total system failures and for return to a previous point in processing after a lapse of time, as from one day to the next. The checkpoint/restart function is partially controlled by the user via menu inputs. Error recovery protects the results of previous processing whenever a program failure occurs by automatically returning the system to prefailure status.

The ERIPS user interfaces with the system by means of a terminal. This terminal includes a graphics or conversational screen on which the menus are displayed, two image screens (one of which displays 16 levels of gray, the other either 8 or 64 colors), a keyboard, and a "joystick" which drives a cursor indicator to the same relative position on all three screens. The physical layout of the hardware is shown in figure 1 and the keyboard layout in figure 2.

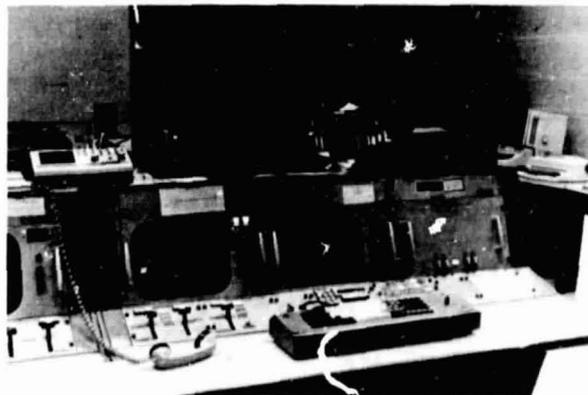


FIGURE 1.—ERIPS interactive terminal.

^aIBM Federal Systems Division, Houston, Texas.



FIGURE 2.—ERIPS keyboard layout.

MENUS

Purpose

The menu interface in ERIPS was originally established to provide an interactive user with a tutorial display of the options available at each point of interaction and then to respond to selections. The ERIPS menus also display various types of messages and dynamic data. They thus provide the user with a considerable amount of status information to aid in the intelligent selection of a processing option.

This is in sharp contrast to the philosophy of those interactive systems which, like OS-360, depend on the user to request functions at will, with reference to some external documentation such as a user's guide. That approach tends to leave the user tied to at least three sets of documentation: his own objectives for this particular session, notes and hardcopies of what has happened so far, and the user's guide to determine what can be done next. This is simply too large a burden to place on the user of a system with as many capabilities as ERIPS.

Other systems have used the menu concept, including the DRAFT (Display Retrieval and Formatting Technique) and Skylab software systems from which some of the menu and display logic was taken (see ref. 1), but an unusual feature of the ERIPS approach is the way the menus are logically linked to form what is in essence an inverted tree. This makes it easy to show the user only what is needed to continue along the path he has chosen; it also makes it very easy to back up and choose an alternative path.

Display Hardware Considerations

The DRAFT Digital Television Equipment (DTE) terminal hardware for which ERIPS was developed has naturally had a significant effect on the implementation of the menu concept. This terminal has the capability to display lines and a wide range of symbols, but it has no local editing capability—each symbol or line of display must be specifically directed by the software. Forward and backward paging and scrolling are also under direct control of the system software. All these functions, and many others, are incorporated into some of the "intelligent" terminal systems now available at a relatively low cost; use of such terminals could relieve the central processor of much of the display processing overhead.

Another feature of the terminal displays in general is the split screen. This is implemented on current hardware in limited form. The conversational screen, for instance, is segmented into the menu display area, the separately maintained overlay for the menu display, and three one-line message areas. Terminal systems with far more extensive segmentation capabilities now exist and might profitably be used to provide a running message log, operator-generated messages, a scratch pad area, and so forth.

It should be noted, however, that the transportability of a menu-oriented system can be severely diminished by too great a dependence on terminal intelligence, which might not be common to all the installations for which the system is intended. A "lowest common denominator" must then be identified, and the requirement of total control by the central processor has, historically, been very successful (see Johnson).

Menu Definition

Before menus can be used, they must be defined. Since display of a menu involves processing overhead and demands user interaction, the menu's scope must be wide enough to justify its existence. This makes the menu concept impractical for some systems. On the other hand, the menu must not have such a wide scope that it overwhelms the user with its complexity. When this question of scope has been resolved, menu definition can begin.

The ERIPS menus consist of control information, static contents, and dynamic contents. (See reference

edge of the conversational screen. These act like the special function keys found on many terminal systems, except that here it is possible for a key to have a different function on each menu. They always include one for return to the previous menu and alternatives to the end-of-field (EOF) and end-of-transmission (EOT) special action keys. Depending on the applications, there may be up to 15 other functions defined; selection of one of these normally brings up the first menu for that function.

Special action keys.—The terminal keyboard has several keys which are special menu action indicators. The EOF key signals that input for the current field is complete (and, if there is no input, confirms the default values for the field). The EOT key ends any current field and signals that all menu input is complete and ready to be processed. There are forward and backward space keys, and there are keys which clear the areas containing the supervisor and application messages (which also clear automatically on a timed basis). There are other special action keys which deal not with menus but with management of reports and with debug services; these will be discussed later.

Joystick and toggle switch.—The joystick/toggle switch combination is a very powerful portion of the ERIPS input scheme. The joystick drives cursors on all three screens of the user's terminal; the toggle switch, when pointed toward a screen, indicates that the cursor has reached a significant position on that screen. The user points the toggle switch at the conversational screen to make decisions; to move the alpha cursor to the beginning of a field, overriding its automatic placement (and ending any current field without the need for the EOF key); and to identify coordinates for line drawing or other functions. In these cases, the match to the points identified in the menu control information need not be exact but only within approximately two character widths in any direction. Image screen pointing serves similarly for coordinate identification on these screens.

Erasures.—Certain menu inputs serve as erasures of previous actions. Pointing at a previously selected decision box cancels the selection and removes the plus sign. Pointing at the beginning of a field and typing a blank reestablishes the default for that field, while typing alternative characters corrects the field. Backspacing when a line-drawing function is underway erases the most recent line and returns to its starting point. Even the menu itself is effectively erased by use of the "return" special function. Thus,

nearly all menu actions can be undone if the user so desires.

Error Checking

The input error-checking function of ERIPS is extensive and operates at several levels. The system constantly evaluates current status and displays messages when something violates the rules then in effect (or follows an abnormal path of which the user ought to be notified).

The first level of error checking on menu inputs occurs as they are entered, when they are compared with the current set of expectable actions. Types of errors noticed in this fashion include trying to type in too many characters for the receiving field, pointing to an area which was not identified in the control information, or attempting to make inputs when the keyboard is logically locked out. Such inputs are ignored, and an appropriate message is output.

The second level of checking occurs when the user signals end-of-input. At this time, the set of inputs made for the menu is checked to see that all are in the appropriate form and do not violate any of the limits identified in the control information. Errors noticed at this level include attempting to make incompatible decisions, making no choice when one is required, or violating the numeric range associated with a field. Again, appropriate messages are issued, and the system waits for corrected inputs to be made.

Finally, the inputs are checked against the dynamic condition of the system. For instance, if the user's choice of a processing path depends on the existence of data which are not available, a message is generated indicating that either another path must be chosen or the data must be supplied. The user then has to respond appropriately, perhaps by returning through the menu sequence to a point at which the data can be generated, then proceeding in the required fashion until he again reaches the point at which the error was detected.

The intent of all this is to be as forgiving of human error as possible without allowing such errors to jeopardize the integrity of the results.

Menu Use in Noninteractive Mode

The discussion so far has been in terms of interactive use, with an analyst at a terminal responding to

the displays as they appear. In fact, though, this is not the mode in which the system is most often used. The whole menu scheme has been adapted for use in noninteractive (batch) mode, in which the terminal hardware is not even connected to the system. Though the resulting system is certainly unlike any developed strictly for batch purposes, it produces comparable output, and it has the distinct advantage that software modifications are applied to one set of programs rather than to two divergent ones, thus ensuring synchronization and reducing implementation costs.

There are actually two forms of batch mode, both depending on input cards to control system processing. The "process control" batch mode divides the possible system actions into those which are always (or never) to be done and those for which user inputs are required. Many assumptions based on common practice have been made in order to reduce the number of actions in the latter set; there are less than 30 types of input cards, corresponding to about as many menus, and nominal processing for a particular geographic site requires only 3 of these. The user creates a control deck for each site to be processed; normally, many such decks are combined to form a job's input. The off-line process control system checks the syntax and logical compatibility of the cards. If all goes well, it constructs a scenario for the run, drawing from user inputs as required. It produces a data base for the on-line system, along with an abbreviated version of the scenario on a hardcopy printer. (Figure 4 shows a typical batch deck and its scenario.) When the on-line system uses the data base, it repeats the hardcopy scenario as a record of activities, appending a status line which indicates success or failure. This hardcopy, together with other outputs such as film products, is returned to the user.

The other form, "regular" batch, demands very detailed knowledge of the menus and their flow, as the inputs must all be specified at the same level as for an interactive run. Because of this, it is seldom used, but it can provide for accurate repeatability of complex input sequences which do not conform to the assumptions made for process control.

One aspect of running in either batch mode, of course, is that nobody is in a position to act on the "displayed" messages, so the system cannot wait for corrections as it would in interactive mode. Instead, the messages as defined are divided into those after which no meaningful results for the site can be produced ("fatal errors") and those which can be treated

```

NAME=DPY (19-PE12)
MODE=120
GROUP=(73304)
ITER=AUTO/PROG/,120,0.05,1,1,1,0191,1,1,1,ASC/,/0015//

RECOGNITION 1720 (73304) 01
RETRIEVE RECOGNITION AND TRANSMIT SITES - RESULTING NUMBER OF CHANNELS IS 04
RETRIEVE FIELDS, PRINT FIELDS REPORT
NORMAL STATISTICS NOT COMPUTED

ITERATIVE CLUSTERING
OUTPUTSOME - NAME GENERATED BY AUTOMATIC LABELING
AREA DEFINITION - 0000
STARTING VECTOR - 0015
ITERATIVE DESCRIPTIVE PARAMETERS
CO 20 G2 0191
PERCENT 0.05 GNAME 1
ESP 1 ITAM 1
STORM 1 P OF 0 1
PLAN 1 SPL 11/COMBINE SEQUENCE SSC
CHANNELS USED 01 02 03 04

GREEN NUMBER REPORT
REQUESTS DEVIATION REPORT
FEATURE SELECTION OPTION SEPARABILITY ONLY
CHANNELS ALL AVAILABLE
OUTPUT SEPARABILITY REPORT
CLASSIFICATION CATEGORIES ALL TO THE CATEGORY LEVEL
CHANNELS ALL AVAILABLE
OUTPUT CLASS SUMMARY REPORT, CLASS MAP TRACE TAPE
DEAD CONNECTION CATEGORIES-OF-INTEREST ARE ALL CANDIDATE CATEGORIES
DEAD CONNECTION UNIDENTIFIABLE CATEGORY IS 0

```

FIGURE 4.—Batch mode input cards and hardcopy scenario.

as warnings. It should be noted that a fatal error encountered in processing for one site does not stop the job; work continues with the next site.

Activity Tracing

The ERIPS, like any other complex system, must provide for activity tracing to verify that the intended processing path was taken or to show where deviations from the path occurred and how that affected the outcome. The major means of activity tracing in ERIPS are reports and logging, though various other products exist which are outside the scope of this discussion.

Many of the ERIPS functions generate reports, which are formatted for display on the conversational screen. (Figure 5 shows an ERIPS report.) During batch mode operation, these reports go to a microfiche tape. The interactive user can view them whenever he wants by use of the report special action keys (enter/exit report mode, page forward, and page backward) and the report menu. In either mode, and whether they are displayed or not, report pages are written to a log tape when they are formatted.

Also on the log tape are the various application menu displays at the times of significant change—that is, excluding partial input playback but including initial appearance, final input playback, and system-generated messages and dynamic data.

DETAILED CLUSTERING REPORT										
RUN NO.	CLUSTER 1		CLUSTER 2		CLUSTER 3		CLUSTER 4		CLUSTER 5	
	C OF 0/0	S FROM 0								
CHANNELS	MEAN	STD								
1	12.70	0.94	13.33	0.95	13.06	0.99	13.06	0.95	11.32	1.00
2	14.40	0.95	14.52	1.01	17.04	0.93	15.78	0.99	13.26	1.01
3	15.99	0.93	19.09	2.29	15.31	0.95	13.36	0.74	10.11	1.00
4	7.77	0.67	9.29	1.30	9.87	0.99	9.54	0.72	4.93	0.76
DIFF	0.32		1.15		0.32		0.01		0.16	
DISTANCE										
MEAN	2.99		4.13		2.14		2.46		2.98	
STDEV	0.99		0.99		0.99		0.99		0.99	

FIGURE 5.—ERIPS report.

When detailed activity tracing is desired, the log tape can be processed to create a hardcopy version of how the screen appeared (or would have appeared) at various times. (See figure 6 for a typical page of delog.) This processing is generally done for troubleshooting and development testing, not operationally.

CHECKPOINT/RESTART

The complementary functions of checkpointing and restarting the system are mainly automatic, involving the maintenance of disk data sets by ERIPS. Checkpoints are taken at predefined points in a program, such as upon entry to or normal exit from an application, and only one checkpoint disk data set exists for a terminal at any one time. The data saved for a terminal's checkpoint represents its complete environment, so the data can be retrieved to restart the terminal.

One user interface with the restart function is at sign-on time. As soon as the user signs on, a menu appears which asks whether to restart the terminal or disregard any restart data. If restart is requested, and the data exist, the user is in effect returned to the system environment at the time the last checkpoint was taken. This means that if, for instance, a terminal session is interrupted to acquire a full dump for use in debugging, the job can be restarted without having to recreate data.

The other user interface with restart is the writing of a restart tape, which can then be specified during a later sign-on as the source of the restart data. This capability, which is provided as a special function of

certain application menus, can be used to save restart data which would otherwise be lost when the next checkpoint occurred.

ERROR RECOVERY

The error recovery function is one of the most important and unusual features of ERIPS. Most complex software systems respond to serious error by abnormally terminating ("abending") the whole job, leaving the user with the need to start all over again (and maybe again and again as the same error is encountered in different disguises). Through its error recovery procedures, ERIPS drastically reduces the impact on the user.

When a serious system error is encountered, the abending process is intercepted by ERIPS software. If error recovery is desired, as it generally is, a partial dump is produced for debugging later. In interactive mode, the interrupted application is notified that recovery is needed and it restarts itself. Thus, if the user can deduce the cause of the error and avoid it in further processing, the session can proceed normally. In batch mode, the supervisor finds the inputs for the next site and proceeds from there.

```

CONSOLE
0 0 1 1 2 3 3 4 4 5 6 6 7 7 8 9 9 10
0 8 2 8 4 0 6 2 8 4 0 6 7 8 4 0 6 2
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-----
1008
984* 09009 MENU INPUT ACCEPTED * 23/25/11 1
990* 1839
995*
912* ITERATIVE INITIALIZATION MENU
888*
884* INPUT YOUR ITERATIVE PARAMETERS
840*
816*
792*
788* DEFAULT --- R2 8191
744*
720* CO. 60 NMIN2 5----
696*
672* PERCENT 0 95 ITMAX 1----
648*
624* SEP 1---- P OF N 1----
600*
576* S/DMAX 1---- MATCH THRS
562*
528* DLMIN 1---- SPLIT/COMBINE SEQUENCE
504* SSC
480*
456* | | L1 MATCH (IF SELECTED THE MATCH THRS
432* | | ) VALUE MUST BE AN L1 VALUE)
408*
384*
360*
336*
312* STARTING VECTORS
288* | | SELF GENERATING VECTORS
264* | | SUBCLASS STATISTICS VECTORS
240* | |
216* | | VECTORS FROM R/JN NUMBER
192* | |
168* | | EXTERNAL VECTORS
144* | |
120* | | DOTS AS STARTING VECTORS
096* | |
072* | |
048*
024*
000*
EOP
TOT

```

FIGURE 6.—Delog output from log tape.

The error recovery function can be turned off if more information is needed to solve the problem. In this case, the abending process goes to completion, producing a full dump and bringing down the job. Even then, of course, an interactive user can still make use of the restart data when he signs on again.

Error recovery can also be specifically forced by the user by enabling and depressing the "reset" special action key. This is generally done when the user realizes that he has accidentally started some process which cannot or should not complete (such as specifying the number of a read-only tape for a write operation).

CASE STUDY

To see how the man-machine interfaces work in practice, let us follow an interactive user through a terminal session. The purpose of this session is to define some fields based on an image screen display, to classify the image, and when the results look good, to note the field definitions and get a film product of the classification map. Figure 7 illustrates the menu flow.

First, the user signs on for a cold start (no restart data to be used) and enters the Pattern Recognition application. The image selection menu appears, and he selects the IM (Image Merge) special function box. The IM menu appears and asks for a site number and for the acquisition dates to be merged. For purposes of demonstration, suppose that the user has some trouble entering these data. He tries to type the first character of an acquisition date without having shown that the site number is finished, so he gets a terminal control message from the first level of error checking. After correcting this, he enters the dates and selects EOT, signaling the end of the inputs. Now the second level of error checking discovers that a required field (the name to be used in referencing this image) was omitted. This causes the appearance of a supervisor message; the user makes the correction and selects EOT. Finally, when the application software tries to retrieve the data from the Image Data Base, it discovers that one of the acquisition dates was invalid, and an application message is displayed. (The screen now looks like figure 3.) The user corrects this error and selects EOT; this time, the terminal control message says, "Menu input accepted." The image is merged, and the menu reappears in its initial state, ready for another image to be defined. In this case, the user selects the IMD (Image

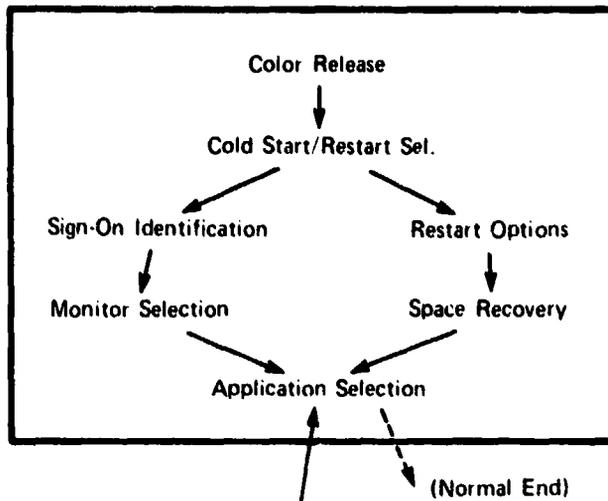
Manipulation and Display) special function box, uses the IMD menu to cause display of the image on his grayshade screen, then selects the RET (return) special function until the Pattern Recognition (PR) image selection menu reappears. The user then enters the image name, and an EOT causes display of the PR process selection menu.

The user chooses the Field Selection process and defines fields using the grayshade screen to connect the points he selects. Occasionally, the field definition is rejected because it has too many vertices (in which case the user can backspace, erasing lines so that a simpler field can be drawn) or because it is an illegal shape (for instance, it does not close and must be redefined). In each case, the appropriate messages are output. When all the desired fields have been defined, the user requests a Field Definition Report, which he then views via the report mode special action keys.

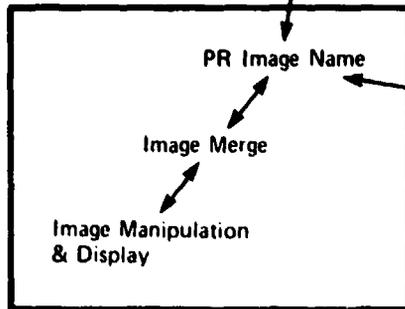
The next step should be the computation of statistics for these fields, but suppose the user forgets and requests classification. He soon gets a message that the statistics are not available; he backs up to request them, then comes back to classification. When he forgets to make his a priori value inputs on the appropriate menu and returns to it to input them, he encounters what turns out to be an application software error. His application abends (as reported by a message on his screen), an abbreviated dump is produced, and the PR process selection menu reappears. He tries classification again the same way, and again error recovery occurs when he retries the a priori menu. At this point, the user disables error recovery (which requires both the "enable" and the "switch recovery mode" special action keys, so that it will not happen accidentally) and goes through the sequence again. This time, the whole job abends when the error is encountered, creating a full dump. (Typically, the user would just write up the problem, and the programmer responsible for solving it would recreate the situation if the abbreviated dump was insufficient.) The user then asks the computer operator to feed the job in again.

When the user signs on again, he requests a restart using the existing data. The merged image, field definitions, and statistics are retrieved automatically, and the user can simply enter the image name for PR and select classification. (This time, he is careful to make his a priori inputs when the menu first appears.) Classification proceeds normally, and the user requests and views a classification summary report. Since everything looks good, he selects the class

Sign-On And Application Control



PR Initialization



PR Processing

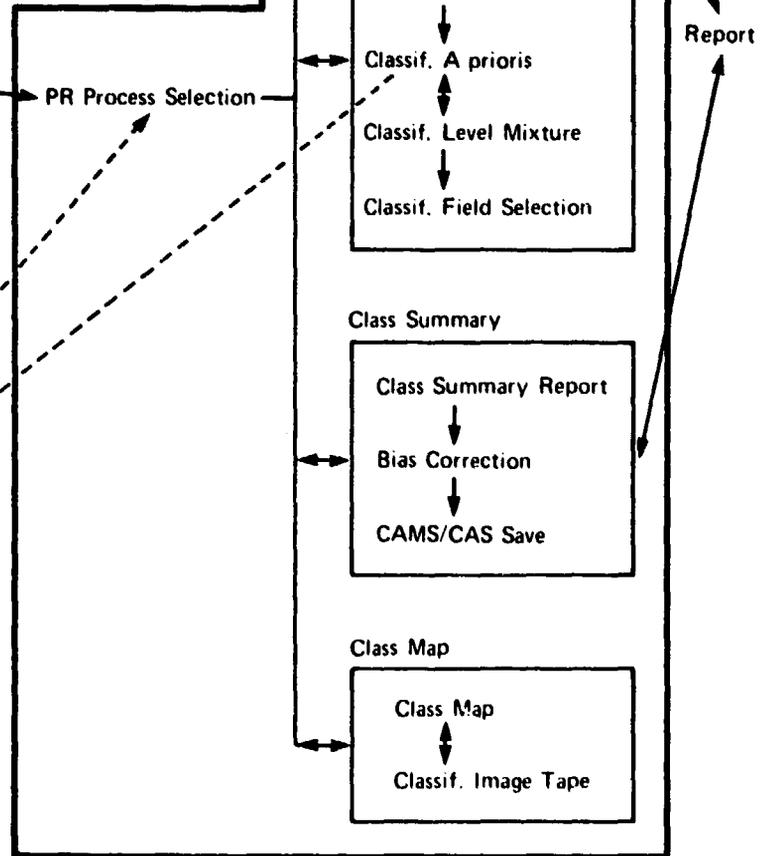


FIGURE 7.—Case study menu flow.

map function, generates a tape from which the film product can be produced, and finishes the job normally.

This case study, while it is by no means typical, has illustrated the way in which the user and the ERIPS software interact. At all times, the user has been given both the tools needed to do the work and the information needed to control the process. He has recovered, with very little pain, from his own errors and (even more significantly) from a software error.

SUMMARY

In summary, the major man-machine interfaces in ERIPS are

1. Menus to display system status and processing path options and to request necessary information
2. Checkpoint/restart to save results between terminal sessions, reducing reworks
3. Error recovery to minimize the impact of serious errors

In implementing these interfaces, several lessons of general interest have been learned. First, a highly interactive system can be made easy to modify by using the menu concept described here. The initial cost of such a scheme is high, since it must include generalized routines for menu definition, management, and data field input/output formatting. Once this has been done, however, the alteration of existing menus and the addition of new ones are simple. A bonus is the localization of a terminal-dependent code into a small set of routines, allowing software transportability and hardware upgrading.

Second, such a system can be compatible with batch mode operations. There are two keys to success here: careful selection of a set of significant input types and development of generalized software to merge the static and dynamic information and feed it into the system.

Third, provision for a batch mode of operation is not optional but imperative when frequently used functions involve a large number of interactive menus. A possible improvement in the current system would be a means by which an interactive user could indicate, at the beginning of each major process, whether the path assumptions made for batch mode operations are applicable. If so, the number of required interactions could be substantially reduced.

Finally, high-level compiler languages such as FORTRAN and PL/1, while they offer some advantages in ease of implementation, are only marginally compatible with error recovery as described here; in addition, the bulky modules they tend to produce can seriously interfere with multiterminal interactive use. We have found that an assembler language with macro capabilities (in our case, Assembler-360 with the High-Level Assembler Language (HLAL) structured programming macros) is almost as easy to implement and avoids these problems (ref. 3).

Throughout the ERIPS system, a primary concern is to require minimum input from the user after supplying him with maximum information, while being as forgiving as possible of human error. Success at this goal plays a major role in the success of LACIE as a whole, since the usefulness of a system depends largely on its usability.

REFERENCES

1. Display Formatting System. Terminal and Data Base Management Operating System, Vol. 1 (User's Guide), Sec. 3, IBM Federal Systems Division, Houston, Texas, Oct. 1971.
2. Large Area Crop Inventory Experiment (LACIE) Programmer's Guide. IBM Federal Systems Division, Houston, Texas, May 1975.
3. High Level Assembler Language User's Guide. IBM Federal Systems Division, Houston, Texas, Mar. 1972.

LACIE/ERIPS Software System Summary

C. L. Johnson^a

ABSTRACT

The Earth Resources Interactive Processing System (ERIPS) software supports the Large Area Crop Inventory Experiment (LACIE) with the analysis of agricultural data sensed by the Landsat spacecraft. Its primary function is the classification of the data on the basis of statistical similarity to those portions which have been identified by analysts. Since its original definition in 1971, ERIPS has been used to develop analysis tools related to that process. This is a summary of the development and capabilities of the ERIPS software system.

INTRODUCTION

The ERIPS software was developed by the IBM Federal Systems Division, Houston, Texas, to support NASA in its Earth resources activities. LACIE/ERIPS executes on an IBM 360/75 mainframe with an attached Goodyear STARAN S-500 special-purpose processor (SPP) at the NASA Johnson Space Center (JSC) Real-Time Computer Complex (JSC Building 30) in Houston. It is used to process the Landsat data to estimate the wheat growing area in several countries. From these estimates, the analysts develop their production predictions.

The LACIE/ERIPS is a large program (approximately 240 000 lines of code on the 360/75 and 19 470 on the SPP) the development of which actually began in 1971, on the basis of the algorithms used by the Purdue University Laboratory for Applications of Remote Sensing System (LARSYS) (ref. 1); it was not associated with LACIE until 1974. During its development, ERIPS has evolved from an interactive system used as a research tool into a system that is primarily used noninteractively on a production mode basis. Many capabilities have been added to the original ones, and significant lessons

have been learned about the implementation of this type of system.

SYSTEM REQUIREMENTS

Originally, ERIPS was planned as a local version of LARSYS which would use available computer and display hardware (refs. 2 to 4). LARSYS was chosen as the base system because it was a powerful, operationally proven system for the analysis of remotely sensed multispectral data. Among its capabilities were

1. Statistical computations of means, standard deviations, covariance matrices, and correlation matrices for data classes
2. Separability measurements for distributed classes and the use of these measurements to select features (spectral channels) for further processing
3. Classification of data by a Gaussian maximum likelihood algorithm
4. Performance evaluations of classification results

In addition to requiring these capabilities, the ERIPS definition called for use of the available Display Retrieval and Formatting Technique (DRAFT) digital television equipment (DTE) terminals as image analysis stations, supporting independent users concurrently. ERIPS would also provide an error recovery capability to reduce the impact of software failures (see the paper by Duprey entitled "Man-Machine Interfaces in LACIE/ERIPS"). Finally, ERIPS was required to operate in a multijobbing environment, since the Real-Time Computer Complex resources were also needed for development of manned space-flight programs (Skylab and Apollo-Soyuz). This latter requirement led to a budget of 250 kilobytes for ERIPS, in a 300-kilobyte region to ensure some future flexibility.

Investigation of this system definition showed that it would not be cost effective simply to modify LARSYS. First, the existing LARSYS was programmed in FORTRAN. The FORTRAN language is

^aIBM Federal Systems Division, Houston, Texas.

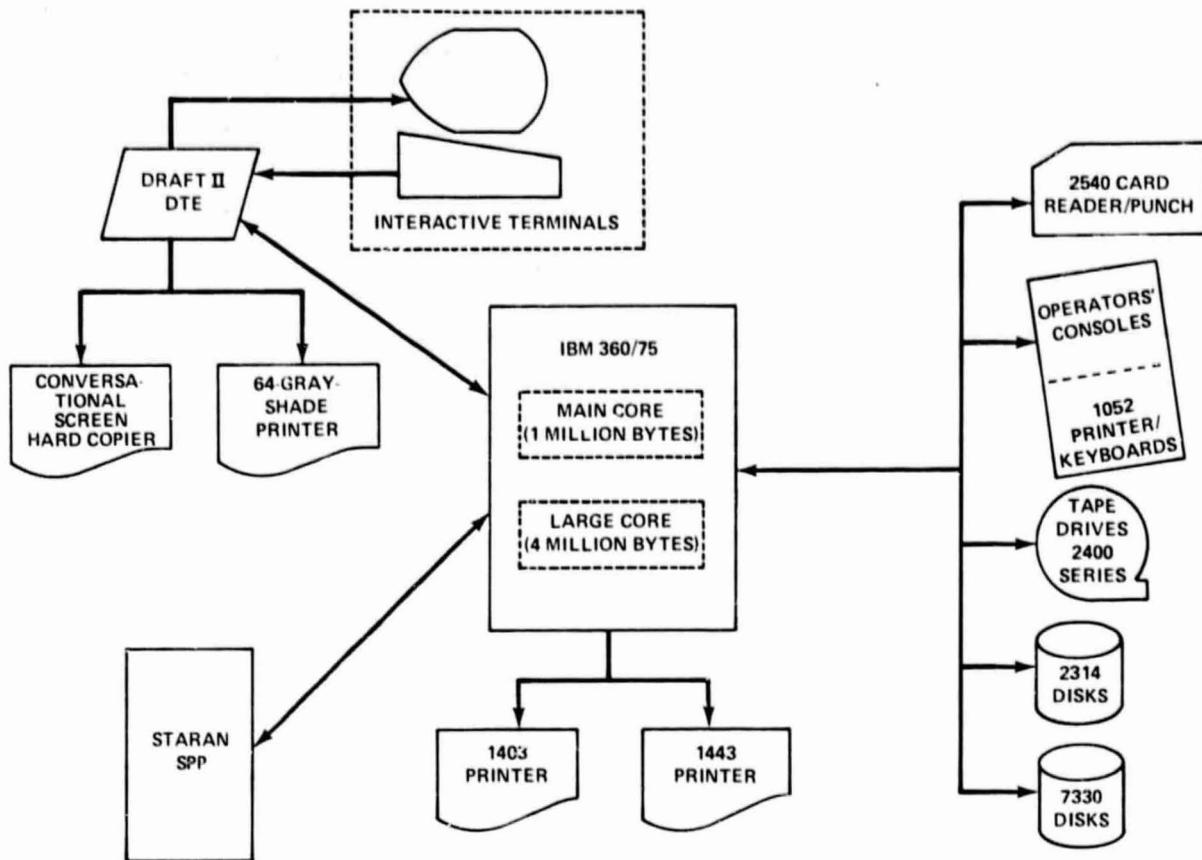


FIGURE 3.—LACIE/ERIPS hardware configuration.

frame image of data from Landsat, takes less than 8 minutes.

Other components of the system besides the SPP have had a significant effect on LACIE/ERIPS development. For instance, careful management of the allocation of core is required. First, the speed at which large core operates is only about one-third that of main core. This feature can cause high-access applications like class summary and feature selection to consume far too much time unless their storage requirements are taken almost entirely from main core. Second, there is always the possibility of fragmenting the available storage into noncontiguous segments that cannot be used to satisfy a core request. Finally, both terminals can be concurrently executing a large application. This condition becomes even more significant if more terminals are to be supported, as is intended for the future. It is very helpful to have a monitoring function, like the advanced statistics collector utility used with LACIE/ERIPS. This utility runs in the background, recording information about storage requests and so



FIGURE 4.—LACIE/ERIPS interactive terminal.

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR



FIGURE 5.—LACIE/ERIPS keyboard.

on. Analysis of these recordings can provide warning of impending problems so that corrective action can be taken.

The terminals have also been important in LACIE/ERIPS development. Since they have very little local intelligence, the ERIPS software includes a considerable amount of overhead for display processing. Each user action (typing a character for a menu field, for instance) must be interpreted by ERIPS. The software must maintain all the data being displayed on the conversational and image screens, including any which have been temporarily displaced so the screen could be used for another purpose (like viewing a report).

SYSTEM FUNCTIONS

The LACIE/ERIPS software is composed of a supervisor and various application programs. This modular structure enables easy addition of new applications as the system grows. LACIE/ERIPS executes under a locally modified OS/MVT (operating system/multitasking with variable number of tasks). Originally, this was the Real-Time Operating System, which has significant modifications and locally added features required to support manned space-flight activities. To meet its transportability goal, ERIPS did not use any of the real-time extensions to the standard operating system; LACIE/ERIPS now operates under the Extended Operating System, which has fewer local modifications.

The supervisor is a sophisticated executive routine that controls all system activities. It provides the interfaces between the user, his equipment, and the application programs. It ensures minimal impact when adding applications to LACIE/ERIPS. Basic

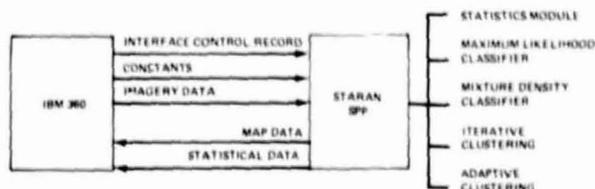


FIGURE 6.—Host-SPP overview.

system initialization, error recovery, and computer sharing by application programs are the system functions supplied by the supervisor. The supervisor isolates the application routines from the characteristics of the external hardware devices. This hardware independence has proved to be efficient in transporting ERIPS to other locations with different equipment. Versions of ERIPS have been installed on IBM 370/158 and 370/168 models for the Earth Resources Laboratory (ERL/ERIPS) and the ERMAN project, using different configurations of disks and terminals.

Some common services are located in the supervisor and used by each application that needs them. These services include menu control, image display to screen, image data access service, dynamic space and device allocation, data logging and delogging, and batch input control.

To give the application programs the capability to read image data from disk, the supervisor incorporates a specialized high-speed image data access method. This method is a very efficient storage space manager, providing direct retrieval of the imagery data from the disk (see ref. 10).

Application programs perform various image processing functions. The ultimate function is to group statistically analyzed imagery data according to specified criteria. Before this process, the image is made available and prepared for the user as needed.

The load application handles image data generated from various sensors. It is designed to accept data from tapes in the LARSYS, Landsat multispectral scanner (MSS), or universal formats and put those data into a standard format for use by the other applications. The image data from these tapes can be put on the system disk packs. In addition to imagery loading, load can scroll a tape-resident image directly onto an image screen, unload a disk-resident image to tape, and display reports on the contents of the image tapes.

Image manipulation and display (IMD) is the application for image manipulations and image displays. This application can display an image on a display screen in a maximum of 16 shades of gray or 8

or 64 colors. The user can view a currently loaded image by supplying information such as image name, first pixel, initial line, and channel number of the data to be displayed. IMD places one image line on the screen at a time beginning at the top, shifting down one line as a new line appears. This procedure continues until the screen is full. Since most images are larger than the amount of data allowed on the screen, the scroll capability allows the additional image data to be viewed. Image data lines are added to the top of the screen line by line when the SCROLL key is depressed. The IMD application provides other services such as an Available Image Report, which displays the names and characteristics of all images resident on disk, and a Latitude/Longitude Report, which gives the latitude/longitude coordinates for the image elements.

The registration application enables user control of the geographic relationship of picture elements (pixels) within an image. It provides two capabilities. First, registration is capable of conforming an input image of a given scene to a reference image of the same scene. Second, it maps an input image onto a predefined latitude/longitude grid. These capabilities are important for several reasons. Registration can be used to remove image distortion introduced by the remote sensor and by the curvature of the Earth. Moreover, two images of the same scene produced by two entirely different sensor devices can be registered together. This capability makes it possible to correlate data from a satellite image with data from an aircraft scanner. Images of the same scene produced at different times can be registered together to permit multitemporal analysis of the scene.

For image-to-image registration, the user identifies one of the two loaded images as the reference image, the other as input. The input image is mapped to conform to the reference image. The user then selects a point on each image which he has determined is common to both. An identification (ID) number is assigned to the point pair. This step is repeated for the desired number of points. Once the point set is determined, map generation processing takes place.

In image-to-Universal Transverse Mercator (UTM) registration, the user defines the geographic boundary of the output image desired. Each picture element is assigned a latitude/longitude. The map generation process takes place after all assignments are made for the desired number of points.

Once the user has decided to generate a mapping

polynomial, the processing for image-to-image and image-to-UTM registration is similar. The image position of each input point is compared to the corresponding reference point. The least squares coefficient for a bivariate N th-order polynomial is calculated. In image-to-UTM registration, the reference point is a geographic coordinate. (See the section entitled "Registration" in the appendix for the algorithm description.)

The image creation application provides the user with three different methods of creating an image: image composition, image difference, and image merge.

Image composition allows the user to combine two images containing the same number of lines and pixels into one image. The most common use of this application is to combine two registered images into one. This combined image can then be used for experiments in multitemporal analysis.

Image difference allows the user to take the difference of two images and form a third image. The two images must contain the same number of lines and pixels. For example, if two four-channel images of the same scene, taken at two different times, were differenced, the resulting image would be the area that changed during the timelag.

Image merge allows the user to juxtapose as many as four separate images with different acquisition dates to create one new image. Resident images from the image data base which are 117 lines and 196 pixels in size are used in the process. The resulting image can be as many as 468 lines long with 196 pixels per line. The channels in this image are numbered consecutively beginning with 1. This capability was not available until the LACIE phase of ERIPS. LACIE/ERIPS also has a delog application which allows the user to receive a printed copy of the menus and reports generated during his terminal run.

Remaining to be discussed is the pattern recognition application. It utilizes numerous programs to acquire classified image data, and, because of its size and its importance to the LACIE/ERIPS system, a separate section is devoted to it.

THE PATTERN RECOGNITION CONCEPT

Pattern recognition is the largest application in the system. Its function is to classify the picture elements (pixels) or group elements (fields) of an image into classes. Several processing steps are performed before the actual classification. First, the user

analyst defines areas of the loaded image to be processed and identifies the materials belonging to each area. These materials can be of an agricultural nature (such as wheat, corn, or soybeans) or any other type of ground data (trees, roads, water, etc.). After the fields of an image have been established, statistical analysis can be done and the selected area of the image classified. If the analyst finds the results unsatisfactory, he may return to any point of the pattern recognition process to redefine and/or recompute the data. The analyst can also choose an unsupervised type of classification without having to train the classifier. The clustering algorithm examines all data elements in the area to be classified and assigns those elements that are spectrally similar to the same class or cluster. The user input parameters control the processing and specify the degree of closeness required. The output clusters can then be used by classification.

The subapplications in pattern recognition perform separate functions which together accomplish the pattern recognition task. Many of the algorithms were taken from LARSYS, but other features and algorithms have been added.

The field selection subapplication gives the user the option of determining fields from an image and assigning attributes to these fields. A field can be defined with a minimum of 2 vertex points (a 1-dimensional line field) and a maximum of 10 vertex points. Field vertices are entered on the image screen via cursor or are typed onto the conversational screen as line/pixel values. The user can view all current field definitions by displaying the Field Definition Report.

The statistics program computes the means, standard deviations, and covariances for each class defined in the system that contains at least one training field. It also performs other statistical manipulations, such as the combining of several classes' statistics into one class, Sun angle correction, mean level adjustments, deletion of class or field statistics, reassignments of fields from one class to another, and the changing of a field's status. (See the sections entitled "Statistics," "Sun Angle Correction," and "Mean Level Adjustment" in the appendix for algorithm descriptions.)

Clustering is a method for grouping data into homogeneous sets. In LACIE/ERIPS, the clustering subapplication partitions a collection of pixels into subsets which have similar spectral signatures. The primary uses of clustering are to assist in defining the boundaries of fields, to evaluate fields according to

homogeneity of data, to collect homogeneous data for fields from nonhomogeneous areas, and to act as a nonsupervised classifier of multispectral data. Two algorithms are used to achieve the results, adaptive and iterative. The user has the option to use the one-pass adaptive algorithm or the multipass iterative algorithm or a sequential combination of the two. (See the sections entitled "Adaptive Clustering," "Iterative Clustering," and "Clustering Report Functions" in the appendix for algorithm descriptions.)

In feature selection, an optimal subset of channels with which to classify an image can be determined. This feature selection process utilizes a separability measure involving the Bhattacharyya distance. Classification time can be greatly lowered with this reduction of dimensionality of the data, although use of the SPP for classification has made this characteristic far less important. The optimal channel subset retains a significant percentage of the separability inherent in all channels of the image. After every execution of feature selection, a resultant best-channel subset is made available to classification. The user has the choice of several processing paths, including the original ERIPS divergence function. (See the sections entitled "Divergence" and "Feature Selection" in the appendix for algorithm descriptions.)

The classification processor in LACIE/ERIPS assigns each pixel of a given field to the candidate class the statistics of which that pixel most nearly represents. Classification is now done in a mixed environment of likelihood density functions, with some summations performed to the class level and others to the category level. (See the sections entitled "Classification (ERIPS)" and "Classification (LACIE)" in the appendix for algorithm descriptions.)

Pattern recognition produces outputs representing the results of the various programs. The report outputs are in the form of imagery data, graphic data, or digital data.

Classification maps and cluster maps are outputs of the classified and clustered image, respectively. Each data element is represented by a symbol on a character map. Each element can also be represented by a gray shade or a color. The gray level or the color is associated with the class that the element has been classified or clustered into.

Other major reports produced by pattern recognition are Bias Correction, Spectral/Trajectory Plots, and Green Number. Bias correction computations are done for classification results and/or clustering

results. (See the section entitled "Bias Correction" in the appendix for algorithm description.) Plots are produced for spectral plots, tabulations, and trajectory plots. The spectral plots provide a two-dimensional plot of the dots per acquisition with their classification symbols and user-assigned labels. The first two Kauth/Thomas coordinates, greenness and brightness, are the axis values. The two tabulations show as many as four acquisitions per chart, ordered by dot number and by the first Kauth/Thomas coordinate (greenness value). The trajectory plot is a plot of the Kauth/Thomas coordinates for each dot in each of the acquisitions. The Green Number Report displays the Kauth/Thomas green number (greenness minus average soil greenness) associated with clusters and dots. It is used by Classification and Mensuration Subsystem (CAMS) analysts to monitor wheat emergence and drought conditions. A greenness and brightness value is displayed for each cluster and/or dot for all acquisitions that are processed.

The LACIE/ERIPS provides several independent subsystems that utilize imagery data and application results. The CAMS/Crop Assessment Subsystem (CAS) interface subsystem¹ gives the user the capability of using pattern recognition results on other systems that can receive the tape inputs. Input data as well as data generated from composition and indexing, statistics, feature selection, clustering, and classification are collected onto disk and then saved on tape. These interface tapes are designed to be as compatible with other systems as possible. The data on the tapes is in American Standard Code for Information Interchange (ASCII) format, with record lengths of 720 bytes.

With the various applications processing large amounts of data, the need for operating speed became a primary concern in LACIE/ERIPS. There were hardware and software constraints which governed time required to complete execution. The addition of the Goodyear special-purpose processor as a parallel processor significantly reduced the computer time required for LACIE/ERIPS operations. To support this function, hardware and software interfaces between the computers had to be satisfied (ref. 11). Five software modules were implemented in the SPP to interface with the software in the IBM 360: the statistics processor, the maximum likeli-

hood classifier, the mixture density classifier, iterative clustering, and adaptive clustering (fig. 6). Five distinct logical units of input/output (I/O) data are specified by these applications for communication between the host (IBM 360) and the SPP. Three logical data units are transferred from the host to the SPP: the Interface Control Record, the input parameters, and the input vectors (imagery data). The remaining two logical units, the output vectors and the output parameters, are transferred from the SPP to the host.

The result of this system interaction was a substantial reduction in total execution time. Improvements in performance times for classification and clustering were most important (ref. 12).

LACIE/ERIPS DATA BASES

The LACIE/ERIPS uses several data bases as storage facilities for the large volumes of information needed to support the system functions. The largest system update came during the 1976 crop year. It contained six functional Information Management System data bases: history, image, fields, process control, status tracking, and results. These data bases eliminated the previous ones developed in support of crop year 1975. During that time, image tapes from the NASA Goddard Space Flight Center (GSFC) were first introduced. GSFC tapes were multifile, universal format tapes containing preprocessed imagery data acquired from Landsat-1 scanners. Since many tapes were needed to support LACIE/ERIPS, the concept of an image and a field data base was developed to handle the data on the GSFC tapes. The system stored and processed data collected for 4193 sample segments with an average of 4.5 acquisitions per segment. A sample segment corresponds to a ground area of about 6 by 5 nautical miles, or 117 lines by 196 pixels.

The history data base contains sample segment identification, GSFC controlling information, and acquisition history. The identification and GSFC controlling information includes the sample segment ID, type, country, crop type, biological window, film flags, and color codes. The acquisition history includes data quality information, a tape index, and an image data base index. This data base is updated by the composition and indexing subsystem. It contains a maximum of 7 million bytes of data.

The image data base contains the data on the GSFC imagery tapes for an entire growing season.

¹CAMS/CAS Interface Tape Interface Control Document JSC-09866, Johnson Space Center, Houston, Texas, Feb. 1978.

Both header and imagery are stored so that images can be reconstructed into the universal format when unloading to tape. A maximum of 4 four-channel images is stored for an ordinary sample segment, and a maximum of 16 four-channel images is stored for training and intensive study sample segments. The image data base consists of many physical data bases because of its size, which is 2.9 billion bytes of data, maximum. Composition and indexing also maintains this data base.

The field definitions for the LACIE/ERIPS sample segments are stored on the field transaction data base. Information stored on this data base describes field locations, field types, and the category and class and/or subclass associated with each field. The data base also contains default a priori and threshold values. All data can be retrieved during a LACIE execution on a per-segment basis to provide data for statistics computations and classification processing. The field transaction data base subsystem consists of the physical data base along with the programs necessary to create, update, maintain, and report on this data base.

The process control data base contains information which defines the processing to be done in batch mode on each site. These data are written out to the data base by the process control subsystem and deleted after the site is processed by the batch production system. The batch production system reads the process control data base and generates a series of simulated menu inputs in the manner specified by the process control information. These inputs are then passed to the LACIE/ERIPS supervisor for input to the pattern recognition application. The maximum size of this data base is 1.6 million bytes.

The status tracking and mensuration results data bases are not maintained. After their development, it was determined that they did not meet the design goals set for them. Status tracking was developed to store a history or tracking of the LACIE/ERIPS production jobs and the products used or produced by the system. The mensuration results data base was designed to contain the results obtained each time a site was processed through the interactive or batch system.

The most recent data base to be developed, the dot data base, contains the data for each fixed set of 209 pixels or dots for each sample segment. (Dots represent every tenth pixel on every tenth line of a LACIE image.) Information stored on this data base indicates the location of the dot, the category, and the usage. Dots are used in pattern recognition as

starting vectors, labeling vectors, or bias correction vectors.

CONCLUSION

The ERIPS has undergone many changes since its original implementation in 1972, as can be seen in figures 1 and 2. During the process, much experience has been gained in handling large volumes of data and providing analysis aids.

A fundamental design conclusion reached very early in the process was that an assembler language with macrocode capabilities, particularly when combined with a preassembler that recognizes structuring macrocodes and processes them before assembly, has almost all the advantages and none of the drawbacks of the compiler languages. Both FORTRAN and PLI were eliminated from consideration as primary implementation languages, mainly because of the large module sizes they produced. Also, in these languages, the interfaces with the nonstandard I/O packages, error recovery, and the use of large core storage capacity are cumbersome. Thus, the use of these languages has been restricted to certain computational subroutines (FORTRAN) and formatting-intensive routines (PLI); the rest of ERIPS is coded in High Level Assembler Language (HLAL).

Another basic design decision which has proved very successful was the development of the nonstandard I/O packages. The image direct access method provides efficient access to the multispectral imagery data on whatever basis it happens to be needed—whole segments, specific channels, particular image lines, even line-skipping and pixel-skipping patterns. The extended access method provides the protocol required for communication with the terminal hardware. Both packages have been successfully transported with ERIPS to other mainframe/disk/terminal configurations. ERL/ERIPS, for instance, has been implemented on several different models of IBM 360 and 370, under various operating systems, using 3330 and 3330-11 disks, and communicating with Ramtek terminal hardware (ref. 13).

During the life of the system, not all the application changes have been related to implementation of new functions. Several applications have disappeared, fallen into disuse, or become of less importance. Bhattacharyya chaining, for instance, which used a Bhattacharyya distance function to form

"chains" of clusters having similar distance characteristics, proved useless for analytic purposes. The divergence technique for reducing the number of channels to classify was largely replaced by feature selection, which has itself been eliminated from the common processing path, since the time it consumes outweighs the time saved in classification with use of the SPP. Feature selection has thus become primarily a reporting tool the analyst can use to refine his inputs to the rest of the system. Mensuration has disappeared from ERIPS, as has status tracking. All this is further proof of the need for modularity of system structure, which in ERIPS is a natural consequence of the menu concept.

An important shift of emphasis in the mode of processing has also occurred. Originally, ERIPS was purely an interactive system. By 1975, however, a noninteractive mode had become imperative, and most of the system's work is now of this type.

Though the experimental phase of LACIE is now over, with transition to a multicrop analysis requirement beginning, the history of ERIPS gives some clue to the future of large-scale image processing systems. There has been a continual development of ways in which the computer can be used to provide analysis tools, and there is no reason to believe that the last such tool has been found.

REFERENCES

1. Landgrebe, D. A.: LARSYSAA: A Processing System for Airborne Earth Resources Data. LARS Information Note 091068. Purdue University.
2. Felix, M. R.: The Earth Resources Interactive Processing System. IBM Federal Systems Division, Houston, Texas, 1974.
3. Detchmندی, D. M.; and Wylie, A. D.: RTCC Requirements for Implementation of a Clustering Program in the ERIPS. NASA Johnson Space Center, Houston, Texas, Jan. 1973.
4. ERIPS Program Documentation. NAS 9-996, IBM Federal Systems Division, Houston, Texas, 1972.

5. LACIE-2 Design Review. NAS 9-14350, IBM Federal Systems Division, Houston, Texas, Nov. 1974.
6. LACIE-6 Detailed Design Specification. NAS 9-14350, IBM Federal Systems Division, Houston, Texas, 1977.
7. LACIE-7 Detailed Design Specification. NAS 9-14350, IBM Federal Systems Division, Houston, Texas, Oct. 1977.
8. Requirements for a Large Area Crop Inventory Project. NASA Johnson Space Center, Houston, Texas, 1974.
9. LACIE Baseline Document, Vol. 1. NAS 9-14350, IBM Federal Systems Division, Houston, Texas, 1975.
10. Pape, A. E.; and Truitt, D. L.: ERIPS Image Data Access Method (IDAM). IBM Federal Systems Division, Houston, Texas, 1976.
11. Interface Control Document NASA/STARAN S-500. NAS 9-14591, Goodyear Aerospace Corporation, Akron, Ohio, Feb. 1976.
12. Wylie, A. D.: Independent Software Verification of LACIE IV System. NASA Johnson Space Center, Houston, Texas, May 1976.
13. ERL/ERIPS Summary. IBM Federal Systems Division, Houston, Texas, Sept. 1975.
14. LACIE Program Documentation, Vol. 1. NAS 9-14350, IBM Federal Systems Division, Houston, Texas, 1975.

ACKNOWLEDGMENTS

Many people have contributed significantly to the success of LACIE/ERIPS. Though it is not practical to recognize each individual, we would like to acknowledge the following people as key contributors to the program.

NASA JSC

J. Leroy Hall
John C. Lyon
Alan D. Wylie
John L. Parker
Thomas G. Price

Other

George A. Austin

IBM

Bernard H. Berman
Richard M. Morrison
John A. Hooten
Gordon E. Myers
David King
Maurice Carr
Curt W. Blessitt
James P. Johnson
Michael R. Felix

Appendix

LACIE/ERIPS Algorithm Descriptions

The descriptions of algorithms used in LACIE/ERIPS contained in this appendix are from reference 14.

STATISTICS

1. Field means for a channel i and training field f .

$$M_{f,i} = \frac{1}{N_f} \sum_{k=1}^{N_f} X_{f,k,i}$$

where $M_{f,i}$ is the mean of channel i data in field f , $X_{f,k,i}$ is the k th pixel value for channel i in field f , and N_f is the number of elements in f . Means are computed for each channel in the image.

2. Field covariance matrix element. An element of the covariance matrix for field f represents the covariance between a pair (i, j) of channels for data taken over all elements of f . These elements are computed

$$\Gamma_{f,i,j} = \frac{1}{N_f} \sum_{k=1}^{N_f} (X_{f,k,i} - M_{f,i})(X_{f,k,j} - M_{f,j})$$

3. Standard deviation. The standard deviation $\sigma_{f,i}$ of channel i for field f is given by

$$\sigma_{f,i} = (\Gamma_{f,i,i})^{1/2}$$

4. Correlation matrix. The correlation matrix (normalized covariance) element is computed for each covariance element for all fields:

$$C_{f,i,j} = \frac{\Gamma_{f,i,j}}{\sigma_{f,i} \sigma_{f,j}}$$

5. Class means.
Adding field statistics to a class

$$M_n = \left(\frac{P_f}{P_n}\right) M_f + \left(\frac{P_c}{P_n}\right) M_c$$

where P_f is the population of the field, P_n is the population of the resulting class, and P_c is the population of the current class.

Deleting field statistics from a class

$$M_n = \left(\frac{P_c}{P_n}\right) M_c - \left(\frac{P_f}{P_n}\right) M_f$$

6. Class covariances.
Adding a field to a class

$$\Gamma_n = \left(\frac{P_f}{P_n}\right) \left[\Gamma_f + (M_n - M_f)(M_n - M_f)^T \right] + \left(\frac{P_c}{P_n}\right) \left[\Gamma_c + (M_n - M_c)(M_n - M_c)^T \right]$$

Deleting a field from a class

$$\Gamma_n = \left(\frac{P_c}{P_n}\right) \Gamma_c - \left(\frac{P_f}{P_n}\right) \left[\Gamma_f + (M_c - M_f)(M_c - M_f)^T \right] - \left[(M_c - M_n)(M_c - M_n)^T \right]$$

DIVERGENCE

The divergence calculation utilizes a class distance measure found in Kullback's "Information Theory

and Statistics." For classes x and y in an n -channel environment,

$$D(x,y) = \sum_{i=1}^n \sum_{j=1}^n \left\{ \left[\Gamma_{x(i,j)} - \Gamma_{y(i,j)} \right] \cdot \left[\Gamma_{y(i,j)}^{-1} - \Gamma_{x(i,j)}^{-1} \right] + \left[\Gamma_{x(i,j)}^{-1} + \Gamma_{y(i,j)}^{-1} \right] \cdot \left[M_{x(i)} - M_{y(i)} \right] \left[M_{x(i)} + M_{y(i)} \right] \right\}$$

FEATURE SELECTION

Given subclasses i and j with mean vectors M_i and M_j , covariance matrices Γ_i and Γ_j , and a priori values q_i and q_j , the Bhattacharyya distance between subclasses i and j is given as

$$B_{ij} = \exp \left[\frac{1}{4} M_{ij}^T (\Gamma_i + \Gamma_j)^{-1} M_{ij} - \frac{1}{2} \ln \left(\frac{|\Gamma_i + \Gamma_j|}{2^n |\Gamma_i| |\Gamma_j|} \right) \right]$$

where $M_{ij} = M_i - M_j$ and n = number of channels used. The associated separability between subclasses i and j is

$$S_{ij} = (q_i q_j)^{-\frac{1}{2}} B_{ij}$$

CLASSIFICATION (ERIPS)

Maximum likelihood classification evaluates the quadratic form

$$C(A)_{k,i} = 0.5 \ln |\Gamma_k| + \left\{ (\mathbf{x} - M_k)^T (\Gamma_k^{-1}) (\mathbf{x} - M_k) \right\}$$

where $|\Gamma_k|$ is the determinant of the covariance for class k , Γ_k^{-1} is the inverse covariance matrix, \mathbf{x} is the

vector of pixel values, and M_k is the mean vector for class k .

CLASSIFICATION (LACIE)

The density function used in the assignment decision for classifying pixels to a subclass involves the following contribution from a given subclass

$$P(\mathbf{X}, k, i, j) = \left(\frac{q_i}{N_k C_i} \right) (2\pi)^{-\frac{d}{2}} |\Gamma_{k,i,j}|^{-\frac{1}{2}} \cdot \exp \left[-\frac{1}{2} (\mathbf{x} - M_{k,i,j})^T \Gamma_{k,i,j}^{-1} (\mathbf{x} - M_{k,i,j}) \right]$$

where \mathbf{X} is the vector representation of the pixel to be classified, k is the category identifier, i is the class identifier, j is the subclass identifier, q_i is the a priori fraction corresponding to category k , N_k is the number of classes in category k , C_i is the number of subclasses in class i of category k , d is the dimension of pixel vector \mathbf{X} , $M_{k,i,j}$ is the mean vector of subclass (k,i,j) , and $\Gamma_{k,i,j}$ is the covariance matrix of subclass (k,i,j) .

In the nominal default classification to the category level, the following sum is computed for each category k :

$$P(\mathbf{X}, k) = \sum_{i=1}^{N_k} \sum_{j=1}^{C_i} P(\mathbf{X}, k, i, j)$$

The pixel is then assigned to the subclass within that category which has the maximal $P(\mathbf{X}, k, i, j)$.

The subclass assignments are saved, together with the following associated likelihood value:

$$\min \{ A \cdot (\mathbf{x} - M_{k,i,j})^T \Gamma_{k,i,j}^{-1} (\mathbf{x} - M_{k,i,j}) - 255 \}$$

where A is a preset tabular conversion factor.

For any category k for which classification to the

class level has been specified, N_k sums of the form

$$P(\mathbf{X}, k, i) = \sum_{j=1}^{C_i} P(\mathbf{X}, k, i, j)$$

would enter the overall consideration for largest density sum as stated previously. The specific subclass assignment and associated likelihood value would be derived in the same manner.

REGISTRATION

Two bivariate polynomials relate the points in a reference (output) image (X, Y) to corresponding preimage points in an input image (U, V) :

$$U = F(X, Y) = \sum_{i=0}^n \sum_{j=0}^n a_{ij} X_i Y_j$$

$$V = G(X, Y) = \sum_{i=0}^n \sum_{j=0}^n b_{ij} X_i Y_j$$

The least squares method is used to compute the coefficients a_{ij}, b_{ij} from the set of user-designated points of coincidence in the input and reference images.

ADAPTIVE CLUSTERING

The adaptive clustering algorithm generates clusters by cycling once through all fields to be clustered, assigning pixels in small homogeneous strips to cluster centers, which are continually being modified by statistically merging the pixel strips into the clusters. After analyzing all pixels, the cluster mean vectors are frozen and the data are passed

again to assign the pixels to these fixed cluster centers generating the final statistics.

1. Strip formulation. If $V_j(i)$ equals the i th component of the j th vector to be assigned, and S is a strip refinement parameter, then the local group or strip is defined by vectors $V_{j+k}, k = 0, 1, \dots, L$, where L is the last S for which

$$|V_j(i) - V_{j+k}(i)| < S$$

is valid for all values of i . After generating the local subgroup, its mean is computed.

2. Sequential search.

$$\left. \begin{array}{l} R \\ R1 \end{array} \right\} \text{refinement parameters}$$

The sequential search computes the distance between the mean of the local subgroup and each of the cluster means. The search terminates whenever this distance is less than $R3 = M1 \cdot R1$ ($0 < M1 \leq 1$), where $M1$ is a control parameter. The cluster means are searched in the order of their populations. Three outcomes are possible: (a) the subgroup is assigned to the first cluster for which the distance is less than $R3$; (b) the subgroup is assigned to the nearest cluster when the distance to the nearest cluster is greater than $R3$ but less than R ; or (c) the subgroup is used to begin a new cluster; that is, the distance to the nearest cluster is greater than R . After assignment of the strip—cases (a) and (b)—the mean and population count are updated.

3. Cluster merging. The cluster merging process operates by computing the distance between the nearest pair of cluster means. If this distance is less than a threshold C , then the two means are averaged into one. The nearest distance between clusters is recomputed, and the merging process continues until all the clusters are separated by C or more. The merging operation is performed when the counter NMC of the number of clustered points since the last merger exceeds the threshold NMT (system parameter).

4. Deleting clusters. The test for deleting clusters is made when the counter NEC exceeds the threshold NET (system parameter); NEC is the number of clustered points since the last deletion process. All clusters with less than $NMIN1$ points (system parameter) are deleted.

ITERATIVE CLUSTERING

The iterative clustering algorithm generates clusters by cycling a variable number of times through all fields to be clustered. On each pass through the data, the cluster centers are fixed and all pixels now in small homogeneous strips are assigned to the fixed clusters on a closeness basis. The pixels assigned to a certain cluster form new statistics for the cluster to be held as fixed for the next pass. Only between passes may the number of clusters fluctuate. Fewer clusters may result if "combine" logic is exercised; more if "split" logic is performed. Either split or combine is done after each pass until the last pass which determines final statistics.

1. Cluster splitting. In splitting a cluster, the channel with the largest variance (σ_j^2) is determined. If the standard deviation σ_j exceeds the threshold $T1$ (system parameter), the cluster is split along channel j alone into two subclusters. Assuming an N -channel vector space, let $M_i, i = 1, \dots, N$ denote the mean vector for the initial cluster; $M1_i, i = 1, \dots, N$ denote the mean vector for the first subcluster; and $M2_i, i = 1, \dots, N$ denote the mean vector for the second subcluster. SEP denotes a user-specified system parameter defining the separation of the new cluster means from that of the original cluster. Then the splitting process generates the two subclusters $M1$ and $M2$ in a manner such that

$$M1_i = M_i \quad i \neq j$$

$$M1_i = M_i + SEP \cdot \sigma_j \quad i = j$$

$$M2_i = M_i \quad i \neq j$$

$$M2_i = M_i - SEP \cdot \sigma_j \quad i = j$$

2. Cluster combining. On a combining iteration, each cluster is limited to combining with at most one other cluster. The process begins with computing the weighted distance between a cluster and each of the remaining cluster means. When the weighted distance is less than a threshold $T2$, the two respective means are averaged (weighted average) together. The mean averaging effectively combines two clusters into one cluster for the next pass of the data. The dis-

tance computations and thresholding continue until all the original clusters are tested.

CLUSTERING REPORT FUNCTIONS

1. Intercluster Distance Report uses the following distance formulas to calculate the distance between each pair of clusters:

Adaptive

For clusters I and J and channel L .

$$D_{IJ} = \sum_{L=1}^{NC} |M_{IL} - M_{JL}|$$

where NC is the number of channels.

Iterative

$$D_{IJ} = \left[\sum_{L=1}^{NC} \frac{(M_{IL} - M_{JL})^2}{\sigma_{IL} \sigma_{JL}} \right]^{1/2}$$

where M is the mean and σ is the standard deviation.

2. Cluster match routine. The nearest subclass defined in the subclass statistics table to a cluster is determined by using the match formula

$$D = \sum_{L=1}^{NC} \frac{(M_{IL} - M_{NL})^2}{\sigma_{IL} \sigma_{NL}}$$

where M_{IL} is cluster I mean for channel L . M_{NL} is the mean of the N th subclass in channel L , and σ_{IL} is the standard deviation. The distance is calculated over all channels and the nearest subclass is found. The final distance is the square root of the smallest D . (Note: The clustering algorithms described here are those originally implemented on the IBM 360/75. Some changes have been made to take advantage of the parallel SPP processing, but the effect remains the same.)

SUN ANGLE CORRECTION

If the user has selected the Sun angle correction option, the subclass means μ and covariance matrices Γ will be modified as follows:

$$\mu' = A_1\mu + A_2$$

$$\Gamma' = A_1\Gamma A_1^T$$

where μ' is the modified mean of the subclass, μ is the mean of the subclass, Γ' is the modified covariance matrix, Γ is the covariance matrix of the subclass, and A_1 and A_2 are diagonal constant matrices the elements of which are functions of the training segment and recognition segment Sun angles.

MEAN LEVEL ADJUSTMENT

If mean level adjustment has been specified by the user, the following computations must be performed. Compute the mean vector μ_I for the segment I to be processed, and the mean vector for the training segment J . If Sun angle correction for J has not been performed, then compute

$$\mu'_J = A_1\mu_J + A_2$$

The mean level adjustment vector is then

$$\Delta\mu = \mu'_J - \mu_J$$

and the resultant corrected mean vector is

$$\mu_I = \mu_I + \Delta\mu$$

BIAS CORRECTION

The following indicates the calculations used in bias correction.

Quantity

Meaning

p

The total number of bias correction categories.

B

The number of pixels considered. For classification bias correction, this is the total number of pixels used in classification, minus the number thresholded-out, minus the number in the category under consideration ("A"). For cluster bias correction, it is the number of pixels in clusters not labeled "A."

C_q

The label of the q th bias correction category.

N_q

The number of pixels classified (for classification bias correction) or clustered (for cluster bias correction) into category C_q .

N_{p+1}

The number of pixels classified (or clustered) into a category other than "A" or the bias correction set.

n_q

The number of bias correction vectors classified (or clustered) into category C_q .

n_{p+1}

The number of bias correction vectors classified (or clustered) into a category other than "A" or the bias correction set.

m

The total number of bias correction vectors.

$m_{q,r}$

The number of bias correction vectors labeled C_q by the user and classified (or clustered) into the r th bias correction category.

$m_{q,p+1}$

The number of bias correction vectors labeled C_q by the user and classified (or clustered) into a category other than "A" or the bias correction set.

$\alpha_{q,r}$

$m_{q,r}$ divided by n_r (0 if $n_r = 0$). Here, r varies from 1 to $p+1$.

$P(C_q)$

Bias corrected classified (or cluster) percentage for category C_q .

$$P(C_q) = 100 \sum_{r=1}^{p+1} (N_r/B) (\alpha_{q,r})$$

$P(C_{p+1})$

Bias corrected classified (or cluster) percentage for categories other than "A" or the bias correction categories.

$$P(C_{p+1}) = 100 - \sum_{q=1}^p P(C_q)$$

B_q

Beta value for category C_q .

$$B_q = (100N_q/B) - P(C_q)$$

Quantity	Meaning	Quantity	Meaning
$\alpha_{q,r}$	The r th component of the variance for category C_q (r varies from 1 to $p+1$); 0 if $n_r < 2$. Otherwise,	$U(C_q)$	Uncorrected classified (or cluster) percentage for category C_q
	$V_{q,r} = (100N_r/B)^2 (\alpha_{q,r}) (1 - \alpha_{q,r}) / (n_r - 1)$	$U(C_q)$	$U(C_q) = N_q/B$
$Var(C_q)$	Variance for category C_q (unreliable in any $n_r < 2$).	$U(C_{p+1})$	Uncorrected classified (or cluster) percentage for categories other than "A" or the bias correction categories.
	$Var(C_q) = (1/m) \left(\sum_{r=1}^{p+1} V_{q,r} \right)$	$U(C_{p+1})$	$U(C_{p+1}) = 100N_{p+1}/B$
B_q	Bias percentage range. The lower value is	%DO	Percentage classified "designated other" (classification bias correction only).
	$B_q - 1.645 \sqrt{Var(C_q)}$	%DO	$\%DO = 100 \times \text{number of DO pixels} \div B$
	and the upper value is	$I(D)$	Uncorrected percentage of unidentifiable pixels (classification bias correction only).
	$B_q + 1.645 \sqrt{Var(C_q)}$	$I(D)$	$I(D) = 100 \times (\text{number of pixels classified } I(D) + \text{number of pixels thresholded-out} + \text{number of pixels classified or clustered into category "A"}) \div 22932$

Considerations for Design of Future Research and Development Interactive Image Analysis Systems

T. B. Wilkinson^a

INTRODUCTION

The Earth Observations Division (EOD) at the NASA Johnson Space Center (JSC) will be shifting emphasis from the quasi-production programs such as LACIE to a more basic research and development (R&D) role to provide the necessary technology for the 1980's. Current workload forecasts for future programs show significant increases in the amount of imagery data that will have to be processed and analyzed.

An interactive approach to image analysis provides the responsiveness and adaptability required by the multiuser, multidiscipline programs of the 1980's. An interactive system puts the "man in the loop" and provides for almost immediate human judgment decisions relative to spatial data. A human observer typically makes many decisions during an interactive session based on his visual perception. The same decisions, however, may become cumbersome to make by way of machine processing. When cost, flexibility, and throughput are considered, it is apparent that the interactive image analysis approach best meets future requirements. To meet these requirements, the Earth observations interactive image analysis capability must be significantly enhanced.

The design of future interactive image analysis systems must consider the changing nature of the problem. An R&D environment requires a highly flexible system as opposed to the more limited flexibility of a production system.

Design considerations must include the increased processing requirements imposed by the addition of a thermal channel to Landsat-3 and the increased number of spectral channels with significantly higher spatial resolution provided by the Landsat-D thematic mapper. Other design considerations must include the rapidly changing technology in memories

and special-purpose processors. The analyst-machine interface and the human factors involved are often overlooked; however, they are considered to be significantly important for future systems.

Consideration of these and other factors has evolved to a basic conceptual approach for the design of future image analysis systems for the 1980's.

CURRENT ENVIRONMENT IN EOD

The JSC overall capability for processing and analyzing remotely sensed data is currently based on a number of special-purpose stand-alone systems. Each of these hardware/software systems was initially implemented to provide some aspect of the fast growing technology for processing and application of remotely sensed Earth resources data. The two existing interactive image analysis systems at JSC are categorized as special-purpose stand-alone systems.

The two systems are the Earth Resources Interactive Processing System (ERIPS) and the Image-100 system. ERIPS, a multiuser system, provides both a batch and an interactive capability. ERIPS provides the analyst with 2 high-resolution black-and-white displays (16 shades of gray) and their associated conversational monitors for interactive control via a command/prompt menu structure. The analyst is also provided with 2 color displays providing a maximum of 64 colors. Color imagery display and control are also provided via a conversational monitor. The basic image analysis hardware is a modified digital television (TV) equipment (DTE) cluster originally designed for display and control applications in the JSC Mission Control Center. The computational capability for ERIPS is provided by an IBM 360-75 computer and a STARAN special-purpose processor. Large-volume storage required by the imagery and ancillary data base is provided by as many as 42 high-density disk drives. ERIPS uses a large number of special-purpose software modules to provide the

^aLockheed Electronics Company, Houston, Texas.

clustering, classification, and display manipulation capabilities required by LACIE. ERIPS is the primary LACIE "production" system and is designed to achieve a high throughput rate. The software used by ERIPS for LACIE production is primarily structured for one specific set of tasks or algorithms and is cumbersome to use as an R&D tool. The basic hardware capabilities of ERIPS are limited when compared to more current systems designed specifically for interactive image analysis.

The other major stand-alone image analysis system is the General Electric Image-100. The Image-100 computational capability is provided by a Digital Electronics Corporation (DEC) programmed data processor (PDP) 11-45 computer. The Image-100 is a 5-channel interactive system that provides a color display of 512 by 480 picture elements (pixels) with as many as 256 intensity levels per channel. The normal configuration is to use four channels for video and the fifth channel for graphics. The fifth channel consists of eight 1-bit graphics planes and therefore provides eight graphics or theme tracks. The Image-100 provides some of the more sophisticated display manipulation capabilities not found in ERIPS but does not provide the throughput rate required for production. The Image-100 has had only limited use in LACIE and has been more of an R&D tool in LACIE. The Image-100 has been heavily used in the development of new procedures or techniques, such as an interactive maximum likelihood classification procedure. The Image-100 is also being used for other programs such as the Forestry Applications Program (FAP) and the Regional Applications Project (RAP). FAP and RAP are essentially pilot projects with heavy emphasis on R&D and have no major production requirements. Although providing a configuration more conducive to R&D, the Image-100 is only a single-user system and cannot easily be expanded to accommodate the increased number of spectral channels that will be used in Landsat-3 and in the Landsat-D thematic mapper. The Image-100 is also limited in its computational power; thus, clustering and classification processes are time-consuming.

In summary, the two existing interactive systems are adequate for their current tasks, but this will not be the case for the 1980's.

FUTURE ENVIRONMENT IN EOD

Forecasts for the next 10 years indicate that six programs will comprise the major workload: LACIE,

LACIE Transition, FAP, RAP, the Food Multicrop Program, and the Joint Soil Moisture Experiment. Starting with fiscal year 1979, these programs will be integrated into one unified program: the Global Food and Fiber Information System.

At first glance, it may not appear that the workload is increasing significantly, but this is far from being the case. The Landsat-3 will have an additional channel for thermal data, whereas the Landsat-D thematic mapper will also have the thermal channel and an additional spectral channel or possibly even two additional spectral channels. The significant change in the Landsat-D thematic mapper will be increased resolution—almost three times greater than the resolution of the current Landsat. If the number of Landsat acquisitions is assumed to remain constant, the increase in the number of channels and resolution represents a sevenfold increase in data alone. It should be rather obvious that the batch processing and manual photograph interpretation of film segments, such as in the LACIE program, could not accommodate the future environment. The analyst of the 1980's must be provided with a more efficient means of performing his work, and the implementation of a highly flexible interactive image analysis system will provide the necessary means.

A future interactive image analysis system must provide a multiuser, multiprogram capability. It must be a totally integrated system with all users having full capability to acquire common imagery data bases and to perform clustering and classification procedures quickly. The system should provide the capability to perform R&D work and to handle quasi-production work efficiently. The hardware and software must not be program-dependent; i.e., one analysis console should not be for LACIE, another for FAP, another for RAP, etc. The system should be capable of handling the variable size data sets of the various programs and should allow the analyst to select parametric classification techniques such as maximum likelihood and mixture density or non-parametric techniques such as parallelepiped, decision tree, and table look-up. The future environment will be one in which virtually all image analysis will be performed interactively with the key considerations being flexibility and speed.

It is projected that future programs will have a much higher involvement by the academic community. The increased involvement will probably include the use of JSC facilities by members of the academic community. The future interactive image analysis systems must be designed to accommodate the divergent backgrounds of the users. Future

systems must have a simplified analyst-machine interface so that the user may work with the system after minimal training. It should be possible for the user who has little computer experience to work effectively on digital image processing and analysis tasks.

IMPROVED TECHNOLOGY

Microprocessors

Some of the recent image analysis system designs are using microprocessors as an integral part of the system. Microprocessors such as the LSI-11 are being used to provide interactive control and processing and thereby to free the host computer for more complex operations such as classification and clustering. Microprocessors are also being used as the nucleus of an image array processor in which entire images instead of single pixels can be manipulated and displayed arithmetically at video frame rates. It appears that the relatively low cost of microprocessors, coupled with faster memory cycle times, will make the microprocessor extremely attractive for use in future image analysis systems.

Memory

Perhaps the most significant improvements have been in the area of refresh memory devices. Initially, most image analysis systems used video disks to refresh the display. The electromechanical video disks normally rotated at 1800 or 3600 rpm, which provided a memory latency of 33 or 16.6 milliseconds. In addition to being slow, the memory was subject to wear and alignment problems. The video disk refresh memory typically uses fixed read/write heads which "fly" a few microinches above the disk surface. The heads have a finite life and must be replaced periodically.

The next major improvement in refresh memory was the use of solid-state memory for the charge-coupled device (CCD). The CCD memory is a serial shift register and configuration designed to be plug-compatible with disk refresh memory systems. The CCD refresh memory overcomes the problems of an electromechanical device but still has the same memory latency characteristics.

The current state of the art is the random access memory (RAM), which, like the CCD, is a solid-state device but does not have the latency constraints. RAM refresh memory is very high speed

with access times typically less than 500 nanoseconds. The cost of RAM's was initially quite high; however, costs have continued to decrease with increased availability to the point that RAM refresh memory is on the order of 0.5 cent per bit and is projected to go to 0.2 cent per bit within the next year or two. RAM refresh memories are rapidly approaching the point at which the card, connectors, and switches, rather than the memory chips, become cost considerations.

Display Devices

For many years, the limiting factor in an image analysis system has been the color cathode-ray tube (CRT). The conventional shadow mask color CRT has improved over the years in terms of brightness and more consistent colorimetry because of the new phosphors; however, spatial resolution has remained about the same. The resolution limitations posed by the shadow mask design have prompted other approaches to high-resolution color such as the multilayer beam penetration tube. The beam penetration CRT involves the switching of different anode voltages to excite a particular phosphor layer. This approach has had only limited success because of the problems encountered when switching 15- to 20-kilovolt pulses.

The biggest improvement has been the introduction of a high-resolution shadow mask color CRT. A gravure quality shadow mask CRT with a 0.3-millimeter pitch (triad spacing) is currently being produced. This shadow mask provides approximately five times the dot density of conventional shadow mask CRT's and is being used to provide 1024- by 1024-pixel color displays. The closer spacing of the color triads also provides a display free from moiré patterns.

Monitor circuitry has also been improved to keep pace with the improved CRT's. This improvement has included the use of individual operational amplifiers in the convergence circuitry to minimize interaction and to increase stability. A major improvement has been in the area of stabilizing color temperature by a beam-controlled feedback circuit, which automatically adjusts the monitor color temperature to a fixed reference during each vertical blanking interval.

ANALYST-MACHINE INTERFACE

The LACIE has provided only a limited amount of experience with interactive image analysis

systems; however, the knowledge gained is sufficient to provide guidelines for future systems. The analyst-machine interface appears to be an area that is frequently overlooked or that is given a relatively low priority in the overall systems considerations. This interface must be carefully considered in future system designs because lack of attention to this aspect will adversely affect both the productivity and accuracy of the image analysis and the classification process.

Human Factors

Perhaps the most frequently expressed concern for future interactive systems is in the area of human factors. The analyst is not only concerned with the implementation of the imagery display, the ancillary information display, and the interactive controls but also with the work environment. Consideration should be given to reducing the fatigue of an analyst, who may typically work at an interactive display console for periods of as long as 8 hours. Viewing conditions are of paramount concern and will be addressed first.

The television broadcasting industry has long been concerned with viewing conditions in control rooms where, for critical evaluation of picture quality, all factors affecting the perception of colors and brightness should be closely controlled. The analysis of multispectral sensor data usually involves the display of the data in pseudocolor or false color rather than the true color used in the broadcast industry. The maintenance of the colors and brightness perceived by the analyst is equally as critical because of the decisions that must be made in the classification process based on the color perceived. The Canadian Broadcasting Corporation has prepared several papers dealing with the control room environment which have been published in the Journal of the Society of Motion Picture and Television Engineers (ref. 1). The recommendations made are largely appropriate for consideration in the design of future interactive image analysis systems. Some of the pertinent considerations are as follows.

1. Chromaticity of color picture monitor screen at reference white: The screen chromaticity at reference white should be D6500 ($X=0.313$, $Y=0.329$) with a tolerance of ± 200 K along the daylight locus.

2. Luminance of color picture monitor screen at reference white: The screen luminance at reference white should be 20 ± 2 footlamberts.

3. Viewing distance: The viewing distance requirements differ from those of the broadcasting industry and will vary depending on the type of analysis being performed.

- a. In general, the analyst should be positioned so as to view the display from a distance of not less than one nor greater than four times the height of the monitor picture.

- b. The analyst should be placed so that his angle of view is no greater than 30° from a line normal to the face of the monitor.

4. Light surround: Light surround is defined as the light, visible to the analyst, from a plane or from behind a plane coincident with, and surrounding but not including, the viewing screen. Light surround requirements are as follows.

- a. Light surround should be provided outside the monitor screen mask and over an area at least eight times the area of the monitor screen.

- b. Light surround should have a luminance of 3 ± 1 footlamberts.

- c. Light surround should have a chromaticity matching the color monitor screen reference white, thus providing a fixed color reference for the analyst.

5. Monitor screen mask: The monitor screen should be framed by a narrow black matte mask.

6. Analysis console room decor: The viewing room in which the analysis console is located should have a decor that gives a generally matte impression without the use of dominant colors.

- a. The ambient light on the monitor screen should be kept to the lowest possible level. Specular reflections must be avoided.

- b. Light sources within the room should be of a similar color temperature to that of the color monitor reference white.

- c. Desk surfaces used by the analyst should be illuminated with lighting of the cool white fluorescent type and adjusted so that the luminance of white paper on the desk falls between the limits of 6 to 10 footlamberts.

- d. The desk surfaces should give a generally neutral matte impression without the use of dominant colors.

Another problem often encountered by the analyst is that of noise. The efficiency and accuracy of an analyst is greatly diminished when in a noisy environment for significant periods of time. Background noise sources generally fall into two basic categories, machine-induced and human. Machine-induced noise sources may include airconditioning systems, blowers, and computer peripherals such as line printers and disks. Human noise sources include

conversations of other workers, telephone conversations, and general office noise.

It is recommended that the analysis console be placed in a room sufficiently isolated from noise sources so that a noise criterion of no greater than NC-35 can be obtained. The NC-35 level is based on studies performed by Beranek (ref. 2).

The analyst should have control over the temperature of the room in which the analysis console is located so that he may adjust it to suit his personal preference without affecting other areas.

Simplified Operator Interaction

A second major consideration in the analyst-machine interface is the simplification of operator interaction. In looking at future R&D interactive image analysis system users, one sees a more diverse group representing numerous disciplines with heavy involvement from the academic community. As the multiuser, multiprogram environment projected for the 1980's approaches, it becomes apparent that the interactive image analysis system must be a "tool" which can be easily used with a minimum amount of training. The interactive image analysis system should not require the analyst to have extensive programming skills or a detailed knowledge of computer architecture. The interaction between the analyst and the system should not require the typing of each operator-initiated command. The use of typed commands should only be required in certain special cases such as program development.

A more simplified approach to operator interaction is the use of an interactive processing monitor to govern all terminal interactive processing. In essence, the analyst should have a monitor at each display console that presents a menu from which the analyst can select options and control the processing sequence. The monitor on which the menu is displayed should be considered as a conversational monitor because it will prompt the analyst and serve as the primary communication link with the image analysis system. The conversational monitor should be either of the storage tube type or a high-resolution raster scan. The conversational monitor is primarily an alphanumeric (A/N) display, and special attention must be given to the display if a raster scan system is employed. Flicker becomes a problem in a standard system of 525 lines per frame and 30 frames per second because of the high-contrast display (typically black and white) and the rate at which the display is refreshed. Flicker can be reduced to an ac-

ceptable level by employing a repeat-field display so that a given pixel is repeated in both fields (odd and even), thus providing a 60-hertz refresh rate rather than a 30-hertz refresh rate. A repeat-field approach, although attractive from the flicker reduction standpoint, means that vertical resolution must be sacrificed. A repeat-field display has one-half the vertical resolution of a non-repeat-field display so that if there are 480 active lines in a 525-line display, one would have a vertical resolution of only 240 lines, which is insufficient for the display of small alphanumeric characters. It is therefore recommended that a repeat-field display of 1024 lines per frame be used because it has significantly reduced flicker and the necessary vertical resolution.

The analyst should have the capability to select menu options via such devices as a graphics tablet, a trackball, a lightpen, a joystick, or similar devices. The menu should have a hierarchical processing structure that uses monitor programs, processing programs, and subprograms. The system should include an interactive processing monitor that controls menu generation, interrogation and editing, and image display and manipulation.

The design of the system should allow the analyst to override the computer control of the processing and image display and manipulation from the console. All manual controls provided to an analyst should provide a positive indication of their status such as pushbutton indicators which illuminate when depressed. Visual indication must also be provided at the console to ensure that the analyst knows whether he is operating under computer control or in the manual mode.

Any action initiated by the analyst should result in some positive visual indication within 15 seconds that the command has been accepted and that processing is in progress. If the system is "busy" and cannot process the request entered by the analyst, there should be a visual indication that the command has been accepted but that processing will be delayed.

CONCEPTUAL SYSTEM DESIGN

The development of an optimum design configuration for interactive image analysis would require extensive modeling of each candidate configuration. At this point in time, the "hard" requirements for a future interactive analysis system are not sufficient to allow any meaningful modeling. The main emphasis of this paper is to present certain concepts that should be addressed rather than a detailed

design of such a system. Two basic system approaches that illustrate these design concepts are described in the following subsections: a "centralized system" and a "distributed system." Several factors suggested their consideration: the projected workload, the probable availability of a large-scale computer such as an IBM 360-75, the number of image analysis terminals designed to interface with midrange computers such as the PDP 11-70, and the evaluation of existing image analysis systems.

Studies performed by the MITRE Corporation (ref. 3) indicate that a maximum of 12 image analysis terminals would be required in the future. However, an initial configuration of six image analysis terminals would meet the workload requirements of the early 1980's. The conceptual approaches are therefore based on a 6-image-analysis-terminal (IAT) configuration with possible expansion to as

many as 12 image analysis terminals. The conceptual approaches also assume the availability of the STARAN parallel processor or equivalent device.

Centralized System

The centralized system concept as shown in figure 1 employs a direct interface between the large-scale computer (IBM 360-75 class) and the image analysis terminals. In this configuration, the image analysis terminals are connected directly to the selector channels via model 2701 data adapter units. The mass data storage facility and the STARAN processor are interfaced to the selector channels with the low-speed peripherals connected to the multiplexer channels. Each selector channel (SC) is capable of maximum data rates ranging from 1.3 to 1.85 million

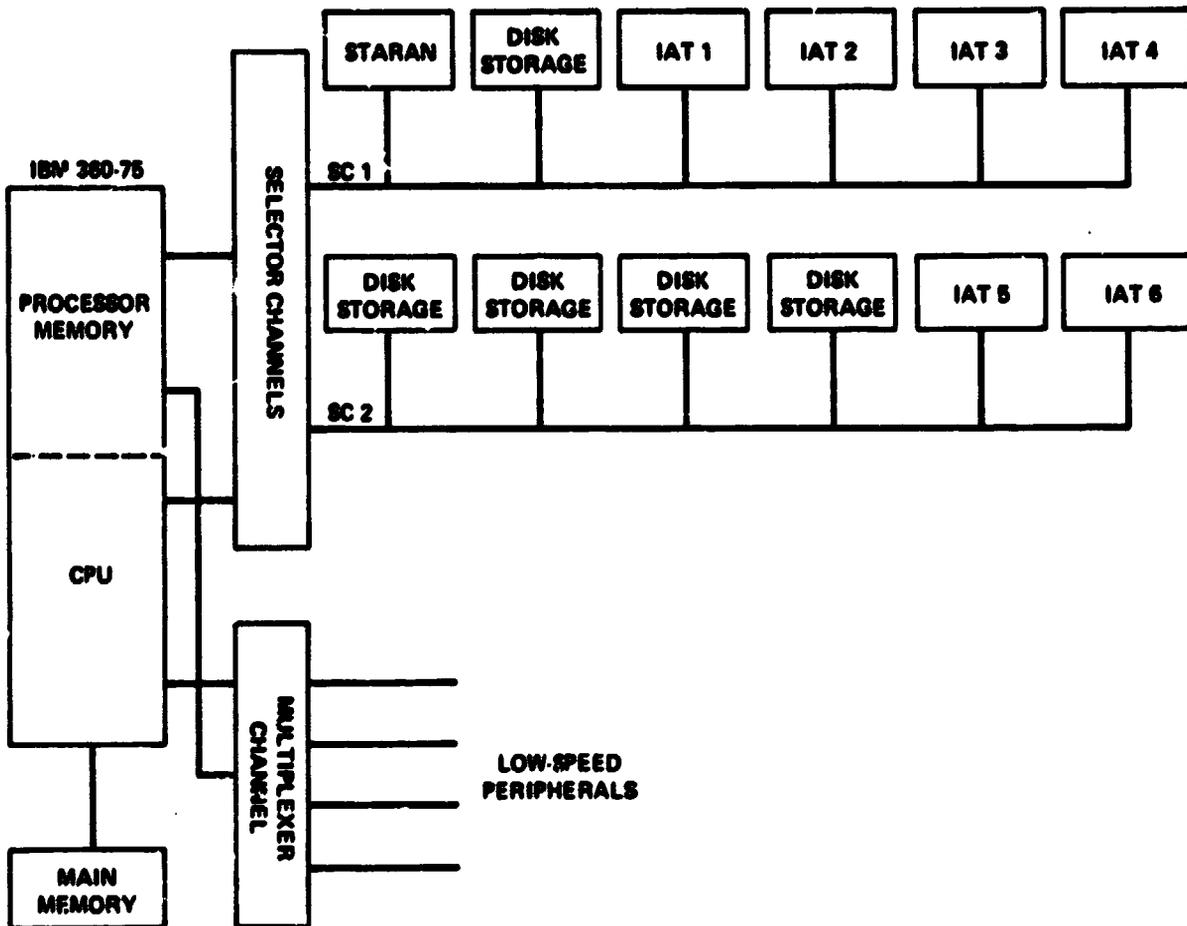


FIGURE 1.—Centralized system.

bytes per second. Each selector channel attaches as many as 8 input-output control units and can address as many as 256 input-output devices. Only one input-output device per selector channel can transmit data at any given time, and no other input-output device on the channel can transmit data until all data are handled for the selected device. The centralized system could easily interface to as many as 12 image analysis terminals; however, system throughput becomes a matter of concern. Although selector channel bandwidth appears to be adequate, a single central processing unit (CPU) is used to perform all computations, to control data movement to and from the array processor, and to control data movement to and from the image analysis terminals. Normal data movement on the selector channels does not appear to be a problem, but the situation in which two analysis terminals may simultaneously request maximum likelihood classification or any other time-consuming computational process may affect the interactive capability of the system. It may well be that this configuration would be adequate to provide the desired interactive image analysis capability; but this adequacy can only be determined by development of an accurate model of the projected workload, which is beyond the scope of this paper.

The centralized configuration offers the following attractive features.

1. It is straightforward.
2. It is the least expensive to implement.
3. All software is in one computer.
4. Lower maintenance and operational costs are involved.

Some of the disadvantages of the configuration include the following.

1. No stand-alone capability exists; it is completely dependent on main CPU and associated communications channels.
2. Software modifications are required as image analysis terminals are added.
3. Loading may slow system throughput to the point of not being truly interactive.

Distributed System

The distributed system concept as shown in figure 2 employs the use of a "control computer" between the large-scale computer (IBM 360-75 class) and the image analysis terminals. Interface of the mass storage facility and the array processor to the large-scale computer is similar to that in the centralized

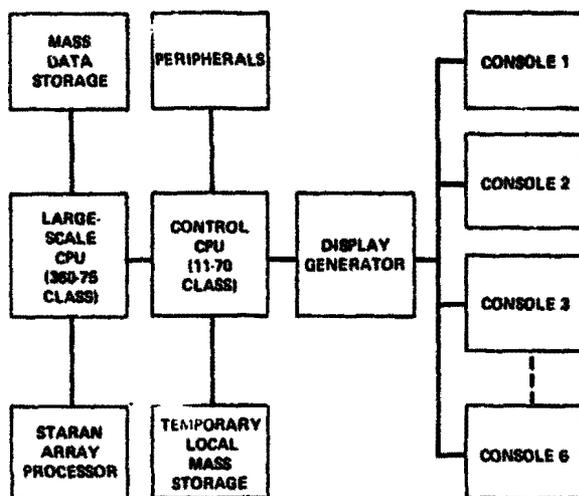


FIGURE 2.—Distributed system.

system. The control computer provides the interface to the image analysis terminals and to its own temporary local mass storage and peripherals. The functioning of the distributed system can best be understood by a more detailed analysis of the control computer.

Control Computer

The control computer is envisioned as a midrange, highly interactive machine such as a PDP 11-70. Its primary functions are interacting with the analysis consoles, moving data between the other system elements, storing and retrieving bulk data, and monitoring the overall functioning of the system. The control computer also provides minimal computational support for the image analysis system, but it is assumed that most bulk data processing will be accomplished in the special-purpose processor and the large-scale computer.

In concept, most of the actions by the control computer will be in response to inputs from an analysis console. The inputs from the analysis consoles may be made via alphanumeric terminals associated with the analysis consoles. The emphasis in all cases will be to provide a near-real-time response to inputs. In practicality, this means that a response must be available to the analyst within a few seconds after initiation. To be a truly interactive system, the control computer should execute most tasks within 30 seconds. Complex tasks that may require more

than 3 to 5 minutes for completion should be scheduled for deferred or background execution while the console is made available for further operations.

The control computer will provide the capability to input data directly from normal computer tape or high-density digital tape (HDDT). The primary data source, however, will be the data base and data base management system resident in the CPU's of the main computational facility. Through the data base management system, the control computer will have access to imagery data bases. Transmission bandwidth limitations, however, make real-time access to imagery data impractical. In fact, the size of the imagery data base is so large as to make on-line storage impractical. The EOD systems and facilities workload requirements forecast prepared in April 1977 shows a LACIE Phase II imagery data base of 8.5×10^8 bytes and a LACIE Phase III imagery data base of 15.5×10^8 bytes (ref. 4). The numbers should not be considered absolute; they are given only to indicate the magnitude of the imagery data base.

To operate in a truly interactive mode, the analyst will know in advance imagery data requirements in an interactive session and will submit a request for the required data to be available for the session. The control computer will acquire the appropriate data bases and download the data into local mass storage on a non-real-time schedule. During the interactive session, data will be acquired from the temporary local mass storage and transferred in real time to the displays and the special-purpose processor. The sizing of temporary local mass storage is expected to be sufficient to support 1-day (24-hour) transactions. The availability of temporary local mass storage also makes the system semi-independent and much less sensitive to communications failures. Additionally, most of the interactive image analysis system demands on other elements of the data processing system can be scheduled at periods of low activity and more available communications channels.

No current system configuration exactly duplicates the proposed distributed system. However, the Atmospheric and Oceanographic Information Processing System (AOIPS) at the NASA Goddard Space Flight Center (GSFC) has some similarities (ref. 5), and the performance achieved at GSFC may be used as a baseline. The GSFC AOIPS uses an IBM 360-91 for large-scale computational capability and a DEC PDP 11-70 interfaced to two image analysis terminals. Each IAT contains five channels of RAM

refresh memory and has a dual interface to the PDP 11-70 in which high-volume data are transferred on the high-speed (H/S) bus and low-volume data and control signals are transferred on the unibus. GSFC has determined that the current AOIPS configuration as depicted in figure 3 will provide from 50 to 70 percent of the total theoretical system input-output bandwidth of 5.8 megabytes per second. If it is assumed that each refresh memory will be loaded with a 512-line by 512-pixel image, there will be a data transfer of 262 144 bytes per channel since a pixel is 1 byte. If it is further assumed that the data are formatted and stored on a disk such as an RP06 (which has a data transfer speed of 1.2 microseconds per byte) a single channel (512 by 512) requires a bandwidth of approximately 0.325 megabyte per second. A transfer of four channels into refresh memory requires a bandwidth of approximately 1.3 megabytes per second. The 10-channel configuration of AOIPS requires a bandwidth of 3.25 megabytes per second and nominally provides for the transfer of approximately 2 full television images per second. An extrapolation of this capability can be extended to future systems. Assume a system having 12 image analysis terminals, each with 5 refresh memories (512 by 512). The theoretical bandwidth excluding overhead would be 12 consoles by 1.625 megabytes (5 channels), or 19.5 megabytes. If a nominal bandwidth of 3.5 megabytes is assumed, the system throughput would be reduced by a factor of 6. The throughput figures assume that one console requests at a given time and that all data transfers to the first requesting terminal are completed before servicing a second requesting terminal. It should be made clear that the throughput addressed thus far is only for movement of data and does not address the time required for computational processes such as classification and clustering.

The distributed system is similar to the centralized system in that the repetitive operations required for classification and clustering will be accomplished by an array processor such as the STARAN. A detailed analysis of array processor performance in image analysis applications is not considered necessary because several papers addressing this topic have been presented in the past 3 years.

Based on performance achieved with the AOIPS, it would appear that a single control computer such as a PDP 11-70 could support six image analysis terminals in an interactive mode. The system throughput rate would be approximately six times as slow but would still provide acceptable response

from the main CPU, set into operation, and left to run independently until the results are available. The results are downloaded into the control computer and then routed to the appropriate display console. When two or more consoles request operations requiring the special-purpose processor simultaneously, the requests will be queued on a priority basis. In all other cases, requests will be processed in the order received. It is unlikely that a sufficiently large number of analysts simultaneously initiating contending requests will cause an unacceptable waiting time.

Image Analysis Terminals

The image analysis terminals will consist of two basic components: a display generator and display and analysis consoles. The image analysis terminal concept is identical for both the centralized and the distributed systems and differs only in the display generator/computer interface.

Display Generator

The centralized display generator that services all the consoles is the core of the interactive image analysis system. The display generator will contain the refresh memories for all consoles and the cursor generation, hardware classifier, character generation, and digital-to-analog (D/A) conversion circuitry. The logical and physical architecture of the display generator will be highly modular and will use standardized building blocks. As currently envisioned, the standard building block will be a module that contains between 1 and 8 RAM refresh memories (512 by 512 by 8), computer interface, hardware classifier, and video processing circuitry. The exact configuration of the display generator module will be determined by the configuration of the interfacing console. The centralized display generator configuration is illustrated in figure 4.

The centralized approach has several advantages.

1. Each individual terminal need not be implemented with a refresh memory for the maximum number of channels it is ever expected to handle. The centralized system can instead be implemented with the most likely number of channels that can be expected to be in use at any given time. Channels can then be added or deleted to a given terminal, based on the requirements of the task to be performed.

2. The centralized system is also easier to maintain and service. If any individual channel should fail, modules from another channel can be "borrowed" while diagnosis and repair are performed on the faulty unit.

3. The centralized concept also simplifies the data distribution problem to the console displays. The display generator will output video signals that will be routed to the console displays via a coaxial cable. The consoles can be located as far as several hundred feet from the display generator without any degradation to the signal. For cable runs of less than 1000 feet, simple unbalanced cable equalizers can be used at the console to restore a flat frequency response and reject any low-frequency noise that may be induced in the cable run. For longer cable runs, balanced transmission techniques can be used.

4. All other display generator/console interfaces are low-speed control and status indicator circuits and are not constrained by critical cable lengths. The longer cable runs may require amplification and redrive circuitry, but this is not considered a major problem.

Display and Analysis Consoles

The display and analysis consoles will employ a modular design consisting of standard modules. The two basic categories of standard modules are display and control.

The display modules include a color monitor (512 by 512 pixels), a conversational monitor (black-and-white repeat field), a high-resolution monitor (1024 by 1024), and a light table using an illuminant having the same color temperature as the color monitor. The control modules include a keyboard, target or cursor control, a display control and status module, function buttons, and a communications panel.

The display and control modules are used in a building block approach to provide the required console configuration to perform a specific task or function. As programmatic requirements change, console modules may be added or deleted as required with minimal impact on the overall system.

Preliminary studies indicate that the consoles may fall into four basic categories: analysis console; screening and editing console; registration, mensuration, and correlation console; and systems manager console. The analysis console is a full capability console with two color displays, a conversational monitor, a light table, a keyboard, and appropriate control

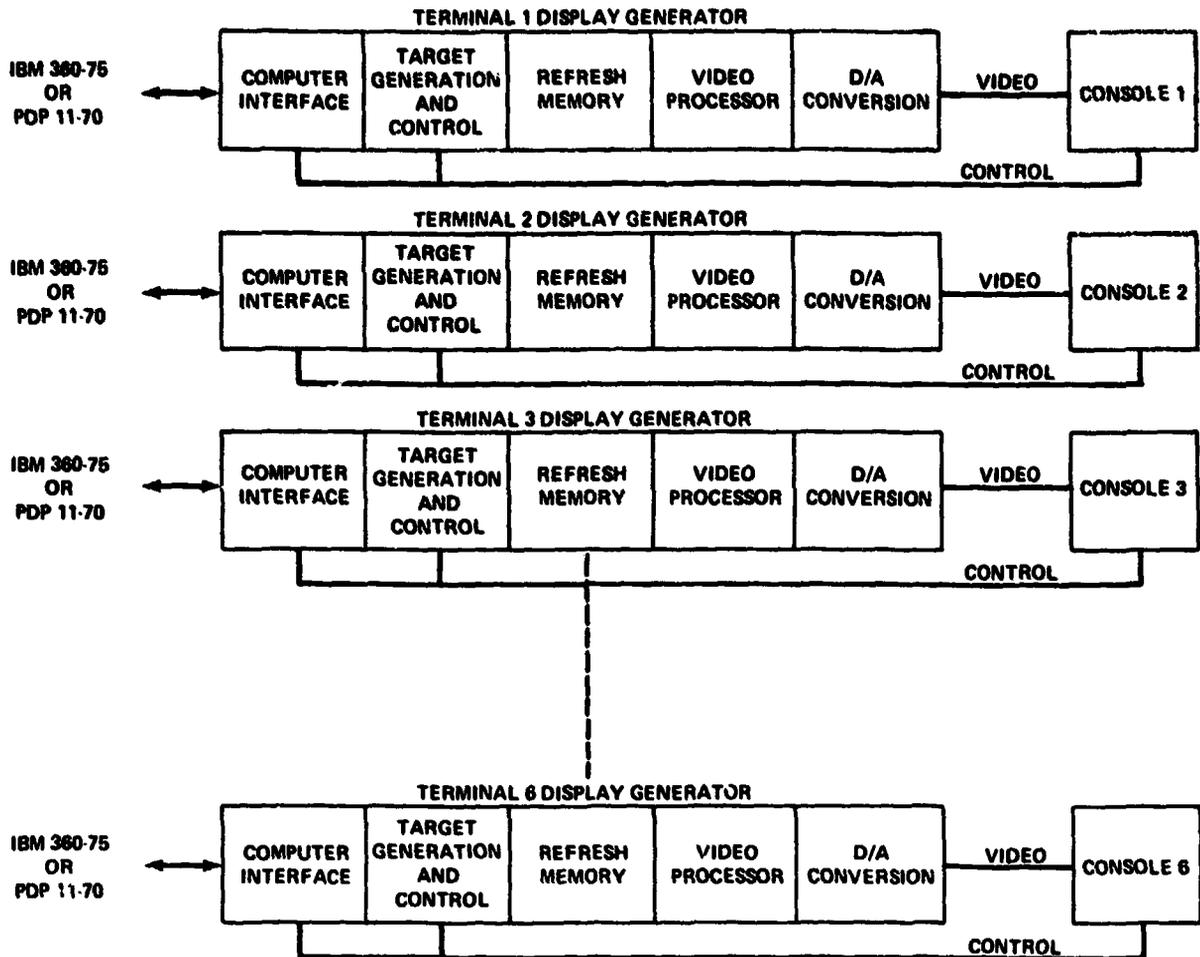


FIGURE 4.—Centralized display generator.

modules; whereas a screening and editing console may use only a single color display. The registration, mensuration, and correlation console capability is primarily centered around the high-resolution (1024 by 1024 pixels) black-and-white display. The systems manager console provides the capability to monitor the activity of all consoles. The systems manager is capable of calling up any display from any console to ensure quality control or to aid in maintenance. The systems manager console is the means by which system resources are allocated and tracked.

SUMMARY

Interactive image analysis systems currently at

JSC have been used to perform a limited number of specific tasks; however, future requirements indicate that a more general image analysis capability must be provided. The design of such a system will involve a thorough analysis of the environment projected for the 1980's. The design must address the analyst-machine interface in detail because this is often a limiting factor in interactive image analysis systems. The design of future systems must involve careful evaluation of state-of-the-art technology as well as improved analysis techniques.

In summary, it is not possible to address all the considerations affecting the design of future image analysis systems; however, this paper has attempted to identify those key considerations for a future JSC image analysis system.

ACKNOWLEDGMENTS

The author extends special appreciation to Jack Robertson of Lockheed Electronics Company for his contributions to this paper.

REFERENCES

1. Glover, S.: A New Approach to Television Studio Control Room Design. J. SMPTE, Jan. 1967.
2. Beranek, Leo: Noise and Vibration Control. McGraw-Hill, 1971.
3. Henderson, R. G.; and Lemmon, A. V.: A Study of Digital Image Analysis Systems. MTR-74, The MITRE Corp., Jan. 1977.
4. Stevens, W. A.; and Vitellaro, J. A.: EOD Systems and Facilities Workload Requirements Forecast. LEC-10522, Lockheed Electronics Co., Houston, Tex., Apr. 1977.

5. Bracken, P. A.; Dalton, J. T.; Billingsley, J. B.; and Quann, J. J.: Atmospheric and Oceanographic Information Processing System (AOIPS) System Description. NASA Goddard Space Flight Center (Greenbelt, Md.), Mar. 1977

BIBLIOGRAPHY

- Adams, John; and Wallis, Robert: New Concepts in Display Technology. IEEE Computer, Aug. 1977.
- Fink, Donald G.: Television Engineering. McGraw-Hill, 1952.
- Gambino, Lawrence A.; and Schrock, Bryce L.: An Experimental Digital Interactive Facility. IEEE Computer, Aug. 1977.
- Rohrbacher, Donald; and Potter, J. L.: Image Processing With the STARAN Parallel Computer. IEEE Computer, Aug. 1977.
- Zwick, Daan M.: Television Receiver White Color: A Comparison of Picture Quality With White References of 9300 K and D6500. J. SMPTE, Apr. 1973.

Very High Speed Processing: Applicability of Peripheral Devices to Pixel-Dependent Tasks

J. C. Lyon^a

ABSTRACT

The LACIE was representative of applications users of Landsat data; it was distinguished in the context of the present paper primarily by the quantity of data processed in the project. This data volume was anticipated, before project inception, to exceed the processing capacity of existing support systems, particularly in the performance of LACIE implementations of classical pattern-recognition functions; viz, iterative clustering and maximum-likelihood classification. This paper describes the early options studied in the satisfaction of LACIE computational demands and the ultimate selection and development of an array processing solution to the problem. The economic justification, as a function of required multitemporal Landsat analysis, is provided for this approach; the suitability of such processors for LACIE and other applications is discussed.

THE IMAGE PROCESSING ENVIRONMENT: PIXEL PROCESSING IN LACIE

Digital image processing typically involves use of computationally intensive techniques in both data preparation and data analysis applications. Examples of such techniques are geometric and radiometric corrections, other filtering applications, some data clustering procedures, and various statistically based classifiers. These procedures are characterized by a relatively large number of arithmetic operations to be executed for each picture element in an image or image subset. Since images tend to be composed of

many pixels, the performance of any of these procedures can demand an astronomical number of instruction executions. Even in relatively limited applications, conventional serial processing devices have sometimes proved to be inadequate image analysis vehicles, either yielding unacceptable interactive response times or monopolizing system resources. The result has been a proliferation of special-purpose equipment designed to assume the processing load associated with computationally burdensome algorithms. The implementation of a programmable parallel processor, peripheral to a large mainframe, as the solution of the "pixel processing" problem posed by LACIE is discussed herein.

To justify this implementation, the first major portion of the paper will be a review of the conditions driving the decision to modify the system. This review will be effected through (1) a detailed examination of the behavior of the pre-LACIE system when performing a representative computationally bound (compute-bound) application and (2) a qualitative survey of considered processing alternatives to the existing unsatisfactory mainframe performance. In the second major section, the performance and economic justification of the selected special-purpose device as applied to LACIE will be discussed and the general applicability of special devices to the general image processing problem will be noted.

Finally, it should be observed that the basic intent of the implementation, to improve system throughput in the face of very large data volumes, was achieved. Essentially, in most applications, the system performance was converted from a central processing unit (CPU)-bound state to an input-output (I/O)-bound state. This conversion, however, exposed the I/O and data management/traffic control of increasingly large and diverse data sets as a critical problem for future large-scale remote-sensing applications.

^aNASA Johnson Space Center, Houston, Texas.

PROBLEM STATEMENT: PRE-LACIE AND EARLY LACIE CAPABILITIES AND LIMITATIONS

System Description

In late 1973, the Earth Resources Interactive Processing System (ERIPS), an existing NASA Johnson Space Center (JSC) image analysis facility, was identified as the baseline from which the LACIE production classification system was to be developed. This proposed use of ERIPS, which had been conceived only as an interactive research tool, was based primarily on the existence within the system of much suitable applications software, an expandable hardware/software configuration, and schedule constraints precluding total design and development of a new support system. The choice of ERIPS was satisfactory, even if not ideal, but extensive enhancements¹ were required in a number of areas. A block diagram of the system as it existed in 1973 is shown in figure 1, with later additions for LACIE indicated.

Of prime concern to the use of this system within LACIE was its throughput, constrained both by factors resolved as described elsewhere¹ and by the more fundamental CPU limitation exposed by computationally bound routines invoked frequently in analysis. Within the principal ERIPS subsystem (Pattern Recognition) to be used by LACIE, the

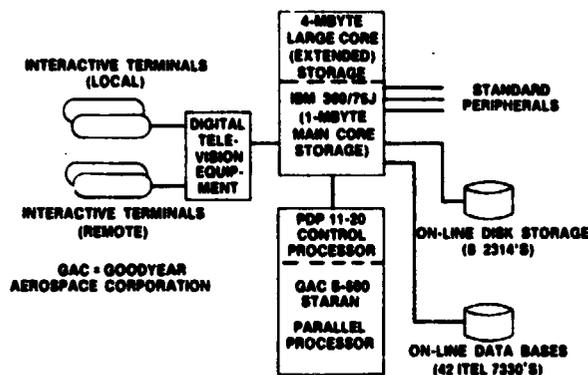


FIGURE 1.—LACIE/ERIPS facility.

¹For a description of these enhancements, see the following LACIE Symposium papers: C. L. Johnson, "LACIE/ERIPS Software System Summary"; L. E. Westberry, "The LACIE Data Bases: Design Considerations"; and Barbara B. Duprey, "Man-Machine Interfaces in LACIE/ERIPS."

modules exhibiting this characteristic were iterative clustering (a derivative of the ISOCLS algorithm), feature selection (a weighted divergence routine), and maximum-likelihood classification. These processors are described in the paper by Johnson; here, the example of maximum-likelihood classification is adequate to quantify the extent of the CPU-driven problem and will be used subsequently for continuity.

The ERIPS Maximum-Likelihood Classifier: LACIE Implications

The objective of maximum-likelihood classification is to minimize the function:

$$H_c(\mathbf{x}) = S_c + \frac{1}{2}(\mathbf{x} - \mu_c)^T \Lambda_c^{-1}(\mathbf{x} - \mu_c) \quad (1)$$

where S_c is a class constant associated with each class C , Λ_c^{-1} is the N -channel inverse covariance matrix of class C , μ_c is the N -channel mean vector for class C , \mathbf{x} is the N -channel observation vector (pixel) under test, and H_c is the likelihood measure that the pixel \mathbf{x} belongs to class C . If $H_c \leq H_k$ for all $k \neq c$, then \mathbf{x} is assigned to class C . The essential limitations within ERIPS were typical to many classification schemes; viz, $N \leq 30$ and $C \leq 20$ (in LACIE, $N \leq 16$ and $C \leq 60$).

Various implementations of equation (1) have been adopted at image processing installations; most follow rational considerations of maximization of processing efficiency given a specific equipment and software architecture. ERIPS was no exception; considerable attention was paid to code optimization in the IBM 360/75 computer. The following steps were adopted.

1. Reformulate equation (1) into a computationally efficient expression:

$$H_c(\mathbf{x}) = S_c + \sum_{i=1}^N \sigma_i \sum_{j=1}^i \Lambda_{ij}^{-1} \sigma_j \quad (2)$$

where σ_i is the standard deviation in channel i of class C . The covariance matrix is, of course, lower triangular. This restatement is typical.

2. Replace original FORTRAN code with assembly language.

3. Improve disk access methodology—A local (ERIPS) disk access method was developed to exploit image organization characteristics. This procedure, the Image Data Access Method (ref. 1), permitted retrievals from the image storage medium (IBM 2314 disks) at essentially full disk rates and thus reduced I/O waits at any point to a minimum.

4. Serialize representation of equation (2) in core to avoid CPU cycles in resolution of branch addresses.

5. Exploit register-to-register arithmetic and eliminate register-to-storage arithmetic (slower) where possible.

Given the preceding considerations, the classification process (eq. (2)) was reduced essentially to its arithmetic components with minimal system overhead or time-consuming addressing and logical operations. Consequently, the arithmetic required for the solution of equation (2) was given by

Adds: $(N^2 + N + 1)C$ per vector

Multiplies: $1/2(N^2 + 3N)C$ per vector

consisting of the bulk of CPU cycles (0.75 second in the ERIPS CPU) used in the classification process.

The number of operations required on a single pixel under this breakdown for representative ERIPS test cases is shown in table I. Using a LACIE sample segment (for comparison with later sections) consisting of 22 932 pixels, it can be seen that the number of arithmetic operations for this "inner loop" of classification can approach 10^9 , given the LACIE experience of approximately 40 classes defined per

TABLE I.—Maximum-Likelihood Classification Instruction Executions Per Pixel in the Quadratic Form

Parameter	Values				
No. of channels	4	4	16	16	16
No. of classes	10	20	10	20	60
No. of adds	250	500	2890	5780	17 340
No. of multiplies	140	280	1520	3040	9 120

sample segment.

Table II affords a more global look at the classification processes other than the inner loop and shows the significance of the quadratic form to the complete process. It is easily seen that the classifier is heavily compute-bound and that necessary I/O operations, represented only in lines 1, 2, 5, and 6, are negligible contributors to the process. The "Gen stats" entry is associated with the preparation of statistics for the classifier from the original training set, and is, with the quadratic form evaluation, almost totally compute-bound.

To exemplify the limiting time resource, consider the problem of classifying into 10 classes an image of 7.5 million pixels in 4 channels (a Landsat-2 frame). This problem would require nearly 4 hours of CPU time under the indicated breakdown of the table.

The timing figures developed here were applied to the LACIE anticipated workload; similar values were obtained for the other processes which would

TABLE II.—IBM 360/75 Maximum-Likelihood Performance, June 1973^a

[Computed timings for LACIE segments, sec]

Process	Time, μ sec	C=10;	C=20;	C=10;	C=20;
		N=4	N=4	N=16	N=16
System overhead	3000L	0.351	0.351	0.351	0.351
Data movement	$(6.49N + 8.82)LP$.797	.797	2.583	2.583
Gen stats	$5.03C^2LP$	11.535	46.140	11.535	46.140
Quadratic form	$[7.46 + N(4N + 16.21)]CLP$	31.256	62.512	296.010	592.020
Store best Q	$2.92CLP$.700	1.400	.700	1.400
Store best C	$5.9LP$.135	.135	.135	.135
Total		44.780	111.335	311.314	642.629
Percent in quadratic form (pure compute)		70	56	95	92

^aImage containing L lines, P pixels/line, N channels, with C classes.

contribute significantly to a typical LACIE scenario. The conclusion, based on a project peak of 4800 sample segments, was that some 60 hours of IBM 360/75 time would be required daily to perform the assignable ERIPS functions. Such equipment time, although theoretically available locally within the five 360/75 CPU's of the installation, was clearly impractical (ref. 2). An alternative was indicated.

Processing Alternatives

The problem identified previously, common in some degree to most image processing installations, can be addressed in several ways. Some of these techniques were unsuitable for LACIE application but are included here for completeness. Given a constrained CPU, the following alternatives may merit consideration under given assumptions for some applications.

1. Technique improvements/alternatives (software)

a. A variety of improvements to the classification process time can be achieved without loss of accuracy; the measure of improvement is strictly data-dependent but does not generally exceed a factor of 4 or 5 for the most successful, vector classification. This procedure involves the retention of classification assignment for pixels through maximum likelihood in an ordered table. Subsequent pixels (vectors) are interrogated for presence in the table. If present, the assignment is made on the table class; if not, classification proceeds normally. This procedure works well on single-acquisition Landsat images in agricultural areas in which data spread is small and duplication of pixels is a frequent occurrence. Unfortunately, in multitemporal applications (LACIE), the procedure is essentially unusable; the likelihood of, for example, duplication of 16-channel vectors is small in typical images.

b. Another alternative is the treatment of the classification expression (eq. (2)) as a running sum of the calculated components for a pixel (x) against a class (C). Periodically in this summing procedure, a test of the sum against the best H_c to date is made; if the running sum exceeds this best value, processing is terminated on the class and the pixel is advanced to examination against the next class. Typically, the performance gain in this approach may be 50 percent. The total system improvement afforded by this technique, however, was inadequate to meet LACIE requirements.

c. Other techniques which offer varying degrees of promise to specific users as computational reduction devices include linear (fixed-point) classifiers and the parallelepiped classifier implemented in the General Electric Image-100 system. These techniques cause some degradation in the statistical basis of the assignment process (or information loss) and in their existing forms were unsuitable for LACIE. It is not impossible that better understanding of the data may lead to increased use of similar procedures in the future. The performance advantage of such algorithms is a function of their computational simplicity when compared to maximum likelihood.

d. In the data clustering application (see the paper by Johnson) used in ERIPS, a derivative of ISOCLS, a large number of iterations could be required before the desired stabilization of clusters was achieved. In the aggregate, these passes required CPU resources nearing those for classification. An adaptive clustering procedure was adopted; in this procedure, cluster centers were dynamically adjusted and clusters created during a pass, and the results of one such pass were input to the iterative clustering process. If the adaptive control parameters are properly chosen, the entry vectors to iterative clustering can be close approximations to the results of several iterative passes under the same initial conditions. Such intelligent preprocessing has the potential, never successfully adapted to LACIE, of reducing clustering processing requirements. Similar methods continue to imply attractive computational reductions.

e. Feature (or channel) selection is often employed to identify channels containing maximum useful information on the separability of defined classes. If the number of channels entering classification can be reduced, the number of arithmetic operations is quadratically reduced according to equation (2). Unfortunately, the computation time associated with most viable feature reduction techniques is great. In anticipation of later results, it is observed that both the divergence and the Bhattacharyya distance processors implemented in ERIPS and LACIE are burdensome and, as well, computationally intractable for meaningful development on special-purpose equipment.

f. Finally, it is noted that the most significant prospects for computational gains are likely to come from data compression and from demonstrations that the information loss from application of some compression procedure is minimal. No such pro-

cedure, however, has yet been shown applicable (workable) in the LACIE type of problem, although some progress has been made.

2. Equipment modification or expansion

a. An obvious route to increased capability is a more powerful mainframe; however, it will be shown subsequently that the devices of paragraph 2c generally obviate this consideration in the relevant signal processing environment.

b. Special (nonprogrammable) hardware has been implemented in numerous installations which enable performance of certain image analysis processes essentially at I/O rates. The singular disadvantage of such equipment is its inflexibility. The characteristic trend within LACIE, recognized at the outset, was continuing algorithmic modification and development of new processing procedures as the project evolved. "Boxes" capable of supporting maximum likelihood, for example, at extremely high rates would shortly have outlived their usefulness in the project, although this type of equipment can be cost-effective in facilities with definable charters implying long-term stability. The performance of such hardware classifiers is, nevertheless, exceptional. Several existing devices can perform the full Landsat frame classification exercise described earlier in less than 60 seconds, given appropriate I/O ports and drivers.

c. The final option to be considered, and the one selected for ERIPS augmentation, is division of labor. This concept can be realized in the general sense by any distributed or networking approach to a configuration; in the present context, however, the limitations of the conventional serial processor precluded inexpensive solutions to the LACIE problem by use of several serial devices. The availability of programmable high-speed computers of exotic architectures offered another alternative. Traditionally, these devices had their origins in signal processing applications such as radar, sonar, and geophysical data reduction. Normal applications included the Cooley-Tukey fast Fourier transforms, simultaneous solution of many differential equations, or various image enhancement or filtering operations. Several architectural concepts have been exploited for dramatic performance improvements in these compute-bound applications; in all such architectures, the conventional serial CPU is replaced by several (or many) arithmetic/logical units, arranged either in parallel or in series. More complex architectures generally derive from these two functional types; however, several implemented

systems can be viewed only functionally, and not necessarily electronically, as parallel or serial. The focus of the second portion of this paper will be the use of one such architecture in LACIE, with some discussion of different approaches for other applications.

PERIPHERAL PROCESSOR IMPLEMENTATION

Projected LACIE workload models and algorithm characteristics were jointly examined during 1974 to define features suitable for implementation in a peripheral device. The utility criterion was used by the investigated algorithm of the mainframe CPU, a direct product of its CPU requirement per use and its frequency of use. It was concluded that the bulk of algorithms or calculations which dealt with manipulations only of statistics or data sets in the aggregate were best left in the mainframe and that all processes requiring direct pixel operations in quantity would be transferred to the special-purpose device. These processes included the following specific modules.

Statistics

The statistics module was invoked both for direct computation of training field means and covariances and for establishing these same measures on the results of clustering runs (as described subsequently). It was not anticipated that the LACIE statistics processor would be an excessive user of system CPU resources, but rather that the in-line nature of its use within clustering would tend to complicate and reduce efficiency of the clustering procedure if statistics were left in the host mainframe. The computations of interest were

$$\mu_x(c) = \frac{1}{M} \sum_{k=1}^M x_k(c)$$

$$\Lambda_{xy}(c) = \frac{1}{M} \sum_{k=1}^M x_k x_y - \mu_x \mu_y$$

where M is the population of the class (field) under investigation, and the other terms are as defined

earlier (for eq. (1)). It was intended that many (≤ 60) such classes be processed simultaneously; i.e., accumulations to be made would be performed on fields in tandem as the processor advanced through the image only once.

Adaptive Clustering

The adaptive clustering procedure, briefly mentioned earlier, provided a means of grouping similar vectors from measurement space (where similarity is determined through a selected N -space metric). The technique was never exploited in the LACIE production (batch) system, although it was frequently invoked for continuing research into data behavior. The adaptive clustering algorithm is relatively complex and, because of its low significance to the project, will not be described here. However, it is germane to note that the algorithm, as ultimately implemented on the parallel processor chosen, was modified *conceptually* to exploit parallel features in the equipment architecture. Essentially, an algorithm which had been a serial, spectral clustering technique was converted into one which implicitly incorporated spatial data characteristics.

Iterative Clustering

The iterative clustering algorithm provides a means both for assigning measurement vectors to clusters and for evolving the statistical description of the reference clusters. The algorithm determines the "distance" of each measurement vector (or a set of such vectors) from the mean vector of each cluster and assigns each measurement vector to the "nearest" cluster. The statistics of all measurement vectors assigned to a particular class are determined and are used to modify the original clusters and cluster statistics. When the tasks described previously are accomplished, the algorithm is considered to have undergone one "pass." Usually, several passes are executed before the iterative clustering process is terminated.

The implementation of this algorithm using the peripheral processor required interpass operations by the host. The prime expression to be resolved by the peripheral processor was

$$D_c(x) = \sum_{i=1}^N |x_i - \mu_i(c)|$$

for each cluster center C . If $D_c(x) < D_k(x)$ for all $k \neq c$, then x was assigned to C .

Maximum-Likelihood Classification

The maximum-likelihood algorithm was described earlier. The computational kernel and the retention and generation of a classification map under defined thresholding conditions were viewed as appropriate for conversion to the special-purpose device.

Mixture Density Classification

The mixture density (sum of likelihoods) classification algorithm is similar to the maximum-likelihood algorithm. The distinction is a derivative of the class statistics definition made in each case. The maximum-likelihood classification algorithm uses a set of class statistics (mean and covariances) obtained for the population of the class as a whole; the mixture density function is formulated to treat a class as a union of independent subclasses, each of which is described as a population having a complete set of class (subclass) statistics. This representation tends, under careful preprocessing and definition of subclasses, to separate a population consisting of a multimodal distribution into several unimodal distributions and to improve the performance of the classification algorithm. This algorithm supplanted maximum likelihood as the principal LACIE classifier. The computation of the likelihood expression was complicated by the requirement to compute an exponential. Let

$$Q_k = \sum_{i=1}^N \sigma_i \sum_{j=1}^I \Lambda_{ij}^{-1} \sigma_j$$

and

$$H_k = S_k \exp(-Q_k/2)$$

Then,

$$H_c = \sum_{k=1}^M H_k$$

where Q_k is the quadratic form for subclass k , M is the number of subclasses in the class C , and other terms are as defined previously. The pixel is assigned to the class C if $H_c > H_d$ for $c \neq d$.

SPECIAL-PURPOSE PROCESSOR DESCRIPTION

The algorithms defined previously and performance benchmarks consistent with LACIE throughput requirements were specified in a competitive request for proposal in September 1974. In March 1975, award was made to the Goodyear Aerospace Corporation, which had proposed a Staran S-500 system manufactured by Goodyear as the LACIE special-purpose processor (SPP). The SPP was installed in December 1975 and made available for LACIE production use in March 1976.

The SPP (STARAN) system is based on a computer organization in which many identical operations are executed simultaneously; that is, it is a "single instruction stream, multiple data stream" processor. For example, in the SPP, an "add" operation can be executed simultaneously for 512 pairs of numbers. The parallel execution of an operation for many data pairs is made possible by employing many processing elements (512).

A top-cut diagram of the SPP mainframe is shown in figure 2. It consists of a conventionally addressed control memory for program storage and data buffering, a control logic unit for sequencing and decoding

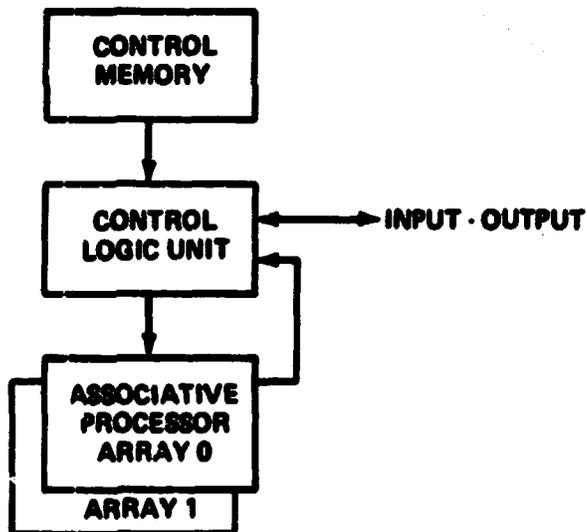


FIGURE 2.—SPP top-cut architecture diagram.

instructions from control memory, and two associative array modules.

The high processing and throughput speeds achieved by the SPP resulted from the capabilities of the associative array (fig. 3). Each SPP array contains 256 simple processing elements. All processing elements (PE's) perform the same operation at the same time, but each processing element acts on independent data. Thus, in each SPP array, 256 independent data streams can be processed.

Array memory used to support the PE's is composed of 256 words having 256 bits. Multiple access paths between the PE's and the bit memory locations provide ready access to 256 different bit patterns in the array. Two access "stencils" are shown in figure 3.

To further enhance the data routing capability of an array module, an alignment, or permutation, network in the machine provides a flexible interconnection between processing elements. The multiple processing elements, the multidimensional access memory, and the permutation network give the SPP the flexibility to be useful for a wide set of problems.

The LACIE algorithms were all well suited to the SPP architecture because they had an inherent parallelism resulting from a given computation being performed on all picture elements (pixels) in an image. Since computation associated with each pixel was the same for a given algorithm, it could be implemented in a single-instruction stream. The LACIE algorithms thus fit the single-instruction, multiple-data-stream concept that is part of the SPP architecture. It should be observed that, although the SPP is a

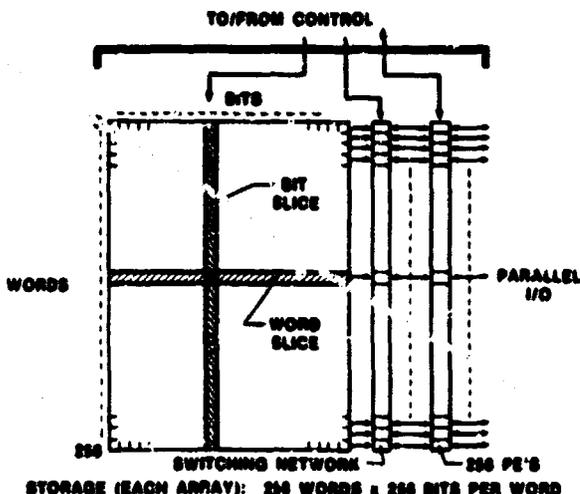


FIGURE 3.—SPP associative array.

full array processor, the LACIE applications exploited the architecture primarily as a strictly parallel device. The details of the SPP architecture and arithmetic operations are described elsewhere (ref. 2); it will suffice here to describe briefly the LACIE installation.

In the configuration of figure 1, an SPP interface unit (IU) is the communications link between the SPP and the IBM 360/75, which shall subsequently be identified as the host. The host port is a conventional high-speed (to 1.2 Mbytes/sec) IBM 2860 selector channel. The IU contains, on the SPP side, an I/O controller (IOC) interfacing to (1) a buffered I/O channel to the SPP control memory used in LACIE for large data transfers (imagery) and (2) an external function channel, which generally carries control signals governing processing sequences. The IU interface to the host is through an interface module (IFM), which is a functionally self-contained module that communicates with and transfers data to and from the host channel. The module handles all control line sequences required for selection and command execution on the host channel. There are two IFM's of identical design, one for each host channel. Each can operate independently of the other; however, only one can be logically connected to the IOC at a time. Again, additional details are specified in reference 2.

Performance Summary

The LACIE performance advantages of the SPP over the 360/75 are functionally dependent on (1) algorithm organization (the ability to exploit parallelism), (2) number of data channels, (3) number of signatures (classes/clusters), (4) number of pixels (vectors) per quantum of system workload (job), (5) SPP setup time (formatting of vector transfers to and from the SPP), and (6) data base retrieval rates. The effects of these drivers are mutually dependent and difficult in many cases to distinguish. The sampling of results provided subsequently will be generally treated in terms of these driving functions, with only a few specific comments in order as they relate to computational idiosyncracies of the individual algorithms. Some preliminary remarks follow.

1. To repeat earlier comments, a LACIE image consists of 22 932 data vectors or as many as 4 sets of such (4-channel) vectors. A maximum of 60 signatures for classification or clustering may be defined; practically, these values remain between 30

and 40. Other system or algorithm delimiters are generally exercised across their entire range. Extensive testing of the SPP software in the production environment confirmed both logical and performance timing behavior of the system throughout the range of software specifications.

2. The historical driver of the 360/75-based LACIE/ERIPS performance was, as stated, the CPU. In the SPP configuration, principal limitations on throughput are, in practice, the retrieval functions from the imagery storage medium, the IBM 2314 disks. Only on jobs of significant complexity, specifically classification exercises on 12 channels or greater with discrimination of more than 20 classes, does the system perform in an SPP CPU-bound state. Development of an imagery data retrieval technique (ref. 1) has ensured optimal exploitation of the disks for the peculiarities of the LACIE application, but the disks generally remain the system driver. Direct access to the imagery on the ITEL 7330 data base would permit significant throughput improvements for most LACIE jobs; such implementation may be made at a later date, as necessary, but current performance (although suboptimal because of I/O) satisfied existing constraints on resources.

3. SPP arithmetic is field-length dependent in performance characteristics. The LACIE applications specifications dictated effective equivalence with 360/75 floating-point arithmetic results for purposes of continuity; this stringent requirement on the SPP, which was achieved, is not statistically justified on the basis of measurement vector variance, and legitimate results of processing can be obtained by way of shorter fields than those employed with significant performance advantage.

4. In a comparison of pre-SPP and post-SPP timing, the control base was modified to some extent in software that could have affected 360/75 applications performance; that is, certain 360/75 system software routines were optimized at the time of SPP implementation. These changes could, to some extent, be reflected in the timing figures given subsequently for pre-SPP algorithms, but the figures shown display pre-SPP results without such system changes. Further, as cited earlier, the adaptive clustering algorithm was extensively and theoretically modified when incorporated into the SPP; the objective was to maximize the benefits of parallelism and to use spatial as well as spectral data characteristics. The result has been a technique of improved convergence and stability, but no direct (timing) performance comparisons can be made with pre-SPP results. Com-

ments and conclusions pertinent to each implemented algorithm are given in the following paragraphs.

Statistics

Statistical processing ordinarily occurs fairly rapidly in the LACIE system and was included in the SPP development for consistency with the notion that all pattern-recognition processors of a pixel-dependent type would be SPP-resident. Also, the STATS routine is invoked in the body of ITCLUS; SPP implementation reduced organizational complexities. LACIE characteristics, however, include occasional and numerous small (<20 pixel) fields on which processing must be performed; SPP performance is severely compromised by system overhead on such jobs. Occasionally, SPP STATS is slightly slower even than the 360/75 STATS, but the SPP rate has never been less than 90 percent of 360 rates (on tasks of 4 to 5 seconds). On larger fields and on large channel set jobs, the SPP performance advantage reaches about 3 to 1; but, because the process rarely requires more than 20 seconds on the 360 in the most complex LACIE cases, 360/75 execution would not be deleterious to the system.

Clustering

An adaptive/iterative clustering exercise was defined for a benchmark as follows: 500 × 200 (10⁵) vectors, 16 channels, to be distinguished into 10 clusters in an artificial data set; results: non-SPP required 35.1 minutes, SPP required 37 seconds, a per-

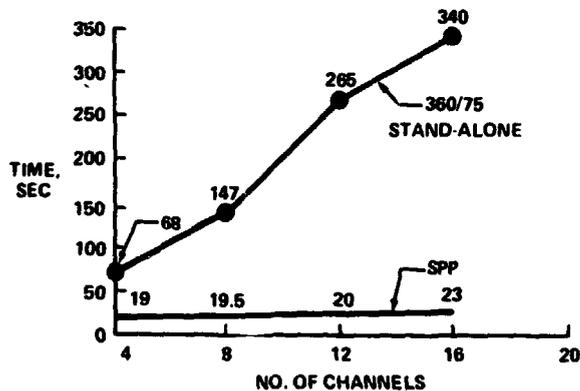


FIGURE 4.—Iterative clustering timings.

formance gain of 57 to 1. Figure 4 shows typical LACIE results for 22 932 vectors, under various channel set sizes and with implicitly discriminated clusters. Performance gains, less than for the benchmark, reflect system overhead penalties for smaller data sets but demonstrate the I/O constraints driving the SPP on complex applications and significant performance improvement (as great as 15 to 1) normally experienced.

Classification

A classification benchmark was defined as follows: MAXLIK, 4 channels, 10 classes, 2340 × 3240 vectors (7.58 million pixels); results: pre-SPP—105 minutes, SPP—8.15 minutes, a performance gain of 13 to 1. Figure 5 shows MIXDEN results on LACIE images of 22 932 vectors under various channel set sizes and defined signatures. As in clustering, system overhead diminishes performance factors on smaller segments of data, although the trends are clearly I/O driven. MAXLIK, essentially identical to MIXDEN organizationally, produces timings approximately 20 percent less for both SPP and non-SPP.

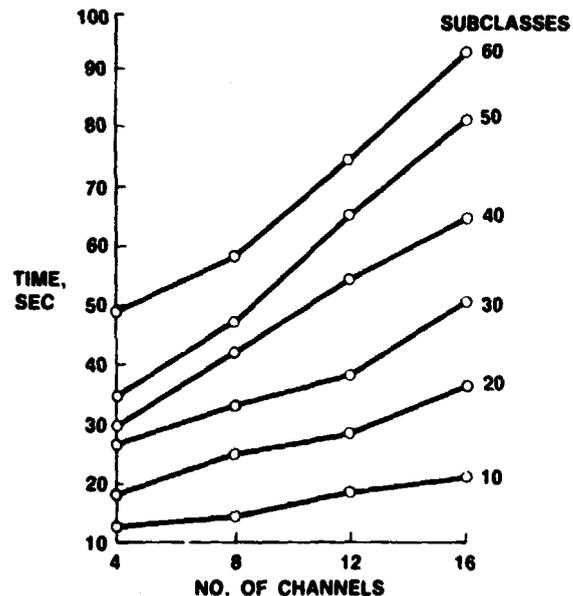


FIGURE 5.—LACIE classification timings with SPP (mixture density).

CONCLUSIONS

The SPP has satisfied and exceeded performance specifications originally defined. The system performance can be significantly improved, when necessary, by modifications in the host data retrieval technology without impact to the SPP software or addition of arrays. Within the LACIE context, the most tangible improvements have been in processes (clustering, classification) that were previously prohibitively expensive users of host resources. Because of host I/O constraints, the statistics function on the SPP, as anticipated, offered relatively little improvement except in exotic test cases involving large data sets.

Table III is a summary of the performance gains associated with the SPP, both as a function of SPP execution only and in terms of total system throughput, for representative LACIE jobs. It is to be recalled that some LACIE jobs are neither suitable for nor implemented on the SPP but are still considerable users of system resources. Such tasks obviously should be examined in detail for performance improvements in production environments.

As anticipated before the SPP procurement, additional requirements, both modifying existing algorithms and proposing entirely new analytic techniques, were developed for LACIE support. Because of serial device limitations, such schemata previously have been useful only on limited amounts of data.

In summary, the LACIE environment, including high throughput requirements in a quasi-production system and a requirements flux in a technologically and theoretically developing discipline, demonstrated the cost-effectiveness and utility of a programmable SPP. It has been shown that the image

TABLE III.—LACIE Throughput Improvements^a

[Time, sec. per segment]

No. of channels	360/75	With SPP	Ratio	Only SPP tasks
4	386	201	1.92	3.6
8	921	260	3.54	7.5
12	1941	323	6.00	13.25
16	2738	396	6.91	14.8

^a Assuming 40 classes per segment, SPP performance of STATS, ITCLUS, and MIXDEN modules.

processing tasks conventionally considered unmanageable (in quantity) are tractable with such devices as described here or available from other sources. It is believed that the primary foci of attention in image processing in the long term should not be, in general, the computational tasks, but rather data management, storage, and traffic control for large numbers of large data sets.

REFERENCES

1. Pape, A. E.; and Truitt, D. L.: The Earth Resources Interactive Processing System (ERIPS) Image Data Access Method (IDAM). Symposium on Machine Processing of Remotely Sensed Data, June 29, 1976, Purdue University, West Lafayette, Indiana.
2. Ruben, S.; Faiss, R.; Lyon, J.; and Quinn, M.: Application of a Parallel Processing Computer in LACIE. 1976 International Conference on Parallel Processing, Wayne State University, Detroit, Michigan, August 1976.

A Look at Computer System Selection Criteria

E. W. Poole,^a F. L. Flowers,^a and W. I. Stanley^a

ABSTRACT

The question of computer system selection criteria is growing more complex as the cost of centralized systems decreases and the performance of distributed systems increases. In many cases, the discussions become emotional and evaluations are made on criteria which do not address the technical merits of a system solution to a specific problem.

Identifying the criteria involved in the selection process is not difficult. The complexity arises in objectively evaluating various candidate configurations against these criteria based on the user's specific requirements. This paper describes a process (model) which can be used to formalize the selection process. The process consists of two major steps.

1. Verification that the candidate configuration is adequate to meet the user's processing requirements
2. Determination of an overall system evaluation rating based on cost, usability, adaptability, availability, etc.

An Earth resources data system of the future has been used as an example in the application of the weighting factors to a set of user requirements, and the LACIE/Earth Resources Interactive Processing System (ERIPS) package at the NASA Johnson Space Center (JSC) provides an example of the applied procedure. This approach does not eliminate all judgment from the procedure and therefore is still subject to some discussion. It does force the discussion away from emotional arguments into an orderly set of decisions which provides a specific solution to the problem being addressed.

^aIBM Federal Systems Division, Houston, Texas.

INTRODUCTION

The Problem Statement

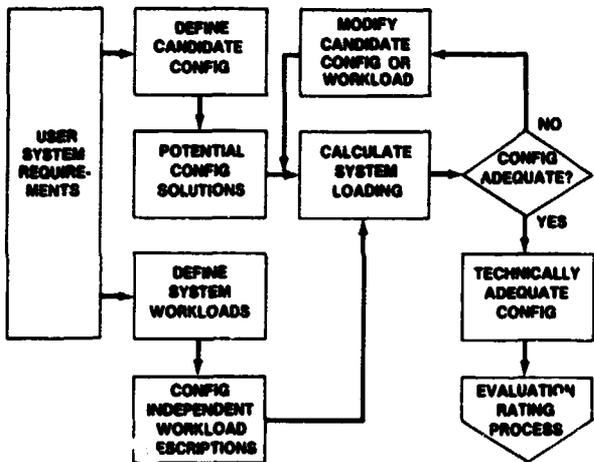
Discussion of computer system selection criteria is complex today and is growing more complex as technological advances affect price and performance of computing systems. The problem is further complicated by the fact that there are few system designers or analysts who can speak as experts on both centralized and distributed system architectures. Discussions on selection criteria tend to be composed of some fact and of some emotion and are of limited value. The burden of selecting a computer system falls to the buyer. His ability to judge the technical merits of a particular hardware architecture for an application will be tested by many discussions which require the use of every available tool to establish the relevance of a solution to the problem being addressed. In today's bag of tools, high value is placed on historical precedence and on rules of thumb derived before the advent of current technology. Extrapolating partly applicable models to produce evaluation data can magnify the problem by comparing unlike quantities as though they were alike instead of realizing that the real data point from this activity is qualitative.

Identifying the parameters involved in the selection process is not difficult since most system designers will agree on what they are. The complexity arises in developing the model that will generate the weighting factors needed to evaluate candidate configurations for solving a particular problem. Since the basic question will become more and more common, a model which can be used in the selection process has been developed.

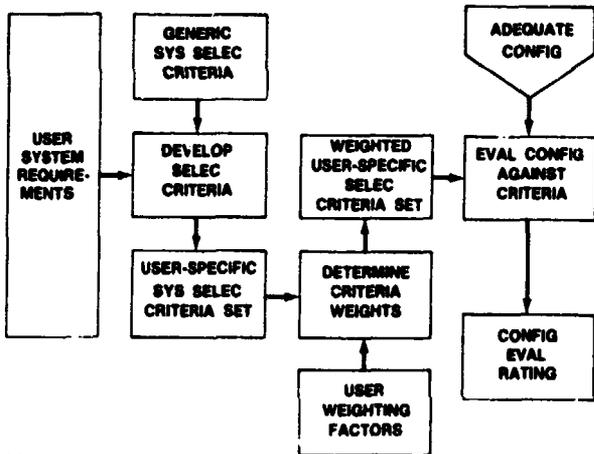
The parameters involved in the evaluation process and the impact of the technical approach selected on their relative importance to the solution will be reviewed. The decisionmaking process is structured into an orderly review which provides a specific solution to the problem being addressed and reduces the emotional aspects of this activity.

System Selection Process Definition

The system selection process is a two-step activity in this approach to the evaluation of candidate configurations for the solution of a stated problem (fig. 1). Step I is the testing of the candidate configuration for technical adequacy. This step involves significant requirements analysis for effectiveness. The problem



(a)



(b)

FIGURE 1.—System selection process flow diagram. (a) Step I. (b) Step II.

to be solved must be divided into functional elements which are described in terms of data in and out, instructions executed, and overall size. The functions are mapped into processes and the probability of a function executing in a process is addressed. The processes are then incorporated into workloads which reflect user activities, and workload baselines are established.

Candidate configurations are then defined and system loading for the various workload baselines is evaluated against them. Acceptable configurations are retained, and others may be redefined and evaluated again. Only candidate configurations which support system definitions will proceed beyond this step in the process.

The establishment of the technical adequacy of a candidate configuration in this manner provides a system-level solution set, the elements of which can be compared using other selection criteria. Configurations which do not meet the performance requirements or the system definition are eliminated from further consideration.

Step II of the system selection process is the derivation of an evaluation rating for each technically adequate candidate configuration. The most important element of this step is the definition of a set of well-described, detailed selection criteria which can be viewed objectively and which specifically address the user requirements to be supported.

The user of a system who has a requirement for real-time processing is concerned with a different set of capabilities from that of the interactive researcher or the time-sharing inventory control administrator. Regardless of the user, the following generic criteria set will support the derivation of a user-specific system selection criteria set.

1. Adequacy—Does this configuration fulfill the requirements?
2. Cost—Over the life of the system, is the cost reasonable?
3. Adaptability—Is the system capable of growth to new applications?
4. Availability—Is the system maintainable and are partial support options possible?
5. Transportability—Is the system composed of standard hardware and software which is generally available?
6. Usability—Can the users, developers, and operators perform their activities on the system with ease?

The user-specific system selection criteria set is categorized to assist in the weighting process. Each

category is assigned a weight, and subfunctions within each category are also assigned weighting factors.

Numerical evaluation ratings are applied to the various elements of the selection criteria for the acceptable candidate configurations. Weighted evaluation ratings are summed for a category, and the results are weighted appropriately. The resultant ratings are then summed to produce an overall evaluation rating for a configuration which can be evaluated against ratings of other candidate configurations. Then, based on system availability on a schedule consistent with project needs, a selection can be made.

A summary of the methodology is presented in the following section. Detailed descriptions of Steps I and II are contained in appendix A and appendix B, respectively. A comprehensive example demonstrating the configuration adequacy step is contained in appendix C. This example is an analysis of the LACIE/ERIPS currently operational at JSC.

THE EVALUATION PROCESS

When attempting to select a data processing system to address a particular problem, the buyer quickly discovers that the process is very complicated. Many factors must be analyzed and evaluated to make an intelligent decision. Examples are the following.

1. Architecture—There are usually several architectural approaches available to solve a problem. A number of questions must be answered. Should the system processing be performed by a centralized computer only or should some processing be distributed throughout the system? Should the system support on-line and batch processing concurrently? Should remote terminals be supported? If so, what kind of communication capabilities are required? How much intelligence should be distributed to communications processors? Should the data base be centralized or distributed? What kind of data base management capabilities are required?

2. Hardware manufacturers—There are many data processing equipment manufacturers, each offering a wide range of hardware capabilities and options. The buyer must decide whether his job can be done with the proposed hardware. For example, does each processor have the necessary processing speed? Are the main storage and the auxiliary storage of adequate size and speed? Are the on-line storage devices of adequate size and speed? Can the

job be handled with standard interface devices? If not, what are the requirements for any special-purpose devices?

3. System software—The vendor-supplied software, such as operating systems, utilities, data base managers, and communications controllers, varies widely in functions and capabilities. Some vendors offer limited or no software support for their hardware. The buyer must determine whether the appropriate functions are provided and the extent to which they meet his requirements.

4. Application software—There are many application software packages available for sale or lease which may be suitable for performing part or all of the buyer's data processing. Each package must be reviewed to determine the functions available, the extent to which they meet the requirements, performance and usability characteristics, etc.

5. Intangible attributes—There are many intangible attributes of a system that must be evaluated; for example, maintainability, reliability, availability, flexibility, operability, usability, and ease of technology transfer.

6. Cost—There is a wide range of costs for both hardware and software products. The buyer must evaluate cost against the functions provided, the functions required, and the functions potentially desirable in the future. He must also evaluate cost against the intangible attributes mentioned previously.

To further complicate the process, evaluation of various capabilities frequently produces an answer which is not clearly black or white but rather a shade of gray. Evaluations and opinions of several experts in different areas are usually required. Even after a thorough evaluation is completed, it is often difficult to summarize conclusions and evaluation ratings in a manner that can be easily understood by the decisionmaking levels of management.

It is obvious that there is a need for a systematic and simplified approach to the problem of computer system selection. This paper describes a technique for formalizing the process by creating a step-by-step sequence of activities. The reader should not mistake this technique (model) for an "automatic selection program." Human judgment is still required. The model simply requires that judgment be made in an orderly fashion and reduces evaluation of various elements to a common unit of measure. The end result is a single numerical evaluation rating for the system.

The model actually consists of two separate but

related models, each having a different objective. The models are referred to as Step I and Step II and are defined as follows.

1. **Step I: Configuration Adequacy**—This model predicts system performance at a fairly gross level. Based on a predefined workload, loading on each system component is calculated. Configurations incapable of handling the workload are either eliminated from further consideration or modified until they are adequate. All candidate configurations passing this test are then carried to Step II.

2. **Step II: Evaluation Rating**—This model calculates an evaluation rating for the system based on the selection criteria defined by the user. The selection criteria are carefully structured in a hierarchical fashion and a relative importance (weight) assigned to each. The user then evaluates each item in the lowest level of the criteria hierarchy on a scale from 1 to 10, thus converting all evaluations to a common unit of measurement. The model then calculates the weighted evaluation rating for each element of the selection criteria and sums them in a hierarchical manner. The final result is a single numerical evaluation rating.

Each of these steps is discussed in more detail in the following sections.

The reader should understand that the models discussed here are neither discrete simulation models nor analytical models. Thus, they are not intended for detailed system design evaluation and performance prediction. Instead, they are at a higher level, intended to help the user answer those questions he encounters when trying to select a configuration. The models should help eliminate those configurations that cannot handle the job and quantitatively evaluate the relative merits of those that can.

Step 1: Configuration Adequacy

The first step in evaluating a configuration is to determine whether it is adequate to handle the processing load. The objective is to eliminate from consideration those configurations that do not meet performance requirements. Each component of the system will be analyzed to determine the adequacy of such parameters as processor speed, memory size, data transfer rates, and on-line storage capacity. There are three major activities in this step.

1. **Define workload**—Determine the amount of processing that must be performed during a specific time period.

2. **Define candidate configurations**—Diagram the configuration, identifying each hardware/software component and the characteristics of each.

3. **Calculate system loading**—Compute consumption of system resources, elapsed times, and component utilizations and then analyze total system performance.

Each of these activities is described in greater detail in the following sections. To understand the text, however, it is important that the reader first understand the definition of terms used throughout the discussion. Table I contains a list of terms, the definition of each, and examples. A step-by-step detailed description of all three major activities is contained in appendix A.

Workload definition.—The first activity in determining configuration adequacy is to define the system workload for typical time periods to be analyzed. This workload will act as the yardstick against which each candidate configuration is measured. This activity is performed only once and the results saved for later use when system loading is

TABLE I.—Definition of Terms

Term	Definition	Examples
Resource	An element of the computer system required to perform a function	Central processing unit, direct access storage device, memory
Resource usage variables	Parameters on which the amount of resource consumed is dependent	Number of fields, number of classes
Function	A logical user action for which resource usage can be defined in a manner independent of the configuration	Clustering, classification, dot summary
Process	A typical sequence of functions that represents the activity of a particular user	Production analyst, researcher, software developer
Benchmark	A combination of processes (users) that is a typical representation of the total system workload during a specific period	2 production analysts, 1 researcher, 3 software developers, etc.

calculated. The workload is defined in terms of system resources consumed. This definition process is performed in a hierarchy of three levels: functions, processes, and benchmarks. A flow diagram depicting these actions is shown in figure 2.

Function definition: The configuration evaluator first looks at all work to be performed by the system and identifies each logical user action, called a function. The amount of system resources used by each function must then be defined in terms which are independent of a particular configuration. For example, the processor execution time is defined as the number of machine instructions that must be executed to perform that function rather than a fixed process time. The model converts this value to process time based on the instruction execution rate of the processor being analyzed. Similarly, the amount of input data and output data is expressed in total bytes rather than in some configuration-dependent measurement such as number of records or number of tracks.

In most cases, the amount of system resources consumed by a function is not a constant but rather a variable dependent on certain input parameters. For example, the amount of processing performed for the "classification" function may depend on the

number of fields, subclasses, and channels (fig. 3). The amount of main storage necessary to execute this function may depend on the number of fields. Therefore, resource usage for these functions must be defined by resource usage variables like those shown in the example in figure 3.

Process definition: The next level in the hierarchy is the "process." The process is a typical sequence of functions which represents the activities of a particular user session. The objective is to identify a set of "typical" users which can be used to construct the total workload. The system evaluator must identify all potential users of the system and the characteristics of each. For this discussion, a "user" may be thought of as a work session by an individual. If the same person uses the system in a significantly different manner at various times, then each session might be represented by a different process definition. For example, one user may be a research analyst evaluating various classification techniques. One work session may be a batch job run consisting of a compile-load-and-go sequence. At another time, the same person may perform some actual classification exercises in an interactive mode at a display. These two sessions would be represented by different process descriptions.

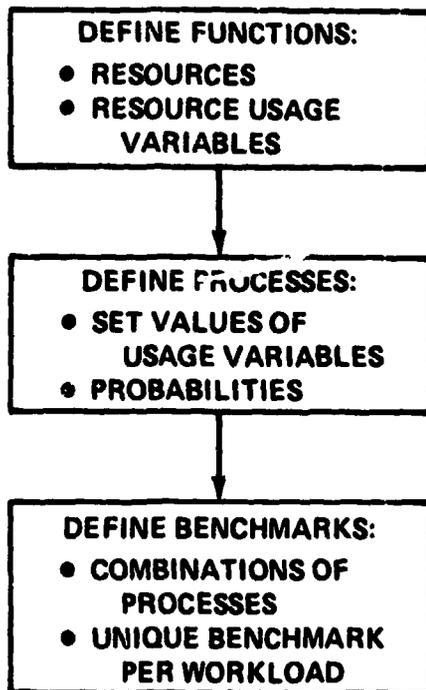


FIGURE 2.—Workload definition flow diagram.

FUNCTION	RESOURCE DEFINITION VARIABLES
CLASSIFICATION	CPU: $608\ 000 + (\text{NO. FIELDS}) \times (45\ 000 + (\text{NO. SUBCLASSES}) \times (42\ 500 + (\text{NO. CHANNELS}) \times 2500))$
	MEMORY: $16\ 700 + 80 \times (\text{NO. FIELDS})$
I/O:	$46\ 250 + 22\ 932 \times (\text{NO. CHANNELS}) + 240 \times (\text{NO. SUBCLASSES}) + (\text{NO. FIELDS}) \times (120 \times (\text{NO. SUBCLASSES}))$

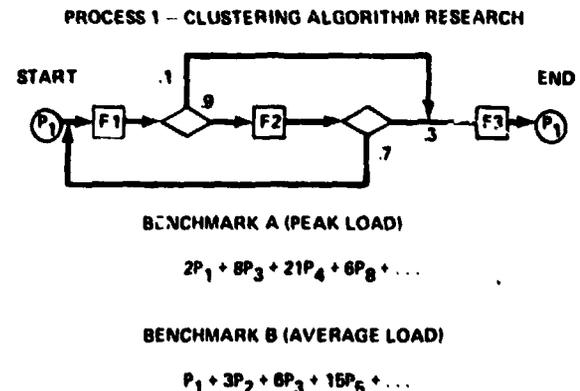


FIGURE 3.—Workload definition example.

After identifying the processes, the evaluator must determine the characteristics of each. The following information must be specified.

1. **Functions**—Identify the functions executed by this process and the order in which they are executed. This can be done by a functional flow diagram as shown in figure 3.

2. **Probabilities**—Since a work session does not always consist of a fixed sequence of functions, the evaluator must determine the probability of moving from one function to the next. This information is used by the model to determine the number of times each function is executed.

3. **Parameters**—The value of each parameter in the resource usage variables must be set. For example, the number of fields, subclasses, channels, etc., must be determined for this user session.

Benchmark definition: The highest level in the hierarchy of workload definition is the total system workload for a specific time period. This is called a benchmark. In the model, it is represented as a combination of processes (users). The evaluator must first determine which time periods he wishes to analyze. He would normally select a period that represents the average workload and another that represents the peak workload. If so, he will define a separate benchmark for each period.

Since all the detailed information is defined at the function and process levels, the benchmark definition is simply a specification of the number of times each process is executed. It can be specified as a mathematical expression as illustrated in figure 3.

Candidate system definition.—The second activity in determining configuration adequacy is the definition of candidate configurations. It consists of diagramming the configuration, identifying the hardware components, determining hardware characteristics, assigning maximum utilization values, and assigning functions to the appropriate hardware components. Each of these actions is depicted in figure 4.

The candidate configurations to be evaluated would normally consist of the solutions proposed by the various responding vendors. The buyer may, however, choose to determine the system architecture himself and specify the desired configuration in the request for proposal. In this case, each vendor would merely specify the equipment to use for each system component. This approach has several disadvantages of which the buyer should be aware.

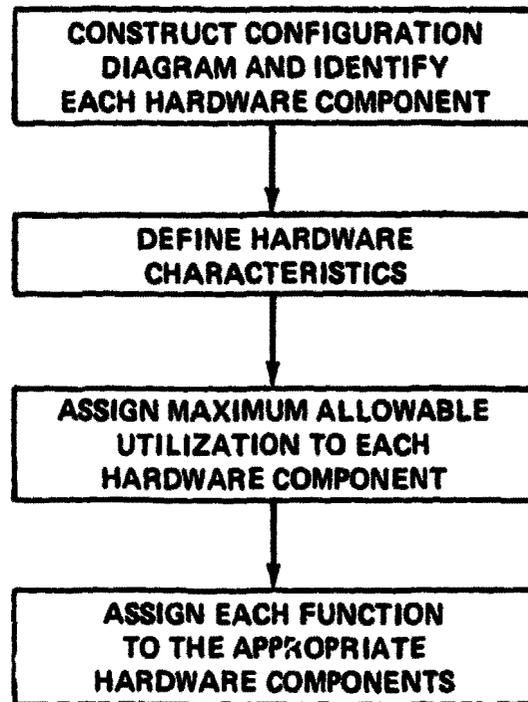


FIGURE 4.—Configuration definition flow diagram.

1. It discourages individual vendors from seeking the best architecture to address the problem.
2. It may not allow vendors to propose the hardware/software systems containing recent advances in the state of the art.
3. It usually results in more expensive configurations submitted by vendors.

Configuration diagram: The first action is to create a detailed diagram of the configuration showing each system component and its interface with other components. Each component must be identified by its name, model number, etc. For each processor, the main storage size and secondary storage size, if applicable, must be indicated. The diagram must also identify the number and type of each data channel, input/output (I/O) device, terminal, etc. Any special-purpose equipment and/or interface should be clearly identified.

Hardware characteristics: The characteristics of each hardware component identified previously must now be determined. The type of information needed includes the following.

1. Instruction execution rate for each processor
2. Memory speed
3. Data channel transfer rates

4. Disk characteristics such as rotation time, average seek time, data transfer rate

5. Display characteristics such as screen size, data transfer rate

These data are used by the model in calculating elapsed execution times and component utilizations.

Maximum utilization: To avoid overloading any hardware component, the user may specify a maximum allowable utilization percentage for each component. The model will not allow that component utilization to exceed the input value. The model first computes elapsed execution times and component utilizations for the specified workload without regard to this maximum value. A check is then made to determine whether the calculated value exceeds the input maximum. If so, the elapsed times are adjusted to bring the component utilization down to this maximum value.

Function assignment: By analyzing the configuration and the functions to be performed, the evaluator must now assign each function to the appropriate hardware component. This is a trial and error process and represents only the initial best estimate of function assignment. If the configuration is shown to be inadequate, some functions may have to be reassigned.

System loading calculation.—The last activity in determining configuration adequacy is the calculation of system loading and analysis of the results. This procedure includes the calculation of resource consumption for each process, calculation of total resource utilizations and elapsed execution times, adjustment of elapsed times if any resources are overused, and analysis of total system performance. These actions are shown as a flow diagram in figure 5. Each is discussed in detail in appendix A and in summary form in the following paragraphs.

Resource consumption calculation: The objective of resource consumption calculation is to determine the amount of each system resource consumed by the benchmark being analyzed. Resource consumption is first calculated by function, then by process, and finally for the total benchmark; the value of each resource usage variable is calculated using the specified input parameters for each process. System overhead is estimated and added to the totals. The result is the total active or "busy" time of each resource.

Elapsed time/utilization calculation: The evaluator is now ready to determine the elapsed execution

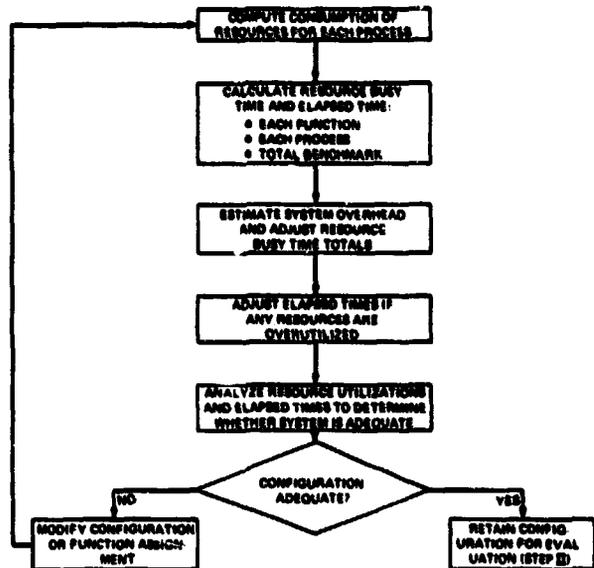


FIGURE 5.—System loading calculation flow diagram.

time and resource utilizations. To obtain the total elapsed execution time for the entire benchmark, each function must first be evaluated individually to determine stand-alone execution times. This procedure requires estimation of all I/O inactive or "wait" time based on the I/O device characteristics, data base design, etc. It also involves consideration of simultaneous use of multiple resources that may be used by a function.

The elapsed time for each process can now be obtained by simply summing the elapsed times for each function in that process. Similarly, elapsed execution time for the total benchmark is calculated by summing the times for the appropriate processes. The percentage utilization can now be calculated for each system resource by dividing the total busy time (for a resource) by the total elapsed time.

Elapsed time adjustment: If any resource utilization exceeds the maximum value (percentage) specified by the evaluator, then elapsed times for all processes using the overused resources must be increased until each resource utilization value drops below its maximum. A "stretchout factor" is computed to determine the amount of the increase for each process. This calculation is described in detail in step 31 of appendix A. After elapsed times have been adjusted, elapsed times for the benchmark and resource utilizations must be recalculated.

System performance analyses: The evaluator can

now make a judgment about the adequacy of the candidate configuration by first making a reasonableness check on the resource utilizations. Low utilizations across all workloads might suggest that the configuration is overpowered. Maximum utilizations on a majority of resources may mean the hardware is not fast enough.

If resource utilizations appear reasonable, the evaluator can compare the total benchmark elapsed execution time to the predetermined time required. Elapsed times of each process and function should also be analyzed. If the configuration is judged to be adequate to handle the workload, it is saved for further analysis in Step II. If not, the evaluator has three alternatives.

1. Augment the configuration in those places where resource utilization is very high.
2. Modify assignment of functions to system components.
3. Eliminate the configuration from further consideration.

Step II: Evaluation Rating

The second step in evaluating a system is to evaluate it against a set of detailed selection criteria to obtain an overall system rating. The objective of this model is to reduce the evaluations of all the various system elements to a common unit of measurement so that numerical methods can be used to combine them. The evaluation is based on a set of user-defined selection criteria representing the requirements for this system. Evaluation is performed on such items as cost, maintainability, usability, flexibility, and operability. Only those configurations judged to be technically adequate in Step I are evaluated in Step II. The entire process and the relationship of Steps I and II are shown in figure 1.

Three major activities are involved in this model.

1. Develop selection criteria—Develop a detailed set of criteria against which each candidate configuration will be evaluated and determine the relative importance of each item.
2. Determine numerical ratings—Evaluate the system against each element of the selection criteria and assign numerical ratings for each element.
3. Calculate evaluation rating—Calculate the weighted evaluation rating for each level in the selection criteria hierarchy and also for the total system.

A detailed description of these activities is contained in appendix B. The following sections contain a summary-level discussion of each.

Selection criteria development.—The first activity in Step II is to develop a detailed set of selection criteria against which each configuration will be evaluated. This activity consists of three actions.

1. Develop the user-specific selection criteria list.
2. Separate criteria into categories.
3. Determine the relative importance of each item and assign a weighting factor.

This process is depicted in flow-chart form in figure 6.

User-specific selection criteria: By analyzing system requirements, the evaluator must develop a detailed set of selection criteria specific to his system. These criteria are developed by reviewing the generic set of criteria, selecting the applicable groups, and then expanding them to the appropriate detail. The detail is acquired by using the hierarchical decomposition technique. Each major area is analyzed individually and broken down into its elements, each of which is in turn broken down into its subelements. This process is repeated until each selection category is decomposed into a set of subelements which can be clearly evaluated for each candidate configuration. There is no restriction on the number of levels of decomposition. This hierarchy forms the logical levels for later combination of weighted evaluation ratings.

Figure 7 is an example for an Earth resources data system. Shown are five major areas of selection criteria and several elements of each. Each element must be analyzed by subelements before ratings can be generated. Figure 8 is an example of criteria ratings and also of the level of the criteria subelements that might be evaluated.

Criteria categorization: After definition of his selection criteria, the user may find it helpful to categorize them in a different manner. This step is optional and simply reorders the criteria set in a way which better relates selection criteria to system requirements or simplifies the weighting process described subsequently. Figure 9 is an example of criteria categories for an Earth resources data system.

Criteria weighting: The user is now ready to determine the relative importance of each criterion by assigning a weighting factor (a percentage of the total) to each category, element, subelement, etc., in the selection criteria set, starting at the highest level

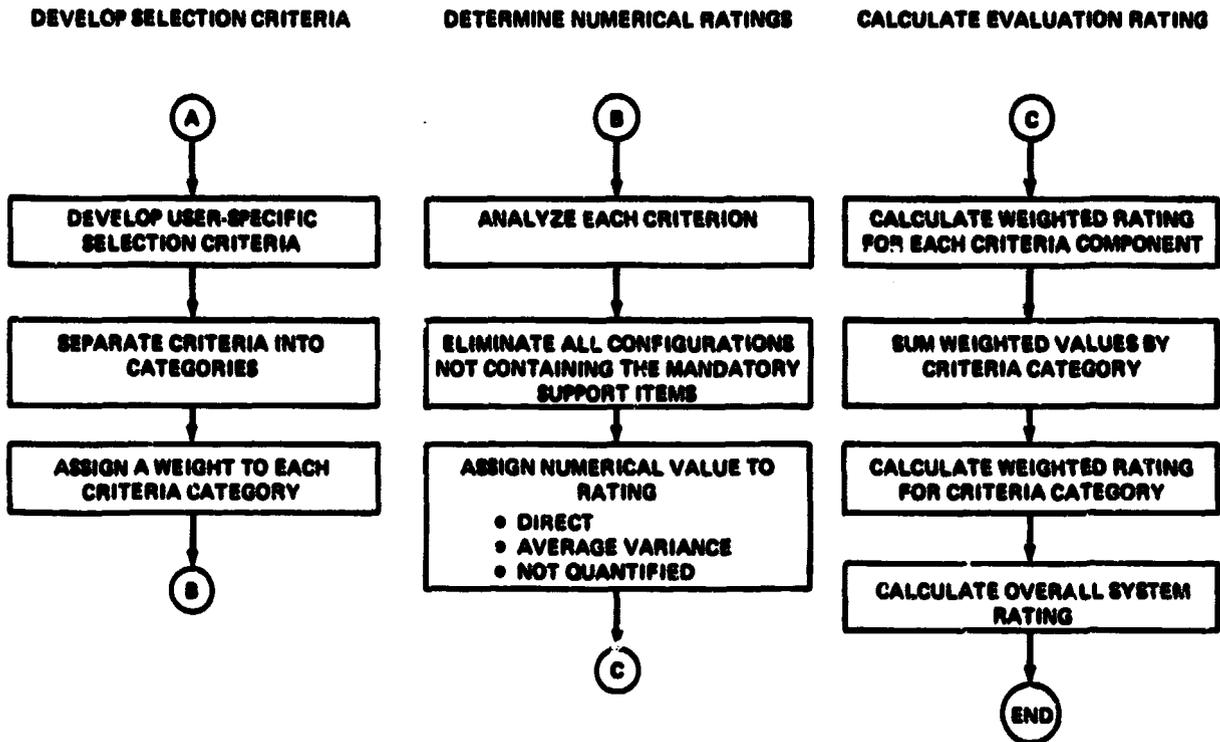


FIGURE 6.—Evaluation rating model flow chart.

in the hierarchy and working downward. The weighting factors are assigned such that the total for each level in the hierarchy is 100 percent. That is, the sum of all categories (top level) must be 100. For each category, the element weights total 100. Within each element, the subelement weights total 100, etc. An example of criteria weighting is shown in figure 10.

Numerical rating determination.—The second activity in evaluating a configuration is to determine its relative merit or value for each criterion. This procedure requires analysis and judgment by the evaluator. The relative value of each item must be determined on a scale from 1 to 10. These data are then used in calculating the weighted values for each criterion. This activity requires that the evaluator (1) eliminate any configuration that does not contain the mandatory support items and (2) assign a numerical rating value for each criterion. Figure 6 shows these actions in flow-chart form.

FOR EACH TECHNICALLY ADEQUATE CONFIG EVALUATE:

- COST
 - INITIAL DEV COSTS (NONRECURRING)
 - OPS COSTS (RECURRING - 10-YR LIFE CYCLE)
 - CONFIG FLEXIBILITY
- EXPERIMENTAL FLEXIBILITY
 - NEW TECHNIQUE/TECHNOLOGY EVAL FLEXIBILITY
 - EASE OF TECHNOLOGY ACCESS
- TECHNOLOGY TRANSFER
 - RELEVANCE TO USER REQS
 - TRANSITION AND OPS COSTS
 - SUPPORT TECH TO AID NEW USER
- CONFIGURATION USABILITY
 - USER PRODUCTIVITY
 - EVOLUTIONARY DEV CAPABILITY
 - OPS ACCEPTABILITY
- SCHEDULES
 - TIMELY DEV SUPPORT
 - PLANABLE CONFIG EXPANSION

FIGURE 7.—Example of user-specific selection criteria for an Earth resources data system.

CRITERION	WEIGHT, PERCENT	RATING
GENERAL SUPPORT CAPABILITIES	30	
• OPERATIONAL MANAGEMENT AND FLEXIBILITY	30	
A. EXPERIMENTAL FLEXIBILITY	20	
- SUPPORT HIGH LEVEL LANGUAGES		9
- EASY ADDITION OF NEW PROGRAMS		8
- TRANSFER PROGRAMS/DATA TO/ FROM REMOTE LOCATIONS		2
- EASY CREATION OF NEW DATA SETS		9
- REMOTE SOURCE CODE EDITING		8
- REMOTE JOB INITIATION		8
- SUPPORT MULTI USERS SIMULTANEOUSLY		8
B. TERMINAL CONTROL PROGRAM		

FIGURE 8.—Example of selection criteria rating.

CRITERIA CATEGORY	WEIGHT, PERCENT
1. COST (10-YR LIFE CYCLE)	50
2. INTERACTIVE SUPPORT CAPABILITIES	20
• OVERALL SYSTEM ARCHITECTURE	40
• EXPANDABILITY	20
• AVAILABILITY/MAINTAINABILITY	40
3. GENERAL SUPPORT CAPABILITIES	30
• OPERATIONAL MANAGEMENT AND FLEXIBILITY	50
• EXPANDABILITY	30
• TRANSPORTABILITY OF TECHNOLOGY	20

FIGURE 9.—Example of selection criteria categories for an Earth resources data system.

Configuration elimination: There are certain functions and support items in every data processing system which are mandatory. Any system not containing these items should be considered nonresponsive. The objective of this action is to eliminate from further consideration any configuration which does not contain the mandatory items. The evaluator may, of course, allow the vendor an opportunity to supply the missing capabilities.

Rating value assignment: The evaluator must now judge the configuration's merit for each selection criterion. The rating is done on a scale of 1 to 10, where 10 is the best or highest rating. Three methods of arriving at the value are described in appendix B.

1. **Direct method**—To be used if a rating can be easily assigned with no ambiguity. These ratings are generally, but not necessarily, linear.

2. **Sample variance method**—To be used if ratings cannot be directly assigned to the values being rated. Basically, the method presumes that all points being rated are valid points but probably not extreme points. This presumption does not preclude the assignment of a 1 or a 10 rating to a point, but it makes it more difficult.

3. **Not-quantified method**—To be used if the value being rated is subjective rather than objective; that is, a direct numerical value cannot be assigned. Typically, no more than three ratings are used.

Value	Rating
Above average	5
Average	5
Below average	2

An example of criteria weighting and rating is contained in figure 8.

Weighted rating calculation.—The final activity in evaluating the system is to calculate weighted rating values for the selection criteria. This activity uses the criteria weights and rating values determined previously. The output is a weighted rating for each criterion and each hierarchical level.

At this point, the evaluation is a matter of simple multiplication and addition. Starting at the lowest (subelement) level in the hierarchy, the weighted rating is calculated for each criterion by multiplying the weight (converted from a percentage to a decimal) by the rating. The results are then summed to the next highest level, which is in turn multiplied by that weighting factor and accumulated into the next highest level. The process is repeated until the top of the hierarchy is reached and a single system rating is obtained. An example demonstrating this process is shown in figure 10.

SUMMARY

The complex world of computer systems selection will not suddenly become simple. Technological advances will accelerate the trend to distribute functions among nodes of a widely scattered network. As communication devices become more sophisticated, reliable, and generally available, the intensity of arguments for centralized versus distributed solutions to a problem will begin to fade.

These changes in environment will only serve to

CRITERIA	WEIGHT	RATING	WEIGHTED RATING
I. CATEGORY I	70	7.5	5.2
A. ELEMENT IA	40	5.3	2.1
1. SUBELEMENT A	30	8	2.4
2. SUBELEMENT B	40	2	.8
3. SUBELEMENT C	30	7	2.1
B. ELEMENT IB	60	9.0	5.4
1. SUBELEMENT D	50	10	5.0
2. SUBELEMENT E	50	8	4.0
II. CATEGORY II	30	5.3	1.6
A. ELEMENT IIA	25	3.5	.9
1. SUBELEMENT F	80	4	3.2
2. SUBELEMENT G	10	2	.2
3. SUBELEMENT H	10	1	.1
B. ELEMENT IIB	50	5.8	2.9
1. SUBELEMENT I	40	9	3.6
2. SUBELEMENT J	40	3	1.2
3. SUBELEMENT L	20	5	1.0
C. ELEMENT IIC	25	5.9	1.5
1. SUBELEMENT M	90	6	5.4
2. SUBELEMENT N	10	5	.5
TOTAL SYSTEM	100	6.8	6.8

FIGURE 10.—Example of weighted rating calculation.

complicate the problem of the decisionmaker who is evaluating a set of proposed solutions to his computer system problem. Since the number of available options is growing and the relative importance of each of them to the user involved is different, the most objective approach possible, to the selection process, is demanded.

The procedure described is not trivial in that each step requires significant time and effort to accomplish. The process provides a focus on the elements of the selection process which are required to make an objective evaluation; in particular, the system loading analysis provides a pass/fail evaluation of candidate configurations which will eliminate exposure to inadequate system solutions and their attendant loss of flexibility and cost overruns. It also highlights underuse which might add unwarranted cost. The evaluation rating requires that project-level decisionmakers be identified and involved in the establishment of the weighting algorithms as they apply to the user community which the computer

system is being selected to support. Parochial considerations must be identifiable and capable of being evaluated. Ill-defined criteria will be eliminated from consideration.

This procedure ensures an equitable evaluation of technically adequate candidate configurations for a computer system application. It lends itself to use by buyers for evaluation and vendors for system determination. The system loading model and a model which supports sensitivity analysis of evaluation criteria weighting are relatively simple computer implementations in a programming language.

The need for a formalized computer selection process based on quantifiable criteria will grow as the number of alternative solutions continues to grow. The approach presented addresses this need. It does not eliminate all judgment from the procedure and therefore is still subject to some discussion. It does force the discussion away from emotional arguments into an orderly set of decisions which provides a specific solution to the problem being addressed.

C-5

ACKNOWLEDGMENTS

The authors wish to express their sincere appreciation to others who assisted in the development of this paper. C. C. Smith and L. E. Westberry defined, performed, and documented the model example described in appendix C. Also, D. L. Truitt developed the Evaluation Rating Model techniques and concepts described in appendix B. Appreciation is also expressed to J. C. Lyon and D. Hay of the NASA Earth Observations Division for their guidance, support, and constructive criticism throughout the conception and writing of this paper.

BIBLIOGRAPHY

- Adkins, A.W.; Ahr, H. A.; Emry, C. A.; and Sanders, R. L.: POCC Configuration Study. IBM Shuttle Payloads Data Processing. IRAD Final Report, 1977.
- Cypser, R. J.: Communications Architecture for Distributed Systems. Addison-Wesley Publishing Company, Inc., 1978.
- Hays, M. L.; O'Connor, M. F.; and Peterson, C. R.: An Application of Multi Attribute Utility Theory: Design to Cost Evaluation of the U.S. Navy's Electronic Warfare System. DDI Tech. Rep. DT/TR 75-3, 1975.
- Kemeny, J. G.; and Snell, J. L.: Finite Markov Chains. D. Van Nostrand Company, Inc., 1960.

Appendix A

Computer Sizing Evaluation Process

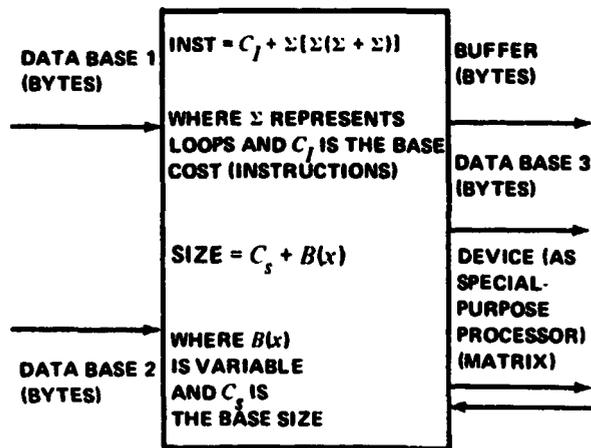
This appendix provides a 36-step instruction set to the computer sizing evaluation process. It includes sample worksheets and the definitions required to apply it to a specific problem. This process provides the technical adequacy evaluation of a candidate configuration described as Step I of the computer system selection process.

NOTE: Step 1 should include application use of operating system services, access methods, and data managers. Any operating systems or control program overhead not specifically invoked by an application function can be accounted for in a later step.

PART I—THE PROCESS

A. Define Computer Functions and Interfaces

- | <i>Step</i> | <i>Description</i> |
|-------------|--|
| 1 | Describe coded functions in terms of data input (data base, bytes), instructions executed (loop sizes, loop control variables, base cost), data output (buffers, data bases, bytes), and size (bytes, variables controlling size). |
| 2 | Represent with hierarchical input/processing/output diagram and identify all variables affecting computer resources. |



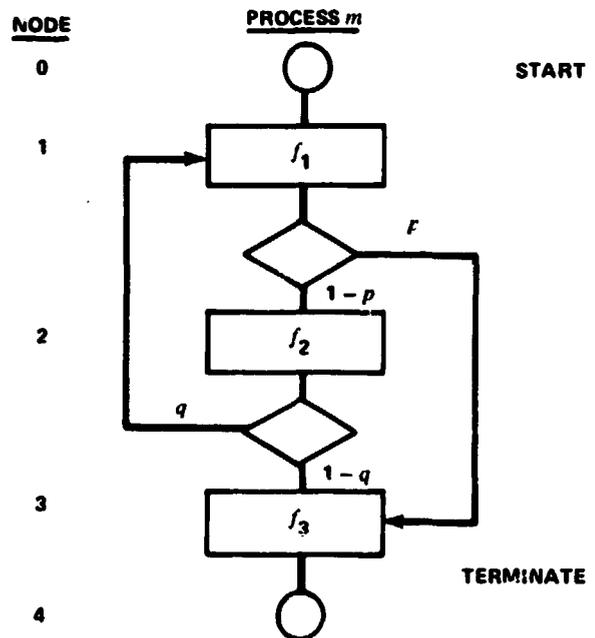
- 3 Repeat steps 1 and 2 for all functions to be included in the following processes.

B. Define Processes and Process Resource Usage

- | <i>Step</i> | <i>Description</i> |
|-------------|--|
| 4 | Establish flow-chart functional relationships for a process (<i>m</i>) (function = nodes in flow). |

NOTE: Refer to Part III for assistance in calculating steps 5 and 6.

- 5 Add probability decisions for the desired optional paths based on interactive analyst decisions.



where node n represents function n , node $n + 1$ terminates the process m , and p and q are probabilities of branching.

- 6 Define a transition matrix using the probabilities of moving from any present node (i) to any other node (j) on the very next step, and find the utilization of all nodes using the theory of Markov chains (example in Part III) or another suitable method.
- 7 Repeat the steps in category B for all processes to be defined in category C workloads.

C. Define Workload for Computer System

<i>Step</i>	<i>Description</i>
8	Identify computer users (including batch) that would simultaneously be active on the computer system.
9	Assign process descriptions (from Part I-B) that might best describe each user's activity while active on the computer system. This step assumes that a reasonable test for a computer configuration is the successful execution of several benchmark workloads.
10	Specify workload parameters that dictate how each process within a workload will exercise the functions within the processes; for example, segments of data processed, multispectral bands per segment.
11	Repeat the preceding steps for all processes within a workload and select (specify) response times for each process.
12	Repeat the preceding steps for all benchmark workloads that are judged reasonable tests of computer configuration suitability.

D. Define Candidate Configurations

<i>Step</i>	<i>Description</i>
13	By inspection of resources used by procedure and by judging the total workload from multiple users (includ-

ing batch), structure a configuration that appears to offer the capacity required for the benchmark workloads (consider bus loading for distributed systems).

- 14 Describe each hardware component and data path in terms of the parameters used to represent the functions (i.e., bytes transferred per second, instructions executed per second, etc.).
- 15 Assign a maximum utilization threshold to each hardware component such that there will be spare capacity to handle estimating errors and peak loading conditions. (Application + system overhead + estimating contingency + spare capacity for peak loading = 100-percent utilization.)
- 16 Assign each function within each process to appropriate hardware components (central processing unit (CPU), direct access storage device (DASD), etc.) and assign data flow to/from all functions to a data path. (Use the work from steps in categories A and B.)
- 17 Repeat this process for all reasonable configuration alternatives.

E. Calculate System Loading

<i>Step</i>	<i>Description</i>
18	Select a configuration (from steps in category D) and a specific benchmark (from steps in category C). For example, select a centralized architecture and a benchmark representative of a development and test environment.
19	Compute the consumption of resources for each function in terms of bytes of data transferred plus number of input/output (I/O) accesses for data bases, CPU instructions executed, etc. Category C defines the workload parameters, whereas category A defines the effect those parameters have on utilization of resources.

NOTE: Refer to tables and discussion in Part II to aid with the following steps.

- 20 Calculate resource busy time (save in table A-I) required to support a single execution of each function n of process m within the selected workload. (Use hardware configuration characteristics from category D.)
- 21 Using judgment, measured data, or other sources, evaluate each function (individually) to determine stand-alone elapsed execution time. This step involves consideration of simultaneous execution of multiple resources that may be used by the function. (Use table A-I as an aid.)
- 22 Calculate the use of computer resources by each function n within process m by multiplying the individual resource use times by the function node utilization factors calculated in step 6, category B. Update table A-I.
- 23 Sum the use of individual computer resources by each function to obtain resources used by process m (table A-II)
- 24 Determine stand-alone process elapsed time by summing the serial elapsed time component from each function.
- 25 Calculate the average utilization of computer resources for the process m executing in a stand-alone environment. Use table A-II as an aid.
- 26 Repeat steps 19 through 25 for each process within the selected workload.
- 27 Calculate the instantaneous use (average utilization) of each computer resource in support of the selected workload. Use table A-III as an aid.

TABLE A-I.—Function Resource Utilization and Elapsed Time^a

[Function n of process m]

Resource busy time (diagonal)	Stand-alone single function execution					Serial resource elapsed time	Contribution to process		
	R_1	R_2	R_3	...	R_r		Node utilization (step 6)	Adjusted busy time	Adjusted function elapsed time
R_1	T_1					T_1	U_{nm}	T_1'	
R_2		T_2	$t_{2,3}$			T_2	U_{nm}	T_2'	
R_3	$t_{3,1}$		T_3			$T_3 - t_{3,1} - t_{2,3}$	U_{nm}	T_3'	
⋮						⋮	⋮	⋮	
⋮						⋮	⋮	⋮	
R_r					T_r	T_r	U_{nm}	T_r'	
	Average function elapsed time					Σ_{fn}	U_{nm}		

^aWhere R_r = a hardware resource r used by function n

T_r = the time R_r is in use for a single execution of fn

t_{ij} = units of time resource i time T_i overlapped execution with another resource j or j (where j was initiated prior to i)

U_{nm} = utilization of function n by process m (step 6)

Σ_{fn} = stand-alone elapsed time for a single execution of fn

T_r' = resource busy time for R_r for fn within process m

$F_{nm} = U_{nm} \Sigma_{fn}$ = elapsed time component of fn for process m

TABLE A-II.—Process Resource Usage and Elapsed Time^a

[Resource utilization for process m]

Process m functions	Adjusted resource busy time (table A-I)					Adjusted function elapsed time
	R ₁	R ₂	R ₃	...	R _r	
f ₁	T ₁₁ '	T ₁₂ '	T ₁₃ '	...	T _{1r} '	F _{1m}
f ₂	T ₂₁ '	T ₂₂ '	T ₂₃ '	...	T _{2r} '	F _{2m}
f ₃	T ₃₁ '	T ₃₂ '	T ₃₃ '	...	T _{3r} '	F _{3m}
⋮	⋮	⋮	⋮	...	⋮	⋮
f _n	T _{n1} '	T _{n2} '	T _{n3} '	...	T _{nr} '	F _{nm}
Total	$\sum_1^n T_{1i}'$	$\sum_1^n T_{2i}'$	$\sum_1^n T_{3i}'$...	$\sum_1^n T_{ri}'$	ΣF_m
Utilization	P _{m1}	P _{m2}	P _{m3}	...	P _{mr}	

^aWhere R_r = a resource in process m
 f_n = a function in process m
 T_{nr}' = R_r busy time for function n within process m
 $\sum T_{ri}'$ = total R_r busy time for process m
 P_{mr} = $\sum T_{ri}' / \Sigma F_m$ = utilization of R_r, or the probability that R_r is busy during the execution of process m
 ΣF_m = stand-alone elapsed execution time for process m

TABLE A-III.—System Resource Usage^a

Processes in the workload	Resources used by processes				
	R ₁	R ₂	R ₃	...	R _r
Process 1	P ₁₁	P ₁₂	P ₁₃	...	P _{1r}
Process 2	P ₂₁	P ₂₂	P ₂₃	...	P _{2r}
Process 3	P ₃₁	P ₃₂	P ₃₃	...	P _{3r}
Process m	P _{m1}	P _{m2}	P _{m3}	...	P _{mr}
Total	$\sum_1^m P_{1i}$	$\sum_1^m P_{2i}$	$\sum_1^m P_{3i}$...	$\sum_1^m P_{ri}$
System overhead	O ₁	O ₂	O ₃	...	O _r
Threshold value	V ₁	V ₂	V ₃	...	V _r
Stretchout factor	S ₁	S ₂	S ₃	...	S _r

^aWhere process m = processes in the selected workload
 R_r = all computer resources used by m processes
 P_{mr} = the utilization or probability of use of any resource r by any process m
 ΣP_{ri} = the instantaneous utilization of computer resources by the workload
 V_r = the "threshold value" or the maximum permitted instantaneous utilization of a resource
 S_r = the stretchout factor for resource r (used to approximate process response time)
 O_r = the fraction of utilization of resource r to be expected (an estimate) for the selected workload being evaluated

- 28 Estimate (by means of experience with similar systems or a calculation) the system overhead use of computer resources to support such system functions as workload and configuration management. Add this information plus the process response criteria (step 11, category C) to table A-III.
- 29 Compare resource utilizations with the maximum permitted utilization of resources by application processes. (Maximum utilization for all processes equals threshold utilization less system overhead utilization, or

$$\left[\sum_1^m P_r \leq (V_r - O_r) \right]$$

for any resource r in the workload of m processes.)

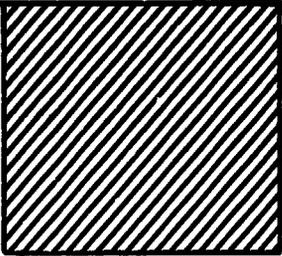
- 30 Calculate a "stretchout factor" (S_r) for all overused resources in order to extend the process elapsed times of all processes using the resource. See calculation in table A-III.
- 31 If any resources are overused, approximate the average elapsed execution time for all processes in the selected workload by recalculating (extending) each resource time using a stretchout factor (S_r) for the involved resources (r).

$$S_r = \frac{\sum_1^m P_r}{V_r - O_r} \quad \text{if } \sum_1^m P_r + O_r > V_r$$

$$S_r = 1 \quad \text{if } \sum_1^m P_r + O_r \leq V_r$$

TABLE A-IV.—Function Elapsed Time for Selected Workload^a

{Function n for process m}

Resource busy time (diagonal)	In a workload single function execution					Serial resource elapsed time	Contributor: to process	
	R ₁	R ₂	R ₃	...	R _r		Node utilization (step 6)	Recalculated function time
R ₁	T ₁					T ₁ (S ₁)		U _{nm} U _{nm} Σ f _n ' = F _{nm} '
R ₂		T ₂	t _{2,3}			T ₂ (S ₂)		
R ₃		t _{3,1}	T ₃			T ₃ S ₃ - t _{3,1} (S ₁) - t _{2,3} (S ₂)		
⋮						⋮		
R _r					T _r	T _r (S _r)		
Resource stretchout factor (from tables)	S ₁	S ₂	S ₃	...	S _r	Σ f _n '		

^aSee table A-I for definition of symbols. By proceeding along the diagonal, calculate the "serial resource elapsed time" by stretching out the time T_i that the resource R_i is busy for a single function execution by its "stretchout factor" S_i and subtracting any portion of that time t_{ij} which overlapped the execution of other resources (same row or same column) listed to the left and above the diagonal at point T_j (where i - 1 to n). Scale each amount of overlapped resource time t_{ij} (where j is the resource that initiated execution first) by selecting a stretchout factor S_j (where S_j = S_i if t_{ij} is in the same row, and S_j = S_j if t_{ij} is in the same column as T_j). In the example shown in this table, the unoverlapped component of T₃ = T₃(S₃) - t_{3,1}(S₁) - t_{2,3}(S₂).

Use tables A-IV and A-V as aids.

- 32 Repeat for all combinations of configurations and workloads.

NOTE: Step 31 uses a stretchout factor to force reasonable utilization of resources; i.e., equal to or less than the specified threshold value. However, some resources queuing should be expected and the average process response times will (most likely) be greater than the final value calculated (step 31). Therefore, based on reasonable resource utilization (70 percent or less) and guided by queuing theory calculations for mean waiting time in a queue,

$$T_w = \frac{\rho}{2(1-\rho)} T_s$$

where T_w is time spent waiting for resource, ρ is the resource utilization after step 31 has stretched out process

time, and T_s is a constant resource service time or average time per use. The average process time should be bounded by the time calculated in step 31 to twice (2 times) that number. Individual process times may well vary beyond those bounds.

F. Readjust Configurations

Step	Description
33	Compare utilizations to reasonableness (low utilizations across all proposed workloads suggest lower performance, less expensive hardware).
34	Compare final, recalculated average process elapsed times (table A-V) for all workloads to the expected response, specified by end-users. For process times that exceed expectations, either (1) change the proposed configuration by increasing resource performance or (2) separate multifunction use of resources to avoid contention and pro-

cess elapsed-time stretchout. (Configuration resources with the highest stretchout factors should be considered first.)

- 35 Repeat steps in categories D and E based on adjustments from steps 33 and 34.
- 36 Retain configurations that will handle the benchmark workloads satisfactorily as candidate configurations that can be further evaluated on the basis of other selection criteria, such as cost and operations flexibility.

TABLE A-V.—Recalculated Process Elapsed Time^a

[Process *m* for selected workload]

Function number by process	Processes in selected workload				
	Process 1	Process 2	Process 3	...	Process <i>m</i>
<i>f</i> ₁	<i>F</i> _{1,1}	<i>F</i> _{1,2}	<i>F</i> _{1,3}	...	<i>F</i> _{1,<i>m</i>}
<i>f</i> ₂	<i>F</i> _{2,1}	<i>F</i> _{2,2}	<i>F</i> _{2,3}	...	<i>F</i> _{2,<i>m</i>}
<i>f</i> ₃	<i>F</i> _{3,1}	<i>F</i> _{3,2}	<i>F</i> _{3,3}	...	<i>F</i> _{3,<i>m</i>}
⋮	⋮	⋮	⋮	⋮	⋮
<i>f</i> _{<i>n</i>}	<i>F</i> _{<i>n</i>,1}	<i>F</i> _{<i>n</i>,2}	<i>F</i> _{<i>n</i>,3}	...	<i>F</i> _{<i>n</i>,<i>m</i>}
Recalculated process time	$\sum_{i=1}^n F_{i,1}$	$\sum_{i=1}^n F_{i,2}$	$\sum_{i=1}^n F_{i,3}$...	$\sum_{i=1}^n F_{i,m}$
Twice the process time	$2(\sum_{i=1}^n F_{i,1})$	$2(\sum_{i=1}^n F_{i,2})$	$2(\sum_{i=1}^n F_{i,3})$...	$2(\sum_{i=1}^n F_{i,m})$
Expected response <i>E</i> _{<i>m</i>}	<i>E</i> ₁	<i>E</i> ₂	<i>E</i> ₃	...	<i>E</i> _{<i>m</i>}
Delta response ΔE_m	ΔE_1	ΔE_2	ΔE_3	...	ΔE_m

^aWhere $\sum_{i=1}^n F_{i,m}$ is the sum of all *n* function elapsed times for process *m*
 $2(\sum_{i=1}^n F_{i,m})$ is the maximum of the range in which the true average process *m* response times should fall
*E*_{*m*} is the expected response time or the required response time that should be met if the system meets end-user processing needs
 $\Delta E_m = 2\sum_{i=1}^n F_{i,m} - E_m$ is a difference between actual and expected response times for each process *m*

PART II—SAMPLE WORKSHEETS FOR PART I-E

Based on the parameters specified in the benchmark workload, the computer resources used by any function may vary. For example, the number of instructions to handle a record of data may be relatively constant but the total cost of the record-processing function will be a multiple of the number of input records. Steps 19 and 20 use workload parameters plus configuration rate (speed) characteristics to derive the common unit of time for expressing use of computer resources (i.e., resource busy time). Record calculations on diagonal of table A-1.

Step 21 considers the overlap of resource usage (if any) and uses this information to approximate the average function elapsed time if executed in a stand-alone environment (no other interfering activity in the computer). Table A-1 provides a way to express total time *T_r* by resource (CPU, I/O device, etc.) on the diagonal and the overlapped units of time (*t_{ij}*) within the row or column for the particular resource *R*. Record units of time that resource *R* overlapped execution time with previously considered resources by recording the amount of the overlap to the left of the diagonal if the resource overlapped was first to initiate or by recording the amount of the overlap above the diagonal if resource *R* began executing first. If there are *n* functions in *m* processes, there will be *nm* function resource utilization tables (table A-1).

By proceeding along the diagonal, calculate the "serial resource elapsed time" by subtracting any time *t_{ij}* that resource *r* overlapped execution with another resource (same row to left of the diagonal or same column above the diagonal). Table A-III is an example for *T₃*. Calculate all rows, record results in the column labeled "elapsed time," and sum the "elapsed time" column.

Since a specific process *m* may not use function *n* just once and only once, the function node utilization was calculated in step 6, category B. Record the utilization factor *U_{nm}* in the proper column and multiply the factor *U_{nm}* by the resource busy time (diagonal) for each row (resource).

To find an adjusted elapsed-time component for the process *m*, use the factor *U_{nm}* to multiply the stand-alone average elapsed time for function *n*.

Steps 23 to 25 require the calculation of process elapsed time and the average utilization of each

resource used by the process if the process was executing in a stand-alone environment. Table A-II aids this calculation process. If there are m processes in the workload, there will be m process resource usage tables (table A-II).

The next steps assume that given the selected workload, any mixing of resources is possible and could occur without ordering. Also, over any extended period of time, a sustained utilization of resources would occur and could be calculated for the combined processes in the workload by summing P_{nr} for all r 's given infinite resource capacity. Obviously, the data provides only utilization in a stand-alone environment. However, by summing all average resource utilizations across all processes, one can determine whether or not the workload (the combined processes) overloads the computer hardware resources. If not, the summed utilizations offer a reasonable approximation of hardware loading (excluding operating system configuration management or error management functions). If overloads do occur (i.e., application processes plus operating system usage exceeds the threshold utilization), then there will be conflicts for resources that will surely stretch out process average response time. Steps 29 to 31 provide an approximate way of evaluating resource loading and resulting impact to process response time.

Use table A-III to record resource utilization by process, to calculate system resource utilization, and to calculate a "stretchout factor" (S_r) affecting process response time for all overused resources. Next, use the "threshold value" V_r (step 30) for calculating a multiplier to adjust average process response time for all resources exceeding the threshold.

The maximum permitted utilization of a resource could be reached: first, because the application processing required the resource and, second, because of system overhead utilization O_r resulting from shared use of the resource to manage the workload and the hardware configuration (step 31). Therefore, the stretchout factor for resource r is determined as follows:

$$S_r = \frac{\sum_1^m P_r}{V_r - O_r}$$

where O_r is the overhead use of resource and r is re-

quired to support configuration and workload management.

$$S_r = \frac{\sum_1^m P_r}{V_r - O_r} \quad \text{if } \sum_1^m P_r + O_r > V_r$$

$$S_r = 1 \quad \text{if } \sum_1^m P_r + O_r \leq V_r$$

Record either 1 or the calculated value for the "stretchout factor" S_r in table A-III.

The following is a method for recalculating approximate average process response time. The recalculation process requires that both the function elapsed times (a table A-I calculation) and the process elapsed times be reevaluated. Table A-IV provides a means of structuring the recalculation of function elapsed time. Similar to table A-I, there may be nm uses of table A-IV to represent n functions of m processes. If any function of any process does *not* use an overutilized resource (i.e., $S_r = 1$ for all r 's) then table A-IV is unnecessary and the adjusted function time $U_{nm} \sum_{fn}$ may be immediately recorded in table A-V as the recalculated function elapsed time F_{nm}' .

Sum the serial resource elapsed time to obtain an elapsed time for the particular function n . This elapsed-time figure represents a single execution in the selected workload environment. Factor the result by U_{nm} or the utilization of function n for process m and record the recalculated function time F_{nm}' in table A-V. Repeat for all functions using overused resources for each process in the selected workload.

Table A-V is simply a means of recording function elapsed time components of the process elapsed time in a column. Summation of the columns provides a new approximate process elapsed time ($\sum_1^n F_m'$ summed for all n functions of process m).

By summing the recalculated function times for a process, reduction of the configuration throughput (for the workload being considered) has been forced to the extent that resource utilization is reasonable. The configuration can be expected to complete work no faster than the average process times suggest: i.e., if it takes 20 minutes to handle a segment at each

analyst station, then only three segments per hour will be processed at each analyst station. However, in performing the preceding calculations, waiting time for resources has not been considered. From queuing theory,

$$T_r = T_w + T_s$$

where T_r is the average response time for a resource, T_w is time spent in a queue waiting for the resource to be free, and T_s is the average service time the resource spends on a request.

It is obvious that $T_w = 0$ for all previous calculations, although excessive queueing should be avoided by generally expecting resource utilization to be 70 percent or less and by expecting processes within the workload to use any resource serially within the process. Nevertheless, some queueing between processes will occur. If one assumes that resources will be used in a relatively constant way (i.e., $T_s = \text{constant}$), then the following queueing theory formula applies.

$$T_w = \frac{\rho}{2(1-\rho)} T_s$$

where T_w and T_s are as defined previously and ρ is the resource utilization ($\rho = \text{threshold value or less}$). At approximately 70-percent utilization ($\rho = 0.7$) of any resource, T_w is approximately equal to T_s .

At this point, one could recalculate the "serial resource elapsed time" column of table A-IV or simply assume that, as an approximation, the average response time of any process m falls between the value already calculated in table A-V and twice that value. In table A-V, a row for two times $\sum F_m'$ has been left for each process in the workload.¹

If differential response $\Delta E_m > 0$, then there is a configuration problem for this workload. Another hardware alternative should be considered. More specifically, all resources used by process m for which $\Delta E_m > 0$ should be examined to determine whether there are overutilizations (table A-III) that can be solved by faster hardware components or by limiting usage by other processes and/or the system overhead functions.

If E_m falls between $\sum F_m'$ and twice that value, then there is a potential configuration problem. Proceed with judgment.

Given new approximate average process elapsed times, utilizations of hardware resources that were not overused could be recalculated. Repeat the final calculation for table A-II using process elapsed times from table A-V instead of \sum_m to obtain

$$P_{mr} = \frac{\sum_1^n T_r'}{\sum_1^n F_m'}$$

for all processes. Recreate table A-III and sum to find approximate utilization. By inspection, one might determine that some resources are clearly underused. A reduction in performance (if possible, considering other workloads) could mean a reduction in cost.

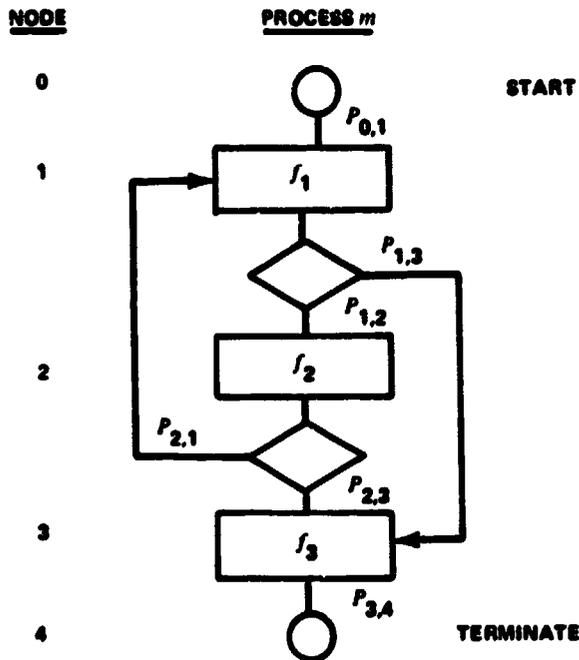
PART III—IMPLEMENTATION OF STEPS 5 AND 6 OF PART I-B

In any process m that offers the opportunity to repeat or to skip functions in that process, one faces the necessity of calculating the use (or utilization) of any specific function (node). If the function is repeated, it is reasonable to expect that computer resources are used more and the process will take longer. If the function is skipped, then the opposite is true. The following is a simple example with a method of calculating the expected (average) utilization of functions within an interactive process. These results for utilization U_{nm} for n functions of process m are used in tables A-I and A-IV.

The solution for function utilization is necessary only if a process has probabilistic branching; i.e., no branches or predetermined branching implies that each function executed is used once and $U_{nm} = 1$ for all functions (1 to n). The example uses a solution to a Markov chain; however, any other means (as simulation, a calculation of probabilities, etc.) would be appropriate. In the example, there are three func-

tion nodes in the process m plus a terminating node 4 ($n + 1$). There are two decision points at which a terminal user might alter the flow through the process. Each probability P_{ij} that the process will flow from node i to node j on the very next step must be determined by judgment or from past experience (measurements). Assignment of transition probabilities is done by assuming that the process is now at node i and deciding (without regard to and independent of past branching) the likelihood of either proceeding to the next function (node) or taking a branch.

In this example (and for all processes considered by this system sizing procedure), the probability of



beginning at the first function node f_1 is one (i.e., $P_{0,1} = 1$). Also, the probability that the process will complete and terminate is one (i.e., $P_{n,n+1} = 1$, or, for this example, $P_{3,4} = 1$). A solution can be implemented by simply building a matrix to represent the probability of moving from node i to node j for all combinations of $n + 1$ nodes and solving for the utilization of each function (1 to n).

EXAMPLE

		TO NODE			
		1	2	3	4
FROM NODE	1	0	$P_{1,2}$	$P_{1,3}$	0
	2	$P_{2,1}$	0	$P_{2,3}$	0
	3	0	0	0	1
	4	0	0	0	1

In the transition matrix, each row sums to 1:

$$\sum_{j=1}^{n+1} P_{ij} = 1$$

where $i = 1, 2, \dots, n, n + 1$. Although the process m is always begun at node 1 from a zero node outside the process and will terminate at node $N + 1$, the solution for utilization considers only the n nodes of the process (in this case, three nodes representing three functions). Therefore, in this example, matrix P is defined without either node 0 or node $n + 1$ represented.

$$P = \begin{bmatrix} 0 & P_{1,2} & P_{1,3} \\ P_{2,1} & 0 & P_{2,3} \\ 0 & 0 & 0 \end{bmatrix}$$

Using the theory of Markov chains, the utilization at each node is taken from the matrix A defined by

$$A = (I - P)^{-1}$$

where A is the resulting matrix, I is the identity matrix (an n by n matrix with all zeroes except for ones on the main diagonal), and P is a transition

matrix for process m , $(I - P)^{-1}$ is the inverse of the difference between matrices I and P .

Once the matrix A has been determined, the utilization of each function (assuming process m begins at function 1) is the value of the corresponding element in row 1 of the matrix A .

$$U_1 = a_{1,1}$$

$$U_2 = a_{1,2}$$

$$U_3 = a_{1,3}$$

where a_{ij} is the element in the j th column of the first row of A and, in general, $U_n = a_{1,n} = U_{nm}$ for process m . If matrix $(I - P)^{-1}$ does not exist (i.e., the inverse cannot be obtained), matrix P defines an indefinite loop such that the process will never terminate.

The described procedure gives the expected utilization at each node. In cases for which the maximal loading of the nodes is important, one would augment this procedure by calculating the variance of the utilization at each node. Appropriate equations for this calculation may be found in "Finite Markov Chains" by Kemeny and Snell (see bibliography).

Appendix B

Selection Criteria and Weighting Process

This appendix provides a process by which acceptable configurations can be compared. It also depicts the authors' application of the process to an Earth resources data system.

PART I—THE PROCESS

Once a set of configurations capable of performing the required workloads is found, it is necessary to establish a process which compares them. Since each configuration can perform the required functions, this choice will be made on the basis of system characteristics other than technical feasibility. This appendix presents an approach to finding an evaluation rating and the application of the approach to the Earth resources data system example.

In selecting comparison criteria, it is necessary to have well in mind the characteristics and needs of the intended user. Some criteria apply universally (e.g., cost), whereas others are applicable to only a small fraction of users (e.g., real-time support facilities). The weights assigned to the various criteria will also vary from user to user. A user with very clear-cut applications firmly in mind will place great emphasis on the capability of a potential system to support that application and extensions of it, while a user with somewhat ill-defined future plans will emphasize adaptability and flexibility.

In the case of the Earth resources data system example, the intended user is assumed to have two major uses for the system: as an analysis tool to evaluate algorithms and techniques and as a demonstration tool to apply already developed algorithms and techniques on a repetitive basis to show feasibility in a production environment. Since these algorithms cannot be known beforehand, this user also needs good incremental development tools. It is assumed that some of the analysis and possibly the demonstration work will be done interactively. Based on these user characteristics, three major criteria for evaluating the candidate systems have been selected: cost, interactive support capabilities, and general support capabilities. The weights chosen make cost equal to the sum of the others and make general support more important than interactive support, corresponding to the assumed relative utilization.

When defining the components of each criterion, the objective is a set of descriptions of functions and/or facilities that are as explicit as possible. The definitions presented here should permit an evaluator to precisely rate a system on the basis of quantitative measurements rather than subjective assessment. The following points describe the method to be used.

1. Three major areas were established and weighted.

- a. Cost—50-percent weighting

b. Interactive support capabilities—20-percent weighting

c. General support capabilities—30-percent weighting

Each of these areas is defined in detail subsequently.

2. With the exception of cost, the measurable components of the major areas were defined and also weighted. Specific items were defined that would permit an evaluator to assign ratings to individual components without having a detailed knowledge of how the measured value for that component was derived.

3. A rating system of 1 to 10 was defined, where 10 is the best rating that can be assigned.

Several methods of assigning numerical ratings can be used. In general, the method chosen for assigning a particular rating depends primarily on the numerical values that are being rated. The methods to be used are described as follows.

1. Direct method—To be used if a rating can be easily assigned to the values being rated. It should be noted that the ratings are generally linear, though not precisely.

2. Sample variance method—To be used if ratings cannot be easily assigned to the values being rated. Basically, the method presumes that all points being rated are valid points but probably not extreme points. This presumption does not preclude the assignment of a 1 or a 10 rating to a point, but it makes it more difficult. The method consists of the following steps.

a. An average value (mean) is computed.

$$P_A = \frac{\sum_{i=1}^n P_i}{n}$$

b. The average variance between the points and the mean is computed.

$$V_A = \frac{\sum_{i=1}^n (P_A - P_i)^2}{n}$$

c. Maximum and minimum values are arbitrarily assigned as

$$\text{Max} = P_A + 2 \sqrt{V_A}$$

$$\text{Min} = P_A - 2 \sqrt{V_A}$$

d. The difference between the maximum and the minimum points is uniformly divided into 10 segments, and each segment is assigned a rating. The ratings may be either increasing or decreasing (e.g., higher cost would be rated lower). An example of this method follows.

VALUE RANGE	14.5	15.2	15.9	16.4	17.0	17.5	18.2	18.9	19.4	20.0	20.8
POINTS			P_1	P_2					P_3		
RATING	10	9	8	7	6	5	4	3	2	1	

The interim calculations that lead to the rating scale are not shown. However, note that the point with the higher value receives the lowest rating. The final rating would be $P_1 = 8$, $P_2 = 7$, and $P_3 = 2$.

3. Not-quantified method—To be used if the value being rated is subjective rather than objective; that is, a direct numeric value cannot be assigned. Typically, no more than 3 values and ratings should be used, as shown in the following example.

Effort required to connect new terminal

Value	Rating
Minimal	8
Medium	5
Significant	2

PART II—COST ANALYSIS

Cost is defined as being the total estimated cost for a 10-year period, including recurring and non-recurring costs. The importance of the cost criterion is inherent in the 50-percent weighting factor.

Aside from the fact that budget limitations are always an important factor, cost is especially important in this evaluation because of the way it is used. Each configuration is sized to meet the requirements of a user workload. Consequently, cost inherently includes a number of other functions and/or facilities that could be used as evaluation criteria. An example might be the programming language facilities of different operating systems. This item is not considered outside of the cost criterion because most operating systems support the necessary language requirements. It is important to determine that all the systems to be tested are capable of performing all required functions. The cost criterion does not allow the direct trade-off of function for cost.

Because all dollars expended on a system are treated as being of equal value, no weights are assigned to components of costs. Some examples of the cost items to be included are as follows.

1. Initial purchase costs
 - a. Equipment
 - b. Installation
 - c. Custom and purchased software
 - d. Engineering
2. Maintenance costs
 - a. Software (in-house or vendor supplied)
 - b. Hardware (in-house or vendor supplied)
3. Operation cost
 - a. Personnel
 - b. Expendables (paper, film)
 - c. Electrical power

PART III—INTERACTIVE SUPPORT CAPABILITIES

The interactive support capabilities criteria measure the capability of the configuration to support interactive analysis activities. The components of these criteria are as follows.

1. Overall system architecture (40-percent weighting)—This is an evaluation of how well the system is structured to support interactive use. Specific items to be considered are the ease of use of the analyst/user interface language and how much flexibility it provides for responding to errors or interim results. The adequacy of the system with regard to physical accessibility, ease of operation, and response time for interactive use will also be evaluated.

2. Expandability (20-percent weighting)—This is an evaluation of the system capacity to support

future additional interactive use. Specific items to be considered are availability of central processing unit (CPU), memory, and input/output (I/O) resources, ease of adaptation of hardware and software to additional terminals of a new or different type, and ease of connection to remote interactive users.

3. Availability (40-percent weighting)—This is an evaluation of the probability of error conditions occurring and how well the system responds to errors. Specific items to be considered are the likelihood of a hardware error interrupting a terminal session, the ease of reconnection, and the amount of processing which must be repeated after a hardware or software error.

In assigning weights to these internal components, it was believed that current usability was of overriding importance; therefore, the factors making the system easy to use and available for use were weighted most heavily.

PART IV—GENERAL SUPPORT CAPABILITIES

The general support capabilities criteria measure the capability of the configuration to support noninteractive use. The components of these criteria are as follows.

1. Operational manageability and flexibility (50-percent weighting)—This is an evaluation of how well the configuration allows for control and management of its resources. Specific items to be considered are the existence of an operational interface to control access to specific hardware, software, or data bases; support for new program development including testing and incremental release capabilities; and the flexibility of the system to manage several changing workloads.

2. Expandability (30-percent weighting)—This is an evaluation of the system capability to support additional noninteractive use. Specific items to be considered are availability of CPU, memory, and I/O resources to support additional workloads, the ease with which additional resources could be added, and ease of connection to remote users. The capability of the system to support program development is measured as part of the operational manageability.

3. Transportability of developed technology (20-percent weighting)—This is an evaluation of how well the system can be duplicated and adapted to the needs of other users. Specific items to be considered are the proportion of the cost of the system ex-

pended for special engineering or installation support activities, the ease with which the system input/output and data base formats can be adjusted to a new user, and the availability of support for all software used within the system.

Again, in assigning weights, ease of current use

has been heavily favored. It is believed that the success of a multiuse system such as the Earth resources data system will greatly depend on its capacity to respond to a changing environment and to management of that change.

Appendix C

Sample Application of the Configuration Adequacy Model

This appendix contains an example of the computer sizing evaluation process presented in appendix A. It applies the methodology to solving a hypothetical problem that is related to the Earth Resources Interactive Processing System (ERIPS) currently in use at the NASA Johnson Space Center. The hypothetical problem that is addressed is "Given a 'batch production' job, determine the elapsed time to complete that job and leave 40 to 50 percent of the central processing unit (CPU) available for other applications." This example is divided into seven parts, and the various parts address specific ideas presented in appendix A.

PART I—FUNCTION DEFINITION

The LACIE/ERIPS batch production user environment was specified as the target for investigation. It is necessary to define all processing functions associated with this environment and the resource usage variables for each. Steps 1 to 3 of the process covered in appendix A are depicted here.

A flow chart of the baseline environment provides insight and serves as an aid for identifying the functional elements of the system. This flow chart is developed as though the ERIPS is a centralized system. Even though it is known that the special-purpose processor (SPP) is part of the configuration, the concern here is only with identifying the functions which must be performed and the cost associated with those functions as they relate to ensuring that 40 to 50 percent of the host CPU is available to other jobs. This objective can be accomplished by treating the SPP simply as an input/output (I/O) device to which some functions write and others read. Unique

functions are identified with dotted lines in figure C-1.

Hierarchical input/processing/output (HIPO) representations of functions and their cost in resource variables were developed. Figure C-2 represents this activity.

PART II—PROCESS DEFINITION

Steps 4 to 7 of the computer sizing evaluation process are depicted in this section. A batch process is defined for a representative production user, and the probability of that user executing functions in a particular sequence is taken into consideration; then, resource utilization is computed on the basis of the probability (fig. C-3).

PART III—WORKLOAD DEFINITION AND SELECTED BASELINE

Steps 8 to 12 of the computer sizing evaluation process are depicted in this section. The workload defined is a batch production workload, which consisted of the processing requested on a site. In table C-1, the parameters to be input to each function within the process are specified.

PART IV—CONFIGURATION

Steps 13 to 17 are depicted in this section. Since this example involves the answering of a question about the existing configuration, it is the only candidate configuration. This system, with the appropriate

hardware capabilities, is shown in the configuration block diagram in figure C-4. All functions are assigned to the host CPU and the SPP is simply treated as an I/O device for the reasons discussed earlier.

PART V—INITIAL SYSTEM LOADING CALCULATIONS

Steps 18 to 31 are represented in this section. The system loading chart presented in table C-II is only for the clustering and products function. It represents the resource requirements of this function for a single pass through the transition matrix for the process. Table C-III represents the elapsed time for a single pass through the matrix for all functions contained in the process. This same chart must be developed for each function contained in the process.

Resource busy times and elapsed execution times for each process function are summarized in table C-III. Resource utilizations were computed as described in steps 22 to 25. The results of these computations show that the expected CPU utilization in a stand-alone environment will be 72 percent, which exceeds the 50- to 60-percent guideline (table C-IV). Thus, it is necessary to proceed with steps 27 to 31,

which involve the stretchout computations. Subsequent calculations use 50 percent as the target CPU utilization.

PART VI—STRETCHOUT

The CPU stretchout factor was applied to each function in the process as is shown for the clustering process and products function, and a new table was developed using the stretched out adjusted function elapsed times (table C-V).

PART VII—SUMMARY

Using the data from the stretchout computations contained in Part VI, a new table was developed which details the results to be expected from the stretchout (table C-VI). Upon completion of the first stretchout computation, it appears that the elapsed time to complete the batch production process and leave 40 to 50 percent of the CPU available for other jobs has increased from 476.5 seconds to 839.12 seconds (table C-VII). This increase in elapsed time results in a CPU utilization of 56 percent, which is within the 50- to 60-percent target.

TABLE C-I.—Benchmark Case Study

<i>Function</i>		<i>Process control parameter description</i>
<i>F_{B2}</i>	Image merge	3 acquisition 12 channel
<i>F_{B3}</i>	Signature extension	X ^a
<i>F_{B3}</i>	Field retrieve	1 field
<i>F_{B4}</i>	Field report (both)	Both
<i>F_{B6}</i>	Clustering	All channel, RSEG (DO, DU)
<i>F_{B6}</i>	Select Sun angle	RSEG
<i>F_{B6}</i>	Iterative clustering	L1, dots as starting vectors
<i>F_{B6}</i>	Cluster report	X
<i>F_{B6}</i>	Store statistics	X
<i>F_{B6}</i>	Detailed report	X
<i>F_{B6}</i>	Distance table	X
<i>F_{B6}</i>	Cluster map	Conditional
<i>F_{B6}</i>	Cluster map	Image tape
<i>F_{B6}</i>	Cluster dot report	X
<i>F_{B6}</i>	Green number report	Both
<i>F_{B10}</i>	Mean/standard report	Both
<i>F_{B11}</i>	Feature selection	Separability acquisition
<i>F_{B11}</i>	A prioris computation	X
<i>F_{B11}</i>	Feature selection report	X
<i>F_{B12}</i>	Classification	W/O replacement, all subclass, compute a prioris, RSEG
<i>F_{B12}</i>	Class summary report	No overrides
<i>F_{B12}</i>	Class map	X
<i>F_{B12}</i>	Bias correction	Both
<i>F_{B13}</i>	Spectral trajectory plots	X

^aFunctions marked "X" have no parameters.

TABLE C-II.—Function Resource Utilization and Elapsed-Time Table

<i>Resource (a)</i>	<i>Stand-alone single function execution</i>				<i>Serial resource elapsed time</i>	<i>Contribution to process</i>		
	<i>SC1</i>	<i>SC2</i>	<i>MCO</i>	<i>CPU</i>		<i>Node utilization</i>	<i>Adjusted busy time</i>	<i>Adjusted function elapsed time</i>
SC1	48.75				48.75	1.0	48.75	
SC2		9.02			9.02	1.0	9.02	
MCO			9.31		9.31	1.0	9.31	
CPU		1.05		99.58	98.53	1.0	98.53	
Average function elapsed time					165.61	1.0		165.61

^aSelector channel 1 = SC1, selector channel 2 = SC2, byte multiplexer channel 0 = MCO, central processor busy time = CPU.

TABLE C-III.—Resource Usage and Elapsed Time for Batch Process

Batch process functions	Adjusted resource busy time				Adjusted function elapsed time
	SC1	SC2	MCO	CPU	
FB1	0.85	1.61	0	3.25	5.71
FB2	2.63	1.41	0	6.71	10.75
FB3	0	.06	0	.15	.21
FB4	1.01	.07	0	1.94	3.02
FB5	0	0	0	0	0
FB6	48.75	9.02	9.31	98.53	165.61
FB7	0	0	0	0	0
FB8	0	0	0	0	0
FB9	0	0	0	0	0
FB10	9.50	.05	0	7.40	16.95
FB11	10.81	.13	0	158.74	169.68
FB12	29.9	2.5	6.5	64.9	103.8
FB13	.31	0	0	.47	.78
Total	103.76	14.85	15.81	343.14	476.50
Utilization	.22	.03	.03	.72	

TABLE C-IV.—System Resource Usage

Processes in the workload	Resources used by processes			
	SC1	SC2	MCO	CPU
Batch production ^a	0.22	0.03	0.03	0.72
Total	.22	.03	.03	.72
System overhead	.00	.00	.00	.15
Threshold value	.70	.70	.70	.50
Stretchout factor ^b	1	1	1	2.06

^aWith current LACIE system, interactive users and batch production are mutually exclusive.

$$^b S_{CPU} = \frac{0.72}{(0.50 - 0.15)} = 2.06$$

TABLE C-V.—Function Elapsed Time for Selected Workload

Resource busy time	Single function execution (in a workload)				Serial resource elapsed time	Contribution to process
	SC1	SC2	MCO	CPU		
SC1	48.75				48.75	Node Recalculation function time
SC2		9.02			9.02	
MCO			9.31		9.31	
CPU				98.53	98.53	
Resource stretchout factor	1.0	1.0	1.0	2.06	270.05	1.0

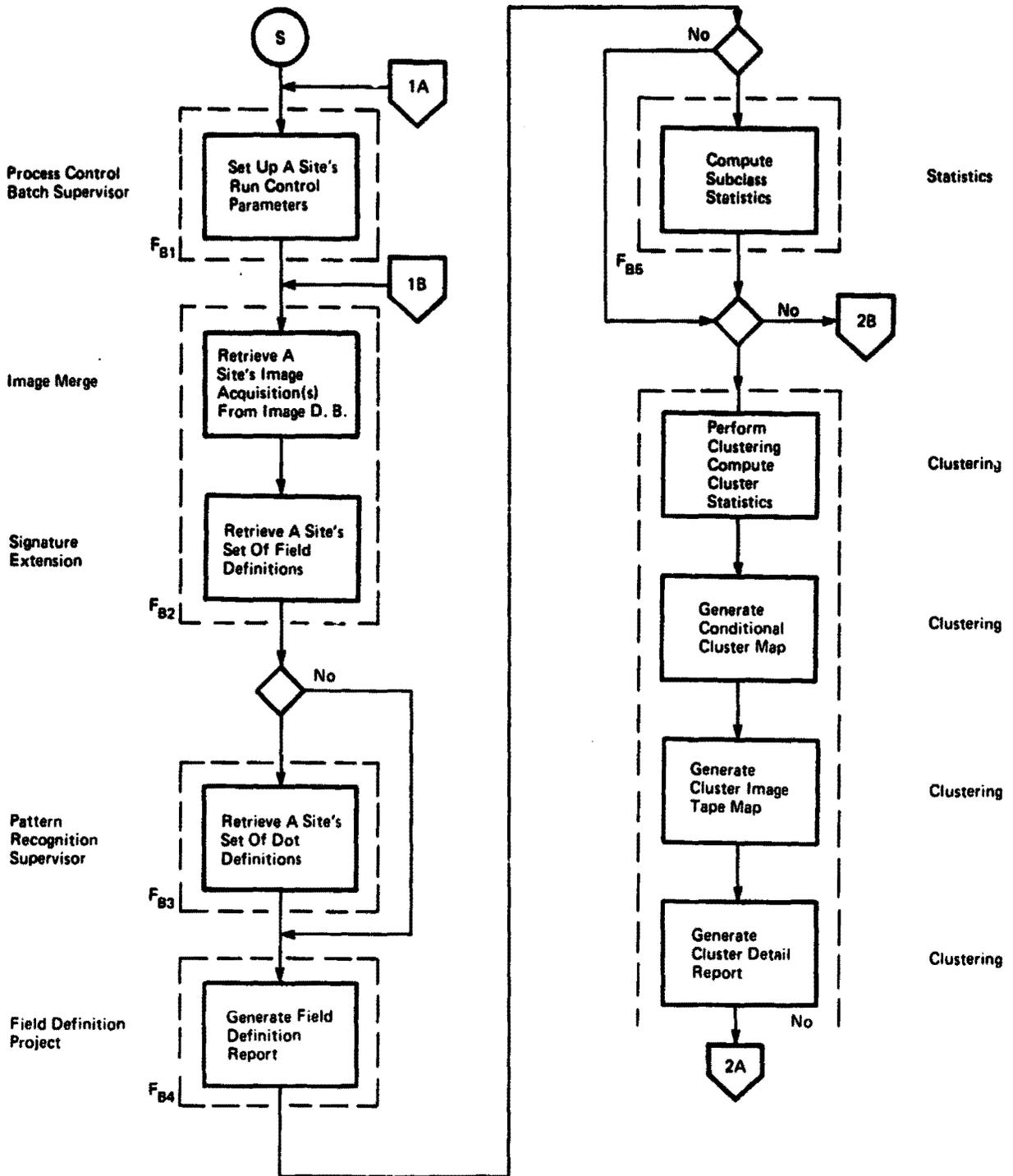
TABLE C-VI.—Modified System Resource Usage

Processes in the workload	Resources used by processes			
	SC1	SC2	MCO	CPU
Batch production	0.22	0.03	0.03	0.41
Total	.22	.03	.03	.41
System overhead	.00	.00	.00	.15
Threshold value	.70	.70	.70	.50
Stretchout factor	1	1	1	1

TABLE C-VII.—Modified Resource Usage and Elapsed Time for Batch Process

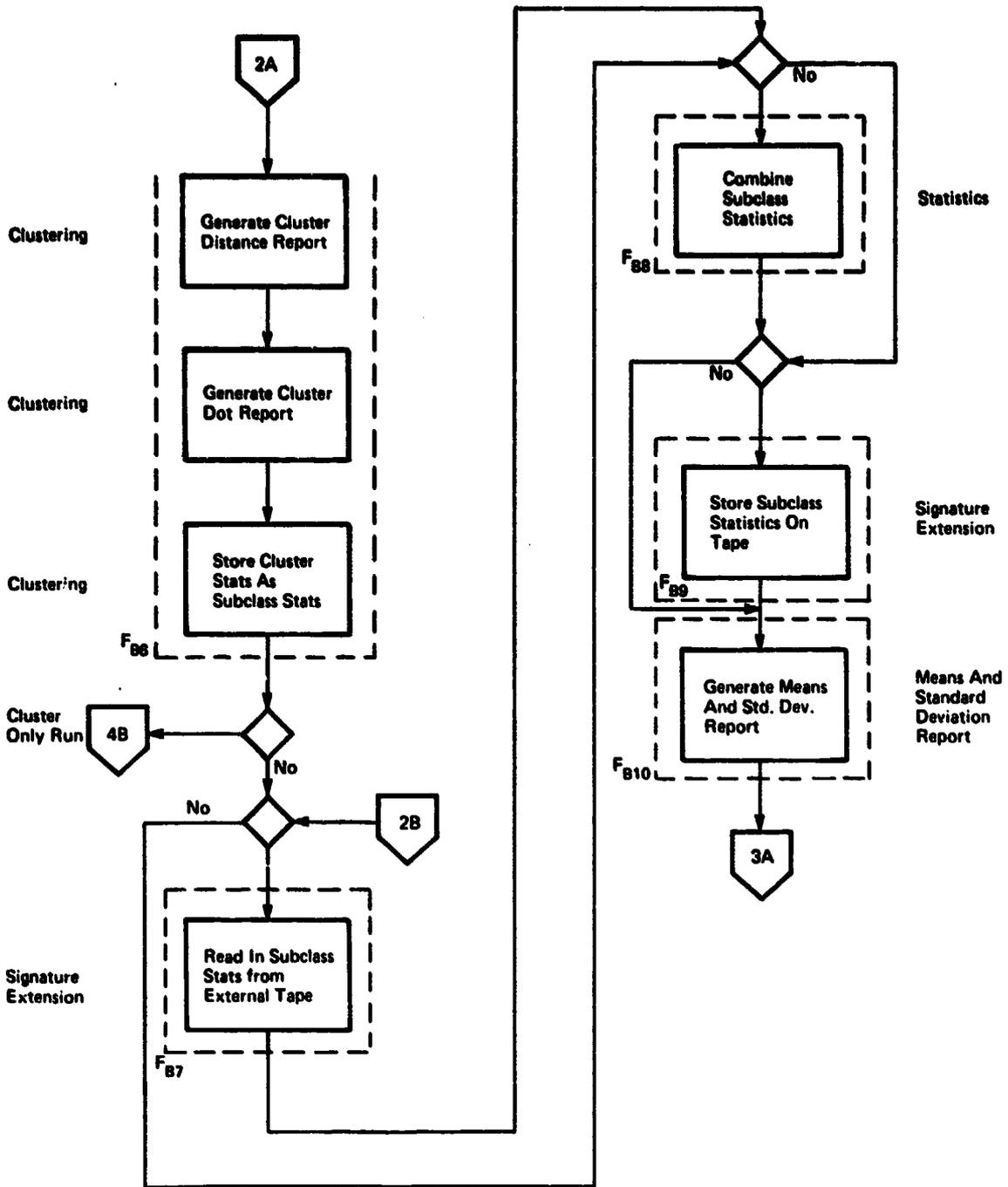
Batch process functions	Adjusted resource busy time				Adjusted function elapsed time
	SC1	SC2	MCO	CPU	
FB1	0.85	1.61	0	3.25	9.16
FB2	2.63	1.41	0	6.71	17.86
FB3	0	.06	0	.15	.37
FB4	1.01	.07	0	1.94	5.08
FB5	0	0	0	0	0
FB6	48.75	9.02	9.31	99.58	270.05
FB7	0	0	0	0	0
FB8	0	0	0	0	0
FB9	0	0	0	0	0
FB10	9.50	.05	0	7.40	24.79
FB11	10.81	.13	0	158.74	337.94
FB12	29.9	2.5	6.5	64.9	172.59
FB13	.31	0	0	.47	1.28
Total	103.76	14.85	15.81	343.14	839.12
Utilization ^a	.22	.03	.03	.41	—

^aUtilization figure excludes system overhead



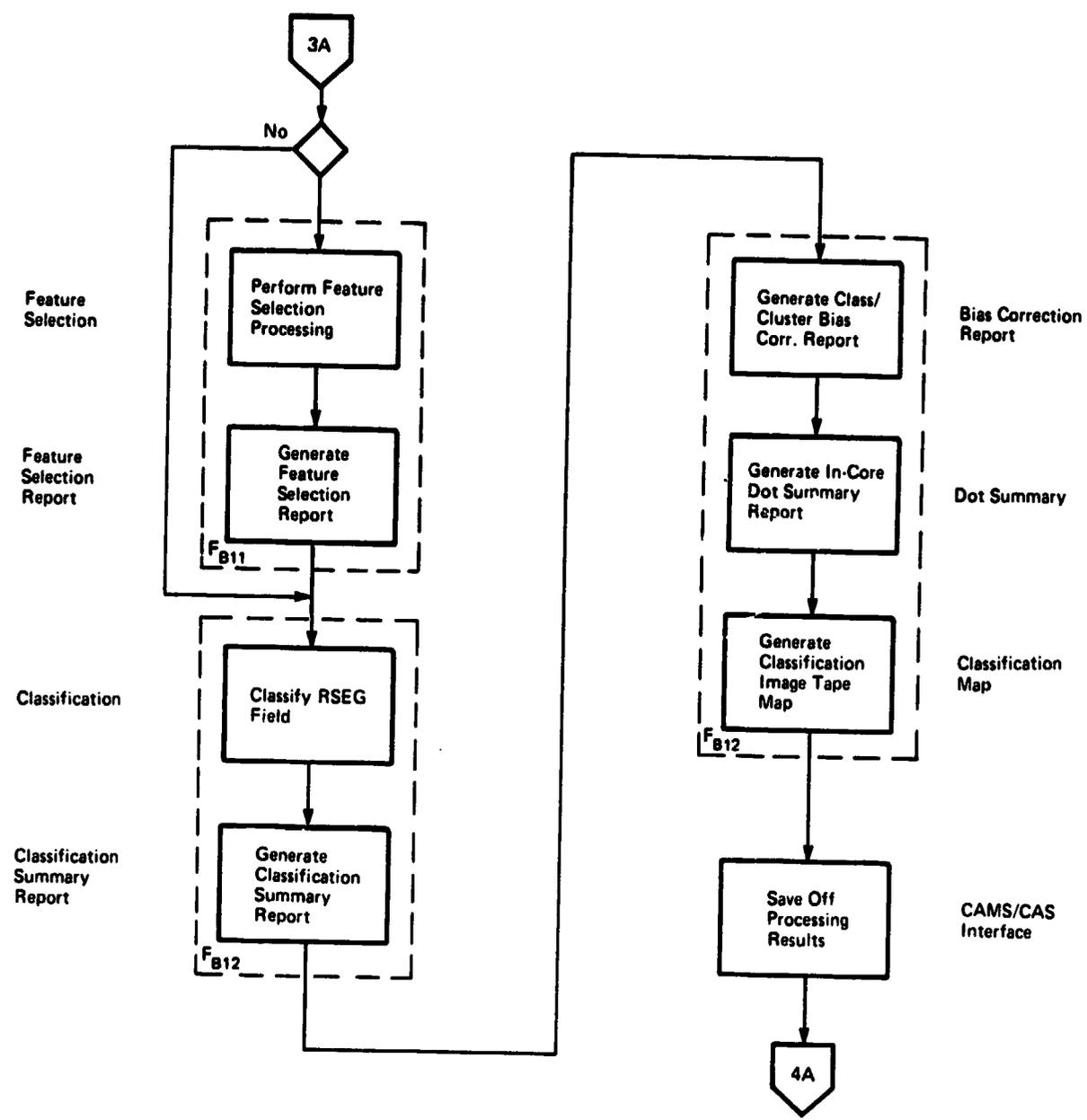
(a)

FIGURE C-1.—Baseline environment flow diagram. (a) Functions F_{B1} to F_{B6} . (b) Functions F_{B6} to F_{B10} . (c) Functions F_{B11} and F_{B12} . (d) Function F_{B13} .



(b.)

FIGURE C-1.—Continued.

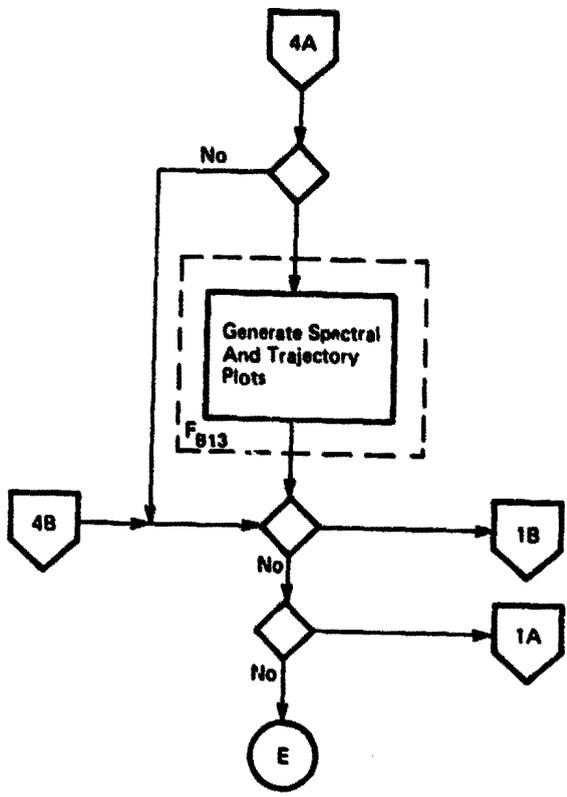


(c)

FIGURE C-1.—Continued.

Trajectory Plots

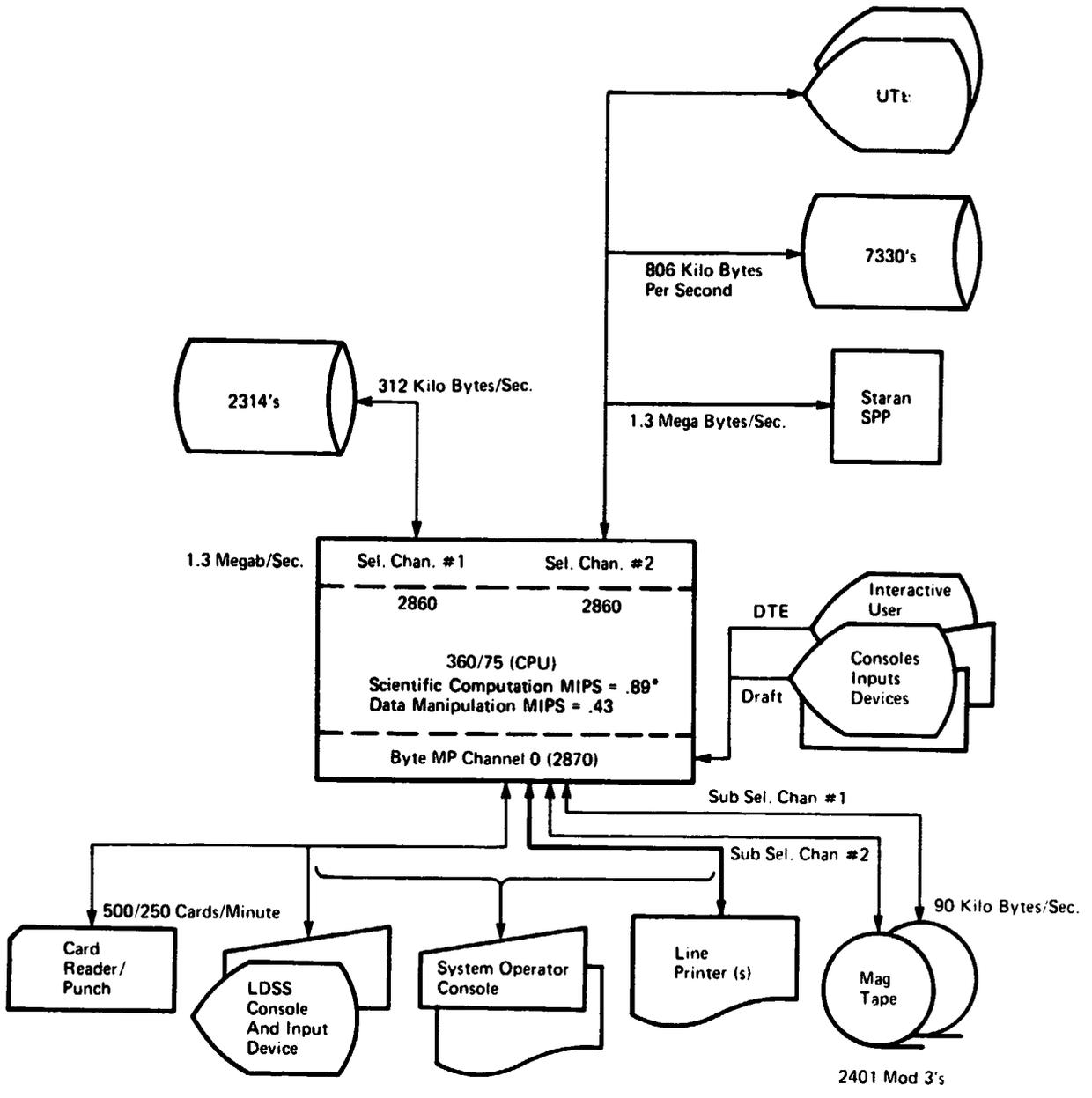
More Than One Site/Same Processing Requirements
New Processing Requirements



(d)

FIGURE C-1.—Concluded.

CS



*MIPS Rates Were Measured For LACIE Implemented Code.

FIGURE C-4.—Block diagram of current LACIE configuration.

Cost and Performance Characteristics of Data System Configurations for Processing Remotely Sensed Data

P. J. Gregor^a and J. F. Spletzer^a

INTRODUCTION

The objective of this paper is to explore some alternative approaches to constructing a large remote-sensing data system. The discussion focuses mainly on the cost and performance implications of using a collection of "small" computers versus using one large computer. Several implications of the use of large remote-sensing data systems in general are discussed.

The discussion begins with a consideration of the first step in the planning for any large data system: definition, both functional and quantitative, of the expected use of the data system. Several possible large data system architectures and some cost factors that may influence the selection of an architecture are identified. Cost and performance of architectures based on a collection of small computers versus architectures based on a single large computer are compared. Finally, some conclusions concerning data system architecture are drawn which may benefit members of the Earth resources community who are contemplating acquisition of a large data system. Because of the MITRE Corporation's recent experience with planning for a data system at the NASA Johnson Space Center (JSC), the proposed Earth Resources Data System (ERDS) will be discussed extensively as an example of a large data system.

CURRENT TRENDS

A trend in the Earth resources remote-sensing community is toward the construction of "large"

data systems. Established users of remote-sensing technology have been replacing small single-user research and development (R&D) data systems with larger, multiuser, multifaceted data systems. Three factors have combined to bring about this trend: (1) successful experience with remote-sensing technology (such as the near-operational production estimation procedures of LACIE), (2) improving price and performance characteristics of computers, and (3) availability of large volumes of data. The trend toward large systems is expected to continue as more users plan to take advantage of Landsat-D's thematic mapper. (Beginning in 1981, the thematic mapper will provide almost an order of magnitude more data per day than does the currently used multispectral scanner on Landsat-2.)

Several Earth resources data systems which are planned or under construction can be cited as examples of the trend toward large systems. The U.S. Department of Agriculture (USDA) is planning a User Advanced System that will tentatively support 15 interactive image analysis stations plus additional data processing (ref. 1). The NASA Goddard Space Flight Center (GSFC) has issued a specification for a Landsat-D Assessment System that will have considerable processing capability (ref. 2). The Canada Centre for Remote Sensing is currently completing a facility designed to support multiple users (G. Willoughby, personal communication). A final example of this trend is the experience here at JSC. In the past, LACIE has required the use of several computers in dispersed locations. JSC is currently planning a new, unified ERDS to support its continuing remote-sensing activities.¹ In all four of these cases,

^aThe MITRE Corporation, Houston, Texas.

¹ADP Acquisition Plan for an Earth Resources Data System for the Johnson Space Center. Unpublished document, JSC, Nov. 9, 1977.

the use of a single image display device connected to a small minicomputer was an early step in the development of remote-sensing activities. Now, the data processing requirements of each of these organizations considerably exceed the capacity of a single small computer.

DATA SYSTEM USAGE

A necessary first step toward the acquisition of a new data system is the determination of the extent to which the system will be used. The use of an existing data system is a good basis for making projections about the use of a new system. When such information is not available, planning becomes rather difficult. Fortunately, such information generally should exist in the Earth resources community, for large data systems characteristically represent growth from a small data system.

The use of a data system can be categorized by the type of "activities" supported and quantified in terms of what system "resources" are required. At the planner's discretion, activities can represent Earth resources functions (image preprocessing, image classification, data base management) or computer system functions (edit, compile, execute). Creation of several different categorizations of system use (system workload) can yield insight into how the work may be allocated to the resources of a large data system.

In an Earth resources data system, the basic system resources can include

1. General-purpose computational processor(s)
2. Special-purpose computational processor(s)
3. Alphanumeric (A/N) terminals
4. Image analysis stations

The set of resources required in an Earth resources data system is to some extent unique. A special-purpose processor (SPP) is required to quickly perform large numbers of parallel pixel-oriented calculations, and color displays provide the required interaction between the analyst and the imagery.

The units of utilization of each resource are generally those measured by the accounting system used on the existing or proposed data system. Examples of utilization units of a general-purpose processor (GPP) are the System Resource Unit of the Control Data Corporation's operating systems and the Standard Unit of Processing (SUP) hour of UNIVAC's Exec-8 operating system. Both units represent weighted sums of measures of a program's

use of the system's central processing unit (CPU), main memory, and input/output devices. A weighted sum approach is necessary to capture the complexity of a multiuser, multiqueue general-purpose processor.

Accounting systems for special-purpose processors have not been developed; thus, their use must be roughly measured in "connect hours" (actual hours that a particular machine is, or will be, used). Use of alphanumeric and image terminals is similarly measured in connect hours. A simple connect-hour measurement for these devices is appropriate because of their single-user, single-queue nature.

The identification of user activities and resources required can be easily represented in matrix format. Table I represents such a matrix constructed for JSC's Earth resources data processing activities in fiscal year 1977. The activities are grouped initially by the data system component used to support them. The last category in the table, "Supporting computers," represents a variety of eight IBM, UNIVAC, and DEC computers located throughout the United States.

The utilization unit for general-purpose processors was SUP hours per week. Use of each system component was originally measured in units unique to that component. Then, with results from benchmarks and comparison performance data published in the literature, these levels of use were reexpressed as SUP hours per week to present a total picture of system use. Such a conversion also enabled comparisons of the relative importance of each component of the system.

The categorization of activities in table I is based on accounting subdivisions used at JSC. Other ways of categorizing activities are possible; table II is an example. Table II reveals an interesting characteristic of JSC's Earth resources computer workload. Routine LACIE activities, such as classification and production estimation, represent only a small portion of the workload in the general-purpose and alphanumeric resource categories. (They are, however, major consumers—greater than 50 percent—of SPP and image terminal resources.) The workload is instead heavily oriented toward software development; quality assurance; and research, test, and evaluation (RT&E) activities.

The choice of categories for activities and the distribution of the workload among the activities are important considerations when planning a large data system, especially when a multicomputer configura-

TABLE I.—JSC Earth Resources Workload for Fiscal Year 1977

Activities	General-purpose processing, av SUP hr/wk	Special-purpose processing, av STARAN connect hr/wk	Alphanumeric terminal use, av connect hr/wk	Image analysis terminal use, av connect hr/wk
IBM 360/75				
LACIE:				
Segment analysis	7.7	12.2		44.8
Data base maintenance	4.3			
Quality assurance	3.0	5.0		
Software development	16.4	6.3		
Research, test, and evaluation (RT&E)	1.6			
DEC PDP 11M5				
LACIE:				
Automated Status and Tracking System	7.0		22.8	
Classification and Measurement Subsystem, Crop Assessment Subsystem, etc.	8.0		82.1	22.6
Exploratory studies ^a				22.2
Software development	.2		25.4	38.4
RT&E	9.7			3.5
System support	10.4		24.4	
Supporting computers				
RT&E	9.0		26.3	
Exploratory studies	2.1		22.0	
Data formatting	10.0		13.6	
Software development	2.9		39.2	
Yield modeling	5.0			
Quality assurance	5.3			
Aircraft program	4.0			
Flow control	7.0			
Other Laboratory for Applications of Remote Sensing	3.7			
Total	117.3	23.5	255.8	131.5

^aExploratory studies related to the National Forestry Applications Program and the Texas Applications Program.

tion is being considered. As shown in the next section, the number, size, and type of categories can form the basis for choosing the number, size, and type of computers to use in the data system.

DATA SYSTEM ARCHITECTURE

Until recently, the projected workload largely

determined the data system architecture. Small workloads were accomplished on small machines, available from the minicomputer vendors, whereas large workloads dictated large machines. As minicomputers have become more powerful (i.e., have more main memory, including cache; more sophisticated operating system; wider choice of word length—16, 24, and 32 bits; wider options for "netting" several computers), the size of the workload no

TABLE II.—Resources Required by Activity

Activity	Resource, percent, for —			
	GPP	SPP	A/N terminal	Image terminal
Software development and quality assurance	23.1	48.0	25.4	29.2
RT&E	25.6		10.1	2.7
System support	8.5		9.4	
LACIE production	29.1	52.0	41.0	51.2
Exploratory studies ^a	1.9		8.6	16.9
Other	11.8		5.5	
Total	100.0	100.0	100.0	100.0

^aExploratory studies related to the National Forestry Applications Program and the Texas Applications Program.

longer is the only factor in choosing the data system architecture.

The effect of this freedom of choice on the Earth resources community has been the selection of a variety of architectures for a variety of data system requirements. Prominent in the large machine camp are the NASA Jet Propulsion Laboratory, which has an IBM 360/65 (ref. 3), and the Purdue University Laboratory for Applications of Remote Sensing (LARS), which has an IBM 370/148 (ref. 4). A multicomputer configuration of five to six small- to medium-scale minicomputers has been proposed for the USDA's User Advanced System (ref. 1), and GSFC will have a collection of four minicomputers once the Landsat Assessment System (ref. 2) is interfaced to the existing Atmospheric and Oceanographic Information Processing System facility (ref. 5). Planning studies for ERDS have considered both multicomputer and single-computer configurations (ref. 6).

Generally, the available architectures for the previously described kinds of workloads are of two basic types: (1) a large single-machine configuration (fig. 1), or (2) a multiple small to medium machine configuration (fig. 2). The large single-machine configuration is centered around a mainframe from the IBM 360-370 series (more recently 303x series), the UNIVAC 1100 series, or some comparable product

line.² One or more SPP's and alphanumeric and image analysis terminals are attached to the mainframe as required. The size of the data system is determined by the workload. The general-purpose component's (the mainframe) size is determined by the number of SUP hours required, the number of special-purpose processors by the number of connect hours needed, and the alphanumeric and image terminals by both the number of connect hours and the number of shifts during which the user is willing to "staff" the terminals. Requirements ranging from 70 SUP hours per week and 9 terminals to 600 SUP hours per week and greater than 30 terminals can be accommodated by the single large computer architecture. A mass data storage facility (MDSF) of greater than 30 billion bytes and the data base management software to manage this mass storage are available large machine options.

The multiple-machine configuration of figure 2 is another option for accommodating large workloads. Generally, in this architecture, the workload is distributed functionally among the multiple computers; i.e., a computer and its associated facilities are dedicated to a project such as LACIE or to a function such as data acquisition, data management, or image display. The separation between functions, and thus computers, may be based on passive requirements (e.g., two functions require little or no interchange of information and thus may be separated) or active requirements (e.g., two functions may interfere with each other if they are not separated). The major active requirement for separation of functions in data systems generally is that development and production activities should not be supported by the same machines. This reasoning holds if the development activities center on modifying the operating system (and are thus likely to cause system "crashes") and if the production activities involve large on-line software systems (such as an airline reservation system) or time-oriented batch systems (such as billing or payroll). This type of reasoning would not generally apply in an Earth resources setting.

Figure 3 is an example of a functional allocation of an Earth resources workload. The JSC workload categories presented in table II have been converted to specifications for a series of four computers. In

²References to vendors are for illustrative purposes only and do not constitute an endorsement or recommendation.

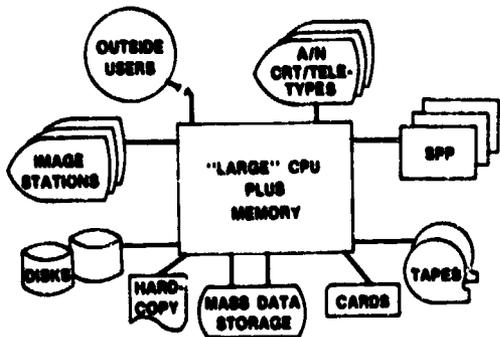


FIGURE 1.—Single large-machine architecture.

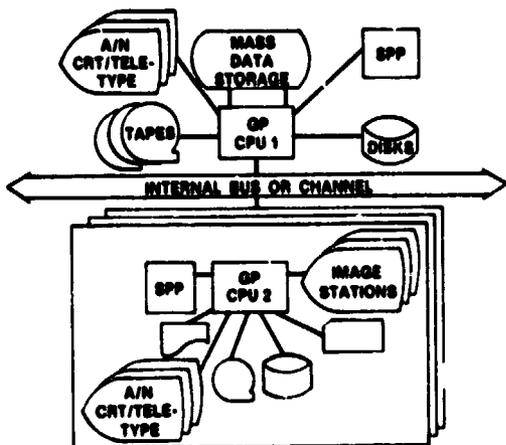


FIGURE 2.—Basic multiple-machine architecture.

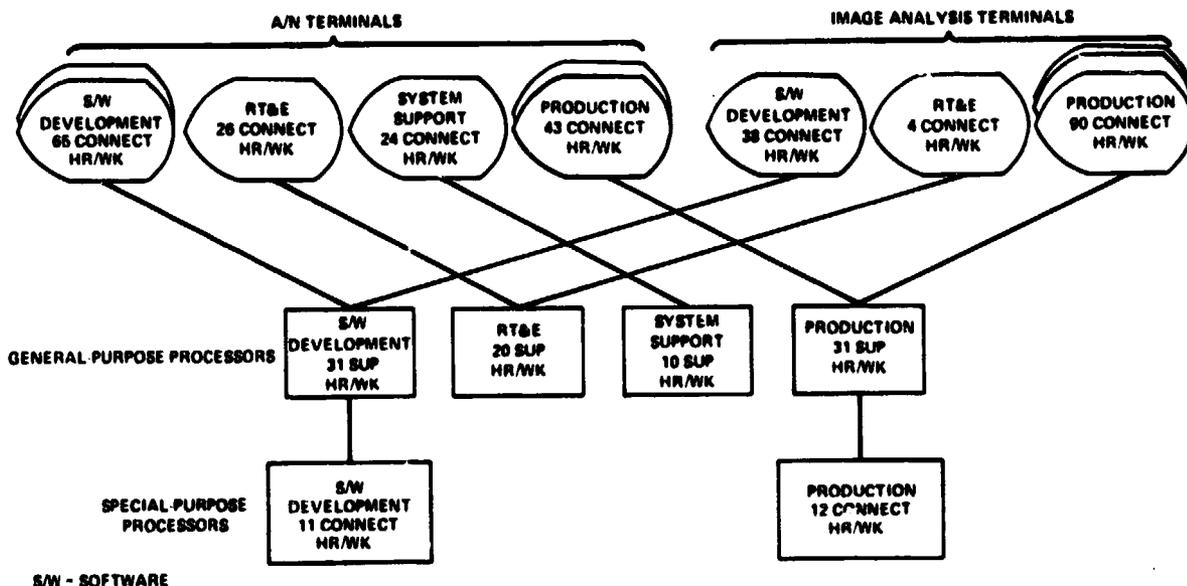


FIGURE 3.—Allocation of workload to resources.

such a system, the software development component would require

1. A general-purpose computer with a capacity of at least 27 SUP hours per week³
2. A special-purpose processor available 11 hours per week
3. An alphanumeric terminal available two shifts per week (or two terminals, each available one shift per week)
4. An image analysis terminal available one shift per week.

Although one would probably not choose to dedicate expensive resources to the software development function (since software development loads tend to vary widely over time), one might well choose to dedicate physical resources to a longer term and more predictable "production" activity. In any case, the example illustrates that workload description and architecture selection are strongly interrelated.

Figure 4 depicts a system which is a slight modification to the system in figure 2. Here, resources are pooled, intercommunication between

³The specified capacity should, in fact, be higher than the planned average week usage to accommodate peaks in activity.

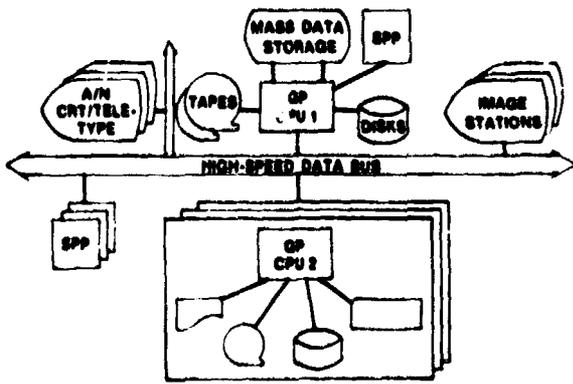


FIGURE 4.—Bus-oriented multiple-machine architecture.

processors is provided by a wide-bandwidth communications bus, and terminals are permitted to interact with any of the processors. Processors may be of the same or different size and model, depending on the sophistication of the communications protocol. The goal of this architecture is to add to a multicomputer configuration some of the capability for resource sharing among functions that is inherent in a single-computer system. The degree of sharing can become quite high if considerable investment is made in operating system development, as some prototype systems have shown (ref. 7).

DATA SYSTEM FACTORS

The one-time and recurring cost factors associated with a new large data system are as follows.

1. One-time costs
 - a. Hardware purchase and development
 - b. System software purchase and development
 - c. Applications software development and conversion
 - d. Facility modifications
 - e. Communications installation
 - f. Training
 - g. Procurement and system engineering support
 - h. System integration and test
2. Recurring costs
 - a. Hardware maintenance and software lease
 - b. Operations, system management, and support
 - c. Communications
 - d. Consumables

Frequently, only the one-time costs are considered in planning a system and choosing an architecture. It should, however, be realized that over the expected 7- to 10-year life of a data system, the recurring costs of maintenance and operations generally represent the majority of the investment in the system. Thus, to estimate the cost-effectiveness of a planned large system, *all* cost factors should be considered.

Despite the last admonition, only some of the previously mentioned cost factors are considered in this paper. The factors discussed are hardware purchase, hardware maintenance, software conversion, and operations. These factors were chosen because they represent major cost items, because they may be affected by choice of architecture, or because they are costs that are usually underestimated.

To compare the hardware costs of data systems of different architectures requires identification of specific candidate processors. Once candidate machines have been selected, the determination of hardware purchase and maintenance costs is rather straightforward. Most hardware vendors will gladly supply detailed pricing information on their products. Other sources of cost data are industry-survey publications (e.g., ref. 8) and, for systems to be procured by the Government, price schedules of the General Services Administration. Studies by the MITRE Corporation have shown that, in general, annual maintenance costs for large mainframes are approximately 5 percent of the initial purchase price, whereas maintenance of minicomputers annually costs 10 to 15 percent of the purchase price.⁴

The difficulty in determining hardware costs lies in the selection of appropriate processors to be evaluated. The main criterion of appropriateness is that the candidate machine must have the computational capacity to process the desired workload. A machine's capacity can be estimated in a number of ways, varying from "hands-on" testing (application-specific benchmarks and general-purpose benchmarks) to literature review (industry surveys and benchmark reports). Table III summarizes some of the information on the capacities of various

⁴This can be seen by comparing vendor price lists. It has also been demonstrated empirically by a survey of 20 systems throughout JSC, as disclosed by S. Berthiaume in a briefing entitled "Institutional Data Systems Division Long-Range Planning Review," presented to C. C. Kraft, Director of JSC, on December 8, 1977.

machines. The table compares various processors in terms of both the number of SUP hours they can provide in an 18-shift week and the average quantity of concurrently active terminals they can support during the day shift. The table was compiled from several sources including benchmark runs (refs. 9 to

12), published articles, and discussions with hardware vendors.

The information in table III was used to identify candidate EKDS processors in the MITRE study and can be used in other studies. For example, suppose that an Earth resources workload has been identified that requires 200 SUP hours per week, allocated as follows:

1. 50 SUP hours of routine analysis and classification (production work)
2. 30 SUP hours of software development
3. 20 SUP hours of quality assurance activities
4. 40 SUP hours of data base maintenance
5. 50 SUP hours of experimental, scientific, numerical processing

A potential configuration for such a workload would be as follows:

1. One "large" computer (with a capacity of approximately 90 SUP hours) with a mass data storage facility for the scientific processing and data base activities
2. One "small" computer (with a capacity of approximately 55 SUP hours) for the production activities
3. A second "small" computer (55 SUP hours) for the software development and quality assurance activities

Table III reveals that a system consisting of an IBM 370/148 and two SEL 32/75's should meet the user's needs. IBM and SEL vendors could then be contacted for detailed pricing information. The user could also investigate the costs of establishing communications among the three computers.

A cost item that is frequently not considered and is generally underestimated is software conversion. In moving from a small system to a large one, the user generally wishes to retain all the capabilities he had before. If the new hardware and the operating system are not compatible with the old, the applications software must be converted. A recent study for the Central Computational Facility at JSC projected that conversion of its code to a new system would cost approximately \$2.60 per line of code (ref. 13). This cost was found to be in line with other Government conversion efforts, which ranged from \$2 to \$7 per line of code. The total impact of this cost item on a particular Earth resources center planning a new system will depend on the center's previous investment in software. However, users should be aware that the actual costs of the conversions reviewed in reference 13 ranged from \$900 000 to \$5 000 000 and could well exceed hardware costs in some cases.

TABLE III.—Capacities of Several Classes of Computers

Computer class	Capacity range		Typical systems	Vendor
	Batch ^a	CI ^b		
A	15 to 30	4 to 6	360/30; 360/40; 360/50; 370/135; 370/138	IBM
			PDP 11/45	DEC
B	45 to 55	7 to 9	370/145	IBM
			PDP 11/70	DEC
			32/75	SEL
C	70 to 90	9 to 10	360/65; 360/67; 360/75; 370/148; 370/155	IBM
			AS-4	ITEL
			U1106; U1100/11	UNIVAC
D	90 to 110	13 to 14	U1108; U1100/21	UNIVAC
E	130 to 175	18 to 30	370/158; 3031	IBM
			AS-5	ITEL
			U1110 (2 x 2); U1100/41; U1100/42	UNIVAC
F	220 to 270	>30	370/158 AP; 370/168; 3032	IBM
			AS-6	ITEL
			470/V5	AMDAHL
			U1100/81	UNIVAC
G	375 to 470	>30	370/168 AP; 3033	IBM
			470/V6	AMDAHL
			U1100/82	UNIVAC
H	500 to 600	>30	470/V7	AMDAHL

^aEquivalent Batch M. P./hr/wk
^bConcurrently active average terminals

The final cost item to be considered here is operations. Large data systems require people to run them. Computer operators are needed for scheduling jobs, mounting tapes, and servicing the card reader and printer. Clerks are needed for accepting card decks from users and for distributing output. Systems programmers are required for maintaining the operating system and other software packages. Consultants should be available to aid users having problems with the data system. Training courses may be offered for new users. Systems analysts are necessary to prevent bottlenecks in the system. Earth resources systems will likely have a large data base and thus will need a data base administrator to control the structure and to maintain the integrity and security of the data base. And, of course, there is the need for supervisors and data center administration.

All these people represent a considerable cost. These costs may be "buried" by shifting some, or all, of the necessary activities to the user (e.g., forcing the user to teach himself how to use the system or to mount his own tapes on a minicomputer). They will still, however, represent an expense: for when an analyst is mounting a tape, he is not performing those activities for which he was hired. Thus, more analysts will have to be hired to accomplish the same total analysis workload.

ARCHITECTURE SELECTION CRITERIA

After identification of expected data system usage, hardware alternatives, and system costs, the next step in system development is to select a system architecture. The selection can be made using several criteria. Common selection methods rely on the following criteria or on some weighted combination of these criteria:

1. Cost-effectiveness analysis (Which system is the least expensive way to support a specific workload?)
2. Qualitative factors (Which system is easiest to use? Which offers the most services?)
3. Specific characteristics (Is the system's word length at least 24 bits?)
4. Performance factors (Which system has quickest response or turnaround times?)

Each of these sets of factors can be subdivided into a larger number of categories. For example, one author cites 50 possible performance measures of computer

service delivered through a remote terminal.⁵ Selection methods can thus get quite involved.

As system development proceeds from feasibility study through evaluation of vendor proposals, the selection process can require very detailed information on the characteristics of both the workload and the candidate configurations. Certainly, any proposed architecture should be studied analytically or by simulation model to determine necessary resource capacities and communications bottlenecks. This evaluation is especially important in the bus-oriented multimachine architecture, where bus contention may become the factor that controls system throughput. For example, in one particular case (ref. 14), several processor, memory, and input/output buses were required to minimize contention delays in a fairly small multimachine configuration.

The choice of an architecture for an Earth resources data system can also be determined by factors not directly related to either the Earth resources project to be supported or the hardware being considered. For instance, procurement regulations may determine the nature of a data system by specifying what type of approval is necessary for purchase of computer hardware. Large expenses (a large mainframe) may require high-level approval. Small expenses (a minicomputer) may require only lower level approval. Thus, the construction of a configuration of minicomputers over a span of several years may be the easiest way for an agency to obtain a large data system. (It should be noted that distribution of costs over time is not a legitimate reason for the purchase of a series of minicomputers, since the cost of a large mainframe can also be distributed over a period of several years.)

A second example of an externally imposed decision criterion would be a company-wide or organization-wide decree to purchase only a certain type or size of computer. The idea behind such a decree might be to facilitate movement of equipment from location to location as projects end and begin.

In the example presented in the following section, a cost-effectiveness analysis approach was used to evaluate the candidate configurations. External factors such as those cited were not considered, although they can easily influence the design of any Earth resources data system.

⁵M. D. Abrams and S. Treu, "A Methodology for Interactive Computer Services Measurement." In communications of the ACM, to be published.

SAMPLE COMPARISON OF CANDIDATE CONFIGURATIONS

It may be instructive to demonstrate the concepts of the preceding sections by briefly describing preliminary planning for a proposed data system for future JSC resources programs (the ERDS).

Proposed Earth resources programs at JSC in the late 1970's through the mid-1980's suggest a workload of 200 to 400 SUP hours per week and the need for 14 to 20 terminals. The range in the workload is due to program options, uncertainty about which programs will be funded, and speculation about how Landsat-D data will be used to achieve various program objectives.

To support this workload, data systems using each of the three architectures described earlier were proposed. Because of the workload range, four options of each of the three architectures were studied. In order to achieve a realistic comparison of architectures, the specific hardware selected represented, where possible, computer vendors and models that are currently in use for Earth resources applications.

In particular, the SEL 32/75 was chosen as a representative 32-bit minicomputer for Earth resources applications. (The less powerful SEL 32/55 is installed at the Earth Resources Observation System Data Center in Sioux Falls, South Dakota, as the EROS Digital Image Processing System.)

The IBM 303x series architecture (and the plug-compatible AMDAHL central processor) was chosen as representative of large mainframe Earth resources hardware. Such hardware supports Earth resources applications at several locations (e.g., at LARS).

For the multimachine architecture option requiring a communications bus, MITRE chose the off-the-shelf Network Systems Corporation bus as the intercomputer interface. This communications system offers a wide-band data path of 1.5 to 50 megabits per second (depending on the bus length), which can be used to interconnect computers and high-speed peripherals such as disk and image analysis terminals.

Each configuration was priced according to vendor price lists for purchase (one-time) and maintenance (recurring). Tables IV and V show the costs for each of the three architectures at the four possible design points: (1) 190 SUP hours per week (configuration A), (2) 245 SUP hours per week (config-

uration B), (3) 300 SUP hours per week (configuration C), and (4) 410 SUP hours per week (configuration D).

Estimated operations costs are given in table V. These costs included the use of either two or three operators per shift (depending on the complexity and size of the data system) on an 18-shift/week basis, system management and support services, and consumables.

Integration costs are not included in this example because of the somewhat arbitrary nature of the available estimates and the lack of a rigorous estimation procedure. Moreover, integration costs should be architecture-dependent (i.e., integrating a multicomputer, multivendor system should cost more than integrating a single large-computer system); however, there was no clear, precise way of quantifying this dependence. One-time integration costs will, however, be large. One rule of thumb states that the cost of integration will be about 40 percent of the cost of initial hardware and software acquisition.

The life cycle costs of each data system configuration, priced using the one-time and recurring cost factors discussed previously, are presented in table VI. The last column of this table presents a dollars-per-SUP hour figure of merit for each configuration. This figure of merit is plotted versus configuration size and architecture in figure 5. Several conclusions are immediately evident:

1. An SUP hour costs approximately \$10 less on the large single-machine architecture than on the basic multimachine architecture.
2. An SUP hour costs approximately \$20 less on the large single-machine architecture than on the bus-oriented multimachine architecture.
3. The cost per SUP hour in all three architectures decreases as the configuration grows.

Considering the cost data in the light of both the high costs of operating the dispersed data system currently in use for Earth resources work at JSC and the R&D, software-development-oriented nature of the JSC workload, it was recommended that JSC consolidate its Earth resources data processing and establish a new data system using the large mainframe architecture. These, however, are not blanket recommendations for all Earth resources data systems; a particular Earth resources organization may have legitimate reasons for selecting any of the three architectures despite the trends shown in figure 5.

TABLE IV.—Details of ERDS Hardware Purchase and Annual Maintenance Costs^a

(Fiscal year 1977)

(a) Large single computer

Item	Purchase costs, thousands of dollars				Recurring maintenance costs, thousands of dollars, for configuration—	
	Configuration A or B		Configuration C or D		A or B	C or D
	Cost	Comments	Cost	Comments		
CPU, channels, console, power	2550	470/V5 with 3 megabytes	3280	470/V6II with 4 megabytes	92	103
Paging disks, tapes, CRT's, card and paper units	523	IBM products	523	IBM products	24	24
SPP	273	2 AP-190L	410	3 AP-190L	17	26
MDSF	1234	100 × 10 ⁹ bytes; IBM 3851-B2	1234	100 × 10 ⁹ bytes; IBM 3851-B2	35	35
Image terminals	1425	5 512 by 512; 2 1024 by 1024	1425		114	114
Communication lines	—	Local terminal lines	—	Negligible cost	—	—
Totals	6005		6872		282	302

(b) Multiple computers

Item	Purchase cost, thousands of dollars, for configuration—								Recurring maintenance costs, thousands of dollars, for configuration—			
	A		B		C		D		A	B	C	D
	Cost	Comments	Cost	Comments	Cost	Comments	Cost	Comments				
CPU, channels, console, power	1376	1 370/148; 2 SEL 32/75	1558	1 370/148; 3 SEL 32/75	1740	1 370/148; 4 SEL 32/75	2104	1 370/148; 6 SEL 32/75	137	181	225	312
Peripherals	606		783		960		1314		46	65	84	121
SPP	229	1 AP-190L; 1 AP-120B	320	1 AP-190L; 2 AP-120B	320	1 AP-190L; 2 AP-120B	412	1 AP-190L; 3 AP-120B	14	19	19	24
MDSF	1234	IBM 3851-B2	1234		1234		1234		35	35	35	35
Image terminals	1425	5 512 by 512; 2 1024 by 1024	1425		1425		1425		114	114	114	114
Intercomputer bus ^b	1007	1 controller; 13 interfaces	1037	14 interfaces	1067	15 interfaces	1127	17 interfaces	132	132		132
Terminal bus ^b	102		102		102		102		7	7	7	7
Totals	5979		6459		6848		7718		485	553	616	745

^aCosts are based on vendor price lists for both purchase and maintenance, maintenance of CPU, channels, console, and power at 24-hour on-call rates (i.e., 1.4 times 8-hour on-call rate), and other maintenance on 8-hour on-call rate.

^bIncluded only for bus-oriented multicomputer configuration.

TABLE V.—Estimated Operations Cost

Operations category	Cost per MYE	No. MYE required	Total cost/year
Direct operations (operators)	\$25 000	^a 8 or ^b 12	\$200 000 or \$300 000
System management (supervisors, quality control, performance analysis, resource allocation, user services)	40 000	11	\$440 000
Systems support			
System programmers	45 000	5	225 000
System software lease			75 000
Consumables			195 000

^a8 man-year equivalents (MYE) required for configurations A and B
^b12 MYE required for configurations C and D

TABLE VI.—Life Cycle Costs

(Fiscal year 1978)

System	Config-uration	Cost, thousands of dollars					10-yr processing, \$/CPU hr	Figure of merit, \$/CPU hr
		Purchase	10-yr maintenance	10-yr direct operations	10-yr other ^a	10-yr total		
Bus-oriented multimachine	A	5979	4850	^b 2000	9350	22 179	95 × 10 ³	233.46
	B	6459	5530	^b 2000	9350	23 339	122.5	190.52
	C	6848	6160	^c 3000	9350	25 358	150	169.05
	D	7718	7450	^c 3000	9350	27 518	205	134.23
Single-machine	AB	6005	2820	^b 2000	9350	20 175	122.5	164.69
	CD	6872	3020	^c 3000	9350	22 242	205	108.50
	~D	7500	3400	^c 3000	9350	23 250	300	77.50
Basic multimachine	A	4870	3460	^b 2000	9350	19 680	95	207.16
	B	5320	3930	^b 2000	9350	20 600	122.5	168.16
	C	5679	4770	^c 3000	9350	22 799	150	151.99
	D	6489	6060	^c 3000	9350	24 899	205	121.46

^aIncludes system management (11 MYE at \$40 000/MYE = \$440 000), system support (5 MYE at \$45 000/MYE = \$225 000), system software lease (\$75 000), and consumables (\$195 000) = \$440 000 + \$225 000 + \$75 000 + \$195 000 = \$935 000

^b8 MYE
^c12 MYE

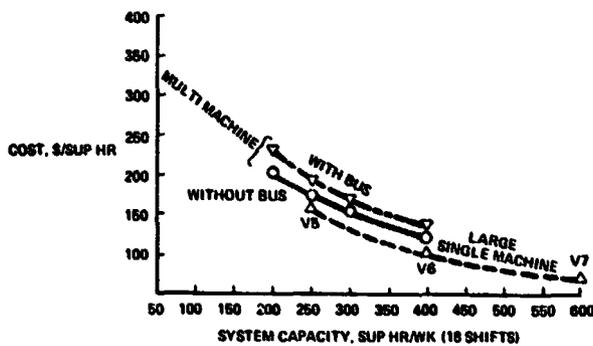


FIGURE 5.—Cost/performance comparison of architectures.

CONCLUSIONS

Data systems in support of remote-sensing applications are getting larger. Many applications can no longer be satisfied with a single minicomputer in a convenient laboratory operated prime shift by the analyst. As the trend continues, the Earth resources community will be forced to consider both the complexities and the efficiencies offered by large data systems, operated three shifts a day, by operations personnel. Data centers in support of Earth resources applications will appear, and users will be increasingly separated from the actual computational resources. On the other hand, image terminal resources will become increasingly available in user work areas or even at the user's own desk.

This environment already exists and is accepted by the low-bandwidth alphanumeric terminal user. The rapid advances in raster-scan image terminal technology, accompanied by the rapid decline in cost of the refresh memory required by raster-scan systems, make the extension of this environment available even now to Earth resources users.

Significant economies of scale will result if the diverse data systems supporting Earth resources at an installation are combined into a single data center. This paper has presented several architectures and associated costs for the large data system supporting such a data center. Recurring cost factors (maintenance and operations) currently slightly favor the single large-machine architecture, but other factors may dictate the choice of one of the two multimachine architectures discussed.

The construction of a quantitative workload model in support of a data system acquisition

(ERDS) has been demonstrated; any user contemplating a data system acquisition should do the same.

REFERENCES

1. A System/Subsystem Design for the USDA Which Will Provide a User Advanced System. SISO-TR633, Ford Aerospace and Communications Corp., Sept. 1976.
2. Specification for the Landsat-D System. GSFC-430-D-100, NASA Goddard Space Flight Center, July 1977.
3. Green, W. B.: Applications of Interactive Digital Image Processing to Problems of Data Registration and Correlation. 1978 National Computer Conference (Anaheim, Calif.), June 1978.
4. Spencer, P. W.: LARSYS User's Manual. Purdue University Laboratory for Applications of Remote Sensing, Dec. 1974.
5. Bracken, P. A.; et al.: Atmospheric and Oceanographic Information Processing System (AOIPS) System Description. GSFC x-933-77-148, NASA Goddard Space Flight Center, Mar. 1977.
6. Gregor, P. J.; et al.: Requirements and Costs for an Earth Resources Data System. MTR-4709, The MITRE Corporation, Nov. 1977.
7. vanDam, A.; and Stankovic, J.; eds.: Computer, vol. 11, Jan. 1978. (Entire issue concerned with distributed processing.)
8. Datapro 70: The EDP Buyer's Bible. Datapro Research Corporation, 1978.
9. Hillman, H. C.: A Performance Analysis of Several Candidate Computers for NASA's Flight Planning System. MTR-4599, The MITRE Corporation, Mar. 1977.
10. Hillman, H. C.; and Williams, J. N.: A Performance Analysis of Candidate Minicomputers for NASA's Flight Planning System. MTR-4716, The MITRE Corporation, Apr. 1978.
11. Matuck, G. N.: Comparative Analysis of Proposed Data Processing Systems for the NASA ODRC. MTR-4597, The MITRE Corporation, Nov. 1976.
12. H. A. C. Amdahl 470 V/6 Evaluation. Hughes Aircraft Company Computing and Data Processing Corporate Technical Services, Nov. 11, 1976.
13. Kincy, W. P. Jr.; et al.: Conversion of CCF Univac Software. MTR-4710, The MITRE Corporation, Jan. 1978.
14. Heart, F. E.; et al.: A New Minicomputer/Multiprocessor for the ARPA Network. In AFIPS Conference Proceedings, Vol. 42, 1973.

Experiment Results and Accuracy

FOREWORD

This session presents a detailed account of the LACIE results. Comprehensive assessments of LACIE performance in production, area, and yield estimation are described. Emphasis is placed on an assessment of the accuracy of the LACIE estimates in terms of error source isolation and the resulting corrective measures applied throughout LACIE.

The LACIE was conducted in three phases. Phase I, the 1974-75 crop year, was devoted to developing the experimental apparatus; assembling data bases of historical agronomic and weather data; developing sampling approaches, yield models, and specific procedures for handling and analyzing Landsat data; and training people. Preliminary testing and evaluation of the U.S. Great Plains region was also accomplished during this phase.

During Phase II, the 1975-76 crop year, the technology as modified in Phase I was evaluated again in the U.S. Great Plains region, in the prairie provinces of Canada, and in both a spring wheat and a winter wheat region in the U.S.S.R. Exploratory studies of wheat identification and yield model tests were conducted in five other wheat regions: India, the People's Republic of China, Australia, Argentina, and Brazil.

In Phase III, crop year 1976-77, the evaluation in the U.S. Great Plains region was repeated, and the region covered in the U.S.S.R. was expanded to produce total country estimates. The coverage in Canada was reduced to 30 segments. The Canadian investigators collected ground observations for further evaluation of the problems identified in Phase II. Changes made before and during the 1976-77 crop year were thought to comprise significant improvements. These included the implementation of a new machine classification process (known as Procedure 1), an improved stratification of the regions to be inventoried, relocation of selected samples, and revised wheat yield models.

The papers in this session provide the detailed results from the three phases of LACIE. Included are

papers describing the growing conditions under which LACIE estimates were made and the accuracy and performance of the production, area, yield, and crop growth stage estimates.

The three crop years of LACIE have been marked by a wide variety of weather conditions in the regions of interest. "The LACIE Crop Years: An Assessment of the Crop Conditions Experienced in the 3 Years of LACIE" describes the wheat-growing conditions in each crop year for each region for which LACIE estimates were made. "Application of Landsat Digital Data for Monitoring Drought" describes the drought-monitoring capability of Landsat data when used with the wheat growth stage at the time of the Landsat acquisition.

The three papers describing "LACIE Area, Yield, and Production Estimate Characteristics" present the LACIE estimates made during each year for the U.S. Great Plains, the U.S.S.R., and Canada. The papers compare official country estimates to the LACIE estimates and evaluate the LACIE estimates with respect to the 90/90 criterion. They describe the estimates with respect to the scope, reporting schedule, sampling scheme, and associated problems of each LACIE phase.

A more thorough evaluation of the area estimation error sources in the U.S. Great Plains region is given in "Accuracy and Performance Characteristics of LACIE Area Estimates." This paper presents the results of the more detailed investigations based on ground observations obtained in the United States during the three phases of LACIE.

"Accuracy and Performance of LACIE Yield Estimates in Major Wheat Producing Regions of the World" briefly reviews the yield modeling approach and evaluation methodology. The results of testing and evaluation of the operational yield models are presented with emphasis on the U.S. Great Plains region using the most recent 10 years of historical data as an independent test set. The modifications of the models and the reasons for changes are also addressed in the order of occurrence throughout LACIE.

"Accuracy and Performance of LACIE Crop Development Models" describes the application of such models to LACIE needs and gives an evaluation of their performance. The form of the original spring wheat development models is briefly discussed, and the modifications required for winter wheat are presented. Further desired improvements are described in light of the performance evaluation conducted during LACIE.

Finally, **"Economic Evaluation: Concepts, Selected Studies, System Costs, and a Proposed Pro-**

gram" presents a conceptual framework for estimating the value of improved information and an overview of completed studies focused on identifying and quantifying benefits resulting from improved information. Comparisons of the costs of current and satellite-based crop information systems are made. The shortcomings of current systems, which are the strengths of a satellite-based system, are illustrated. In addition, a proposed economic evaluation program for a satellite-based crop production estimation system is discussed.

The LACIE Crop Years: An Assessment of the Crop Conditions Experienced in the 3 Years of LACIE

J. D. Hill^a and D. R. Thompson^b

INTRODUCTION

The LACIE undertook the task of testing, evaluating, and developing the technology needed to utilize remote sensing and associated information for assessing potential wheat production globally. The project extended over three crop years, from fall plantings in 1974 to summer harvests in 1977. The project scope was confined to the U.S. Great Plains during the 1974-75 crop year, then expanded to include Canada and the U.S.S.R. during the latter phases. As one would expect, LACIE encountered a variety of growing conditions under which to develop and test its technical approach.

In order to assess the LACIE results, it is important to keep in mind the crop growing conditions under which the results were obtained. For this reason, a crop condition assessment team was organized as an ad hoc group drawing on the agronomic, meteorological, and other expertise in the various project elements. The team used meteorological data and Landsat spectral data from the growing regions to make their assessments of the conditions and document where anomalies such as drought, floods, and freezes were having an impact on the crop yield and appearance. Weather data available to make the assessments included precipitation totals and average temperatures for periods of a month, as well as for shorter periods of 7 or 10 days.

In the United States, the weekly rainfall and temperature data were used to estimate soil moisture, which was then related to crop needs by a Crop Moisture Index (CMI). This index relates the available water to the usual supply for each week during the growing season. The index has been normalized so that indexes in the range from -1.0 to 1.0 repre-

sent typical moisture supplies. Larger positive indexes indicate surplus water, while larger negative values represent deficiencies. Weekly maps of the CMI were used to infer the general moisture situation in the various crop growing regions of the United States.

Landsat color-infrared full-frame images were used to determine the existence of drought and assess its areal extent. The Landsat digital data was transformed into a Green Index Number (GIN), which was useful in defining the degree of drought and the extent of drought-stricken regions during the final two phases of LACIE. The GIN is a technique utilizing transformed Landsat digital data for detection of agricultural vegetative water stress. It provides a procedure whereby Landsat data from a LACIE segment can be classified as drought affected or not.

As might be expected, the multiyear time period and the large area with which LACIE concerned itself presented opportunities for extremes to be encountered. The project made no attempt to set the extremes apart as nonrepresentative but considered them as cases where the technology would have its most stringent performance tests. The period from 1974 to 1977 was one in which the LACIE countries experienced a balance of extremes, ranging from serious drought in South Dakota during 1976 to overabundant rain in the European U.S.S.R. during 1977. The following discussion will provide more detail about the specific growing seasons encountered in each LACIE country.

PHASE I

The first phase of LACIE consisted of assembling candidate technology for both crop acreage and yield estimation. The test region covered the winter wheat area of the U.S. Great Plains planted in the fall of

^aNOAA Environmental Data and Information Service, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

1974 and the spring wheat area planted in the spring of 1975. Within this growing region, there are many factors which affect plant density and the final crop yields. For winter wheat, moisture is critical when plants are established in the fall and when regrowth begins following dormancy. Water is also critical to the development of spring wheat; the crop depends heavily on pre-season-stored moisture as well as precipitation after planting. For both crops, moisture plays a key role in accurate identification in the Landsat imagery since the analysts rely extensively on characteristic unstressed crop signatures.

Winter Wheat

During the fall of 1974, the U.S. Great Plains had near normal soil moisture at planting. Subsequent rainfall was adequate for establishment in the Southern Great Plains, but dryness was notable in the winter wheat areas of northeastern Colorado, Nebraska, and Montana. The winter wheat crop entered dormancy in good condition throughout the Southern Great Plains and fair to good condition in northern areas.

Winter temperatures were near to or slightly below normal, and cold injury was minimal when compared to that of other years. Some wind damage was notable, however, in parts of eastern Colorado and western Nebraska where fall rains had been below normal and the dry soil was prone to blowing. Wind erosion was also reported in western Kansas and the Panhandle portions of Oklahoma and Texas.

Across the entire Great Plains, cool temperatures from March through May slowed regrowth of the wheat after dormancy. Rainfall from March through May was generally 100 to 150 percent of normal throughout the Great Plains, except for Nebraska, southeastern Colorado, and the western Texas Panhandle, which received little more than half the normal amount. Those areas were further affected by wind damage during May. Greenbugs and local soil disease were reported in central portions of the winter wheat area, but there were no widespread serious disease or insect problems. In general, by spring, most of the wheat was in fair to good condition, except in Colorado where moisture remained critically short.

Just before harvest, late spring and early summer showers created problems. On the first of June, heavy thunderstorms caused some flooding in southwestern and central Oklahoma, while hail was

responsible for lodging in parts of Texas. Harvest was slowed by rains in Kansas, Oklahoma, and Texas, with hail damage and lodging reported to be greater than normal. Rains in June were timely enough, however, to aid the Colorado wheat, which was still in the grain-filling stage.

The crop season LACIE encountered in the U.S. winter wheat area during 1974-75 was ideal for the first experience of the project. A wide variety of crop conditions was experienced, but there were no widespread problems of a catastrophic nature. Good weather for establishment and postdormancy growth allowed a normal progression of crop signatures, and the technology was tested in what might be considered a "most likely" case.

Spring Wheat

In the spring wheat area of the Northern U.S. Great Plains, seeding was delayed 2 to 3 weeks by frequent rain and wet fields. When the wheat was finally planted, it emerged in fair to good condition, but the ample soil moisture promoted shallow root development. Timely rains maintained good moisture through June, with very heavy rains occurring over eastern North Dakota and western Minnesota on June 28-29. Considerable local flooding occurred in portions of the Red River Valley. During July, hot, dry weather stressed the shallow-rooted crops and forced early maturity. July temperatures averaged 6° (F) above normal in some areas. Frequent showers recurred at harvest and may have caused some lodging losses during that critical period.

Even though some local flooding, insect, hail, drought, and disease problems were reported, they were not widespread episodes likely to cause anomalous crop appearance over large areas. The official USDA yield estimates indicated near-normal crop conditions. These estimates are shown in table I.

PHASE II

U.S.S.R.

During Phase II of LACIE, the technology to estimate wheat production was applied in the principal winter and spring wheat growing areas of the

**TABLE 1.—1975 (LACIE Phase I)
U.S. Great Plains Official Wheat Yields^a**

State	1975 yield, bulacre	4-year (1972-75) average yield, bulacre
<i>Winter wheat</i>		
Texas	23.0	26.4
Oklahoma	24.0	24.4
Kansas	29.0	31.5
Colorado	22.5	24.3
Nebraska	32.0	34.4
South Dakota	30.0	30.4
Montana	35.0	30.3
<i>Spring wheat</i>		
Minnesota	31.0	32.4
North Dakota	29.9	23.0
South Dakota	18.0	18.5
Montana	25.8	23.2

^aUSDA Economics, Statistics, and Cooperative Service.

U.S.S.R. These indicator regions are shown in figure 1 and comprised approximately 83 percent of total Soviet winter wheat and 37 percent of the spring wheat production in 1971, the latest year for which reliable statistics were available.

Winter wheat.—In the winter wheat area, dry fall weather allowed planting to proceed on schedule. During crop establishment, weather conditions re-

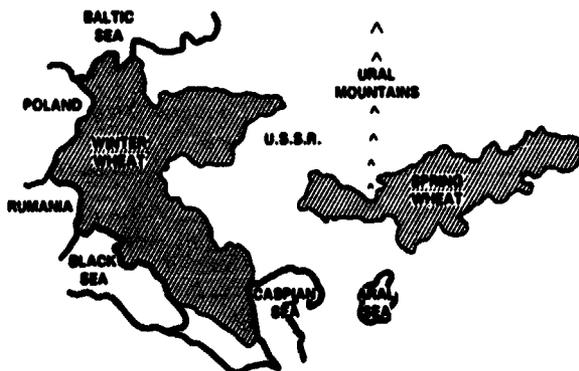


FIGURE 1.—U.S.S.R. winter and spring wheat indicator regions.

mained dry and soil moisture became short. A region of the important Ukraine area had less than 25 percent of normal soil moisture at the end of October.

During the winter dormancy period, precipitation was near to or slightly below normal; however, snow cover was more extensive than usual. By the end of January, snow cover began retreating from the southern portion of the winter wheat area, and extremely cold temperatures throughout the first week in February caused cold injury to the exposed crop in the region east of the Black Sea.

After dormancy, the winter wheat area received ample rainfall and mild temperatures provided favorable conditions for crop development. The adverse conditions of fall establishment were completely reversed, and there were no serious soil moisture shortages during the heading and grain-filling periods. The nearly ideal spring weather was reflected in the final average yield for Soviet winter wheat which was 27 ql/ha, a near record. These very high yields show how winter wheat can rebound in the spring after undergoing very poor conditions during establishment. This situation indicates caution must be used in making judgments of final crop yield on the basis of early-season growing conditions since the wheat plant can recover from extreme conditions if not completely killed.

Spring wheat.—The spring wheat area did not receive the early rains which benefited the winter wheat crop. Both April and May were drier than normal in the area east of the Ural Mountains. By mid-June, soil moisture was below normal in all but the central portion of the growing area LACIE was investigating. During the entire period from April when the crop was being established through July when it was heading, rainfall was less than normal except in the extreme northern and eastern sections of the indicator region. The dryness persisted through the harvest season and minimized possible harvest losses.

Spring wheat grown in the European portion of the U.S.S.R., outside the LACIE test region, experienced better moisture distribution than that in the Asiatic portion throughout the entire growing season. As a result of more favorable conditions over this major spring wheat producing area, yields for the entire Soviet spring wheat crop averaged 13 ql/ha, a near record. No detailed yield statistics are available yet to characterize the impact of the dryness on the 37 percent of the crop for which LACIE prepared estimates or to verify whether it was truly an anomaly within the country.

Canada

During Phase II, LACIE technology was tested in the spring wheat growing area of the Canadian prairie provinces. This area is contiguous to the northern border states of the United States and covers the southern portions of Alberta, Saskatchewan, and Manitoba, from the Canadian Rockies to Lake Winnipeg. The region is outlined in figure 2.

Wheat grown in the Canadian prairie provinces is normally planted during April and May and depends on both pre-season and growing-season moisture. During 1976, precipitation was near normal from January through June and provided adequate soil moisture reserves.

During the planting season, rainfall was near or below normal and most of the crop was seeded without major delays. Precipitation during the April-May establishment period was less than half of normal, but rains improved in June and encouraged good growth.

As the wheat progressed into the critical heading and grain-filling stages, July precipitation became very erratic; however, the early-season moisture was apparently sufficient to support the crop. Wheat stands were excellent, with only light infestations of disease and insects.

By late July, the wheat heads were filling well, but hot, dry weather in early August caused premature ripening in parts of Saskatchewan and Manitoba. Generally dry conditions favored harvest during August throughout all regions except portions of

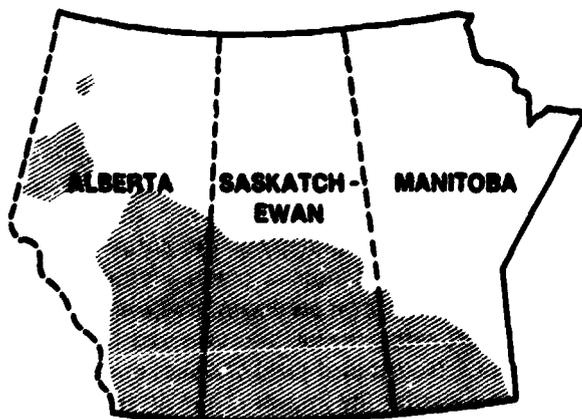


FIGURE 2.—Outline map of the Canadian prairie provinces showing wheat area shaded.

Alberta. Heavy rains and hail in the Peace River region during late August caused extensive lodging.

The 1976 spring wheat growing season in the Canadian prairie provinces was very good. Moisture was never seriously limiting and insect and disease damage was minimal. There were no widespread harvest losses. As a result of these good growing conditions, yields of spring wheat were above average. The final 1976 yields officially released by the Canadian government are shown in table II.

U.S. Great Plains

Winter wheat.—The winter wheat area of the U.S. Great Plains from South Dakota to the Texas Panhandle had rainfall that was below normal during the summer of 1975. Consequently, soil moisture was short during fall planting. From September through November, rainfall in the growing region totaled 2 to 4 inches, which was generally only 50 to 75 percent of normal (fig. 3). Wheat seeding proceeded on schedule, but emergence was below normal because of the dryness. Some rainfall was received during November, but cooler than average temperatures limited plant growth and the chance for improved establishment. Wheat condition varied widely from state to state, but the area most seriously affected by drought was southwestern Kansas, southeastern Colorado, and the Panhandle regions of Oklahoma and Texas.

Through the winter months, the weather continued to be dry with little or no rainfall reported in the Southern Great Plains. From December through February, little or no snow cover was received and the dry soil became vulnerable to serious erosion. Farmers in Colorado, Kansas, western Oklahoma,

TABLE II.—Official 1976 Canadian Yields With Comparisons^a

Province	Final yield, bu/acre		Average yield, bu/acre, 1965-74
	1976	1975	
Manitoba	27.1	25.2	25.3
Saskatchewan	31.3	25.6	23.2
Alberta	32.7	29.9	26.1

^aStatistics Canada.



FIGURE 3.—U.S. precipitation (percent of normal), September to November 1975.

and the Texas Panhandle began to abandon large areas and plow the land to control wind erosion. The dryness was accompanied by temperatures which averaged 4° to 6° above normal for some areas. This not only increased the effect of the dryness but encouraged the development of insects, particularly greenbugs.

Following dormancy, crop conditions were fair to poor in the Southern Great Plains; however, the northern growing regions had received normal fall rainfall and winter snows, which provided excellent conditions for winter wheat. Very warm temperatures from February through March pushed crop development ahead of normal, and the Texas crop ripened under considerable moisture stress. Ample rains began in early April at about the last opportunity to benefit wheat in Oklahoma and Kansas. Conditions improved, but stands were thin and weeds became competitive.

The unusual growing conditions caused the wheat to appear much different from normal in the Landsat imagery, and analysts were not able to correctly identify many of the fields, particularly in Oklahoma. As

a result, the acreage estimates were consistently below the actual values in that area. The LACIE yield estimates were also low, particularly in the Panhandle regions of Oklahoma and Texas. This was due to the large amount of dryland area abandoned, leaving a greater than normal proportion of the production to the high-yielding irrigated area.

Rains benefited wheat in the Northern Great Plains during April, and the entire area experienced cool May temperatures. The cooler weather prolonged the grain-filling period and reduced moisture demands. By mid-May, the general moisture situation had improved except in eastern Colorado and western Texas, which were still experiencing shortages. The cool weather brought a late freeze to eastern Kansas on May 3, when temperatures fell to 28° F. The wheat was in the jointing-heading stage and losses were estimated to be as much as 50 percent in some counties. A late freeze also occurred in South Dakota during mid-May, causing severe damage to winter wheat, which was mostly in the boot stage.

The cool temperatures continued from May into



FIGURE 4.—Crop Moisture Index for June 19, 1976.

June and slowed ripening in the Central Plains. Harvest activity fell slightly behind normal as a result. Rains caused some local harvest delays and were responsible for losses in southeastern and south-central Kansas on July 1. Heavy rain on that date caused flooding and lodging of the grain in that area. In most other areas, the winter wheat harvest season experienced dry weather to expedite the combining with a minimum of harvest losses.

Spring wheat.—In the spring wheat areas of the Great Plains, dry April weather provided favorable planting conditions. The dryness persisted into May and slowed early plant development, while frost in mid-May damaged young top growth. The spring wheat area received less than 75 percent of the normal May precipitation. As a result, establishment of

the crop was only fair; however, rains developed in mid-June to aid all areas except South Dakota and southwestern Minnesota. Good moisture prevailed over Montana, North Dakota, and part of Minnesota through the heading and grain-filling stages. The remainder of the area highlighted by large negative crop moisture indexes in figure 4 remained drought stricken, and, when pastures began to fail, many farmers cut their wheat for hay or turned cattle into it for grazing.

Harvest conditions were excellent and ample periods of dry weather allowed combining to proceed faster than normal. Lodging and other harvest losses were minimized. For South Dakota, harvest conditions could be of little importance since the severe midsummer dryness had already devastated the

**TABLE III.—1976 (LACIE Phase II)
U.S. Great Plains Official Wheat Yields^a**

<i>State</i>	<i>1976 yield, bushels</i>	<i>4-year (1972-75) average yield, bushels</i>
<i>Winter wheat</i>		
Texas	22.0	26.4
Oklahoma	24.0	24.4
Kansas	30.0	31.5
Colorado	21.5	24.3
Nebraska	32.0	34.4
South Dakota	18.0	30.4
Montana	32.0	30.3
<i>Spring wheat</i>		
Minnesota	32.4	32.4
North Dakota	24.7	23.0
South Dakota	13.2	18.5
Montana	29.4	23.2

^aUSDA Economics, Statistics, and Cooperative Service.

spring wheat crop in portions of that state. The relative impact of the drought in South Dakota is indicated by the reduced yields of both the spring and winter wheat for the state, as shown in table III. Statewide, the yields were 40 percent below recent averages; however, they were reduced by as much as 60 percent in some northeastern sections where the drought was most severe.

In the affected area of South Dakota, corn, pasture, and other crops were as severely decimated as the spring wheat. At the time when the spring wheat is normally ripening and other crops are still green, much of the Landsat imagery presented a colorless appearance for all crops. When the drought-stricken crops took on the appearance of ripened grain, the analysts classified an excessive number of fields as wheat, thus causing a large overestimate of spring wheat area in South Dakota.

Development of Drought-Monitoring Capability

The Phase II experience of LACIE in the U.S. Great Plains was the project's first encounter with widespread serious drought. It tested the capability of the technology to produce yield estimates repre-

sentative of severely stressed crops and to classify wheat when its characteristic spectral appearance was changed by drought. Fortunately, sufficient weather data were available in the United States to characterize the degree and extent of drought. However, it is possible that routine weather data available from some foreign areas may not allow such precision. An effort was undertaken to develop the Landsat data as a tool for monitoring the extent and severity of drought in such an area.

Once the problem area was defined from meteorological data, Landsat color-composite transparencies, prepared from band 4 (0.5 to 0.6 micrometer), band 5 (0.6 to 0.7 micrometer), and band 7 (0.8 to 1.1 micrometers) of the satellite's multispectral scanner, were used to refine the initial problem area delineated from meteorological data. A total of 33 Landsat full-frame images (100 by 100 nautical miles) were required for the Southern U.S. Great Plains analysis and 14 Landsat images for the South Dakota analysis. These color transparencies were evaluated by comparison to Landsat imagery for essentially the same date in previous years and also to previous 9-day acquisitions for the current year. Both Landsat-1 and Landsat-2 were used to acquire 9-day coverage for the drought analysis. Normal healthy green vegetation on the ground is recorded on the Landsat color composites as bright red. As moisture stress reduces the vigor of the vegetation on the ground, the Landsat-recorded signatures correspondingly decrease in redness from what one normally sees at the same crop stage. Thus, by relating the lack of redness in the signature to the red signatures that should have been present, the areal extent of drought was delineated on a mosaic of Landsat images over the potential drought area.

The drought-affected area in the U.S. Southern Great Plains was determined from Landsat data to be located in the southwestern corner of Kansas, in southeast Colorado, and in the Oklahoma and Texas Panhandles. The areal extent of the affected area as of April 12, 1976, is shown in figure 5.

The drought severity within the area was rated subjectively by comparing the 1976 and 1975 Landsat imagery. These ratings correlate well with the acreage losses estimated from ground-based observations. The CMI for April 10, 1976, also verified that this general area was undergoing moisture stress (fig. 6).

The initial drought-affected area in the Northern Great Plains, as determined from full-frame images, was located within South Dakota. From April 18 to

June 12, 1976, the area appeared to be deteriorating, but the full-frame imagery did not indicate severe effects. The June 11 to 13 overpass showed the effects of the drought were becoming pronounced. The drought-afflicted area delineated at this time continued to expand until the July 8 to 11 overpass when it stabilized (fig. 7). From this overpass, the drought area was rated subjectively as having been severely or moderately affected by the dryness.

The July 10, 1976, Landsat 100- by 100-nautical-mile image (fig. 8) shows the western edge of the severe drought damage in South Dakota. There is a lack of red signatures on the right side of the image when compared to a July 7, 1975, image (fig. 9). The 1975 image shows red signatures, especially in the natural drainageways, that are not in the 1976 image.

The drought analysis of the Great Plains indicated that the Landsat data contained meaningful information about moisture stress. To automate the analysis, the digital data from Landsat was transformed into a quantitative measure of greenness called the green index number or GIN, which has since proved useful

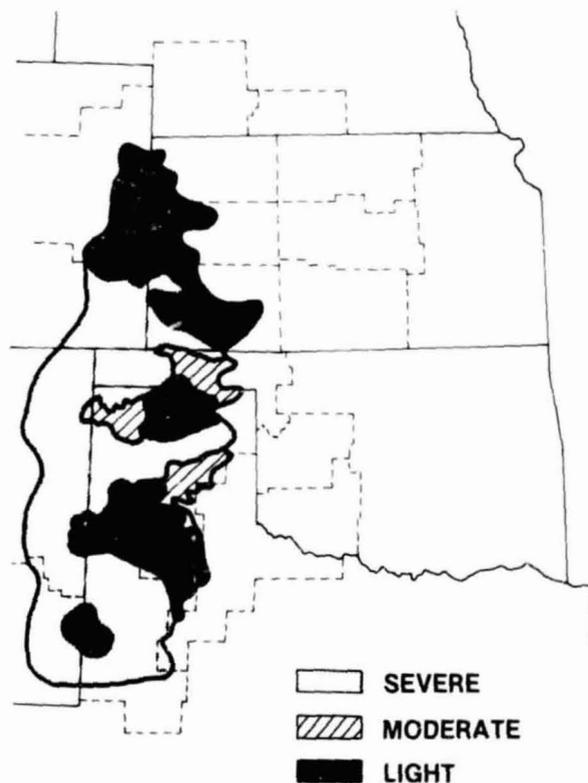


FIGURE 5.—Drought conditions determined from full-frame Landsat imagery for April 12, 1976.

in the analysis of drought conditions in other countries. The specific technical approach and results are presented in the Experiment Design Section.

PHASE III

U.S. Great Plains

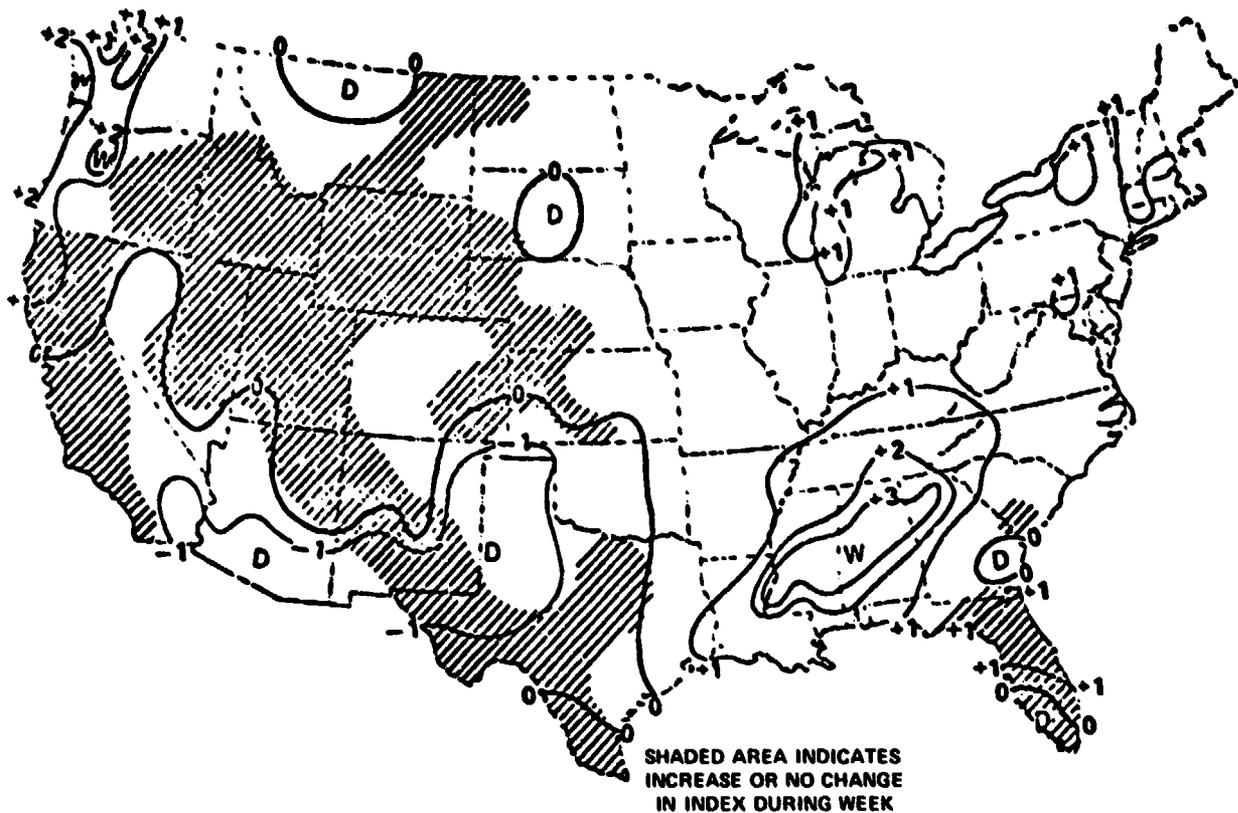
Winter wheat.—The U.S. Great Plains winter wheat region was dry before fall planting, but a series of timely rains replenished topsoil moisture at planting. Wheat generally had adequate moisture for germination and emergence; however, cold weather in October caused the wheat to enter dormancy early with little vegetative growth. The winter period was colder than normal with variable snow cover and below normal precipitation. During February, temperatures across the winter wheat region averaged above normal and encouraged early green-up. Conditions were conducive to rapid increases in ground cover during March and April as continued warm temperatures were accompanied by timely spring rains.

Moisture was ample in most states during the critical grain-filling stage. The only notable exception was Colorado, where dryness stressed the wheat and reduced potential yield. Temperatures across the Great Plains were very warm during April and May, ranging as much as 10° above normal in some areas.

No widespread adverse weather occurred over Texas and Oklahoma during harvesting. In Kansas, heavy rains affected the eastern sections of the state during the third week in June and some hail was reported.

In the northern winter wheat states of Montana and South Dakota, growing conditions were highly variable. The extreme dryness which affected South Dakota in 1976 caused the crop to be planted with little moisture available for germination and early growth. The wheat entered dormancy in poor condition and was susceptible to further stand reduction by winterkill. The lack of vigor suspected in South Dakota winter wheat was confirmed by analysts who reported poor signatures on the Landsat imagery obtained from that region after dormancy. Timely showers during spring and early summer improved the condition of the wheat that survived.

The Montana wheat had better moisture conditions for establishment than the South Dakota wheat and received the benefit of early spring showers;



- | | |
|---|---|
| ABOVE 3.0 - EXCESSIVELY WET, SOME FLOODING | 0 TO -1.0 - TOPSOIL MOISTURE SHORT, RAIN NEEDED |
| 2.0 TO 3.0 - TOO WET, SOME STANDING WATER | -1.0 TO -2.0 - TOO DRY, DETERIORATING PROSPECTS |
| 1.0 TO 2.0 - FAVORABLE, SOME FIELDS TOO WET | -2.0 TO -3.0 - TOO DRY, YIELD PROSPECTS REDUCED |
| 0 TO 1.0 - FAVORABLE FOR NORMAL GROWTH | -3.0 TO -4.0 - POTENTIAL YIELDS CUT BY DROUGHT |
| | BELOW -4.0 - EXTREMELY DRY, MOST CROPS RUINED |

FIGURE 6.—Crop Moisture Index for April 10, 1976.

however, the showers became less reliable during June. The GIN was used to estimate drought intensity on 9- by 11-kilometer sample segments in eastern Montana. On May 20, GIN indicated moisture stress was present. The CMI map indicated abnormally dry conditions at that time only in eastern Montana, but by late June, dryness was prevalent over almost the entire state. This intensified moisture stress during what is usually the heading stage caused significant reductions in yield in Montana.

In summary, during Phase III, conditions were

such that the U.S. Great Plains experienced no widespread disease or insect problems. Moisture stress was a localized problem which did not affect the major winter wheat growing region after dormancy. Even though the crop had poor conditions for overwintering, growing season weather was generally favorable after green-up.

Winterkill.—Conventional and satellite sources of meteorological data were monitored and the data analyzed throughout the late fall, winter, and early spring to delineate areas where the wheat crop was exposed to frigid temperatures with minimal or no

snow cover—the criteria for winterkill. These criteria were obtained from a search of pertinent literature describing the tolerance of wheat to cold. Although the critical limits depend on variety, plant moisture content, and several other factors, a value of -20°C is generally an appropriate threshold. The only Great Plains potential winterkill locations that were determined to be without snow cover during the occurrence of critically cold temperatures were extreme northern Kansas, Nebraska, and the central area of South Dakota, where fall dryness had already caused poor wheat-stand development. Field reports received from this area later in the season indicated wheat had been affected by winterkill and some fields seriously thinned or abandoned.

Spring wheat.—Much of the U.S. spring wheat region entered the 1977 growing season year with a serious deficiency of subsoil moisture carried over from the previous fall. During most of the crop season, however, timely showers provided adequate rainfall to promote good growth in South Dakota, Minnesota, and most of North Dakota. Only Montana experienced widespread drought, although the western portion of North Dakota had less severe moisture shortages. During June and July, when the wheat in the Northern Great Plains was heading and

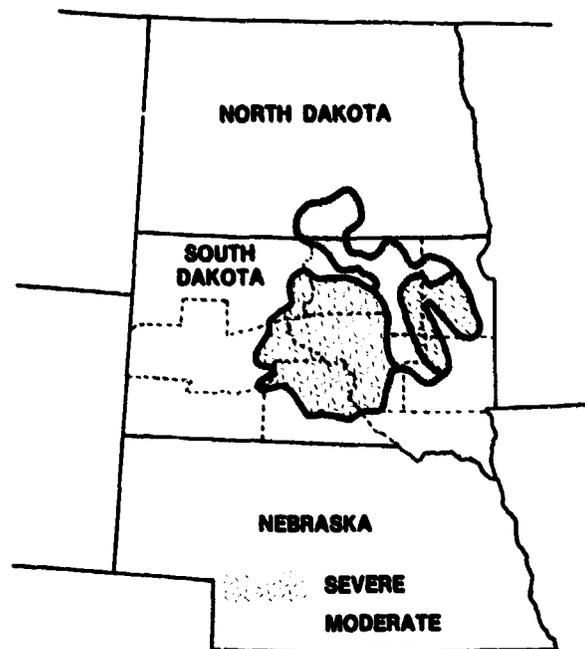


FIGURE 7.—Drought conditions determined from full-frame Landsat imagery for July 8 to 11 and 17 to 20, 1976.

filling, the GIN was used to evaluate the area for indications of stress. Most of Montana's spring wheat, except isolated pockets in the southeastern corner and northcentral portion, exhibited signs of drought stress, but this was the only state where such conditions were noted.

U.S.S.R.

Winter wheat.—The U.S.S.R. winter grain planting season was characterized by ample soil moisture and the early onset of cold temperatures. Warm and moist September weather promoted the initial establishment of the crop in many areas, but growth was quickly limited by temperatures in October which fell below the optimum needed for vigorous development. In many areas, the cold temperatures occurred within 2 to 3 weeks after normal planting dates while, in a few, temperatures sufficiently cold to bring on dormancy developed at normal planting time. Landsat imagery acquired in late fall, before snow cover, indicated weak, spotty wheat signatures when compared with imagery from previous years, thus confirming the poor growing conditions.

January was colder than normal in the winter wheat region; however, temperatures near or above normal during December and February caused the 3-month period to average warmer than the long-term mean. The mild temperatures in February were an indicator of the warmer than normal weather which was to continue through the spring season. In most crop regions, this was accompanied by adequate and timely precipitation. The seasonal weather typical of the important winter wheat growing region of the northeast Caucasus is shown in figure 10. It illustrates the early onset of cold fall temperatures and the warmer than normal weather during February, March, and early April which pushed crop development ahead of normal.

Winterkill.—During December, precipitation was near normal and produced adequate snow cover over all of the winter wheat region with the exception of the extreme southern portions. On January 2 and 5, temperatures dropped below -20°C in the areas with poor snow cover and may have caused loss of some plant stands. The regions affected were Krasnodar Kray and the northeast Caucasus, along with small portions of the eastern Ukraine and the Lower Volga.

In the region where cold injury was suspected, the Landsat imagery was reviewed closely to determine



FIGURE 8.—Drought effect on Landsat image 5448-16135 acquired on July 10, 1976, over central South Dakota.

whether the wheat's characteristic appearance was altered in an identifiable way. It would be suspected that damaged wheat would be thin and irregular after green-up unless it was severely injured, in which case the land would be plowed and planted to another crop. One of the segments studied is shown in figure 11, where the bottom panel is a machine classification map on which the fields identified as wheat are a light color. Wheat planted in several fields during the fall shows the brightest red color where it has emerged and is growing on October 7. One of the fields, outlined in white, shows establishment in the

fall but has been dropped from the inventory on April 4, possibly because it has been damaged and does not exhibit the characteristic green-up after dormancy.

Figure 12 is imagery of a segment located on the eastern edge of the delineated potential winterkill area. The good response seen in this segment during dormancy when there is snow in the fence rows indicates good fall establishment. No effects of the cold temperatures during January can be seen in the winter grain response in the spring acquisition. Clearly, these two examples show that evidence of

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR



FIGURE 9.—Normal Landsat image 2166-16493 acquired on July 7, 1975.

winterkill on the satellite imagery is very subtle and requires extensive analysis to assess the degree or extent of damage.

Spring wheat.—The U.S.S.R. spring wheat growing region experienced dry weather during the spring months, and planting progressed in a timely manner. The moisture supply throughout the remainder of the growing season varied considerably between regions, however. July was a particularly important period as the wheat moved through the critical grain-filling stage. During July, the western portion of the

spring wheat area received approximately 100 to 125 millimeters of rainfall, almost twice the normal amount.

In the area from the Middle Volga region eastward, moisture was most variable. July rainfall across the northern portion of the spring wheat belt totaled 50 to 100 millimeters—near-normal precipitation. Temperatures also averaged near normal there.

The southern portion of the belt received less rainfall with crop region averages ranging generally from 25 to 75 millimeters. Temperatures in those

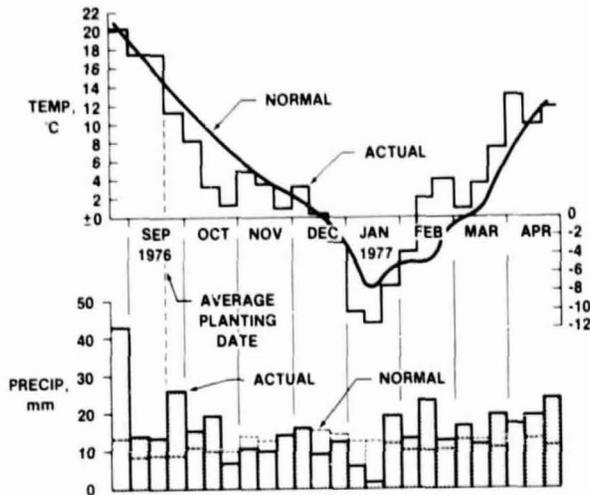


FIGURE 10.—Climograph for crop region 10, northeast Caucasus region of the U.S.S.R.

areas are normally warmer, however, than in the northern portion of the spring wheat area and make precipitation less effective. Some reports had indicated below normal soil moisture existed in that area early in the season and the lack of any significant opportunity to renew the supply meant that plant stress was likely. Imagery and digital data obtained from Landsat overpasses were used to evaluate whether stress was actually present in the area from the Urals eastward.

The GIN program developed by LACIE was run on two separate Landsat overpasses in the U.S.S.R. spring wheat region during July when the crop was in the critical jointing-heading stage. Data for June 23 and July 2 to 18 (fig. 13) indicated that much of the U.S.S.R. spring wheat region was undergoing stress at that time. The next Landsat pass during July 19 to 30 (fig. 14) indicated that stress conditions were still present. The total area of probable moisture stress for July, using the combined data, is shown in figure 15. These data imply that during July, much of the U.S.S.R. spring wheat in the area east of the Urals and generally south of a line from Orenburg to Omsk experienced moisture stress.

Landsat color-infrared full-frame 100- by 100-nautical-mile segments of four separate areas were also examined for indications of distinct changes in the spectral quality that would imply variations in crop vigor or crop development stage. Apparent within the imagery was an obvious darkening of the soil where scattered showers had recently occurred, indicating that rainfall was not altogether absent.

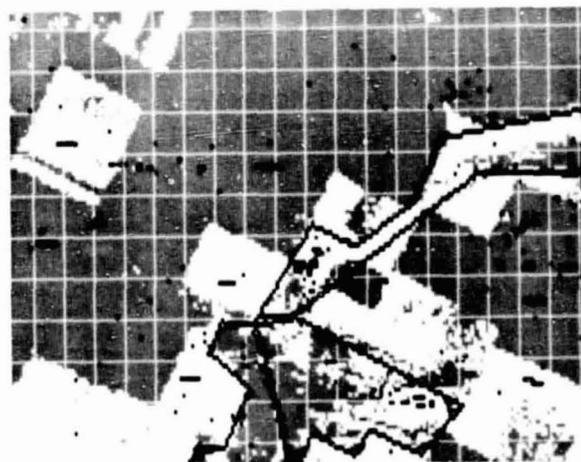
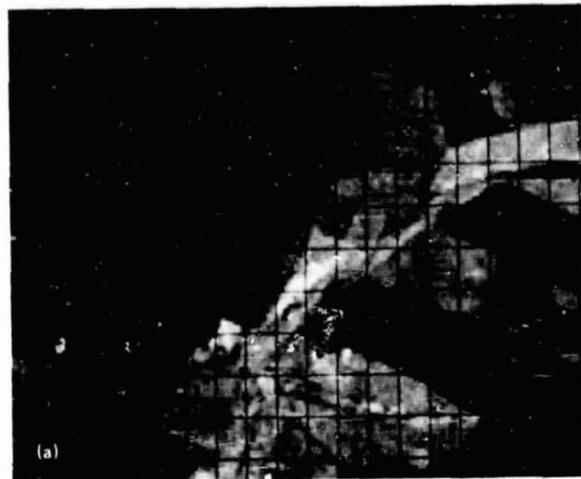


FIGURE 11.—Example of possible winterkill, northeast Caucasus, U.S.S.R. (a) Partial fall emergence, October 7, 1976. (b) Spring greening up, April 4, 1977. (c) Machine classification map, April 4, 1977 (light color = wheat, bordered area = city).

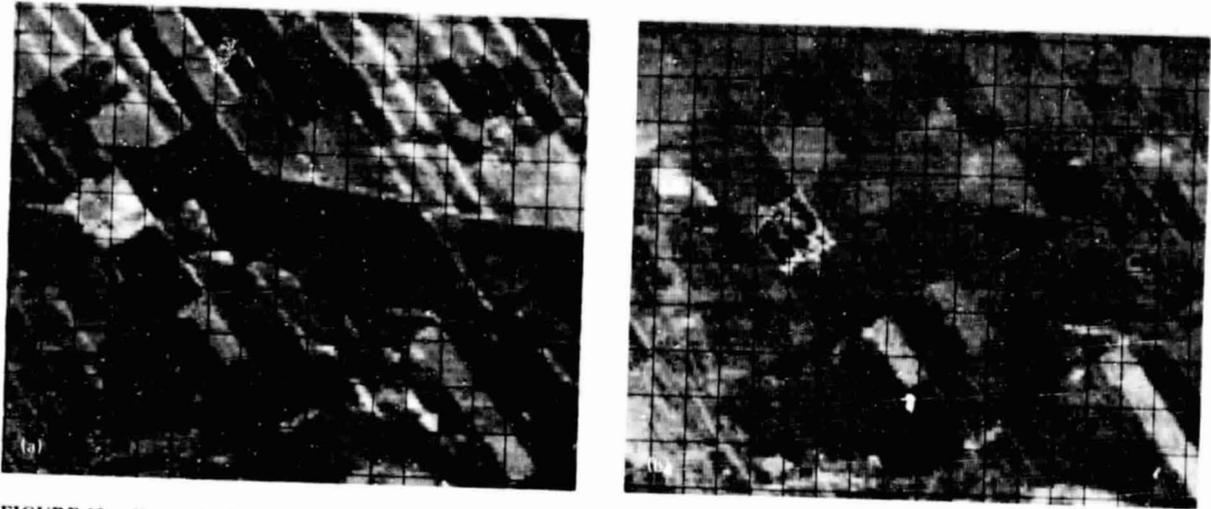


FIGURE 12.—Example of no winterkill, Krasnodar Kray, U.S.S.R. (a) Dormant, light snow, January 7, 1977. (b) Spring greening up, April 6, 1977.

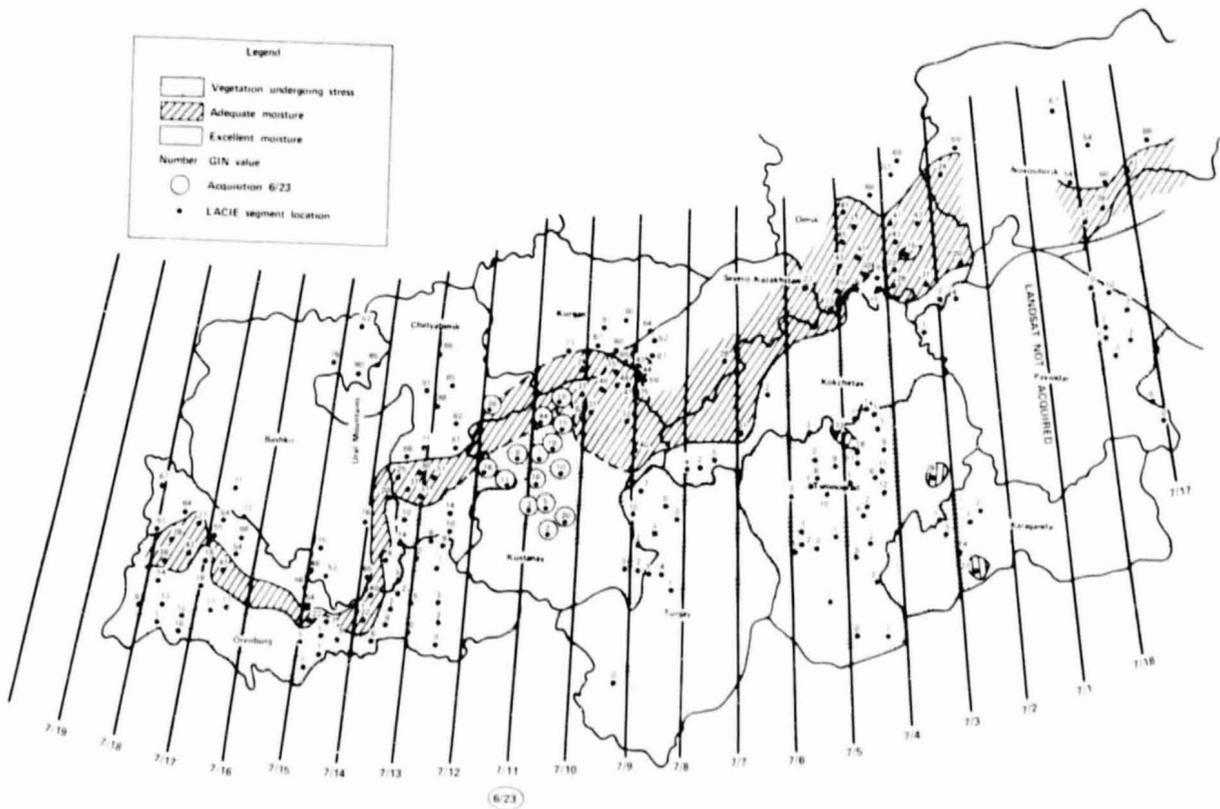


FIGURE 13.—Moisture conditions over U.S.S.R. spring wheat from the LACIE GIN monitoring program; Landsat data acquired June 23 and July 2 to 18, 1977.

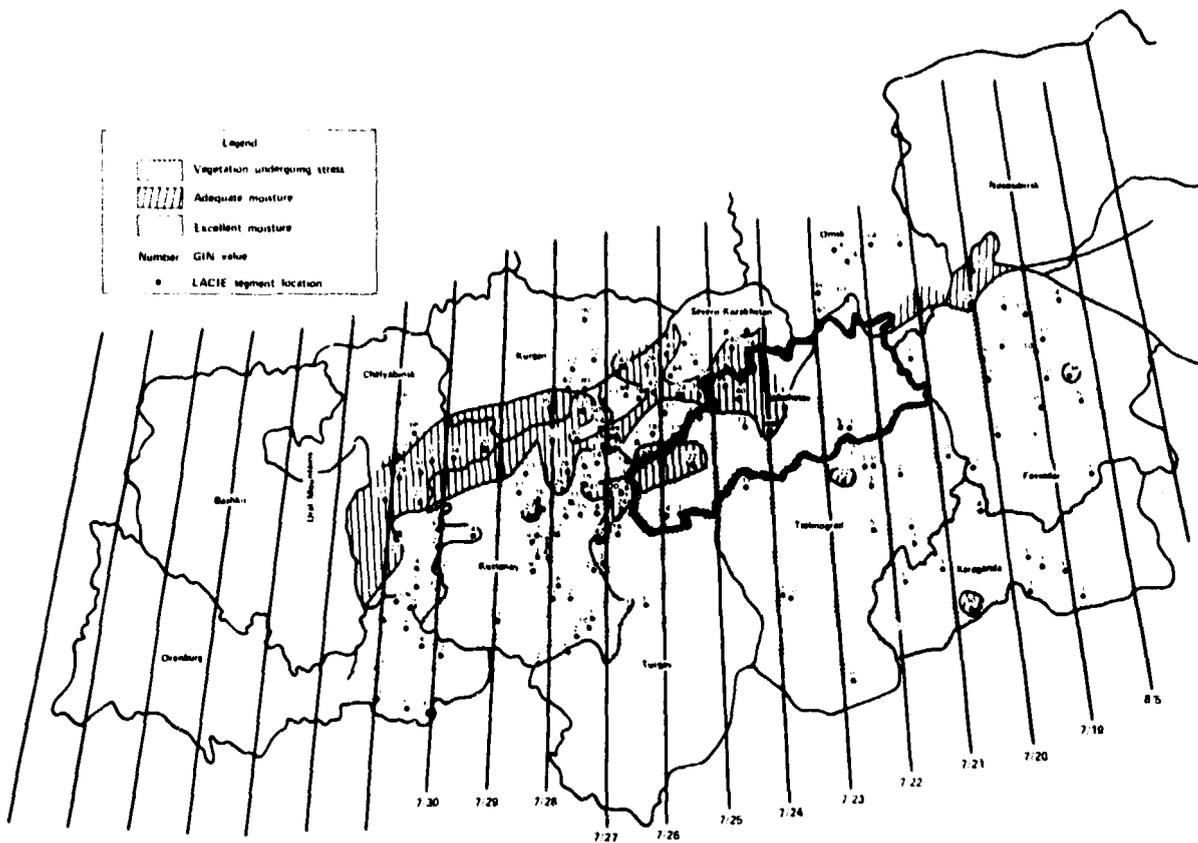


FIGURE 14.—Moisture conditions over U.S.S.R. spring wheat from the LACIE GIN monitoring program; Landsat data acquired July 19 to 30, 1977.

A Landsat overpass acquired on July 4, 1977, is shown in figure 16. This area extends from north to south between the cities of Omsk and Tselinograd. The vigorous red crop signatures in the north lose much of their intensity as one moves southward. In the central portion, very few field patterns can be identified. An acquisition showing the central portion 1 year earlier on July 25, 1976, indicates that cultivated fields do exist throughout this area and should be apparent. Clearly, the spring wheat did not develop well early in the season and stands do not exhibit the signature one might expect at a time when the crop should be near heading. Rains later in July could have caused a slight improvement in the crop condition, but significant reduction in potential yield had already occurred when moisture shortages limited development of the plant stands.

The sequence of Landsat imagery acquired from the stressed region shows abnormal crop signatures, and alternative cropping practices may have been im-

plemented in some areas. In figure 17, a series of acquisitions for a 5- by 6-nautical-mile segment in southern Kustanay is shown which indicates some likely problems. On April 29, the fields in the lower portion are being prepared; on June 4, they have emerged but show weak color-infrared signatures. On July 28, much of the crop has matured, while the same area on August 1, 1976, was still showing active growth. Early maturity is a common response when wheat is under moisture stress.

In figure 17, there is an isolated field in the lower right which shows rapid growth between June 4 and July 28. This is probably a field that was abandoned during June and planted to a crop such as millet, which could be harvested for silage before a freeze. However, abandonment and reseeded was not a widely identifiable practice this year.

The Landsat imagery contains additional information indicative of wheat condition in the dry areas. In figure 18, the large area of small grains in the upper

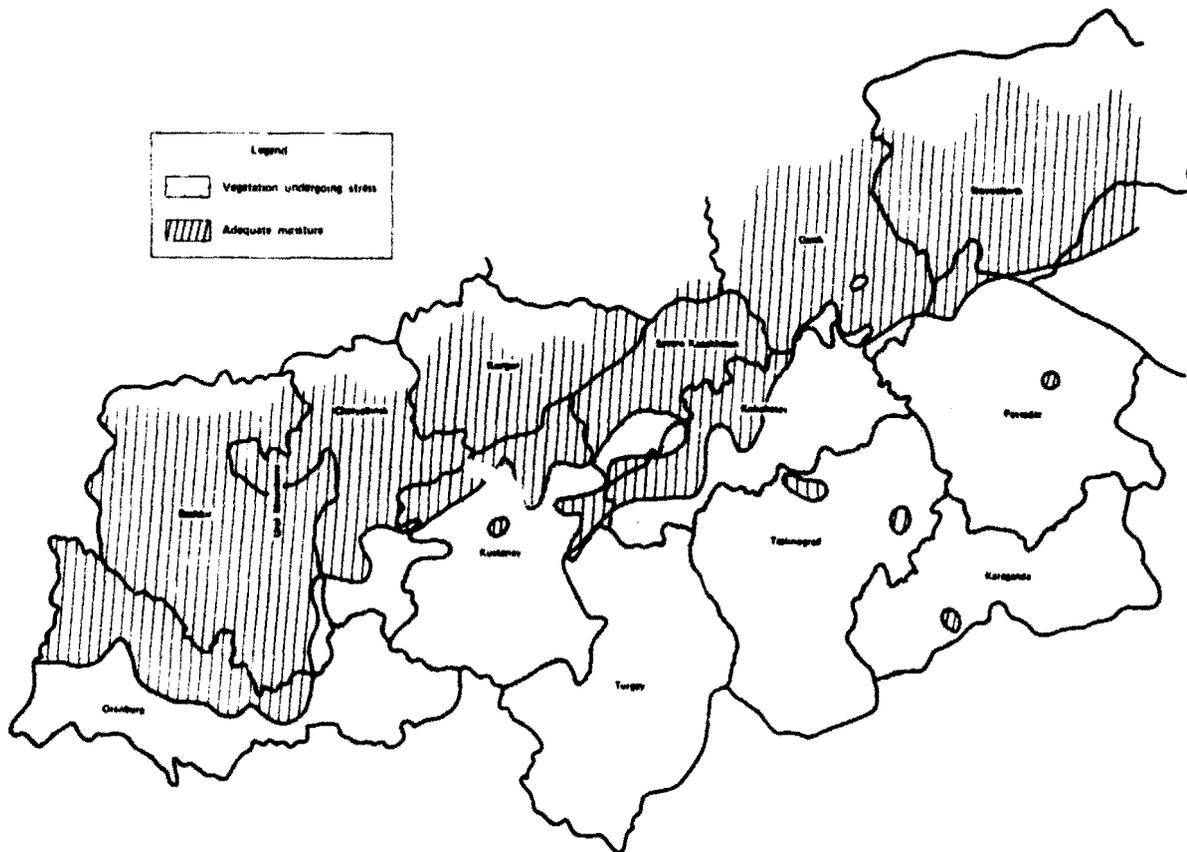


FIGURE 15.—Moisture conditions over U.S.S.R. spring wheat from the LACIE GIN monitoring program for July 1977.

central portion of the segment appears to be progressing past the heading stage on July 27 and presents a dark-red signature. By mid-August, these should be ripe and nearing harvest; however, the August 14 acquisition indicates some red color is still present, a sign of active growth. This is likely the result of vegetation in the form of weeds and secondary tillers from the wheat plants responding to the rains which occurred. The rain probably had the effect of slowing maturity as well. In some areas of the oblast, August rainfall totaled up to 300 millimeters, which is about 500 percent above normal.

Only in isolated cases does this secondary wheat growth produce a significant amount of additional grain; more often it joins with the weeds to interfere with combining operations. In fields which are allowed to stand while secondary growth develops, some loss of standing grain will occur, causing a possible reduction in yield. The extensive presence

of weeds in the fields reduces the quality of the wheat that is harvested.

In summary, the wheat crop in the area east of the Ural Mountains endured the poorest growing conditions of any area harvested during 1977. In the southern portion of the New Lands, drought appears to have made a serious impact; some fields were apparently overseeded, while those left for harvest were probably of low quality.

SUMMARY AND CONCLUSIONS

The LACIE experience in major wheat-producing regions of the world encompassed a wide variety of crop growing conditions. These represented the spectrum of growing conditions likely to be encountered by an operational system using the LACIE technology to perform timely monitoring of global wheat production. It is clear that an abundance of in-

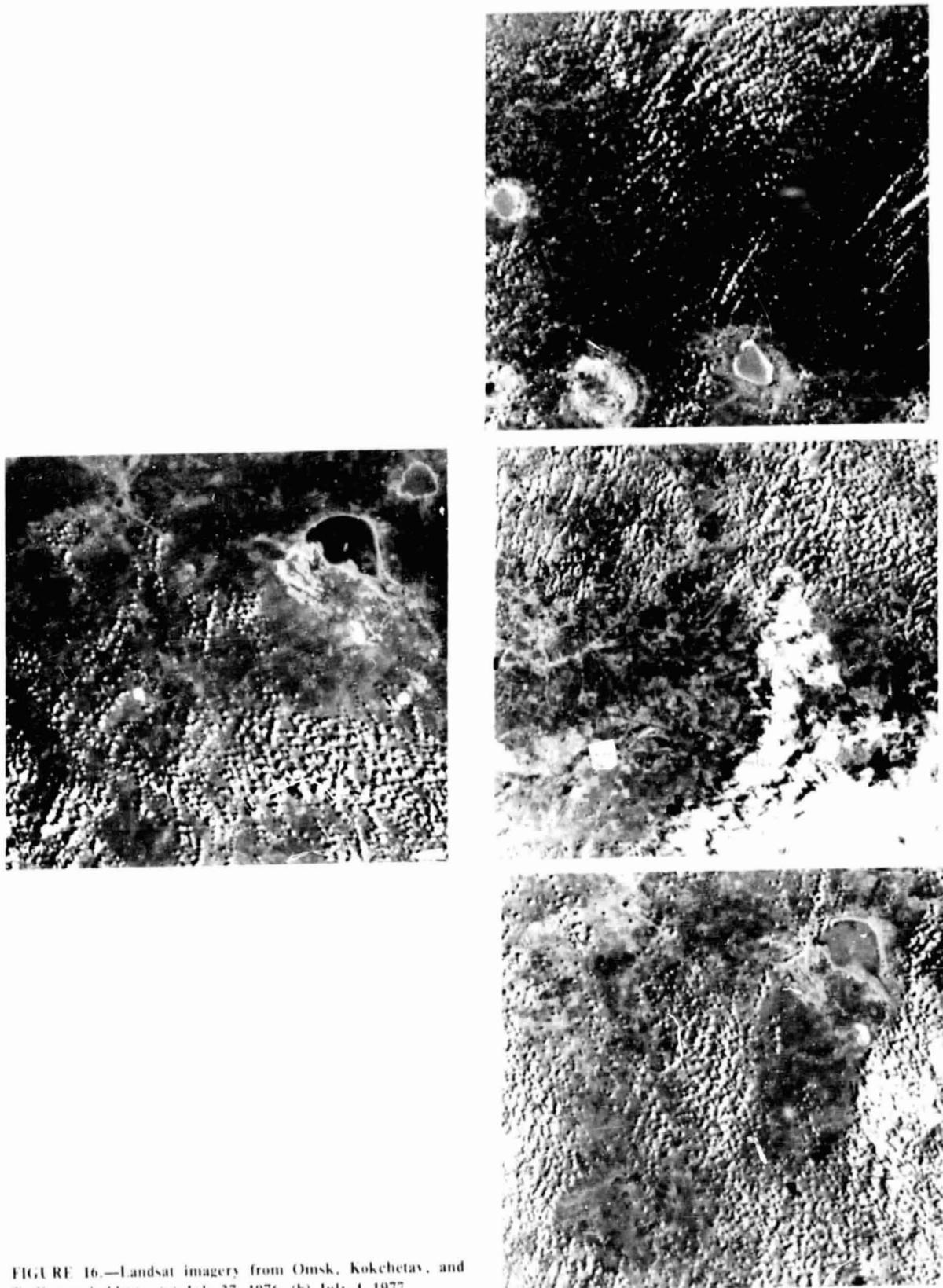


FIGURE 16.—Landsat imagery from Omsk, Kokchetav, and Tselinograd oblasts. (a) July 27, 1976. (b) July 4, 1977.

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

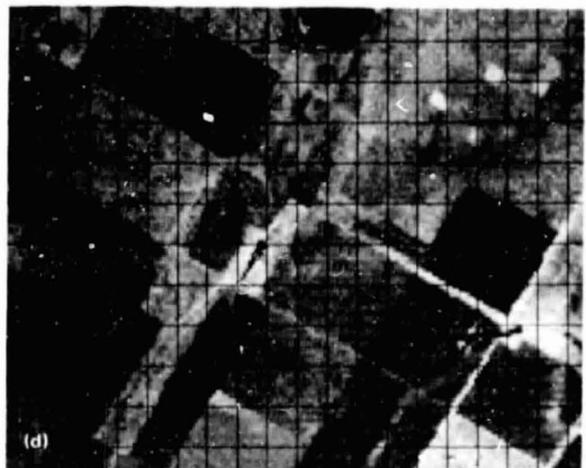
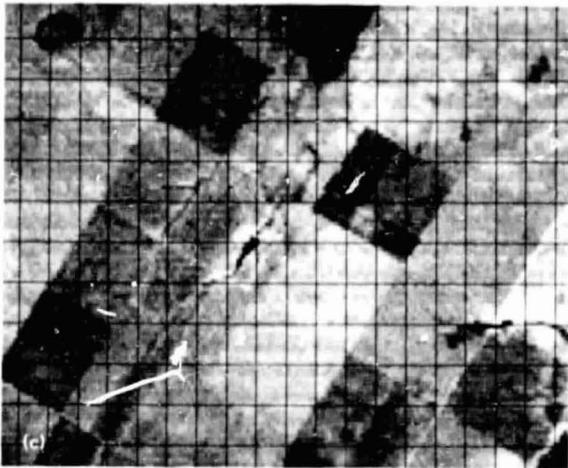
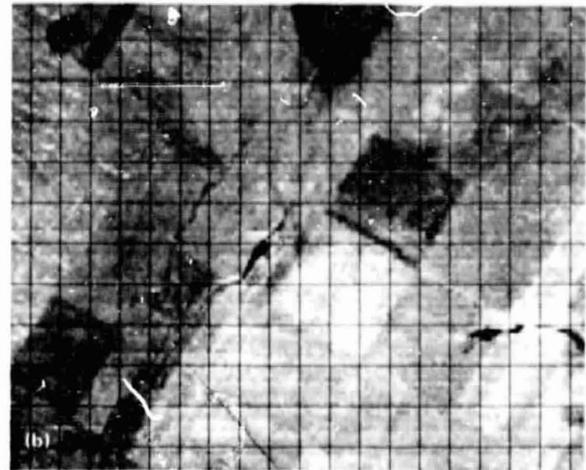
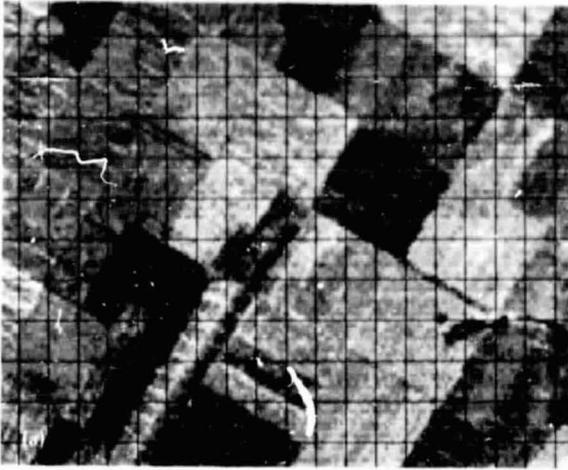


FIGURE 17.—Landsat 5- by 6-nautical-mile segment in southern Kustanay oblast. (a) April 29, 1977. (b) June 4, 1977. (c) July 28, 1977. (d) August 1, 1976.

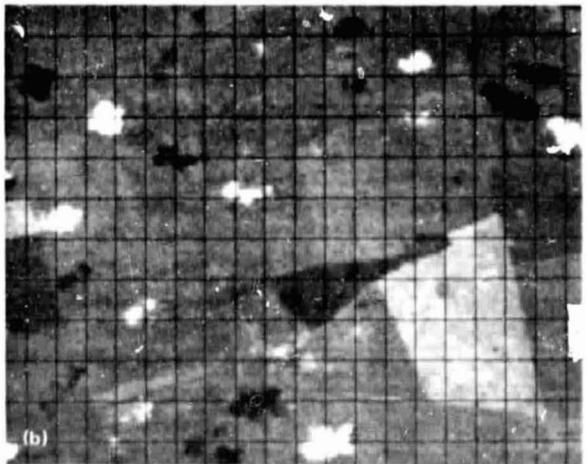
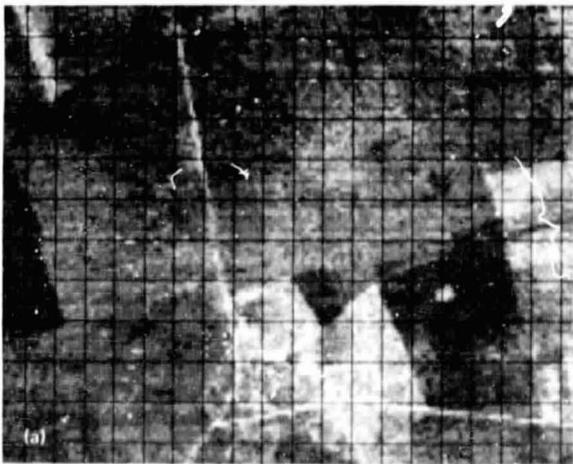


FIGURE 18.—Landsat 5- by 6-nautical-mile segment in northeastern Kustanay oblast. (a) July 27, 1977. (b) August 14, 1977.

formation is available from the meteorological and Landsat data which can be used to infer likely crop condition. The experience in the U.S. Great Plains indicates that such inferences truly reflect the actual condition of the wheat growing there. Such inferences do not follow directly from the data, however, and considerable agronomic insight and ancillary data are required to make them meaningful and reliable.

The application of Landsat digital data to crop condition assessment through development of the GIN represents a major application of that information source for a purpose other than crop identification and acreage mensuration. It has a distinct advantage over the meteorological data in that it provides continuous spatial coverage, whereas the weather observations are only samples at discrete points. However, the use of the GIN also requires care to avoid confounding crop conditions with crop phenological development. The GIN, meteorological data, crop development models, and ancillary historical data comprise a powerful combination of information sources to be exploited for making qualitative

assessments of crop vigor. The LACIE experience has focused primarily on qualitative assessments of moisture stress and temperature effects, but the information sources can be extended to inferring the presence or absence of other detrimental crop influences such as disease, insects, or extreme desiccation.

The data sources available to make assessments of crop conditions offer varying timeliness advantages but these can be complementary to each other. The meteorological data acquired each day provide a definition of environmental conditions at specific locations and an early warning when extremes are exceeded. The Landsat data provide a measure of the plants' integrated response to those conditions over a period of several days. The present interval of 18 days between passes of each Landsat satellite over a point on the Earth may not be timely enough to provide rapid verification and define the areal extent of adverse conditions, however, thus establishing a need for Landsat data acquisition at shorter intervals of 9 or 6 days.

Application of Landsat Digital Data for Monitoring Drought

D. R. Thompson^a and O. A. Wehmanen^b

SUMMARY

A technique utilizing transformed Landsat digital data for detection of agricultural vegetative water stress was developed during the 1976 South Dakota drought. The procedure was expanded to the U.S. Great Plains during 1977 to evaluate the technique for detecting and monitoring vegetative water stress over large areas. This technique, the Green Index Number, used Landsat digital data from 5- by 6-nautical-mile sampling frames (segments) to indicate when the vegetation within the segment was undergoing stress. Segments were classified as either moisture stressed or normal using remote-sensing techniques combined with a knowledge of the crop condition. The remote-sensing-based information was compared to a weekly ground-based index (the Crop Moisture Index) provided by the U.S. Department of Commerce. This comparison demonstrated good agreement between the 18-day remote-sensing technique and the weekly ground-based data. The procedure developed over a small geographic area (South Dakota) for detecting moisture stress was adapted to a larger geographic region (the U.S. Great Plains).

INTRODUCTION

Landsat color-infrared images were used in the LACIE during the 1976 droughts in the U.S. Great Plains to determine the areal extent of these droughts (refs. 1 and 2). The use of Landsat images for drought monitoring is dependent on the subjective judgment of an analyst-interpreter in deciding

that a region is or is not drought affected. During the analysis of the U.S. Southern Great Plains drought, studies were started using Landsat digital data from LACIE sample segments for quantifying the subjective judgment of the analyst-interpreter (ref. 1). During the drought in the U.S. Northern Great Plains, a technique utilizing transformed Landsat digital data for detection of agricultural vegetative water stress was developed (refs. 2 and 3). The procedure, which was developed over a small geographic region, was expanded to selected LACIE sample segments throughout the U.S. Great Plains during the 1977 crop year. The remote-sensing technique, the Green Index Number (GIN), was compared to a weekly ground-based index, the Crop Moisture Index, provided by the U.S. Department of Commerce. This paper presents the approaches used for and the results from the GIN monitoring program.

APPROACH

The GIN concept for detecting and monitoring drought uses Landsat multispectral scanner (MSS) values for LACIE sample segments (figs. 1 and 2) and a knowledge of the wheat growth stage at the time of the Landsat acquisition.

Using ideas presented by Kauth and Thomas (ref. 4), a screening number, the GIN, was developed during the 1976 drought in South Dakota. The GIN is a value designated to summarize the condition of vegetation within a sampling frame. It is based on the ability to detect the area of growing vegetation using all four Landsat bands and to measure the area within the region. In approximate terms, the GIN is the percentage of land in an area with a "healthy" cover of vegetation.

The procedure developed during the 1976 South Dakota drought was different from the procedure applied to the selected subset of LACIE sample seg-

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Inc., Systems and Services Division, Houston, Texas.

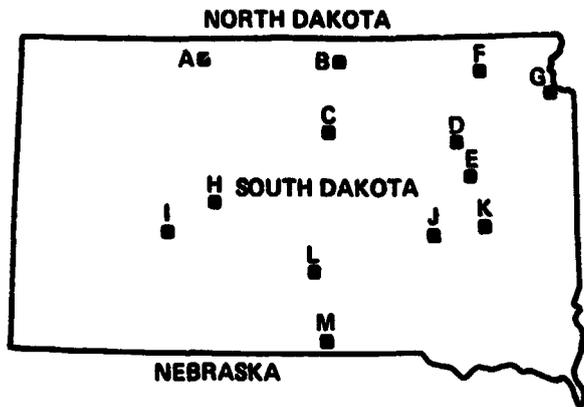


FIGURE 1.—Map of South Dakota showing locations of LACIE sample segments.

ments throughout the U.S. Great Plains, which were used to test the adaptability of the procedure developed over a small geographic area (South Dakota) to a larger geographic area (the U.S. Great Plains). The final GIN value is essentially the same value if computed by the two different methods. Both the original method and the present method of computing the GIN will be presented.

1976 GIN CALCULATION

The GIN is defined as follows. First, the data in the segment acquisition are summarized by clustering using the Iterative Self-Organizing Clustering System (ISOCLS) algorithm as implemented on the Earth Resources Interactive Processing System (ERIPS) on the special-purpose processor. (This parallel processor clusters a segment in approximately 30 seconds.) The clustering procedure summarizes the segment in 20 or fewer cluster means in the four Landsat channels. The count of picture elements (pixels) belonging to each cluster is also calculated. Each mean vector x^i is then transformed by

$$y^i = Ax^i + b^i \quad (1)$$

where

$y^i = \begin{bmatrix} y_1^i \\ y_2^i \\ y_3^i \\ y_4^i \end{bmatrix}$ a vector representing the Landsat-1 version of the Kauth-Thomas transformation of x^i (6); the subscript number indicates the Landsat channel, and the superscript is the cluster number.

$$A = \begin{bmatrix} 0.4326 & 0.6325 & 0.5857 & 0.2641 \\ -0.2897 & -0.5620 & 0.5995 & 0.4907 \\ 0.8242 & 0.5329 & -0.0502 & 0.1850 \\ 0.2229 & 0.0125 & -0.5431 & 0.8094 \end{bmatrix}$$

$$b^i = (0.45, -1.50, 10.61, 2.22)$$



FIGURE 2.—Map of U.S. Great Plains showing locations of LACIE sample segments.

Each vector is inspected automatically, and any vector having values unreasonable for agricultural data is discarded using the following procedure.

A cluster y^i is accepted as good only if

$$\left. \begin{array}{l} 30 < y_1^i < 110 \\ -10 < y_2^i \\ -10 < y_3^i \\ -10 < y_4^i < 10 \end{array} \right\} \quad (2)$$

The greenness level m of the soil line then is estimated by the minimum second-channel value y_2^i for acceptable clusters. That is,

$$m = \text{MIN } y_2^i \quad (3)$$

i is good

Then the green number g^i is computed for each cluster by

$$g^i = y_2^i - m \quad (4)$$

The value of g^i should be a good measure of the green vegetation present on the pixels in cluster i . Experience from test sites with spectral plots in the brightness-greenness plane and the corresponding imagery led to the following assumptions.

1. $g^i = 0$ indicates bare soil.
2. $g^i = 5$ indicates a trace of vegetation.
3. $g^i = 15$ indicates good cover of vegetation.

Operating on these assumptions, level 15 was chosen and GIN was defined to be the percentage of pixels in the entire image within clusters having green numbers greater than 15. This value was considered somewhat unstable for consecutive-day data where, for example, a cluster would slip from 15.1 one day to 14.9 the next day; therefore, cubic weighting was added to smooth this calculation in the following manner.

Let a cluster have green number g^i and contain n pixels. Define the weighting factor w by

$$\begin{aligned} w &= 0 \text{ if } g^i = 11 \\ w &= 1 \text{ if } g^i = 17 \end{aligned} \quad (5)$$

Otherwise,

$$w = \frac{1}{2} + \left[\frac{g^i - 14}{4} \right] \left[1 - \frac{(g^i - 14)^2}{27} \right]$$

The cluster is counted as having $w \cdot n$ pixels with green numbers greater than 15. This curve makes a smooth transition from full counting to not counting as the green number decreases.

The GIN then is an estimate of the percentage of pixels in a Landsat scene having green numbers high enough (≥ 15) to indicate full cover of green vegetation. It is computed using only Landsat data. A sample spectral plot of green numbers versus brightness is given in figure 3.

1977 GIN CALCULATION

The GIN is computed on a LACIE sample segment, an area (5 by 6 nautical miles) with 22 932

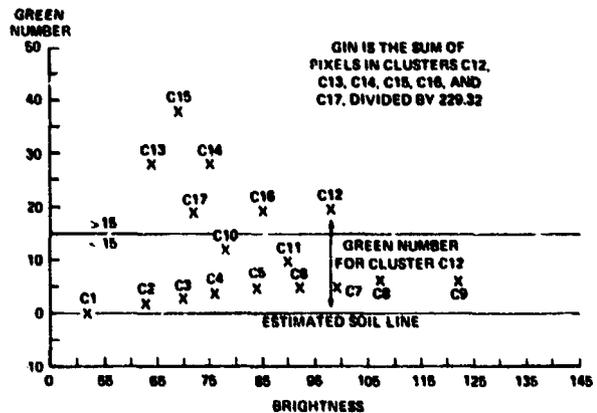


FIGURE 3.—Sample spectral plot of cluster statistics.

Landsat image elements or pixels. The data are processed using an automated screening procedure that rejects pixels with values which are unreasonable for agricultural area (because of clouds, water, or bad data). This procedure was established empirically after inspecting many LACIE segments. The procedure for computing GIN is defined as follows.

For an observation X , where

$$X = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix}$$

Z is computed, where

$$Z = \begin{pmatrix} Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \end{pmatrix}$$

$$Z = RX \quad (6)$$

where

$$R = \begin{bmatrix} 0.433 & 0.632 & 0.586 & 0.264 \\ -0.290 & -0.562 & .600 & .491 \\ -0.824 & .533 & -0.050 & .185 \\ .223 & .012 & -0.543 & .809 \end{bmatrix}$$

Each vector is inspected automatically, and any vector having values unreasonable for agricultural data (because of clouds, water, or bad data) is discarded using the following procedure.

A pixel is accepted as good only if

$$\begin{aligned} Z_1 &< 100 \\ -8 &< Z_2 \\ -19 &< Z_3 \\ -5 &< Z_4 < 15 \end{aligned} \quad (7)$$

Once the screening has been performed, the histogram of the value Z_2 , truncated to integer, is accumulated for good pixels. Z_2 is defined as the greenness channel in the Kauth-Thomas transformation and is a weighted difference between the spectral values in the infrared and visible channels. Bare soil has a low greenness level, which changes with haze level and sample segment location. The soil greenness s_1 for each segment is estimated to be the greenness of the pixel that is greener than only 228 (approximately 1 percent) of the other good pixels. The greenness is zero; thus, the green number g for a pixel is $g = Z_2 - s_1$. The green number contains information about green vegetation. To compute GIN, the pixels with $g \geq 15$ are counted, divided by 22 932 (the number of pixels in a scene), and multiplied by 100 to obtain the percentage. The level 15 was observed empirically to represent healthy green agricultural vegetation. The GIN is an estimate of the percentage of pixels in a Landsat 5- by 6-nautical-mile scene having green numbers high enough (≥ 15) to indicate full cover of green vegetation. It is computed using only Landsat data.

It was determined during the 1976 South Dakota drought that a plot of GIN versus acquisition time for a normal, predominantly wheat segment should follow a predetermined curve (fig. 4). If an observed point for a segment fell into the shaded region, the segment was classified as drought affected. The bounds for the shaded region were defined empirically with t defined as the approximate spring emergence in days. For different areas or years, the shaded area is moved from side to side to match the green-up curve. The bounds for the shaded region are such that a normal agricultural segment greens up at a rate greater than 0.5 percent per day for 40 days and exceeds a GIN value of 20 for 30 days. After this time, the wheat in a region has completed grain fill.

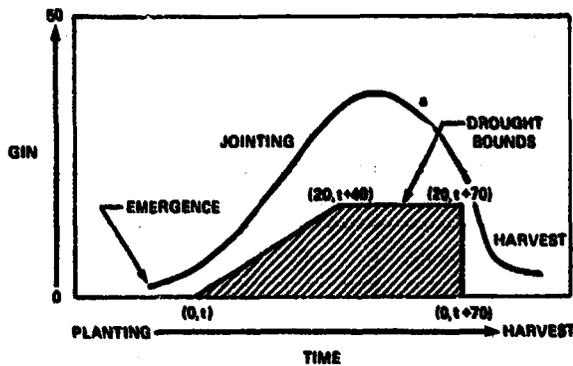


FIGURE 4.—Plot of GIN versus time for a normal, predominantly wheat segment.

ing and is harvested and the greenness of the region may decrease. Each acquisition for a segment was plotted and the segment was classified as moisture stressed or not moisture stressed. For comparison, these segments were also classified using the Crop Moisture Index (CMI) for Crop Reporting Districts (CRD's). The remote-sensing classification procedure of GIN was evaluated against the CMI, which measures the degree to which moisture requirements of growing crops were met during the previous week. The index is computed from average weekly values of temperature and precipitation. Along with previous soil moisture condition and current rainfall, the temperature and precipitation values are used to calculate the actual moisture loss. If the potential moisture demand or potential evapotranspiration exceeds available moisture supplies, actual evapotranspiration is reduced and the CMI gives a negative value. However, if moisture meets or exceeds demand, the index is positive. The CMI represents the average conditions over a several-county region (CRD); so local moisture conditions may vary because of differences in rainfall distribution or soil types. The specific type of agriculture is not considered in the CMI, but it assumes a water-use curve typical of the leaf area index of the crops which predominate in the region. A CRD was classified as drought affected if its CMI fell below -0.5 for 2 consecutive weeks. Both classifications were restricted to similar time frames. It was possible for a segment to start normal and then undergo moisture stress. If, in any one instance, the GIN classification did not agree with the CMI, the GIN was considered as not agreeing with the CMI (tables I and II).

TABLE I.—Results of GIN and CMI Classifications^a

Segment	1975		1976	
	GIN	CMI	GIN	CMI
A	W	W	—	D
B	—	W	W	D
C	W	W	D	D
D	W	W	D	D
E	W	W	D	D
F	W	W	W	D
G	W	W	W	D
H	W	W	D	W
I	W	W	—	W
J	W	W	D	D
K	W	W	D	D
L	D	D	D	D
M	—	D	W	D

^aD = drought conditions; W = normal conditions; and — = no data.

RESULTS

South Dakota, 1976

The data used in this study consisted of all LACIE segments in South Dakota which had at least 5 percent wheat as measured by the LACIE Classification and Mensuration Subsystem (CAMS) in the 1976 growing season. This definition yields 17 segments with 34 possible classifications or segment years. (A segment year is defined as an observation of one segment for a growing season.) Of the 31 classifications, 12 had either insufficient data during the growing season or data that were inaccessible for other reasons. The final data set contained 22 segment years for 13 LACIE segments (fig. 1 and table I). The contingency table (table III), which applies the two classification methods to the 22 good segment years, shows that the classifications based on the CMI and GIN are related. It was concluded that the GIN observations at appropriate phenological stages are detecting moisture through crop responses.

An inspection of the five disagreements on the classification results (table I) disclosed that on two segments, the GIN algorithm was confused by a lake. Three of the segments were on the edge of their

TABLE II.—Results of GIN and CMI Classifications^a

Segment	1977		1976	
	GIN	CMI	GIN	CMI
1560	W	W	D	W
1605	W	W	W	W
1388	W	W	—	—
1521	W	W	W	D
1595	W	W	W	W
1596	W	W	W	W
1056	W	W	D	D
1084	W	W	D	D
1015	D	W	D	D
1020	W	W	D	D
1642	W	W	W	D
1644	W	W	W	D
1503	W	D	D	D
1592	W	W	W	W
1557	D	D	W	W
1260	W	W	D	D
1694	W	W	D	D
1670	W	W	D	D
1227	W	W	W	W
1231	D	W	D	W
1233	W	W	W	W
1238	W	W	W	W
1166	W	W	W	W
1264	W	W	D	D
1047	—	—	D	D
1048	W	W	D	D
1008	W	W	D	D
1010	W	W	D	D
1857	W	W	W	W
1861	W	W	D	W
1645	W	W	W	D
1884	W	W	W	W
1887	W	W	W	W
1892	D	D	W	W
1601	D	D	W	W
1538	D	D	W	W

^aD = moisture-stressed conditions, W = normal conditions, and — = no data.

TABLE III.—Contingency Table of GIN and CMI Classification Methods^a

		CMI		
		Normal	Dry	
GIN	Normal	10	4	14
	Dry	1	7	8
		11	11	22

^a $\chi^2 = 7.07$ with 1 degree of freedom; $\hat{p} = 0.9918$ = level of significance.

CRD's, and the CMI classifications may not have reflected the actual conditions in these segments. One segment was located on the eastern edge of the CRD where heavy rains occurred at the weather station located in the Black Hills (in the western part of the CRD), causing a possible incorrect condition to be reflected by the CMI.

Examples of the segment classification procedure are shown in figures 5 and 6. The GIN indicates that 1975 was normal for the entire crop season for segment J (fig. 5). In 1976, the GIN indicated that by May 24 there was moisture stress in segment J. This indicates that the GIN detected vegetation moisture stress at the same time as the CMI. Segment L (fig. 6) experienced drought as indicated by both the GIN and the CMI during the 1975 crop year. In 1976, the GIN indicated moisture stress on May 26, which was confirmed by the CMI.

U.S. Great Plains, 1977

The data used in this study consisted of 36 LACIE segments located throughout the U.S. Great Plains wheat growing region (fig. 2). These segments contained at least 5 percent agricultural cropland. The final data set consisted of 70 segment years for the 36 LACIE segments (fig. 2).

The final data set of 70 segment years shows good agreement between the remote-sensing (GIN) classification and the CMI classification (table II). The contingency table (table IV), which compares

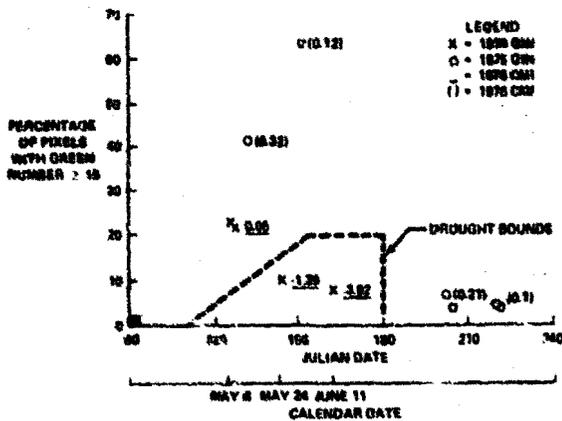


FIGURE 5.—Graphic plot of GIN versus time with CMI values for segment J.

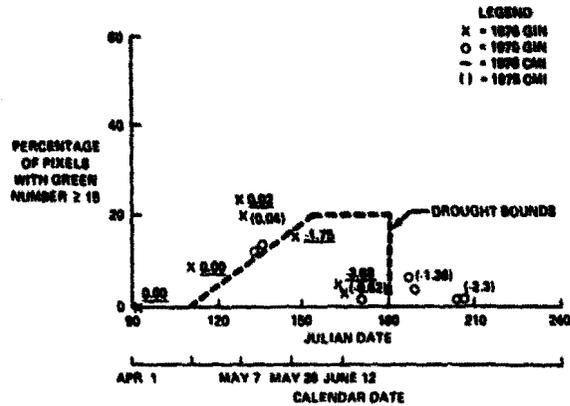


FIGURE 6.—Graphic plot of GIN versus time with CMI values for segment L.

TABLE IV.—Contingency Table of GIN and CMI Classification Methods Throughout the USGP^a

		CMI		
		Normal	Stressed	
GIN	Normal	43	5	48
	Stressed	6	16	22
		49	21	70

^a $\chi^2 = 27.89$ with 1 degree of freedom; $\hat{p} = 0.999$ = level of significance

the two classification methods, shows that the classifications based on the CMI and GIN are the same 85 percent of the time. It was concluded that the GIN is detecting moisture through crop responses and that this procedure, which was developed over a small region, is extendable to larger areas.

An inspection of the 11 disagreements on the classification results (table II) disclosed that the soil types at segment locations related to 5 of the disagreements have different water-holding capacities

than those used in the CMI model. Also, rainfall patterns produced amounts at the segment location which differed from the amounts recorded at the weather stations used in computing the CMI. This is reflected also in the other six disagreements, which occurred in segments that were located on the edge of the CRD; thus, the CMI does not necessarily represent the conditions that existed at the segment location.

CONCLUSIONS

A technique was developed, using Landsat digital data from 5- by 6-nautical-mile sample segments, which indicates when agricultural vegetation is undergoing moisture stress. A relation between this technique, which utilizes remote sensing, and a ground-based criterion (the CMI) has been shown. The remote-sensing procedure was shown to be expandable to a larger geographic area and repeatable for different areas and years. Thus, in areas of the world where ground truth is not available or reliable, it is possible to detect and determine the areal extent of moisture stress using Landsat data in an automatic mode. The GIN is now automatically calculated for all LACIE segments as they are loaded into the data base. The procedure has been implemented on the U.S. Department of Agriculture (USDA) system.

While this procedure was developed for detecting

and monitoring moisture stress, variations of the idea were implemented in on-line CAMS processing. These included the green number and brightness for each of the 209 grid interactions, green number versus brightness scatter plots, and trajectory plots of green number and brightness. Thus, a procedure which was developed for a particular application has had a key role in helping to solve other LACIE problems, such as aiding the analyst in verifying the consistency of dot labels (Procedure 1) and separation of wheat from other small grains.

REFERENCES

1. Thompson, D. R.: Results of LACIE Integrated Drought Analysis (Southern Great Plains Drought, 1975-76). LACIE-00424, NASA Johnson Space Center, Houston, Texas, 1976.
2. Thompson, D. R.: Results of Drought Analysis (South Dakota Drought 1976). LACIE-00437, NASA Johnson Space Center, Houston, Texas, Sept. 1976.
3. Thompson, D. R.; and Wehmanen, O. A.: The Use of Landsat Digital Data to Detect and Monitor Vegetation Water Deficiencies. Eleventh International Symposium on Remote Sensing of Environment (Ann Arbor, Mich.), Apr. 1977, pp. 925-931.
4. Kauth, R. J.; and Thomas, G. S.: The Tasseled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue Univ. (West Lafayette, Ind.), IEEE cat. 76, CH1103-1-MPRSD, July 1976.

LACIE Area, Yield, and Production Estimate Characteristics: U.S. Great Plains

Duane L. Marquis^a

OVERVIEW

The objective of the LACIE is to estimate production of wheat on a country-by-country basis. LACIE was designed to meet U.S. Department of Agriculture (USDA) needs in areas where ground-truth information is not readily available. However, in order to test the design (to determine the accuracy and reliability), an area where comparison information was available was chosen. This area was the nine states of the U.S. Great Plains (Colorado, Kansas, Minnesota, Montana, Nebraska, North Dakota, Oklahoma, South Dakota, and Texas). LACIE was not designed to improve the accuracy of the U.S. crop reports.

In 1974, the U.S. Great Plains (USGP) accounted for over 64 percent of the U.S. winter wheat area (56 percent of production), over 93 percent of the U.S. spring wheat area (89 percent of production), and 73 percent of all wheat area in the U.S. (64 percent of all wheat production). By 1977, these percentages had all increased. These nine states represented the wide range of soil types, climatic conditions, topography, cultural practices (such as crop rotation, strip cropping, irrigation, summer fallowing), and crop varieties which was needed to test the adequacy of the design (accuracy of the technology). Of the nine states that comprise the Great Plains region, five states (Colorado, Kansas, Nebraska, Oklahoma, and Texas) are almost entirely producers of winter wheat; two states (Minnesota and North Dakota) are producers of spring wheat; and two states (Montana and South Dakota) produce significant amounts of both winter and spring wheat. As the LACIE technology changed, comparison information had to be readily available to assess the capability of the new technology.

^aNASA Johnson Space Center, Houston, Texas.

To estimate wheat production on a country basis, the country is subdivided into areas (strata) where yield and the prevalence of wheat planted are relatively uniform. Yield and the areal extent of wheat within each stratum are estimated by independent methods and then multiplied together to obtain production at the stratum level. The production estimates in each stratum are then added to obtain the production estimates at other geographical or political levels. In addition, area and yield are aggregated to determine wheat area and yield at other hierarchical levels within the country.

The LACIE was designed as a three-phase operation to cover three global crop seasons. The U.S. Great Plains has played a significant role in all three LACIE phases. The next three sections will describe each phase in more detail as it pertains to the U.S. Great Plains. Items discussed include scope, sampling strategy, Landsat and yield data, estimates reported, accuracy of the estimates, and technical issues raised in each phase.

PHASE I

Scope

Phase I of LACIE had as its major objective the development and testing of a system that used Landsat data as the primary input for estimating wheat area in selected regions of the United States. Phase I also was devoted to yield model development and to determining the feasibility of estimating production in the United States. During the Phase I developmental period, techniques and operational procedures were constantly being evaluated. Based on such evaluations, several procedural changes were implemented.

Sampling

Sample segments were allocated to the county level based on the total area used for wheat production as reported in the 1969 Census of Agriculture. On the basis of these 1969 data, counties within the Great Plains were classified into three categories:

1. Group I counties—Those counties that produced sufficient wheat in 1969 to justify the allocation of one or more sample segments.

2. Group II counties—Those counties for which the historical wheat area in an individual county did not justify the allocation of a sample segment; however, when counties in the same Crop Reporting District (CRD) were combined, the allocation of one or more sample segments was justified. Sample segments in Group II counties were allocated on their probability proportional to size and were referred to as PPS segments. In PPS sampling, the probability that a particular sample unit will be selected is proportional to some measure of size associated with it.

3. Group III counties—Those counties in which historical wheat area did not justify allocation of any sample segments.

The above sampling procedure allocated 411 sample segments, 5 by 6 nautical miles in size, to the nine Great Plains states—359 Group I segments and 52 Group II segments (table I). (For a more detailed discussion of the sampling methodology employed, see LACIE C00200, Vol. IV, Rev. C, Oct. 1977.)

TABLE I.—Number of Sample Segments Allocated by State, by Crop Type, and by Group; U.S. Great Plains, Phase I

State	Number of segments allocated					
	Winter wheat			Spring wheat		
	Group I	Group II	Total	Group I	Group II	Total
Colorado	27	5	32			
Kansas	73	11	84			
Minnesota				8	5	13
Montana	14	2	16	43	1	44
Nebraska	28	7	35			
North Dakota				65	0	65
Oklahoma	33	7	40			
South Dakota	9	1	10	21	2	23
Texas	38	11	49			
Total	222	44	266	137	8	145

Landsat Data

The collection of Landsat data for each segment was scheduled to coincide with the major wheat growth stages. Due primarily to cloud cover, the number of usable images was substantially reduced. On the average, 2.3 acquisitions per segment were obtained.

At the beginning of Phase I, acquisitions were assigned to average or nominal biowindows based on historical wheat growth stage (crop calendar) data. Biowindow 1 included Robertson crop growth stages 1 and 2 (planting to jointing); biowindow 2 included stage 3 (jointing to heading); biowindow 3 included stage 4 (heading to soft dough); and biowindow 4 included stages 5, 6, and 7 (soft dough through harvest). When the crop calendar models based on actual development became operative in May 1975, biowindows were updated at the CRD level; a second updating of the biowindows was performed at the end of the season based on actual crop development. Actual biowindow definitions remained constant for biowindows 2 through 4, but biowindow 1 was adjusted to eliminate growth stages prior to Robertson stage 2.3. Growth stage 2.3 corresponds to the minimum plant cover for detection.

In each segment for which Landsat data were acquired, the Classification and Mensuration Subsystem (CAMS) estimated the proportion of area within the segment that was devoted to wheat or small grains production. These estimates were internally evaluated by CAMS for classification accuracy and transmitted to the Crop Assessment Subsystem (CAS) where they were used to estimate wheat acreage at the stratum (CRD), the zone (state), and the region (Great Plains) levels.

After all Landsat data for Phase I had been processed, procedures used in CAMS were reviewed and ultimately revised based on the experience obtained during Phase I operations. A decision was made to use the "new CAMS procedures" to take a retrospective look at the Landsat data acquired in Phase I. These data were reprocessed and analyzed using a multitemporal approach to provide the "best possible estimate" of wheat proportion in all segments for which data were available. The proportion estimates were set up as the CAMS rework data base, and aggregations were performed. The results of these aggregations are referred to as the "at-harvest" estimates. These estimates cannot be compared to the real-time estimates reported from April through August because in the earlier estimates (1) bare

ground was included as potential wheat and (2) Landsat data in biowindow 1 were acquired before wheat greened up enough to allow positive identification. Table II shows the number of segments allocated and the number of segments used by crop type and by state for the aggregation of the reworked data.

Estimates

Table III shows the results of the aggregation based on the CAMS rework data with PPS segments used in the aggregation plus the corresponding acreage estimates reported by the Crop Reporting Board of the Economics, Statistics, and Cooperatives Service (ESCS) (formerly the Statistical Reporting Service (SRS)) of the U.S. Department of

TABLE II.—Number of Sample Segments Allocated and Number Used for the CAMS Rework Aggregation, by State and by Type of Wheat

State	Segments allocated		CAMS rework segment	
	Group I	Group II	Group I	Group II
<i>Winter wheat</i>				
Colorado	27	5	21	3
Kansas	73	11	50	5
Nebraska	28	7	19	4
Oklahoma	33	7	23	6
Texas	38	11	23	5
Subtotal	199	41	136	23
<i>Spring wheat</i>				
Minnesota	8	5	5	4
North Dakota	65	0	42	0
Subtotal	73	5	47	4
<i>Mixed wheat</i>				
Montana	57	3	36	3
South Dakota	30	3	21	2
Subtotal	87	6	57	5
<i>Total wheat</i>				
Total segments	359	52	240	32

Agriculture. The five-state winter wheat area was within 1 percent of the USDA final estimate for 1975. (The five states are Colorado, Kansas, Nebraska, Oklahoma, and Texas. The seven winter wheat states include Montana and South Dakota with the five winter wheat states. The four spring wheat states are Minnesota, Montana, North Dakota, and South Dakota.) The Oklahoma estimate was

TABLE III.—LACIE At-Harvest Estimates With PPS Segments Included in the Aggregation Compared to 1975 Final Estimate of Harvested Acres by USDA

State	LACIE estimate, thousands of acres	CV, ^a percent	1975 final USDA estimate, ^b thousands of acres	Relative difference, ^c percent
<i>Winter wheat</i>				
Colorado	3 058	20.8	2 470	+19.2
Kansas	12 940	7.1	12 100	+6.5
Nebraska	2 657	28.0	3 070	-15.5
Oklahoma	6 906	11.2	6 700	+3.0
Texas	4 218	32.6	5 700	-35.1
Subtotal	29 779	7.0	30 040	-0.9
<i>Spring wheat</i>				
Minnesota	^d 2 150	15.7	2 787	-29.6
North Dakota	^d 5 853	14.8	10 090	-72.4
Subtotal	^d 8 003	11.6	12 877	-60.9
<i>Mixed wheat</i>				
Montana	^d 3 999	25.9	4 975	-24.4
South Dakota	^d 4 154	17.7	2 965	+28.6
Subtotal	^d 8 153	15.6	7 940	+2.6
<i>Total wheat</i>				
Total estimates	45 935	5.7	50 857	-10.7

^aCoefficient of variation = $\frac{\text{Sample standard error}}{\text{LACIE estimate}} \times 100$

^bCrop Production 1977 Annual Summary, Crop Reporting Board, ESCS, USDA, CrPr2-1(78), Jan. 16, 1978

^cRelative difference = $\frac{\text{LACIE} - \text{SRS}}{\text{LACIE}} \times 100$

^dThe CAMS rework resulted in winter wheat and spring small-grains proportion estimates. These estimates were adjusted (trained) at the segment level for oats, barley, and flax harvested in 1974

closest to the USDA state estimate at 3 percent above. Then followed Kansas at 6.5 percent above, Nebraska at 15.5 percent below, Colorado at 19.2 percent above, and Texas at 35.1 percent below.

For the mixed wheat states, LACIE's Montana estimate of total wheat was 24.4 percent below the USDA estimate, and the South Dakota estimate was 28.6 percent above. In the spring wheat states, the Minnesota estimate was 29.6 percent below, and the North Dakota estimate was 72.4 percent below the USDA estimate. Investigations into the causes of these errors are discussed in more detail in the next section.

Accuracy of the Estimates

In Phase I, the LACIE accuracy goal was applied at the national level. The LACIE accuracy goal, referred to as the 90/90 criterion, specified that the at-harvest wheat production estimator be within 10 percent of the true value 90 percent of the time. In order to assess whether the acreage estimates would support this criterion, it was assumed that area and yield estimators were unbiased and independent and that the coefficient of variation of the yield predictor was equal to that of the acreage estimator. Under these assumptions, the 90/90 criterion would be satisfied if the coefficient of variation of an acreage estimator, at the national level, was less than 4.3 percent.

The coefficient of variation of the acreage estimator at the U.S. Great Plains level was estimated to be 5.7 percent and was projected to be 3.7 percent at the national level. However, a significant difference of -10.7 percent was observed between the LACIE estimate for the U.S. Great Plains and the corresponding USDA estimate, indicating a negative bias. Since the projected coefficient of variation of 3.7 percent was less than the required 4.3 percent, some bias was tolerable, and it was inferred that the LACIE acreage estimator marginally supported the 90/90 accuracy goal. However, problems with sampling and classification encountered during Phase I indicated that improvements could and should be made before concluding that the LACIE acreage estimator met the 90/90 criterion.

A significant contributor to the wheat acreage underestimate at the Great Plains level was the underage (-72.4 percent) observed for North Dakota spring wheat (see table III). Analyses based on comparisons of segment wheat proportion estimates with

corresponding ground-observed proportions for 20 sample segments and with historical wheat proportions for the corresponding counties indicated that sampling was the major problem rather than classification. Additional segments were added in North Dakota to alleviate this problem in Phase II.

The comparisons of segment wheat proportion estimates with ground-observed wheat proportion estimates also indicated a difficulty in differentiating wheat from other related small grains. This required that wheat area estimates be obtained by ratioing small grains area estimates in accordance with the historic prevalence of these crops. Also, wheat identification was found to be more difficult in regions of marginal wheat production, small fields, or large amounts of confusion crops.

Technical Issues

There were technical problems which arose during Phase I. Those described in this section are the major ones affecting the results obtained in Phase I.

Wheat/small grains separation.—A major source of error found in Phase I was that spring wheat could not be reliably distinguished from other spring small grains, although spring small grains could be distinguished from other crops. For winter wheat, the major source of error appeared to be classification error in marginal areas, where confusion crops such as alfalfa were in abundance. The fact that wheat could not be consistently and accurately separated from other small grains leads to a second issue.

Wheat/small grains ratioing.—If total small grains proportion estimates were the products generated by CAMS for the spring and mixed wheat areas, then CAS was forced to develop wheat estimates from these small grains estimates. The procedure used was to "ratio" the spring small grains estimate to obtain a spring wheat estimate. This ratio was based on some historical proportion of spring wheat to spring small grains or to total small grains in the state. This ratioing technique introduced errors into the system since the current-year ratio of wheat to small grains is not likely to be the same as the ratio in some year past.

Group II sampling.—A review of the LACIE procedures in early August 1975 pinpointed the fact that a relatively large part of the positive bias (overestimation) observed in the acreage estimates made to that time was directly attributable to classification of wheat in the Group II (PPS) sample segments. These overestimates indicated a potential source of

error in the PPS approach. While such a sampling strategy is conceptually sound from a statistical point of view, questions were asked about the practicality (applicability) of the approach since CAMS tended to overestimate wheat area in these low-density wheat areas and the PPS approach requires an unbiased estimate.

The major inference drawn from experiences with the PPS segments in Phase I was that the PPS segments appeared to have characteristics (such as low percentage wheat) that made accurate classification more difficult than classification in Group I segments. It was also found that the aggregation logic was particularly sensitive to errors in acreage estimates for the PPS segments.

Early-season estimation.—Because an estimate of wheat production early in the crop year was considered especially valuable, it was a project concern to produce estimates as early as possible. During Phase I, an attempt was made to arrive at an area estimate using fall data, which showed little wheat emerged. The approach was to classify areas of seedbed preparation or bare soil as "potential wheat." However, fall plowing and seedbed preparation were conducted in many areas for purposes other than planting wheat, and thus the LACIE estimate was initially considerably higher than the USDA estimate. Other major causes of the high estimates, in addition to this "potential wheat" problem, seemed to be (1) cases in which wheat could not be separated from small grains and other crops and (2) cases in which classifications would be made with the estimate results in three overlapping classes (i.e., winter wheat, spring wheat, and total wheat).

Sampling.—Certain problems were found in sampling. One was placing samples in nonagricultural areas because of a lack of full-frame Landsat data to support the proper delineation of such areas. Another problem concerned the assumption that counties were relatively homogeneous. Actual experience did not support this assumption. Landsat data provided the basis for the delineation of areas to be sampled.

PHASE II

Scope

The Phase II scope of LACIE in the U.S. Great Plains "yardstick" region was to test the total system;

that is, to estimate area, yield, and production components, with emphasis on early-season estimates. The plan called for the generation of area, yield, and production estimates for the seven-state area from February through October, four-state estimates from July through October, and nine state estimates from July through October. However, insufficient Landsat data in Montana and South Dakota caused the February through May estimates for winter wheat to be made for the five states of the Southern Great Plains only. Spring wheat was not detectable until the mid-June Landsat pass. With the 30-day period from acquisition to receipt by CAS, the first spring wheat and total wheat estimates were delayed until August.

Sampling

The sampling strategy used was the same as that used in Phase I with the exception of the redelineation of agricultural and nonagricultural areas in North Dakota, which resulted in moving some segments. Also, as mentioned earlier, 20 sample segments were added in North Dakota. Thus, Phase II included 431 sample segments in the Great Plains region.

Landsat Data

In Phase II, all Landsat acquisitions were utilized. Table IV shows the number of segments allocated by state and by crop type plus the number of sample segments used in each of the CAS reports.

At the five-state level, of those segments allocated, the number of segments used ranged from over 54 percent for the February report to 97 percent for the end-of-season report. For the seven-state area, the number of segments used ranged from over 81 percent in June to 96.5 percent by the end of season. At the four-state level, usable acquisitions were obtained from 48.3 percent of those segments allocated in August to over 90 percent by the end of season, and for the nine states combined, the usable acquisitions ranged from 76.1 percent in August to 94.4 percent by the end of season. CAS Annual Report 03, December 15, 1976, contains tables showing the average percent wheat per segment used in each report, the average number of elapsed days from Landsat acquisition until receipt in CAS, and the distribution of segments used in each report by month of acquisition.

Figure 1 shows the number of classifications passed as wheat or small grains by monthly report for the nine states, the five states, and the seven states. All the classifications for the four spring wheat states were small grains. Most of the classifications were made for small grains, even for the five states.

Yield Data

Yield estimates were received in CAS on the fourth working day of the month. The yield models are documented in LACIE-00431 (June 1975). The models were developed at the state level and were run at the CRD level. Data bases for the CAS software used the individual CRD-level yield estimates.

The yield models were developed from a yield and climatic data base of approximately 45 years. The yield data were USDA yield per harvested acre, and the climatic data were National Oceanic and Atmospheric Administration (NOAA) monthly climatic division averages of precipitation and temperature. A piecewise linear curve was used to model the technology trend. A more detailed explanation of the models is contained in the document referred to above.

Estimates of Area, Yield, and Production

The discussion of the Phase II results refers to the estimates as revised in CAS Annual Report 03, December 15, 1976. The revised estimates resulted

TABLE IV.—Number of Sample Segments Allocated and Number of Segments Used in Each Report, by State and by Crop Type; U.S. Great Plains, Phase II

State	Number of segments allocated	Number of segments used by report										
		Feb.	Mar.	Apr.	May	June 8	June 29	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>												
Colorado	32	13	25	25	26	26	28	30	31	32	32	32
Kansas	84	43	61	62	70	75	77	78	78	81	81	81
Nebraska	35	13	21	22	27	30	31	32	32	33	33	33
Oklahoma	40	30	36	36	38	38	40	40	40	40	40	40
Texas	49	31	42	44	47	47	47	47	47	47	47	47
5-state total	240	130	185	189	208	216	223	227	228	233	233	233
Montana	38					10	17	21	22	35	35	36
South Dakota	10					8	9	9	9	9	9	9
2-state total	48					18	26	30	31	44	44	45
7-state total	288					234	249	257	259	277	277	278
<i>Spring wheat</i>												
Minnesota	13								10	10	11	11
Montana	22								14	19	20	20
North Dakota	85								31	67	79	79
South Dakota	23								14	16	19	19
4-state total	143								69	114	129	129
<i>Total wheat</i>												
9-state total	431								328	391	406	407

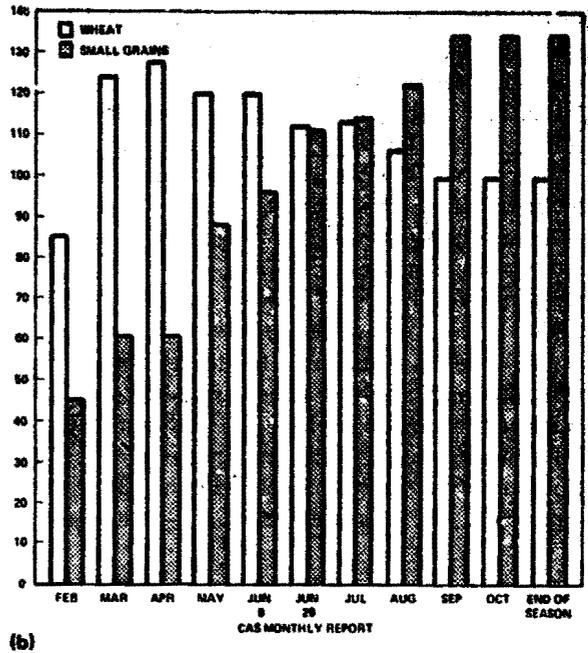
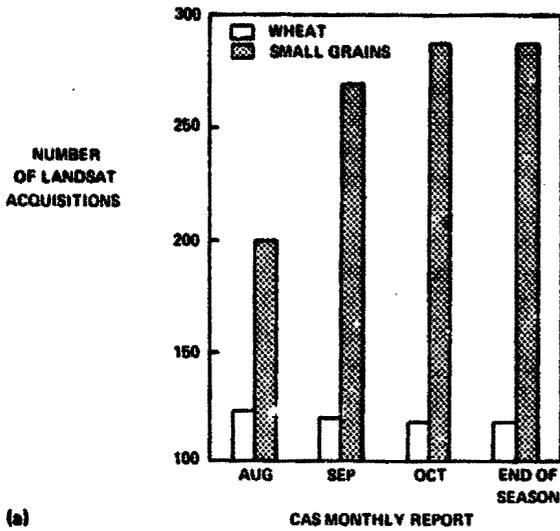
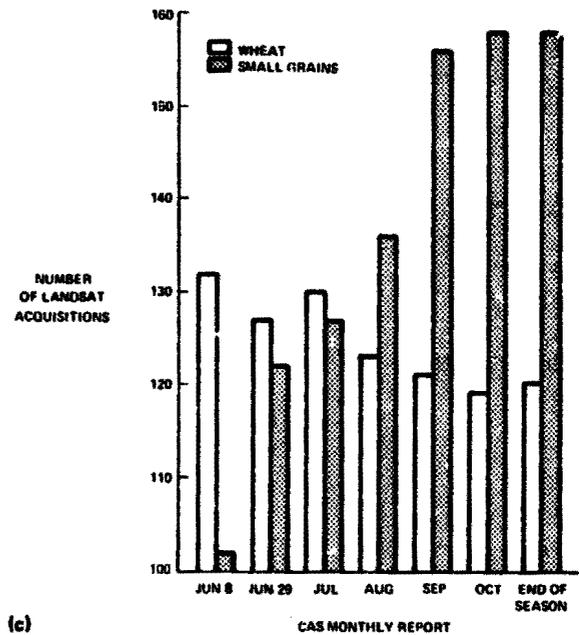


FIGURE 1.—Distribution of Landsat acquisitions by the number classified as wheat and the number classified as small grains for each CAS Monthly Report, Phase II. (a) U.S. Great Plains wheat, nine states; (b) U.S. Great Plains winter wheat, five states; (c) U.S. Great Plains winter wheat, seven states.

from the use of the CAS technology available at the end of Phase II to provide a consistent set of estimates. Known data errors were also corrected. Appendix A contains the Phase II estimates of area, yield, and production by state, by crop type, and by report. Tables A-I, A-II, and A-III contain the revised area, production, and yield estimates. Tables A-IV, A-V, and A-VI contain the corresponding coefficients of variation for the revised estimates. Tables A-VII, A-VIII, and A-IX contain the real-time LACIE estimates. However, the statistics reported in real time during Phase II were not correct and will not be presented here.

U.S. Southern Great Plains winter wheat.—Figure 2 shows the revised monthly LACIE estimates of area, production, and yield for the five states compared to the corresponding monthly USDA estimates. The area estimate in February was 22.7 million acres. The area estimates for Nebraska and Colorado were quite high when compared to historical data, and the estimates for Kansas, Oklahoma, and Texas were low by



(c)

the same comparison. The average estimated percent wheat per segment was 13.7 for the five states. With additional segments acquired, the April estimate was 21.8 million acres. This decrease occurred because the Colorado and Nebraska estimates dropped to

near their historical averages. The Kansas and Texas estimates increased somewhat. The average percent wheat per segment for the five states was 14.5 in April. The area estimate reached its peak at 26.7 million acres in late June. The average percent wheat per segment for Colorado, Nebraska, and Texas declined after this report, while Kansas and Oklahoma percent wheat averages increased. The end-of-season area estimate was 25.8 million acres compared to the final USDA estimate of 27.65 million acres (a -7-percent relative difference).

The production estimate for the five states was 626.1 million bushels in February (table A-II). The estimate declined to 564.1 million bushels because of both the area estimate changes discussed above and the reductions in yield model estimates. By the late June report, the estimate was 706.2 million bushels.

This increase resulted from large area increases in Kansas, Oklahoma, and Texas, plus yield increases in all states except Nebraska. The final LACIE estimate was 686.2 million bushels compared to the USDA final production estimate of 739.6 million bushels (a -7.8-percent relative difference).

The derived yield estimate for the five states was 27.6 bushels per acre in February. (Yield is derived by dividing the production estimate by the area estimate. This yield is a weighted average of all the yield model estimates.) Yield model estimates in Colorado and Kansas declined by nearly 2 bushels per acre, and the Nebraska estimate declined by 3 bushels per acre by May (table A-III). This resulted in a five-state yield estimate of 25.3 bushels per acre. By the end of the season, the Kansas, Nebraska, Oklahoma, and Texas yields increased, resulting in a five-state

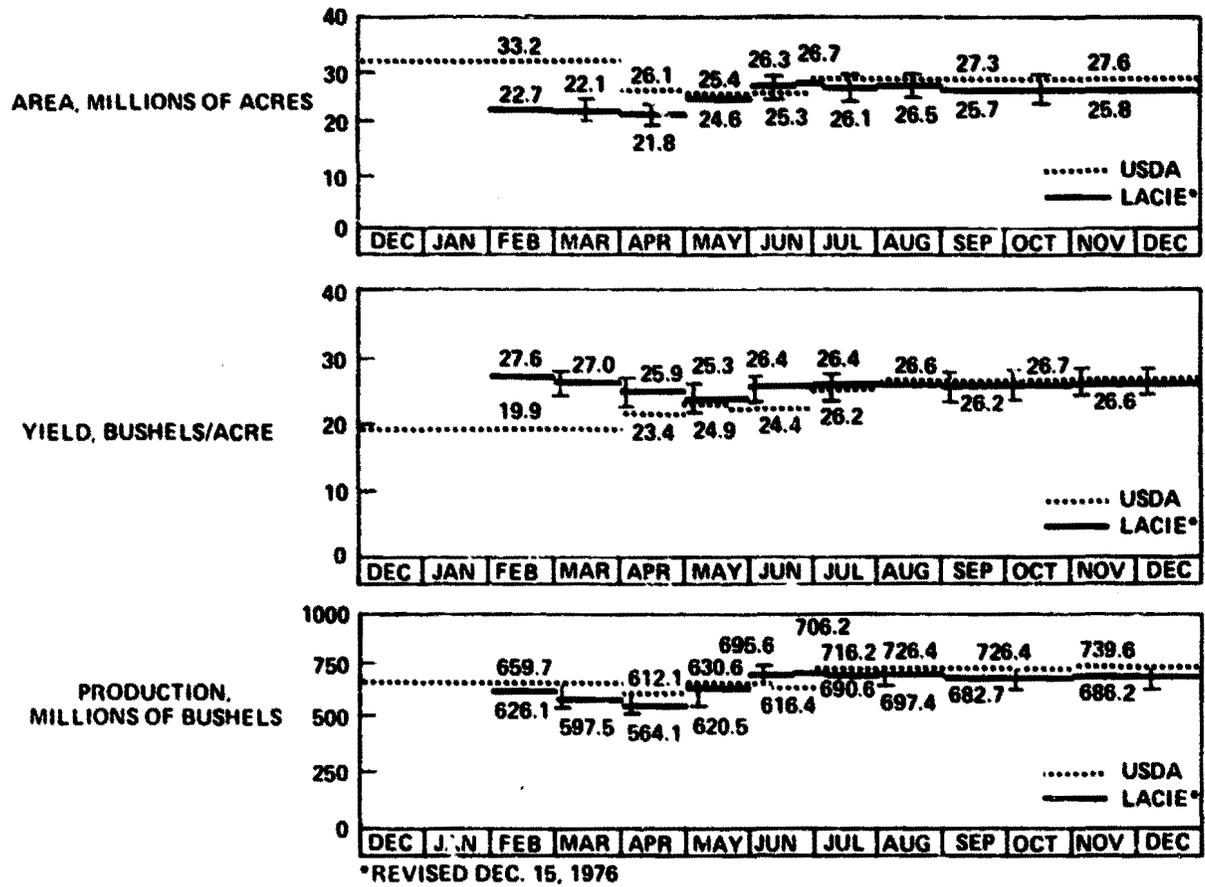


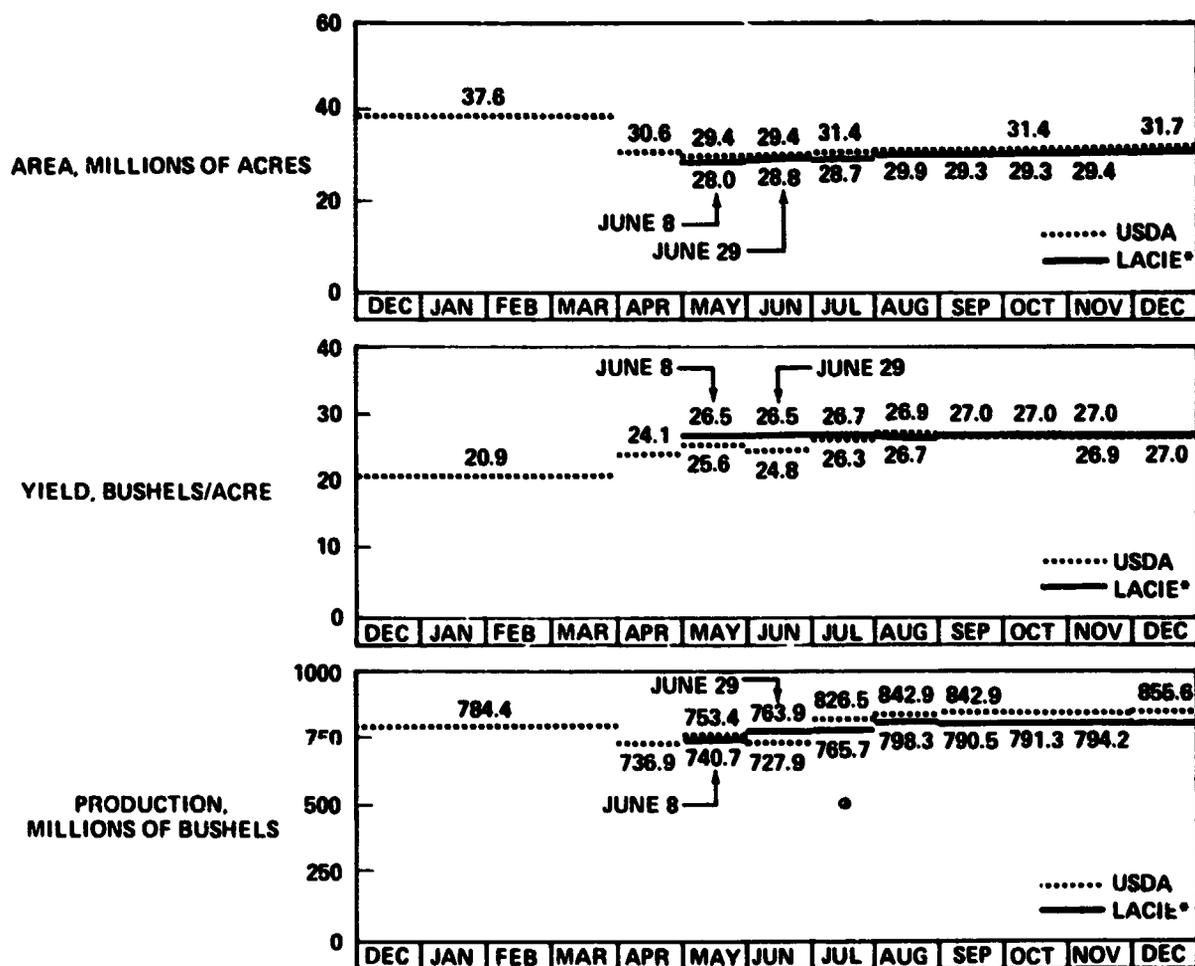
FIGURE 2.—Monthly comparison of LACIE and USDA estimates; Phase II; Southern Great Plains winter wheat states (Colorado, Kansas, Nebraska, Oklahoma, and Texas).

yield estimate of 26.6 bushels per acre. The final USDA estimate was 26.7 bushels per acre (a -0.4-percent relative difference).

U.S. Great Plains winter wheat.—The first estimates for the seven states were made in June. Figure 3 shows the revised LACIE area, production, and yield estimates compared to the USDA estimates. The June area estimate was 28 million acres. The estimates for Nebraska and South Dakota were overestimates, while the estimates for Oklahoma, Texas, and Montana were underestimates (table A-I). The estimate increased to 29.9 million acres in August due primarily to the area increases in Montana and

South Dakota. The final LACIE estimate was 29.4 million acres compared to the final USDA estimate of 31.7 million acres (a -7.8-percent relative difference).

The initial seven-state production estimate was 740.7 million bushels (table A-II). The estimate increased to 798.3 million bushels by August due primarily to area and yield estimate increases in Montana and South Dakota (tables A-I and A-III). The final estimate declined to 794.2 million bushels, due primarily to a decline in the Nebraska area estimate. The final USDA estimate was 855.6 million bushels (a -7.7-percent relative difference).



*REVISED DEC. 15, 1976

FIGURE 3.—Monthly comparison of LACIE and USDA estimates; Phase II; winter wheat, seven states (Colorado, Kansas, Nebraska, Oklahoma, Texas, Montana, and South Dakota).

The June derived yield estimate was 26.5 bushels per acre. The estimate increased to 27.0 bushels per acre by September due mainly to yield model estimate increases in Nebraska, South Dakota, and Montana. The final estimate was 27.0 bushels per acre, the same as the final USDA yield estimate.

U.S. Great Plains spring wheat.—The first spring wheat estimates made by LACIE were in August. Figure 4 shows the revised LACIE area, production, and yield estimates compared to the USDA estimates. The initial area estimate was 13.2 million acres. The estimates for Minnesota, Montana, and North Dakota were under their respective historical estimates (table A-1). By September, the area esti-

mate increased to 15.6 million acres due largely to increases in Minnesota and North Dakota. The average percent wheat per segment increased to 18.5 from 14.4 in August. The area estimate decreased in October due to a drop in the area estimate for Minnesota. The end-of-season estimate increased to 15.63 million acres due to area increases in Montana and North Dakota. The average percent wheat per segment was 19.5. The final USDA area estimate was 19.8 million acres (a -26.9-percent relative difference).

The August spring wheat production estimate was 347.4 million bushels (table A-II). By September, the estimate was 409.4 million bushels. The increase was

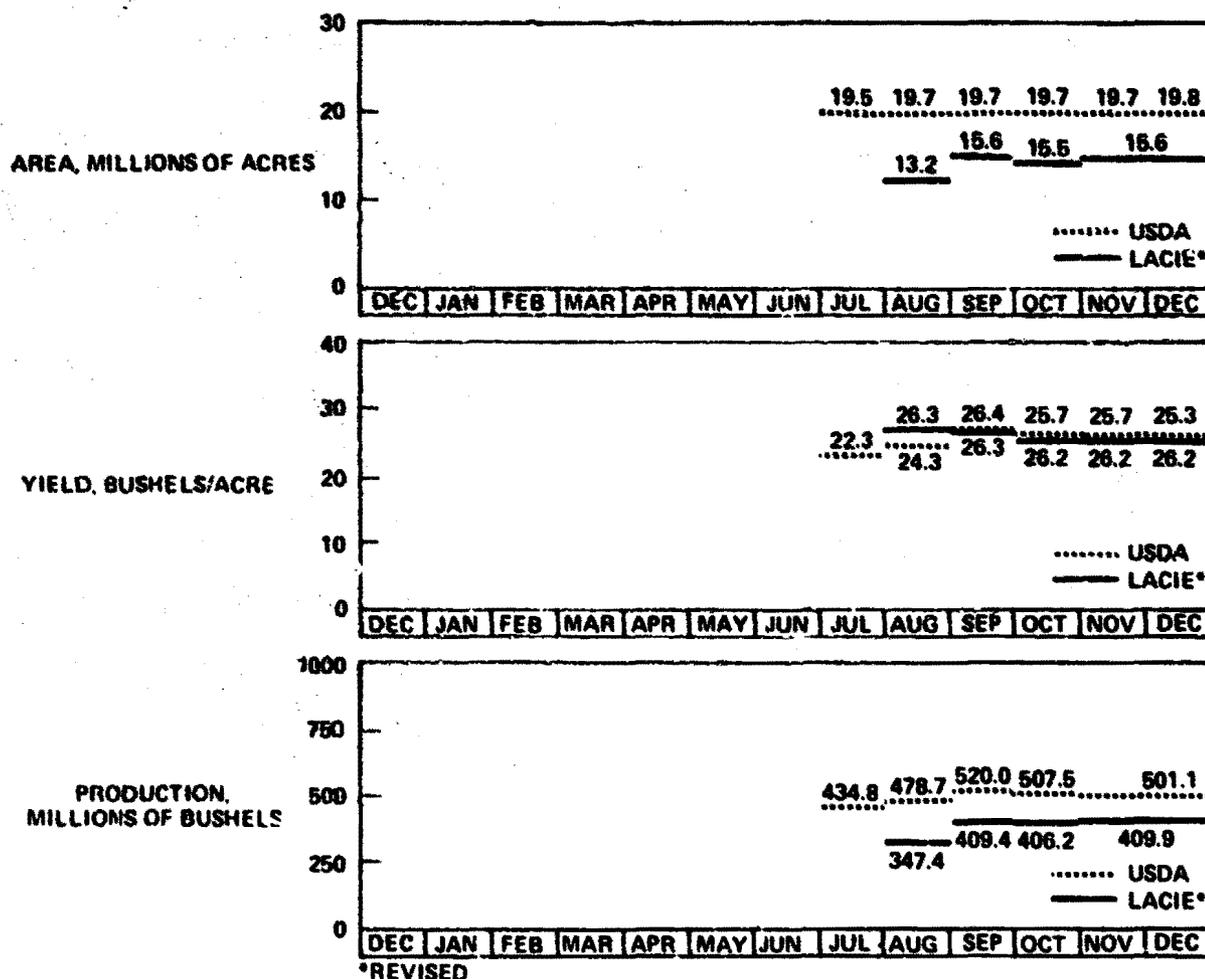


FIGURE 4.—Monthly comparison of LACIE and USDA estimates; Phase II; spring wheat, four states (Minnesota, Montana, North Dakota, and South Dakota).

due primarily to the area estimate increases since only the Montana spring wheat yield increased substantially (tables A-I and A-II). In October, the production estimate was 406.2 million bushels; the decrease was due to the Minnesota area estimate decline. The final production estimate by LACIE was 409.9 million bushels, the increase due entirely to area estimate increases. The final USDA estimate was 501.1 million bushels (a -22.2-percent relative difference).

The derived yield for the four spring wheat states was 26.3 bushels per acre in August and September and 26.2 bushels per acre in October and for the final LACIE estimate (table A-III). The final USDA estimate was 25.3 bushels per acre (a relative difference of 3.4 percent).

U.S. Great Plains total wheat.—The first LACIE

estimate of total wheat for the nine states in the U.S. Great Plains came in August (fig. 5). The revised estimate was 43.1 million acres (table A-I). In September and October, the estimate was 44.8 million acres. The increase was due to the spring wheat estimates. The final LACIE area estimate was 45.0 million acres. The final USDA estimate was 51.5 million acres (a -14.4-percent relative difference).

The initial production estimate was 1145.8 million bushels (table A-II). In September, the estimate was 1199.9 million bushels; in October, it was 1197.5 million bushels; and the final estimate was 1204.1 million bushels. The final USDA estimate was 1356.7 million bushels, for a relative difference of -12.7 percent.

The initial derived yield estimate was 26.6 bushels per acre (table A-III). The final LACIE yield esti-

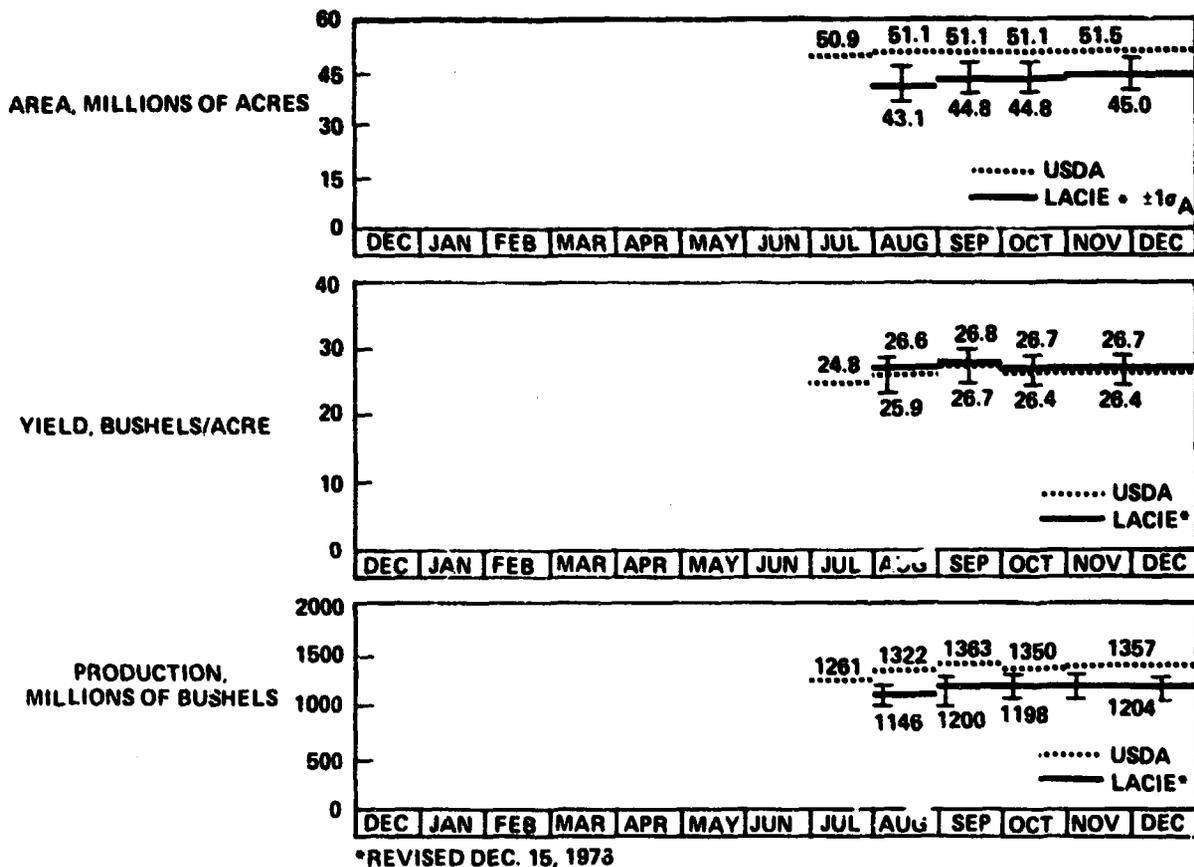


FIGURE 5.—Monthly comparison of LACIE and USDA estimates; Phase II; U.S. Great Plains, nine-state total (Colorado, Kansas, Nebraska, Oklahoma, Texas, Montana, North Dakota, South Dakota, and Minnesota).

mate was 26.7 bushels per acre compared to the final USDA estimate of 26.4 bushels per acre (a 1.1-percent relative difference).

Accuracy of the Estimates

In Phase II, LACIE estimates were made for area, yield, and production. Generally the yield estimates were quite close to USDA estimates and were considered satisfactory. However, the area and production estimates at the U.S. Great Plains level were low compared to the USDA estimates due primarily to significant underestimates of spring wheat area in the four U.S. Northern Great Plains states and to a significant underestimate of winter wheat area in Oklahoma.

An evaluation of the LACIE total wheat production estimator for the USGP in terms of the 90/90 criterion indicated that the coefficient of variation of the estimator, calculated to be 5 percent, was sufficiently small to tolerate a relative bias of 4 percent. However, a relative difference between the LACIE production estimate and the USDA estimate of -12.3 percent indicated that the relative bias of the production estimator was likely to be larger than was tolerable. An estimate of the bias using ground-truth information from the blind sites also indicated that the relative bias was larger than the tolerable 4 percent, supporting the difference observed from the USDA estimate. As a result, it was concluded that the 90/90 criterion was not met. It was inferred, however, based on the blind site analysis, that an accuracy goal of 90/75 was achievable.

For winter wheat production in the USGP, the LACIE estimate was not significantly different from the USDA estimate. However, significant acreage underestimation problems were indicated for Oklahoma, a problem not observed in Phase I. During Phase II, Oklahoma and other states of the Southern Great Plains experienced generally dry conditions through April 1976. These conditions created poor wheat stands and caused these wheat signatures to differ significantly from those of normal wheat. In some cases, sparsely vegetated fields were not detected as "emerged" acreage in the Landsat or on the aircraft ground-truth color-infrared imagery. April rains greatly improved the wheat stands; however, the drought-altered growth cycle misled the analysts to believe the late-recovering wheat to be spring-planted crops.

For spring wheat production in the USGP, a significant difference was observed between the LACIE and USDA estimates. The major contributors were spring wheat acreage underestimates for Minnesota and Montana. As was indicated in Phase I, spring wheat could not be reliably differentiated from some other spring small grains. As a result, historic ratios of spring wheat acreage to spring small grains acreage were used to obtain spring wheat acreage estimates. This introduced additional error into the spring wheat acreage estimates, particularly in Phase II when the planting of wheat in preference to other small grains greatly increased from previous years. Blind site comparisons of LACIE small grains proportion estimates with ground-truth proportion estimates also indicated a tendency towards underestimation of small grains proportions. Both ratioing and classification were significant contributors to the underestimation problem.

The small grains proportion underestimation was partially due to the strip-fallow cropping practice in the spring wheat region, particularly in Montana. Strip-fallow fields, some of which are small compared to Landsat resolution, are difficult to detect and measure on the imagery. In Minnesota, underestimation generally occurred in segments with high spring wheat density. Analysis indicated that unusual wheat signatures, partially due to color distortions in the Landsat imagery, were the major cause of the underestimation. Analysis also indicated that sampling was a problem in Minnesota. The reallocation in Phase III resulted in an increase from 13 segments to 47 segments.

Regarding the performance of the first-generation yield models developed during LACIE Phase I and implemented in Phase II, tests of them by comparison with 10 years of historic data indicate adequate performance in estimating wheat yields when aggregated to the USGP level. At state levels, however, investigations have indicated the need to improve yield predictions for extreme weather conditions. For example, 1975-76 was an extremely dry year for South Dakota, and USDA estimated the spring wheat yield at 10.9 bushels per acre and the winter wheat yield at 18.0 bushels per acre. The LACIE South Dakota yield models, on the other hand, estimated 17.2 and 31.6 bushels per acre for spring wheat and winter wheat yields, respectively. Even if a zero value of precipitation had been entered in the spring wheat model, the estimate would have been 13.0 bushels per acre. This indicates the

inadequacy of these yield model forms to reflect the total dynamic range of the plant's response to its environment.

Technical Issues

As a result of the LACIE experience through Phase II, several technical issues remained which required further study.

Differentiation of small grains.—Wheat was not reliably differentiated from other small grains. In Phase I, an analysis of North Dakota blind sites revealed that barley was not being reliably distinguished from spring wheat. In addition, other crops such as alfalfa and pasture became confusion crops. Efforts were begun late in Phase I to develop improved analysis procedures—procedures that would take advantage of any spectral separability that exists between the crops. For Phase II, however, the classification and mensuration procedures were used to estimate total small grains, and ratios based on historic proportions of spring wheat to small grains were used to convert the Landsat-based small grains estimates to wheat estimates.

Historic ratios of wheat to small grains.—In Phase II, the ratios from the latest year for which data were available at the county level were used to estimate wheat, given the Landsat-based estimate of small grains. In most cases, the current-year prevalence of wheat was not the same as the given historic year. In the United States, the current-year wheat ratio averaged about 10 percent higher than the historic ratio used. Thus, the use of the historic ratios contributed to the underestimation problem.

A second issue concerning ratios is that the historic ratio that was used to derive the wheat estimate from the CAMS small grains estimate needed to include as "small grains" all crops that CAMS included in their Landsat-based estimate of small grains. It was not readily apparent from the Landsat imagery what crops were included. Also, for the ratios to be more effective, it was necessary for the CAMS analyst to conscientiously identify all "small grains" on the imagery.

Improved yield models.—While the yield models had performed well in several regions, they tended to underestimate or overestimate yields in regions encountering extreme weather conditions. While the extreme weather conditions had been somewhat local within the LACIE regions, the models were not expected to perform well in a year for which a coun-

try was subjected to extreme conditions over a majority of its regions.

A second problem with the yield models resulted from the overlapping of the models in some areas. For example, during the development of the models, the Nebraska Panhandle area was used for both the "Badlands" model and the Nebraska model. This overlapping, plus the fact that the models were developed at the state level and used at the CRD level, introduced significant correlation problems with the statistical descriptors of the yield estimate.

Sampling mixed spring and winter wheat areas.—In LACIE Phases I and II, segments in areas containing both spring and winter wheat were designated winter or spring in proportion to the historic percentage of winter or spring wheat grown in the county. Once these segments were so designated, each segment was analyzed only for spring or only for winter small grains acreage and data were collected only during the growing season appropriate to either the winter or the spring grain crop calendar, but not both. However, a mixed area by definition has a probability of both winter wheat and spring wheat being grown in a sample segment area. Thus, it appeared logical to collect Landsat data for both winter and spring wheat in all sample segments allocated to mixed wheat areas throughout the complete growing season. This was, in fact, implemented for Phase III.

Sampling strategy.—The initial sample segment allocation was based on 1969 Census of Agriculture data. However, significant shifts in production patterns for wheat since 1969 resulted in an apparent undersampling in Minnesota. In addition, analysis of Landsat imagery indicated a need for further delineation of agricultural/nonagricultural areas.

PHASE III

Scope

The scope of Phase III in the U.S. Great Plains region was similar to that of Phase II—to estimate area, yield, and production components. The plan called for the generation of the estimates for the seven winter wheat states from February through October, the four-state spring wheat estimates from July through October, and the nine-state total wheat estimates from July through October. The July spring wheat and total wheat estimates were not released in real time but were released with the an-

nual report in December. The Phase III scope was further expanded to include parallel evaluations of the second-generation acreage sampling and yield estimation technology over portions of the yardstick region.

A reallocation was performed in the U.S. Great Plains for Phase III (to be discussed later). This modified allocation was not completed prior to the Phase III data order submission in August 1976 for the 1977 crop year. Therefore, the Phase II sample segments were ordered at that time. The initial LACIE Phase III crop report in February 1977 was based on these Phase II sample segments acquired through December 1976. The Phase III sample locations were completed and ordered retrospectively on January 31, 1977. The new segments acquired through December 1976 were processed, and a 601-segment "February" report, replacing the earlier 431-segment report, was generated on April 6, 1977.

Sampling

The first-generation sampling strategy utilized in Phases I and II was designed to achieve a 2-percent sampling error at the U.S. country level. This sampling strategy was modified in Phase III to achieve a 5-percent coefficient of variation (CV) of the production estimate for the U.S. Great Plains region. This CV would permit the 90/90 criterion to be met with a reasonable degree of bias in the production estimate.

The sampling strategy consisted of a two-stage stratified probability sample in which substrata (counties) were the primary sampling units and the 5- by 6-nautical-mile segments were the secondary units. Sample segments were allocated to the counties on the basis of weights which were a function of (1) the agricultural area in the county, (2) the within-county standard deviation of small grains area from segment to segment, (3) the classification error variance, (4) the county yield, and (5) the county yield prediction error. Depending on these weights, the counties were designated as Group I (high sampling rate), Group II (low sampling rate), and Group III (not sampled). (See Appendix B, LACIE C00200, Vol. IV, Rev. C, CAS Requirements Document, for more details.)

The segment-to-segment small grains standard deviation was determined from small grains identified in the Landsat imagery used during Phase II. The assumption was made that the distribution of

wheat was proportional to the distribution of small grains. Full-frame Landsat data were used to determine the agricultural areas.

The reallocation resulted in an increase in sample segments from 431 to 601. In Phase III, data were collected in the mixed wheat areas for the total wheat growing season—essentially the entire year. Thus, the assumption was that a mixed wheat area has a probability of both winter and spring wheat being grown in a sample segment. Table V shows the initial Phase III allocation of sample segments by state and by crop type.

As will be discussed later, problems occurred during Phase III operations which caused the allocation to be modified to reflect wheat rather than small

TABLE V.—Number of Sample Segments Allocated by State, by Crop Type, and by Group; U.S. Great Plains, Phase III

State	Segments allocated			Ending Phase III
	Group I	Group II	Total	
<i>Winter wheat</i>				
Colorado	28	4	32	31
Kansas	103	18	121	121
Nebraska	40	27	67	56
Oklahoma	42	4	46	46
Texas	11	27	38	35
5-state total	224	80	304	289
Montana	79	1	80	58
South Dakota	37	19	56	21
2-state total	116	20	136	79
7-state total	340	100	440	368
<i>Spring wheat</i>				
Minnesota	46	12	58	47
Montana	79	1	80	48
North Dakota	103	0	103	103
South Dakota	37	19	56	37
4-state total	265	32	297	235
<i>Total wheat</i>				
9-state total	489	112	601	557

grains and the segments in the mixed areas of Montana and South Dakota to be redesignated as spring only, winter only, mixed spring and winter, or not to be used. The last column of table V reflects these changes during Phase III. Thus, at the close of Phase III, the allocation was modified to 557 sample segments for the nine states.

Landsat Data

In Phase III, all acquisitions for sample segments in the U.S. Great Plains were analyzed by CAMS. Over 5300 acquisitions were acquired for the 601 segments allocated. Over a third of these acquisitions were usable in CAS for aggregation purposes.

In Phase III, a thresholding procedure was developed to eliminate, from consideration in the acreage estimation procedure, estimates from seg-

ments suspected of being incompletely emerged. This procedure consisted of monitoring the rates of change of segment wheat percentages between classification dates for a given segment. At the average date when the rate of change was small, the crop growth stage was computed, and all segment wheat percentage estimates based on Landsat acquisitions before that growth stage were deleted from the acreage estimation procedure. This procedure was shown to decrease the magnitude of the acreage underestimate. The underlying assumption was that it was more accurate to estimate an area using a Group III ratio than to use incomplete data.

Also, during Phase III, a procedure was developed to "screen" the segment wheat estimates. This screening procedure was aimed at identifying the outlier segments and eliminating these segments from the area aggregation procedure.

Table VI shows the number of segments used by

TABLE VI.—Number of Sample Segments Used by State for Each Monthly Report; U.S. Great Plains, Phase III^a

State	Feb. CMR	May CMR	June CMR	July CMR	Aug. CMR	Sept. CMR	Oct. CMR	End of season
<i>Winter wheat</i>								
Colorado	25	22	22	21	26	25	24	24
Kansas	82	98	104	96	103	107	108	106
Nebraska	41	38	40	29	31	40	39	39
Oklahoma	35	39	40	35	37	38	41	42
Texas	25	30	30	24	28	28	29	29
5-state total ^b	208	227	236	205	225	238	241	240
Montana	30	28	29	27	39	39	43	43
South Dakota	6	3	7	9	12	13	14	15
2-state total	36	31	36	36	51	52	57	58
7-state total	244	258	272	241	276	290	298	298
<i>Spring wheat</i>								
Minnesota				22	30	33	37	38
Montana				5	23	30	32	32
North Dakota				13	39	62	70	73
South Dakota				5	24	26	32	35
4-state total				45	116	151	172	178

^aMay and June estimates are screened estimates only, remaining estimates are screened and thresholded estimates

state and by crop type for each report during Phase III, as revised in CAS Annual Report 05, December 22, 1977. For the Southern Plains states, of the segments allocated, the number of segments used ranged from 71 percent for the July report to over 83 percent for the October report. For the seven winter wheat states, the number of segments used, as a percent of segments allocated, ranged from 66 percent in the July report to 81 percent in October. For the four spring wheat states, the percentage of segments allocated that were used ranged from 19 in July to over 75 by the end of season. CAS Annual Report 05 contains tables showing the average percent wheat used in each report, the distribution of segments used by percent wheat classified, the number of usable segments available, the number of usable segments available but not used, the average acquisition date and segment distribution by month of acquisition, the average number of elapsed days and the segment distribution by number of elapsed days, and the segment distribution by biostage.

All classifications were small grains classification from which the wheat was estimated using wheat to small grains ratios. For the five states, the ratios were based on the latest historical data available at the county level. For the four Northern Plains states, the ratios were based on econometric modeling to estimate the current-year ratios of wheat to small grains.

Yield Data

Yield estimates were received in CAS on the fourth working day of the month. The yield models were reworked to eliminate the area overlap that had caused problems with statistics during Phase II. In addition, the results of the yield model when run at the state level were applied to all CRD's within the model boundaries. This was a change from Phase II where the state model was run using CRD-level inputs to derive individual CRD-level yield estimates. The Phase III yield models are documented in LACIE 00431, Rev. A, June 1977. The models were developed from historical data as described previously for Phase II.

Estimates of Area, Yield, and Production

The discussion of the Phase III results refers to the estimates as revised in CAS Annual Report 05, December 22, 1977. These estimates were revised

using final Phase III CAS technology to provide a set of consistent estimates. Known data errors were also corrected. Appendix B contains the Phase III estimates of area, yield, and production by state, by crop type, and by report. Tables B-I, B-II, and B-III contain the revised area, production, and yield estimates, respectively, and tables B-IV, B-V, and B-VI contain the coefficients of variation for the revised estimates. Tables B-VII, B-VIII, and B-IX contain the estimates reported throughout Phase III.

U.S. Southern Great Plains winter wheat.—Figure 6 shows the revised monthly LACIE estimates of area, production, and yield for the five states compared to the corresponding monthly estimates of the USDA. The area estimates started at 18.5 million acres in February. The estimates for Kansas, Oklahoma, and Texas were extremely low in comparison to the historical data (table B-I). In May, the estimates increased to 26.3 million acres. Area estimates increased in all five states due mainly to increased percentage wheat estimates. The Colorado estimate was too high, while the Kansas, Oklahoma, and Texas estimates remained lower than historical data indicated was average. By July, the estimate had reached a peak of 30.8 million acres. The average per segment percentage wheat also peaked in July at 24.9 percent. The Colorado, Nebraska, and Texas estimates were higher than their historical averages, Kansas was slightly high, and Oklahoma remained slightly low. The final LACIE estimate was 29.5 million acres. The Colorado and Nebraska estimates remained high and the Oklahoma estimate low. The average per segment percentage wheat was 24.3 percent. The final USDA area estimate was 28.8 million acres (a 2.5-percent relative difference).

The initial production estimate for the five states was 472.9 million bushels (table B-II). In May, the estimate was 662.1 million bushels. The increase was due to the area estimate increases since the average yield declined from February to May. The production estimate peak occurred in July at 785.1 million bushels. Again, area estimate increases were primarily responsible for the production increase, although some individual state yields did increase. The final LACIE estimate was 752.0 million bushels, with the decline attributed to the decreased area estimate. The final USDA estimate was 797.2 million bushels (a 6.0-percent relative difference).

The derived yield estimate was 25.5 bushels per acre in February's report (table B-III). The estimate dropped to 25.1 bushels per acre during May and June and increased to 25.5 bushels per acre in July.

The final USDA estimate was 27.7 bushels per acre (a -8.6-percent relative difference).

U.S. Great Plains winter wheat.—The first estimates for the seven winter wheat states were made in February (table B-I). Figure 7 shows the revised LACIE area, production, and yield estimates with comparisons to the USDA estimates. The initial area estimate was 21.45 million acres. All state estimates except for those from Nebraska and South Dakota were underestimated, with the Kansas, Oklahoma, and Texas estimates being the largest underestimates. The area estimate increased to 30.8 million acres in May. The Colorado, South Dakota, Nebraska, and Montana estimates were over their respective historical averages. The Kansas, Oklahoma, and Texas estimates remained lower than normal. The area estimate reached its peak in July at 35.4 million acres. The area estimates for South Dakota, Nebraska, Colorado, Kansas, and Texas

were all higher than normal, with the South Dakota estimate three times the actual USDA estimate. The final LACIE estimate was 33.8 million acres compared to the final USDA estimate of 32.3 million acres (a 4.6-percent relative difference).

The initial seven-state production estimate was 551.5 million bushels (table B-II). This estimate increased to 787.1 in May due primarily to the area estimate increase. The production estimate reached a peak in July, again due primarily to the area estimate changes. The final LACIE estimate was 865.9 million bushels; the decrease was due to area estimate changes. The final USDA production estimate was 895.4 million bushels (a -3.4-percent relative difference).

The initial yield estimate for the seven states was 25.7 bushels per acre (table B-III). The estimate declined to 25.5 bushels per acre in May and June. The final LACIE estimate was 25.6 bushels per acre,

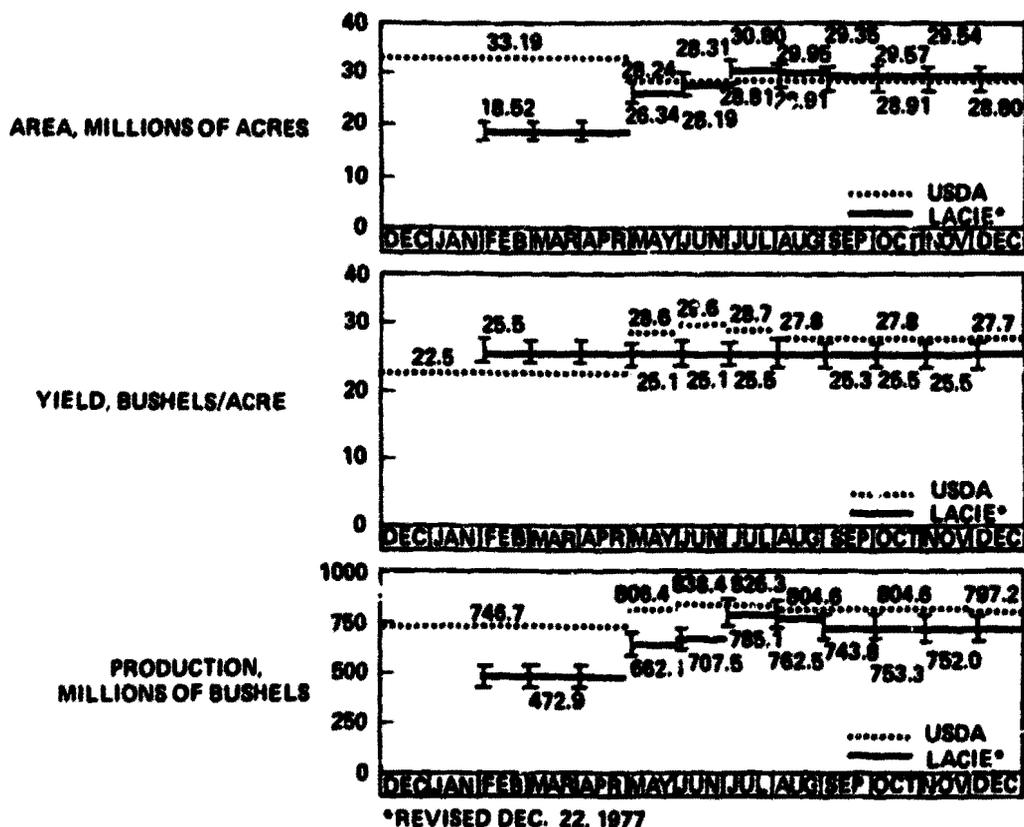


FIGURE 6.—Monthly comparison of LACIE and USDA estimates; Phase III; Southern Great Plains winter wheat states (Colorado, Kansas, Oklahoma, Nebraska, and Texas).

compared to 27.7 bushels per acre for the final USDA estimate (a relative difference of -8.5 percent).

U.S. Great Plains spring wheat.—The first spring wheat estimates of area, production, and yield were made in July (fig. 8). The initial area estimate was 14.7 million acres (table B-1). The area estimates for Minnesota and South Dakota were below their actual levels. The North Dakota estimate was within 1 percent of the actual harvested area. The August estimate increased to 16.0 million acres due primarily to the area estimate increase in South Dakota. The Minnesota estimate remained lower than the USDA estimate. The final LACIE estimate dropped to 15.64

million acres due to decreases in Minnesota and South Dakota. The final USDA estimate was 16.97 million acres (a -8.5-percent relative difference).

The initial spring wheat production estimate was 363.7 million bushels (table B-II). This estimate was low due to both area and yield underestimation. The August estimate was highest at 374.5 million bushels. This increase from July was due to area estimate increases since the yield estimates declined. The final LACIE estimate was 366.4 million bushels. Again, the change was due to the area estimates. The final USDA estimate was 460.6 million bushels (a -25.7-percent relative difference).

The derived yield for the four states was 24.8

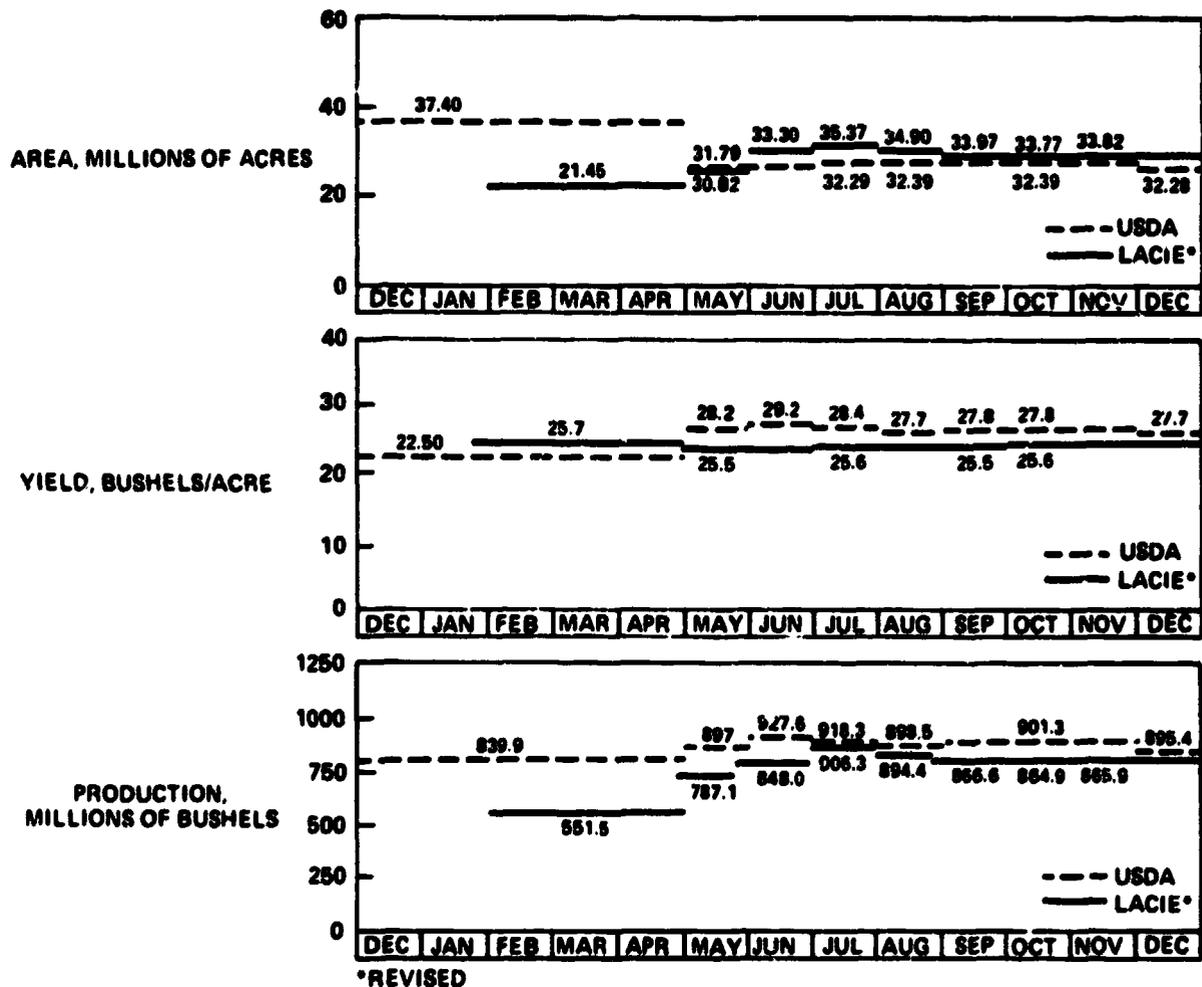


FIGURE 7.—Monthly comparison of LACIE and USDA estimates: Phase III; winter wheat, seven states (Colorado, Kansas, Nebraska, Oklahoma, Texas, Montana, and South Dakota).

bushels per acre in July (table B-III). The estimate was 23.4 bushels per acre in August, 23.6 bushels per acre in September, and 23.4 bushels per acre in October and for the final estimate. The final USDA yield estimate for the four states was 27.1 bushels per acre (a -15.8-percent relative difference).

U.S. Great Plains total wheat.—The initial area estimate for the total wheat in the nine states was 50.0 million acres (fig. 9 and table B-I). The winter wheat portion was overestimated by 3.0 million acres (8.7 percent) while the spring wheat was underestimated by 2.3 million acres (-15.8 percent). The August estimate was 50.9 million acres, resulting from the winter wheat overestimate of 2.6 million

acres (7.5 percent) and the spring wheat underestimate of 0.9 million acres (-5.9 percent). The final LACIE estimate was 49.46 million acres and resulted from the winter wheat overestimate of 1.5 million acres (4.6 percent) and the spring wheat underestimate of 1.3 million acres (-8.5 percent). The final USDA estimate was 49.2 million acres (a 0.4-percent relative difference).

The July production estimate for all wheat was 1270.0 million bushels, 6.8 percent below the final USDA estimate of 1356.0 million bushels (table B-II). The winter wheat production estimate was 10.9 million bushels (1.2 percent) above the final USDA winter wheat estimate, and the spring wheat estimate

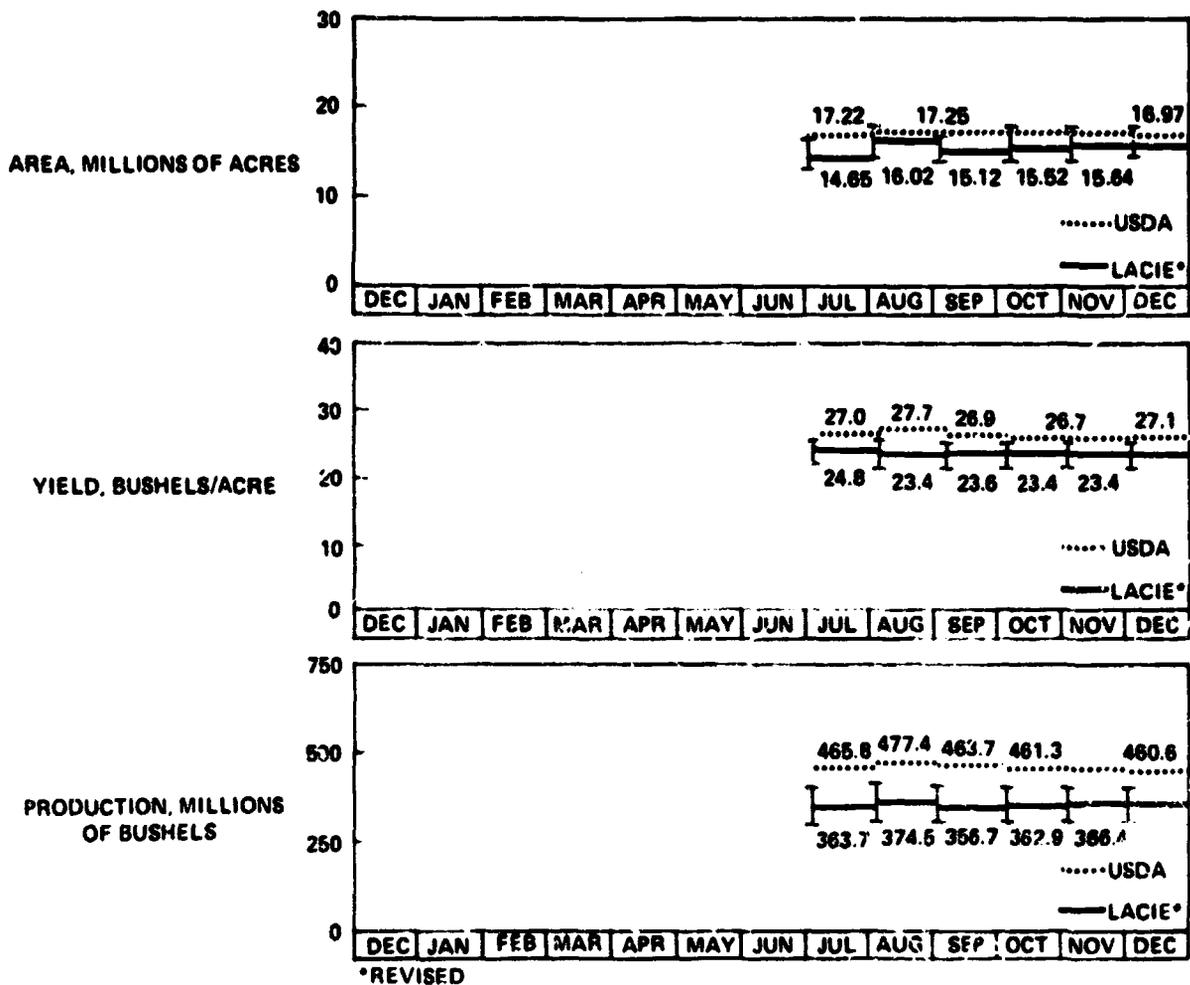


FIGURE 8.—Monthly comparison of LACIE and USDA estimates; Phase III; spring wheat, four states (Minnesota, Montana, North Dakota, and South Dakota).

was 96.6 million bushels (26.6 percent) below the USDA final spring wheat estimate. The August production estimate was 1269.0 million bushels, 6.9 percent below the final USDA estimate. The August winter wheat production estimate was 1.0 million bushels (0.1 percent) below the USDA estimate, and the spring wheat estimate was 85.7 million bushels (22.9 percent) below the USDA estimate. The final LACIE production estimate was 1232.3 million bushels, 10.0 percent below the final USDA estimate. The LACIE winter wheat estimate was 29.5 million bushels (3.4 percent) below the final USDA estimate, and the spring wheat estimate was 94.3 million bushels (25.7 percent) below the final USDA estimate.

The nine-state yield estimate was 25.4 bushels per acre in July and 24.9 bushels per acre in all other months (table B-III). The final USDA estimate was 27.5 bushels per acre (a relative difference of -10.4 percent).

Accuracy of the Estimates

In Phase III, LACIE estimates were made for area, yield, and production. The final USDA total wheat production estimate for the USGP at 1.36 billion bushels was 10.0 percent above the final LACIE estimate of 1.23 billion bushels. This relative difference of -10.0 percent indicates the presence of

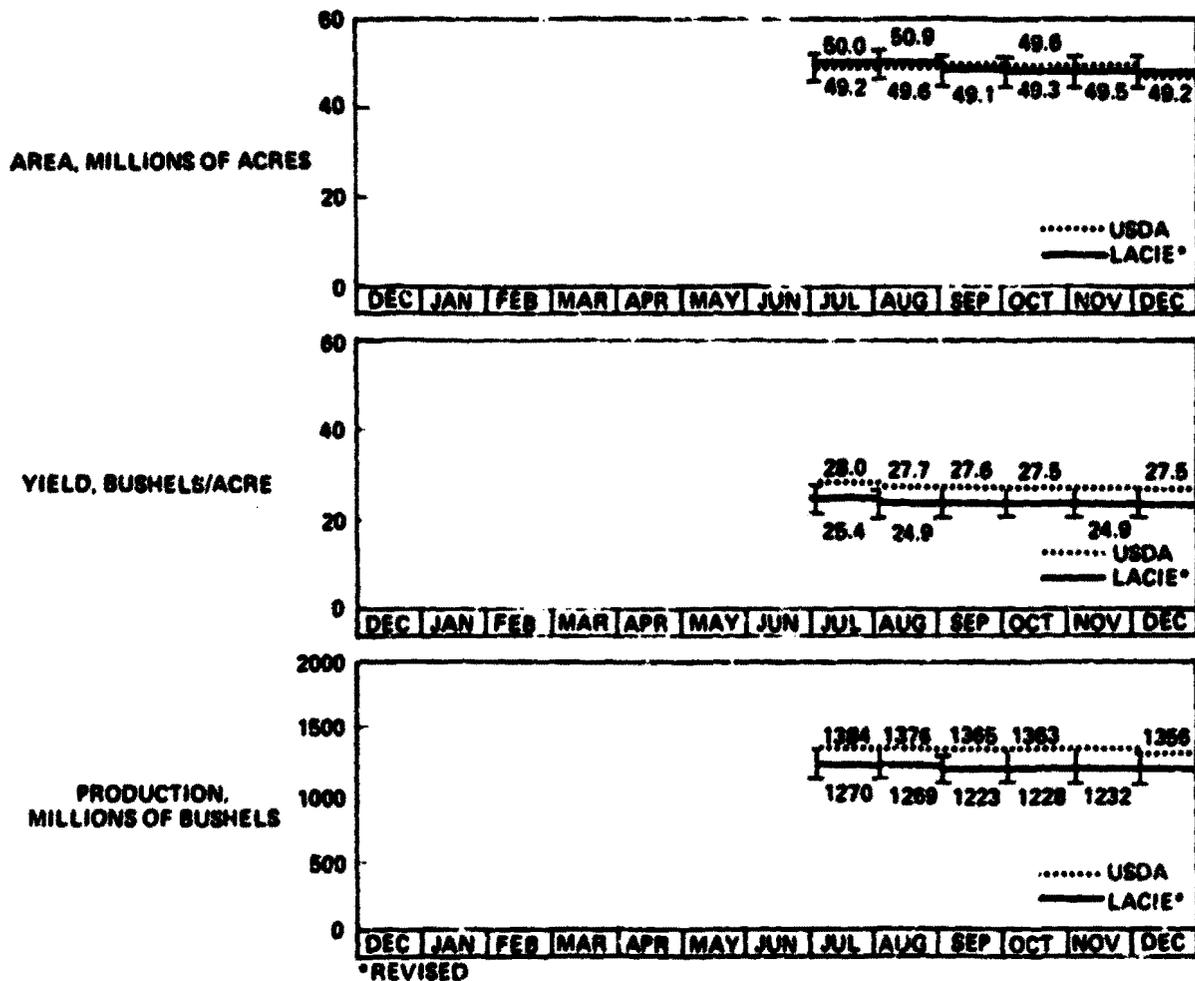


FIGURE 9.—Monthly comparison of LACIE and USDA estimates: Phase III; U.S. Great Plains, nine-state total (Colorado, Kansas, Nebraska, Oklahoma, Texas, Montana, Minnesota, North Dakota, and South Dakota).

a negative bias in the LACIE production estimator. The coefficient of variation of the LACIE production estimator was calculated to be 4.8 percent. This would allow a relative bias between -4.2 percent and 3.4 percent and still satisfy the 90/90 criterion. If the true relative bias of the LACIE production estimator is -4.2 percent, then the probability of observing a relative difference less than or equal to -10.0 percent is only about 11 percent. Thus, it is concluded that there is probably a bias in the LACIE estimator large enough to cause more than 1 of 10 estimates to fall outside the ± 10 -percent accuracy bounds required by the 90/90 criterion. However, even with a relative bias as large as -10 percent, the variability of 4.8 percent is small enough to produce estimates within ± 15 percent in 9 of 10 years; i.e., a 90/85 estimator. Thus, it would appear from these analyses that, while the LACIE estimator of USGP total wheat production did not satisfy the 90/90 criterion, it only marginally missed doing so.

For winter wheat production in the USGP, the LACIE estimate was not significantly different from the final USDA estimate. The LACIE winter wheat acreage estimate was in close agreement with USDA figures; however, for the first time in LACIE, the yield model predictions were consistently below the USDA estimates. The biggest contributors to the yield underestimate for both the five- and the seven-state regions were the Oklahoma and Texas yield models. Investigations indicated two primary factors contributing to these underestimates.

In both states, the technology trend term was selected such that no average increase in yield occurred due to technology since 1960. On the contrary, ancillary data show that an irrigated winter wheat area in Texas now produces almost 25 percent of the total winter wheat acreage for Texas. Nearly all of this additional irrigated acreage has been introduced since 1960. The weather terms in the Texas model did not alter the yield estimate significantly from trend. Therefore, it is likely that the constant trend since 1960 is a major contributor to the underestimate for Texas winter wheat yield.

In Oklahoma, both the weather terms and the constant trend term were factors in the underestimate. The model underestimate in Oklahoma resulted mainly from below-normal precipitation between August and February (over the winter period), a March precipitation deficit relative to potential evapotranspiration, and an above-average May precipitation. The weather factors which most likely contributed to the improved Oklahoma yields and

which were overlooked by the LACIE yield model were the above-normal April temperatures and precipitation and the temporal distribution of the May precipitation in Oklahoma.

The April temperatures were about 5° F above normal (upper 60's) in Oklahoma, which would make them nearly ideal for wheat. Three inches or more of well-distributed precipitation occurred in April and 4 inches fell in May. Good April rainfall amounts following moisture-deficit periods, such as those which occurred during the previous winter months and even the previous season, typically give an extra stimulus to yield by encouraging more extensive crop rooting. This results in improved utilization of nutrients when moisture becomes available.

The monthly averaging of precipitation in the Oklahoma model also created an unrealistic response to the rather well-distributed May rainfall, which nearly doubled the average May precipitation. Since wheat in Oklahoma is harvested at the end of May and the first of June, large amounts of rainfall near the end of May tend to reduce yields. However, a majority of the 1977 May precipitation came in mid-May, with lesser amounts in late May. The mid-May precipitation came during the heading-to-ripening period for Oklahoma winter wheat and thus contributed to increased yields, as opposed to the decrease predicted by the LACIE models.

For spring wheat production in the USGP, the LACIE estimate was significantly smaller than the USDA estimates and the 90/90 accuracy goal was not supported. The underestimation of spring wheat production was primarily due to yield underestimation, although area was significantly underestimated also. However, significant improvement was realized in the spring wheat area estimate (as compared to USDA) over Phase I and II results.

The LACIE spring wheat yield estimates were significantly smaller than the corresponding USDA estimates throughout Phase III, primarily due to underestimates of Minnesota and Montana spring wheat yields. These errors were due, in part, to trend terms which failed to account for new varieties of wheat in Minnesota and for increased fertilizer usage in Montana during the past 5 years.

Technical issues

At the completion of Phase III processing, the following technical issues were apparent.

Wheat/small grains separation.—As has been a

2-6

problem throughout LACIE, CAMS analysis has not reliably, nor consistently, been able to differentiate between wheat and small grains. However, for the first time during LACIE, a technique for separating wheat and small grains was tested in North Dakota. The results appear encouraging thus far, pending completion of a more thorough evaluation.

Improved yield models.—The first-generation yield model estimates were noticeably below the USDA estimates of yield. Although the 10-year tests and the 3 years of LACIE operations indicate that the yield models performed adequately (in terms of satisfying the 90/90 criterion), investigations of model performance at the subregional levels have indicated that the models could and should be improved. In a year with extended episodic conditions, the yield models are not adequately responsive to extreme climatic conditions, and, during such years, considerable yield estimation error can result.

Sampling in mixed wheat regions.—During Phases I and II, segments were allocated on the basis of total wheat statistics, and areas containing both spring and winter wheat were designated either winter or spring depending on the historical predominance of winter or spring wheat in the county. Once designated, each segment was analyzed for one crop type only, and data were collected only during the appropriate growing season. This strategy created problems for those segments which contained significant amounts of both winter and spring wheat.

In Phase III, Landsat data were collected in the mixed wheat areas for the total growing season of both winter and spring wheat. This collection scheme acquired satellite data to estimate both spring and winter wheat grown in all segments.

However, in Phase III a problem occurred in South Dakota, a mixed wheat state. The real-time LACIE estimates for South Dakota winter wheat were significantly larger than one would expect on the basis of historical data. An investigation disclosed the fact that many South Dakota segments contained almost no winter wheat. In low-density segments, the CAMS errors tend to be overestimates. While the absolute overestimate was not large, the relative overestimate was quite large. This relative overestimate resulted in the large South Dakota overestimate.

The problem was not caused by the mixed wheat sampling plan but resulted from (1) the sampling strategy based on total small grains which resulted in

sample segments being placed in low wheat density/high small grains density areas, (2) the CAMS procedures indicating that both winter and spring wheat must be found in all segments, and (3) the CAMS inability to accurately identify wheat in low wheat density regions.

Ratio modeling.—The CAMS inability to identify wheat resulted in the use of ratios of wheat to small grains to estimate wheat from satellite-based small grains estimates. During Phase II, the ratios were based on data for the most recent year for which statistics were available at the county level. However, the recent predominance of wheat was probably significantly different from the current year. This resulted in the introduction of another error into the estimation process. At the end of Phase II, an effort was initiated to econometrically model the current-year ratios of wheat to small grains for the four Northern Plains states. The accuracy of the models in predicting current-year ratios is unknown at present, but one point is certain—ratio modeling is no substitute for the ability to identify crops.

Sampling based on total small grains.—The Phase III sampling strategy was based on the incidence of total small grains. The assumption was that the distribution of wheat was proportional to the distribution of small grains. This assumption was in error in large areas of the U.S. Great Plains and resulted in the placement of segments in areas of low wheat density, high small grains density. The low wheat density segments then resulted in area estimation problems as discussed previously.

Data editing techniques.—During Phase III, two data editing procedures were applied in the U.S. Great Plains. Thresholding was intended to eliminate segment data that were suspected of being biased low because of an early acquisition date. (An early acquisition date was synonymous with low ground cover and therefore marginal signatures on the imagery.) The screening procedure was applied to determine whether the CAMS classification result was significantly different from the estimates of segments in the same group stratified by historical data.

Neither procedure was subject to rigorous evaluation, and both should have further testing and verification before being used again. In any event, such procedures should be considered as data analysis procedures to flag anomalies, which should then be sent back to the CAMS for further investigation.

RECOMMENDATIONS

At the conclusion of LACIE we should ask ourselves one question: can we separate wheat, consistently and accurately, from other crops without the use of ancillary data or ratioing techniques? The answer is NO. As we proceed to other crops, it should be obvious that the most important goal must be the crop separation/identification.

A second priority item should be the development of more accurate yield models. These models must be able to estimate yields accurately through extended episodic events. The use of Landsat data in the models is a worthy goal.

Finally, as LACIE moves on to other crops, it becomes more important to develop a sampling strategy that is both accurate for the particular crops of interest and efficient in terms of multiple-crop uses of each segment.

Appendix A

LACIE Phase II Estimates

TABLE A-1.—LACIE Area Estimates for Phase II, by Crop Type, by State, and by Monthly Report, as Revised in CAS Annual Report 03, December 15, 1976, U.S. Great Plains

(Thousands of acres)

State	CAS monthly report										
	Feb.	Mar.	Apr.	May	June 8	June 29	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>											
Colorado	3 539	2 768	2 768	2 807	2 995	2 960	2 867	2 830	2 704	2 704	2 704
Kansas	8 013	8 536	8 536	9 392	10 535	10 744	10 795	10 932	10 989	10 989	11 125
Nebraska	4 500	3 632	3 583	3 653	4 104	4 178	4 133	4 086	3 399	3 399	3 399
Oklahoma	3 499	3 450	3 450	3 897	4 148	4 182	4 025	4 305	4 261	4 261	4 261
Texas	3 170	3 725	3 479	4 810	4 556	4 619	4 314	4 310	4 344	4 344	4 344
5-state total ^a	22 721	22 111	21 816	24 559	26 338	26 683	26 134	26 463	25 697	25 697	25 833
Montana					488	885	1 044	1 911	2 103	2 131	2 079
South Dakota					1 159	1 210	1 482	1 482	1 452	1 452	1 452
2-state total ^a					1 647	2 095	2 526	3 393	3 555	3 583	3 531
7-state total ^a					27 986	28 778	28 660	29 856	29 252	29 280	29 363
<i>Spring wheat</i>											
Minnesota								1 741	2 551	2 198	2 198
Montana								1 127	1 291	1 487	1 516
North Dakota								8 161	9 650	9 735	9 856
South Dakota								2 169	2 095	2 079	2 079
4-state total ^a								13 198	15 586	15 499	15 650
<i>Total wheat</i>											
9-state total ^a								43 054	44 838	44 779	45 013

^aTotals may not add correctly because of rounding.

TABLE A-II.—LACIE Production Estimates for Phase II, by Crop Type, by State, and by Monthly Report, as Revised in CAS Annual Report 03, December 15, 1976, U.S. Great Plains

(Thousands of bushels)

State	CAS monthly report										
	Feb.	Mar.	Apr.	May	June 8	June 29	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>											
Colorado	76 418	60 759	56 089	55 285	61 191	60 500	51 492	50 024	52 924	52 924	52 924
Kansas	258 074	269 638	255 147	283 124	326 677	333 644	334 107	338 078	339 974	339 974	344 472
Nebraska	151 762	124 342	118 458	110 496	128 692	131 019	132 118	130 547	110 972	110 972	110 972
Oklahoma	80 264	76 041	74 823	84 699	94 975	95 667	92 052	98 156	96 491	96 491	96 491
Texas	59 550	66 676	59 559	86 910	84 094	85 324	80 797	80 637	81 312	81 312	81 312
5-state total	626 068	597 456	564 076	620 514	695 629	706 154	690 566	697 442	681 673	681 673	686 171
Montana					13 527	24 808	30 082	55 788	62 877	63 758	62 167
South Dakota					31 553	32 931	45 096	45 096	45 904	45 904	45 904
2-state total					45 080	57 739	75 178	100 884	108 781	109 662	108 071
7-state total					740 709	763 893	765 744	798 326	790 454	791 335	794 242
<i>Spring wheat</i>											
Minnesota								55 490	77 230	66 589	66 589
Montana								29 188	35 064	40 240	41 058
North Dakota								226 034	261 197	263 703	266 529
South Dakota								36 719	35 908	35 675	35 675
4-state total								347 431	409 399	406 207	409 851
<i>Total wheat</i>											
9-state total								1 145 757	1 199 853	1 197 542	1 204 093

TABLE A-III.—LACIE Yield Estimates for Phase II, by Crop Type, by State, and by Monthly Report as Revised in CAS Annual Report 03, December 15, 1976, U.S. Great Plains

[Bushels per acre]

State	CAS monthly report										
	Feb.	Mar.	Apr.	May	June 8	June 29	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>											
Colorado	21.6	22.0	20.3	19.7	20.4	20.4	18.0	17.7	19.6	19.6	19.6
Kansas	32.2	31.6	29.9	30.1	31.0	31.1	30.9	30.9	30.9	30.9	31.0
Nebraska	33.7	34.2	33.1	30.2	31.4	31.4	32.0	32.0	32.7	32.7	32.7
Oklahoma	22.9	22.0	21.7	21.7	22.9	22.9	22.9	22.8	22.6	22.6	22.6
Texas	18.8	17.9	17.1	18.1	18.5	18.5	18.7	18.7	18.7	18.7	18.7
5-state average	27.6	27.0	25.9	25.3	26.4	26.5	26.4	26.4	26.5	26.5	26.6
Montana					27.7	28.0	28.8	29.2	29.9	29.9	29.9
South Dakota					27.2	27.2	30.4	30.4	31.6	31.6	31.6
2-state average					27.4	27.6	29.8	29.7	30.6	30.6	30.6
7-state average					26.5	26.5	26.7	26.7	27.0	27.0	27.0
<i>Spring wheat</i>											
Minnesota								31.9	30.3	30.3	30.3
Montana								25.9	27.2	27.1	27.1
North Dakota								27.7	27.1	27.1	27.0
South Dakota								16.9	17.1	17.2	17.2
4-state average								26.3	26.3	26.2	26.2
<i>Total wheat</i>											
9-state average								26.6	26.8	26.7	26.7

TABLE A-IV.—Coefficient of Variation of the LACIE Revised Area Estimates for Phase II, by Crop Type, by State, and by Monthly Report, CAS Annual Report 03, December 15, 1976, U.S. Great Plains

(Percent)

State	CAS monthly report										
	Feb.	Mar.	Apr.	May	June 8	June 29	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>											
Colorado	26	25	25	24	23	22	25	24	24	24	24
Kansas	12	8	8	6	6	6	6	5	5	5	5
Nebraska	18	13	13	13	12	12	11	11	11	11	11
Oklahoma	24	18	18	16	14	14	15	15	14	14	14
Texas	25	30	20	14	15	14	15	16	16	16	16
5-state CV	9	8	7	6	5	5	5	5	5	5	5
Montana					193	56	52	35	29	28	28
South Dakota					44	34	23	23	23	23	23
2-state CV					65	31	25	22	20	19	19
7-state CV					6	5	5	5	5	5	5
<i>Spring wheat</i>											
Minnesota								40	27	30	30
Montana								28	23	24	22
North Dakota								14	5	5	5
South Dakota								12	13	13	13
4-state CV								10	6	6	6
<i>Total wheat</i>											
9-state CV								5	4	4	4

TABLE A-V.—Coefficient of Variation of the LACIE Revised Production Estimates for Phase II, by Crop Type, by State, and by Monthly Report, CAS Annual Report 03, December 15, 1976, U.S. Great Plains

(Percent)

State	CAS monthly report										
	Feb.	Mar.	Apr.	May	June 8	June 29	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>											
Colorado	33	32	32	31	28	28	30	29	29	29	29
Kansas	17	14	13	12	11	11	11	10	10	10	10
Nebraska	23	19	19	19	17	17	16	16	16	16	16
Oklahoma	29	25	22	21	17	17	18	18	18	18	18
Texas	28	32	22	17	17	17	17	18	18	18	18
5-state CV	11	10	8	8	7	7	7	7	7	7	7
Montana					192	57	53	36	30	29	30
South Dakota					46	37	27	26	26	26	26
2-state CV					63	32	27	23	21	20	20
7-state CV					8	7	7	7	7	7	7
<i>Spring wheat</i>											
Minnesota								42	29	32	32
Montana								29	25	25	24
North Dakota								17	12	12	12
South Dakota								18	19	18	18
4-state CV								13	10	10	10
<i>Total wheat</i>											
9-state CV								6	5	5	5

TABLE A-VI.—Coefficient of Variation of the LACIE Revised Yield Estimates for Phase II, by Crop Type, by State, and by Monthly Report, CAS Annual Report 03, December 15, 1976, U.S. Great Plains

(Percent)

State	CAS monthly report										
	Feb.	Mar.	Apr.	May	June 8	June 29	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>											
Colorado	21	21	21	20	17	17	17	17	17	17	17
Kansas	12	12	10	10	9	9	9	9	9	9	9
Nebraska	14	14	14	14	13	13	12	12	12	12	12
Oklahoma	17	17	14	14	10	10	10	10	10	10	10
Texas	19	18	14	13	12	12	12	12	12	12	12
5-state CV	7	7	6	6	5	5	5	5	5	5	5
Montana					12	12	9	9	9	9	9
South Dakota					15	15	15	14	14	14	14
2-state CV					9	10	9	8	8	8	8
7-state CV					5	5	5	5	5	5	5
<i>Spring wheat</i>											
Minnesota							11	11	11	11	11
Montana							9	9	9	9	9
North Dakota							11	11	11	11	11
South Dakota							14	13	13	13	13
4-state CV							7	7	7	7	7
<i>Total wheat</i>											
9-state CV							4	4	4	4	4

TABLE A-VII.—Real-Time LACIE Area Estimates for Phase II, by Crop Type, by State, and by Monthly Report, U.S. Great Plains

(Thousands of acres)

State	CAS monthly report										
	Feb.	Mar.	Apr.	May	June 8	June 29	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>											
Colorado	3 900	2 755	2 755	2 795	2 969	3 023	2 856	2 851	2 735	2 735	2 704
Kansas	8 413	8 468	8 499	9 463	10 623	10 855	10 937	10 956	10 969	10 989	11 125
Nebraska	5 385	3 750	3 609	3 679	4 111	4 184	4 140	4 092	3 399	3 399	3 399
Oklahoma	3 498	3 433	3 449	3 917	4 148	4 181	4 031	4 311	4 267	4 268	4 261
Texas	3 208	3 947	3 602	5 644	4 578	4 642	4 266	4 313	4 344	4 344	4 344
5-state total ^a	24 404	22 353	21 914	25 498	26 429	26 885	26 230	26 523	25 714	25 735	25 833
Montana					511	836	918	1 448	1 783	2 128	2 079
South Dakota					573	613	783	1 305	1 263	1 415	1 452
2-state total ^a					1 084	1 449	1 701	2 753	3 046	3 543	3 531
7-state total ^a					27 513	28 334	27 931	29 276	28 760	29 278	29 363
<i>Spring wheat</i>											
Minnesota								1 300	2 583	2 192	2 198
Montana								1 205	1 382	1 487	1 516
North Dakota								6 835	9 598	9 600	9 856
South Dakota								1 837	2 063	2 140	2 079
4-state total ^a								11 177	15 626	15 419	15 650
<i>Total wheat</i>											
9-state total ^a								40 453	44 386	44 697	45 013

^aTotals may not add correctly because of rounding.

TABLE A-VIII.—Real-Time LACIE Production Estimates for Phase II, by Crop Type, by State, and by Monthly Report, U.S. Great Plains

[Thousands of bushels]

State	C/S monthly report										
	Feb.	Mar.	Apr.	May	June 8	June 29	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>											
Colorado	84 516	60 500	55 847	54 973	60 634	61 676	51 290	55 697	53 535	53 534	52 924
Kansas	270 415	267 199	253 948	285 572	329 607	337 214	338 940	340 092	339 728	339 974	344 472
Nebraska	168 944	128 366	119 359	111 280	128 890	131 216	132 322	134 040	110 970	110 972	110 972
Oklahoma	80 385	75 765	74 808	85 027	94 962	95 645	92 214	97 663	96 645	96 670	96 491
Texas	60 286	70 637	61 642	101 308	84 470	85 723	79 817	80 798	81 325	81 312	81 312
5-state total	664 546	602 467	565 604	638 160	698 563	711 474	694 583	708 290	682 203	682 462	686 171
Montana					14 154	23 341	26 512	43 528	53 260	63 666	62 167
South Dakota					17 755	18 941	24 360	41 858	39 117	44 722	45 904
2-state total					31 909	42 282	50 872	85 386	92 377	108 388	108 071
7-state total					730 472	753 756	745 455	793 676	774 580	790 850	794 242
<i>Spring wheat</i>											
Minnesota								39 361	78 200	66 404	66 589
Montana								33 411	37 406	40 240	41 058
North Dakota								148 985	259 815	260 198	266 529
South Dakota								236	35 417	36 765	35 675
4-state total								292 093	410 838	403 607	409 851
<i>Total wheat</i>											
9-state total								1 085 769	1 185 418	1 194 457	1 204 093

**TABLE A-IX.—Real-Time LACIE Yield Estimates for Phase II
by Crop Type, by State, and by Monthly Report, U.S. Great Plains**

(Bushels per acre)

State	CAS monthly report										
	Feb.	Mar.	Apr.	May	June 8	June 29	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>											
Colorado	21.7	22.0	20.3	19.7	20.4	20.4	18.0	19.5	19.6	19.6	19.6
Kansas	32.1	31.6	29.9	30.2	31.0	31.1	31.0	31.0	31.0	30.9	31.0
Nebraska	31.4	34.2	33.1	30.2	31.4	31.4	32.0	32.8	32.7	32.7	32.7
Oklahoma	23.0	22.1	21.7	21.7	22.9	22.9	22.9	22.7	22.7	22.7	22.6
Texas	18.8	17.9	17.1	17.9	18.4	18.5	18.7	18.7	18.7	18.7	18.7
5-state average	27.2	27.0	25.8	25.0	26.4	26.5	26.5	26.7	26.5	26.5	26.6
Montana					27.7	27.9	28.9	30.1	29.9	29.9	29.9
South Dakota					31.0	30.9	31.1	32.1	31.0	31.6	31.6
2-state average					29.4	29.2	29.9	31.0	30.3	30.6	30.6
7-state average					26.6	26.6	26.7	27.1	26.9	27.0	27.0
<i>Spring wheat</i>											
Minnesota							30.3	30.3	30.3	30.3	30.3
Montana							27.7	27.1	27.1	27.1	27.1
North Dakota							27.5	27.1	27.1	27.1	27.0
South Dakota							17.0	17.2	17.2	17.2	17.2
4-state average							26.1	26.3	26.2	26.2	26.2
<i>Total wheat</i>											
9-state average							26.8	26.7	26.7	26.7	26.7

Appendix B LACIE Phase III Estimates

**TABLE B-1.—Revised LACIE Area Estimates by State for Each CAS Report in Phase III,
U.S. Great Plains, CAS Annual Report 05, December 22, 1977**

(Thousands of acres)

State	CAS monthly report							
	Feb.	May	June	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>								
Colorado	1 997	3 600	3 608	3 268	3 253	3 059	3 395	3 459
Kansas	6 826	10 439	11 055	12 919	12 579	12 468	12 669	12 494
Nebraska	3 067	3 278	3 839	3 844	3 558	3 130	3 375	3 433
Oklahoma	3 206	4 832	5 228	5 755	5 963	6 083	5 658	5 675
Texas	3 365	4 196	4 462	5 011	4 600	4 613	4 476	4 476
5-state total ^a	18 523	26 345	28 192	30 797	29 953	29 353	29 573	29 537
Montana	2 127	3 369	3 704	2 626	3 355	3 628	3 314	3 371
South Dakota	800	1 107	1 401	1 943	1 594	989	883	912
2-state total ^a	2 927	4 476	5 105	4 569	4 949	4 617	4 197	4 283
7-state total ^a	21 450	30 821	33 297	35 366	34 902	33 969	33 771	33 820
<i>Spring wheat</i>								
Minnesota				2 420	2 553	2 474	2 289	2 344
Montana				1 895	1 942	2 187	2 150	2 174
North Dakota				9 071	9 220	8 523	9 173	9 183
South Dakota				1 269	2 309	1 958	1 909	1 936
4-state total ^a				14 655	16 024	15 142	15 522	15 638
<i>Total wheat</i>								
9-state total ^a				50 021	50 926	49 111	49 293	49 358

^aTotals may not add correctly because of rounding

**TABLE B-II.—Revised LACIE Production Estimates by State for Each CAS Report in Phase III,
U.S. Great Plains, CAS Annual Report 05, December 22, 1977**

(Thousands of bushels)

State	CAS monthly report							
	Feb.	May	June	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>								
Colorado	45 520	81 898	85 314	73 383	73 031	68 675	76 226	77 666
Kansas	199 123	293 385	312 339	372 688	362 866	359 652	365 465	360 410
Nebraska	93 931	102 497	115 745	122 819	114 134	100 106	107 830	109 823
Oklahoma	69 688	102 554	103 413	114 725	119 208	121 845	113 064	113 387
Texas	64 623	81 789	90 667	101 510	93 261	93 510	90 695	90 695
5-state total ^a	472 885	662 123	707 478	785 125	762 500	743 788	753 280	751 981
Montana	56 803	96 173	104 087	69 502	88 789	96 021	87 712	89 224
South Dakota	21 849	28 809	36 457	51 718	43 143	26 760	23 907	24 682
2-state total ^a	78 652	124 982	140 544	121 220	131 932	122 781	111 619	113 906
7-state total ^a	551 536	787 105	848 022	906 345	894 432	866 570	864 900	865 888
<i>Spring wheat</i>								
Minnesota				78 481	80 840	79 043	73 213	74 955
Montana				34 939	34 939	39 357	38 683	39 112
North Dakota				223 257	210 668	197 503	211 253	211 990
South Dakota				26 977	48 075	40 759	39 748	40 309
4-state total ^a				363 654	374 522	356 662	362 896	366 367
<i>Total wheat</i>								
9-state total ^a				1 269 999	1 268 954	1 223 233	1 227 796	1 232 255

^aTotals may not add correctly because of rounding.

**TABLE B-III.—Revised LACIE Yield Estimates by State for Each CAS Report in Phase III,
U.S. Great Plains, CAS Annual Report 05, December 22, 1977**

(Bushels per acre)

State	CAS monthly report							End of season
	Feb.	May	June	July	Aug.	Sept.	Oct.	
<i>Winter wheat</i>								
Colorado	22.8	22.8	23.6	22.5	22.5	22.5	22.5	22.5
Kansas	28.9	28.1	28.3	28.8	28.8	28.8	28.8	28.8
Nebraska	30.6	31.3	30.2	31.9	32.1	32.0	32.0	32.0
Oklahoma	21.7	21.2	19.8	19.9	20.0	20.0	20.0	20.0
Texas	19.2	19.5	20.3	20.3	20.3	20.3	20.3	20.3
5-state average	25.5	25.1	25.1	25.5	25.5	25.3	25.5	25.5
Montana	26.7	28.5	28.1	26.5	26.5	26.5	26.5	26.5
South Dakota	27.3	26.0	26.0	26.6	27.1	27.1	27.1	27.1
2-state average	26.9	27.9	27.5	26.5	26.7	26.6	26.6	26.6
7-state average	25.7	25.5	25.5	25.6	25.6	25.5	25.6	25.6
<i>Spring wheat</i>								
Minnesota				32.4	31.7	31.9	32.0	32.0
Montana				18.4	18.0	18.0	18.0	18.0
North Dakota				24.6	22.8	23.2	23.0	23.1
South Dakota				21.3	20.8	20.8	20.8	20.8
4-state average				24.8	23.4	23.6	23.4	23.4
<i>Total wheat</i>								
9-state average				25.4	24.9	24.9	24.9	24.9

TABLE B-IV.—Coefficient of Variation for the Revised LACIE Area Estimates for Phase III, by State and by Monthly Report, U.S. Great Plains, CAS Annual Report 05, December 22, 1977

(Percent)

State	CAS monthly report							End of season
	Feb.	May	June	July	Aug.	Sept.	Oct.	
<i>Winter wheat</i>								
Colorado	21.0	14.2	13.6	13.4	11.3	10.3	9.9	9.8
Kansas	13.9	6.2	5.8	4.5	4.8	4.5	4.2	4.0
Nebraska	14.9	11.4	9.5	11.6	10.2	9.2	9.6	9.2
Oklahoma	9.6	10.0	9.0	7.1	6.7	7.2	7.7	7.6
Texas	16.7	14.2	12.5	11.6	12.8	12.7	13.7	13.7
5-state CV	7.1	4.5	4.1	3.6	3.6	3.5	3.5	3.4
Montana	21.1	18.8	17.8	9.8	7.9	6.9	7.8	7.9
South Dakota	60.0	43.1	25.0	40.3	38.1	26.5	25.7	25.0
2-state CV	22.4	17.7	14.6	18.1	13.4	7.8	8.2	8.2
7-state CV	6.8	4.6	4.1	3.9	3.6	3.2	3.2	3.2
<i>Spring wheat</i>								
Minnesota				12.2	13.0	11.6	9.9	9.5
Montana				37.2	18.0	12.2	10.3	10.2
North Dakota				10.7	5.7	5.0	4.4	4.4
South Dakota				40.4	13.4	13.1	11.6	9.6
4-state CV				9.2	4.8	4.2	3.6	3.5
<i>Total wheat</i>								
9-state CV				3.4	2.6	2.5	2.4	2.4

TABLE B-V.—Coefficient of Variation for the Revised LACIE Production Estimates for Phase III, by State and by Monthly Report, U.S. Great Plains, CAS Annual Report 05, December 22, 1977

(Percent)

State	CAS monthly report							
	Feb.	May	June	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>								
Colorado	28.0	22.3	20.3	19.8	18.6	17.9	17.7	17.7
Kansas	18.4	12.5	11.5	10.7	10.8	10.7	10.5	10.5
Nebraska	18.1	15.4	14.3	15.0	13.9	13.1	13.4	13.1
Oklahoma	16.7	15.9	14.1	12.7	12.3	12.4	12.9	12.9
Texas	20.2	16.5	14.7	14.0	14.9	14.9	15.7	15.7
5-state CV	9.7	7.4	6.8	6.5	6.4	6.4	6.4	6.3
Montana	30.4	23.1	22.0	15.5	14.4	13.9	14.4	14.4
South Dakota	61.6	46.2	30.8	43.9	41.8	31.9	31.3	30.7
2-state CV	27.9	20.8	18.2	20.7	16.8	12.9	13.1	13.1
7-state CV	9.3	7.1	6.5	6.4	6.1	5.8	5.9	5.8
<i>Spring wheat</i>								
Minnesota				16.1	16.3	15.1	13.9	13.6
Montana				40.0	22.7	18.6	17.4	17.3
North Dakota				16.1	13.7	13.1	13.1	13.0
South Dakota				41.9	17.7	17.5	16.4	15.0
4-state CV				12.1	9.6	9.1	9.1	9.0
<i>Total wheat</i>								
9-state CV				5.0	4.9	4.8	4.9	4.8

TABLE B-VI.—Coefficient of Variation for the Revised LACIE Yield Estimates for Phase III, by State and by Monthly Report, U.S. Great Plains, CAS Annual Report 05, December 22, 1977

[Percent]

State	CAS monthly report							
	Feb.	May	June	July	Aug.	Sept.	Oct.	End of season
<i>Winter wheat</i>								
Colorado	18.9	17.4	15.2	14.8	14.8	14.8	14.8	14.8
Kansas	12.1	10.8	10.0	9.7	9.7	9.7	9.7	9.7
Nebraska	10.5	10.8	10.7	9.7	9.5	9.3	9.3	9.3
Oklahoma	13.8	12.5	11.0	10.7	10.3	10.2	10.4	10.4
Texas	16.5	11.6	10.6	10.8	11.3	11.3	11.7	11.7
5-state CV	6.7	6.1	5.6	5.6	5.6	5.6	5.6	5.6
Montana	22.5	13.7	13.2	12.1	12.1	12.1	12.1	12.1
South Dakota	17.5	18.6	18.6	18.9	18.5	18.5	18.5	18.5
2-state CV	16.3	11.1	10.7	10.1	9.9	10.2	10.2	10.2
7-state CV	6.3	5.5	5.1	5.1	5.1	5.1	5.1	5.1
<i>Spring wheat</i>								
Minnesota				12.8	11.6	11.2	10.8	10.7
Montana				14.9	14.0	14.0	14.0	14.0
North Dakota				15.1	12.8	12.3	12.4	12.4
South Dakota				12.1	11.6	11.6	11.6	11.6
4-state CV				10.5	8.6	8.3	8.5	8.4
<i>Total wheat</i>								
9-state CV				3.9	3.9	4.2	4.3	4.3

TABLE B-VII.—Real-Time LACIE Area Estimates for Phase III, by Crop Type, by State, and by Monthly Report, U.S. Great Plains

(Thousands of acres)

State	CAS monthly report									
	Feb. ^a	Feb. ^b	Apr. 22	May	June	July	Aug. ^c	Sept. ^d	Oct. ^e	End of season ^f
<i>Winter wheat</i>										
Colorado	2 183	2 135	2 189	3 093	3 065	2 962	3 059	3 059	3 432	3 459
Kansas	6 719	6 491	6 794	10 190	10 915	11 764	12 385	12 501	12 669	12 494
Nebraska	2 977	2 892	3 072	3 169	3 610	3 475	3 423	3 105	3 325	3 433
Oklahoma	2 953	2 943	3 061	4 506	4 875	5 264	5 543	6 074	5 950	5 675
Texas	2 954	3 294	3 517	4 262	4 529	4 511	4 311	4 513	4 581	4 476
5-state total ^g	17 786	17 755	18 633	25 220	26 994	27 976	28 721	29 252	29 957	29 337
Montana	2 763	2 274	2 274	2 973	3 253	3 097	2 746	3 597	3 416	3 371
South Dakota	1 044	1 721	1 721	2 261	2 601	4 629	1 353	1 039	963	912
2-state total ^g	3 807	3 995	3 995	5 234	5 854	7 726	4 099	4 636	4 379	4 283
7-state total ^g	21 594	21 750	22 627	30 453	32 848	35 701	32 819	33 888	34 336	33 820
<i>Spring wheat</i>										
Minnesota							2 238	2 461	2 289	2 344
Montana							1 369	2 187	2 150	2 174
North Dakota							6 761	8 678	9 173	9 183
South Dakota							2 167	2 160	1 909	1 936
4-state total ^g							12 535	15 487	15 522	15 638
<i>Total wheat</i>										
9-state total ^g							45 355	49 375	49 857	49 458

^aCAS monthly report released February 8, 1977; results based on Phase II sampling strategy with 431 sample segments allocated; used data through December 1976

^bCAS monthly report released April 6, 1977; results based on 601-sample-segment Phase III strategy; used data through December 1976; duplicated the February CMR using Phase III sampling strategy.

^cThe results contained in the August 1977 CMR were obtained by redesignating the Montana and South Dakota segments as spring only, winter only, and mixed. In addition, the Landsat data were "thresholded" to eliminate early-season data. See CMR-29, August 10, 1977, for more details.

^dThe results in the September 1977 CMR were obtained by "thresholding" and "screening" the Landsat data to eliminate early-season data and data that were significantly different from historically similar data. See CMR-31, September 9, 1977, for more details.

^eThe results in the October 1977 CMR were obtained by thresholding and screening plus a reallocation of the Phase III segments based on wheat rather than small grains. Thirty-eight segments were dropped. See CMR-33, October 11, 1977, for more details.

^fThe end-of-season results were obtained by screening, thresholding, segment redesignation, and reallocation as described in footnotes c, d, and e above.

^gTotals may not add correctly because of rounding.

TABLE B-VIII.—Real-Time LACIE Production Estimates for Phase III, by Crop Type, by State, and by Monthly Report, U.S. Great Plains

(Thousands of bushels)

State	CAS monthly report									
	Feb. ^a	Feb. ^b	Apr. 22	May	June	July	Aug. ^c	Sept. ^d	Oct. ^e	End of season ^f
<i>Winter wheat</i>										
Colorado	49 772	48 659	49 037	70 357	72 456	66 516	68 682	68 675	77 070	77 666
Kansas	194 220	187 644	190 941	286 373	308 387	339 348	357 263	360 616	365 465	360 410
Nebraska	90 058	88 444	96 579	99 038	108 793	111 903	109 960	99 264	106 120	109 823
Oklahoma	64 391	63 918	64 413	95 560	96 550	104 907	110 463	121 671	119 208	113 387
Texas	56 726	63 305	63 516	83 068	91 965	91 691	87 579	91 594	92 885	90 695
5-state total ^g	455 167	451 970	464 486	634 396	678 151	714 365	733 947	741 820	760 748	751 981
Montana	73 799	60 723	65 712	85 751	91 417	81 983	72 678	95 206	90 411	89 224
South Dakota	28 513	46 978	46 057	58 836	67 685	123 196	36 621	28 130	26 072	24 682
2-state total ^g	102 312	107 701	111 769	144 587	159 102	205 179	109 299	123 336	116 483	113 906
7-state total ^g	557 479	559 672	576 255	778 982	837 254	919 544	843 247	865 156	877 231	865 888
<i>Spring wheat</i>										
Minnesota							71 199	78 744	73 213	74 955
Montana							24 634	39 357	38 683	39 112
North Dakota							157 751	200 529	211 247	211 990
South Dakota							45 103	44 969	39 748	40 309
4-state total ^g							298 686	363 599	362 890	366 367
<i>Total wheat</i>										
9-state total ^g							1 141 933	1 228 754	1 240 121	1 232 255

^aCAS monthly report released February 8, 1977; results based on Phase II sampling strategy with 431 sample segments allocated; used data through December 1976

^bCAS monthly report released April 6, 1977; results based on 601-sample-segment Phase III strategy; used data through December 1976; duplicated the February CMR using Phase III sampling strategy.

^cThe results contained in the August 1977 CMR were obtained by redesignating the Montana and South Dakota segments as spring only, winter only, and mixed. In addition, the Landsat data were "thresholded" to eliminate early-season data. See CMR-29, August 10, 1977, for more details.

^dThe results in the September 1977 CMR were obtained by "thresholding" and "screening" the Landsat data to eliminate early-season data and data that were significantly different from historically similar data. See CMR-31, September 9, 1977, for more details.

^eThe results in the October 1977 CMR were obtained by thresholding and screening plus a reallocation of the Phase III segments based on wheat rather than small grains. Thirty-eight segments were dropped. See CMR-33, October 11, 1977, for more details.

^fThe end-of-season results were obtained by screening, thresholding, segment redesignation, and reallocation as described in footnotes c, d, and e above.

^gTotals may not add correctly because of rounding.

TABLE B-IX.—Real-Time LACIE Yield Estimates for Phase III, by Crop Type, by State, and by Monthly Report, U.S. Great Plains

[Bushels per acre]

State	CAS monthly report									
	Feb. ^a	Feb. ^b	Apr. 22	May	June	July	Aug. ^c	Sept. ^d	Oct. ^e	End of season ^f
<i>Winter wheat</i>										
Colorado	22.8	22.8	22.4	22.8	23.6	22.5	22.5	22.5	22.5	22.5
Kansas	28.9	28.9	28.1	28.1	28.3	28.8	28.8	28.8	28.8	28.8
Nebraska	30.2	30.6	31.4	31.1	30.1	32.2	32.1	32.0	31.9	32.0
Oklahoma	21.8	21.7	21.0	21.2	19.8	19.9	19.9	20.0	20.0	20.0
Texas	19.2	19.2	18.1	19.5	20.3	20.3	20.3	20.3	20.3	20.3
5-state average	25.6	25.5	24.9	25.2	25.1	25.5	25.6	25.4	25.4	25.5
Montana	26.7	26.7	28.9	28.8	28.1	26.5	26.5	26.5	26.5	26.5
South Dakota	27.3	27.3	26.8	26.0	26.0	26.6	27.1	27.1	27.1	27.1
2-state average	26.9	27.0	28.0	27.6	27.2	26.6	26.7	26.6	26.6	26.6
7-state average	25.8	25.7	25.5	25.6	25.5	25.8	25.7	25.5	25.5	25.6
<i>Spring wheat</i>										
Minnesota							31.8	32.0	32.0	32.0
Montana							18.0	18.0	18.0	18.0
North Dakota							23.3	23.1	23.0	23.1
South Dakota							20.8	20.8	20.8	20.8
4-state average							23.8	23.5	23.4	23.4
<i>Total wheat</i>										
9-state average							25.2	24.9	24.9	24.9

^aCAS monthly report released February 8, 1977, results based on Phase II sampling strategy with 431 sample segments allocated, used data through December 1976

^bCAS monthly report released April 6, 1977, results based on 601-sample-segment Phase III strategy, used data through December 1976, duplicated the February CMR using Phase III sampling strategy

^cThe results contained in the August 1977 CMR were obtained by redesignating the Montana and South Dakota segments as spring only, winter only, and mixed. In addition, the Landsat data were "thresholded" to eliminate early-season data. See CMR-29, August 10, 1977, for more details

^dThe results in the September 1977 CMR were obtained by "thresholding" and "screening" the Landsat data to eliminate early-season data and data that were significantly different from historically similar data. See CMR-31, September 9, 1977, for more details

^eThe results in the October 1977 CMR were obtained by thresholding and screening plus a reallocation of the Phase III segments based on wheat rather than small grains. Thirty-eight segments were dropped. See CMR-33, October 11, 1977, for more details

^fThe end-of-season results were obtained by screening, thresholding, segment redesignation, and reallocation as described in footnotes c, d, and e above

LACIE Area, Yield, and Production Estimate Characteristics: U.S.S.R.

J. R. Hickman^a

OVERVIEW

The U.S.S.R. is one of the major food producers of the world; however, because of the climatic differences and fluctuations in the vast area involved (20° to 140° E longitude and 40° to 60° N latitude), the environmental limitations often render the country incapable of maintaining self-sufficiency on a yearly basis. Although the U.S.S.R. is the world's largest producer of wheat (the principal food grain used for human consumption) and caloric intake supplied by grains is declining, the U.S.S.R. is a net importer of this commodity in almost as many years as it has a surplus.

The Soviet position regarding wheat supply and demand is not unique, but four aspects of the supply picture make the wheat position of this country a dominant factor in the world market for grains because of the magnitude of U.S.S.R. import requirements in years of non-self-sufficiency. The four parts, all covered by the umbrella of nonavailability of information, are (1) stock position at any given time, (2) crop condition during the growing season, (3) total production for the year, and (4) magnitude of the total import requirement in years of short supply. The sensitivity of the market to the U.S.S.R. wheat situation and the varied climatic, topographical, agronomic, and cultural features encountered in the wheat-growing area made the U.S.S.R. a natural choice for LACIE's first attempt at crop estimation outside North America.

The objective of this paper is to discuss production, area, and yield estimates in the U.S.S.R. during Phases II and III of LACIE (no estimates were generated in Phase I). For Phases II and III, the following topics are discussed.

1. Scope
2. Sampling strategy
3. Data base

4. Landsat data
5. Yield analysis for winter wheat and spring wheat
6. Area and production estimates for winter wheat and spring wheat
7. Technical issues and problems

In addition, the methods of selecting data (area estimates) in Phase III are discussed; and winter wheat, spring wheat, and total wheat estimates are compared. The accuracy of the winter wheat, spring wheat, and total wheat production, area, and yield estimates for Phase III is discussed.

PHASE I: CROP YEAR 1974-75

The U.S.S.R. work in Phase I centered around constructing an initial historical-statistical data base, locating sample segments within the country, and acquiring multispectral scanner (MSS) data for a subset of the sample for study by the image analyst. No estimates were generated for the U.S.S.R. during this phase.

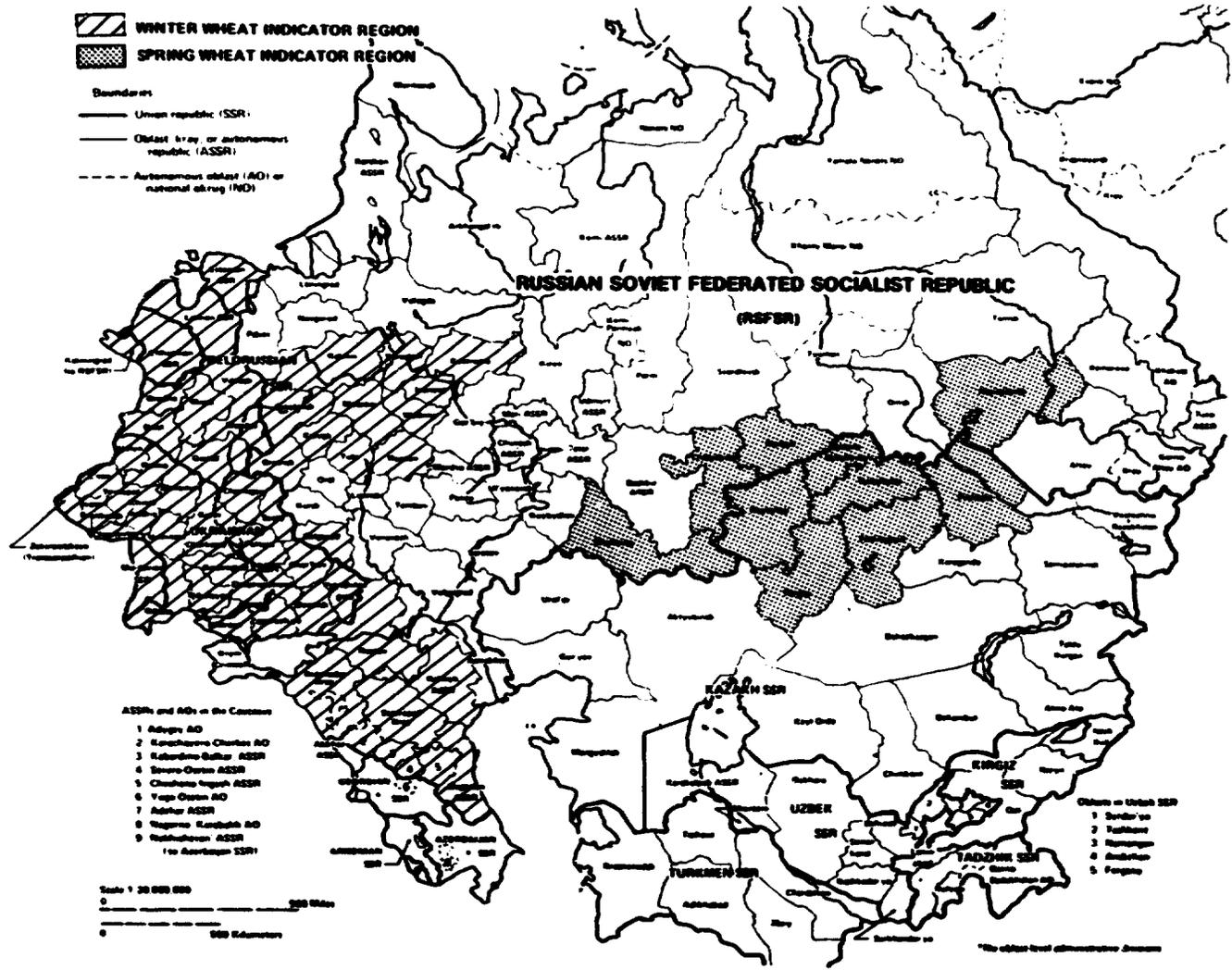
PHASE II: CROP YEAR 1975-76

Scope

The LACIE Phase II effort in the U.S.S.R. was limited to two indicator regions, one in the winter wheat area and one in the spring wheat area (fig. 1).

The winter wheat indicator region includes the Baltics, Belorussia, the Ukraine, Moldavia, the northern Caucasus, central non-Chernozem, and the oblasts of Belgorod and Kalmyk. This area can be expected to give a fairly clear indication of the entire winter wheat situation because it contributes 70.3

^aUSDA Foreign Agricultural Service, Houston, Texas.



REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

FIGURE 1.—U.S.S.R. spring and winter wheat indicator regions.

percent to total area and 82.3 percent to total production, according to 1971 historical data. The percent contribution of area versus production indicates the inclusion of a preponderance of the higher yielding areas in the indicator region. This does not appear to be significant because an examination of additional years of data indicates this is the norm rather than the exception.

The spring wheat indicator region is composed of 10 oblasts—Orenburg, Chelyabinsk, Kurgan, Kustanay, Turgay, Tselinograd, Kokchetav, Severo-Kazakhstan, Pavlodar, and Novosibirsk. These oblasts (except Orenburg) were good testing sites for the LACIE technology because such factors as the short growing season, marginal precipitation for wheat production, the potential for the occurrence of early frost, and other potential wheat problems made them high-risk areas. According to 1971 statistics, the spring wheat indicator region contributed 39.7 percent to the total spring wheat area and 37.2 percent to the total production. The percent contribution of area versus production indicates a slight preponderance of the more marginal yielding areas in this region.

Strategy

The sampling strategy used in LACIE, based on total area and wheat density, allocated a total of 1947 sample segments (5- by 6-nautical-mile samples) for the U.S.S.R. These segments were randomly located in the wheat-producing areas. The winter wheat indicator region contained 385 sample segments; the spring wheat region had 362 segments.

Data Base

The data base can be divided into four principal categories—allocation data, historical-statistical data (crop production, area, and yield), yield data, and spectral data.

Allocation data.—Allocation data include hierarchical identifiers; i.e., codes for country (UR), region (economic region), zone (group of oblasts), and stratum (oblast). This part of the data base also identifies each sample segment (winter wheat or spring wheat) and its position in the hierarchy or its geographical location.

Historical statistical data.—Historical statistical

data include historical production, area, and yield statistics for both spring and winter wheat at the stratum (oblast) level and for each higher level of the hierarchy. Also included in this category are data on derived ratios of wheat to other small grains. The need for this portion of the data base manifested itself early in Phase II (fall 1975) when it became evident that the image analysts (photograph interpreters working with hard-copy MSS data) were unable to separate wheat from rye or barley. Therefore, the classification of the MSS data was for small grains rather than for spring or winter wheat.

The ratios established were for (1) winter wheat to fall-sown small grains (rye, barley, and winter wheat), (2) winter wheat to total small grains (fall-sown rye and barley; spring-sown barley; and oats, spring wheat, and winter wheat), (3) spring wheat to spring-sown small grains (oats, barley, and spring wheat), and (4) spring wheat to total small grains (winter wheat; winter-sown rye and barley; and spring-sown oats, barley, and spring wheat).

Yield data.—The yield data were received from the National Oceanic and Atmospheric Administration (NOAA) Center for Climatic and Environmental Assessment (CCEA) at Columbia, Missouri. The spring and winter wheat yield estimates were then input at the stratum level (oblast) as appropriate; i.e., winter yield, spring yield, or both depending on the class (or classes) of wheat historically produced within a given stratum.

Spectral data.—For each usable acquisition (data of such quality that the classification procedures would yield a reasonable and acceptable estimate of Landsat data for a sample segment), the following information was entered in the data base: segment number; crop year; percentage of winter, spring, and total small grains; crop type; biological growth stage; Landsat acquisition date; and classification date.

Landsat Data

The quantity of the usable MSS data acquired for the U.S.S.R. is very poor compared to that of the U.S. data. The usable Landsat data acquired for the two countries during Phase II are compared in table I.

Although MSS data were acquired in the U.S.S.R. winter wheat area over a much longer period of time (from August 1975 to August 1976), the acquisition rate in the spring wheat area (data acquired from May to August 1976) was significantly above that for the winter wheat. This difference is due to the

TABLE I.—Comparison of Usable MSS Data Acquired for the U.S.S.R. and the U.S. Great Plains in Phase II

Data	U.S.S.R.		United States	
	Winter wheat indicator region	Spring wheat indicator region	Total wheat indicator region	Seven states, U.S. Great Plains
No. of segments allocated	385	362	747	601
No. of usable acquisitions	690	801	1491	1599
Acquisition rate ^a	1.8	2.2	2.0	2.7

^aNumber of acquisitions divided by number of segments allocated

weather conditions in European U.S.S.R. The fall is overcast much of the time, and spring and early summer rains with their accompanying cloud cover limit the gathering of data by Landsat. The acquisition rate for the United States is much improved over the U.S.S.R. rate because (1) more usable data are acquired from fall planting to harvest over the entire area, (2) cloud cover, haze, and other atmospheric interferences are minor compared to similar influences in European U.S.S.R., and (3) snow covers larger areas for longer time periods in the U.S.S.R.

Five reports (crop estimates) were generated in Phase II for the winter wheat indicator region and three for the spring wheat indicator region. The spectral data (number of segment data) used for each report are summarized in table II.

Winter Wheat Indicator Region

General weather/crop situation.—The 1975-76 winter wheat crop started slowly because of subsoil and topsoil moisture shortages at the time of planting. Emerged stands were thin, spotty, and vulnerable to the cold weather that followed. Winterkill took its toll when a February cold snap affected wheat fields lacking adequate snow cover protection. The most damaged areas included the northern and eastern Ukraine and the lower Volga region.

Precipitation amounts remained very light throughout the fall and winter with the exception of a heavy snowfall in January. Spring rains were timely and stimulated vigorous growth during April and

May. Precipitation amounts over the European U.S.S.R. remained generous throughout the growing season, with occasionally threatening floods in the Dnieper River Valley.

Temperatures were generally near normal during the growing season but turned cooler than normal as the wheat crop neared maturity. Continued shower activity over most of the winter wheat area caused concern about the potential damage caused by lodging and sprouting and about the difficulty in getting harvesting equipment into the fields.

Warm and sunny weather in late August allowed harvesting operations to near completion over most of the winter wheat region, only slightly behind the normal harvest calendar date.

Ideal postdormancy growing conditions produced above-normal crop yields from a near loss the previous fall. In contrast, soil moisture conditions for the 1976 planting season (Phase III) were considered ideal and prompted the planting of a record number of hectares to winter grain in the fall of 1976.

Yield analysis.—As a result of the soil moisture situation during the fall of 1975 and the follow-on winterkill problems, the winter wheat yield in the U.S.S.R. began the 1976 season on a down note. April yield model estimates based only on pre-season pre-

TABLE II.—Spectral Data Used in Phase II Reports

Month of report	No. of segments allocated	No. of segment data used in report	Percent of allocations covered
<i>Winter wheat indicator region</i>			
April	385	146	38
May ^a	385	146	38
June	385	197	51
July	385	248	64
August	385	(b)	
September	385	(b)	
October	385	285	74
<i>Spring wheat indicator region</i>			
August	362	265	73
September	362	296	82
October	362	314	87

^aThis report updated the estimate for yield and production only because no new spectral data were received after the May report.

^bWinter wheat reports were not generated this month because of the effort involved in the compilation of the spring wheat data base.

precipitation and temperature predicted yields to be normal or above normal in only 5 of 18 winter wheat crop regions. Yields steadily improved, however, beginning in May and continuing through to harvest. By the final yield model truncation in July, 9 of the 18 winter wheat crop regions registered above-normal yield indications. In contrast to the subnormal winter moisture conditions, heavy May rainfall proved detrimental to yield in the more northerly areas, notably in the Baltics (Estonia, Latvia, Lithuania, and Kaliningrad) and Belorussia. Farther south along the Black Sea in the lower portion of the Ukraine and Krasnodar, lack of April precipitation held yields down. A combination of preseason moisture stress and subnormal temperatures through March led to severe winterkill problems, primarily in the eastern Ukraine and low- Volga regions, where subnormal temperatures lowered prospective yields an average of 5 quintals per hectare.

Winter wheat yields throughout the remainder of the central and western European U.S.S.R. fared exceptionally well in 1976, especially in the prime production regions of the Caucasus and middle Volga area, where output per hectare ranged up to 5 quintals above normal. Yield model calculations indicate that Moldavia, although small in terms of total winter wheat production, ranked above all other Soviet regions in 1976—38.0 quintals per hectare, or 10 percent above normal expectations.

Spring Wheat Indicator Region

General weather/crop situation.—Because of a lingering drought, the central U.S.S.R. spring wheat region had, at harvest time in 1975, drawn soil moisture reserves down to a critically low level. Much hope was therefore given to a plentiful snowfall during the winter months. Winter and early spring moisture, however, was spotty, with precipitation amounts generally less than normal.

Rain and snowfall rates continued to be low during April, and although May brought increased precipitation to most spring wheat areas, dry pockets were widening in parts of the Ural Mountains, Kazakhstan, and western Siberia. Timely early summer rains really turned spring crop prospects around and replenished dwindling moisture supplies in most of the New Lands.

Summer temperatures, averaging 1° to 2° C below normal north of Kazakhstan, aided crop development. Cooler temperatures as harvest approached

slowed maturity in western Siberia, leaving crops vulnerable to an early frost; however, no damage materialized. Harvest weather was most favorable in late August and September, although seasonal showers interfered with harvesting operations in the eastern portion of the spring wheat area.

Yield analysis.—Similar to the yield estimates for the 1976 winter wheat crop, the spring wheat yield indications were below normal at the May planting time because of adverse preseason precipitation conditions. Yield estimates tended to support reported soil moisture conditions; all seven yield models in the spring wheat indicator region indicated below-normal prospects. Areas to the west along the Volga River in the mixed spring and winter grain regions gave somewhat higher yield predictions throughout the year, although unseasonably abundant rainfall detracted from yields in the later stages as harvest approached. At the end of the season, about one-third of the spring wheat model strata estimated above-normal yields, all but one stratum lying outside the indicator region.

The overall seasonal impact of weather on predicted Soviet yields was somewhat mixed. Within the indicator area surrounding the New Lands, yield model estimates were off as much as 5 or 6 quintals per hectare from normal trends in western Siberia primarily because of severe moisture problems occurring through April. Conversely, the yields in Kokchetav and Severo-Kazakhstan oblasts were nearly 2 quintals (or 16 percent) above trend because cumulative precipitation from May through July exceeded normal precipitation by 60 millimeters. This region posted the highest 1976 spring wheat yield in the Soviet New Lands—12.4 quintals per hectare, followed by the near-normal 10.5-quintal-per-hectare yield in the northeastern Urals sector.

Estimates

Winter wheat indicator region.—Five reports were generated for the winter wheat indicator region in the 1975-76 crop year. The area and production estimates showed consistent significant increases with each succeeding report, except for the May report (fig. 2). No additional spectral data were available from those used in the April report; therefore, the May area estimate remained the same as the previous estimate. However, yield estimates were reduced, resulting in a decrease in the May production estimate from the April level. The decreases in

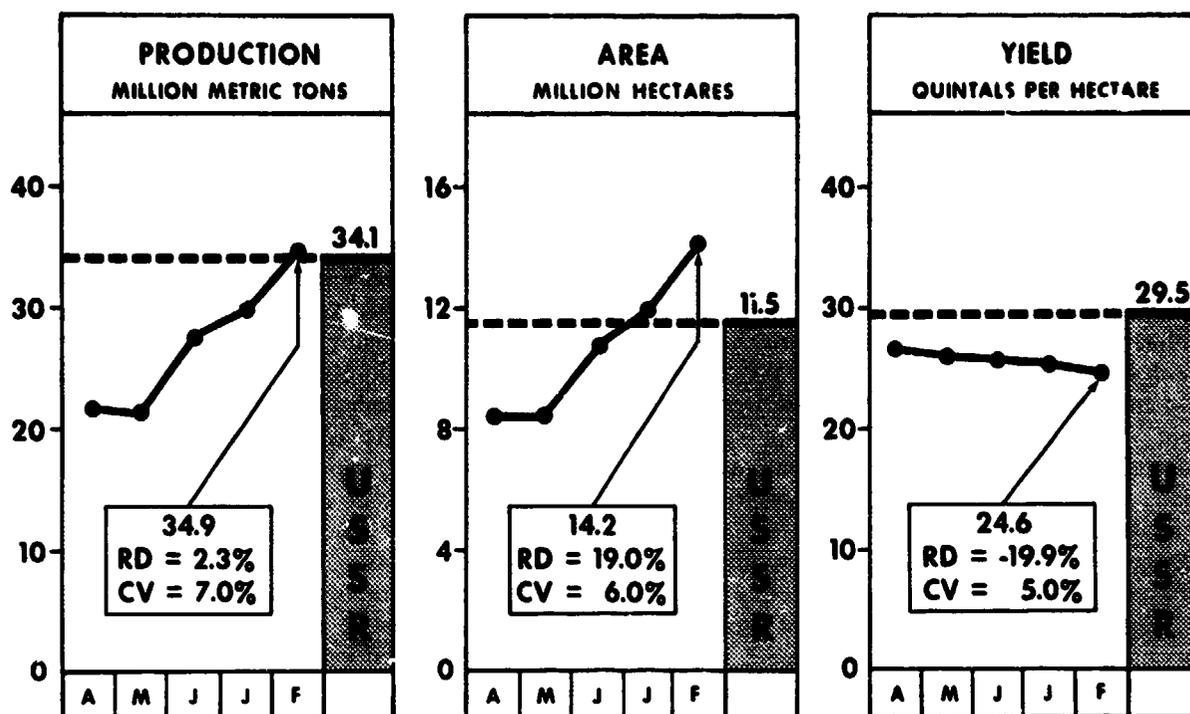


FIGURE 2.—Phase II production, area, and yield estimates of the U.S.S.R. winter wheat indicator region. The U.S.S.R. actual figures are derived from multiple Soviet publications because the U.S.S.R. does not publish production, area, and yield statistics for an area coincident with the indicator regions.

the yield estimates were not uniform for the region; therefore, the reduction in the production estimate was not directly proportional to the decrease in the average yield between the 2 months.

The continued increases in area estimates resulted from improved sample segment coverage and improved classifications of previously worked segments using later acquisitions; however, much of the increase was directly attributable to an inability to separate wheat from other small grains and indirectly to the ratioing technique used. LACIE did not envision the necessity of using a ratio, but, as the experiment progressed, it became apparent that the capability of distinguishing wheat from other small grains in foreign areas was not developed within LACIE. The software for the aggregation was designed specifically for wheat so that a ratio seemed the logical solution to the problem. A ratioing scheme was developed, using an averaging methodology, which appears to function satisfactorily in normal years but which must be modified to perform well in years such as 1975-76 for the U.S.S.R. winter wheat area. Several intervening fac-

tors that year adversely affected the use of ratios in developing percentage of wheat from percentage of small grains; e.g., changes in sown area and abnormal amount of winterkill. In retrospect, it is evident that two major classification errors were made. First, fall-sown small grains were classified when in reality spring-sown small grains made up at least part of the confusion crops. Second, spring-sown or total small grains were classified when crops other than small grains made up a part of the confusion crops. The principal agronomic abnormalities for that year were (1) the expansion of the area devoted to small grains in the winter wheat area and (2) an abnormally large area usually occupied by winter wheat replanted to other small grains when losses from poor germination resulted because of a dry fall and severe winterkill in certain areas of the country.

Spring wheat indicator region.—Three reports were generated for the spring wheat indicator region for the 1975-76 crop year. The area and production estimates showed consistent and expected increases with each succeeding report (fig. 3). Apparently, the ratioing technique was more nearly suited to the

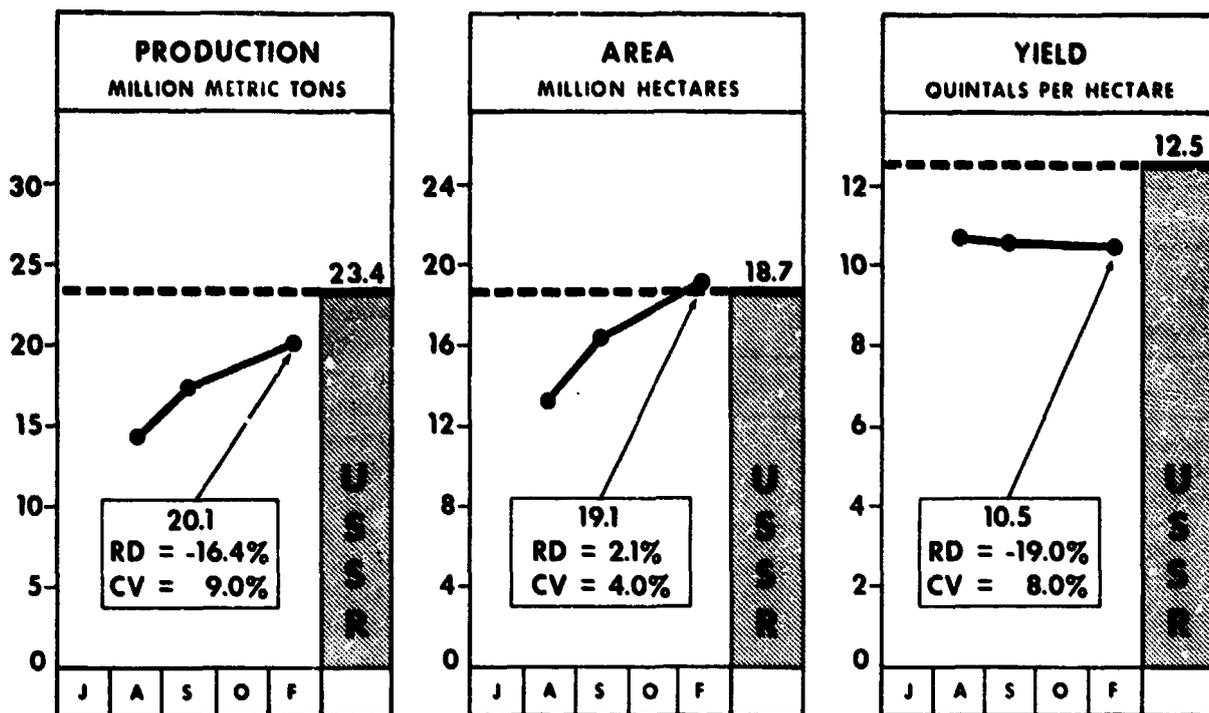


FIGURE 3.—Phase II production, area, and yield estimates of the U.S.S.R. spring wheat indicator region. The U.S.S.R. actual figures are derived from multiple Soviet publications because the U.S.S.R. does not publish production, area, and yield statistics for an area coincident with the indicator regions.

overall spring wheat area since no significant across-the-board abnormalities occurred during this crop year. This was largely due to the preponderance of spring wheat to potential confusion crops (i.e., other small grains) in the spring wheat indicator region. The principal problem with the estimate for the spring wheat indicator region for this year was the timing of the first report. It was anticipated that this report would be released no later than mid-July. However, because of a late season and the amount of time between acquisition and receipt of data by the Crop Assessment Subsystem (CAS), it was mid-August before sufficient data were available for substantive results.

Accuracy of Estimates

Because of the reporting system used by the U.S.S.R., estimation of the bias of estimates at the indicator region level is not possible. However, estimates of precision—coefficient of variation (CV)—can be projected to the country level to determine whether the precision is sufficient to support ac-

curacy goals. The CV of the production estimate at the country level must be less than 6.1 percent to support the LACIE 90/90 accuracy goal (see the paper by Houston et al. entitled "Accuracy Assessment: The Statistical Approach to Performance Evaluation in LACIE"). Table III gives the LACIE production estimates for the two indicator regions with estimated CV's and projections to the country level for those months for which standard statistics were available (ref. 1). These projections, which treat the country-level estimates as having the same characteristics as the indicator region estimates, indicate that the precision at the indicator region levels is more than adequate to support the 90/90 accuracy goal. In fact, these projections indicate that a relative bias in the country-level LACIE production estimator as large as 4 percent can be tolerated and still support the accuracy goal.

The tendency to underestimate winter wheat area at harvest, as was observed in the U.S. Great Plains (Texas, Oklahoma, Kansas, Colorado, Nebraska, and South Dakota) in Phase II, is not indicated in the U.S.S.R., even though the winter wheat signatures in

TABLE III.—LACIE Production Estimates for the Winter and Spring Wheat Indicator Regions

Month	LACIE production estimate, MMT ^a	CV, percent	CV, ^b percent
<i>Winter wheat indicator region</i>			
June	27.8	7.0	3.1
July	30.0	8.0	3.6
October	34.9	7.0	3.1
<i>Spring wheat indicator region</i>			
August	14.3	11.0	4.7
September	17.4	9.0	3.9
October	20.1	9.0	3.9

^aMillion metric tons.

^bCountry-level projection.

the U.S.S.R. were much like those observed in the United States. If anything, a tendency to overestimate is indicated. For some of the drier areas in the U.S.S.R. where the wheat signatures were very weak, late-fall and early-season interpretations of winter grains were difficult because it was hard to determine whether anything was actually growing in many of the fields. This probably led the analysts to confuse natural vegetation and wheat in these areas and hence to tend to overestimate. Spring wheat signatures for the U.S.S.R. did not appear as strong as those for the United States; however, the spring wheat fields in the U.S.S.R. were much larger. Indications are that the spring wheat area estimates at the segment level were much better than those in the United States because of fewer confusion crops, minimal strip-fallow cropping practices, larger fields, and more stable year-to-year ratios of spring wheat to small grains (resulting from more stringent government controls).

Although the LACIE yield estimates did not appear to vary much at the indicator region levels, previous discussion indicates that the yield varied considerably from crop region to crop region in both the winter and spring wheat regions because of weather conditions. The LACIE yield models apparently tracked this variability reasonably well but tended to underestimate. This variability illustrates the need to track both area and yield at the crop region level in order to obtain reliable indicator region and higher level production estimates.

Technical Issues and Problems

Selection of the indicator regions.—Because of resource limitations, it was impossible to work the U.S.S.R. wheat-producing area in its entirety during Phase II. Therefore, it was decided to work indicator regions in the winter and spring wheat areas (see fig. 1) that would be representative of the respective areas. The selection of regions was indeed representative. The winter wheat indicator region contained approximately 70 percent of the area producing winter wheat, which accounts for approximately 82 percent of the production. The spring wheat indicator region contained only 40 percent of the total area devoted to spring wheat, which accounts for approximately 37 percent of the total production but covers the major part of the high-risk or "swing" area producing spring wheat. The unfortunate aspect of this selection is that the U.S.S.R. does not release production, area, and yield data in a manner that allows comparison with the LACIE estimate. From the standpoint of accuracy assessment, a better winter wheat indicator region might have included the Baltics, Belorussia, and the Ukraine to coincide with U.S.S.R. releases.

Sampling.—The U.S.S.R. wheat sampling density was based on total area and wheat density in order to achieve the LACIE 90/90 goal (to be within 10 percent of the actual crop 90 percent of the time on the average). However, the sampling strategy was based on 1971 data that were not fully representative of the 1975 situation; i.e., shifts from wheat to barley and increases in cultivated areas were not reflected. This resulted in an oversample of some areas and an undersample of other areas. The sample consisted of 5-by 6-nautical-mile sample segments randomly placed in agricultural areas as determined by interpretation of Landsat images. Landsat data, however, were not available over the entire area. Where spectral data were missing, historical data were used; however, the historical data again were not truly representative of the current situation and resulted in approximately 35 percent of the sample segments being placed in nonagricultural locations.

Historical data.—Much of the initial planning for the LACIE procedures was based on the United States, and the inconsistencies or lack of needed historical-statistical data in foreign areas necessitated numerous "workarounds" in the initial stages of Phase II for the U.S.S.R. Although the U.S.S.R. reports of statistical data on agriculture are massive, an attempt to develop specific data for the entire coun-

try or to establish relatively long-term trends (5 to 10 years) for specific values usually leads to frustration and to the derivation of data. Across-the-board statistical data at the oblast level are normally not available, and if available, are usually inconsistent. Thus, data for some oblasts may be reported as winter grains and spring grains or total grains, whereas others may define wheat (winter and spring), rye, and barley (winter and spring).

Another constraint in working U.S.S.R. data is that most official releases below the country level are at least 2 years old. As described previously, the sub-country data are inconsistent, and considerable work had to be done to derive representative data for the needed political level, the oblast (no attempt was made to work at the county (rayon) level). It is impossible to analyze the results in the U.S.S.R. in a timely manner because complete data (production, area, and yield estimates) are not released by the U.S.S.R. during the crop year and because production, area, and yield data at the country level are not released until the spring following the crop year of interest.

Landsat collection.—The acquisition rate of usable spectral data for the European U.S.S.R. was 26 percent lower than the rate for the United States (see table I) and Canada because of haze and cloud cover in the fall, snow cover in the winter, and rainy weather in the spring. The weather is a major constraint in an accurate inventory in that it is quite possible that only one acquisition will be obtained during an entire crop year for a given sample segment. If, for example, the acquisition is obtained early in the crop year (in the September-October time frame) and the ground cover is not sufficient for accurate evaluation and classification, the resulting estimate will only reflect a part of the actual crop. This estimate will be carried for the entire crop year and will result in a definite downward bias.

The spring wheat acquisition rate for the New Lands is better except for July. This is not normally a critical time for data collection in this area because the key biostages (emergence, tillering, jointing, heading, and turning) generally occur outside this time frame. However, if the crop is either delayed or accelerated significantly, acquisitions for this time would be crucial for an accurate analysis.

Classification of Landsat data.—One of the most critical issues for the LACIE was the ability to distinguish between crops using the Landsat data. With key acquisitions, it was possible to separate small grains from row crops, hay, and improved pastures,

but LACIE never achieved the capability of distinguishing wheat from other small grains. This inability led to the ratioing technique discussed earlier in this paper. The ratioing technique leaves much to be desired, especially in the U.S.S.R. where the age and completeness of available data make it difficult, if not impossible, to reflect the current situation.

In retrospect, and from a commodity analyst's point of view, it might have been a simpler and more accurate operation if the project had chosen to classify small grains and if the crop analyst had been free to develop a ratioing scheme at the country or regional level. This approach could possibly have allowed more time for research and development on crop separability techniques by the classification component.

Early-season estimates.—The image analysts' procedures during Phase II dictated that only an area showing emergence or growth of small grains be so classified. Therefore, if weather conditions interrupted the seeding of large areas and plant growth was observed in only parts of fields at the time of Landsat data acquisition, the resulting area classification would be low. If no subsequent acquisitions for that sample segment were obtained, this bias would remain in the system throughout the crop year.

Yield estimates.—Yield estimates were obtained from mathematical regression yield models operated by NOAA CCEA at Columbia, Missouri. These models were developed using monthly climatic historical yield data. Deviations from normal temperature and normal precipitation (or a combination of the two as expressed in terms of potential and actual evapotranspiration) result in additions or subtractions from a predicted trend yield. Models requiring monthly accumulations of meteorological data are limited in their capability to respond to abrupt weather extremes, which can rapidly alter the condition of the crop over short periods during critical stages of crop development. The regression-type yield model is also somewhat limited in its capability to respond fully to the impact of either abnormally good or abnormally bad years. Awareness of this limitation makes good analyst judgment paramount when using yield model predictions. Experience during Phase II has shown that regression yield models constantly responded in the proper direction to the weather phenomenon experienced. Trend-term adjustments and some reselection of weather variables, it was later shown, improved the predictive capability of the U.S.S.R. yield models after Phase II. Since more sophisticated phenological modeling type

techniques were not developed for operational application, no alternative to using the regression yield model was available during LACIE.

Cultural practices.—Two cropping practices used in the U.S.S.R. complicated the analysis of its wheat situation. The more significant of the two is the cutting of small grains for "green chop" (harvesting at heading or the soft dough growth stage for animal feed). Therefore, if the last spectral data received for a given area were at the jointing stage and a third of the small grains crop was cut for "green chop," the LACIE final estimate would be biased upward. Although analytical work has not been done on this section, it is suspected that it is a major factor in the LACIE at-harvest estimate of winter wheat area because the practice is more prevalent in the winter-wheat-producing area.

The second practice is overseeding. Poor stands of winter-sown grains (resulting from poor germination, winterkill, or related factors) are normally overseeded; i.e., another grain (normally barley) is sown over the existing crop. Thus, the result is a wheat-barley or rye-barley mix at harvest. If it is a wheat-barley mix and the farm is short in delivering its wheat quota, the mix can be delivered as wheat at a discount depending on the percentage of barley in the wheat. If this type of operation is relatively widespread, it will bias the yield downward because barley yield is normally lower than wheat.

U.S.S.R. reporting of statistical agricultural data.— Besides the delay and inconsistency in reporting, the U.S.S.R.'s report of "bunker weights" is also a problem. Bunker weight is the weight straight from the field with no corrections made for dockage or moisture content. In years of favorable climatic conditions, these "bunker weights" could be fairly representative of the nutritive or feeding value of the harvest; however, in years of poor weather conditions, these values may be inflated unrealistically because of high moisture content, trash, sprouting, disease, etc.

PHASE III: CROP YEAR 1976-77

Scope

The level of activity for the U.S.S.R. was increased from indicator regions in Phase II to total country coverage in Phase III. This coverage automatically increased the segment workload from 747 to 1947

segments, a significant impact on resources. However, the additional workload generated an estimate for which comparable data (official U.S.S.R. estimates) were released by the U.S.S.R. after harvest.

Sampling

In Phase II, the sampling strategy was based on total area and wheat density. The Phase III sampling strategy had the advantage of updated and more accurate historical-statistical data, more and improved spectral imagery of the entire U.S.S.R., and the experience gained in Phase II to use as a base. The sampling strategy was revised for Phase III and was based on agricultural (cropland) area and wheat density. Additionally, a significant number of sample segments in nonagricultural areas during Phase II were relocated to agricultural areas for Phase III.

Data Base

The data base for Phase III was not changed significantly from the Phase II edition, with the exception of the historical-statistical section where updates and corrections were made.

Landsat Data

During LACIE Phase III, the Landsat data for the U.S.S.R. were acquired between August 5, 1976, and September 15, 1977. During the period from August 5, 1976, to November 1, 1977, 8838 acquisitions (imagery of a given sample segment) were examined by the Classification and Mensuration Subsystem (CAMS) and the results were reported to the CAS. This number of acquisitions equates to an average acquisition rate (number of acquisitions divided by number of sample segments) of 4.54 as the project attempted to collect usable data on the total allocation of 1947 sample segments. Although the number of acquisitions (8838) is almost overwhelming in magnitude and may lead one to believe that usable data were collected on 100 percent of the sample segments allocated to the U.S.S.R., the actual coverage of usable imagery amounted to only 78 percent for the entire season. In other words, of the 1947 sample segments allocated to the U.S.S.R., usable data were collected on 1518 segments, or 78 percent of the

allocation. The disposition of the remaining 7320 acquisitions is accounted for by one of the following.

1. The image was not processed because of clouds or image data problems.
2. The image was examined but not classified because of small grains preemergence or dormancy.
3. Several acquisitions of the same segment were available at the time it was worked. The analyst had the option of selecting either the best or the most representative acquisition to process; he would then archive the remainder.
4. The acquisition was processed, but the results were unsatisfactory.
5. Multiple acquisitions resulted in satisfactory classifications for the same sample segment.

Two severe unanticipated data collection constraints—loss of the Pakistani ground station for data collection and failure to acquire early spring spectral data—curtailed the LACIE data collection effort at critical times during the crop year. Undoubtedly, both the coverage and the acquisition rate would have improved if the Landsat data had been collected as originally planned.

The Landsat acquisition history for the U.S.S.R. in Phase III (from August 5, 1976, to November 1, 1977) is summarized in table IV. The early-season, midseason, and late-season periods have been arbitrarily designated to facilitate the preparation of this table. Large areas of the wheat-producing regions were combined by predominant wheat class to further expedite the data compilation. The inclusive dates for early-season, midseason, and late-season acquisitions as they relate to areas defined on the map in figure 4 are as follows:

Area	Wheat class	Season		
		Early	Mid	Late
I	Winter	Aug. 1976 to Mar. 31, 1977	Apr. to May 31, 1977	June 1 to Oct. 31, 1977
II	Winter/ spring	Aug. 1976 to Mar. 31, 1977	Apr. 1 to May 31, 1977	June 1 to Oct. 31, 1977
III	Spring	Apr. 15 to June 30, 1977	July 1 to July 31, 1977	Aug. 1 to Oct. 31, 1977

The exceptions to this legend are Chimkent, Dzhambul, Alma Ata, and Taldy Kurgan in Area I. The acquisition seasons (early, mid, and late) are divided the same as Area III because of the similarity not only in cultural practices but also in acquisition dates.

The sample segment coverage by acquisitions containing usable spectral data was the same for all three

areas (see table IV). This, of course, is an abnormal situation because spectral data are acquired for Areas I and II more than twice as long as for Area III; thus, the possibility of acquiring equal coverage in Area III under normal circumstances is remote. The factors driving this skewed coverage are as follows:

1. A wet fall and an early winter in the winter wheat areas with cloud cover much of the time and early dormancy
2. Failure to acquire spectral data over European U.S.S.R. at the prescribed time in the spring (The first acquisitions were requested for March but actual collection was not begun until about May 1.)
3. A wet spring and summer in European U.S.S.R. with accompanying clouds
4. Unusually favorable climatic conditions for spectral acquisitions over the spring wheat area; i.e., insignificant rainfall in June and most of July and minimal cloud cover

These conditions complicated the analysis of imagery for the European U.S.S.R. because most of the early-season data were acquired extremely early in the season (August/September 1976) and in many cases reflected partial estimates. This led to the assumption that the very early acquisitions did not reflect the actual extent of the area devoted to fall-sown small grains.

The spectral data coverage of the New Lands was much more straightforward; acquisitions were timely and the analysis of the spectral data was an improvement over that of the early-season winter small grains. (The procedures were adjusted to avoid the problems encountered in the analysis of early-season fall-sown grains.)

Spectral coverage during the season differed by wheat class (see table IV). The most complete coverage occurred in the early season in the spring and winter areas, the next best coverage occurred in midseason, and the late-season coverage decreased rather sharply. The coverage of the spring and winter wheat or mixed area showed the opposite situation; coverage in the early season and midseason was approximately equal at one-third of the allocation and improved to one-half in the late season.

The spectral coverage for the three areas (see fig. 4) ranged between 64 and 86 percent for the entire year, except for the Transcaucasus (12 percent), central Asia (43 percent), and the Northwest (0 percent). The central Asia problem was due to a failure of the Pakistani ground station. This station was used to relay most of the spectral data acquired over central Asia in an effort to economize on use of the

TABLE IV.—Phase III U.S.S.R. Landsat Data Review

(Percentage of allocated sample segments for which usable spectral data were received, by crop season and wheat class)

Region	Sample segment allocation	Percent coverage ^a			
		Entire crop year	Season ^b		
			Early	Mid	Late
<i>Winter wheat^b</i>					
1. Baltics	25	64	12	36	24
2. Belorussia	42	79	67	71	26
3. Southwestern Ukraine	117	81	40	54	37
4. Donets-Dnieper	130	72	60	32	40
5. Southern Ukraine	70	94	83	77	13
6. Moldavia	16	94	88	69	38
7. Northern Caucasus	153	83	47	37	59
15. Transcaucasus	26	12	6	7	2
16. Central Asia	23	43	13	26	30
Total	602	78	52	47	38
<i>Winter and spring wheat^c</i>					
8. Central Chernozem	101	69	30	12	47
9. Central non-Chernozem	103	75	17	37	51
15. Volga Vyatsk	34	74	29	24	47
16. Volga	325	83	43	39	55
Total	563	78	35	33	52
<i>Spring wheat</i>					
12. Urals	141	91	70	48	42
13. Kazakhstan	436	71	42	39	20
14. Western Siberia	158	86	59	42	20
17. Eastern Siberia	36	72	19	67	31
18. Far East	10	80	60	20	70
19. Northwest	1	0	0	0	0
Total	782	78	50	42	25
Country total	1947	78	46	41	37

^aThe figure for the early, mid, and late season coverages will not add to crop-year coverage because data were acquired for the same segment in all three seasons in many cases.

^bInclusive dates for early, mid, and late seasons are, for winter wheat, from Aug. 1976 to May 31, 1977, for winter and spring wheat (mixed), from Aug. 1976 to May 31, 1977, and for spring wheat, from Apr. 15 to July 31, 1977.

^cPredominantly mixed area.

satellite recorder. When the station malfunctioned, much of the data for this area was lost. The Transcaucasus problem was more complicated. In a plot of the areas for which the U.S.S.R. spectral data could be "dumped" to the Pakistani and/or Italian station, the Transcaucasus was located in an overlap

area: i.e., both stations would receive the spectral data for this area. As Phase III progressed, it was determined that the strength of the stations was weaker than anticipated and that no data or at best sporadic bits of data were being received for the Transcaucasus. The Northwest coverage was yet

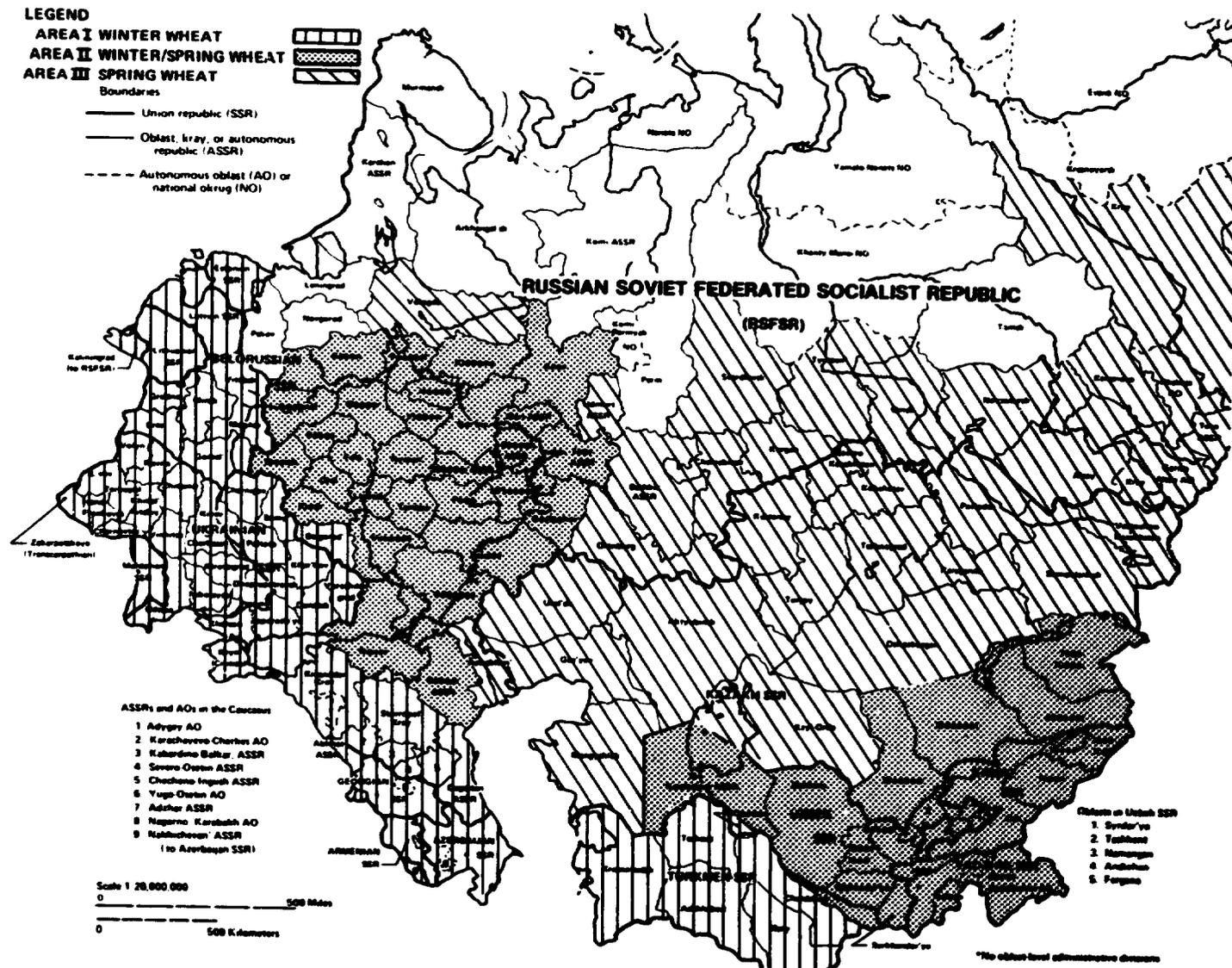


FIGURE 4.—Principal wheat-producing areas of the U.S.S.R. by predominant wheat class.

another problem; cloud cover, snow cover, and the segment's location on the fringe of the collection area combined to negate efforts to acquire usable data for this area.

The experience in Phases II and III led the U.S.S.R. CAS analyst to believe that a 30-percent coverage of the allocation (Phase III allocation of 1947 sample segments) of good usable data (classification) is sufficient to produce a reliable estimate. This assumption is supported in part by project statisticians' determination that the U.S.S.R. was oversampled by approximately 75 percent. The revised sampling developed for the U.S.S.R. to meet the 90/90 criterion of production at harvest allocated approximately 1100 sample segments. The 1100-segment allocation includes a cloud-cover factor based on Phase III acquisition history. Thus, the revised sampling strategy sample segment allocation exceeds the commodity analyst's spectral data acquisition requirements almost twofold.

Yield Analysis

Winter wheat.—The final 1976-77 LACIE winter wheat yield prediction for the U.S.S.R. was 25.6 quintals per hectare, an increase of 1 quintal over the LACIE Phase II (1975-76) crop year estimate of 24.6 quintals covering a reduced indicator region. This is especially significant when one considers that the Phase II coverage consisted mostly of the higher yielding portion of the winter wheat area. The official U.S.S.R. countrywide yield figure for the 1975-76 crop year was 25.9 quintals per hectare for winter wheat. The highest previous Soviet yield was 27.0 quintals, a record set in 1973; the winter wheat yield average since 1970 had been 23.0 quintals.

Individual yield predictions from the 21 crop regions in the U.S.S.R. (fig. 5) are aggregated to obtain the countrywide estimate. Winter wheat yield predictions from the agrometeorological regression yield models used in these crop regions terminate with the July truncation, and any revisions to the combined country estimate since that time are due to area adjustment only.

The improvement in the overall country yield estimate for the 1976-77 crop year over the previous year is reflected in the individual yield stratum results. More than three-fourths of the crop regions with operating yield models in 1976 registered higher yields in 1977. The greatest impact from this increase was in the Ukraine, where 45 percent of the U.S.S.R.

winter wheat crop is produced. Yields in the southern and eastern Ukraine were more than 6 quintals higher than in 1976—the result of milder-than-average winter temperatures and abundant rainfall through the critical spring and early summer periods. The only setback in 1977, compared with a year earlier, occurred along the middle and upper Volga. Here, above-average April temperatures and scanty May rainfall kept yields as much as 7 quintals below the 1976 figure. Winter wheat yields in 1977, compared to trend, also indicated a bumper year. Two-thirds of the 21 crop regions predicted yields above the normal, whereas, in 1976, only slightly more than half the strata were forecasting yields to exceed trend predictions.

Country-level winter wheat yield estimates for the U.S.S.R. appeared to be close to other predictions, official and unofficial, during Phases II and III. No exact check can be performed, however, on stratum-level accuracy because of a 2- to 3-year lag in official publication of regional data. Individually or on a model-by-model basis, the Caucasus-Volga winter wheat covariance model, covering adjacent Crop Regions 10 and 17, may be somewhat suspect. For the 1975-76 and 1976-77 crop years, the model's yields from the lower Volga (Crop Region 17) were consistently below trend throughout the season, whereas predictions for the northeastern Caucasus (Crop Region 10) were above the trend with equal frequency. This could, of course, be entirely possible given the proper meteorological conditions.

Building for a covariance model requires pooling of yield and climatic data over both regions for the data base period from 1958 to 1971. In the operation of the yield model, current area-specific weather is then applied to the individual crop regions. In pooling data for Crop Regions 10 and 17 for the November to March and the April temperature variables, average normal temperatures of -2.98° and 9.17° C, respectively, are indicated. Actual temperature data for the November to March period show the more southerly Crop Region 10 to be above the pooled average, whereas Crop Region 17, lying to the north, deviates below the pooled average. This is not surprising because the north-south span encompassing the two strata covers nearly 500 miles and because temperature averages would normally be expected to be colder inland and farther north (table V). Yields within the two strata are affected accordingly: Crop Region 10 upward because of the above-normal temperatures and Crop Region 17 downward. The April temperature variable reaction in the

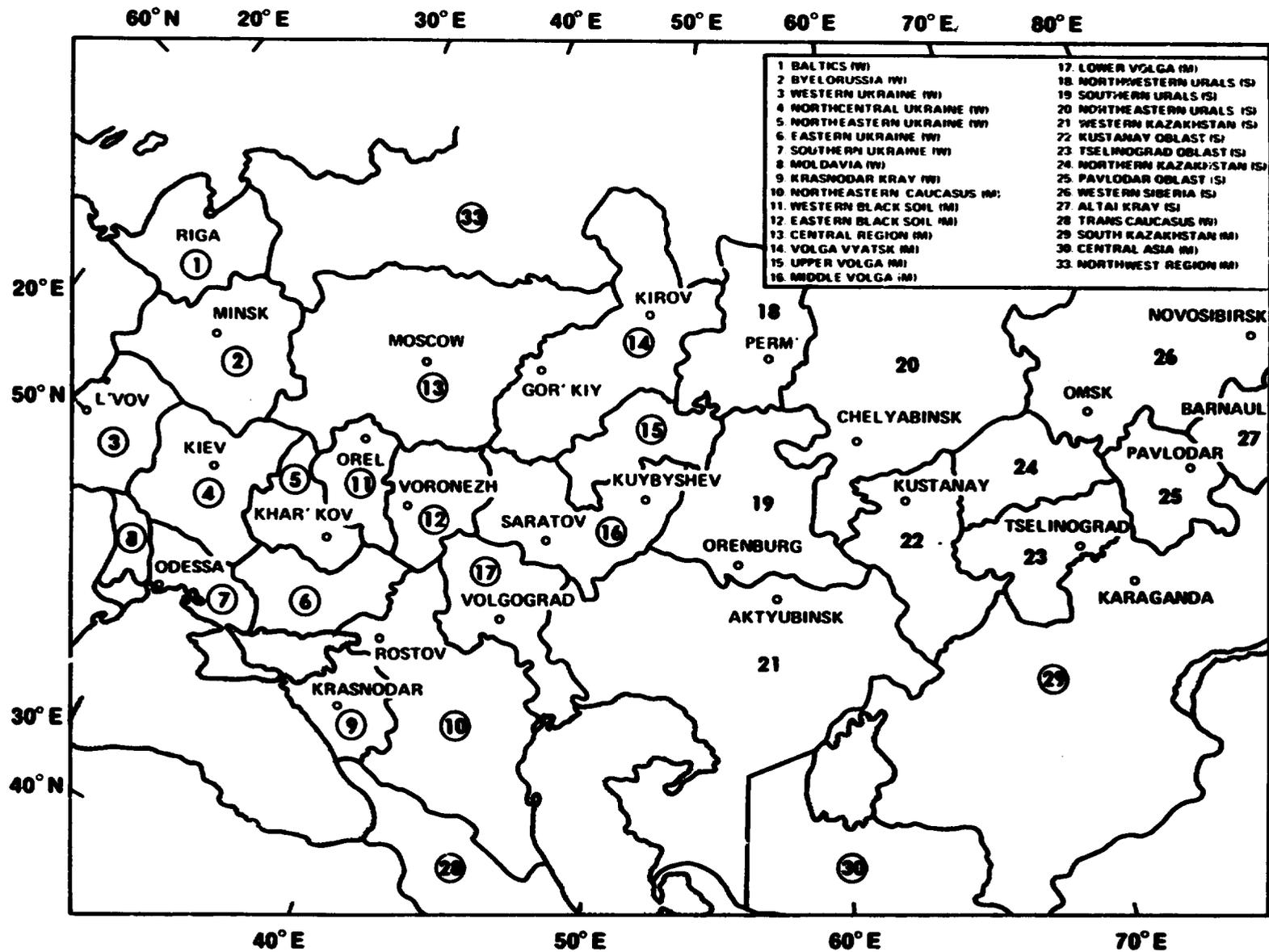


FIGURE 5.—U.S.S.R. winter wheat yield model strata (indicated by circled crop region numbers): W is winter, S is spring, and M is mixed.

TABLE V.—Temperature Variables of the Caucasus-Volga Winter Wheat Yield Model

Crop region	November-March temperature				April temperature				Total temperature variable impact, ql/ha
	Norm	Actual, °C	Deviation from norm	Yield impact, ql/ha	Norm	Actual, °C	Deviation from norm	Yield impact, ql/ha	
1975-76									
10. Northwestern Caucasus	-2.80	-2.0	+0.8	+1.1	+9.40	+11.7	+2.3	+0.8	+1.9
17. Lower Volga	-2.80	-5.4	-2.6	-5.6	+9.40	+9.8	+4	+2	-5.4
1976-77									
10. Northeastern Caucasus	-2.98	-0.3	+2.7	+1.3	-9.17	+11.7	+2.5	+1.1	+2.4
17. Lower Volga	-2.98	-4.2	-1.2	-1.1	-9.17	+10.4	+1.2	+6	-5

model shows effects similar to the November to March period although yield impact is less for the single-month variable. One discernible effect of the temperature pooling over the two regions is that it somewhat destroys the use of trend as a meaningful analytical tool in judging model performance. A region-constant correction of -2.5 quintals applied in 1977 to Crop Region 10 appears to bring the trend yield into comparable range with a straight average yield over the 1958-71 data base period (table VI). Crop Region 17 yield, though, shows a wide divergence between the trend term value as shown in the model and the actual yield average over the same period. Model trend deviations would indicate below-normal yields in 1976 and 1977; however, yields were actually above the 14-year data base average for both years.

In Phase III, one other slight discrepancy was also noted in Crop Region 10, the northeastern Caucasus. Smoothing the November to March temperature over the entire crop region produced an above-average reading of 2.7° C. This subsequently contributed 1.3 quintals to the yield estimate for that period. However, this region was particularly vulnerable to winterkill on January 4 and 5 when temperatures were sufficiently low and the area lacked protective snow cover. Above-normal temperatures generally throughout the 5-month November to March period tended to obscure 2 days of critically low temperatures. This probably underscores even more the need for a good climatic alarm system rather than indicating a yield model defect. Yield monitoring in the Caucasus, where potential win-

terkill situations frequently arise, is particularly crucial because the region is a prime wheat producer and contributes heavily to overall production. The north-

TABLE VI.—Northeastern Caucasus and Lower Volga Historical Winter Wheat Yields

Parameter	Northeastern Caucasus (Crop Region 10)	Lower Volga (Crop Region 17)
5-yr (1960-64) average, ql/ha	15.2	13.6
5-yr (1965-69) average, ql/ha	15.0	14.3
5-yr (1970-74) average, ql/ha	19.9	19.1
10-yr (1965-74) average, ql/ha	17.4	17.4
15-yr (1960-74) average, ql/ha	16.6	16.0
1958-71 average, ^a ql/ha	15.8	14.8
LACIE trend yield, ^b ql/ha:		
For 1976	15.1	20.1
For 1977	15.6	18.2
LACIE final yields, ql/ha:		
For 1967	17.6	16.0
For 1977	20.2	17.0
LACIE deviation from trend, percent:		
For 1976	16.6	-20.4
For 1977	29.5	-6.7
LACIE deviation from 1958-71 average, percent:		
For 1976	11.4	8.1
For 1977	27.8	14.9

^aThe data conform to the same 14 years of yield data used in model development
^bFinal truncation. Model developers assumed trend to be zero over the data base period, however, a region constant correction of -5.0 ql/ha in 1976 and -2.5 ql/ha in 1977 is applied to Crop Region 10

eastern Caucasus is the single largest winter-wheat-producing crop region, contributing nearly 15 percent to the total U.S.S.R. winter wheat crop over the past 5 years.

Spring wheat.—The final LACIE yield prediction for 1977 U.S.S.R. spring wheat leveled at 8.8 quintals per hectare. This estimate was considerably below the 1976 official near-record output of 12.4 quintals. LACIE in that year predicted 10.5 quintals—nearly 2 quintals less than the final U.S.S.R. figure. In 1976, LACIE covered only a reduced spring wheat indicator region lying primarily east of the Ural Mountains. On a strictly comparative basis, the 1977 estimate would be even lower relative to the previous season because the normally higher yielding regions in the west along the Volga River are included in the 1977 estimate. Extremes in U.S.S.R. spring wheat yields in this decade ranged from a high of 13.5 quintals in 1973 down to 7.0 quintals in 1975. The average since 1970 was 11.5 quintals. The 1977 LACIE estimate was lower than for any recent year except the disastrous 1975 Soviet spring crop.

The 16 U.S.S.R. yield models covering the 21 spring wheat crop regions (fig. 6) showed a wide range of predictions for the 1977 season. Yields exceeding 20 quintals per hectare were noted along the western fringes of the spring wheat area in the Black Soil Region. By comparison, a yield of less than 1 quintal was predicted in the more arid central Asia section, where a deficit in precipitation (more than 150 millimeters) minus potential evapotranspiration (PET) during the spring dropped the forecast yield more than 7 quintals below normal expectations. This crop region is relatively insignificant by comparison, producing on an average less than 1 percent of the total U.S.S.R. spring wheat output.

In contrast to the improvement shown in 1977 winter wheat yields compared with those of 1976, the U.S.S.R. spring crop regions predicted a reduction when judged against the previous year's above-average crop. Most of the reduction occurred in Kazakhstan and nearby oblasts in the southern Ural Mountains region (Orenburg and Bashkir), where yield predictions on the average were off by more than 4 quintals compared to the 1976 figures. Above-normal April temperatures were particularly harsh in Crop Regions 22 to 25, centered in northeastern Kazakhstan. Yields in that one month alone were discounted an even 2 quintals across the board. Precipitation minus PET deficits in May and June prevented any subsequent recovery. This four-region

sector alone accounts for 20 percent of the U.S.S.R. spring wheat annual production.

Reduced crop prospects in the central New Lands were somewhat offset by slightly higher predicted yields in the northeastern Urals and western Siberia. The Altai Kray 1977 yield equaled its year-earlier mark of 7.3 quintals. With the exception of the Volga Valley, other peripheral spring-wheat-producing areas to the west and north fared better in 1977. The Volga Valley spring wheat yield declined slightly from its 1976 level because of a pre-season moisture deficit; however, the impact was generally less on spring wheat than was noted for winter wheat yields.

Comparisons of 1977 spring wheat yields to trend also tended to support indications of a deterioration in yield compared with 1976. Exactly half of the 1976 crop regions predicted yields below the normal, whereas, in 1977, more than 60 percent of the 23 spring wheat yield strata were forecasting yields to be less than trend predictions.

No direct check of yield model accuracy can be performed at this time on individual yield models because official Soviet data at the regional level for the 1977 season will not be available until 1979 or 1980. During Phase II, spring wheat models were predicting low—the LACIE indicator region estimate was nearly 2 quintals under the official Soviet figure. An official estimate by the U.S.S.R. Government indicates that, at the country level, the LACIE Phase III yield is 0.9 quintal per hectare low. One feature of the northeastern Kazakhstan and Siberia-Altai spring wheat covariance models might tend to bias the LACIE predictions on the low side. For these two models, covering the all-important Crop Regions 22 to 25 and 26 and 27, respectively, the July precipitation-minus-PET variable contains only the squared departure from normal term with a negative coefficient. (Any combination of precipitation and temperature that produces anything other than a normal precipitation-minus-PET value will detract from yield.) The only way to prevent a dropoff in July yield is for precipitation minus PET to be exactly "normal," in which case its contribution to yield is zero.

One observed feature of the U.S.S.R. spring wheat yield models is the seemingly large influence on monthly yield change caused by the trend and weather coefficients themselves, notably with the July truncation. As a result of July 1977 weather in particular, 14 yield models showed a decrease in their yield prediction, whereas 5 registered gains.

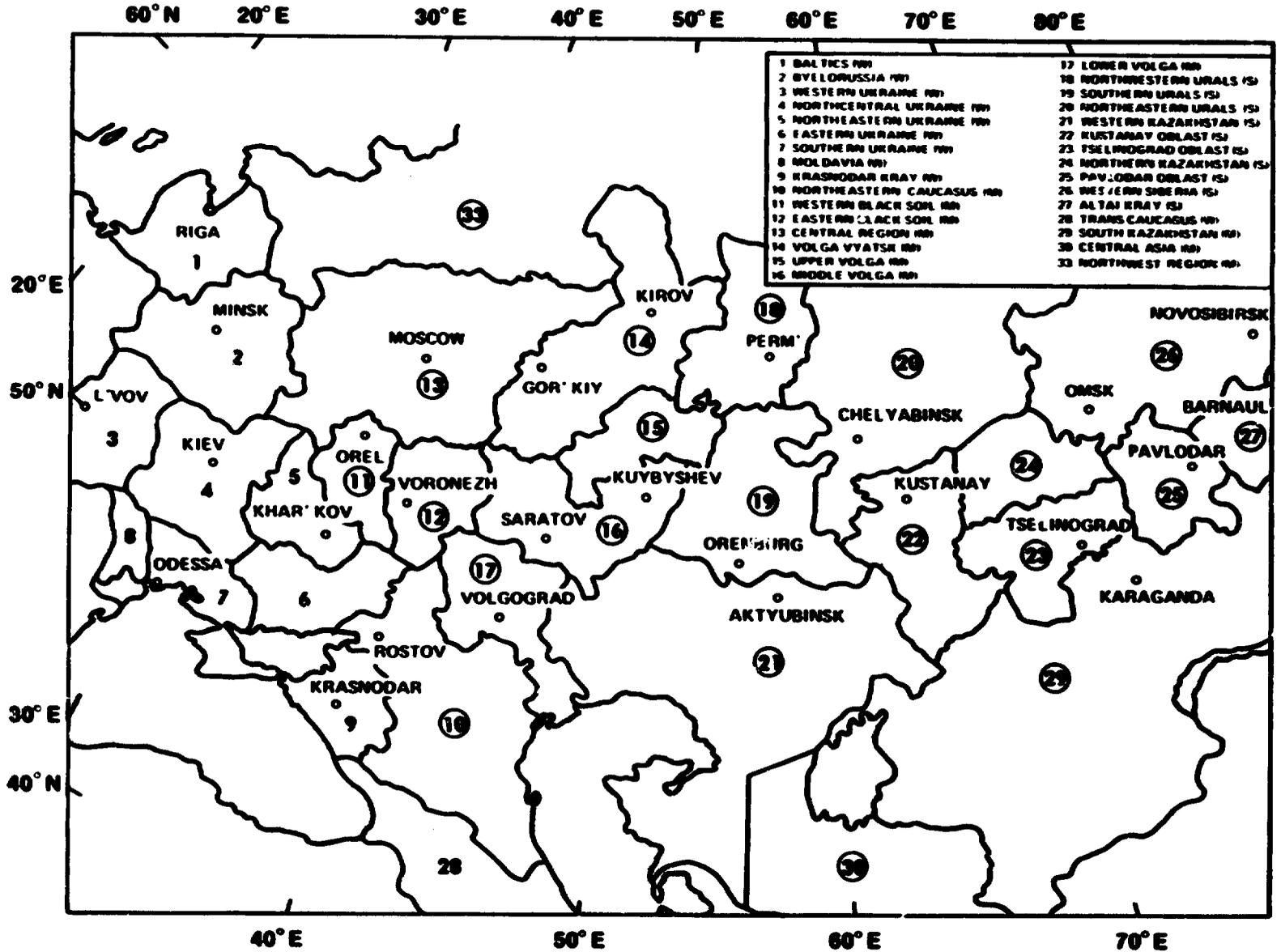


FIGURE 6.—U.S.S.R. spring wheat yield model strata (indicated by circled crop region numbers): W is winter, S is spring, and M is mixed.

However, coefficient adjustments (both trend and weather) between the June and July truncations allowed the July weather variable to be overridden in 10 of the 14 instances. It is particularly confusing to the user when analyzing the effects of weather to find that, on the whole, yield-reducing weather conditions produced overall gains in yield. Table VII provides a breakdown of this type of occurrence by month through the 1977 season for both spring and winter wheat. The greatest frequency is associated with the July spring wheat truncation. July is a particularly crucial month for the Soviet spring wheat crop, and conflicting yield information could possibly lead to questions of yield model credibility. It would seem more desirable to hold all internal adjustments in the model constant and let the weather variables alone shift the yield predictions.

Area Estimates

Although the LACIE area estimates for the U.S.S.R. spring wheat followed predictable patterns and yielded acceptable results, the winter wheat estimates for the most part were unrealistic to the commodity analysts because they tended to be much higher than historical information could substantiate. The principal components of the area estimate are sampling strategy, aggregation formulation, ratio-

ing techniques, and spectral data classification. The sampling strategy and aggregation formulation are considered adequate, and the ratioing techniques and data are the best available, given the currency and accuracy of Soviet agricultural statistics that are required inputs to sampling and aggregation. This opened the possibility of misclassification of spectral data. The following paragraphs address two objective methods of data selection or editing and resultant estimates compared to initial procedures (accepting all spectral classification: at face value).

Methods of data selection.—Three methods of data selection (initial, revised, and final) were used to produce separate sets of country estimates. This section explains the background and development of each method.

Initial method: The initial method of aggregation used the entire population of CAMS aggregatable estimates; i.e., the pure CAMS data were used and no data selection procedure was applied. It should also be pointed out that no "dummy" estimates were input.¹

Revised method: The initial early-season estimates tended to be unrealistically low because a portion of the early-season imagery was acquired before all the wheat had emerged. To reduce the number of partial-emergence area estimates that were generated, the CAS analysts determined the tillering date for each oblast from information obtained through Soviet newspapers and available meteorological data. Area estimates made before the specified tillering dates were thresholded since these data were acquired before full emergence and the data were not representative of the total planted area. Historical data for 1971 were input into the system to cover those areas for which Landsat data were not immediately available. The 1971 data were replaced as Landsat data became available later in the season.

The LACIE winter wheat area estimates in Phase III followed the pattern set in Phase II; i.e., the best estimate was obtained in the June-July time frame but continued to escalate with the receipt of additional spectral data. This escalation occurred pri-

TABLE VII.—Spring and Winter Wheat Yield Models Direction of Response of Yield Prediction to Weather Variables

Factor ^a	Month of truncation (1977)					
	Mar.	Apr.	May	June	July	Aug. Sept.
<i>Spring wheat</i>						
CHG (+) WX (+)		8	7	4	1	1
CHG (-) WX (-)		3	14	5	3	
CHG (+) WX (-)		4	1	10	1	
CHG (-) WX (+)		1	0	1	1	
<i>Winter wheat</i>						
CHG (+) WX (+)	7	11	7	7	1	
CHG (-) WX (-)	3	3	3	2	1	
CHG (+) WX (-)	0	6	2	0		
CHG (-) WX (+)	2	0	1	3		

^aCHG is total yield change, WX is weather variable change.

¹The Phase III CAS software was designed so that at least one stratum (oblast) within a zone (group of oblasts) had an aggregatable classification for a minimum of three sample segments before an estimate could be generated for the zone. To resolve this problem, analysts used historical data as dummy CAMS estimates to generate an estimate for the zone in the revised and final estimating methodology. The zone estimate was the actual historical area figure for the 1971 base year.

marily because the first acquisition of Landsat data for many segments was acquired during biostage 4 or 5 (jointing/heading and soft dough). The escalation is shown in the June, July, and August estimates (fig. 7). The difference between the initial and revised estimates in figure 7 is the result of using biostage 4 and 5 estimates without a previous classification to verify the classification as small grains. At this time of the year, crops such as hay, potatoes, sugar beets, and sunflowers were in direct competition with small grains with respect to spectral detection. In many (if not most) instances, it was impossible, given the current state of the art, to differentiate between small grains and hay or row crops under these circumstances.

The ratioing technique used by LACIE was developed to reduce the winter small grains estimate passed by CAMS to a winter wheat estimate for use in the aggregation. Although these ratios were satisfactory when winter small grains were identifiable on the imagery, this process presented an almost insurmountable problem during the spring and summer when hay and row crops had been planted. If no previous fall or early spring acquisition had been passed for temporal separation of late spring and summer data, the current ratioing technique could not compensate for the increased confusion due to hay or row crops and therefore led to an upward spiral in the succeeding estimates. To correct for this inability to differentiate between small grains and other crops using single acquisitions from specific time periods during the growing season, the commodity analysts used the revised method of data selection. The following criteria were used.

1. If, for a given segment, no acquisition was obtained for biostage 3 (jointing), then acquisitions obtained for biostages 4 (heading) and 5 (soft dough) would not be used for aggregation until they could be processed multitemporally with an acquisition obtained for either biostage 6 (hard dough) or 7 (harvest).

2. If, for a given segment, an acquisition was obtained only for biostage 6 (hard dough) or 7 (harvest) (i.e., no acquisitions were acquired for biostages 3, 4, or 5), then these data would be excluded from the estimate.

Final method: The final aggregation technique was developed in November and December of 1977 as the annual report for the U.S.S.R. was being produced. Close study of the revised approach revealed that, although it resulted in a reasonable country estimate, it distorted some regional estimates to

unreasonable levels and did not use all segment data that were available.

To understand the winter wheat overestimation problems encountered by LACIE, one must understand the plant growth cycle in the U.S.S.R. and the type and amount of image data collected. In the U.S.S.R., winter grain growth in the fall is usually such that significant tillering does not occur in all growing areas before dormancy. The grains are difficult to detect on satellite imagery without sufficient ground cover, which usually occurs at the tillering stage. Therefore, imagery of the spring green-up period is essential for an accurate winter grain estimate because confusion with other crops is minimal.

Multitemporal acquisitions are required to identify and separate crops in the Landsat data. The analyst interpreter uses events in the plant growth cycle such as emergence, heading, and turning correlated with crop calendars to separate winter and spring crops in the imagery. When imagery of the critical events is not acquired or is not available to the analyst, it is not possible to separate winter and spring crops. The separation becomes more exact as more of these events are available to the analyst.

An analysis of Landsat data and the results submitted to CAS by the image analysts revealed that the analysts experienced difficulty in identifying winter grains in areas for which limited Landsat imagery was available. Imagery was not acquired because of clouds, processing problems, and collection problems. In these areas, if a significant amount of spring grains was present, the analyst had little opportunity to separate winter and spring small grains. If the affected areas were designated as winter wheat regions because of a preponderance of winter wheat, the image analysts usually submitted winter grain estimates. When no acquisitions were available between fall tillering for winter grains and jointing for spring grains, LACIE winter grain estimates likely contained both winter and spring small grains.

The sample segment estimates representing the winter wheat were reviewed and the following criteria were applied to resolve the uncertainties created by inadequate acquisition histories.

1. Sample segment winter grain estimates based on Landsat data acquired before winter wheat tillering were not used for aggregation.

2. Sample segment winter grain estimates based on Landsat data acquired between winter wheat tillering and the historical average beginning of spring small grains jointing were used for aggregation. These estimates were considered reasonably ac-

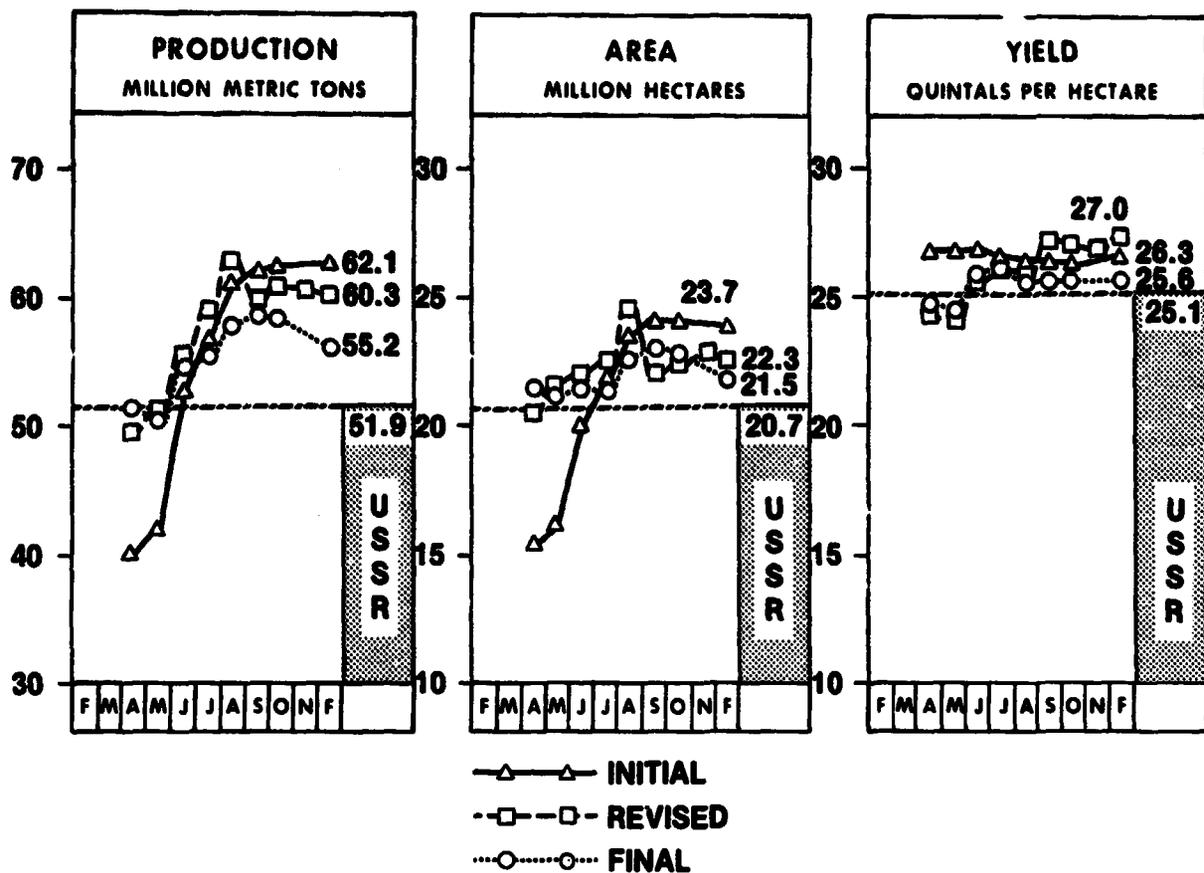


FIGURE 7.—Phase III winter wheat production, area, and yield estimates for initial, revised, and final aggregations.

curate because winter grain emergence was complete and confusion with spring-planted crops was minimal. No estimates based on other critical acquisitions were considered necessary to verify the accuracy of these estimates.

3. Sample segment winter grain estimates based on acquisitions after the historical beginning of spring grain jointing required a second estimate based on acquisitions between winter wheat tillering and spring grain jointing to verify the winter small grains classification before it was used in the aggregation.

4. If a winter grain sample segment estimate satisfied criteria 2 and 3 above but exceeded the previous estimate by an absolute 5 percent, it was converted to a total grain estimate.

For example, if for a given segment, a usable acquisition (after tillering) dated October 30, 1976, was given a winter small grains estimate of 11.0 percent and a subsequent aggregatable acquisition received

on June 3, 1977, was given a winter small grains estimate of 20 percent, the 20-percent winter-sown small grains estimate would not be included in the aggregation. The 20 percent was converted to a total small grains estimate and then included in the aggregation. The June 3, 1977, estimate of winter-sown small grains would have been included in the aggregation if it had not been greater than 16 percent, since it would have been not more than 5 absolute percentage points above the earlier estimate of 11 percent.

When the criteria were applied to the CAMS estimates, many of the estimates that most likely contained both spring and winter grains were converted to total grains estimates. As a result, the winter wheat to small grains (WW/SG) ratio was applied rather than the winter wheat to winter small grains (WW/WSG). The WW/SG ratio was less than or equal to the WW/WSG ratio; therefore, the winter wheat ratioed estimate was lower, as were the final aggregated results. The final approach was desirable

because it avoided data elimination and hence improved the precision of the aggregated regional estimates.

Comparison of estimates involving initial, revised, and final methods of data selection.—Winter wheat, spring wheat, and total wheat estimates as selected by the three methods of editing sample segment data for aggregation are compared as follows.

Winter wheat: During Phase III, winter wheat estimates were submitted to the commodity analyst for aggregation as soon as winter grains were detected on the Landsat imagery. The early acquisitions usually showed only partial emergence, and the estimates did not accurately represent the amount of winter grains planted. In the U.S.S.R., winter grains may not complete tillering in all regions before dormancy occurs. Tillering appears to be the earliest stage in the growth cycle of wheat for positive detection by the LACIE system since this stage provides sufficient vegetative ground cover for identification of plant life.

Because these early estimates were potentially only partial estimates of the emerging winter wheat, aggregated country estimates were biased downward to unsatisfactorily low values. To eliminate these low estimates, an early-season data editing technique was implemented so that segment estimates made before tillering would not be used for aggregation. Figure 7 shows the low April production and area estimates associated with the initial aggregations compared with the higher estimates associated with the revised and final aggregations that use early-season data editing.

As the season progressed, the low early-season estimates were replaced by complete estimates. By the time of the July report, which primarily used spectral data through May 15, the best initial production and area estimates had been derived. The area for the revised and final methods stayed relatively constant from April to July. The increase in production for these estimates was due to increases in the LACIE predicted yields and was not related to shifts in area from low- to high-yielding areas. The initial production estimate rose between April and July because of increases in both area and yield.

After the July report, the initial aggregated area estimates continued to rise to a maximum of 23.8 million hectares, and the production reached a maximum of 62.1 million metric tons at the end of the season. These figures are unrealistically high. As the winter wheat estimates rose, CAS developed and implemented the revised aggregation technique (de-

scribed earlier in this paper). The revised method succeeded in reducing the high initial country area estimate to a reasonable figure; however, the end-of-season estimate was still slightly less than 8 percent too high. After July, the production estimate for the revised method continued to rise but not at the rate of the initial estimate.

With data for the entire year in-house, the CAS analysts studied the revised estimates at the regional level. The regional estimates suggested that, although the country area estimates were reasonable, certain regions that were greatly overstated in the initial estimate remained unrealistically high, whereas certain regions that were unreasonably low were reduced even further. Therefore, the final technique was developed to solve this problem. The deviations of the end-of-season results for the final estimates of area and production compared with the official U.S.S.R. release were 3.9 and -6.4 percent, respectively. Figure 7 shows a rise and fall in area and production between July and the end of the season attributable to CAMS processing procedures. CAMS was required to backlog winter wheat data after July so that spring wheat could be processed. Not until October and November were significant amounts of late-season imagery (biostages 6 and 7) processed. The significant departure in production between the revised and final approaches is partly due to less area in the revised method, but it is primarily due to the final procedure's correction of higher yielding areas.

Spring wheat: The spring wheat revised approach is identical to the initial approach, except that historical data (six segments) have been added to Tyumen¹ and the Northwest, enabling estimates to be generated for these two zones.² The spring wheat final approach used the same data selection criteria as were used for the winter wheat final approach. The additional historical data for Tyumen¹ and the Northwest are included in the final estimate.

The August estimates for area and production are low because sufficient spectral data had not been processed to adequately estimate the entire spring wheat crop by the August cutoff date. The August aggregation contained estimates of only 34 percent of the 1416 allocated spring wheat segments, and many of these estimates were low because of incomplete emergence. As the season progressed, these low esti-

²Under Phase III software, a zone would receive estimates only if at least one stratum within the zone contained a minimum of three segments with aggregatable acquisitions.

mates were replaced by later estimates reflecting the actual area or extent of the small grains with additional segments also being added to the final report. The end-of-season report contained estimates on 69 percent of the 1416 segments and provided a best estimate (final) of 41.4 million hectares and 36.3 million metric tons (fig. 8). These estimates deviated from the official U.S.S.R. announcement by 0.2 and -9.5 percent, respectively.

Total wheat: The total wheat aggregation is simply a combined report of the winter and spring estimates for August through the end of the season. The total wheat initial area estimates (fig. 9) are about 5 million hectares less in August than at the end of the season. This is due to low early-season spring wheat estimates since the winter wheat area only varied by 400 000 hectares during this same time period. The spring wheat area gained 4.5 million hectares from August until the end of the season. The initial production estimates level off from September until the end of the season when the low early-season spring wheat estimates were corrected. Production for

spring wheat and winter wheat varied only 700 000 and 400 000 metric tons, respectively, after the September report.

The revised and final approaches give lower area and production estimates than the initial approach because of reduced winter wheat estimates during this time. These approaches and their effects on winter wheat have been previously discussed. Again, the rise and fall in the production figures by the final approach are due to the backlog of winter wheat data during the season with late-season data being processed only in October and November. This is discussed further in the conclusions section. The best area and production estimates for total wheat are the end-of-season final estimates of 62.9 million hectares and 91.4 million metric tons, which differ from the actual estimates as announced by the U.S.S.R. (62.0 million hectares and 92.0 million metric tons) by 1.5 and 0.7 percent, respectively.

Estimates based on 30-day turnaround time.—To illustrate the importance of spectral data timeliness, all Phase III monthly estimates were recomputed

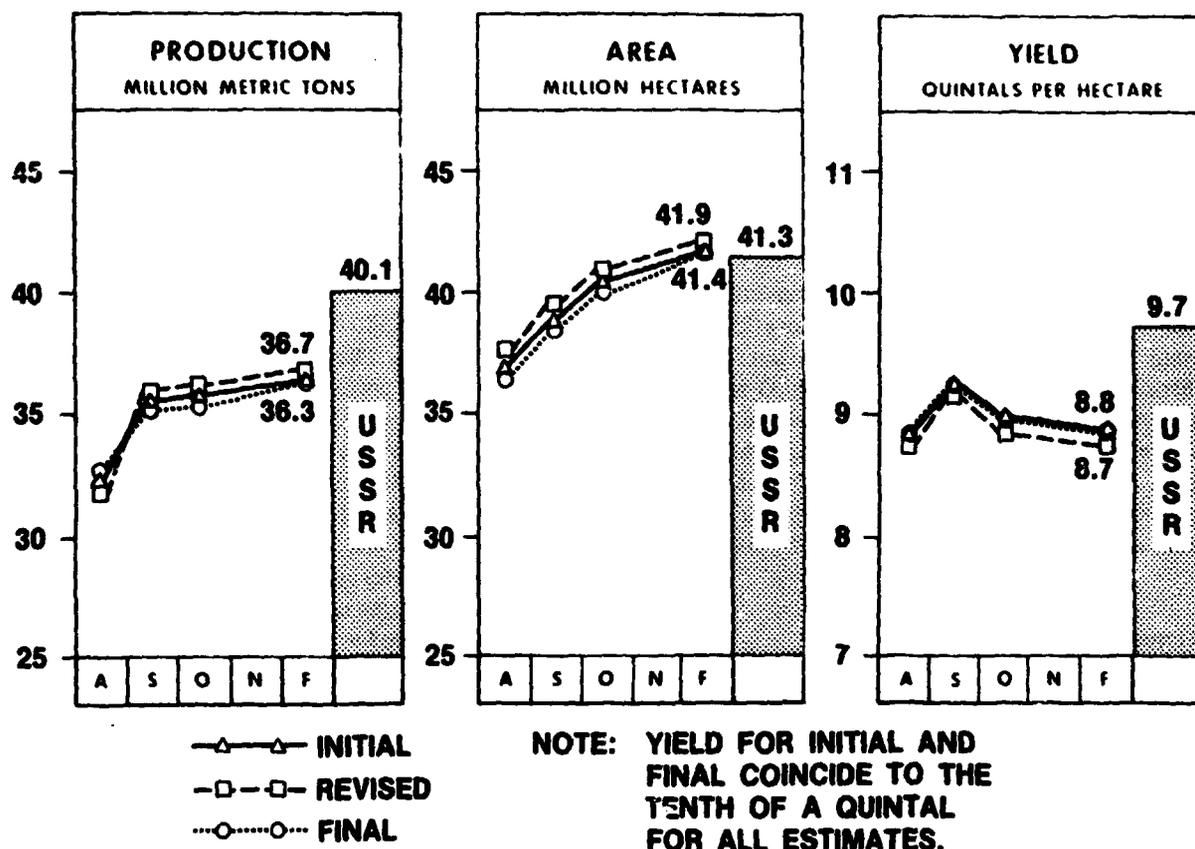


FIGURE 8.—Phase III spring wheat production, area, and yield estimates for initial, revised, and final aggregations.

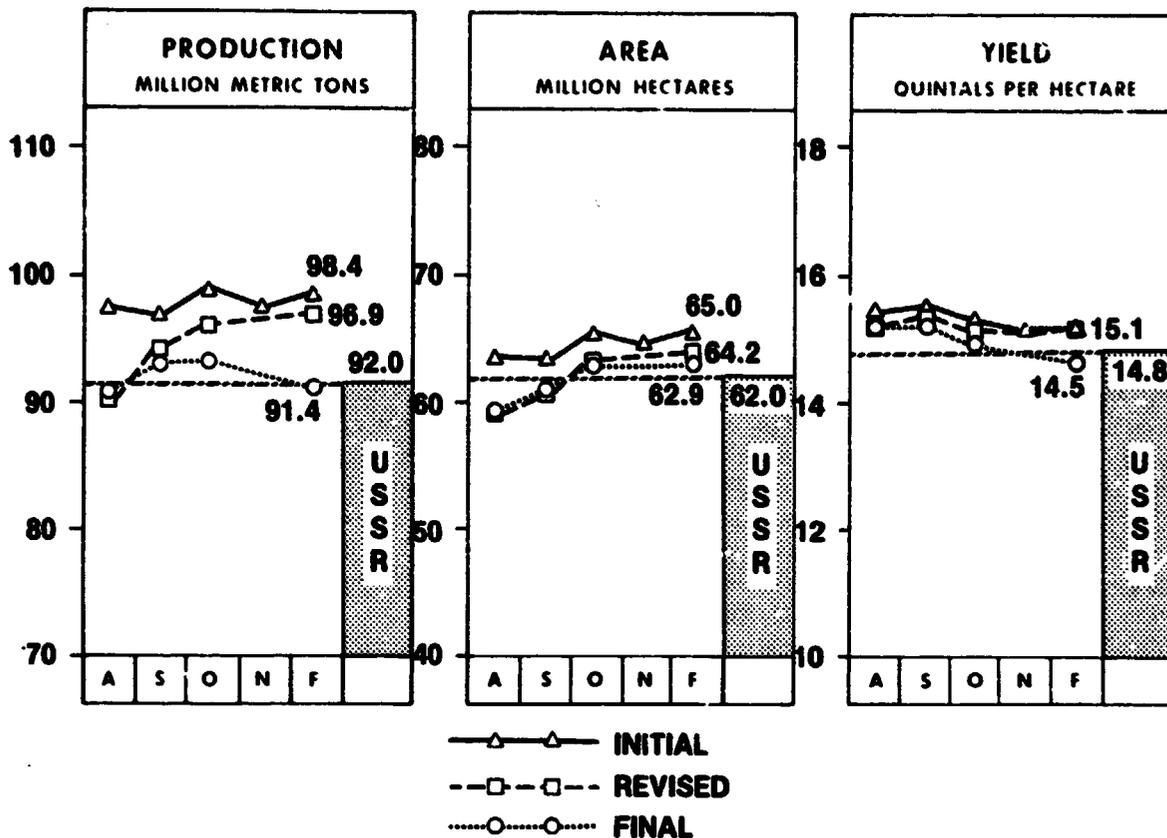


FIGURE 9.—Phase III total wheat production, area, and yield estimates for initial, revised, and final aggregations.

from a 30-day delay average rather than from the real-time average; figures 10 to 12 show the results. The 30-day delay data provide a more logical estimate curve because (1) the yield data are synchronized with the spectral data (both 30 days before the report date) as opposed to the real-time situation where the most current spectral data are normally at least 30 days older than the yield data, and (2) spectral data were aggregated chronologically as they were acquired (rather than backlogging data).

Accuracy of Estimates

The LACIE wheat production estimates for the U.S.S.R. are presented in table VIII with statistics and comparison data. Four types of LACIE estimates were generated in Phase III.

1. Initial estimates used all CAMS-processed acquisitions.

2. Revised estimates (first used in the April estimates) employed a thresholding procedure to eliminate early-season (preemergence) acquisitions and later included a thresholding procedure based on key acquisition to eliminate suspect data.

3. Final estimates, released in the CAS annual report, were recalculated for the entire season using the data editing procedure.

4. Final estimates with a 30-day delay—wherein Landsat data acquired up to 30 days before the report date were aggregated to make area data more directly comparable to yield data and to normalize the processing time—were used. The final estimates with a 30-day delay were released as the official LACIE estimates for Phase III.

Comparison of the LACIE monthly total wheat

TABLE VIII.—Comparison of LACIE and U.S.S.R. Production Estimates

Month of estimate	U.S.S.R. LACIE initial estimate			RD, percent	Value of test statistics	LACIE revised			RD, percent	Value of test statistics	LACIE final			RD, percent	Value of test statistics			
	MMT	Estimate, MMT	CV, percent			Estimate, MMT	CV, percent	Estimate, MMT			CV, percent	LACIE final 30-day delay	RD, percent			Value of test statistics		
<i>Winter wheat</i>																		
April		48.8									51.6	7.0	-0.6	-0.1	51.6	7.0	-0.6	-0.1
May		50.9									50.7	5.5	-2.4	-0.4	52.5	5.2	1.1	.2
June		55.5									54.5	5.8	4.8	.8	56.4	5.7	6.0	1.4
July		50.4									55.5	4.9	6.5	1.3	55.6	4.9	6.7	1.3
August		63.0	4.4	17.6	^a 4.0						57.8	4.4	10.2	^a 2.3	56.3	4.3	7.8	^a 1.8
September		63.9	4.3	18.8	^a 4.4	59.3	4.9	12.5	^a 2.5		58.3	4.3	11.0	^a 2.6	55.2	4.2	6.0	1.4
October		64.0	4.2	18.9	^a 4.5	60.8	4.6	14.6	^a 3.2		58.0	4.2	10.5	^a 2.5	55.2	4.2	6.0	1.4
Final	51.9	62.1	4.1	16.4	^a 4.0	60.3	4.4	13.9	^a 3.2		55.2	4.2	6.0	1.4	55.2	4.2	6.0	1.4
<i>Spring wheat</i>																		
April																		
May																		
June																		
July																		
August		34.6	9.2	-15.9	^a -1.7						32.5				32.8	8.8	-22.3	-2.5
September		37.9	7.2	-5.8	-.8						35.4				37.1	7.3	-8.1	-1.1
October		38.3	7.0	-4.7	-.7						35.7				36.4	7.1	-10.2	-1.4
Final	40.1	36.3	7.1	-10.2	-1.4						36.3	7.2	-10.2	-1.4	36.3	7.2	-10.2	-1.4
<i>Total wheat</i>																		
April																		
May																		
June																		
July																		
August		97.6	4.3	5.7	1.3						90.3				89.1	4.3	-3.3	-0.8
September		101.8				97.2					93.7				92.3	3.9	.3	-.1
October		102.3				99.1					93.7				91.6	3.8	-.4	.1
Final	92.0	98.4	3.7	6.5	^a 1.8	96.6					91.4	3.8	-0.7	-0.18	91.4	3.8	-.7	-.2

^aLACIE estimate is significantly different from the reference standard at the 10-percent level.

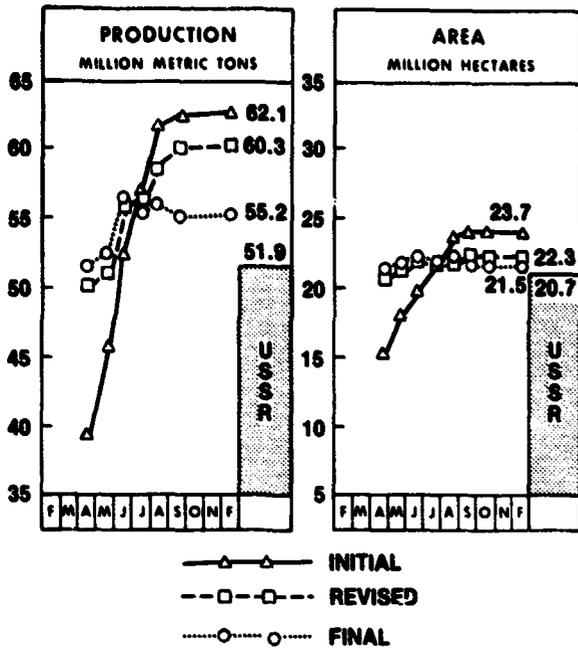


FIGURE 10.—Phase III winter wheat production and area estimates recomputed from a 30-day delay average.

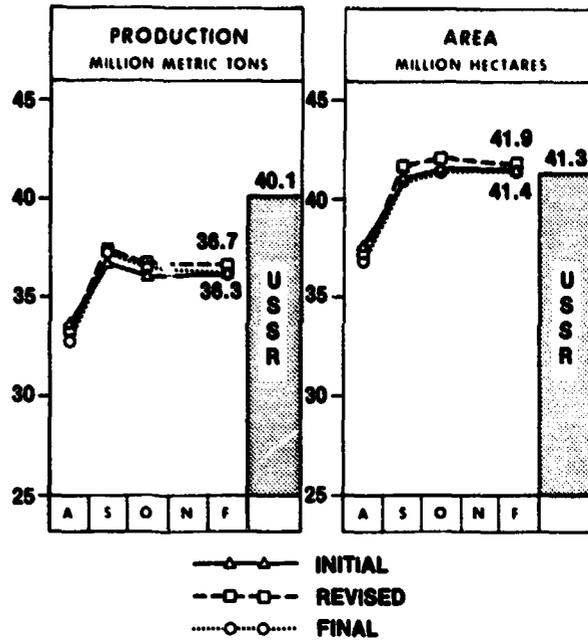


FIGURE 11.—Phase III spring wheat production and area estimates recomputed from a 30-day delay average.

production estimates with the official U.S.S.R. Government's final estimates indicates that the LACIE estimates supported the 90/90 accuracy goal each month from August through the final report. Table IX gives the statistics used to evaluate the 90/90 criterion. It contains the estimated relative difference (RD), the CV for each monthly estimate, the tolerable relative biases given for the observed CV, and the significance level. For example, the RD and the CV for the final LACIE estimates were -0.7 and 3.8 percent, respectively. With a CV of that magnitude, the LACIE total wheat production estimate would support the 90/90 criterion if the relative bias was between the limits of -5.6 and 4.6 percent. Since the estimated relative bias was within this interval, the estimate supported the 90/90 criterion. The last column gives estimates of the probability of observing the RD encountered, given a 90/90 production estimator. It is inferred that if the significance level is greater than 10 percent, the estimator supports the 90/90 accuracy goal.

The results of LACIE Phase III with its revised

TABLE IX.—Evaluation Statistics

Month	RD, percent	CV, percent	Tolerance limits, percent	Significance level, percent
August	-3.3	4.3	(-4.5, 4.0)	50
September	.3	3.9	(-5.6, 4.6)	50
October	-4	3.8	(-5.6, 4.6)	50
Final	-0.7	3.8	(-5.6, 4.6)	50

approach indicate that the accuracy goal of 90/90 was achieved in the U.S.S.R., where the shortfall in the spring wheat crop was identified 3 months before completion of harvest and similar results were achieved in the winter wheat regions. The initial LACIE estimate of 97.6 million metric tons in August was within 6 percent of the U.S.S.R. January 28 figure of 92 million metric tons, and the LACIE final estimate released on January 23 for total wheat production was within 1 percent. Throughout 1977,

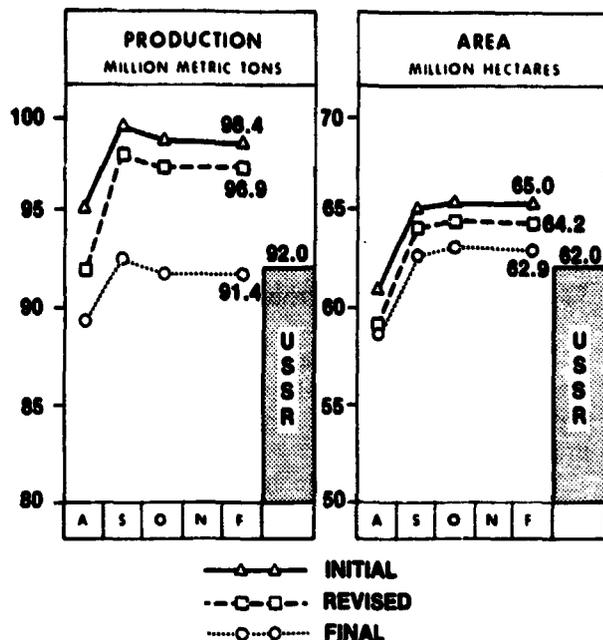


FIGURE 12.—Phase III total wheat production and area estimates recomputed from a 30-day delay average.

implementation problems and data processing backlogs were encountered that resulted in estimation error beyond that which would be encountered in a functioning operational system. Faulty data acquisition orders and inoperative ground receiving stations led to the loss of Landsat acquisitions over a portion of the U.S.S.R. winter wheat region. These lost Landsat data were never received or recorded and so could not be recovered. Data already processed were reevaluated, and the crop analyst's procedures were modified. In December 1977, these data interpretation problems were circumvented, and the LACIE estimates were recomputed using Landsat data assuming a 30-day processing delay operationally. The resulting estimates were released on January 23 before the final Soviet release. In a future operation, such results could be produced as early as August or September. These somewhat improved results were within 3 percent of the Soviet figures in August, some 3 months before harvest.

A detailed examination of the conditions leading to the Soviet shortfall in spring wheat production and the response observed in the LACIE models indi-

cated that the LACIE forecast technology did respond in a timely fashion. In most of the U.S.S.R. spring wheat regions, the growing season experienced temperatures warmer than average. These elevated temperatures led to moisture deficiencies through increased demand on available precipitation. The PET data indicated that the above-normal temperatures in the growing season seriously depleted soil moisture supply throughout the southern portions of the U.S.S.R. spring wheat area. While the northern regions had normal to above-normal moisture in addition to these impacts, the April temperature was nearly 4° C above normal, which theoretically at least would deplete the pre-season soil moisture supply.

An investigation of the Landsat data and the yield model response at subregional levels indicated that the drought conditions were clearly observable in the Landsat data and that the yield models responded by reducing yield estimates in the affected regions. Radiometric measurements by Landsat (green index number), which were known to be related to crop vigor, indicated the southern portions of the spring wheat region were under severe drought conditions. However, in the northern regions, LACIE was forecasting above-normal yields. In the southern regions, LACIE yield models reduced the yield prospects nearly 2 quintals per hectare in response to the high April temperature before the growing season had commenced. The continuing drought reduced the yield nearly 2 more quintals per hectare below the normal 11.5 quintals per hectare.

Production.—For total wheat, there was no significant difference at the 10-percent level between the final official LACIE production estimate and the final estimate released by the U.S.S.R. Government. In fact, the final or official LACIE production estimates for Phase III were consistently between 89.1 and 92.3 million metric tons and were never significantly different from the official Soviet figure of 92.0 million metric tons. The final RD between the official LACIE and U.S.S.R. Government estimates was -0.7 percent. The CV for the official LACIE estimates dropped steadily from 4.3 percent in August to 3.8 percent for the final estimate.

A comparison of the monthly LACIE production winter wheat figures with the official U.S.S.R.

Government figures showed that only in August was the difference significant at the 10-percent level.

The official LACIE spring wheat estimates for September, October, and the final report were not significantly different at the 10-percent level from the official Soviet estimate, but the difference was highly significant in August because of LACIE underestimation.

Area estimates.—The LACIE wheat area estimates for the U.S.S.R. and associated statistics and comparison data are presented in table X. The test statistics for total wheat showed that the differences between the official LACIE estimates and the official Soviet estimate were not significant at the 10-percent level except in August. The underestimation in August was due to the LACIE underestimation of spring wheat area in the first spring wheat area aggregation for the U.S.S.R. in Phase III.

Although a complete set of test statistics was not available for the initial, revised, and final estimates, it is apparent that the final and official estimates were closer to the official U.S.S.R. Government figure than were the initial or revised estimates.

There were marginally significant differences at the 10-percent level between the official LACIE winter wheat area estimates and the official Soviet estimates for August, September, and October in Phase III. The final official LACIE estimate was not significantly different from the official figure released by the U.S.S.R. Government.

The spring wheat statistics and associated comparison data in table X indicate that the official LACIE estimates compared well with the Soviet estimate after August. However, there is a large RD of -12.8 percent for August due to underestimation by LACIE.

Yield estimates.—The LACIE wheat yield estimates for the U.S.S.R. are presented in table XI with the associated statistics and comparison data. The estimates of the precision (i.e., the CV) were not available for the LACIE total wheat yield estimates. However, the official LACIE estimates were quite close to the official U.S.S.R. Government estimate. The difference between the official LACIE estimates for total wheat and the official U.S.S.R. Government figure was never more than 0.4 quintal per hectare.

The final and official LACIE winter wheat yield estimates were closer to the U.S.S.R. Government estimate than were the initial or revised estimates. There was no significant difference at the 10-percent level between the official LACIE and the official U.S.S.R. Government estimates. The absolute

difference between the monthly LACIE estimates and the official U.S.S.R. Government estimate never exceeded 1.1 quintals per hectare.

None of the differences between LACIE spring wheat estimates and the official Soviet estimate were significant at the 10-percent level. A tendency towards underestimation is apparent, however, and has been addressed previously.

Technical Issues and Problems

The technical issues and problems in Phase II were also major constraints in Phase III with the exception of the indicator region and the sampling problem.

Indicator region.—When it was decided to work the entire wheat-producing area of the U.S.S.R. in Phase III, the issue of indicator regions and the associated problem of what to use as a comparison for the LACIE estimates were eliminated. Although the U.S.S.R. announcement of production, area, and yield statistics was released months after the Phase III crop year terminated, these data were irreplaceable in calculating the accuracy and success of LACIE in the U.S.S.R.

Sampling.—Approximately 800 of the total allocation of 1947 sample segments were relocated between Phases II and III. Most of the relocation involved moving sample segments from non-agricultural to agricultural areas; however, some segments were moved to provide more efficient sampling of the agricultural area. The segments were allocated on total land area in Phase II and on agricultural area only in Phase III.

Software problems.—During Phase II, it was decided that zones (see fig. 13) which did not have at least one stratum (oblast) with a minimum of three segments for which usable spectral (Landsat) data had been acquired would not be included in the aggregation or LACIE estimate. The rationale was that Landsat data for less than three segments would not be representative of an area or zone size and therefore could not be estimated by remotely sensed data. This software was in place at the beginning of Phase III. It became immediately apparent that the LACIE estimates for the U.S.S.R. would be unacceptable (early season at a minimum) because of the large number of zones which did not meet this criterion. It was decided to insert historical data in areas (zones) for which no estimate was generated until sufficient usable Landsat data were acquired to meet the necessary criterion.

TABLE X.—Comparison of LACIE and U.S.S.R. Area Estimates

Month of estimate	U.S.S.R. estimate, million hectares	LACIE initial		RD, percent	Value of test statistics	LACIE revised		RD, percent	Value of test statistics	LACIE final		RD, percent	Value of test statistics	LACIE final 30-day delay		RD, percent	Value of test statistics
		Estimate, million hectares	CV, percent			Estimate, million hectares	CV, percent			Estimate, million hectares	CV, percent			Estimate, million hectares	CV, percent		
<i>Winter wheat</i>																	
April		20.3								21.3	6.3	2.8	0.4	21.3	6.3	2.8	0.4
May		21.3	4.6	2.8	0.6					20.9	5.3	1.0	.2	21.8	4.9	5.0	1.0
June		21.8	3.5	5.0	1.4					21.3	4.5	2.8	.6	22.1	4.5	6.3	1.4
July		22.4	3.0	7.6	^a 2.5					21.3	3.8	2.8	.7	21.5	3.6	3.7	1.0
August		24.3	2.4	14.8	^a 6.2					22.5	2.9	8.0	^a 2.8	22.1	2.7	6.3	^a 2.3
September		24.6	2.7	15.9	^a 5.9	21.9	3.5	5.5	1.6	22.8	2.7	9.2	^a 3.4	21.6	2.5	4.2	^a 1.7
October		24.6	2.5	15.9	^a 6.3	22.6	3.3	8.4	^a 2.5	22.7	2.6	8.8	^a 3.4	21.6	2.5	4.2	^a 1.7
Final	20.7	23.7	2.4	12.7	^a 5.3	22.3	2.8	7.2	^a 2.6	21.5	2.5	3.7	1.5	21.5	2.5	3.7	1.5
<i>Spring wheat</i>																	
April																	
May																	
June																	
July																	
August		38.9	4.3	-6.2	-1.4					36.8				36.6	3.6	-12.8	^a 3.6
September		41.0	2.9	-7	-3					38.7				41.1	2.4	-5	^a 2
October		42.6	2.6	3.1	1.2					40.3				41.5	2.3	.5	.2
Final	41.3	41.4	2.4	.2	.1					41.4	2.3	0.2	0.1	41.4	2.3	.2	.1
<i>Total wheat</i>																	
April																	
May																	
June																	
July																	
August		63.2	2.3	1.9	0.8					69.3				58.6	2.6	-5.8	^a -2.2
September		65.6				62.9				61.6				62.7	1.9	1.1	.6
October		67.2				65.2				63.0				63.0	1.8	1.6	.9
Final	62.0	65.0	1.8	4.6	^a 2.6	63.7				62.9	1.8	1.4	0.8	62.9	1.8	1.4	.78

^aLACIE estimate is significantly different from the reference standard at the 10-percent level.

TABLE XI.—Comparison of LACIE and U.S.S.R. Yield Estimates

Month of estimate	U.S.S.R. estimate	LACIE initial		RD. percent	Value of test statistics	LACIE revised		RD. percent	Value of test statistics	LACIE final		RD. percent	Value of test statistics	LACIE final 30-day delay		RD. percent	Value of test statistics
		q/ha	Estimate, CV, percent			q/ha	Estimate, CV, percent			q/ha	Estimate, CV, percent			q/ha	Estimate, CV, percent		
Winter wheat																	
April		24.0								24.3	4.4	-2.9	0.7	24.3	4.4	-2.9	-0.7
May		23.9								24.2	3.2	-3.3	-1.0	24.1	3.1	-3.7	-1.2
June		25.5								25.6	4.3	2.3	.5	25.6	4.2	2.3	.5
July		26.1								26.1	3.9	4.2	1.1	25.9	3.9	3.5	.9
August		25.9	3.4	3.5	1.0					25.6	3.6	2.3	.7	25.5	3.6	2.0	.5
September		26.0	3.6	3.8	1.1	27.0	3.7	7.4	*2.0	25.6	3.5	2.3	.7	25.6	3.6	2.3	.7
October		26.0	3.5	3.8	1.1	26.8	3.6	6.7	*1.9	25.6	3.5	2.3	.7	25.6	3.6	2.3	.7
Final	25.0	26.3	3.6	4.9	1.4	27.0	3.6	7.4	*2.1	25.6	3.6	2.3	.6	25.6	3.6	2.3	.6
Spring wheat																	
April																	
May																	
June																	
July																	
August		8.9	8.7	-9.0	-1.0					8.8				9.0	8.8	-7.8	-0.9
September		9.3	7.1	-4.3	-6					9.1				9.0	7.2	-7.8	-1.1
October		9.0	6.9	-7.8	-1.1					8.9				8.8	7.0	-10.2	-1.5
Final	9.7	8.8	7.0	-10.2	-1.5					8.8	7.0	-10.2	-1.5	8.8	7.0	-10.2	-1.5
Total wheat																	
April																	
May																	
June																	
July																	
August		15.4								15.2				15.2			
September		15.5				15.5				15.2				14.7			
October		15.2				15.2				14.9				14.5			
Final	14.8	15.1				15.2				14.5				14.5			

*LACIE estimate is significantly different from the reference standard at the 10-percent level.



FIGURE 13.—Zone boundaries in the U.S.S.R.

CONCLUSIONS

The major problems with area estimation that have been discussed in this paper were overcome or successfully circumvented in Phase III. This resulted in estimates well within the accuracy tolerance goals of LACIE; however, there are refinements and enhancements that should be implemented to further improve the estimates.

In general, the yield models provided much better estimates in Phase III than in Phase II because of the updated data and the resultant coefficient changes. The models never demonstrated the capability of reflecting the degree of extremely high or low yields but rather indicated the direction of deviation from trend. These models should be upgraded or replaced with models capable of more accurate yield estimates to reduce more accurate production estimates.

Availability and timeliness of Landsat data are very important to accurate area estimates. If the appropriate spectral data coverage had been available at the right time, the data editing procedures, as discussed earlier, could have been avoided.

The consensus of the commodity analysts participating in the U.S.S.R. work is that Phase III U.S.S.R. area results are repeatable given the same Landsat classification and aggregation procedures.

The LACIE U.S.S.R. Phase III results evolved from 3 years of concentrated effort by essentially the same commodity analysts. Moving to a new foreign area would require a minimum preparation time of 1 year before reliable results could be expected.

ACKNOWLEDGMENTS

A. G. Houston of the NASA Johnson Space Center, Houston, Texas, contributed the section entitled "Accuracy of Estimates." L. O. Lovfald of the USDA Economics, Statistics, and Cooperatives Service (Houston, Texas) contributed the section entitled "Yield Analysis." Other personnel having key roles in the work and data analysis described herein are Daniel Bridge of the USDA Foreign Agricultural Service (Houston, Texas) who provided data analysis in Phase III; Delanne Conte of the USDA Foreign Agricultural Service (Houston, Texas) whose assistance was invaluable in developing the U.S.S.R. data base in Phases I and II and data analysis in Phases II and III; Larry Davis of the USDA Foreign Agricultural Service (Houston, Texas) whose analytical abilities and knowledge of remote sensing were instrumental in achieving the high degree of accuracy of the U.S.S.R. estimates in Phase III; and William E. McAllum of the NASA Johnson Space Center who, in his position as Manager of the Crop Assessment Subsystem, provided continuing guidance and support to the commodity analysts. Their contributions are hereby gratefully acknowledged.

REFERENCE

1. LACIE Phase II Crop Assessment Subsystem Annual Report. CAR-1, NASA Johnson Space Center, Houston, Texas, Oct. 22, 1976.

LACIE Area, Yield, and Production Estimate Characteristics: Canada

Delanne Conte,^a A. G. Houston,^b and L. O. Lovfald^c

OVERVIEW

The LACIE project originally planned to generate area, yield, and production estimates in the Canadian Prairie Provinces—the principal wheat-growing region of that country—in both Phases II and III. Because of a change in scope at the beginning of Phase III, the investigation in Canada was reduced to a moderate number of sites where ground truth was collected. In order to support accuracy assessment, a small number of exploratory and intensive test sites were analyzed during all three phases of LACIE, but in-season area, yield, and production estimates were generated only in Phase II.

Phase I

The work in Canada in Phase I was centered around building a historical-statistical data base, locating sample segments within the country, and acquiring multispectral scanner (MSS) data for a subset of the sample for study by the image analyst. There were no estimates generated for Canada during this phase.

Phase II

Scope.—The LACIE Canadian spring wheat region includes the major spring wheat producing region in Canada (fig. 1). This area is comprised of the Provinces of Saskatchewan, Alberta, and Manitoba. This region grows predominantly spring wheat and spring grains with some winter rye scat-

tered throughout the three provinces. Saskatchewan accounts for the major proportion of the three-province total spring wheat production, with approximately 65 percent, while Alberta accounts for 23 percent and Manitoba 12 percent.

Sampling.—The sampling frame used in Canada is the same as that used for sampling in the United States: Group I units which historically account for substantial wheat area and Group II units where segments are allocated on the basis of probabilities proportional to size. The initial sample segment allocation was based on the 1971 census and census district boundaries. The 1971 census was used instead of more current information since it was the only publication that included county-level area data. This choice of a base had a major impact on the aggregation results and caused a number of problems in the analysis and evaluation of the Canadian data. The final section of this paper will describe in detail the major problems associated with the Phase II Canadian reports.

The sample segment allocation for Canada placed 283 segments within the three provinces, with Saskatchewan allocated 170, Alberta 75, and Manitoba 38.

Data base.—The Canadian data base is comprised of five data sets that are necessary to support the aggregation and reporting functions. These five data sets are as follows.

1. Allocation data. This file includes the political hierarchy and associated segment and area descriptors—agricultural land, political area, segment location, and political hierarchy identifiers.

2. Historical data. At a minimum, historical area, yield, and production data were input for the 1971 base year that was used to generate the sample segment allocation.

3. Ratio data. This file was used to ratio the small grains estimates generated from the Landsat data to the wheat estimates that were needed to support

^aUSDA Foreign Agricultural Service, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

^cUSDA Economics, Statistics, and Cooperatives Service, Houston, Texas.

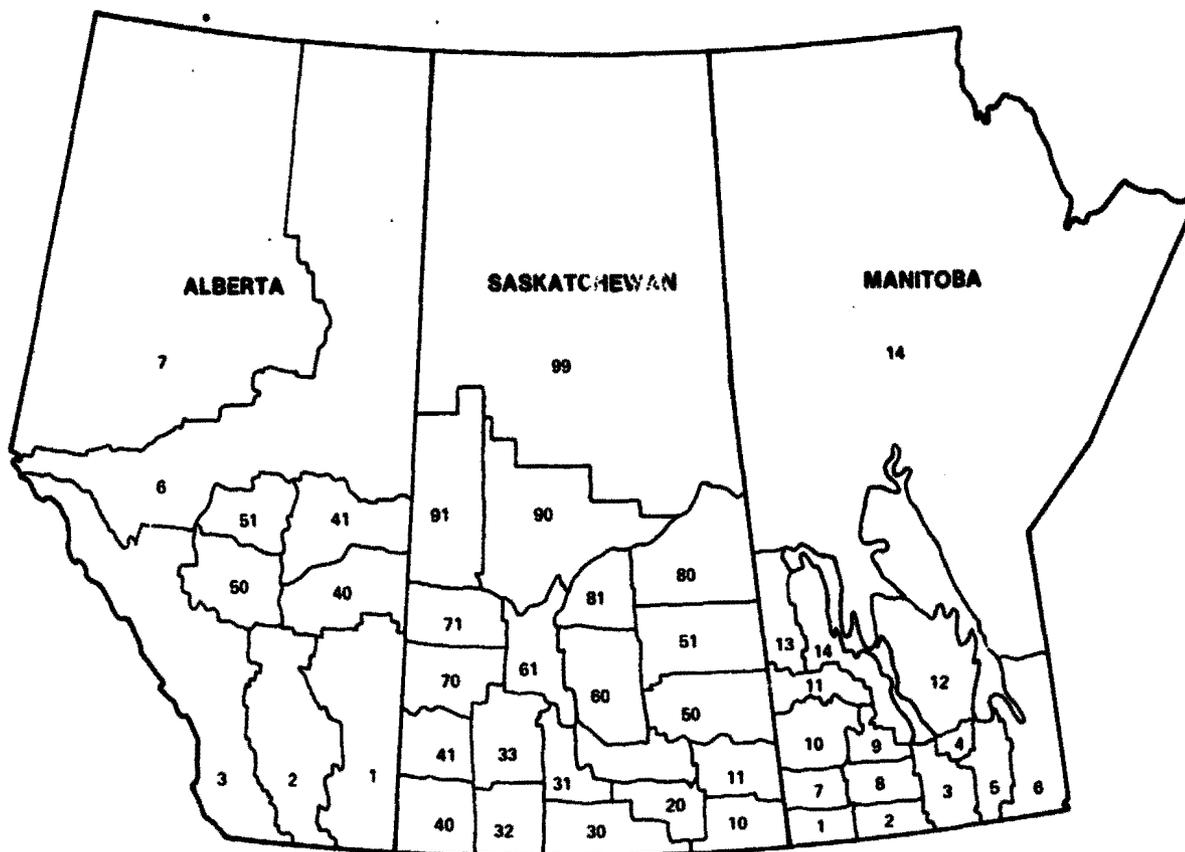


FIGURE 1.—Map of Canadian Prairie Provinces with LACIE strata numbers.

wheat area aggregation. The ratios were based on 1971 data and were constructed for two different data types—spring wheat to spring small grains and spring wheat to total small grains.

4. Landsat data. All segment-level estimates that were generated for Canada were stored in the data base for use in generating aggregated area estimates.

5. Yield data. Yield estimates were generated for each crop district by the National Oceanic and Atmospheric Administration's (NOAA) Center for Climatic and Environmental Assessment (CCEA). These base estimates were stored and used in aggregating yield and production estimates to the country level.

Landsat data.—The final wheat area estimate for the Canadian spring wheat region was based on spectral coverage obtained between May 29 and September 23, 1976, with the majority of acquisitions received between late June and mid-August. All Canadian data were processed as spring small grains or small grains. Ratioing was performed within the

Crop Assessment Subsystem (CAS) to determine the percent wheat.

Landsat acquisition coverage throughout the 1976 crop season was exceptionally good as a result of the relatively cloudfree summer. The overall average acquisition rate was 6 acquisitions per segment, with a usable acquisition rate of 2.8. This netted an end-of-season estimate that contained usable data for 90 percent of the 283 allocated segments. Both Saskatchewan and Manitoba had usable acquisitions that accounted for 90 to 95 percent of the segments for each of the three reports generated during the season. Only Alberta had consistently low coverage rates, with 27 percent usable for the first report, 57 percent for the second, and 79 percent for the final.

The project processed 1704 acquisitions during the 1976 Canadian crop season, mid-May to mid-September. Of these 1704 acquisitions, 254 were used in the final aggregation. The data dropout is accounted for by 913 acquisitions being classified as not usable for aggregation (cloud cover, mechanical problems,

etc.) and 537 acquisitions, which were classified as satisfactory, being replaced by subsequent acquisitions (fig. 2). The classifications for all acquisitions for all reports are shown in table I. Of the 254 acquisitions used in the aggregation, 1 segment was classified as nonagricultural. In the first report, 28 segments were classified or ratioed to 0 percent wheat. By the final report, 12 of these segments had received updated information, leaving 16 segments with a 0 percent wheat estimate.

Eighty percent of all estimates used in the final aggregation included acquisitions for biostage 4 or later. This figure indicates that most of the Landsat analysis included data for the end of the crop's full growth cycle. Throughout the season, approximately 35 days were required to process the data from acquisition to receipt by CAS for use in the aggregation.

Yield data.—The CCEA agrometeorological yield estimates for the 1976 crop year generally confirmed the known meteorological conditions prevalent throughout the three major wheat producing provinces. Much like the summer's weather pattern, yield results were somewhat mixed across the prairies. The drought that dominated the U.S. upper Midwest throughout the summer affected Canadian agriculture as well, driving yield model indications

well below normal along the eastern fringes of the Canadian wheat belt.

Manitoba spring wheat yields apparently experienced the greatest setback, finishing 4 to 6 bushels per acre below average in the east, with all districts predicting below-normal yields. This result represented a gradually deteriorating situation through the summer. Early in the season, crop districts in Manitoba, as well as in most of the other

TABLE I.—Classification of Acquisitions

Classification	No. (percent) of acquisitions in—		
	1st report	2nd report	3rd and final reports
Cloud cover, haze, etc.	38 (5)	68 (6)	91 (5)
Mechanical difficulties	6 (1)	18 (2)	43 (3)
Preemergence	386 (47)	386 (31)	386 (23)
Multiple acquisitions ^a	129 (16)	275 (22)	388 (23)
Unsatisfactory	1 (—)	4 (—)	5 (—)
Satisfactory	254 (31)	489 (39)	791 (46)

^aSeveral passes of one segment were reviewed at one time and only the best acquisition was used to determine the area estimate.

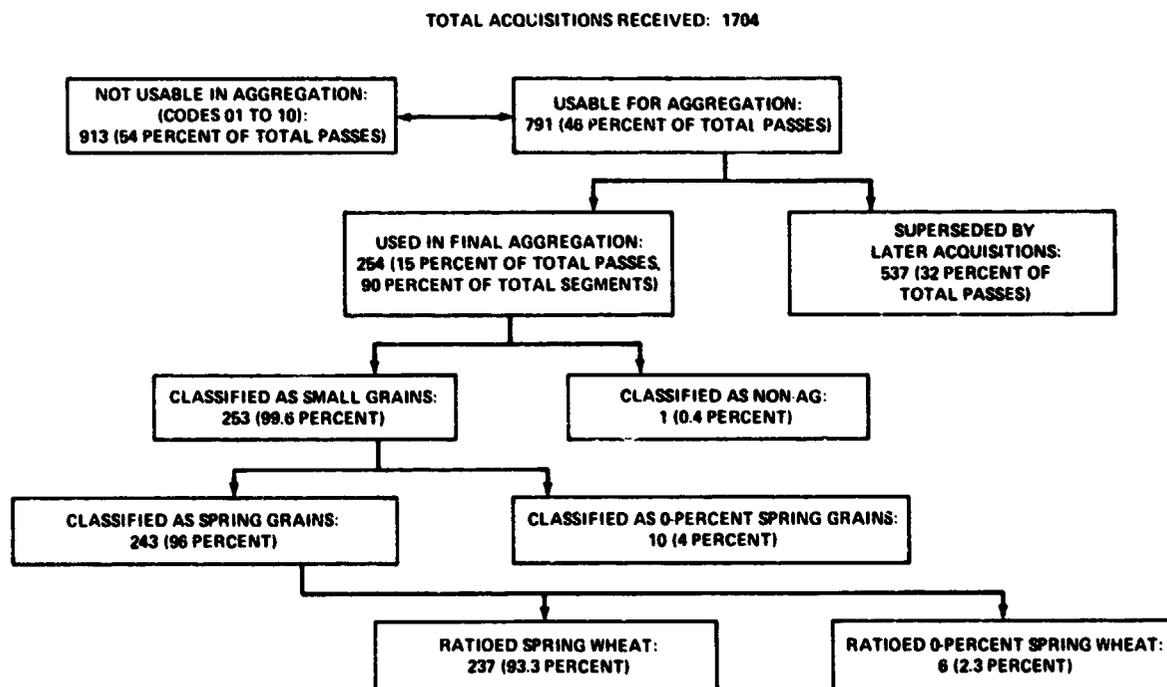


FIGURE 2.—Landsat data flow from CAMS to aggregation for spring wheat in the Canadian Prairie Provinces.

Canadian crop districts, were predicting above-average crop prospects. In early July, stress from insufficient moisture began to appear. From then until the final truncation in August, yield estimates declined steadily throughout the four Manitoba crop districts.

In the Province of Saskatchewan, to the west of Manitoba, yields showed a marked contrast to those in Manitoba, with crop predictions reflecting the prevailing ideal weather that region enjoyed through most of the summer. In the central portion around Saskatoon and westward, yield indications in May were up to 2 bushels per acre below normal, due to pre-season soil moisture deficiency. However, abundant June and July rains rallied crop prospects, and the entire province finished the season with bumper yields—as much as 8 bushels per acre above normal expectations in the north.

The seasonal impact of weather on Alberta wheat production seemed at the end of the season to have been favorable. Two of the three Alberta districts indicated yields slightly above trend with only the northwest Peace River region falling below the norm. Above-normal September temperatures, associated with mild but rainy weather, apparently delayed harvesting in that region, reducing somewhat the final wheat yield and quality. In the more northerly regions, once the crop has passed the frost-susceptible stage, colder temperatures associated with high pressure and clear weather are desirable to speed maturity, freeze down weeds that hamper harvest operation, and thus hasten harvest completion. The eastern and southern Alberta wheat crop, like that of the neighboring portions of Saskatchewan, started the year with slightly below-normal yield expectations, due to pre-season moisture problems, but recovered early and stayed even through most of the summer. Excellent maturing conditions in August and September pushed yields to their final above-trend mark in the major portion of Alberta's wheat producing region.

Phase III

In Phase III, Canada was not worked operationally to provide aggregated area, yield, and production estimates because of the change in scope at the beginning of this phase. A small number of exploratory and intensive test sites were analyzed to support accuracy assessment within the project. This analysis will not be presented in this paper.

ESTIMATES

Three monthly reports and one final report were generated for the Canadian spring wheat region during the 1976 crop year. The area and production estimates showed significant increases with each succeeding report. The final end-of-season area estimate for the three Canadian Prairie Provinces was 20.8 million acres (table II). This is 6.0 million acres (22 percent) below the official Canadian Statistics October estimate of 26.8 million acres. The production estimate was also low—576.3 million bushels versus an official figure of 833.0 million bushels—a 31-percent difference (table II).

The LACIE estimates were based on spectral data acquired from mid-June to mid-September 1976, with the bulk of the data acquired during July and August. For the Canadian Prairie Provinces as a whole, usable Landsat data were acquired for 90 percent of the segments.

The continual increases in the area estimate were due primarily to improved spectral coverage and upgrading of previous acquisitions. The first report was delayed until August because of the lack of adequate spectral coverage to generate an estimate. In that report, usable acquisitions were received for 61 percent of the allocated segments. While this may seem reasonable for a first report, early-season data (i.e., classifications indicating 5 percent or less small grains area) reduced the effective coverage to 43 percent of those segments allocated.

In the second and third reports, spectral coverage improved dramatically to 83 percent and 90 percent, respectively, with a substantial amount of mid- and late-season data being used to generate the estimate. Coverage throughout the three provinces was better than anticipated during this crop season since cloud cover, usually so prevalent during the Canadian crop season, was reduced because of the exceptionally favorable weather conditions.

The area estimates generated by LACIE during Phase II were well below official Canadian reports. In fact, every in-season estimate for the total region during 1976 fell below the 1971 base year used for the aggregation data base (table II). For the province-level estimates, only the third report estimate for Alberta was over the 1971 base level (approximately 9 percent over).

One of the most obvious causes of the underestimate was the use of ratioing to derive wheat estimates from the small grains estimates generated by the Classification and Mensuration Subsystem

TABLE II.—Comparison of August, September, October, and January LACIE Estimates for the Three Canadian Prairie Provinces^a

Province	Historical (1971)	LACIE				Official ^b
		Aug.	Sept.	Oct.	Jan. ^c	
<i>Area, thousands of acres</i>						
Saskatchewan	12 923	9 697	11 202	11 453	13 511	17 400
Alberta	3 443	2 099	3 433	3 767	4 535	5 600
Manitoba	2 519	1 751	2 100	2 125	2 756	3 800
Total	18 885	13 547	16 735	17 345	20 802	26 800
<i>Yield, bushels/acre</i>						
Saskatchewan	26.7	29.5	29.6	29.5	29.5	31.5
Alberta	26.4	23.5	24.6	24.6	25.0	32.5
Manitoba	29.4	23.1	23.3	23.4	23.4	27.1
Average	27.0	27.7	27.8	27.7	27.7	31.1
<i>Production, thousands of bushels</i>						
Saskatchewan	345 000	285 750	331 793	338 354	398 722	548 000
Alberta	91 000	49 235	84 281	92 731	113 216	182 000
Manitoba	74 000	40 400	48 995	49 817	64 376	103 000
Total	510 000	375 385	465 069	480 902	576 314	833 000

^aThese figures are based on a rework of the data performed after the annual report was produced. A number of errors were found in the yield models that required this rework. Yield and production accuracy statistics were not generated during the crop season and were computed during this rework. The rework of the yield estimates did not significantly affect the LACIE results. Figures are not comparable to in-season reported estimates.

^bStatistics Canada, Field Crop Reporting Series, No. 20, Dec. 3, 1976.

^cBased on 1975 data for ratioing CAMS small grains estimates.

(CAMS). 1971 was designated as the base year for Canada because statistics were available at the county level to support the sampling and allocation. When a ratioing procedure had to be implemented, 1971 data were also used to determine the ratios since no other data existed at the county level to support this procedure. Data obtained after 1971 (not at the county level) show a substantial shift in the amount of area planted to wheat versus other grains. Most of the increased wheat acreage has been at the expense of flax, rapeseed, and fallowed land. The government over the past several years has been asking farmers to increase their wheat and feed grain acreage and reduce their fallow acreage, since current research within the country shows that the prevalent summer fallowing practices have tended to waste moisture

and soil resources. When the 1975 ratios were applied to the segment estimates in the final report, the total LACIE wheat acreage increased from 17.3 million acres to 20.8 million acres, an improvement of 13 percent but still 22 percent below official estimates. If the ratio obtained from the 1976 official estimates of wheat and small grains acreage is applied to the LACIE data, an additional improvement of 4 percent is realized.

The use of more current data in the ratioing does, as anticipated, improve the estimate generated by LACIE, but it accounts for less than half the difference between the LACIE estimate and the official estimate. So, while ratioing is a major factor in the underestimate, other factors contributed substantially. All the factors including ratioing thought to

affect the estimate are reviewed in the section entitled "Technical Issues." To date, sampling/classification and Landsat resolution seem to be the other principal causes of the area underestimate.

The yield estimates generated for the 1976 crop season were under the official yields both for the region and for the individual provinces. The regional yield for LACIE was 27.7 bushels per acre compared to the official yield of 31.1 bushels per acre (table II). For the country as a whole, the final estimate reflected the gradual upward trend that the region has been experiencing over the past 10 years, but it did not account for the above-normal yields observed throughout the area. At the province level, the yield models seemed to reflect more strongly the early-season below-normal weather conditions and less strongly the whole-season favorable weather. The yield estimates were above trend only in Saskatchewan; the estimates for Alberta and Manitoba were below both the 1975 estimates and the 10-year average.

The final LACIE production estimate was 31 percent below the official Canadian estimate, 576.3 million bushels versus 833 million bushels. The low area estimate combined with the low yield estimate produced a lower than anticipated production figure. Only Alberta's production figure during the in-season reporting was above the 1971 base year as a result of the area estimate being above that level. The principal cause of this low production estimate was the area figure. Even though the LACIE yield was lower than official projections, it was above the long-term trend.

ACCURACY OF ESTIMATES

Production and Area Accuracy

As mentioned previously, the Canadian Phase II crop reports were regenerated and included yield and production accuracy statistics (see LACIE Phase II, Crop Assessment Subsystem Annual Report, Canada, January 14, 1977, with Addendum dated April 29, 1977). The revision of the production estimates resulted in the following changes (table III): August +3.7 percent, September +2.0 percent, October +0.3 percent, and January +0.2 percent. Since the changes were minor and since these estimates have associated with them the statistics needed for accuracy assessment, the following analysis is based on

these revised estimates. It should be pointed out that the only difference between the October and January estimates is that ratios of spring wheat to small grains based on 1975 Canadian data were used for the January spring wheat area estimates. The October and previous spring wheat area estimates were obtained using 1971 agricultural census data. The census data allowed crop-district-level ratios.

The revised estimates and corresponding coefficients of variation (CV's) for each month are presented in table IV. The relative differences (RD's) shown are with respect to the official Canadian estimates released December 3, 1976, by Statistics Canada. The test statistic indicates whether or not the LACIE estimate is significantly different from the corresponding official Canadian estimate. (For further details on the statistical approach, see the paper by Houston et al. entitled "Accuracy Assessment: The Statistical Approach to Performance Evaluation in LACIE.")

The precision (as measured by the CV) of the LACIE spring wheat production estimates for September, October, and January is sufficient to support the LACIE 90/90 accuracy goal. However, comparisons of each of the monthly production estimates with the official Canadian estimate indicate the presence of a negative bias that is too large to support this goal. Treating the observed RD and CV for the January production estimate as the true parameters of the LACIE production estimator for Canada indicates that a 90/65 accuracy goal is achievable at harvest; i.e., the probability is 90 percent that the at-harvest LACIE spring wheat production estimate is within ± 35 percent of the true Canadian production. This result for Canada, of course, falls far short of the 90/90 goal.

Comparisons of the LACIE area and yield estimates with the corresponding official Canadian estimates indicate that both area and yield errors contributed significantly to the production underestimation. Both area and yield were significantly underestimated each month. However, the area error contributed more to the underestimation of production, as is indicated by inspection of the area and yield relative differences.

The tendency to underestimate spring wheat area was also observed in the U.S. spring wheat region. This underestimation, in both the United States and Canada, is partially the result of the inability to differentiate spring wheat from other small grains. Consequently, historical ratios of spring wheat acreage to small grains acreage were used to obtain

TABLE III.—Comparison of Previously Submitted Canada Reports and Revised Estimates

Province	Area, thousands of acres		Yield, bu/acre		Production, thousands of bushels	
	Original	Revised	Original	Revised	Original	Revised
<i>CAS Annual Report—January</i>						
Saskatchewan	13 546	13 511	29.1	29.5	394 274	398 722
Alberta	4 535	4 535	25.1	25.0	113 919	113 216
Manitoba	2 754	2 756	24.4	23.4	67 226	64 376
Total	20 835	20 802	^a 27.6	^a 27.7	575 419	576 314
<i>CAS Monthly Report—October 19</i>						
Saskatchewan	11 456	11 453	29.1	29.5	333 912	338 354
Alberta	3 767	3 767	24.9	24.6	93 615	92 731
Manitoba	2 124	2 125	24.5	23.4	51 950	49 817
Total	17 347	17 345	^a 27.6	^a 27.7	479 477	480 902
<i>CAS Monthly Report—September 1^a</i>						
Saskatchewan	11 205	11 202	28.9	29.6	324 068	331 793
Alberta	3 433	3 433	23.3	24.6	80 334	84 281
Manitoba	2 099	2 100	24.3	23.3	51 181	48 995
Total	16 736	16 735	^a 27.2	^a 27.8	456 583	465 069
<i>CAS Monthly Report—August 14</i>						
Saskatchewan	9 437	9 697	28.8	29.5	271 418	285 750
Alberta	1 952	2 099	23.2	23.5	45 334	49 235
Manitoba	1 800	1 751	25.2	23.1	45 409	40 400
Total	13 188	13 547	^a 27.5	^a 27.7	362 162	375 385

^aAverage.

spring wheat acreage estimates. These ratios were responsible for a significant amount of the underestimation observed for Canada in the August, September, and October estimates, since a majority of the ratios were developed from 1971 data and the planting of wheat in preference to nonwheat small grains had greatly increased since that time. For example, in the Province of Saskatchewan, which historically produces about 65 percent of the Canadian spring wheat, the ratio of spring wheat acreage to spring small grains acreage increased from 60 percent in 1971 to about 76 percent in 1976, an increase of 16 percent.

Incorporating the use of 1975 ratios of wheat to small grains for the January area estimate made a sig-

nificant improvement over the October estimate, but the January estimate was still significantly smaller than the Canadian estimate. This fact indicates that, as a result of more confusion crops, smaller fields, and a relatively short growing season, the spring small grains area, in Canada as in the United States, is also significantly underestimated. The strip-fallow cropping practice, which effectively creates smaller fields, leads to underestimation, since some of the strip-fallow fields are small compared to Landsat resolution and hence are difficult to detect and measure.

Another potential source of error in the LACIE spring wheat area estimation process is sampling. To date, this particular error has not been quantified for

Canada. The sampling plan was based on 1971 agricultural census data, and, for the same reason that the 1971 ratios of spring wheat to small grains were inappropriate, the sampling scheme employed may also have shortcomings. For example, the increase in the planting of wheat in preference to other small grains probably resulted in counties with sparse wheat acreage in 1971 having significantly more area planted to wheat in 1976. This, of course, would indicate the need for more sample segments in these areas. This particular situation occurred in the United States in Minnesota. The first allocation by

LACIE to estimate spring wheat production in the United States for crop year 1974-75 was based on 1969 agricultural census data and Minnesota was allocated 13 segments. When the sample allocation was redone for the 1976-77 crop year, the allocation was based on the 1974 agricultural census and Minnesota was allocated 47 segments. This increase in sample number for Minnesota was primarily due to the increase in area planted to wheat from 1969 to 1974.

Although spring wheat area appears to be the primary contributor to the production underestimate, the yield was also significantly underestimated.

Yield Accuracy

Since the Canadian government does not publish official yield and production figures until their September report, it was not possible through most of the growing season to pinpoint yield accuracy in the three wheat growing provinces. Yield predictions from the CCEA agrometeorological models were produced beginning in May for each of the four Manitoba, nine Saskatchewan, and three Alberta crop districts. Official Canadian wheat estimates, when they did become available, reported only province-level results. The CCEA yield models, on the other hand, predict only at the crop district level. Therefore, without access to current-year district area data to calculate the province-level yields, it was not possible to precisely track LACIE yields at the province level unless certain assumptions were made about 1976 crop district wheat area and distribution ratios. This was done; namely, the historical 5-year (1971-75) wheat area by distribution by crop district within a given province was assumed to be proportional to the 1976 ratio. Under this assumption then, it was possible to track province yields through the early phases of the season. After the initial LACIE area data became available in August, province-level yields were calculated using these figures and the results were then compared to the official Statistics Canada crop releases for the months of September and October (table V). Table VI gives similar data using the revised CCEA yield predictions and final end-of-season Canadian figures released in December 1976.

Based on fairly conclusive indications from the official Canadian source, it appears the LACIE models underestimated across the board in their initial year of operation. For the three Prairie Provinces com-

TABLE IV.—LACIE Revised Estimates of Spring Wheat Production, Area, and Yield for Canada Compared With Official Country Estimates

<i>(a) Production</i>					
Month	Official, ^a thousands of bushels	LACIE, thousands of bushels	CV, percent	RD, ^b percent	Test statistic (c)
August		375 385	6.5	-121.9	-18.8
September		465 069	5.2	-79.1	-15.2
October		480 902	4.8	-73.2	-15.3
January	833 000	576 314	4.9	-44.5	-9.1

<i>(b) Area</i>					
Month	Official, thousands of acres	LACIE, thousands of acres	CV, percent	RD, percent	Test statistic
August		13 547	5.8	-97.8	-16.9
September		16 735	4.0	-60.1	-15.0
October		17 345	3.1	-54.5	-17.6
January	26 800	20 802	3.2	-28.8	-9.0

<i>(c) Yield</i>					
Month	Official, bu/acre	LACIE, bu/acre	CV, percent	RD, percent	Test statistic
August		27.7	3.6	-12.3	-3.4
September		27.8	3.6	-11.9	-3.3
October		27.7	3.7	-12.3	-3.3
January	31.1	27.7	3.7	-12.3	-3.3

^aStatistics Canada.

^bRelative difference = (LACIE - official) ÷ LACIE × 100

^cThe LACIE estimate is significantly different from the official Canadian estimate.

TABLE V.—Canada: 1976 CCEA/LACIE Yield Estimates With Comparisons

(In bushels per acre)

Province	May	June	July	August		September				October			
	CCEA ^a	CCEA	CCEA	CCEA	LACIE ^b	CCEA	LACIE	Official ^c	RD ^d	CCEA	LACIE	Official ^c	RD ^d
Manitoba	28.0	27.6	27.5	25.5	25.2	24.3	24.4	27.4	-11.0	24.3	24.5	27.1	-9.6
Saskatchewan	24.9	24.5	28.8	28.7	28.8	28.9	28.9	30.2	-4.3	29.1	29.1	31.8	-8.5
Alberta	24.6	26.0	25.6	23.5	23.2	23.4	23.4	31.8	-26.4	25.1	24.9	31.8	-21.7
Three Prairie Provinces	25.3	25.3	28.0	27.1	27.5	27.1	27.2	30.1	-9.6	27.6	27.6	31.1	-11.2

^aCCEA began producing Canadian yield estimates by crop district for each of the three provinces in May 1976. A province-level yield estimate could not be obtained without like area data to aggregate. LACIE area estimates were not forthcoming until August; hence, a scheme to obtain the province yields earlier was initiated using 5-year (1971-75) average crop district wheat area and distribution statistics for each province. For comparison purposes, this technique was continued each month throughout the season.

^bInitial LACIE province yield estimate obtained by aggregating CCEA crop district yield and LACIE crop district area estimates.

^cStatistics Canada, Ministry of Industry, Trade and Commerce, Sept. 10, 1976.

^dRelative difference (in percent) = (LACIE - Statistics Canada) ÷ Statistics Canada × 100.

^eStatistics Canada, Ministry of Industry, Trade and Commerce, Oct. 8, 1976.

TABLE VI.—Canada: 1976 CCEA/LACIE Revised Yield Estimates With Comparisons

(In bushels per acre)

Province	August	September			October			Final		
	LACIE ^a	LACIE	Official	RD ^b	LACIE	Official	RD ^b	LACIE	Official ^c	RD ^b
Manitoba	23.1	23.3	27.4	-15.0	23.4	27.1	-14.0	23.4	27.1	-14.0
Saskatchewan	29.5	29.6	30.2	-2.0	29.5	31.8	-7.2	29.5	31.5	-6.3
Alberta	23.5	24.6	31.8	-22.6	24.6	31.8	-22.6	25.0	32.5	-23.1
Three Prairie Provinces	27.7	27.8	30.1	-7.6	27.7	31.1	-10.9	27.7	31.1	-10.9

^aLeDuc, Sharon. Yield-Weather Regression Models for the Canadian Provinces. LACIE-00433. NASA Johnson Space Center, Houston, Tex., Feb. 1976.

^bPercent.

^cStatistics Canada, Ministry of Industry, Trade and Commerce, Dec. 3, 1976.

bined, latest LACIE calculations place spring wheat yield above the normal, but below Canadian sources, maintaining a pattern established early in the crop season. The margin between the normal yield and the CCEA-predicted yield widened especially during June, largely because of abundant rainfall over nearly 75 percent of Canada's wheat producing region. The average of the Prairie Province yields predicted by the agrometeorological models then held at 27 to 28 bushels per acre through the end of the season, finishing 3 to 4 bushels per acre lower than the official Canadian estimate. September's estimate fell

within 10 percent of that from Statistics Canada; however, a boost in forecast production by the Canadians in October pushed the yield difference to almost 11 percent.

Although Manitoba's official yield (27.1 bushels per acre) was beneath the agrometeorological statistical trend (28.0 bushels per acre), the input of weather data to the model pushed LACIE predicted yields to more than 10 percent below official figures. This would indicate a collective overreaction by the Manitoba models to the known moisture shortage in that region the summer of 1976. Although rainfall

was off as much as 40 percent in parts of the province, the effect was apparently not as extensive as the yield models seemed to indicate. LACIE's final yield predictions for Manitoba were 14 percent below the final Canadian figure.

Across the border in Saskatchewan, overall accuracy improved. September comparisons showed only a 2-percent LACIE underestimation. By October, the margin had widened to just over 7 percent, due mostly to an upward 1.6-bushel-per-acre revision by the Canadians. The Saskatchewan models accurately reflected the season's favorable weather, although collectively they did not react as sharply as conditions appeared to warrant.

Of the yield predictions covering the three Prairie Provinces, the Alberta estimate showed the widest margin of error, underestimating the official spring wheat yield figure of 32.5 bushels per acre by more than 7 bushels per acre. Throughout the season, CCEA-predicted yields hovered around the trend line, checked by below-normal precipitation from May through July throughout 90 percent of the Alberta wheat producing sector. Combined, the three Alberta models appeared to be overly influenced by the July precipitation variable. When this is added to the large negative trend variable adjustment associated with the July truncation, yields in that month alone dropped over 2 bushels per acre, at a time when crops were apparently progressing well.

The CCEA Canadian agrometeorological yield models for 1976 collectively show a significant bias toward underestimation, both where crop conditions were considered good to excellent (Saskatchewan and Alberta) and where they were somewhat less favorable because of insufficient rainfall (Manitoba). In general, the model reaction to weather variables tended toward an overreaction to unfavorable weather and an underreaction to favorable weather (table VII). For the three Prairie Provinces, yield model performance for the 1976 season appeared to be just outside a 90-percent level of accuracy (89.1 percent).

TECHNICAL ISSUES

After an extensive evaluation of the major factors that affected the final LACIE Canadian estimates, two stand out as the principal sources of these lower-than-expected estimates:

1. A change in the area planted to wheat since 1971 which affected the wheat-to-small-grains ratios

TABLE VII.—LACIE Yield Model Reaction to 1976 Meteorological Conditions Compared With Official Source^a

Province	Crop season weather (b)	Yield model response (b)
Manitoba	Unfavorable	Overreaction
Saskatchewan	Favorable	Underreaction
Alberta	Favorable	Underreaction
Three Prairie Provinces	Favorable	Underreaction

^aStatistics Canada, Sept. 10, 1976.

^bDefinitions:

Favorable—model-predicted and official yields above normal trend.

Unfavorable—model-predicted and official yields below normal trend.

Overreaction—model-predicted yield deviates in same direction from normal as official yield but varies by a greater magnitude.

Underreaction—model-predicted yield deviates in same direction from normal as official yield but varies by a lesser magnitude.

used in the current aggregation system (1971 was the base year used for generating the LACIE reports)

2. An overall underestimate due to sampling and classification error

By updating the ratios with 1975 data, the area estimate increased to 20.8 million acres, or 22 percent below the official estimate. Ratios based on the 1976 official estimate improved the LACIE estimate by 4 percent, but this was still substantially below the official figure.

The following sections review the five major problems that were encountered during the Phase II Canadian analysis.

Allocation Data

The first sample segment allocation for Canada was done on the basis of the 1971 census and census district boundaries. The allocation used the sampled portion of the political area covered by the census, not the total area within the subdivision boundaries. To match the yield model geographic boundaries, a second allocation was done using the yield model areas as strata and census subdivisions as substrata. This new allocation showed that about 18 segments of the 283 should have been moved. However, no segments were moved, and the Group I and Group II substrata retained the segments as previously placed.

The LACIE aggregations for Canada in 1976 were done using as strata the crop districts, the boundaries of which essentially coincided with census district boundaries, except in Saskatchewan where the boundaries were slightly different. The yield model areas are aggregates of crop districts, and the calculated yields are assigned to all crop districts (area strata) contained within the yield model area.

The reallocation for Canada using yield strata resulted in the following distribution of segments.

Province	Total	Group I	Group II
Saskatchewan	170	153	17
Alberta	75	74	1
Manitoba	38	14	24

The Group I area estimate was unbiased. The Group II segments presented a problem because the segments were chosen with probabilities proportional to size based on the old allocation and aggregated with weights based on the new aggregation, thus incurring some bias. However, since Saskatchewan and Alberta contain 10 percent or fewer Group II segments, the bias is not considered significant. In Manitoba, 63 percent of the segments were Group II. The census district and crop district boundaries were the same, with yield models covering aggregates of the crop districts. These models provided a single yield estimate at a given time which was used for all crop districts covered by the yield model. Therefore, the aggregations performed during Phase II may be considered to have provided unbiased area estimates and valid area variance estimates. Further, the yield and production variances can be calculated at the yield strata level and higher because the boundaries coincide with multiples of crop districts except for a minor deviation in Alberta.

Historical Data

One of the major problems encountered in working Canada was the lack of historical data at the level at which the original allocation was generated. The original allocation was based on data contained in the 1971 Canadian Census of Agriculture. This census is conducted every 5 years, with publication approximately 2 years after data collection. This publication is the only one the LACIE project has been able to obtain that contains area statistics at the county level (no production statistics are contained in this cen-

sus). All other statistics produced by Statistics Canada are at the province or crop district level.

Most of the data on Canada available to LACIE besides the 1971 census consisted of statistics at the province or crop district level on wheat area and production, with spotty coverage of other crops. No comprehensive set of data for small grains or other crops was available. Data received from Statistics Canada in the late fall of 1976 greatly improved the data available at the province level, but data below that level are still lacking.

The importance of having a comprehensive data set over a period of years cannot be overstated. These data provide a valuable tool for understanding the changing patterns of agriculture. Such understanding serves as a base for comparison of LACIE estimates, helps improve ratioing procedures, and makes it possible to track long-term trends in agricultural policy. The lack of these data has caused major concern, especially in the development of ratios and the analysis of factors affecting the LACIE-generated estimates.

CAS Ratioing Methodology

Ratioing procedure.—During Phase II, Landsat data processed by CAMS was passed to CAS with classifications for spring small grains or small grains. Since the CAS aggregation system was designed to aggregate wheat, a procedure for ratioing was employed to derive wheat from small grains. Preprocessing software was written to accommodate this ratioing procedure. This software was designed to ratio the small grains estimates at the segment level. Two types of ratios were included in this preprocessing step for Canada: (1) spring wheat area to total small grains area and (2) spring wheat area to spring small grains area.

Ratios for Canada were derived using the 1971 census data, since it was the only source available that contained data on other crops at the county level (barley, oats, rye, mixed small grains). For each of the 283 sample segments, ratios were constructed on the basis of the county statistics for each segment. These ratios were used for the three LACIE reports generated in August, September, and October.

In October, additional data were received that updated LACIE statistics for Canada. This information included data through the 1975 crop season on other crops at the Canadian crop district level. These

figures were then used to construct ratios at the crop district level. Instead of each segment having a unique ratio based on statistics for the county in which that segment was located, ratios based on the 1975 data were common for all segments within each crop district. The derived wheat estimates based on this 1975 data were used to generate the final estimate contained in this report.

Changes in wheat area, 1971 to 1976.—One of the most critical elements in determining the reliability of a ratio procedure is the stability or variability of wheat area. Since the late 1960's, some major shifts have occurred in Canadian agriculture that have had decided effects on the ratioing methodology

employed by CAS. Since 1970, there has been a gradual shift toward increased planting of wheat. This is true not only in Canada but in all major wheat-growing countries. In the Prairie Provinces as a whole, this shift has been predominantly at the expense of reduced plantings of flax and rapeseed and reduced summer fallow acreage (tables VIII and IX). Only minor shifts in the proportion of wheat versus other small grains have occurred over the long run, but year-to-year variability is high.

In a review of tables VIII and IX at the province level, one of the most striking items is that, while areas devoted to all small grains and miscellaneous crops were at record levels in 1971, the area planted

TABLE VIII.—Changes in Crop Area in the Three Canadian Prairie Provinces

(In acres)

<i>Crop year</i>	<i>Wheat</i>	<i>Other small grains^a</i>	<i>Flax</i>	<i>Rapeseed</i>	<i>Summer fallow</i>
<i>Saskatchewan</i>					
1965-74 avg.	15 789 800	5 748 900	667 400	1 273 400	17 865 600
1971	12 923 117	8 700 290	924 785	2 736 555	16 559 825
1975	15 200 000	5 850 000	450 000	1 800 000	18 200 000
1976 (est.)	17 400 000	5 410 000	225 000	850 000	18 000 000
<i>Alberta</i>					
1965-74 avg.	4 993 900	7 246 600	307 300	1 068 800	7 238 900
1971	3 443 311	8 800 494	270 753	1 987 625	7 008 714
1975	4 500 000	7 800 000	200 000	1 700 000	6 900 000
1976 (est.)	5 600 000	7 950 000	100 000	850 000	6 700 000
<i>Manitoba</i>					
1965-74 avg.	2 823 400	3 179 500	850 300	309 800	2 878 400
1971	2 519 381	3 932 451	565 551	580 768	2 655 197
1975	3 100 000	2 902 000	750 000	750 000	2 600 000
1976 (est.)	3 800 000	3 052 000	550 000	250 000	2 600 000
<i>Total prairie</i>					
1965-74 avg.	23 607 200	16 175 000	1 825 000	2 652 000	27 982 900
1971	18 885 809	21 433 235	1 761 089	5 304 948	26 223 736
1975	22 800 000	16 552 000	1 400 000	4 250 000	27 700 000
1976 (est.)	26 800 000	16 412 000	875 000	1 950 000	27 300 000

^aOats, barley, rye, mixed small grains

to wheat was significantly reduced compared to recent years. The result is an unusually low wheat-to-small-grains ratio for 1971. In all the comparative data for Canada, 1971 seems to have been an unusual year as far as wheat was concerned.

In Saskatchewan, the wheat-to-small-grains percentage increased 16 percent from 1971 to the 1976 estimate but only 3 percent over the 10-year average. Fallow acreage has remained fairly constant except in 1971 when fallow acreage was substantially lower than the average.

Alberta has the lowest wheat acreage of the three provinces. In 1971, the wheat-to-small-grains ratio was 29 percent; it increased to 41 percent in the 1976

estimate, the same as the 10-year average. Summer fallow, rapeseed, and flax acreage has been continually declining since 1971, giving way to increased acreage planted to small grains.

Manitoba has followed very closely the pattern of Saskatchewan. Wheat acreage increased 15 percent between 1971 and the 1976 estimate. Summer fallow acreage has remained fairly stable since 1971, while wheat acreage increased at the expense of other small grains, flax, and rapeseed acreage.

While there have been dramatic changes in wheat acreage over the past several years, this fact does not fully account for the problems associated with the LACIE-generated estimates. The end-of-season re-

TABLE IX.—Changes in Wheat Area in the Three Canadian Prairie Provinces

Crop year	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Wheat, thou- sands of acres	Total small grains, ^a thou- sands of acres	<u>Col. 1</u> <u>Col. 2</u> percent	Total small grains and other crops, ^b thousands of acres	<u>Col. 1</u> <u>Col. 4</u> percent	Total small grains, other crops, and summer fallow, thou- sands of acres	<u>Col. 1</u> <u>Col. 6</u> percent
<i>Saskatchewan</i>							
1965-74 avg.	15 790	21 539	73	23 591	67	41 357	38
1971	12 923	21 623	60	25 285	51	41 845	31
1975	16 200	21 050	72	23 300	65	41 500	37
1976 (est.)	17 400	22 810	76	23 885	73	41 885	42
<i>Alberta</i>							
1965-74 avg.	4 994	12 240	41	13 617	37	20 856	24
1971	3 443	12 244	28	14 502	24	21 511	16
1975	4 500	12 300	37	14 200	32	21 100	21
1976 (est.)	5 600	13 550	41	14 500	39	21 200	26
<i>Manitoba</i>							
1965-74 avg.	2 283	6 003	47	7 216	39	10 094	28
1971	2 519	6 452	39	7 598	33	10 253	25
1975	3 100	6 002	52	7 527	41	10 127	31
1976 (est.)	3 800	6 852	55	7 684	49	10 284	37
<i>Total region</i>							
1965-74 avg.	23 607	38 012	62	42 554	55	70 537	33
1971	18 885	40 319	47	47 325	40	73 609	26
1975	22 800	39 352	58	45 027	51	72 727	31
1976 (est.)	26 800	43 212	62	46 069	58	73 369	37

^aTotal small grains includes wheat, barley, oats, rye, and mixed small grains

^bOther crops include flax, rape, and buckwheat

port (October) using 1971 ratios produced an area estimate of 17.3 million acres, 35 percent below the official Canadian estimate. Updating the ratio with 1975 data improved the estimate, but it was still 22 percent below the official figure. If the province-level 1976 estimates are used, only a slight improvement (4 percent) is realized. Econometric models were developed to predict the confusion crop ratios for Canada at the crop district level at the end of Phase II; however, because of the changes in project scope, these models were not used in Phase III. Thus, an 18-percent difference between the LACIE estimate and the official estimate remains. In all probability, this difference can be accounted for by sampling and classification error, since all other factors that might contribute to this underestimate have been analyzed.

Sampling and Classification

To date, the Accuracy Assessment Group has not completed the analysis of the Canadian intensive test site data to determine the type and magnitude of sampling/classification problems in Canada. Plans for completion of this analysis are still to be determined.

Some of the preliminary analyses performed on the U.S. spring wheat intensive test sites and blind sites have indicated that the classification analysts have tended to underestimate the amount of small grains within a segment. This tendency also seems to affect the Canadian estimates received during the 1976 crop season.

Interpretation of Canadian Landsat data was more difficult than interpretation of other areas because of the dissected topography, high confusion-crop problems, and the need to perform multitemporal analysis for improved identification. As a result of these major problems, interpretation of Canadian segments requires substantially more data and analysis time to produce a usable estimate. Since multitemporal analysis is almost a necessity in generating an estimate in Canada, a lack of adequate acquisitions can cause serious problems in the interpretation process.

Segment Acquisition and Image Analysis

An analysis of spring small grains area estimates for Canada, which were transmitted to CAS, can be summarized as follows.

1. Of the 170 sample segments allocated to Saskatchewan, 152 had spring small grains area estimates of greater than 0 percent; 7 had 0-percent estimates; and 11 had no acquisitions suitable for interpretation. About 80 percent of the 152 segments (122 segments) had no significant change in estimates between the time of early jointing and heading (about June 15 to July 15) and harvest (about mid-September).

About 10 percent of the 152 sample segments (15 segments) had a significant change in area estimate between the early estimate and the harvest estimate. In most of these segments, the area estimates increased, but in two the revised estimate was smaller. Approximately 10 percent of the 152 sample segments (16 segments) had no acquisitions at harvest.

It is believed that the distribution of overall acquisitions and the lack of at-harvest acquisitions on 10 percent of the sample segments for which estimates were transmitted to CAS had no significant impact on the area aggregation for Saskatchewan.

2. A total of 75 sample segments was allocated to Alberta; of these, 80 percent (60 segments) had area estimates transmitted to CAS. Only one sample segment was estimated to have 0-percent spring small grains. One segment had an increase in area estimate between early processing (about jointing stage) and harvest. Three segments had no at-harvest acquisitions.

There is no apparent reason for these conditions to have had a significant effect on the area aggregation for Alberta.

3. Of the 38 sample segments allocated to Manitoba, 2 segments had no acquisitions suitable for interpretation. The remaining 36 sample segments can be evaluated as follows:

a. About 60 percent (20 segments) had no significant change in estimates between early processing (jointing to heading) and harvest. About 25 percent (9 segments) had significant changes in area estimate between the early estimate and the at-harvest estimate.

b. Approximately 15 percent of the 36 segments (6 segments) had no at-harvest acquisitions and 1 segment had an at-harvest estimate only.

There are no indications that these conditions should have significantly affected the CAS aggregations for Manitoba.

Accuracy and Performance of LACIE Area Estimates

J. F. Potter,^a E. M. Hsu,^a A. G. Houston,^b and D. E. Pitts^b

INTRODUCTION

The accuracy assessment effort is designed to check the accuracy of the LACIE estimates of wheat production, area, and yield throughout the growing season to determine whether the operational procedures are sufficient to satisfy the LACIE project goals and to identify problem areas in the estimation process.

In this paper, the results obtained in assessing the accuracy of the LACIE acreage estimates in the United States are discussed. The accuracy assessment of yield and production estimates is discussed elsewhere in this volume (see the paper by Phinney et al. entitled "Accuracy and Performance of LACIE Yield Estimates in Major Wheat Producing Regions of the World," the paper by Marquis entitled "LACIE Area, Yield, and Production Estimate Characteristics: U.S. Great Plains," the paper by Hickman entitled "LACIE Area, Yield, and Production Estimate Characteristics: U.S.S.R.," and the paper by Conte et al. entitled "LACIE Area, Yield, and Production Estimate Characteristics: Canada").

Although the accuracy assessment studies discussed in this paper are conducted in the U.S. Great Plains region, these studies are also designed to promote the development of procedures that can be used to obtain accurate estimates for other parts of the world.

REGIONS OF THE U.S. GREAT PLAINS

In this paper, results are given for a number of regions within the U.S. Great Plains. These regions are defined as follows.

1. The U.S. Southern Great Plains (USSGP)

region consists of Colorado, Kansas, Nebraska, Oklahoma, and Texas. Only winter wheat estimates have been made for these states.

2. The spring wheat (SW) region consists of Minnesota and North Dakota. These states have very little winter wheat, so LACIE has made estimates for spring wheat only.

3. The mixed wheat (MW) region consists of Montana and South Dakota. These states have both spring and winter wheat.

4. The U.S. Northern Great Plains (USNGP) region consists of the two spring wheat states and the two mixed wheat states.

5. The U.S. Great Plains (USGP) region consists of the nine states of the USSGP and the USNGP.

PHASE I (CROP YEAR 1974-75)

Phase I of the LACIE project concentrated on the estimation of wheat acreage. Yield and production feasibility studies were also carried out, but the accuracy assessment team investigated only the accuracy of acreage estimation.

The 90/90 Criterion

Detailed discussions of the 90/90 criterion and the Phase I estimates are given in the paper by Marquis. It was found that the estimates for winter wheat in the U.S. Southern Great Plains did support the 90/90 criterion but that the total wheat estimates for the U.S. Northern Great Plains and the U.S. Great Plains regions did not.

Area Error Source Analyses

Comparison of LACIE and USDA SRS acreage estimates.—Table I shows the comparison of the LACIE

^aLockheed Electronics Company, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

TABLE I.—Comparison of SRS and LACIE At-Harvest Estimates of Wheat Area^a

Region	n/M ^b	SRS, thousands of acres	LACIE, thousands of acres	RD, ^c percent	CV, ^d percent	Test statistic
<i>Winter wheat</i>						
Colorado	24/32	2 260	3 058	26.1	20.8	
Kansas	55/84	12 100	12 940	6.5	7.1	
Nebraska	23/35	3 070	2 657	-15.5	28.0	
Oklahoma	29/40	6 700	6 906	3.0	11.2	
Texas	28/49	5 700	4 218	-35.1	32.6	
USSGP	159/290	29 830	29 779	-2	7.0	-0.03
<i>Spring wheat</i>						
Minnesota	9/13	2 844	2 150	-32.3	15.7	
North Dakota	42/65	10 213	5 853	-74.5	14.8	
SW states	51/78	13 057	8 003	-63.2	11.6	
<i>Total wheat</i>						
Montana	39/60	4 975	3 999	-24.4	25.9	
South Dakota	23/33	3 003	4 154	27.7	17.7	
MW states	62/93	7 978	8 153	2.2	15.6	
USNGP	113/171	21 035	16 156	-30.2	9.8	e-3.11
USGP	272/411	50 865	45 935	-10.7	5.7	e-1.88
Projected to national	272/637				3.7	

^aLACIE estimates based on CAMS rework data.

^bThe *n* is the number of segments used; the *M* is the number of segments allocated.

^cRelative difference = (LACIE - SRS) ÷ LACIE × 100.

^dCoefficient of variation = standard deviation ÷ LACIE × 100.

^eThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level.

and the U.S. Department of Agriculture Statistical Reporting Service (USDA SRS)¹ estimates. A test statistic is given showing whether the LACIE estimate is significantly different from the USDA estimate. The derivation and interpretation of the test statistic are described in the paper by Houston et al. entitled "Accuracy Assessment: The Statistical Approach to Performance Evaluation in LACIE."

For winter wheat in the USSGP region, the LACIE estimate is very close to the SRS estimate and, according to the statistical test, is not significantly different from it. For spring wheat in the SW states, the LACIE estimate is much lower than the SRS estimate. The LACIE estimate for total wheat in the MW states is slightly higher than the SRS esti-

mate; therefore, the large (and statistically significant) underestimate for total wheat at the USNGP level is due to the large underestimate for the SW states, especially North Dakota. The same is true for the statistically significant underestimate for total wheat at the USGP level. An analysis of the problem in North Dakota showed that the major source of error was sampling.

Study of classification and sampling error using ground-observed proportions.—The expression "blind site" is a designation applied to selected operational segments for which, unknown to the analyst, ground-truth data were acquired for evaluation purposes. The implementation of this approach occurred late in the growing season of LACIE Phase I. Thus, all of the selected sites were in the northern spring wheat regions.

High-resolution color-infrared aerial photography of 29 LACIE segments in North Dakota and Mon-

¹The Statistical Reporting Service has since become a part of the Economics, Statistics, and Cooperatives Service.

tana was acquired in mid-August 1975. Within a few days following the photography, field teams collected ground information for a substantial portion of these segments.

Figure 1 shows plots of the ground-observed segment proportions and the SRS county proportions versus the LACIE proportions for 16 blind sites in North Dakota. All proportions are for small grains. The LACIE estimate is fairly representative of the ground-truth proportion. Indeed, at the 10-percent level of significance, the average LACIE estimate is not significantly different from the average ground-truth proportion. However, it is clear that the LACIE estimate is not at all representative of the SRS county proportions. The ground-observed spring small-grains proportions are 38 percent below the corre-

sponding SRS county spring small-grains proportions. These results indicate that sampling error resulting from nonrepresentative sample segments was the major source of the observed bias in the acreage estimate for North Dakota. Other investigations with full-frame imagery confirmed that agriculture is very heterogeneous in this region and many of the LACIE segments did not adequately represent their counties.

Special Studies

In Phase I, a number of special studies were conducted to investigate various aspects of LACIE procedures. They are described in detail in reference 1. The major studies are summarized as follows.

Study of the effects of site, biophase, and AI.—One study was conducted to investigate the effects of three major factors—site, biophase,² and analyst interpreter (AI)—on errors in the estimation of segment small-grains proportions. All 14 AI's operating within the Classification and Mensuration Subsystem (CAMS) for the LACIE Phase I operations participated in this experiment. The test was run on two intensive test sites (ITS's): segment 1969, Toole County, Montana; and segment 1976, Franklin County, Idaho. These segments were selected because multispectral scanner (MSS) data were available for all four biophases. (Classifications for at least one biophase were missing for all the other ITS's.) Each AI was required to interpret each biophase acquisition for each segment using the Phase I operational procedure. This resulted in a total of 56 small-grains proportion estimates for each segment.

The analysis of the data produced the following results.

1. The error in proportion estimation varied significantly from one ITS to another.
 2. There was a significant difference in the relative performance between AI's from one segment to another.
 3. Use of biophase 1 increased the accuracy for one ITS and decreased it for the other.
- It is important to note that the experiment used only two sites so the results should not be widely applied.

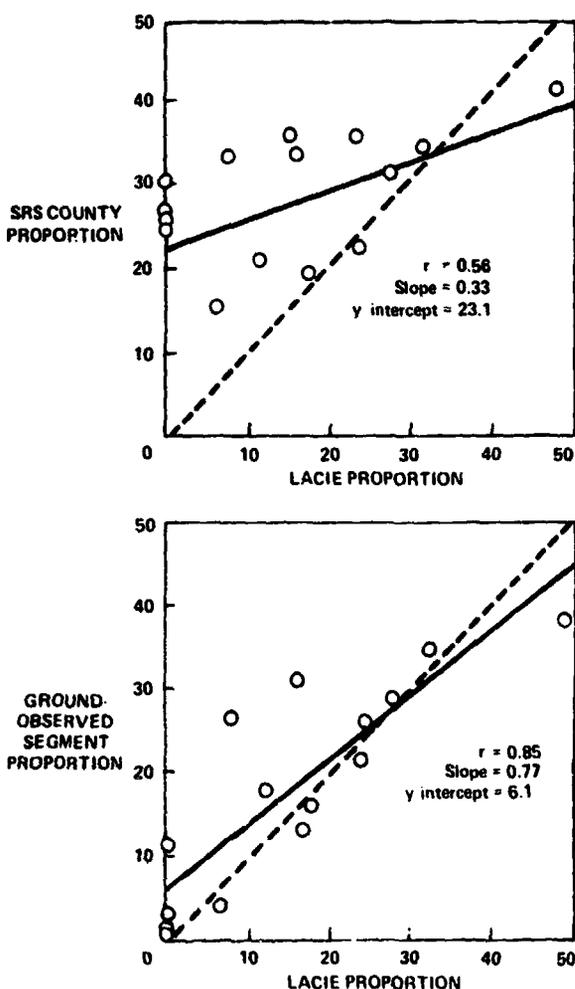


FIGURE 1.—Regression of ground-observed segment proportions and SRS county proportions on LACIE proportions of small grains.

²The four biophases in wheat development are defined as (1) tillering to jointing, (2) jointing to heading, (3) heading to soft dough, and (4) soft dough to harvest.

Four-AI study of the effect of small-grains proportion, amount of training data, and biophase.—In another study, four AI's working independently and using the CAMS rework procedures analyzed all of the acquisitions over the 23 Phase I ITS's that had acquisitions satisfying the CAMS rework criteria. The results were used to study (1) the effect of the proportion of small grains in the segment on proportion error, (2) the effect of the amount of training data on proportion error, and (3) the effect of biophase on labeling accuracy.

The results showed that the proportion of small grains in the segment had a pronounced effect on CAMS proportion error—the sites that were low in small grains tended to be overestimated and the sites that were high in small grains tended to be underestimated. A theoretical explanation of why this effect occurs is given in reference 1.

It was found that only limited information could be gained on the effect on proportion error of the amount of data used to train the classifier because the amount of training data selected by the AI's was very site dependent and proportion error was also very site dependent. It appeared that there was a slight reduction in proportion error as the number of training pixels increased.

In the investigation of the effect of biophase on labeling accuracy, results were obtained for eight biophase combinations. The best combination was biophases 1, 3, and 4. However, these results were not very accurate because only a few sites were averaged for each combination and labeling accuracies varied greatly from site to site.

Crop calendar verification.—To assess the performance of the adjustable crop calendar (ACC), the ACC's for 12 crop reporting districts (CRD's) having intensive test sites were compared to the corresponding historical crop calendars and to the development stages determined by ground observations on the ITS's.

The ACC performance during the jointing-to-soft-dough stage for winter wheat and the planting-to-soft-dough stage for spring wheat in the U.S. Great Plains appeared to be quite good. The biggest discrepancies were at the beginning of the period covered by the ACC—at jointing for winter wheat and at planting for spring wheat. An 8- to 10-day disagreement occurred between the dates the USDA reported for the CRD (which were used as starter dates for the ACC) and the ITS ground-truth data. The ITS ground-truth and ACC output were closest to agreement at the heading and soft-dough stages. Indica-

tions were that more accurate starter dates would have allowed the ACC to perform more accurately throughout the spring and summer.

The results of the study showed that

1. Accurate starter models for spring wheat are vital to good overall performance of the ACC.
2. Proper operation of the ACC for winter wheat before and during dormancy to provide an accurate estimate of jointing in spring is vital to the overall operation of the ACC for winter wheat.

PHASE II (CROP YEAR 1975-76)

As a result of the blind site investigations in North Dakota at the end of Phase I, 20 segments were added to North Dakota for Phase II to alleviate the sampling problem. Also, the analysis of previously unavailable Landsat imagery over large areas showed some sample segments to be in non-agricultural areas. These segments were moved into agricultural areas for Phase II. The blind site investigations also indicated a tendency to underestimate small-grains proportions. Therefore, for Phase II, the number of blind sites was increased substantially in order to investigate classification problems further.

Early in Phase I, it was found that the analysts could not reliably separate spring wheat from other small grains. This procedure required the use of historical ratios of wheat to small grains to obtain a wheat area estimate. In Phase II, the use of historical county-level ratios for wheat area estimation was continued.

In Phase II, LACIE estimates were available for acreage, yield, and production. Most of this section is devoted to a discussion of the accuracy assessment results on acreage estimation in the USGP yardstick region. However, it begins with a brief review of the 90/90 evaluation and the relative contribution of area and yield errors to the production error. A more complete discussion of the yield and production results is given in the papers by Phinney and Marquis, respectively.

The 90/90 Evaluation and Production Sensitivity Analysis

As in Phase I, it was found that the winter wheat production estimates supported the 90/90 accuracy goal but the spring wheat and total wheat production

estimates did not. Tables II and III show the sensitivity analysis results on the effect of errors in area and yield on the variability and bias of the production estimates. (The methods used to determine these results are described in the paper by Houston et al.) Table II indicates that yield error contributed slightly more to production estimate variability than area error; table III indicates that the underestimates in production were due primarily to underestimates of area.

Area Error Source Analyses

Comparison of LACIE and SRS acreage estimates.—These comparisons are designed to monitor how well LACIE is performing throughout the crop year and to detect any problems that may exist. The LACIE and SRS acreage estimates are shown in figure 2 and table IV. In the following discussion, winter wheat is considered first, followed by spring wheat, then total wheat. Figure 2 and table IV are arranged in this order.

For the major regions, a significance test was performed to determine whether the LACIE estimate

was significantly different from the SRS estimate. The test statistic is given in the last column of table IV and the method is described in the paper by Houston et al.

Winter wheat: Figures 2(a) to 2(d) show the acreage estimates for winter wheat. Figure 2(a) shows that the LACIE estimates for the USSGP region were lower than the SRS estimates for every month except June. Statistical tests showed that the LACIE estimates for February, March, and April were significantly different from the corresponding SRS estimates. These lower estimates are expected early in the season, because a number of wheat fields have not yet "greened up" enough to have a characteristic wheat signature. In 1976, this effect was especially apparent in Kansas, Oklahoma, and Texas because these states were affected by drought. In May and June, the LACIE estimate for the USSGP improved and was not significantly different from the SRS estimate from May through the final estimate. In June, it was closer to the final SRS estimate (which held from July on) than to the June SRS estimate. The final LACIE estimate had a relative difference (RD) of -6.3 percent and a coefficient of variation (CV) of 5 percent.

The most serious problem in the USSGP region was the underestimate for Oklahoma, shown in figure 2(b). Blind site investigations indicated that the major source of the underestimate in Oklahoma was analyst-mislabeled fields resulting from early dry conditions and an unusual wheat growth cycle following spring rains. The wheat was late in greening up and had signatures that were quite different from normal wheat. In fact, comparisons of LACIE blind site ground observations, aircraft photography, and analyst labels on a field-by-field basis indicated that the analysts rarely misidentified nonwheat fields as wheat, but the underestimate resulted from labeling wheat fields as nonwheat.

The winter wheat acreage estimates for the two mixed wheat states are shown in figure 2(c). These estimates were very low in June but increased throughout the season. The RD for the final estimate was -14.7 percent and the CV was 19 percent.

Figure 2(d) shows the total USGP winter wheat estimates. The final LACIE estimate had an RD of -7.3 percent and a CV of 5 percent. July was the only month for which the LACIE estimate was significantly different from the SRS estimate. Thus, there was a tendency to underestimate winter wheat but it was significant only for July. This tendency was mainly due to underestimation in Oklahoma.

TABLE II.—Relative Contribution of Area and Yield Errors to Variability of Production Estimate

Region	Total CV, percent	CV, percent	
		LACIE acreages × SRS yields	LACIE yields × SRS acreages
Winter wheat	7.0	4.5	5.3
Spring wheat	10.0	6.3	7.5
Total wheat	5.2	3.7	4.4

TABLE III.—Relative Contribution of Area and Yield Errors to Bias of Production Estimate

Region	Total RD, percent	RD, percent	
		LACIE acreages × SRS yields	LACIE yields × SRS acreages
Winter wheat	-7.2	-7.6	-1.1
Spring wheat	-22.3	-29.1	+6.3
Total wheat	-12.3	-14.9	+1.5

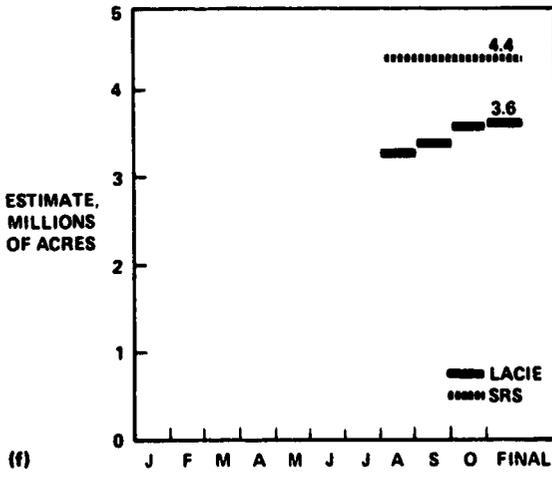
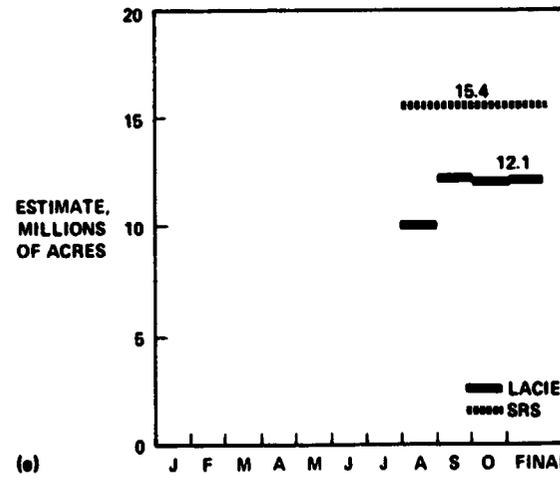
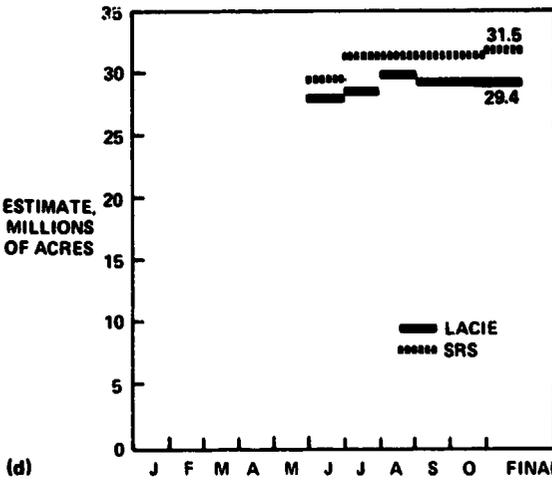
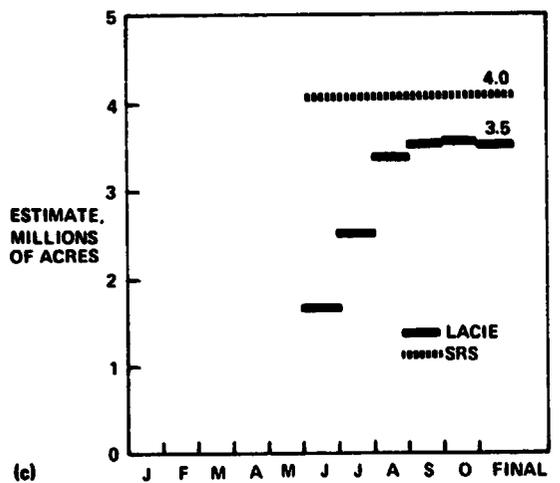
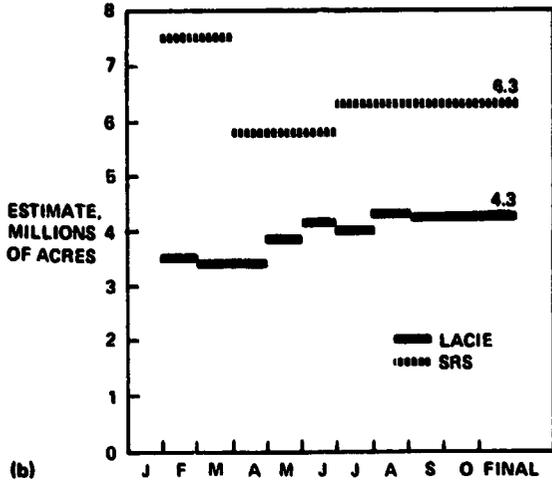
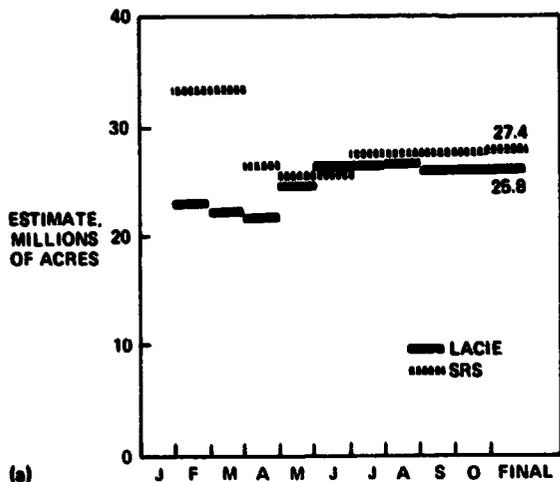


FIGURE 2.—LACIE and SRS acreage estimates. (a) USSGP, winter wheat. (b) Oklahoma, winter wheat. (c) Mixed wheat states, winter wheat. (d) USSGP, winter wheat. (e) Spring wheat states, spring wheat. (f) Mixed wheat states, spring wheat. (g) USSGP, spring wheat. (h) USNGP, total wheat. (i) USSGP, total wheat.

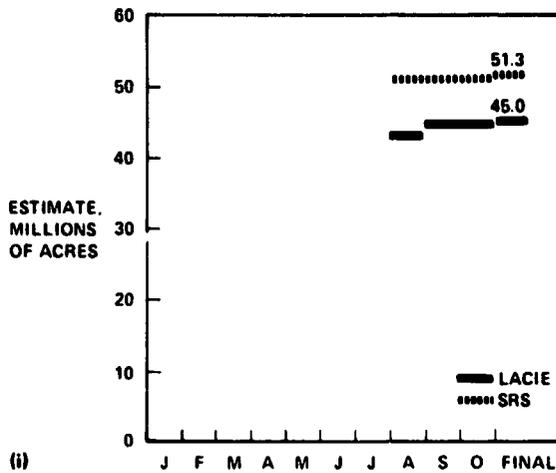
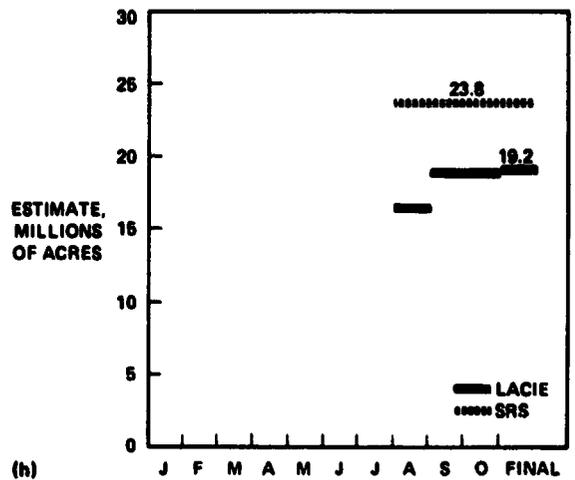
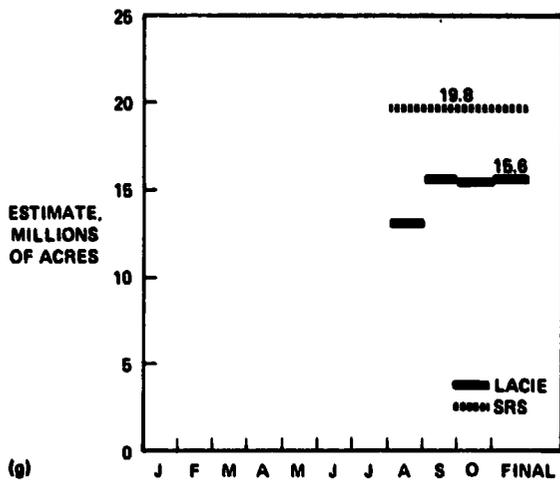


FIGURE 2.—Concluded.

Spring wheat: Figure 2(e) shows the spring wheat estimates in the two spring wheat states, Minnesota and North Dakota. There was consistent underestimation by LACIE, but there was a considerable improvement in September. Part of this improvement was due to a change in the ratios of wheat to small grains that were used to calculate the wheat acreage. For spring wheat, CAMS normally determines only small-grains proportions, and the wheat proportions are then calculated by multiplying these by the historical wheat-to-small-grains ratios for the county in which the segment is located. A change to ratios based on 1975 data accounted for 48 percent of the improvement in North Dakota and 53 percent of the improvement in Minnesota. In North Dakota, a further 36 percent of the improvement was due to the addition of 21 new segments. These new seg-

ments were added to North Dakota to correct the sampling problem identified during Phase I. There was also an undersampling problem in Minnesota, since the spring wheat area had increased from 829 000 acres in 1969 (the year that was used for the sampling allocation) to 2 844 000 acres in 1976. Blind site investigations indicated a number of causes of the underestimate in North Dakota, including poor Landsat resolution of strip-fallow areas, weak or missing signatures, and poor acquisition histories.

Figure 2(f) shows the spring wheat estimates for the two mixed wheat states, Montana and South Dakota. These estimates were consistently low, but they did improve as the season progressed. The improvement was partly due to improved spring-wheat-to-small-grains ratios. The final spring wheat estimate for the mixed wheat states had an RD of -21.1

TABLE IV.—Comparison of SRS and LACIE Acreage Estimates

(a) February^a

<i>Region</i>	<i>n/M</i>	<i>SRS, thousands of acres</i>	<i>LACIE, thousands of acres</i>	<i>RD, percent</i>	<i>CV, percent</i>	<i>Test statistic</i>
<i>Winter wheat</i>						
Colorado	13/32	2 830	3 539	20.0	26	
Kansas	43/84	13 100	8 013	-63.5	12	
Nebraska	13/35	3 400	4 500	24.4	18	
Oklahoma	30/40	7 550	3 499	-90.0	24	
Texas	31/49	6 300	3 170	-98.7	25	
USSGP	130/240	33 180	22 721	-46.0	9	b-5.1

(b) March^a

<i>Region</i>	<i>n/M</i>	<i>SRS, thousands of acres</i>	<i>LACIE, thousands of acres</i>	<i>RD, percent</i>	<i>CV, percent</i>	<i>Test statistic</i>
<i>Winter wheat</i>						
Colorado	25/32	2 830	2 768	-2.2	25	
Kansas	61/84	13 100	8 536	-53.5	8	
Nebraska	21/35	3 400	3 632	6.4	13	
Oklahoma	36/40	7 550	3 450	-118.8	18	
Texas	42/49	6 300	3 725	-69.1	30	
USSGP	185/240	33 180	22 111	-50.1	8	b-6.3

(c) April

<i>Region</i>	<i>n/M</i>	<i>SRS, thousands of acres</i>	<i>LACIE, thousands of acres</i>	<i>RD, percent</i>	<i>CV, percent</i>	<i>Test statistic</i>
<i>Winter wheat</i>						
Colorado	25/32	2 040	2 768	26.3	25	
Kansas	62/84	11 000	8 536	-28.9	8	
Nebraska	22/35	3 400	3 583	5.1	13	
Oklahoma	36/40	5 800	3 450	-68.1	18	
Texas	44/49	3 900	3 479	-12.1	20	
USSGP	189/240	26 140	21 816	-19.8	7	b-2.8

^aThe SRS estimates for February and March are the December 1975 estimates of seeded acreage.

^bThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level.

percent and a CV of 12 percent. The results presented in table IV show that there was an underestimation problem in Montana, where the RD for the final estimate was -54.0 percent and the CV was 22 percent. Investigations indicated that the underestimates in Montana were underestimates of wheat

proportions in strip-fallow areas, which did not classify well because Landsat resolution is not fine enough to resolve the fields.

The monthly estimates for the total spring wheat in the USGP region are shown in figure 2(g). The LACIE estimates were consistently low and were

TABLE IV.—Continued

<i>(d) May</i>						
<i>Region</i>	<i>n/M</i>	<i>SRS.</i> <i>thousands</i> <i>of acres</i>	<i>LACIE.</i> <i>thousands</i> <i>of acres</i>	<i>RD. percent</i>	<i>CV. percent</i>	<i>Test</i> <i>statistic</i>
<i>Winter wheat</i>						
Colorado	26/32	1 900	2 807	32.3	24	
Kansas	70/84	10 800	9 392	-15.0	6	
Nebraska	27/35	2 950	3 653	19.2	13	
Oklahoma	38/40	5 800	3 897	-48.8	16	
Texas	47/49	3 900	4 810	18.9	14	
USSGP	208/240	25 350	24 559	-3.2	6	-0.5
<i>(e) June</i>						
<i>Region</i>	<i>n/M</i>	<i>SRS.</i> <i>thousands</i> <i>of acres</i>	<i>LACIE.</i> <i>thousands</i> <i>of acres</i>	<i>RD. percent</i>	<i>CV. percent</i>	<i>Test</i> <i>statistic</i>
<i>Winter wheat</i>						
Colorado	26/32	1 900	2 995	36.6	23	
Kansas	75/84	10 750	10 535	-2.0	6	
Nebraska	30/35	2 950	4 104	28.1	12	
Oklahoma	38/40	5 300	4 148	-39.8	14	
Texas	47/49	3 900	4 556	14.4	15	
USSGP	216/240	25 300	26 338	3.9	5	-0.8
Montana	10/38	3 020	488	-518.9	193	
South Dakota	8/10	1 040	1 159	10.3	43	
MW states	18/48	4 060	1 647	-146.5	65	
USGP	234/288	29 360	27 985	-4.9	6	-8
<i>(f) July</i>						
<i>Region</i>	<i>n/M</i>	<i>SRS.</i> <i>thousands</i> <i>of acres</i>	<i>LACIE.</i> <i>thousands</i> <i>of acres</i>	<i>RD. percent</i>	<i>CV. percent</i>	<i>Test</i> <i>statistic</i>
<i>Winter wheat</i>						
Colorado	30/32	2 200	2 867	23.3	25	
Kansas	78/84	11 100	10 795	-2.8	6	
Nebraska	32/35	3 000	4 133	27.4	11	
Oklahoma	40/40	6 300	4 025	-56.5	15	
Texas	47/49	4 700	4 314	-8.9	15	
USSGP	227/240	27 300	26 134	-4.5	5	-0.1
Montana	21/38	3 020	1 044	-189.3	52	
South Dakota	9/10	1 040	1 482	29.8	23	
MW states	30/48	4 060	2 526	-60.7	25	
USGP	257/288	31 360	28 660	-9.4	5	^b -1.9

^bThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level

TABLE IV.—Continued

(B) AUGUST

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	RD, percent	CV, percent	Test statistic
<i>Winter wheat</i>						
Colorado	31/32	2 200	2 830	22.3	24	
Kansas	78/84	11 100	10 932	-1.5	5	
Nebraska	32/35	3 000	4 086	26.6	11	
Oklahoma	40/40	6 300	4 305	-46.3	15	
Texas	47/49	4 700	4 310	-9.0	16	
USSGP	228/240	27 300	26 463	-3.2	5	-0.6
Montana	22/38	3 020	1 911	-58.0	35	
South Dakota	9/10	1 040	1 482	29.8	23	
MW states	31/48	4 060	3 393	-19.7	22	
USGP	259/288	31 360	29 856	-5.0	5	-1.0
<i>Spring wheat</i>						
Minnesota	10/13	3 826	1 741	-119.8	40	
North Dakota	31/85	11 540	8 161	-41.4	14	
SW states	41/98	15 366	9 902	-55.2	13	
Montana	14/22	2 315	1 127	-105.4	28	
South Dakota	14/23	2 050	2 169	5.5	12	
MW states	28/45	4 365	3 296	-32.4	12	
USNGP	69/143	19 731	13 198	-49.5	10	b-5.0
<i>Total wheat</i>						
Montana	36/60	5 335	3 038	-75.6	19	
South Dakota	23/33	3 090	3 651	15.4	13	
MW states	59/93	8 425	6 689	-26.0	11	
USNGP	100/191	23 791	16 591	-43.4	9	b-4.8
USGP	328/431	51 091	43 054	-18.7	5	b-3.7

^bThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level.

significantly different from the SRS estimates for every month and for the final estimate. Of the four states contributing to the total spring wheat estimate, only South Dakota's spring wheat acreage was not consistently underestimated. This record indicated a serious underestimation problem for spring wheat. In addition to the reasons given previously, blind site studies indicated that this underestimation was also

due to errors in the ratios of wheat to small grains that were used to calculate the wheat acreage.

Total wheat: Figure 2(h) shows the total wheat estimate in the four-state USNGP region. It was consistently underestimated and was significantly different from the SRS estimate for every month and for the final estimate. The final estimate had an RD of -24.2 percent due to underestimates of spring

TABLE IV.—Continued

(h) September

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	RD, percent	CV, percent	Test statistic
<i>Winter wheat</i>						
Colorado	32/32	2 200	2 704	18.6	24	
Kansas	61/84	11 100	10 989	-1.0	5	
Nebraska	33/35	3 000	3 399	11.7	11	
Oklahoma	40/40	6 300	4 261	-47.9	14	
Texas	47/49	4 700	4 344	-8.2	16	
USSGP	233/240	27 300	25 697	-6.2	5	-0.4
Montana	35/38	3 020	2 103	-43.6	29	
South Dakota	9/10	1 040	1 452	28.4	23	
MW states	44/48	4 060	3 555	-14.2	20	
USGP	277/288	31 360	29 252	-7.2	5	-1.4
<i>Spring wheat</i>						
Minnesota	10/13	3 826	2 551	-50.0	27	
North Dakota	67/85	11 540	9 650	-19.6	5	
SW states	77/98	15 366	12 201	-25.9	7	
Montana	19/22	2 315	1 291	-79.3	23	
South Dakota	18/23	2 050	2 095	2.1	13	
MW states	37/45	4 365	3 386	-28.9	12	
USNGP	114/143	19 731	15 587	-26.6	6	b-4.4
<i>Total wheat</i>						
Montana	54/60	5 335	3 394	-57.2	14	
South Dakota	27/33	3 090	3 547	12.9	12	
MW states	81/93	8 425	6 941	-21.4	9	
USNGP	158/191	23 791	19 142	-24.3	6	b-4.1
USGP	391/431	51 091	44 839	-13.9	4	b-3.5

^bThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level.

wheat in Montana, Minnesota, and North Dakota and of winter wheat in Montana. The CV was 5 percent.

Figure 2(i) shows the total wheat estimate in the nine-state USGP region. The LACIE estimate was consistently low and was significantly different from the SRS estimate for every month and for the final estimate. The final LACIE figure had an underesti-

mate of 2.2 million acres in the winter wheat acreage and an underestimate of 4.1 million acres in the spring wheat acreage.

Studies based on ground-observed proportions.—In Phase II, ground-observed proportions were obtained for 103 winter wheat segments and 33 spring wheat segments called "blind sites" because the AI's did not know the identity of these sites. The ground-ob-

TABLE IV.—Continued

(1) October

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	RD, percent	CV, percent	Test statistic
<i>Winter wheat</i>						
Colorado	32/32	2 200	2 704	18.6	24	
Kansas	81/84	11 100	10 989	-1.0	5	
Nebraska	33/35	3 000	3 399	11.7	11	
Oklahoma	40/40	6 300	4 261	-47.9	14	
Texas	47/49	4 700	4 344	-8.2	16	
USSGP	233/240	27 300	25 697	-6.2	5	-1.2
Montana	36/38	3 020	2 131	-41.7	28	
South Dakota	9/10	1 040	1 452	28.4	23	
MW states	45/48	4 060	3 583	-13.3	19	
USGP	278/288	31 360	29 280	-7.1	5	-1.4
<i>Spring wheat</i>						
Minnesota	11/13	3 826	2 198	-74.1	30	
North Dakota	79/85	11 540	9 735	-18.5	5	
SW states	90/98	15 366	11 933	-28.8	7	
Montana	20/22	2 315	1 487	-55.7	24	
South Dakota	19/23	2 050	2 079	1.4	13	
MW states	39/45	4 365	3 566	-22.4	12	
USNGP	129/143	19 731	15 499	-27.3	6	b-4.6
<i>Total wheat</i>						
Montana	56/60	5 335	3 618	-47.5	12	
South Dakota	28/33	3 090	3 531	12.5	12	
MW states	84/93	8 425	7 149	-17.8	8	
USNGP	174/191	23 791	19 082	-24.7	5	b-4.9
USGP	407/431	51 091	44 779	-14.1	4	b-3.5

^bThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level

served proportions were used to study various aspects of acreage estimation.

Bias due to classification (weighted analysis): Ground-truth information from blind site data obtained at harvest was used to estimate the bias in the aggregated acreage estimates that was due to classification. The procedure is described in the paper by Houston et al. and the results are given in table V. It is called "weighted" analysis because the

proportions for the various segments are multiplied by the weights used in the aggregation. These results show how errors in proportion estimates affect the aggregated acreage estimates.

For winter wheat, there is significant underestimation due to classification in both the USSGP and USGP regions, mainly caused by problems in Oklahoma. Also, there is significant underestimation in both spring wheat in the USNGP region and total

TABLE IV.—Concluded

(i) Final

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	RD, percent	CV, percent	Test statistic
<i>Winter wheat</i>						
Colorado	30/32	2 200	2 774	18.6	24	
Kansas	81/84	11 300	11 525	-1.6	5	
Nebraska	33/35	2 950	3 399	13.2	11	
Oklahoma	40/40	6 300	4 261	-47.9	14	
Texas	47/49	4 700	4 344	-8.2	16	
USGP	233/240	27 450	25 833	-6.3	5	-1.3
Montana	36/38	3 080	2 079	-48.1	28	
South Dakota	9/10	970	1 452	33.2	23	
MW states	45/48	4 050	3 531	-14.7	19	
USGP	278/288	31 500	29 364	-7.3	5	-1.5
<i>Spring wheat</i>						
Minnesota	11/13	3 893	2 198	-77.1	30	
North Dakota	79/85	11 520	9 856	-16.9	5	
SW states	90/98	15 413	12 054	-27.9	7	
Montana	20/22	2 335	1 516	-54.0	22	
South Dakota	19/23	2 020	2 079	2.8	13	
MW states	39/45	4 355	3 595	-21.1	12	
USNGP	129/143	19 768	15 649	-26.3	6	b-4.4
<i>Total wheat</i>						
Montana	56/60	5 415	3 595	-50.6	12	
South Dakota	28/33	2 990	3 531	15.3	12	
MW states	84/93	8 405	7 126	-17.9	8	
USNGP	174/191	23 818	19 180	-24.2	5	b-4.8
USGP	407/431	51 268	45 013	-13.9	4	b-3.5

^bThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level

wheat in the USGP region. These results agree with the SRS comparisons discussed previously.

Bias due to classification (unweighted analysis): This section contains five studies of proportion estimation error. The term "unweighted" is used to indicate that the aggregation weights are not involved.

1. End-of-season winter wheat proportion estimation error. By October, data had been obtained for

103 blind sites in the five-state winter wheat region. An investigation was performed using these data and the CAMS classification results corresponding to the October LACIE estimates. The results are shown in figure 3 and tables VI and VII.

Figure 3 shows plots of the proportion error $\hat{\lambda} - X$ as a function of X , where $\hat{\lambda}$ is the CAMS wheat proportion estimate and X is the ground-truth wheat proportion. These plots are for the five individual

TABLE V.—Estimates of the Bias and Relative Bias of the LACIE Acreage Aggregation Estimates Using Blind Sites

Region	LACIE acreage estimate \hat{A} , thousands of acres	Aggregated acreage bias \hat{B} , thousands of acres	Relative bias \hat{B}/\hat{A} , percent	Standard deviation of \hat{B}	90-percent confidence limits for \hat{B}
<i>Winter wheat</i>					
Colorado	2 704	-26	-1.0	275.6	
Kansas	11 125	-988	-8.9	473.2	
Nebraska	3 399	199	5.9	381.4	
Oklahoma	4 261	-2 583	-60.6	590.9	
Texas	4 344	-483	-11.1	953.9	
USSGP	25 833	-3 881	-15.0	1305.6	(-6 029, -1 733)
USSGP (excluding Oklahoma)	21 572	-1 298	-6.0	1164.2	(-3 213, 617)
Montana	2 079	-913	-43.9	768.9	
South Dakota	1 452	-470	-32.4	255.9	
USGP	29 364	-5 264	-17.9	1536.6	(-7 792, -2 736)
<i>Spring wheat</i>					
Minnesota	2 198	-2 275	-103.5	908.2	
Montana	1 516	-827	-54.6	393.3	
North Dakota	9 856	-2 385	-24.2	801.9	
South Dakota	2 079	-31	-1.8	592.0	
USNGP	15 649	-5 524	-35.3	1404.6	(-7 835, -3 213)
<i>Total wheat</i>					
USGP	45 013	-10 788	-24.0	2078.2	(-14 207, -7 369)

states and the total USSGP five-state region. Points lying above the horizontal line $\hat{X} - X = 0$ correspond to overestimation of wheat proportions by CAMS, and points lying below the line correspond to underestimation.

The plots in figure 3 indicate that there is an overall trend toward negative values of $\hat{X} - X$ as X increases for the five-state region and for each of the individual states except Colorado. In other words, for these regions, CAMS tends to underestimate the true wheat proportion when the true wheat proportion is large. In fact, for $X > 28$ percent, there was

only 1 blind site out of 26 in the 5-state region for which the CAMS result was not an underestimate relative to ground truth. Also, figure 3 indicates that underestimates occurred in Oklahoma and Texas for all values of X . In Oklahoma, 17 of 20 blind sites were underestimated, as were 15 of 19 in Texas.

A statistical analysis of the data shown in figure 3 was performed using the technique described in the paper by Houston et al. The results are shown in table VI. It lists the following factors: (1) the number of blind sites for which data were available for each state or region; (2) the number of segments allocated

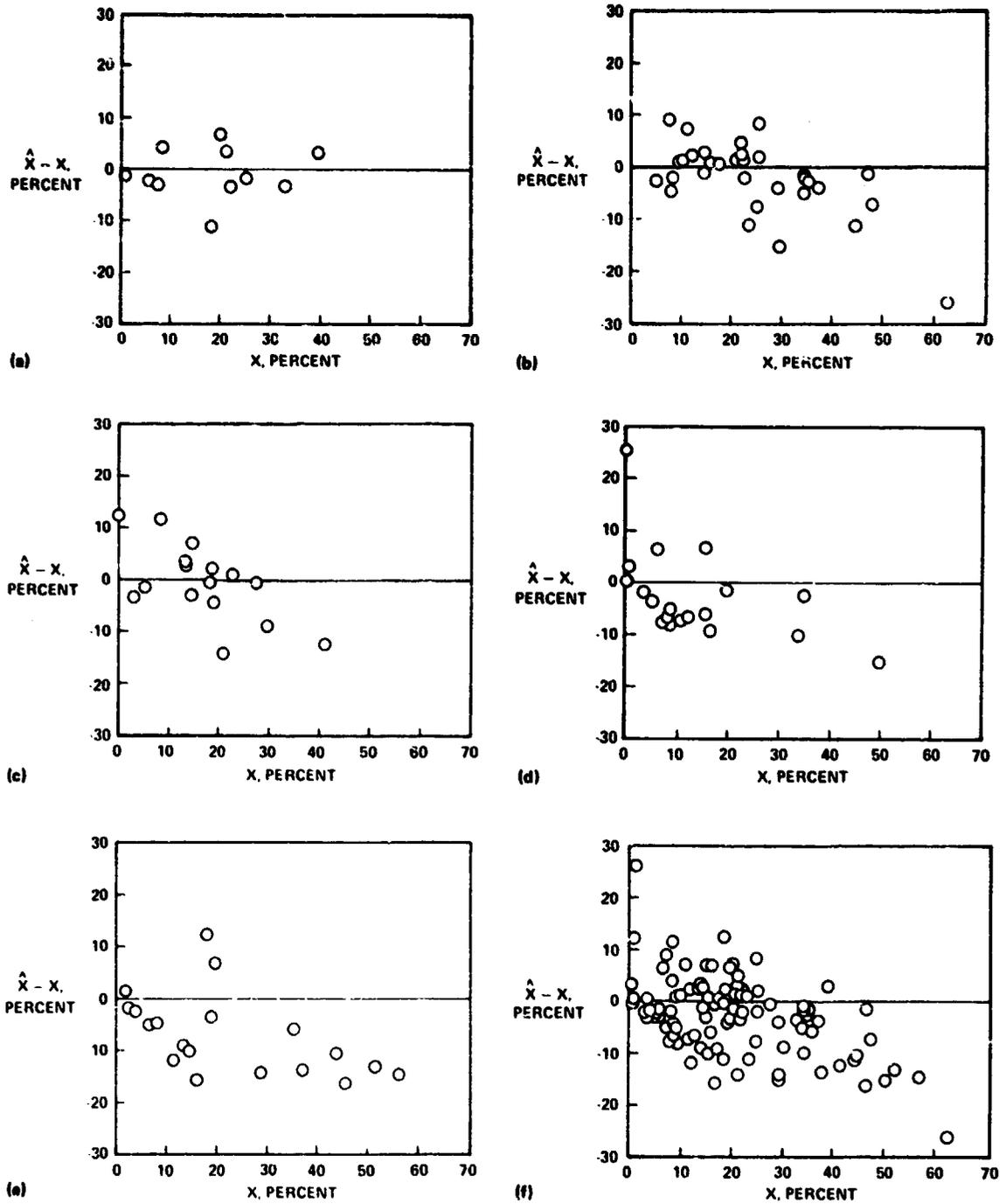


FIGURE 3.—Plots of winter wheat proportion estimation errors versus ground-truth winter wheat proportions for blind sites in the USSGP. (a) Colorado. (b) Kansas. (c) Nebraska. (d) Oklahoma. (e) Texas. (f) USSGP.

TABLE VI.—Winter Wheat Blind Site Results for the USSGP

Region	n ^a	N ^b	\bar{X}	$\bar{\hat{X}}^c$	\bar{D}	S \bar{D}	90-percent confidence limits for μ_D^d
Colorado	13	32	14.7	14.5	-0.1	1.0	(-1.8, 1.6)
Kansas	34	84	23.9	22.3	-1.6	.9	(-3.1, -0.1) ^e
Nebraska	18	35	14.1	14.8	.7	1.1	(-1.2, 2.6)
Oklahoma	20	40	24.7	17.6	-6.6	1.5	(-9.2, -4.0) ^e
Texas	20	49	13.3	11.9	-1.4	1.4	(-3.6, 0.9)
USSGP	105	240	19.1	17.2	-1.9	.6	(-2.9, -1.0) ^e

^aNumber of blind sites.

^bNumber of segments allocated.

^cWinter wheat estimates from the October Crop Assessment Subsystem (CAS) Monthly Report (CMR).

^dPopulation average difference.

^eSignificantly different from zero at the 10-percent level of significance.

TABLE VII.—Comparison of LACIE Estimates and Ground-Observed Proportions in Winter Wheat Blind Sites in the USGP

Month	Number of segments	MSE ^a	\bar{D}^b percent	RMD, ^c percent	Percentage underestimated ^d
February	71	157.5	-6.5	-30.6	83
March	95	112.8	-5.4	-26.2	79
April	95	112.8	-5.4	-26.2	79
May	95	102.5	-4.4	-21.4	75
June	95	89.5	-3.3	-15.7	72
July	95	90.4	-3.4	-16.2	70
August	95	75.0	-3.2	-15.2	71
September	95	65.3	-2.8	-13.3	68
October	95	69.6	-2.8	-13.7	68
Final	95	70.8	-2.7	-13.2	68

^aMSE = $\{\sum(\hat{X}_i - X_i)^2\}/n$ where \hat{X}_i is the wheat proportion estimate for the *i*th segment, X_i is the ground-observed harvested wheat proportion for the *i*th segment, and *n* is the number of segments.

^b $\bar{D} = \{\sum(\hat{X}_i - X_i)\}/n = \bar{\hat{X}} - \bar{X}$.

^cRMD = \bar{D}/\bar{X} .

^dThis column contains the percentage of blind site segments in which LACIE underestimated the wheat proportions.

to each state or region; (3) the average ground-truth wheat proportion \bar{X} ; (4) the average CAMS wheat proportion estimate $\bar{\hat{X}}$; (5) the average difference $\bar{D} = \bar{\hat{X}} - \bar{X}$; (6) the standard error $S_{\bar{D}}$ of \bar{D} ; and (7) 90-percent confidence limits for the average error μ_D .

In order to determine if the population average difference for a particular region is significantly different from zero, one needs only to observe whether the corresponding confidence interval contains zero. If it does, the average difference is not significantly different from zero; i.e., there is insufficient evidence to conclude that there is a bias due to classification error. If it does not contain zero, then

the hypothesis of no bias is rejected at the 10-percent level of significance.

In the following paragraphs, the results presented in table VI are discussed separately for each state and for the USSGP region. The discussion also includes preliminary results from an investigation by CAMS to determine the causes of classification error. At the end of the 1976 crop year, the data for one-half of the blind sites in the USGP were released to CAMS for evaluation of the accuracy and sources of error in the operational analysis during Phase II. These evaluations were carried out in most cases by the analyst who conducted the original interpretation and

classification. In the following paragraphs these studies will be referred to as the "CAMS investigation." The biggest overall problem for both spring and winter wheat was the occurrence of unusual wheat signatures, which were taken to be nonwheat.

The results for Oklahoma (table VI) indicate that there is a large negative bias in the CAMS estimates for the segments allocated to Oklahoma. The CAMS investigation showed that underestimates were due to atypical, weak, and missing signatures, small fields, and spotty stands. Some of these effects were attributed to drought conditions. Only one of the segments checked in the CAMS investigation was overestimated; hail damage of wheat at harvest was the cause of the overestimate. Figure 4 shows an example of wheat signature variability due to drought.

In table VI, it appears that a "significant" bias occurs for the state of Kansas. However, inspection of the data plotted in figure 3 reveals one outlier, a difference of -25.56 percent, corresponding to a ground truth of 61.56-percent wheat. Omitting this one outlier yields an estimate of the bias that is not significantly different from zero. From the CAMS investigation, it was concluded that, in Kansas, overestimates were due to pasture, fallow, and sorghum being included as wheat. Underestimates were usually caused by missed wheat signatures; i.e., wheat signatures that were not included in the training data.

For Texas, 79 percent of the blind sites were underestimated. However, the S_b was so large that

there was insufficient evidence to conclude that a bias existed. Inspection of the data plotted in figure 3 for Texas reveals an outlier, a difference of +25.31 percent, corresponding to a ground truth of 0.69 percent. This extreme overestimate was due to fallow fields and pasture fields which appeared red and tan, respectively, on the imagery and were classified as wheat. The underestimates that occurred for most of the segments were generally due to atypical signatures. Some stands of wheat were spotty.

Neither of the average differences for the other two states, Colorado and Nebraska, were significantly different from zero, nor were any apparent outliers observed. The analysts in CAMS were apparently having some success in identifying wheat for these two states. The CAMS investigation showed that, in Colorado, overestimates were caused by confusion crops such as spring wheat and winter rye being classified as winter wheat; underestimates were caused by missed signatures in drought areas and by strip-cropping areas not resolvable by the Landsat system. In the latter case, the wheat pixels were all essentially border pixels and therefore many were misclassified as nonwheat.

In Nebraska, overestimates were caused by atypical wheat signatures and small fields. Underestimates in Nebraska were due to missed signatures, the absence of key acquisitions such as biowindow 2, some narrow fields that were missed, and some wheat fields that were not identified on the imagery.

At the USSGP five-state level, there was sufficient

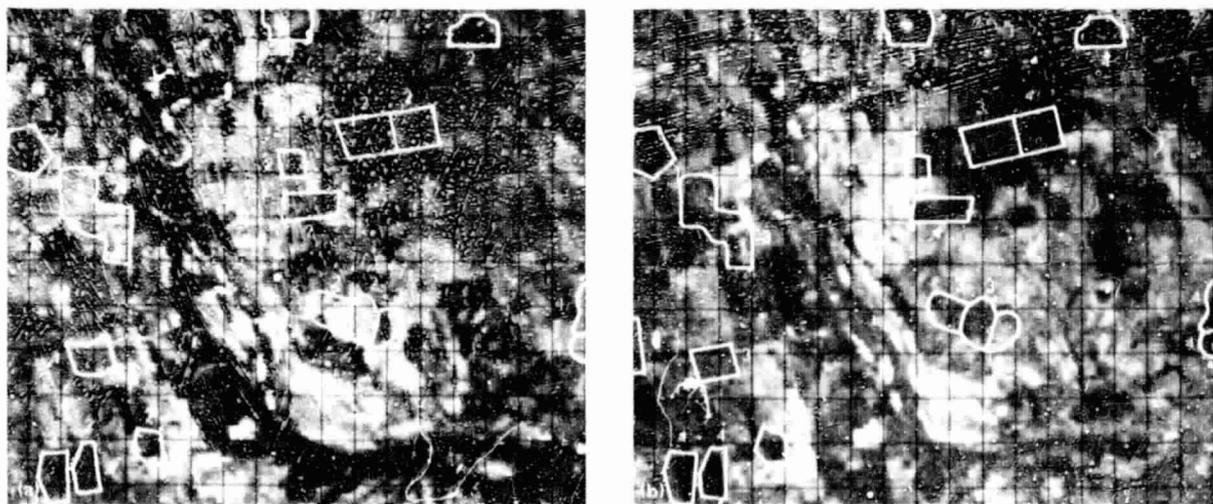


FIGURE 4.—Wheat signature variability due to drought—segment 1232, Kiowa, Oklahoma, 1976 crop year. The numbers indicate the biowindows. (a) November 24, 1975. (b) April 16, 1976.

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

evidence to conclude that the CAMS wheat proportion estimates were significantly different from the ground-observed wheat proportions at the 10-percent level, mainly as a result of problems in Oklahoma. These results agree with those obtained in the weighted analysis.

2. Variation of winter wheat proportion error throughout the season. Table VII presents the results of a blind site investigation to study the variation of classification error throughout the season.

At the time this investigation was performed (December 1976), all blind site data were available, but not all of the segments could be used since CAMS estimates for the whole season were not available for all of them. It is, of course, desirable that the same number of segments be used for each month. It was found that 95 segments had data for March through the end of the season but only 71 segments had data for February.

In table VII, four quantities relating to the classification error are given: (1) the mean-squared error *MSE*, (2) the mean difference \bar{D} , (3) the relative mean difference *RMD*, and (4) the percentage of the segments in which LACIE underestimated the at-harvest wheat proportions. There was a declining trend in the *MSE* throughout the season. The final figure represents a 55-percent reduction in error from that of the February estimate.

The \bar{D} and the *RMD* showed the same behavior; i.e., a general reduction in the size of the error as the season progressed. These errors were all negative, indicating underestimates by LACIE. From February through the final estimate, there was a 58-percent reduction in the magnitude of the \bar{D} and a 57-percent reduction in the magnitude of the *RMD*.

The percentage of segments underestimated by LACIE also decreased throughout the season, from 83 percent in February to 68 percent for the final estimate. All these estimates thus indicate a general improvement in the CAMS estimates as the season progressed.

3. End-of-season spring wheat proportion estimation error. The spring wheat blind site investigation was conducted for 38 segments in the four USNGP states of Minnesota, Montana, North Dakota, and South Dakota. Figure 5 shows plots of the proportion error $\hat{X} - X$ as a function of X , where \hat{X} is the CAMS wheat proportion estimate and X is the ground-truth wheat proportion.

The plots in figure 5 show a tendency toward underestimation in every state except South Dakota. Twenty-nine of the thirty-eight sites in the USNGP were underestimated by CAMS. In the plot for the USNGP, there appears to be a slight dependence on the value of X (i.e., the underestimates seem to be greater for larger values of X), but this trend is less pronounced than that shown in figure 3 for the USSGP.

The statistical analysis of these data is presented in table VIII. The quantities listed are the same as those in table VI.

For the blind sites in the USNGP, the analysis indicated a significant bias in the CAMS wheat proportion estimates. These results agree with those of the weighted analysis. Table VIII shows that the LACIE acreage estimates were low for all of the states except South Dakota, the only state for which the average difference is not significantly different from zero at the 10-percent level of significance. For Minnesota, Montana, and South Dakota, the number of data points was small. Therefore, no inference about the population average difference between CAMS estimates and ground-truth proportions should be made.

In Minnesota, underestimation generally occurred in segments with very high wheat density and was caused by unusual wheat signatures on the imagery. There is some evidence that these unusual signatures were the result of color distortions in the Landsat imagery processing.

In Montana, underestimation was usually due to strip-fallow areas that were not appropriately classified. Some overestimates were due to hay being classified as wheat, even though the two were not confused in the training fields.

In South Dakota, both overestimates and underestimates were caused by drought conditions. There was noticeable difference between the Landsat data for this area and for the USSGP. In the spring, wheat and small grains appeared very similar to pasture, alfalfa, and corn on the imagery because of the stress caused by drought. At harvest time, some corn was grazed or cut for silage and some alfalfa was cut and, because of the drought, never reappeared. In both cases, it was difficult to distinguish these crops from harvested small grains. Many small grains were not harvested but were fall-plowed and could not be distinguished from harvested small grains by CAMS; therefore, wheat was overestimated. Underestimates

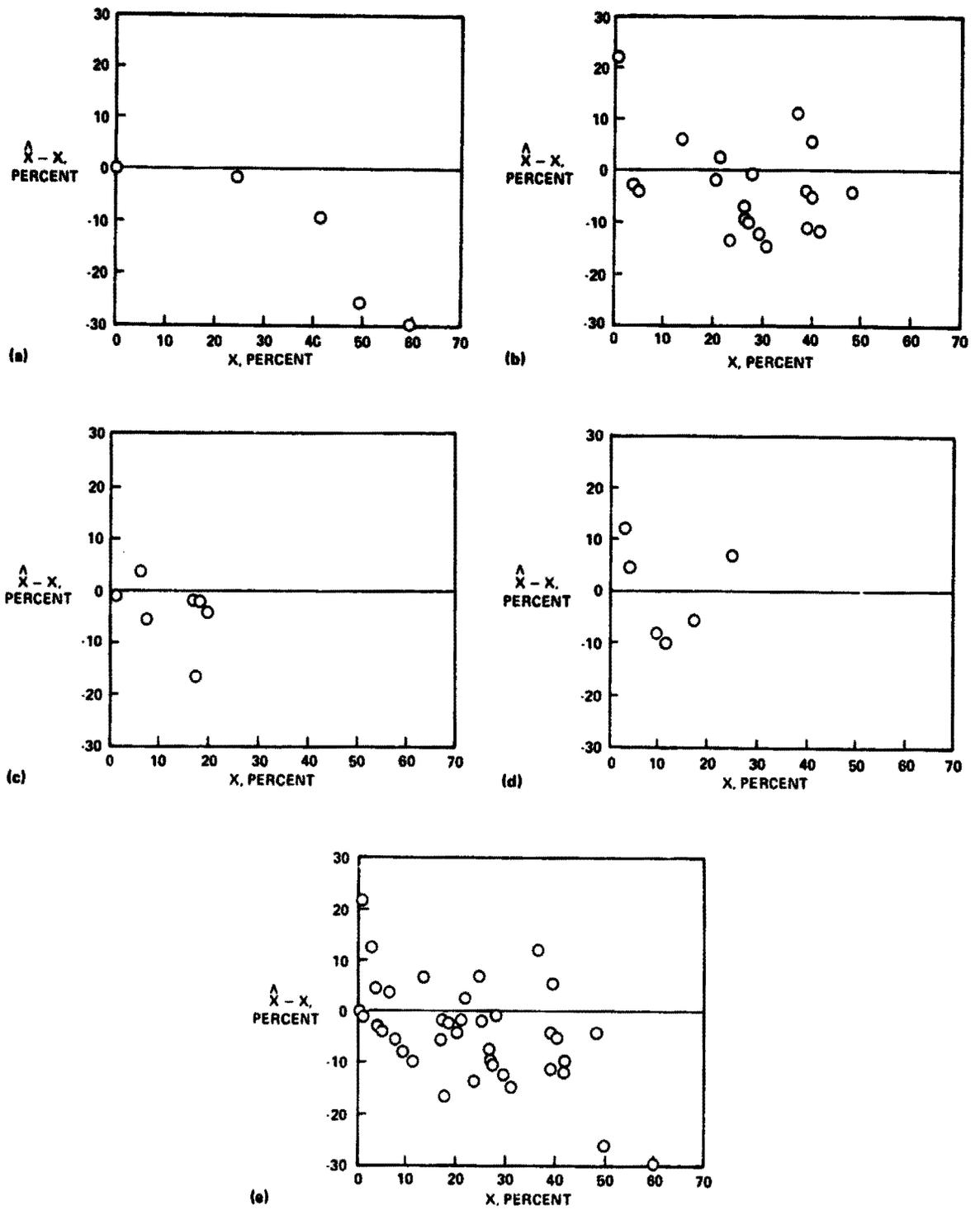


FIGURE 5.—Plots of spring wheat proportion estimation errors versus ground-truth spring wheat proportions for blind sites in the USNGP. (a) Minnesota. (b) North Dakota. (c) Montana. (d) South Dakota. (e) USNGP.

TABLE VIII.—Spring Wheat Blind Site Results for the USNGP

Region	n	M	\bar{X}	\hat{X}^a	D	SD	90-percent confidence limits for μ_D
Minnesota	5	13	35.4	22.8	-12.6	5.0	(-20.8, -4.4) ^b
North Dakota	20	85	27.1	23.0	-4.1	1.95	(-7.6, -0.7) ^b
Montana	7	22	12.7	9.1	-3.6	2.0	(-6.9, -0.3) ^b
South Dakota	6	23	11.3	11.5	.1	3.0	(-4.9, 5.1)
USNGP	38	143	23.1	18.6	-4.5	1.4	(-6.7, -2.2) ^b

^aFinal estimates from the CAS annual report for the 1976 crop year.

^bSignificantly different from zero at the 10-percent level of significance.

were due to missing signatures from poor stands of small grains and poor acquisition histories.

The CAMS investigation found many factors that contributed to the underestimate in North Dakota. Among these were (1) strip-fallow areas unresolvable by the Landsat system, (2) weak or missing signatures, (3) poor color balance in Landsat images due to the transformation that is applied to the Landsat data before the images are made, (4) the absence of early biowindow acquisitions, (5) the omission of some late-planted spring wheat because its signature was behind the jointing signature being indicated by the adjustable crop calendar, and (6) problems in choosing training fields caused by small fields or the absence of identifiable field patterns.

4. Contribution of the classification and ratio errors to the ratioed wheat proportion estimation error at the segment level. The CAMS makes estimates of the small-grains proportion \hat{X}_i for each segment i and, subsequently, the Crop Assessment Subsystem (CAS) obtains wheat proportion estimates by multiplying the \hat{X}_i by ratios \hat{R}_i of the wheat-to-small-grains proportions for the counties in which the segments are located. These county-level ratios were determined from the 1975 SRS estimates. In this section, the blind site data are used to compare the error incurred by using these ratios to the error incurred by misclassification of small grains. The method used is described in the paper by Houston et al.

Table IX presents the numerical results obtained for 37 spring wheat blind sites for Phase II in Minnesota, Montana, North Dakota, and South Dakota. It can be seen that the reduction in bias is slightly larger when there is no ratioing error than when there is no small-grains classification error. On the other hand, a much larger reduction in mean-squared error is obtained when there is no small-grains

classification error than when there is no ratioing error. This indicates that the variability in spring wheat proportion estimation errors is primarily due to classification of small grains. The historical wheat-to-small-grains ratios, however, introduced more bias than did small-grains classification errors.

5. Variation of spring wheat proportion error throughout the season. Table X shows the results of a blind site investigation of the variation of classification error throughout the season. Only 33 of the 38 segments were used. The definitions of the quantities listed are the same as those given in table VII.

The mean-squared classification error dropped from 158.5 in August to 110.1 at the end of the season—a decrease of 30 percent. The average difference was negative for all months, indicating that the wheat proportions were consistently underestimated throughout the year. The magnitude of the errors declined 45 percent in the period from August to the final estimate. In spite of these reductions, there was still substantial underestimation at the end of the season. At that time, the wheat proportion in 79 percent of the sites was still being underestimated by LACIE.

Acree estimation bias due to nonsampled and non-responsive areas.—In order to investigate bias due to nonsampled and nonresponsive areas, an aggregation was performed in which the CAMS proportion for each allocated segment was replaced by the 1975 SRS county wheat proportion for the county containing that segment. In table XI, the results of this “mock aggregation” are compared with the SRS estimates for winter wheat in the USSGP and total wheat in the USNGP and the USGP. The RD at the USGP level is 0.8 percent. This is an estimate of the relative bias due to Group II estimation and Group II₁ ratioing of those counties not allocated segments (see the paper

TABLE IX.—Phase II Final Results for Spring Wheat Blind Sites in the USNGP

Category	\bar{D} , percent	Standard error of \bar{D} , percent	Reduction in bias, percent	90-percent confidence limits for bias	MSE	Reduction in MSE, percent
Phase II final result	-5.2	1.3	—	(-7.4, -3.1)	104.5	—
No ratioing error	-2.2	1.2	57.7	(-4.3, -0.2)	78.6	24.8
No classification error	-3.1	.6	40.4	(-4.0, -2.1)	25.7	75.4

TABLE X.—Measurements of Classification Error for Spring Wheat (LACIE Estimates Versus Ground-Observed Proportions) Over all Available Blind Sites in the USGP

Month	Number of segments	MSE	\bar{D} , percent	RMD, percent	Percentage underestimated ^a
August	33	158.5	-9.29	-41.6	88
September	33	120.1	-5.72	-25.6	82
October	33	115.3	-5.38	-24.1	79
Final	33	110.1	-5.05	-22.6	79

^aThis column contains the percentage of blind site segments in which LACIE underestimated the wheat proportion.

TABLE XI.—Acreage Estimation Bias Due to Nonsampled Areas

Region	M	1975 SRS, thousands of acres	Mock aggregation, thousands of acres	RD, percent
USSGP	240	29 748	30 422	2.2
USNGP	191	21 035	20 768	-1.3
USGP	431	50 783	51 190	.8

by Hallum et al. entitled "Sampling, Aggregation, and Variance Estimation for Area, Yield, and Production in LACIE" for definitions of Group II and Group III). For practical purposes, this RD is negligible and shows that the underestimates in the LACIE estimates were not caused by the methods used to estimate nonsampled areas. In fact, not all the allocated segments were processed during Phase II for various reasons. Those areas which had an allocated segment that was not processed are called non-responsive areas.

Table XII shows the aggregation of county SRS estimates for crop year 1974-75 for all segments processed (394) during LACIE Phase II (1975-76). Since 91.4 percent of the allocated segments were processed in Phase II, table XII differs only slightly

from table XI. The RD at the USGP level is 0.1 percent. Therefore, practically speaking, the relative bias due to Group II estimation and Group III ratioing (of both those counties whose allocated segments were lost to nonresponse and those counties not allocated segments) is negligible.

When the results of table XI are combined with those of table XII, an estimate of the relative bias due to the Group III ratioing of the 37 counties whose allocated segments were lost to nonresponse is -0.7 percent, which is also negligible.

When the results of table XII are compared with the results of table XI, the RD between the mock aggregation estimate of wheat acreage at the USGP level and the SRS estimate is negligible whether or not segments are lost to nonresponse. Also, the esti-

TABLE XII.—Acreage Estimation Bias Due to Nonsampled and Nonresponsive Areas

Region	N	1975 SRS, thousands of acres	Mock aggregation, thousands of acres	RD, percent
USSGP	233	29 748	30 208	1.5
USNGP	161	21 035	20 637	-1.9
USGP	394	50 783	50 845	.1

mate of bias due solely to segments lost to non-response is negligible, indicating that nonresponse is probably not introducing a bias.

Special Studies

Several special studies were performed in Phase II to investigate various aspects of LACIE proportion estimation procedures. These are described in detail in reference 1 and are summarized below.

Dependence of CAMS error on acquisition date.—Two investigations were carried out to determine the relationship between the latest available acquisition and proportion estimation error. In the first, the errors for blind site wheat proportions in the USGP were studied as a function of the month of the latest acquisition used by CAMS to obtain their estimate of wheat proportions. All of the winter wheat blind sites in the USGP for which data were available were used. Spring wheat was not studied because sufficient ground-truth data were not available. Table XIII gives the mean-squared error, the bias, and the standard deviation of the errors for each month from November 1975 to July 1976. The errors were relatively large for estimates made with the latest acquisitions being from November through February. They decreased sharply with March acquisitions and remained relatively small through the end of the season.

In the second study, the CAMS proportion errors for intensive test sites were plotted as a function of the date of the last acquisition used to classify the data. This was done separately for spring and winter wheat. The plots are displayed in figure 6. For winter wheat, the estimates based on very early acquisitions (before December) had very large errors (mostly underestimates). For later acquisitions, the errors were smaller. However, there was no well-defined depen-

dence on acquisition date. For spring wheat, there was a tendency toward underestimation for early acquisitions and overestimation for late acquisitions.

Effect of biophase on proportion estimation.—A study was conducted in Phase II to investigate the effect of various biophase combinations on proportion estimation. Table XIV gives estimates of the bias and standard deviation of the proportion errors that were obtained from blind sites analyzed using the various biophase combinations. Only winter wheat blind sites in the USGP were used. Spring wheat blind sites were not studied because sufficient data were not available. The best results were obtained using data from the biophase combinations 1-2 and 1-2-3. In every case studied, the magnitude of the bias and the standard deviation were increased by adding biophase 4 data, except for the combination 1-3 where the magnitude of the bias increased but the standard deviation remained the same. These results indicated that better estimates might be obtained if data from biophase 4 were not used.

Study of labeling and classification errors.—An investigation of labeling and classification accuracy was conducted for 14 winter wheat intensive test sites and 10 spring wheat intensive test sites. The data consisted of 15 wheat fields and 15 nonwheat fields in the ground-observed area of each ITS. These fields were used to determine the probability of correct classification (PCC) by comparing the classification results for these fields with ground truth on a pixel-by-pixel basis. Labeling error was studied by determining the percentage of training fields in the ground-observed area that was labeled correctly.

For winter wheat, it was found that both labeling accuracy and PCC were considerably higher for non-

TABLE XIII.—Full-Month Classification Error for Winter Wheat

Acquisition period	MSE	Bias, percent	Standard deviation, percent	No. of sites
November	120.1	-4.5	10.1	36
December	161.8	-5.0	11.8	47
January	114.9	-5.5	9.3	61
February	123.5	-5.7	9.6	60
March	80.5	-1.3	8.9	64
April	45.2	-3.3	5.9	63
May	70.2	-9	8.4	82
June	84.3	-2.9	8.8	88
July	48.3	-6	7.0	58

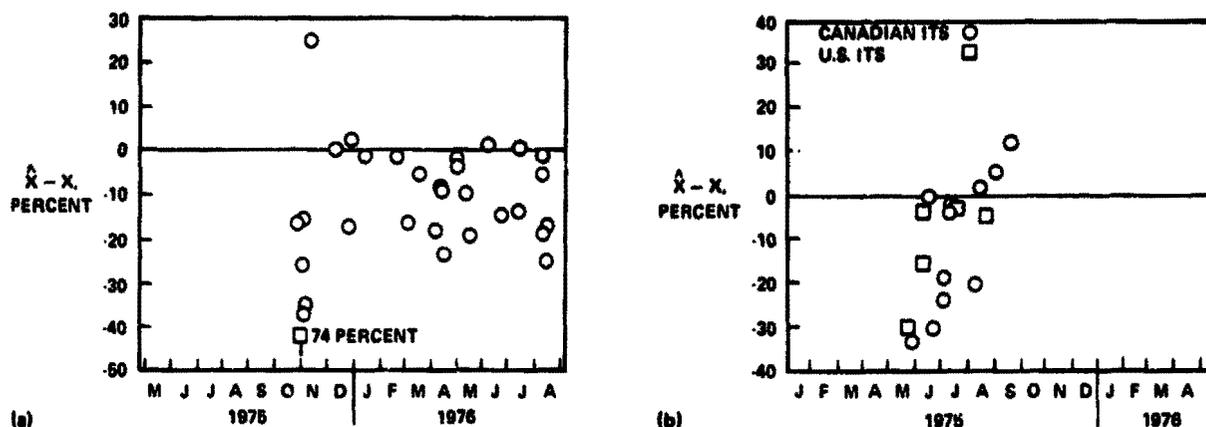


FIGURE 6.—Plots of CAMS error as a function of acquisition date. (a) Winter wheat. (b) Spring wheat.

small-grains than for small grains. For non-small-grains, the PCC and labeling accuracy were 87 percent and 95 percent, respectively; for small grains, the PCC was 63 percent and the labeling accuracy was 86 percent. The results suggested that the lower value for the PCC for small grains was because the analyst missed some small-grains signatures. This problem was probably the major cause of the underestimation in Phase II.

For spring wheat, only PCC was studied. The average value for small grains (81.9 percent) was smaller than the average for non-small-grains (93.4 percent), but the difference was less than that observed for winter wheat.

Adjustable crop calendar error.—The ACC is designed to indicate to the CAMS analyst the growth stage of wheat and other crops in the segments being analyzed. It can therefore be expected to have a considerable impact on the accuracy of the CAMS estimates.

A study was performed to determine the accuracy of the ACC by comparing it with ground-observed growth-stage data over eight ITS's in Texas and Kansas. In most cases, the LACIE growth stage estimate was behind the ground-observed growth stage and the difference increased as the season progressed. By June, all ACC predictions were behind the ground-observed stages.

Subsequently, an investigation was performed to determine whether crop calendar error had an influence on the accuracy of CAMS estimates. The classification errors were correlated with crop calendar errors for 9 winter wheat sites and 12 spring wheat sites. Significance tests applied to the correlation coefficients indicated that no significant correlation existed between crop calendar error and classification error.

PHASE III (CROP YEAR 1976-77)

The Phase II blind site results indicated a tendency to underestimate winter wheat proportions and a significant underestimation of spring wheat proportions. These, of course, led to a negative bias in the area and production estimates for total wheat. Thus, at the beginning of Phase III, the sample strategy for the U.S. Great Plains was revised to achieve a production estimator CV of 5 percent in order to allow for some bias and still meet the 90/90 accuracy goal.

The spring wheat blind site analyses indicated that a major portion of the negative bias in the spring wheat proportion estimates was due to the historical ratios of spring wheat to small grains used in reducing small-grains proportion estimates to spring wheat

TABLE XIV.—Classification Error by Biowindow Combination (Winter Wheat)

Combination	Bias, percent	Standard deviation, percent	No. of sites
1	-2.5	9.2	117
1-2	-8	6.8	72
1-3	-5.1	6.6	19
1-2-3	.8	4.9	32
1-4	-6.1	14.1	19
1-2-4	-2.0	7.9	33
1-3-4	-5.5	6.6	17
1-2-3-4	1.1	5.1	31

proportion estimates. Therefore, a task was initiated early in Phase III to develop econometric models for forecasting these ratios with the intent of eliminating or reducing this bias (see the paper by Umberger et al. entitled "Econometric Models for Predicting Confusion Crop Ratios" for a detailed description of these models).

During Phase III, a new classification procedure, Procedure 1, was introduced to address other problems identified as a result of LACIE experience through Phase II. Procedure 1 provided the first capability to process multigate Landsat acquisitions in a high-throughput mode. The difficulty in obtaining accurate area estimates in regions with small fields had been identified as a major problem. With the single-pixel training approach used in Procedure 1, it was believed that more accurate wheat proportion estimates could be made.

With the advent of the new approach, the blind site program was expanded in Phase III for more detailed classification error analyses. Correct labeling at the pixel level is the key to the success of Procedure 1. Therefore, the blind site program was modified to allow a comparison of analyst pixel labels with ground-observed crop types.

In Phase III, LACIE estimates of wheat area, yield, and production continued to be made in the U.S. Great Plains region. This section is devoted to the area error source analyses conducted in the U.S. Great Plains region during Phase III. However, a brief review of the 90/90 evaluation and relative contributions of area and yield errors to the production error is presented first. The descriptions of the yield and production results are detailed in the papers by Phinney and Marquis, respectively.

The 90/90 Evaluation and Production Sensitivity Analysis

As in Phases I and II, the LACIE Phase III winter wheat production estimate for the USGP supported the 90/90 accuracy goal. On the other hand, the USGP spring wheat production estimate did not support the 90/90 accuracy goal because of a large negative bias in the estimate. As a result of the underestimation for spring wheat, the final LACIE total wheat production estimate only marginally supported the 90/90 accuracy goal.

The results of a sensitivity analysis to determine the relative contributions of area and yield errors to the production error are given in table XV. Unlike

the Phase II results, these results indicate that the total wheat production underestimation was primarily the result of yield underestimation. The large negative bias indicated for the spring wheat production estimate is primarily attributed to yield underestimation, although area was also significantly underestimated.

Area Error Source Analyses

The purpose of these analyses was to quantify the error components and then to determine the causative factors in the LACIE estimation process. The general approach in the U.S. Great Plains in LACIE Phase III was to compare the LACIE area estimates to various reference standards. The reference standards included the ground-observed data for a random sample of the LACIE operational segments, the historical SRS county-level area estimates, and the current SRS state-level area estimates.

Comparison of LACIE and SRS acreage estimates.—For the USGP-7 region (composed of the seven major winter wheat producing states in the USGP), the first LACIE winter wheat area estimate was significantly lower than the corresponding SRS area estimate (see fig. 7 and table XVI). The next LACIE estimate (presented in the May report) was not significantly different from the SRS estimate, as the LACIE estimate increased by more than 9 million acres and the SRS estimate decreased by 5.6 million acres. The increase in the LACIE estimate was a result of increased emergence and ground cover of the wheat, and the SRS decrease is attributed to the difference between planted area and area for harvest. During the remainder of Phase III, the SRS winter wheat area estimate for the USGP-7 region remained essentially unchanged. The LACIE estimate was significantly higher than the SRS figure

TABLE XV.—Relative Contributions of Area and Yield Errors to Bias in the Production Estimate

Region	Total RD. percent	RD. percent	
		LACIE acreage, SRS yield	LACIE yield, SRS acreage
Winter wheat	-3.4	+4.9	-8.9
Spring wheat	-25.7	-12.3	-15.5
Total wheat	-10.0	-.2	-10.9

in July and August (RD's of 8.7 and 7.2 percent, respectively). The September, October, and December LACIE USGP-7 area estimates tended to be high but were not significantly different from the corresponding SRS estimates.

At the state level, the primary winter wheat area estimation problem occurred in Colorado (final RD of 26.3 percent); Colorado was the only state in which the RD was consistently larger than in Phase II. Blind site investigations indicated sampling may be a problem in Colorado. Investigations were continuing at the time of this writing. Initial large underestimates in Oklahoma similar to those observed in

Phase II lessened as the season progressed.

The CV's for all states in the USGP-7 except South Dakota were smaller than those of Phase II, indicating an overall higher degree of reliability for the LACIE Phase III area estimates.

The LACIE Phase III spring wheat area estimates were first available in July. As shown in figure 7, the LACIE spring wheat area estimate for the USNGP was lower than the corresponding SRS estimate in every month. However, the RD's were generally much improved over those of Phase II. The final RD was -8.5 percent for Phase III as compared to -26.3 percent in Phase II.

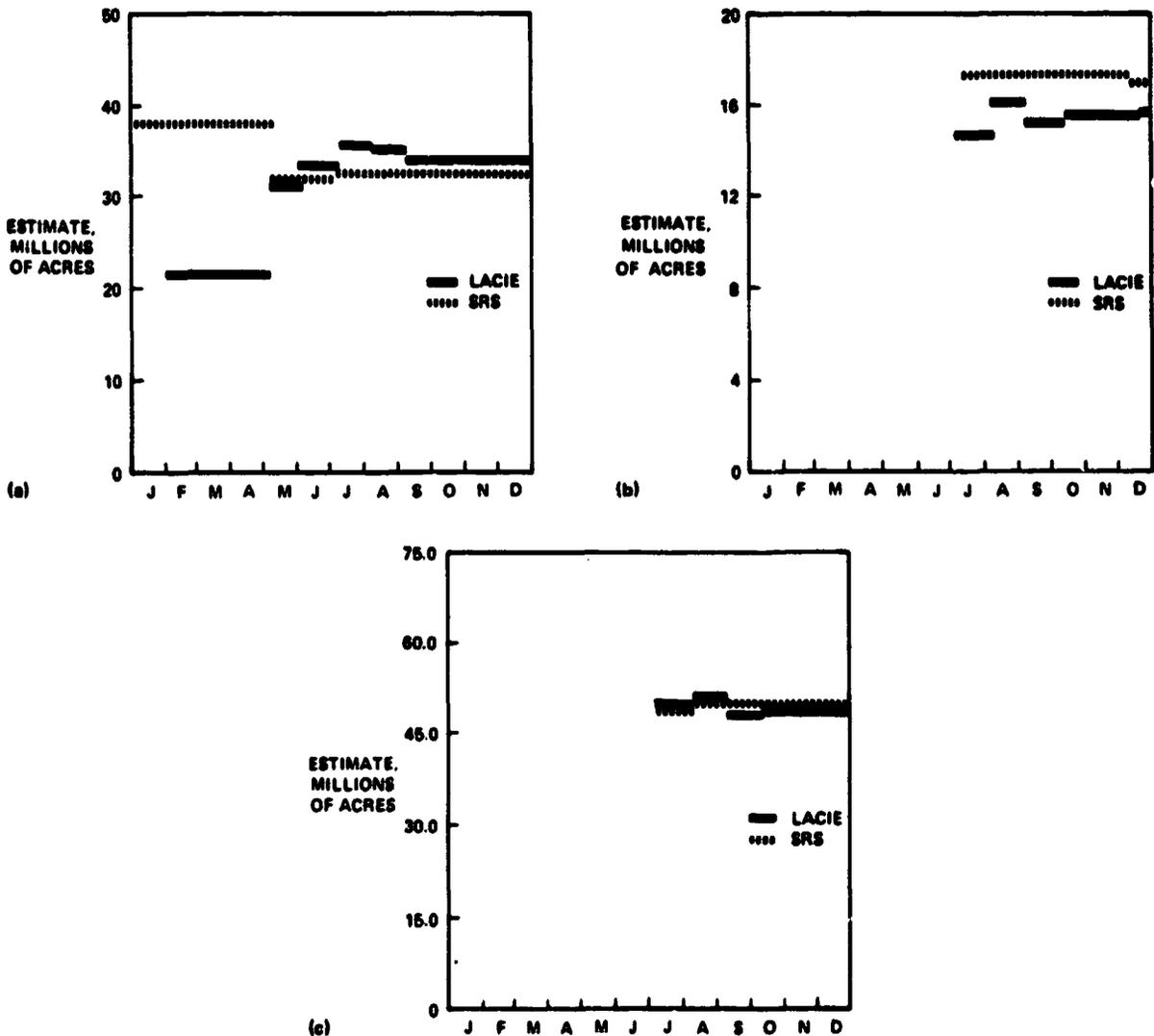


FIGURE 7.—LACIE and SRS acreage estimates (SRS estimates through April 22 are for seeded acres released on December 22, 1976). (a) USGP-7, winter wheat. (b) USNGP, spring wheat. (c) USGP, total wheat.

An underestimation problem in Minnesota, although much less severe than in Phase II, was the major problem for spring wheat area estimation in the USNGP during Phase III. Blind site investigations indicated that the major labeling error source in Minnesota was boundary pixels. Boundary pixels cause spectral and spatial confusion between small-

grains fields and non-small-grains fields.

The CV's of the LACIE spring wheat area estimates for Phase III were generally smaller than those of Phase II. The CV's of the Minnesota estimates showed the greatest reduction from Phase II levels, although they were among the largest for the USNGP states in all Phase III reports. For the

TABLE XVI.—Comparison of LACIE and SRS Area Estimates

(a) February^a

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	LACIE CV, percent		RD. percent		Test statistic
				1977	1976	1977	1976	
<i>Winter wheat</i>								
Colorado	25/31	2 140	1 997	21.0	26	-37.2	30.0	
Kansas	82/121	13 200	6 888	13.9	12	-91.6	-63.5	
Nebraska	41/56	3 300	3 067	14.9	18	-7.6	24.4	
Oklahoma	35/46	7 800	3 206	9.6	24	-143.3	-90.0	
Texas	25/35	6 150	3 365	16.7	25	-82.8	-98.7	
USSGP	208/289	33 190	18 523	7.1	9	-79.2	-46.0	^c -11.15
Montana	30/58	3 050	2 127	21.1	NA ^b	-43.4	NA	
South Dakota	6/21	1 160	800	60.6	NA	-45.0	NA	
MW states	36/79	4 210	2 927	22.4	NA	-43.8	NA	
USGP-7	244/368	37 400	21 450	6.8	NA	-74.4	NA	^c -10.94

(b) May

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	LACIE CV, percent		RD. percent		Test statistic
				1977	1976	1977	1976	
<i>Winter wheat</i>								
Colorado	22/31	2 290	3 600	14.2	24	36.4	32.3	
Kansas	98/121	12 000	10 439	6.2	6	-15.0	-15.0	
Nebraska	38/56	3 050	3 278	11.4	13	7.0	19.2	
Oklahoma	39/46	6 500	4 832	10.0	16	-34.5	-48.8	
Texas	30/35	4 400	4 196	14.2	14	-4.9	18.9	
USSGP	227/289	28 240	26 345	4.5	6	-7.2	-3.2	-1.06
Montana	28/58	2 800	3 369	18.8	NA	16.9	NA	
South Dakota	3/21	750	1 107	43.1	NA	32.2	NA	
MW states	31/79	3 550	4 476	17.7	NA	20.7	NA	
USGP-7	258/368	31 790	30 821	4.6	NA	-3.1	NA	-68

^aSRS production through April released on December 22, 1976

^bData not available

^cThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level

TABLE XVI.—Continued

(c) June

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	LACIE CV, percent		RD, percent		Test statistic
				1977	1976	1977	1976	
<i>Winter wheat</i>								
Colorado	22/31	2 360	3 608	13.6	23	34.6	36.6	
Kansas	104/121	12 000	11 055	5.8	6	-8.5	-2.0	
Nebraska	40/56	3 050	3 839	9.5	12	20.6	28.1	
Oklahoma	4/46	6 500	5 228	9.0	14	-24.3	-39.8	
Texas	30/35	4 400	4 462	12.5	15	1.4	14.4	
USSGP	236/289	28 310	28 192	4.1	5	-4	3.9	-0.10
Montana	29/58	2 800	3 704	17.8	193	24.4	518.9	
South Dakota	7/21	680	1 401	25.0	43	51.5	10.3	
MW states	36/79	3 480	5 105	14.6	65	31.8	-146.5	
USGP-7	272/368	31 790	33 297	4.1	6	4.5	-4.9	1.10

(d) July

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	LACIE CV, percent		RD, percent		Test statistic
				1977	1976	1977	1976	
<i>Winter wheat</i>								
Colorado	21/31	2 360	3 268	13.4	25	27.8	23.3	
Kansas	96/121	12 300	12 919	4.5	6	4.8	-2.6	
Nebraska	29/56	3 050	3 844	11.6	11	20.7	27.4	
Oklahoma	35/46	6 500	5 755	7.1	15	-12.9	-56.5	
Texas	24/35	4 600	5 011	11.6	15	8.2	-8.9	
USSGP	205/289	28 810	30 797	3.6	5	6.5	-4.5	1.81
Montana	27/58	2 800	2 626	9.8	52	-6.6	-189.3	
South Dakota	9/21	680	1 943	40.3	23	65.0	29.8	
MW states	36/79	3 480	4 569	12.1	25	23.8	-60.7	
USGP-7	241/368	32 290	35 366	3.9	5	8.7	-9.4	2.23

¹The LACIE estimate is significantly different from the SRS estimate at the 10-percent level.

USNGP, the Phase III CV was, on the average, about 40 percent smaller than that of Phase II.

The LACIE total wheat area estimates for the USGP region (available from July onward) were not significantly different from the corresponding SRS estimates in any reporting period of Phase III. In fact, the RD between the two estimates stayed between -1.1 percent and 2.5 percent over the entire season. Also, the CV of the USGP total wheat area

estimate was considerably smaller than in Phase II. The final Phase III CV was 2.4 percent, compared to a final CV of 4 percent in Phase II.

Studies based on ground-observed proportions.—In Phase III, near-harvest ground observations were obtained and analyzed for 92 winter wheat segments and 53 spring wheat segments. This section contains the results of studies based on this data set.

Bias due to classification (weighted analysis): This

TABLE XVI.—Continued

(d) July

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	LACIE CV, percent		RD, percent		Test statistic
				1977	1976	1977	1976	
<i>Spring wheat</i>								
Minnesota	22/47	3 202	2 420	12.2	NA	-32.3	NA	
North Dakota	13/103	9 500	9 071	10.7	NA	-4.7	NA	
SW states	35/150	12 702	11 491	8.9	NA	-10.5	NA	
Montana	5/48	2 185	1 895	37.6	NA	-15.3	NA	
South Dakota	5/37	2 332	1 269	40.4	NA	-83.8	NA	
MW states	10/85	4 517	3 164	27.7	NA	-42.8	NA	
USNGP	45/235	17 219	14 655	9.2	NA	-17.5	NA	^c -1.90
<i>Total wheat</i>								
Montana	30/73	4 985	4 521	9.9	NA	-10.3	NA	
South Dakota	13/45	3 012	3 212	17.9	NA	6.2	NA	
MW states	43/118	7 997	7 733	23.3	NA	-3.4	NA	
USNGP	78/268	20 699	19 224	16.1	NA	-7.7	NA	
USGP	283/557	49 509	50 021	3.4	NA	1.0	NA	0.29

^cThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level.

section presents the results of the weighted analysis of the aggregated acreage estimates to determine the bias that was due to classification. A weighted average of the differences between the at-harvest wheat proportion estimates and the ground-observed wheat proportions is obtained, where the weights are those used in the LACIE aggregation process. Table XVII presents the results of the weighted analysis. The results indicate the presence of a negative bias in the LACIE at-harvest area estimation process due to winter and spring wheat proportion estimation errors at the segment level.

Bias due to classification (unweighted analysis): This section presents the results of three segment-level wheat proportion estimation error investigations based on comparisons of LACIE wheat proportion estimates with corresponding ground-observed wheat proportions. The term "unweighted" is used to indicate that the analyses do not involve the expansion factors, or weights, from the aggregation logic.

1. Winter wheat proportion estimation error. Blind site results for the investigation of winter wheat proportion estimation errors for the USGP-7 region are shown in figure 8 and table XVIII. The LACIE proportion estimates used are from the Phase III CAS Annual Report, December 22, 1977. Figure 8 shows plots of the proportion estimation error ($\hat{X} - X$) versus X for the February, July, and final CAS reports, where \hat{X} is the LACIE harvested wheat proportion estimate and X is the ground-observed harvested wheat proportion. Points lying above the horizontal line $\hat{X} - X = 0$ correspond to overestimates and points lying below the line correspond to underestimates of wheat proportions.

Table XVIII contains the results of the statistical analysis of the winter wheat blind site data. The following factors are listed: (1) the average wheat proportion estimate \bar{X} , (2) the average ground-observed wheat proportion \bar{X} , (3) the average difference $\bar{D} = \bar{X} - \bar{X}$, (4) the standard error of the average difference $S_{\bar{D}}$, and (5) the 90-percent confi-

TABLE XVI.—Continued

(e) August

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	LACIE CV, percent		RD, percent		Test statistic
				1977	1976	1977	1976	
<i>Winter wheat</i>								
Colorado	26/31	2 360	3 253	11.3	24	27.5	22.3	
Kansas	103/121	12 300	12 579	4.8	5	2.2	-1.5	
Nebraska	31/56	3 050	3 558	10.2	11	14.3	26.6	
Oklahoma	37/46	6 500	5 963	6.7	15	-9.0	-46.3	
Texas	28/35	4 700	4 600	12.8	16	-2.2	-9.0	
USSGP	225/289	28 910	29 953	3.6	5	3.5	-3.2	0.97
Montana	39/58	2 800	3 355	7.9	35	16.5	-58.0	
South Dakota	12/21	680	1 594	38.1	23	57.3	29.8	
MW states	51/79	3 480	4 949	13.4	22	29.7	-19.7	
USGP-7	276/368	32 390	34 902	3.6	5	7.2	-5.0	62.00
<i>Spring wheat</i>								
Minnesota	30/47	3 202	2 553	13.0	40	-25.4	-119.8	
North Dakota	39/103	9 530	9 220	5.7	14	-3.4	-41.4	
SW states	69/150	12 732	11 773	5.3	13	-8.1	-55.2	
Montana	23/48	2 185	1 942	18.0	28	-12.5	-105.4	
South Dakota	24/37	2 332	2 309	13.4	12	-1.0	5.5	
MW states	47/85	4 517	4 251	11.0	12	-6.3	-32.4	
USNGP	116/235	17 249	16 024	4.8	10	-7.6	-49.5	-1.58
<i>Total wheat</i>								
Montana	52/73	4 985	5 296	6.4	19	5.9	-75.6	
South Dakota	30/45	3 012	3 904	8.6	13	22.8	15.4	
MW states	82/118	7 997	9 200	12.8	11	13.1	-26.0	
USNGP	151/268	20 729	20 973	9.2	9	1.2	-43.4	
USGP	376/557	49 639	50 926	2.6	5	2.5	-18.7	0.96

⁵The LACIE estimate is significantly different from the SRS estimate at the 10-percent level

dence limits for the population average difference μ_D . The formulas for calculating these factors are given in the paper by Houston et al.

To infer whether the population average difference for a particular state or region is significantly different from zero, one may simply check whether the corresponding 90-percent confidence interval contains zero. If it does, the population

average difference is not significantly different from zero; that is, there is insufficient evidence to conclude that there is a bias due to proportion estimation error. If the confidence interval does not contain zero, the hypothesis of no bias is rejected. The test is performed at the 10-percent level of significance.

The plot for February winter wheat shows that, early in the 1977 season, there was a tendency for the

TABLE XVI.—Continued

(f) September

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	LACIE CV, percent		RD, percent		Test statistic
				1977	1976	1977	1976	
<i>Winter wheat</i>								
Colorado	25/31	2 360	3 059	10.3	24	22.9	18.6	
Kansas	107/121	12 300	12 468	4.5	5	1.3	-1.0	
Nebraska	40/56	3 050	3 130	9.2	11	2.6	11.7	
Oklahoma	38/46	6 500	6 083	7.2	14	-6.9	-47.9	
Texas	28/35	4 700	4 613	12.7	16	-1.9	-8.2	
USSGP	238/289	28 910	29 353	3.5	5	1.5	-6.2	0.43
Montana	39/58	2 800	3 628	6.9	29	22.8	-43.6	
South Dakota	13/21	680	989	26.5	23	31.2	28.4	
MW states	52/79	3 480	4 617	7.8	20	24.6	-14.2	
USGP-7	290/368	32 390	33 969	3.2	5	4.6	-7.2	1.44
<i>Spring wheat</i>								
Minnesota	33/47	3 202	2 474	11.6	27	-29.4	-50.0	
North Dakota	62/103	9 530	8 523	5.0	5	-11.8	-19.6	
SW states	95/150	12 732	10 997	4.6	7	-15.8	-25.9	
Montana	30/48	2 185	2 187	12.2	23	.1	-79.3	
South Dakota	26/37	2 332	1 958	13.1	13	-19.1	2.1	
MW states	56/85	4 517	4 145	9.0	12	-9.0	-28.9	
USNGP	151/235	17 249	15 142	4.2	6	-13.9	-26.6	^c -3.31
<i>Total wheat</i>								
Montana	53/73	4 985	5 815	6.0	14	14.3	-57.2	
South Dakota	33/45	3 012	2 947	11.0	12	-2.2	12.9	
MW states	86/118	7 997	8 762	13.4	9	8.7	-21.4	
USNGP	181/268	20 729	19 759	8.7	6	-4.9	-24.3	
USGP	419/557	49 639	49 111	2.5	4	-1.1	-13.9	-0.44

^cThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level.

proportion of wheat in the segments to be underestimated by a greater margin for segments with larger proportions of wheat. This trend became less pronounced as the season progressed, and it appears to be insignificant in the July and final plots for winter wheat.

The results in table XVIII indicate the presence of a negative bias in LACIE winter wheat proportion

estimates for the USGP-7 region for each month shown. This indicates that, for these blind sites, the proportion of winter wheat for the USGP-7 region was underestimated in each reporting period. However, the wheat proportion estimation error decreased in magnitude each month, starting with May and ending in August. From August through the final reporting month, there was a slight increase

TABLE XVI.—Continued

(R) October

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	LACIE CV, percent		RD, percent		Test statistic
				1977	1976	1977	1976	
<i>Winter wheat</i>								
Colorado	24/31	2 360	3 395	9.9	24	30.5	18.6	
Kansas	108/121	12 300	12 669	4.2	5	2.9	-1.0	
Nebraska	39/56	3 050	3 375	9.6	11	9.6	11.7	
Oklahoma	41/46	6 500	5 658	7.7	14	-14.9	-47.9	
Texas	29/35	4 700	4 476	13.7	16	-5.0	-8.2	
USSGP	241/289	28 910	29 573	3.5	5	2.2	-6.2	0.63
Montana	43/58	2 800	3 314	7.8	28	15.5	-41.7	
South Dakota	14/21	680	883	25.7	23	23.0	28.4	
MW states	57/79	3 480	4 197	8.2	19	17.1	-13.3	
USGP-7	298/368	32 390	33 771	3.2	5	4.1	-7.1	1.28
<i>Spring wheat</i>								
Minnesota	37/47	3 202	2 289	9.9	30	-39.9	-74.1	
North Dakota	70/103	9 530	9 173	4.4	5	-3.9	-18.5	
SW states	107/150	12 732	11 462	4.0	7	-11.1	-28.8	
Montana	33/48	2 185	2 150	10.3	24	-1.6	-55.7	
South Dakota	32/37	2 332	1 909	11.6	13	-22.2	1.4	
MW states	65/85	4 517	4 059	7.7	12	-11.3	-22.4	
USNGP	172/235	17 249	15 522	3.6	6	-11.1	-27.3	^c -3.08
<i>Total wheat</i>								
Montana	58/73	4 985	5 464	5.5	12	8.8	-47.5	
South Dakota	38/45	3 012	2 793	9.9	12	-7.8	12.5	
MW states	96/118	7 997	8 257	12.2	8	3.1	-17.8	
USNGP	203/268	20 729	19 719	7.7	5	-5.1	-24.7	
USGP	444/557	49 639	49 293	2.4	4	-7	-14.1	-0.29

^cThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level

each month in the magnitude of the wheat proportion estimation error for the USGP-7 region. Inspection of figure 8 for the final estimates indicates that two outliers were the main cause of the increase.

Although the average winter wheat proportion estimation errors for the individual states in the USGP-7 tended to be negative, they decreased in magnitude as the season progressed. The number of

states with a population average difference that was not significantly different from zero at the 10-percent level increased from two in February to six in October. In the February and the final report, the average proportion estimation error for Oklahoma was nearly twice as large as the average for the other states in the USGP-7 region. The proportion estimation error for Oklahoma from May through October

TABLE XVI.—Concluded

(h) Final

Region	n/M	SRS, thousands of acres	LACIE, thousands of acres	LACIE CV, percent		RD, percent		Test statistic
				1977	1976	1977	1976	
<i>Winter wheat</i>								
Colorado	24/31	2 550	3 459	9.8	24	26.3	18.6	
Kansas	06/121	12 100	12 494	4.0	5	3.2	-1.6	
Nebraska	39/56	2 950	3 433	9.2	11	14.1	13.2	
Oklahoma	42/46	6 500	5 675	7.6	14	-14.5	-47.9	
Texas	29/35	4 700	4 476	13.7	16	-5.0	-8.2	
USSGP	240/289	28 800	29 537	3.4	5	2.5	-6.3	0.74
Montana	43/58	2 800	3 371	7.9	28	16.9	-48.1	
South Dakota	15/21	680	912	25.0	23	25.4	33.2	
MW states	58/79	3 480	4 283	8.2	19	18.7	-14.7	
USGP-7	298/368	32 280	33 820	3.2	5	4.6	-7.3	1.4
<i>Spring wheat</i>								
Minnesota	38/47	3 222	2 344	9.5	30	-37.5	-77.1	
North Dakota	73/103	9 150	9 183	4.4	5	.4	-16.9	
SW states	111/150	12 372	11 527	4.0	7	-7.3	-27.9	
Montana	32/48	2 260	2 174	10.2	22	-4.0	-54.0	
South Dakota	35/37	2 336	1 936	9.6	13	-20.7	2.8	
MW states	67/85	4 596	4 110	7.0	12	-11.8	-21.1	
USNGP	178/235	16 968	15 638	3.5	6	-8.5	-26.3	2.43
<i>Total wheat</i>								
Montana	57/73	5 060	5 545	5.4	12	8.7	-50.6	
South Dakota	41/45	3 016	2 848	9.1	12	-5.9	15.3	
MW states	98/118	8 076	8 393	11.7	8	3.8	-17.9	
USNGP	209/268	20 448	19 921	7.6	5	-2.6	-24.2	
USGP	449/557	49 248	49 458	2.4	4	.4	-13.9	0.17

^cThe LACIE estimate is significantly different from the SRS estimate at the 10-percent level

does not appear to be significantly different from the other states' estimates. The reason for this is that the two outliers referred to previously were in Oklahoma. One was acquired for the October analysis (note the increase in \bar{D} and $S_{\bar{D}}$ from September to October for Oklahoma in table XVIII) and the second was acquired for the final analysis (note the further increase in \bar{D} and $S_{\bar{D}}$ from October to final

for Oklahoma).

Figure 9 displays plots of proportion estimation error versus ground-observed proportion for each state in the USGP-7 winter wheat region, using the final LACIE proportion estimates. The two outliers are again apparent in the plot for Oklahoma. Investigation of these two blind sites indicated that there was no Landsat acquisition during the tillering-

TABLE XVII.—Estimates of LACIE Acreage Estimation Bias Due to Classification

Region	n/N ^a	LACIE area estimate \bar{A} , thousands of acres	Bias B, thousands of acres	Standard deviation of \bar{B}	Relative bias, percent	CV, percent	Test statistic
<i>Winter wheat</i>							
Colorado	11/24	3 459	-567	340	-16.4	9.8	
Kansas	24/106	12 494	-1161	476	-9.3	3.8	
Nebraska	16/39	3 433	-218	227	-6.4	6.6	
Oklahoma	15/42	5 675	-831	442	-14.6	7.8	
Texas	9/29	4 476	-141	708	-3.2	15.8	
USSGP	75/240	29 537	-3049	1104	-10.3	3.7	b-2.8
Montana	14/43	3 371	157	222	+4.7	6.6	
South Dakota	3/15	912	-451	491	-49.5	53.8	
USGP-7	92/298	33 820	-3213	1181	-9.5	3.5	b-2.7
<i>Spring wheat</i>							
Minnesota	11/38	2 344	-770	356	-32.8	15.2	
Montana	9/32	2 174	-780	425	-35.9	19.5	
North Dakota	21/73	9 183	-1442	535	-15.7	5.8	
South Dakota	12/35	1 936	-672	499	-34.7	25.8	
USNGP	53/178	15 638	-3653	916	-23.4	5.9	b-4.0
<i>Total wheat</i>							
USGP	145/449	49 498	-6440	1441	-13.0	2.9	b-4.5

^aThe n is the number of blind sites in the region, the N is the number of acquired segments in the region
^bIndicates classification bias is significantly different from zero

to-heading stage of wheat and, as a result, the analyst mislabeled most of the wheat pixels as non-small-grains. Excluding these two outliers yields an average proportion estimation error of -0.8 with a standard error of 1.4 for the remaining 13 blind sites, and the negative bias is no longer indicated.

Two other states with seemingly large standard errors of the average differences for the final estimates are Texas and South Dakota. The large standard error is expected for South Dakota, as only three blind sites are available. However, there are nine blind sites in Texas, and inspection of the plot for Texas reveals one outlier that is an extreme overestimate. Omitting this outlier yields an average difference of -4.5 with a standard error of 1.6, indicating a negative bias in the Texas winter wheat proportion estimates. Investigation of this site indicated an acquisition pattern similar to that of the two Oklahoma outliers. In this case, though, missing a key acquisition

led to overestimation rather than underestimation. This indicates that when a key acquisition is missing, a proportion estimate should not be made since positive identification of pixel labels is very difficult.

2. Spring wheat proportion estimation error. Figure 10 and table XIX contain spring wheat proportion estimation error results that are analogous to the winter wheat results contained in the preceding section.

The downward trend that was evident in the February plot of winter wheat proportion estimation error versus the ground-observed proportion of winter wheat is also seen in the July spring wheat plot. This means that the problem of underestimating the proportion of wheat early in the season in segments with larger proportions of wheat exists for spring wheat as well as for winter wheat. There was a gradual improvement in the LACIE estimates of the proportion of spring wheat (in the segments with

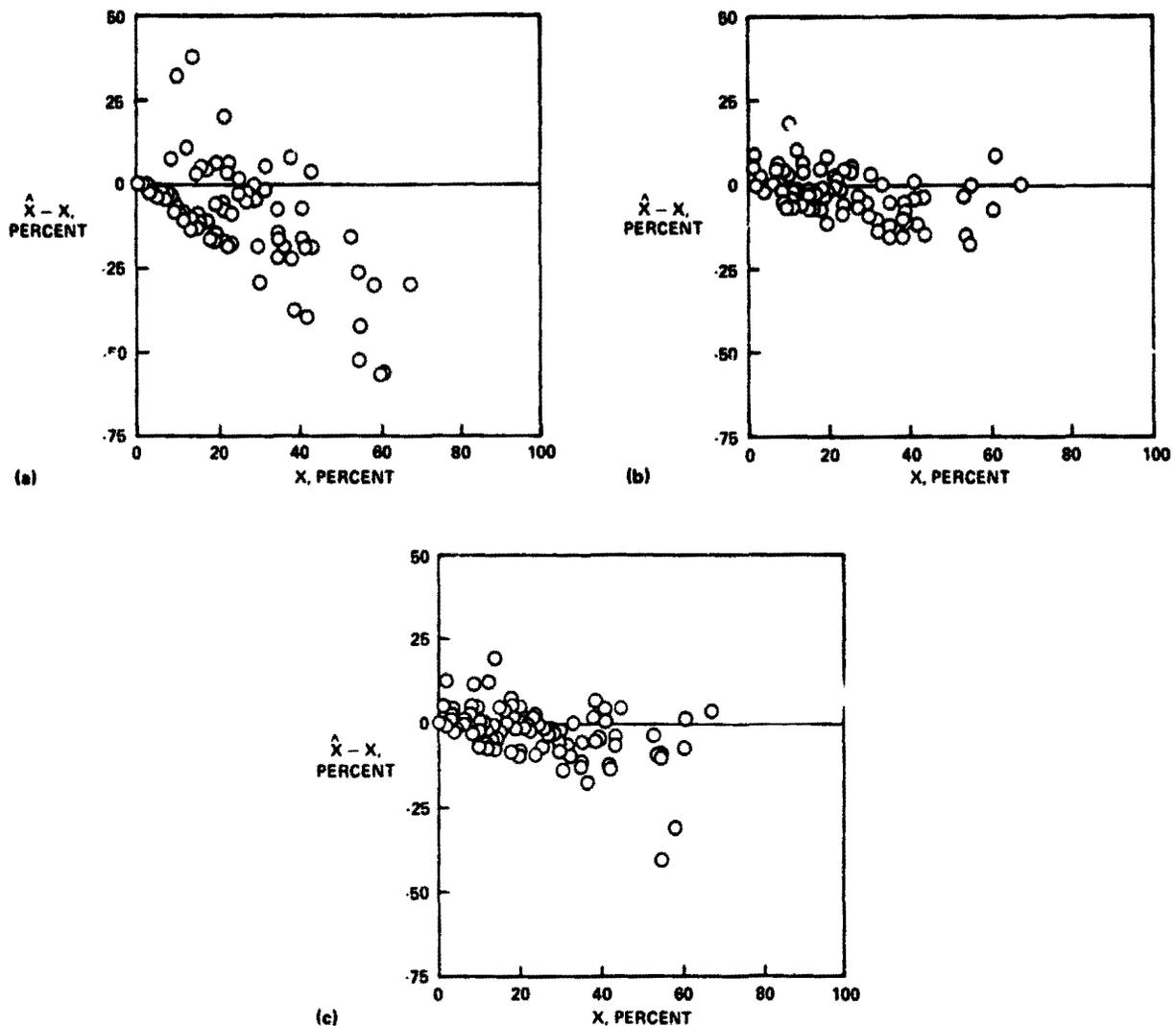


FIGURE 8.—Plots of proportion estimation errors versus ground-observed proportions for winter wheat blind sites. (a) February. (b) July. (c) Final.

large proportions of spring wheat) as the season progressed, but the trend is still present in the final spring wheat plot.

The average wheat proportion error for spring wheat had a tendency to be negative. The average spring wheat proportion estimation error for the USNGP region was negative for each month and, except for July, the population average differences were significantly different from zero at the 10-percent level (see table XIX). This sequence of negative average wheat proportion estimation errors for the USNGP region increased in magnitude from July through September and decreased slightly in the October and final reports. From August through the

final report, the average proportion estimation errors for Montana and South Dakota were not significantly different from zero at the 10-percent level. In July, South Dakota had an average wheat proportion estimation error that was significantly different from zero at the 10-percent level. There were no data for Montana in July.

Figure 11 displays the plots of proportion estimation error versus ground-observed proportion for each state in the USNGP spring wheat region. There are no obvious outliers for any of the states. In each state, though, the tendency to underestimate the larger proportions is apparent.

3. Relative contribution of the classification and

TABLE XVIII.—Winter Wheat Blind Site Results

Region	n/M	\bar{X}	\bar{Y}	\bar{D}	SD	90-percent confidence limits for μ_D
<i>February</i>						
Colorado	10/31	12.9	22.3	-9.5	1.8	(-12.7, -6.3) ^a
Kansas	19/121	14.9	30.2	-15.3	3.9	(-22.0, -8.6) ^a
Nebraska	16/56	20.8	17.7	3.1	3.0	(-2.2, 8.3)
Oklahoma	14/46	17.0	36.8	-19.9	4.2	(-27.3, -12.5) ^a
Texas	9/35	15.3	25.6	-10.3	3.4	(-16.6, -4.0) ^a
Montana	7/58	8.8	14.7	-6.0	1.9	(-9.7, -2.2) ^a
South Dakota	2/21	7.9	11.3	-3.4	2.6	(-19.9, 13.2)
USGP-7	77/368	15.6	25.3	-9.8	1.7	(-12.6, -7.1) ^a
<i>May</i>						
Colorado	10/31	15.4	22.3	-6.8	2.2	(-10.9, -2.7) ^a
Kansas	23/121	22.1	30.6	-8.5	2.6	(-12.9, -4.1) ^a
Nebraska	16/56	13.9	17.1	-3.2	1.8	(-6.4, -0.1) ^a
Oklahoma	15/46	25.3	34.3	-9.0	3.4	(-15.0, -3.1) ^a
Texas	10/35	19.4	23.4	-4.0	2.5	(-8.6, 0.6)
Montana	5/58	12.6	17.2	-4.6	2.9	(-10.7, 1.6)
South Dakota	2/21	6.2	11.3	-5.1	4.3	(-32.1, 21.9)
USGP-7	81/368	18.9	25.4	-6.5	1.1	(-8.4, -4.6) ^a
<i>June</i>						
Colorado	10/31	16.2	22.3	-6.1	2.4	(-10.4, -1.8) ^a
Kansas	25/121	22.3	29.0	-6.7	2.4	(-10.8, -2.6) ^a
Nebraska	17/56	18.0	16.7	1.3	1.6	(-1.5, 4.1)
Oklahoma	15/46	26.1	34.3	-8.2	3.2	(-13.9, -2.6) ^a
Texas	10/35	20.2	23.4	-3.2	2.5	(-7.8, 1.4)
Montana	5/58	14.5	17.2	-2.6	2.8	(-8.6, 3.4)
South Dakota	3/21	5.7	7.8	-2.1	3.7	(-12.8, 8.6)
USGP-7	85/368	20.1	24.6	-4.5	1.1	(-6.3, -2.6) ^a
<i>July</i>						
Colorado	7/31	18.4	19.3	-0.9	1.4	(-3.6, 1.8)
Kansas	21/121	26.5	29.1	-2.7	1.4	(-5.2, -0.2) ^a
Nebraska	14/56	16.4	17.0	-0.6	1.9	(-4.0, 2.8)
Oklahoma	13/46	31.4	35.2	-3.8	1.7	(-6.8, -0.8) ^a
Texas	8/35	21.4	25.5	-4.1	2.5	(-8.8, 0.5)
Montana	8/58	11.3	15.3	-4.0	1.6	(-7.1, -0.9) ^a
South Dakota	3/21	7.2	7.6	-0.4	1.0	(-3.3, 2.6)
USGP-7	74/368	21.7	24.2	-2.5	0.7	(-3.7, -1.3) ^a

^aSignificantly different from zero at the 10 percent level

ratio errors to the ratioed spring wheat proportion estimation errors. As in LACIE Phase II, the LACIE Phase III wheat proportion estimates for a segment were obtained by multiplying the small-grains proportion estimate obtained in CAMS by a wheat-to-small-grains ratio. The wheat-to-small-grains ratios

used in Phase III were obtained from econometric models at the CRD level in the USNGP. The purpose of this section is to provide the results of a sensitivity analysis used to determine the contributions of classification and ratio errors to proportion estimation errors at harvest.

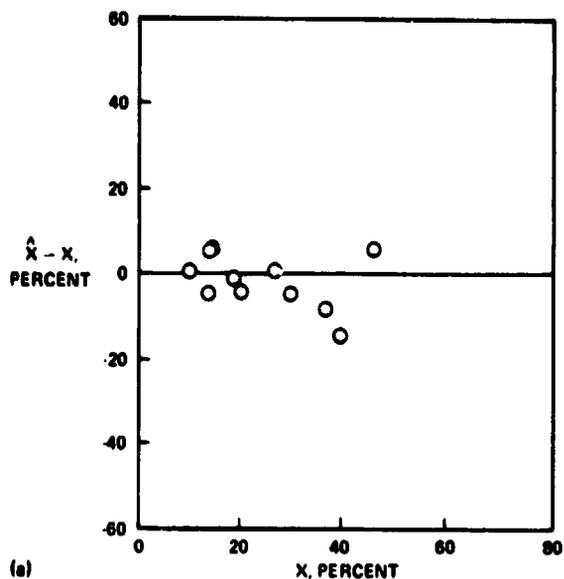
TABLE XVIII.—Concluded

Region	n/M	\bar{X}	\bar{Y}	D	SD	90-percent confidence limits for μ_D
<i>August</i>						
Colorado	10/31	19.9	21.3	-1.4	1.8	(-4.7, 1.8)
Kansas	22/121	28.0	30.6	-2.6	1.3	(-4.8, -0.4) ^a
Nebraska	14/56	15.5	16.2	-.8	1.3	(-3.0, 1.5)
Oklahoma	13/46	35.3	36.9	-1.6	1.6	(-4.4, 1.2)
Texas	9/35	22.4	25.2	-2.8	2.8	(-8.1, 2.5)
Montana	12/58	11.8	14.0	-2.2	1.2	(-4.5, 0.0)
South Dakota	3/21	7.0	7.6	-.6	.8	(-3.1, 1.9)
USGP-7	83/368	22.4	24.2	-1.8	.6	(-2.8, -0.8) ^a
<i>September</i>						
Colorado	11/31	17.3	20.2	-2.9	1.6	(-5.8, -0.1) ^a
Kansas	23/121	28.0	30.5	-2.5	1.1	(-4.4, -0.5) ^a
Nebraska	17/56	13.7	16.0	-2.3	1.1	(-4.2, -0.4) ^a
Oklahoma	13/46	36.3	36.9	-.5	1.6	(-3.4, 2.4)
Texas	9/35	22.6	25.2	-2.6	2.9	(-8.0, 2.8)
Montana	12/58	12.8	13.6	-.7	1.0	(-2.6, 1.1)
South Dakota	3/21	5.0	7.6	-2.6	2.6	(-10.1, 4.9)
USGP-7	88/368	21.7	23.7	-1.9	.6	(-2.9, -0.9) ^a
<i>October</i>						
Colorado	11/31	17.8	20.2	-2.4	1.7	(-5.4, 0.7)
Kansas	24/121	27.0	29.4	-2.4	1.1	(-4.3, -0.6) ^a
Nebraska	16/56	15.7	18.0	-2.2	1.3	(-4.5, 0.1)
Oklahoma	14/46	34.8	38.2	-3.4	2.8	(-8.3, 1.6)
Texas	9/35	22.7	25.2	-2.5	2.9	(-7.9, 2.9)
Montana	14/58	13.6	13.4	.1	1.0	(-1.7, 1.9)
South Dakota	3/21	5.0	7.6	-2.6	2.6	(-10.1, 4.9)
USGP-7	91/368	21.9	24.0	-2.1	.7	(-3.3, -1.0) ^a
<i>Final</i>						
Colorado	11/31	17.8	19.8	-2.0	1.5	(-4.7, 0.7)
Kansas	24/121	26.5	29.3	-2.8	1.1	(-4.7, -0.9) ^a
Nebraska	16/56	16.5	18.0	-1.5	1.1	(-3.5, 0.5)
Oklahoma	15/46	34.4	40.2	-5.8	3.1	(-11.3, -0.4) ^a
Texas	9/35	22.7	24.3	-1.6	2.9	(-6.9, 3.7)
Montana	14/58	13.7	13.4	.3	1.1	(-1.7, 2.2)
South Dakota	3/21	5.0	7.6	-2.6	2.6	(-10.1, 5.0)
USGP-7	92/368	22.0	24.4	-2.4	.7	(-3.6, -1.2) ^a

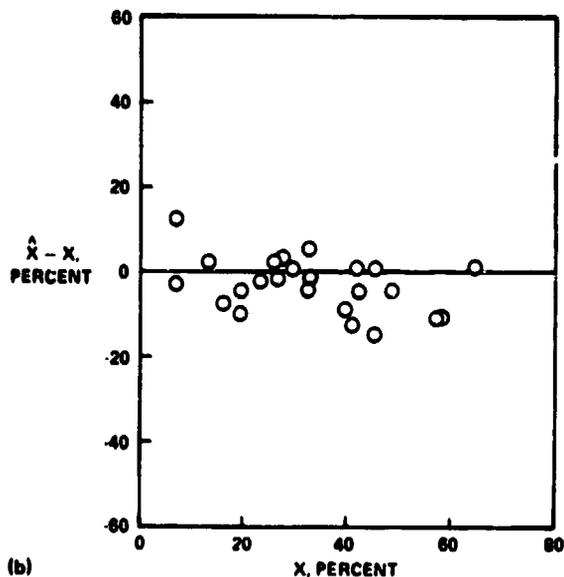
^aSignificantly different from zero at the 10-percent level

This analysis was made for 33 blind sites in the spring wheat region of Minnesota and North Dakota, using the at-harvest LACIE proportion estimates. These are the two states for which a negative bias was indicated for the final spring wheat proportion estimates (see table XIX). The results of this analysis

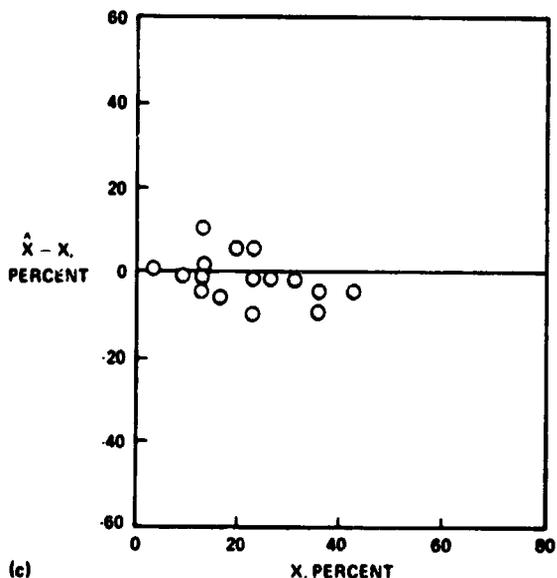
are presented in table XX. The line labeled "No ratioing error" was obtained using the LACIE small-grains proportion estimate with the corresponding ground-observed ratio of spring wheat to small grains. Likewise, the line labeled "No classification error" was obtained using the LACIE estimate of the



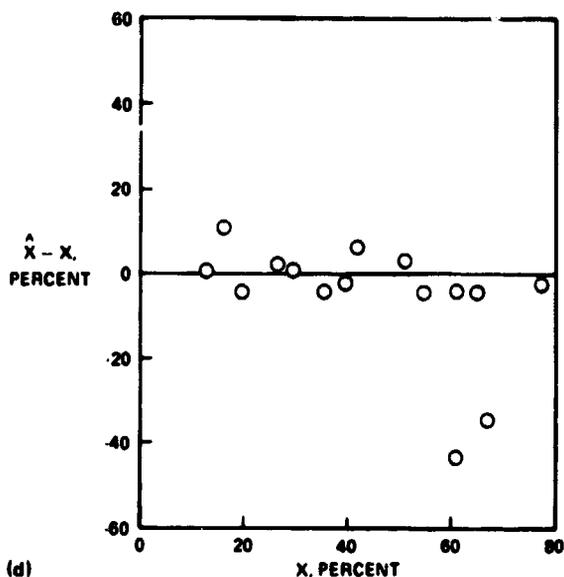
(a)



(b)



(c)



(d)

FIGURE 9.—Plots of at-harvest proportion estimation errors versus ground-observed proportions for winter wheat blind sites by state. (a) Colorado. (b) Kansas. (c) Nebraska. (d) Oklahoma. (e) Texas. (f) Montana. (g) South Dakota.

spring-wheat-to-small-grains ratio (forecast from a CRD-level econometric model) and the corresponding ground-observed proportion of small grains.

The results indicate that classification error was the major contributor to both the bias and the mean-squared error of the total spring wheat proportion estimation error in North Dakota and Minnesota in Phase III. Comparison of these results with those of Phase II (table IX) indicates a significant reduction in total mean-squared error from Phase II to Phase

III. The primary reason for this reduction was an increase in small-grains classification precision in Phase III as indicated by the reduction in mean-squared error from 78.6 in Phase II to 33.4 in Phase III for proportion estimates with no ratiing error (i.e., classification errors only). This is at least partially due to the use of Procedure 1 in Phase III. The mean-squared error for proportion estimates with no classification error was about the same for Phases II and III. However, the use of econometric models in

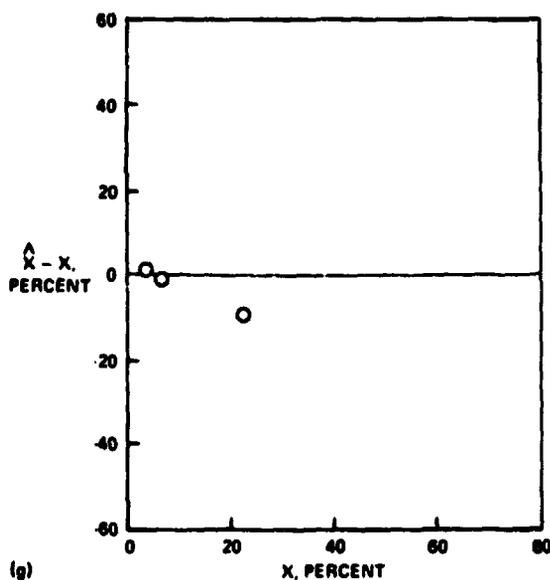
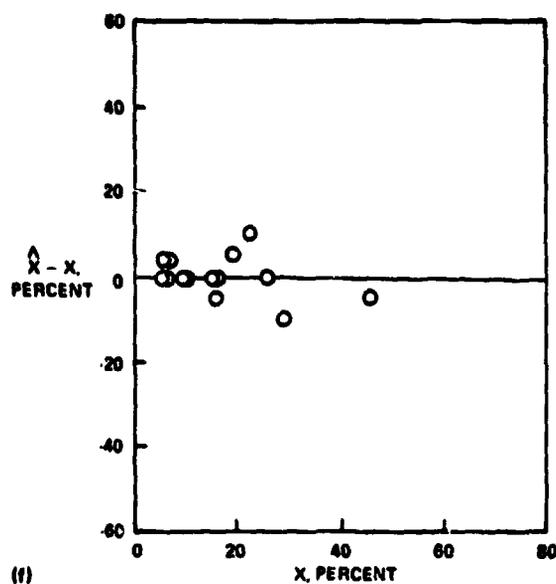
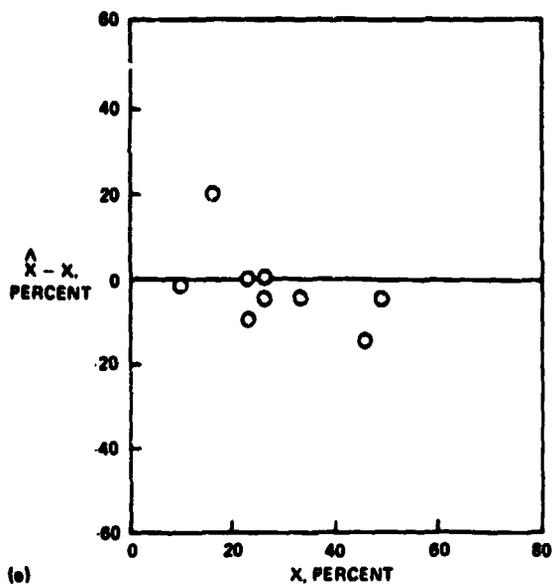


FIGURE 9.—(concluded).

Phase III for forecasting the spring-wheat-to-small-grains ratios apparently reduced the bias considerably from that obtained in Phase II using historical ratios. In Phase III, there was an estimated bias of -1.3 percent as compared to an estimated bias of -3.1 percent in Phase II for proportion estimates with no classification error (i.e., ratioing errors only).

Contribution of sampling and classification errors to the variability of area estimates: This study was

performed for the purpose of measuring the contributions of classification and sampling errors to the within-stratum area variability and estimating the classification and sampling error contributions to the CV's of the regional area estimates. Since the proportion estimates used in this section are for ratioed wheat (winter or spring), the classification error referred to herein is actually compounded with the ratio error.

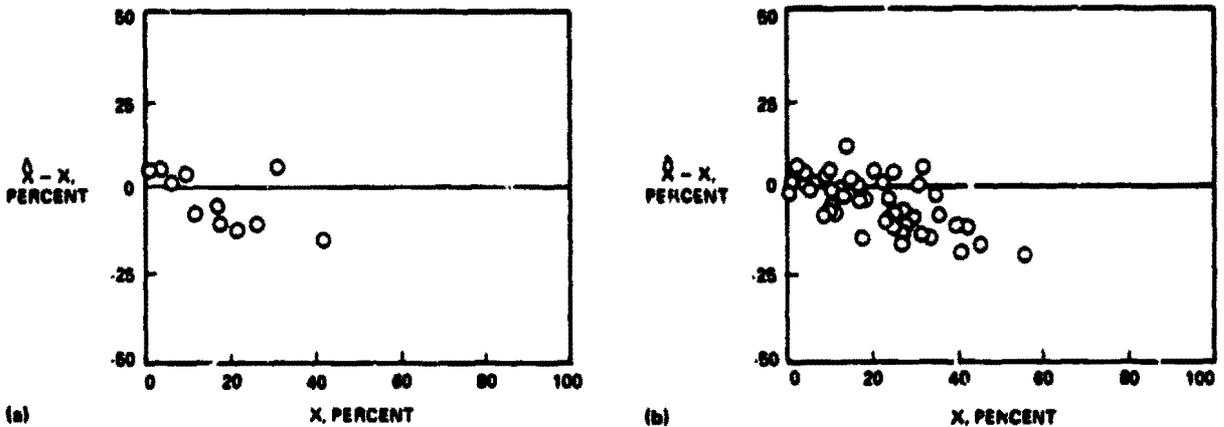


FIGURE 10.—Plots of proportion estimation errors versus ground-truth proportions for spring wheat blind sites. (a) July. (b) Final.

To estimate the within-stratum area variances resulting from classification and sampling errors, the following three basic regression models are constructed.

1. True segment proportion versus historical stratum proportion
2. LACIE segment proportion versus ground-truth segment proportion
3. LACIE segment proportion versus historical stratum proportion

These regression models are used to obtain, respectively, an estimate of sampling's contribution to the variance, an estimate of classification's contribution to the variance, and an estimate of a linear combination of the classification and sampling variances. The maximum likelihood estimation technique, assuming normality and that the regression models in 1, 2, and 3 are applicable, is then used to obtain maximum likelihood estimates of the contributions of sampling and classification to the area variance. A detailed description of this approach is presented in the paper by Houston et al. Table XXI gives the results of this analysis made of 449 Phase III operational segments, 146 of which were blind sites.

These results show that the sampling CV is larger than the classification CV for winter, spring, and total wheat area estimates. The implication is that sampling contributes slightly more to the area variance than does classification. Moreover, winter wheat has smaller CV's for both classification and sampling than does spring wheat; that is, there is less variability in the winter wheat area estimates than in the spring wheat area estimates for the USGP region. The sampling CV for the total wheat area estimate is

1.9 percent, which is well within the sampling accuracy goal of 2.3 percent.

Acreeage estimation bias due to nonsampled and non-responsive areas.—In order to investigate bias due to the ratio estimation procedure used to estimate the wheat area in nonsampled and nonresponsive areas in the United States, aggregations were performed in which the LACIE proportion estimate for each segment was replaced by the corresponding 1976 SRS county wheat proportion. Table XXII contains the results of this "mock aggregation" for all allocated segments and the comparisons with 1976 SRS estimates. The RD at the USGP level is -2.5 percent, indicating a possible small negative bias due to the Group II and Group III ratio estimation procedure used for those counties not allocated segments. This is larger than the observed RD of 0.8 percent obtained in a similar study of the Phase II sample segment allocation to the U.S. Great Plains (see table XI). The Phase II allocation was based on wheat production for an epoch year, whereas the Phase III allocation was based on small-grains production for an epoch year.

An investigation was undertaken to determine the allocation that would have resulted from using the epoch-year wheat production rather than the epoch-year small-grains production. It was found that 32 currently designated Group III counties should have been Group I or Group II counties and that 16 currently designated Group I counties and 43 currently designated Group II counties should have been Group III counties. The decision was made to redesignate the 16 Group I and 43 Group II counties as Group III counties. This caused the original alloca-

TABLE XIX.—Spring Wheat Blind Site Results

Region	n/M	\bar{X}	\bar{Y}	B	SD	90-percent confidence limits for μ_D
<i>July</i>						
Minnesota	6/47	9.1	11.1	-2.0	2.5	(-7.1, 3.0)
Montana	0/48	—	—	—	—	—
North Dakota	2/103	32.6	36.8	-4.2	10.3	(-69.2, 60.8)
South Dakota	3/37	11.2	15.8	-4.6	4.9	(-18.9, 9.8)
USNGP	11/235	13.9	17.1	-3.1	2.3	(-7.3, 1.0)
<i>August</i>						
Minnesota	10/47	17.3	22.6	-5.2	2.4	(-9.6, -0.9) ^a
Montana	4/48	4.2	11.7	-7.5	5.9	(-21.3, 6.3)
North Dakota	8/103	24.4	27.3	-2.8	3.4	(-9.4, 3.7)
South Dakota	9/37	9.8	11.3	-1.6	2.0	(-5.3, 2.1)
USNGP	31/235	15.3	19.1	-3.8	1.5	(-6.3, -1.3) ^a
<i>September</i>						
Minnesota	11/47	19.0	23.7	-4.7	2.3	(-8.8, -0.6) ^a
Montana	7/48	9.9	12.1	-2.2	2.4	(-6.8, 2.4)
North Dakota	17/103	20.9	25.7	-4.8	1.7	(-7.8, -1.8) ^a
South Dakota	9/37	8.4	11.3	-2.9	2.5	(-7.6, 1.8)
USNGP	44/235	16.1	20.1	-4.0	1.1	(-5.8, -2.2) ^a
<i>October</i>						
Minnesota	12/47	18.6	22.9	-4.3	2.2	(-8.2, -0.4) ^a
Montana	9/48	11.9	15.7	-3.8	2.3	(-8.1, 0.5)
North Dakota	20/103	21.0	25.1	-4.0	1.5	(-6.6, -1.5) ^a
South Dakota	9/37	7.9	9.4	-1.5	2.3	(-5.8, 2.8)
USNGP	50/235	16.4	20.1	-3.6	1.0	(-5.2, -2.0) ^a
<i>Final</i>						
Minnesota	12/47	18.5	22.9	-4.4	2.2	(-8.3, -0.5) ^a
Montana	9/48	12.0	15.2	-3.2	2.3	(-7.5, 1.0)
North Dakota	21/103	21.3	24.7	-3.4	1.4	(-5.8, -1.1) ^a
South Dakota	12/37	8.0	11.1	-3.1	1.6	(-6.4, 0.2)
Total	54/235	16.2	19.7	-3.5	.9	(-5.0, -2.1) ^a

^aSignificantly different from zero at the 10-percent level

tion to the United States of 601 segments to be reduced to 557 segments. The results in table XXII are for the 557 segments after redesignation. It was infeasible at the time to allocate more sample segments to the 32 Group III counties that should have been Group I or Group II counties. The use of the Group III estimator to estimate their wheat area accounts for at least part of the observed difference.

Table XXIII contains the results of aggregating the 1976 SRS county wheat proportions for each segment acquired and processed for each Phase III monthly estimate made except the final. The result for the final estimate is expected to be similar to that for the October estimate. The difference between the mock aggregation and the SRS estimate in this study is due to error in the Group II and Group III ratio

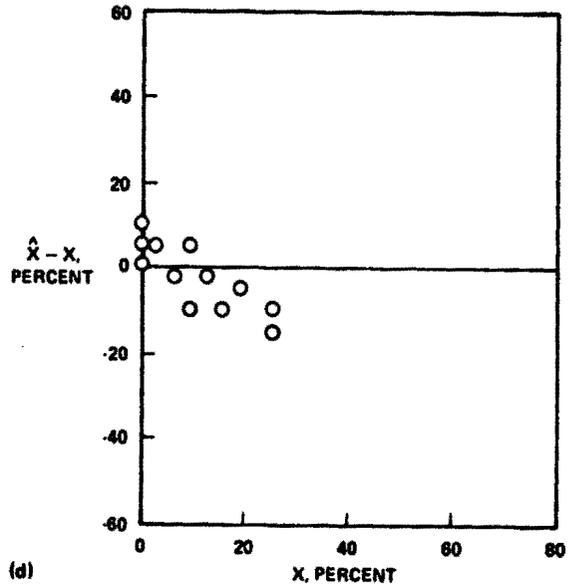
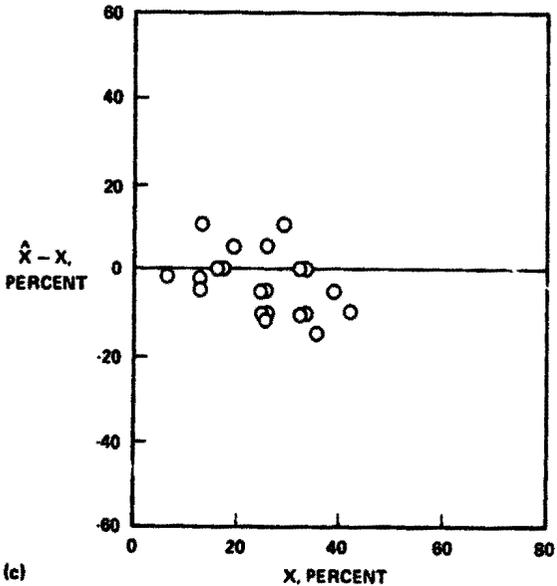
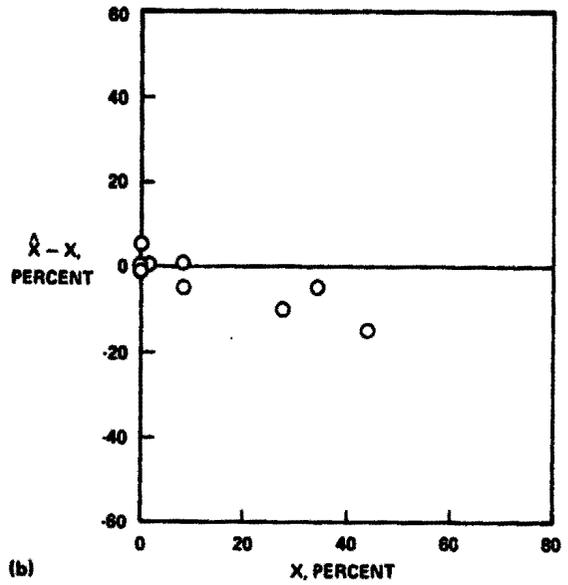
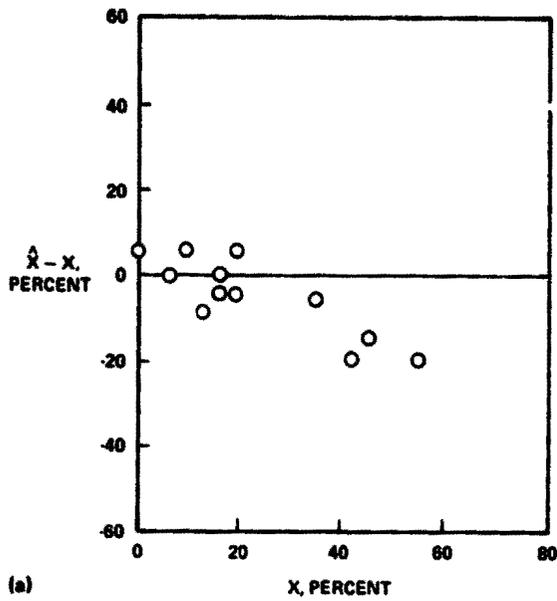


FIGURE 11.—Plots of at-harvest proportion estimation errors versus ground-observed proportions for spring wheat blind sites by state. (a) Minnesota. (b) Montana. (c) North Dakota. (d) South Dakota.

estimation procedure used for both those counties not allocated segments and those counties whose allocated segments were lost to nonresponse.

The results indicate that the error due to the ratio estimation of the nonsampled and nonresponsive areas for each month during Phase III is about the same as that due to nonsampled areas only. This in-

dicates that the error due to Group II and Group III ratio estimation of areas lost to nonresponse is negligible. However, the results do suggest the presence of a small negative bias in the ratio estimation technique applied to the nonsampled areas, particularly in the winter wheat region.

TABLE XX.—Relative Contribution of Classification and Ratio Errors to Final Phase III Spring Wheat Proportion Estimation Errors

Category	n/M	\bar{D} , percent	$S\bar{D}$, percent	Reduction in bias, percent	90-percent confidence limits for $\mu\bar{D}$	MSE	Reduction in MSE, percent
Final Phase III result	33/150	-3.8	1.1	—	(-5.7, -1.8)	68.3	—
No ratioing error	33/150	-2.5	.8	34.2	(-3.9, -1.2)	33.4	51.1
No classification error	33/150	-1.3	.8	65.8	(-2.7, -0.02)	26.9	60.6

TABLE XXI.—Contribution of Sampling and Classification Errors to Variability of Area Estimates

Crop	Within- stratum area variance	Variance component		Percentage error		Area CV, percent	Classification CV, percent	Sampling CV, percent
		Classi- fication	Sampling	Classi- fication	Sampling			
Winter wheat, USGP-7	104.1	41.6	62.5	40	60	3.2	2.0	2.5
Spring wheat, USNGP	65.6	26.2	39.4	40	60	3.5	2.3	2.8
Total wheat, USGP	100.4	39.6	60.8	40	60	2.4	1.5	1.9

TABLE XXII.—Acreage Estimation Bias Due to Nonsampled Areas

Region	M	1976 SRS, thousands of acres	Mock aggregation, thousands of acres	RD, percent
Winter wheat, USGP-7	368	31 500	30 478	-3.4
Spring wheat, USNGP	235	19 768	19 527	-1.2
Total wheat, USGP	557	51 268	50 005	-2.5

Special Studies

The method of estimating wheat proportions using LACIE Procedure 1 requires a shift from the labeling of fields to the labeling of individual grid intersection dots or picture elements (pixels). Previously, the analyst-interpreter (AI) could select the desired fields for labeling. Procedure 1 requires that the AI label pixels from a fixed list of randomly selected pixels taken from the 209 intersections of

the grid overlay of the imagery. The accuracy of Procedure 1 depends to a large extent upon the AI's ability to discern accurately which of these pixels are small grains. Using the software system described in the paper by Pitts et al. entitled "Accuracy Assessment System and Operation," the AI labels and the corresponding ground-observed labels can be compared to evaluate the dot labeling accuracy. The results of such studies are presented in this section (see reference 2 for more detail).

Analyst dot labeling accuracy.—The results presented in this section are from a comparison of ground-observed and analyst-designated labels of dots from 51 blind sites located in North Dakota, Minnesota, Montana, Colorado, and Oklahoma. These dots are type 2 dots, which are those dots used to perform the stratified area estimation part of Procedure 1. The accuracy of the segment-level proportion estimate is critically dependent on the labeling accuracy of these type 2 pixels.

Table XXIV presents the at-harvest total omission and commission error rates (as a percentage of total pixels labeled in a state) for each state as well as the omission and commission error rates for the three major error sources identified in Phase III. Omission error is the result of mislabeling small-

TABLE XXIII.—Acreage Estimation Bias Due to Nonsampled and Nonresponsive Areas

<i>Region</i>	<i>N/M</i>	<i>1976 SRS, thousands of acres</i>	<i>Mock aggregation, thousands of acres</i>	<i>R.D. percent</i>
<i>February</i>				
Winter wheat, USGP-7	244/368	31 500	30 408	-3.6
<i>May</i>				
Winter wheat, USGP-7	256/368	31 500	30 737	-2.5
<i>June</i>				
Winter wheat, USGP-7	272/368	31 500	30 556	-3.1
<i>July</i>				
Winter wheat, USGP-7	241/368	31 500	30 978	-1.7
<i>August</i>				
Winter wheat, USGP-7	276/368	31 500	30 678	-2.7
Spring wheat, USNGP	116/234	19 768	19 934	.8
Total wheat, USGP	376/557	51 268	50 612	-1.3
<i>September</i>				
Winter wheat, USGP-7	290/368	31 500	30 641	-2.8
Spring wheat, USNGP	151/234	19 768	19 523	-1.3
Total wheat, USGP	419/557	51 268	50 164	-2.2
<i>October</i>				
Winter wheat, USGP-7	298/368	31 500	30 475	-3.4
Spring wheat, USNGP	172/234	19 768	19 548	-1.1
Total wheat, USGP	444/557	51 268	50 023	-2.5

grains pixels as non-small-grains; commission error is the result of mislabeling non-small-grains pixels as small grains.

The results in table XXIV show that the omission error is consistently larger than the commission error by state and by error source. This occurrence typically leads to underestimation of the small-grains proportion in a segment and, in fact, is what was found in the blind site analyses of proportion estimation error described previously.

Abnormal signatures, boundaries, and inadequate acquisitions were found to be the three major sources of labeling error. "Abnormal signatures" refers to a signature (small-grains or non-small-grains) that, under the conditions believed by the AI to be occurring in the segment, is not the expected signature or does not follow the expected temporal sequence. "Boundaries" are made up of two types of pixels, border pixels and edge pixels. A border pixel is one which presents an interpretation problem because its signature is spectrally mixed; that is, it represents both a small-grains area and a non-small-grains area. An edge pixel is one for which the signature is spatially mixed; that is, on the acquisitions used by the AI for proportion estimation, the edge pixel moves at least once from a small-grains field to a non-small-grains field because of misregistration. "Inadequate acquisitions" refers to labeling errors that occur because the AI attempts to label a segment when key acquisitions are missing. The AI is usually guessing for many of the pixels in this case and probably should not pass an estimate. This particular error occurred in only one or two blind sites per state; however, when it occurred, both the labeling error and the proportion estimation error were large. For example, the 3.0-percent omission error due to inadequate acquisitions for the 11 blind sites in Oklahoma came from one segment. This particular segment accounts for one of the two extreme underestimates in Oklahoma referred to previously in the proportion estimation error analysis. The other outlier in Oklahoma was not included in this study but it had the same acquisition history.

Labeling errors in the "other" category include clerical errors and inconsistent labeling errors. Inconsistent labeling occurs when an AI has labeled several pixels correctly and then incorrectly labels one or two pixels following the same temporal sequence in the same segment.

Note that nonresolvable small-grains strip-fallow pixels were excluded from the study of Montana. These are pixels for which the MSS resolution is not

TABLE XXIV.—Phase III Label Error Causes

[Percentage of total pixels labeled]

Cause of error	State (number of blind sites)									
	North Dakota (18)		Minnesota (6)		Montana ^a (10)		Colorado (6)		Oklahoma (11)	
	OM ^b	COM ^c	OM	COM	OM	COM	OM	COM	OM	COM
Abnormal signatures	4.4	0.5	2.6	0.3	1.4	0.9	2.8	—	3.3	1.4
Boundaries	3.2	.7	4.0	1.1	1.0	.6	2.3	0.8	2.2	.8
Inadequate acquisitions	1.5	1.0	—	—	.5	—	—	—	3.0	—
Other	2.1	.8	2.5	1.2	1.9	.6	.9	—	1.4	3.3
Total errors	11.2	3.0	9.1	2.6	4.8	2.1	6.0	.8	9.9	5.5

^aNonresolvable small-grains strip-fallow pixels excluded.

^bOmission error rate.

^cCommission error rate.

fine enough to show the strips in the Landsat imagery. The signature is integrated for the whole field and hence cannot be called a boundary-type signature. Either label, small grains or non-small-grains, could be considered correct for these areas. Hence, because of the inability to characterize the error, they were omitted. In Montana, 10.3 percent of the labeled pixels fell in these nonresolvable strip-fallow areas. Of these, 54 percent were labeled non-small-grains and 46 percent were labeled small grains by the AI's. This is fairly good since one would expect 50 percent of these areas to be small grains.

About 11 percent of the pixels labeled in Montana fell in resolvable strip-fallow areas. The relatively low error rate for boundaries (1.0-percent omission, 0.6-percent commission) indicates that the analysts labeled quite accurately in these areas. Overall, the Montana small-grains signatures were found to be quite good. There were very few abnormal signatures and there was good separation of the small-grains and non-small-grains signatures. Recalling the proportion estimation error study, neither the winter wheat nor the spring wheat blind site analysis indicated a bias for the Montana proportion estimates.

Excluding the outlier for Oklahoma, the largest total labeling errors in the study were for Minnesota and North Dakota. These errors were primarily due to omission errors for abnormal signatures and boundaries. In the spring wheat proportion estimation error study, these were the only two states for which a negative bias was indicated. The large errors

of omission apparently caused this proportion estimation underage.

Figure 12 contains an example displaying the two largest sources of omission errors in Minnesota and North Dakota. The blind site is located in Grant County, Minnesota. The pixels identified as 1, 2, and 3 are examples of a border pixel, an edge pixel, and an abnormal signature, respectively. (The upper left corner of the grid intersection is designated as the exact location of the pixel.)

Pixel 1 lies on the border between a spring wheat field and a sunflower field. From the ground-truth map, it was determined that the pixel contained more spring wheat than sunflowers, but the analyst labeled the pixel as non-small-grains. The more accurate ground-truth determination is possible because the ground observations are made at a sub-pixel level, one-sixth the size of a pixel. The evaluator thought that the AI should have labeled the pixel as small grains because close inspection of the imagery revealed that, in the heading acquisition, the pixel was more red than green and, in the turning acquisition, it was more green than red.

Pixel 2 is a classic example of an edge pixel. In the heading acquisition, the pixel is on a road. In the turning acquisition, the pixel is in a spring wheat field. The turning acquisition was the base acquisition for this segment. When an AI works a segment, he selects one of the acquisitions to be the base acquisition. This means that the pixels are to be labeled as to their location in the base acquisition. The grid

intersections of the other acquisitions are registered to the base acquisition for labeling. In this example, pixel 2 was labeled as non-small-grains, but since the turning acquisition was the base acquisition, it should have been labeled as small grains. This may have been a clerical error.

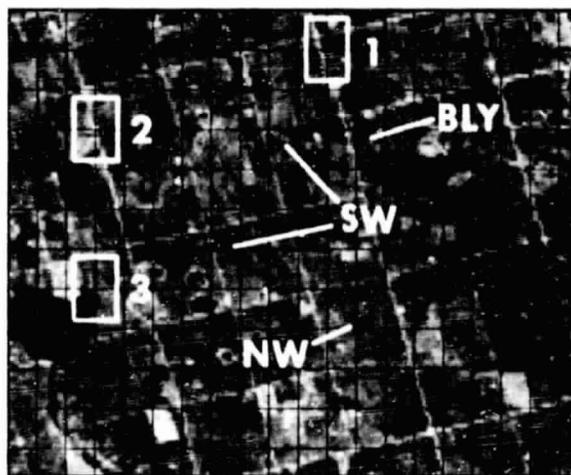
Pixel 3 provides an example of an abnormal signature. It is green in the heading acquisition and red in the turning acquisition. However, this pixel lies on the edge of a small body of water. The ground truth indicated that the wheat field came right up to the edge of the water. The AI labeled the pixel as non-small-grains. The evaluator thought that the AI believed the pixel to be grass growing on the edge of the water. The evaluator determined that the pixel was actually spring wheat, as indicated by the ground truth, but the development of the spring wheat in this pixel had been delayed because of excess moisture and was still in the heading stage although the majority of the wheat in the segment was in the turning stage.

Effects of AI, acquisition history, and bias correction on proportion estimation error.—The Image 100 processor and data from eight U.S. blind sites were used in an experiment wherein each site was analyzed by three AI's to give a "raw" and a "bias-corrected" esti-

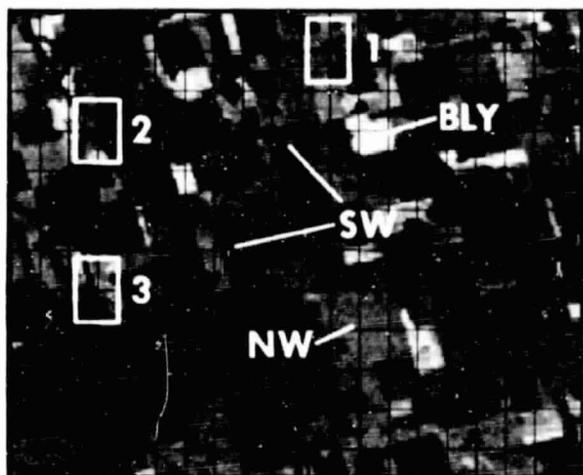
mate of the proportion of small grains in each segment. The segments were of two types; namely, those having acquisitions in all four biophases and those having only early-season acquisitions. The segments were selected at random from the blind sites for which detailed ground truth was available.

The objectives of the experiment were (1) to evaluate the performance of Procedure 1 in terms of absolute proportion estimation error and its repeatability with different AI's, (2) to make comparisons between "bias-corrected" and "raw" Procedure 1 estimates, and (3) to determine whether the performance was better when acquisitions from all biostages were used than when only the early-season acquisition was used.

The third objective could not be properly achieved because of the small number of segments used (four of each type). It was later estimated that to make effective comparisons of this type in a fully nested design, one would need about 10 times as many segments. The efficiency of the test could be improved if the *same* segments were analyzed first using only early-season acquisitions and then using all acquisitions; however, there would be potential biasing problems in such replication if the same AI analyzed the segment under both the early-season and the full-



HEADING — JUNE 23, 1977
SW — RED
NW — GREEN



TURNING — JULY 29, 1977
SW — GREEN
NW — RED

1. BORDER PIXEL — SPECTRAL CONFUSION OF SW AND SUNFLOWERS
2. EDGE PIXEL — SHIFTS FROM ROAD (HEADING) TO SW (TURNING)
3. ABNORMAL SIGNATURE — EXCESS WATER RETARDED SW DEVELOPMENT

FIGURE 12.—Phase III omission labeling error examples.

season conditions. If different AI's performed the analysis, the potentially large variability, as found in the experiment reported here, would further increase the number of segments required.

Table XXV shows the absolute proportion estimation error $|\hat{X} - X|$, where X is the ground-truth small-grains proportion and \hat{X} is the analyst's estimate of X , for the various treatment combinations. Averages are blocked off from the basic data; for example, the average absolute error for AI "B" on early-season segments was 11.6 for the raw estimate and 11.8 for the bias-corrected estimate. The average absolute error on all segments was 7.9 for raw estimates and 11.1 for bias-corrected estimates. The average absolute error for all three AI's was 12.8 for raw early-season estimates, 6.3 for raw full-season estimates, and 9.5 for all eight segments with raw estimates. The grand mean was 10.0.

The most obvious feature of table XXV is the large variability between AI's and between segments. If this variation is taken to be typical, then future experiments should be designed so that segments and AI's are "crossed" with treatments as much as possible; that is, each segment should be worked by each AI using each treatment.

Analysis of variance was used to test for the effects of AI's, time (i.e., early season versus all acquisitions), method (raw versus bias correction), and their interactions. The results led to the following conclusions.

1. The large disparity between data from various AI's was not consistent over segments; i.e., an AI would do better on one segment than on another.

2. There was no significant difference between methods; i.e., the use of bias correction just traded one random error for another of comparable magnitude.

3. Any test involving acquisition history was not significant.

As stated earlier, these tests had extremely low power because of insufficient numbers of segments to account for the large AI-to-AI and segment-to-segment variability.

Summary of Phase III

The Phase III results indicate that significant improvement has been realized over Phase I and Phase II results because of the LACIE area estimation technology improvements. The incorporation in Phase III of the ratios of wheat to small grains as forecast using the econometric models proved to be much better than using historical ratios. The new classification procedure, Procedure 1, apparently helped increase the precision of small-grains proportion estimates, particularly in the spring wheat area. The increased precision in classification, together with the achievement of the Phase III goal of a 2.3-percent sample error, resulted for the first time in a total

TABLE XXV.—Image 100—Procedure 1 Data

$|\hat{X} - X|$ (small grains)

Acquisition history	Segment	Raw				Bias correction				Overall average
		Analyst			Average	Analyst			Average	
		A	B	C		A	B	C		
Early season only	1642	16.5	10.8	2.0		18.9	8.7	16.8		
	1651	11.4	18.5	21.3		5.6	18.3	19.7		
	1660	9.7	14.6	30.3		8.0	11.9	19.5		
	1662	8.4	2.5	7.0		1.6	8.2	1.5		
Average		11.5	11.6	15.2	12.8	8.5	11.8	14.4	11.6	12.2
Full season	1603	0.8	1.4	0.9		1.4	1.4	2.0		
	1614	5.2	10.6	31.7		9.7	32.9	32.6		
	1637	1.3	.3	15.1		7.2	5.0	14.0		
	1656	1.7	4.7	2.4		2.7	2.5	2.5		
Average		2.2	4.3	12.5	6.3	5.3	10.5	12.8	9.5	7.9
Overall average		6.9	7.9	13.8	9.5	6.9	11.1	13.6	10.5	10.0

wheat area estimate for the United States for which the 90/90 hypothesis could not be rejected.

The expanded blind site program proved to be extremely useful for evaluating the area estimation technology in Phase III and is expected to be invaluable for future technology advancements. The major sources of labeling error—abnormal signatures, boundaries, and inadequate acquisitions—were identified through the use of the ground data acquired and processed in Phase III. As a result, classification procedures have already been modified to eliminate segment estimates based on poor acquisition histories. The abnormal signature and boundary problems are still under investigation. Potential solutions are being investigated at this writing.

CONCLUSIONS

The 3 years of area estimation results in the U.S. yardstick region support the U.S.S.R. and Canadian experience. In Phase II, the wheat area for Canada was grossly underestimated. Problems were apparently due to incorrect ratios of wheat to small grains and omission errors in small-grains classification. The omission errors are thought to be the result of boundary pixels and abnormal signatures. The boundary pixels are due to small fields (e.g., strip-fallow cropping practice) and nonhomogeneous fields. The abnormal signatures are like those experienced in the USNGP spring wheat area during LACIE.

In the U.S.S.R., the fields are very large (average about 500 hectares) and the cropping practices appear to be more uniform than in the United States. The large fields result in fewer boundaries, and the uniform cropping practices result in fewer abnormal signatures. This situation should result in better proportion estimates at the segment level, according to the Phase III blind site analyses, and hence in improved wheat area estimates for the U.S.S.R.

For regions such as Canada with a variety of cropping practices (e.g., irregular planting, grazing, and

stripping) that result in many boundaries and abnormal signatures and the prevalence of many confusion crops, it is believed that improvements in the area estimation technology are required. However, it is expected that significant advances in the Landsat scanner and in Landsat data processing will improve the area estimation capability for these regions.

REFERENCES

1. LACIE Phase I and Phase II Accuracy Assessment Final Report. LACIE-00450, NASA Johnson Space Center, Houston, Tex., Apr. 1978.
2. LACIE Phase III Unscheduled Interim Accuracy Assessment Report. LACIE-00473, NASA Johnson Space Center, Houston, Tex., July 1978.

ACKNOWLEDGMENTS

Acknowledgment is due R. S. Chhikara and A. H. Feiveson for their statistical expertise, which contributed to the methodology for various error source investigations. For their long hours digitizing the blind site ground-truth data, the efforts of the personnel of the Cartographic Laboratory are deeply appreciated; these people included J. Bowens, K. Bray, F. Colleen, Y. Dorsey, C. Guttrich, E. Mayer, F. Mayo, D. McKay, M. Rader, A. Sanchez, D. Taylor, R. Vela, and J. Villarreal. For the creation of the ground-truth data base and its manipulation, the authors are indebted to C. W. Ahlers, G. Badhwar, J. G. Carnes, J. E. Hartman, and W. R. Johnson. For their data analysis contributions, thanks are due F. Barbee, W. M. Belcher, R. E. Cheffin, M. Ferguson, and M. L. Mannen. The authors also appreciate the programing and aggregation logic expertise provided by C. J. Liszcz and M. A. Mendlowitz. The analyst labeling expertise provided by N. J. Clinton is gratefully acknowledged. Finally, the authors appreciate the editorial efforts of M. F. McKay.

Accuracy and Performance of LACIE Yield Estimates in Major Wheat Producing Regions of the World

D. E. Phinney,^a R. G. Stuff,^b A. G. Houston,^b E. M. Hsu,^a and M. H. Trenchard^a

INTRODUCTION

The LACIE wheat production for a specific region is calculated as the product of the total area of wheat harvested and the average yield per unit area in the region. Although Landsat data are used in making area estimates, their use in yield estimation either alone or in combination with conventional meteorological data is still in the developmental stage. The current LACIE yield model makes an independent estimate based on weather variables obtained from ground reports. These weather observations are provided by the national weather service in each country and are transmitted internationally by the synoptic network of the World Meteorological Organization (WMO).

The LACIE yield models for the United States (ref. 1) represent the "first generation" of yield models designed for large-area application. These models and their development are described in detail in the paper by Strommen et al. entitled "Development of LACIE CCEA-I Weather/Wheat Yield Models." However, a brief review will be given of some salient points which materially affect the performance of these models.

The models were derived, using multiple linear regression, from historical time series of selected weather variables and yield. The resulting models are area specific with one model for each region. The candidate weather variables were functions of monthly mean air temperature and monthly total precipitation. The final selection of parameters used in a given model was based partly on statistical considerations and partly on agronomic interpretation of the critical weather factors for the modeled area (see

the paper by Strommen et al.) The methodology used to develop the models is shown schematically in figure 1.

Figures 2 and 3 show the modeled areas for the United States and the U.S.S.R. Spring and winter wheat were modeled separately, resulting in 14 models for 12 areas in the United States and 44 models for 33 areas in the U.S.S.R.

Data from 1932 to 1976 were used to construct a data base for U.S. model development and evaluation. The regional yields were aggregated from U.S. Department of Agriculture (USDA) Statistical Reporting Service (SRS)¹ crop reporting district (CRD) data. The weather data used for model development consisted of averages of temperature and precipitation for climatological divisions. Weighted regional averages, based on 1973 acreage distributions for the U.S. models, were calculated for each weather variable.

In foreign areas, the length of available historical records varied greatly. Yields were modeled for political subdivisions which correspond to the official reporting scheme for the area. The handling of the meteorological data also varied from country to country (see the plenary paper by Strommen et al. entitled "The Impact of LACIE on a National Meteorological Capability").

A trend component of the year-to-year variation in yield has long been recognized and has been attributed to technological factors such as improved varieties, increased fertilization, and changing cultural practices. The LACIE yield models use a linear trend based on year which is fit piecewise as shown for the North Dakota spring wheat model in figure 4.

^aLockheed Electronics Company, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

¹The Statistical Reporting Service (SRS) has since become part of the Economics, Statistics, and Cooperatives Service (ESCS).

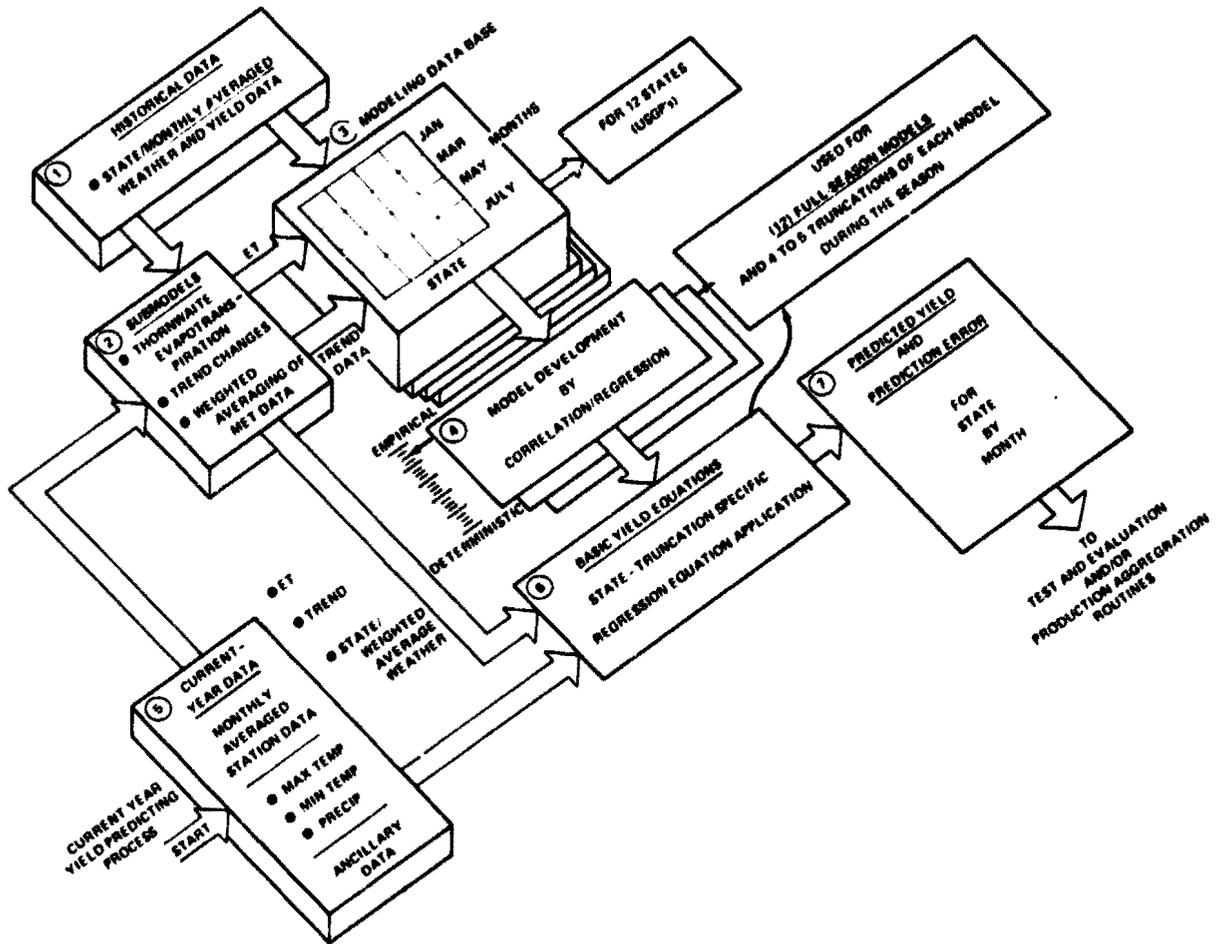


FIGURE 1.—Center for Climatic and Environmental Assessment (CCEA) first-generation wheat yield model.

For use in predicting yields through the crop season, the LACIE yield models were used with coefficients which were estimated using only those weather variables available up to the time of the estimate.

TECHNICAL APPROACH

Test of 90/90 Criterion

The goal of LACIE was to predict wheat production at harvest, over large areas, to within 10 percent of the true value 90 percent of the time. This was referred to as the 90/90 criterion.

An evaluation of the yield models in the context

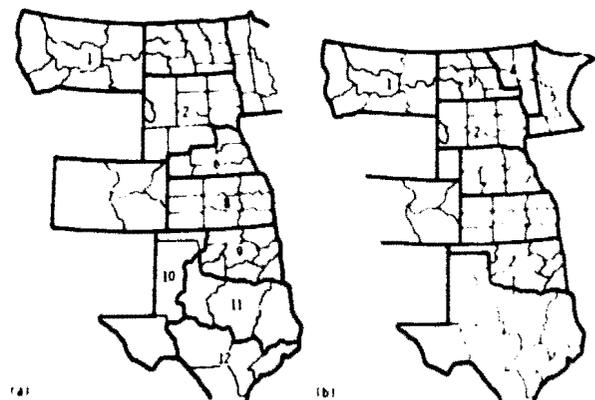


FIGURE 2.—Map showing boundaries for U.S. Great Plains winter and spring wheat models as given by the CCEA. (a) Winter wheat model boundaries. (b) Spring wheat model boundaries.

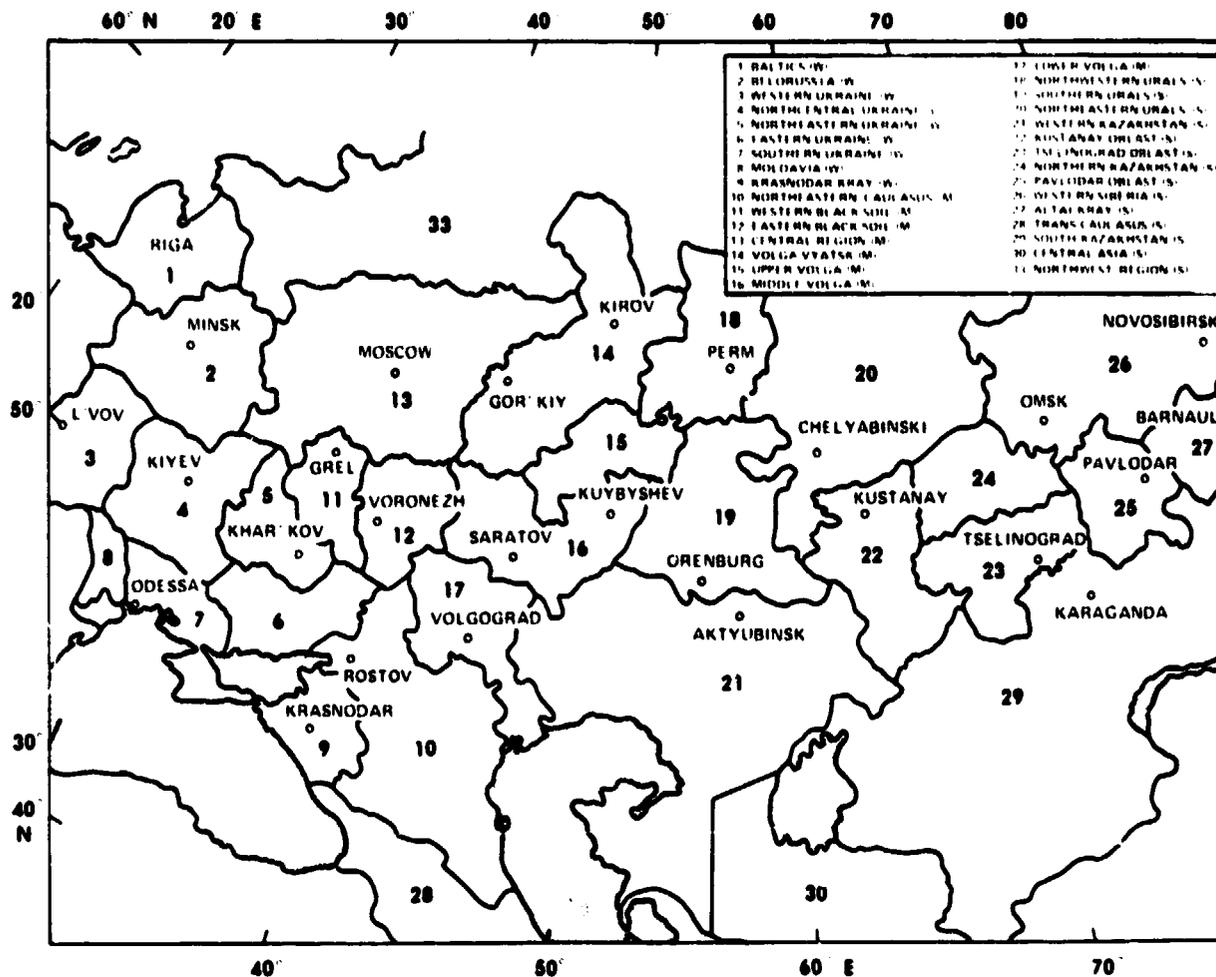


FIGURE 3.—Map showing U.S.S.R. crop regions covered by the spring and winter wheat regression models (regions 31 and 32 are not shown). W is winter wheat; S is spring wheat; and M is mixed winter and spring wheat.

of the 90/90 criterion can be carried out independently of acreage estimation errors by using the reference standard acreage estimates together with the yield model prediction. As shown below, the 90/90 criterion for a production estimate with both acreage and yield errors is equivalent to a 90/93 criterion for a production estimate with only yield errors. A 90/93 criterion specifies that the production estimate, with no acreage errors, be within 7 percent of the true production with a probability of at least 90 percent.

The 90/90 criterion for production may be written as

$$\text{Probability} \left(\left| \frac{\hat{P}}{P} - P \right| < 0.1P \right) > 0.9$$

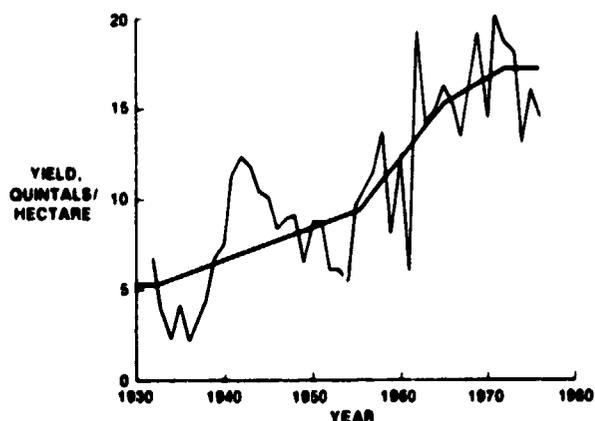


FIGURE 4.—Graph of yields and modeled trend for North Dakota spring wheat (1932-76).

where \hat{P} is the LACIE estimate of wheat production and P is the true wheat production.

Assuming that the yield and acreage estimates are independent and are unbiased estimates of the true yield and acreage and that the production estimates are normally distributed, it has been shown (ref. 2) that the probability statement can be written in terms of the variance $\sigma_{\hat{P}}$ of the production estimate.

$$\text{Pr} \left(\frac{|\hat{P} - P|}{\sigma_{\hat{P}}} < \frac{0.1P}{\sigma_{\hat{P}}} \right) > 0.9$$

It can be shown that the variance of the production estimate may be estimated from the variance $\sigma_{\hat{P}}$ of a production estimate made using the actual acreage and the yield estimate.

$$\sigma_{\hat{P}} = \sqrt{2} \sigma_{\hat{P}_0}$$

It then follows that

$$\text{Pr} \left(\frac{|\hat{P}_0 - P|}{P} < 0.0707 \right) > 0.9$$

That is, the 90/90 criterion for a production estimate with both acreage and yield errors is equivalent to a 90/93 criterion for a production estimate with only yield errors.

A random variable Z can be defined as follows:

$$Z = \frac{|\hat{P}_0 - P|}{P} - 0.0707P$$

An indicator function $\Psi(Z)$ is defined such that

$$\Psi(Z) = 1 \text{ if } Z < 0$$

$$\Psi(Z) = 0 \text{ if } Z > 0$$

The test statistic $\Sigma(Z)$ can be evaluated from binomial tables for a significance level α of 0.07 and a

number of samples equal to 10. For example, for a 10-year test, if $\Sigma\Psi(Z) \geq 8$, the hypothesis that the 90/90 criterion has been supported is not rejected.

The same approach may be used to test a single model by modifying the random variable.

$$Z = \frac{|\hat{P}_* - P|}{\sigma_{\hat{P}_*}} - \frac{0.0707P}{\sqrt{\delta_R}}$$

where δ_R is the fraction of the total production contained in the modeled region. Accepting the 90/90 criterion test for a single model is equivalent to saying that the model performance is acceptable providing that all other models which comprise the total area are similar and that the model errors are not correlated.

Ten-Year Tests

An evaluation of the yield models was made by obtaining 10 years of yield predictions and corresponding prediction error estimates, where the predictions were obtained using a "bootstrap" procedure. In this procedure, predictions were made for a particular year, then the resulting yield and weather for that year were used to recalculate the model coefficients to predict for the next year. The yield predictions were compared with the reference standard over the 10 years at the level at which the model is developed. The reference standard in the United States was the USDA SRS estimates. In foreign countries, official country estimates were used as the reference if available; otherwise, USDA Foreign Agricultural Service (FAS) estimates were employed. These comparisons were used to determine whether biases were indicated and where improvements were needed.

For those areas with relatively short historical records, the predictive model was developed using data from all years except the year to be estimated. Permutation of the test year resulted in a set of quasi-independent estimates. This was called the "jack-knife" test.

The squared prediction errors estimated for each of the 10 predictions are compared with the observed mean squared error over the test set for each yield model zone. This comparison indicates any shortcomings in the estimator for the prediction error.

PHILOSOPHIC APPROACH

Models were developed and tested during LACIE for Argentina, Australia, Brazil, Canada, India, the U.S.S.R., and the United States. All models were subjected to historical tests. The models for Canada, the U.S.S.R., and the United States were also used in LACIE operations. The objective of the model evaluations was to determine whether it was possible to provide yield estimates with sufficient accuracy and reliability to improve predictive abilities, particularly in foreign applications.

However, in foreign areas, it was difficult to isolate the error sources such that an orderly development of yield technology could be carried out. The specification of trend, the density of input meteorological data, and the reliability of even official yield statistics were all confounded in most foreign situations. Thus, the U.S. Great Plains (USGP) was selected as a "yardstick" region. By focusing the available LACIE yield evaluation resources on the USGP, it was possible to understand more fully the strengths and weaknesses of operational yield models.

PHASE I (1975 CROP YEAR) EVALUATION

Yield models were developed during Phase I for regions covering the nine states of the USGP. These models were applied at the CRD level as well as at the regional level. During evaluation of these models, it was found that there were no significant differences between regional predictions obtained directly from the regional models and applying the regional model at the CRD level with individual CRD weather (ref. 2). Exploratory studies indicated that models derived and applied at the individual CRD had the potential for improved model performance due to greater homogeneity of weather and yield data. However, limitations in resources prevented taking advantage of the potential inherent in modeling smaller areas.

The model estimates were aggregated to the USGP level and evaluated using the 90/90 test criterion. The results shown in table I indicate that the Phase I models did not support project accuracy goals. The individual models were also evaluated for their performance over the 10-year period. Results are shown

in table II. All models, except North Dakota and Kansas, support the 90/90 objective when projected to the USGP level.

Besides being evaluated against the 90/90 criterion, models were examined to determine their

TABLE I.—Ten-Year Bootstrap Test for the U.S. Phase I Yield Models Aggregated to the USGP by Year With 90/90 Criterion Test

Year	SRS, bu/acre	LACIE, bu/acre	Error ^a	Z	$\Psi(Z)$ ^b
1965	24.0	23.5	0.6	-39841205	1
1966	22.5	24.9	-2.4	29256745	0
1967	21.5	20.5	1.0	-20336126	1
1968	26.0	24.8	1.2	-26350536	1
1969	28.1	30.5	-2.3	11994856	0
1970	28.2	28.2	-1	-59542908	1
1971	30.8	28.1	2.7	16716330	0
1972	29.3	29.1	-1	-64278092	1
1973	30.8	36.7	-5.9	146531492	0
1974	23.8	28.4	-4.6	135728192	0

^aMean error = -1.0 t/acre, RMSE = 2.77 bu/acre
^b $\Psi(Z) = 5$, reject 90/90

TABLE II.—Ten-Year Bootstrap Test (1965-74) for the U.S. Phase I Yield Models With 90/90 Criterion Test by Model Region

Model	Crop	Mean error, bu/acre	RMSE, bu/acre	Support 90/90
Montana	SW	0.4	2.40	Yes
North Dakota	SW	-2.3	4.55	No
Red River	SW	-2.6	4.69	Yes
South Dakota	SW	0	2.24	Yes
Montana	WW	.7	3.71	Yes
Badlands	WW	1.9	5.30	Yes
Nebraska	WW	2.2	4.42	Yes
Colorado	WW	.3	4.33	Yes
Kansas	WW	-2.1	7.19	No
Oklahoma	WW	-1.7	3.41	Yes
Panhandle	WW	-4	3.29	Yes
Texas Low Plains	WW	1.4	3.08	Yes
Total	SW	-2.0	3.51	
Total	WW	-5	3.51	
Total	W	-1.0	2.77	

ability to respond adequately to extreme weather conditions. Based on subjective analysis of a series of descriptive statistics, 10 of the 12 models were judged to lack adequate sensitivity variations in weather variables (ref. 2).

PHASE II (1976 CROP YEAR) RESULTS

Model Modifications

As a result of the Phase I evaluation, the U.S. models were modified before use in Phase II operations. Two major types of modifications were implemented (ref. 3). The first change involved limiting the allowable range of values for each meteorological variable. In the event that the observed total precipitation exceeded the 90th percentile of historically observed values, the total was reduced to the 90th-percentile value. For temperatures, observations occurring outside the historical 95th and 5th percentiles were assigned the value at the appropriate percentile.

The second modification of the yield models for Phase II was based on an analysis of factors, including cropping practices and fertilizer application, which are inherent in the trend term. The models for Texas, Oklahoma, and the Texas-Oklahoma Panhandle assumed that the slope of the trend line was zero after 1960. The remaining models assumed that the trend did not continue to increase after 1972. The original models had extended the trend through the prediction year.

Historical Testing

The improvement in the overall performance of the Phase II models can be seen through the results of the evaluation tests. Table III presents the results of an 11-year bootstrap test of yield models aggregated to the USGP level. In contrast to Phase I, the Phase II models supported the 90/90 goal. Both the mean error and the root-mean-square error (RMSE) of the aggregated results were sharply reduced. Table IV shows that all the individual models supported the 90/90 criterion. The improved performance can be seen by comparing the results presented in table IV with those given in table II. Kansas and North Dakota showed marked improvement.

An analysis of the relation between the estimated variance of the prediction and the observed mean

square error is shown in figure 5. Overall, the variance estimates appeared reasonable. The models for Kansas, Oklahoma, and Texas had greater than expected numbers of cases which were outside the calculated 90-percent confidence interval, a fact which suggests a probable underestimate of variance for those models.

TABLE III.—Eleven-Year Bootstrap Test for the U.S. Phase II Yield Models Aggregated to the USGP by Year With 90/90 Criterion Test

Year	SRS, bu/acre	LACIE, bu/acre	Error ^a	Z	$\Psi(Z)$ ^b
1965	24.0	24.4	-0.3	-49025712	1
1966	22.5	24.2	-1.7	3007423	0
1967	21.7	22.2	-.5	-42796920	1
1968	26.0	24.4	1.7	-7025946	1
1969	28.4	29.3	-.9	-37918618	1
1970	28.2	26.9	1.3	-21651936	1
1971	30.8	28.1	2.6	15931400	0
1972	29.3	29.1	.2	-64047763	1
1973	30.8	29.9	.6	-56037379	1
1974	23.8	27.2	-3.4	82279019	0
1975	26.8	27.4	-.6	-64975437	1

^aMean error = -0.1 bu/acre, RMSE = 1.68 bu/acre
^b $\Psi(Z)$ = 0, except 90/90

TABLE IV.—Eleven-Year Bootstrap Test (1965-75) for U.S. Phase II Yield Models With 90/90 Criterion Test by Model

Model	Crop	Mean error, bu/acre	RMSE, bu/acre	Support 90/90
Montana	SW	0.7	2.16	Yes
North Dakota	SW	-2.5	3.42	Yes
Red River	SW	-2.0	3.96	Yes
South Dakota	SW	-.3	2.45	Yes
Montana	WW	1.0	3.37	Yes
Badlands	WW	1.6	5.00	Yes
Nebraska	WW	2.7	4.23	Yes
Colorado	WW	-.5	4.55	Yes
Kansas	WW	-.3	3.72	Yes
Oklahoma	WW	1.6	3.00	Yes
Panhandle	WW	1.1	3.23	Yes
Texas Low Plains	WW	.2	2.59	Yes
Total	SW	-1.6	2.70	
Total	WW	.7	1.80	
Total	W	-.1	1.68	

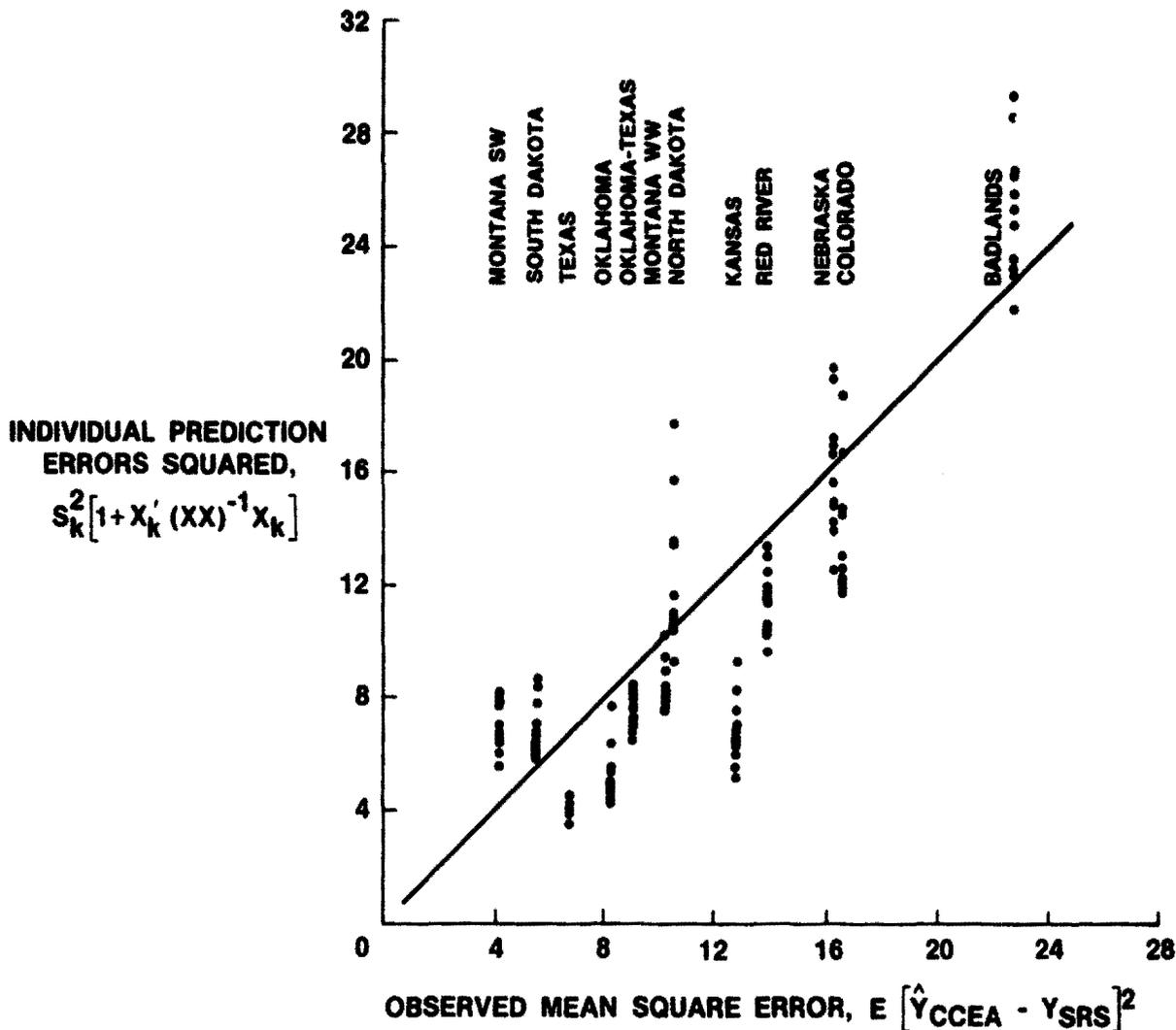


FIGURE 5.—Comparison of estimated and observed variances for yield predictions.

Significant bias, detected by a t-test on the mean error, was found for the North Dakota, Nebraska, and Oklahoma models. This bias was attributed to differences between the zone boundaries used for testing (fig. 2) and those used for developing the models (ref. 3). The areas not included in model development had relatively small acreages of wheat but may have contributed to the observed bias. More significant were those areas, indicated by hatching in figure 6, which were used in the development of adjacent models. For example, inclusion of the low yield in the Nebraska Panhandle in the development of the Nebraska model would result in low estimates when applied to the rest of Nebraska.

Operational use of the yield models for the USGP area during Phase II gave extremely promising results. Table V shows a comparison of the LACIE and SRS yield estimates for the end of season. The results are given by state together with aggregated figures for the spring and winter wheat regions and for the USGP. With the exception of South Dakota, where the LACIE yield estimates did not completely capture the full effects of a severe drought, the performance was remarkable. Reexamining the test results for these models (table III) reveals that there were only 2 years during the 1965-75 period in which the Phase II accuracy was equaled or exceeded.

Table VI gives a month-by-month comparison of

the LACIE and SRS yield estimates for spring, winter, and total wheat for the USGP. As can be seen, the SRS estimates rise steadily and converge with the relatively constant LACIE estimates.

Historical studies of the LACIE yield models for Canada (ref. 4) and for the U.S.S.R. (ref. 5) were conducted. Evaluation of the model tests indicates that the 90/90 criterion was supported at the country level. All Canadian models and all winter wheat models individually supported 90/90. However, 7 of

23 U.S.S.R. spring wheat yield models were judged inadequate.

Operational Testing

Limited testing of the LACIE models for Canada and the U.S.S.R. was carried out in an operational mode. Table VII compares the LACIE yield estimates with those provided by the USDA FAS. The operational test covered the Canadian prairies, representing 16 CRD's, and two indicator regions in the U.S.S.R., covering 36 districts.

The lack of detailed regional figures in foreign areas makes meaningful evaluation of these estimates difficult. This points to the very real need to base the primary determination of the capabilities of the LACIE yield models on their performance in the USGP area, where a detailed error analysis can be carried out.

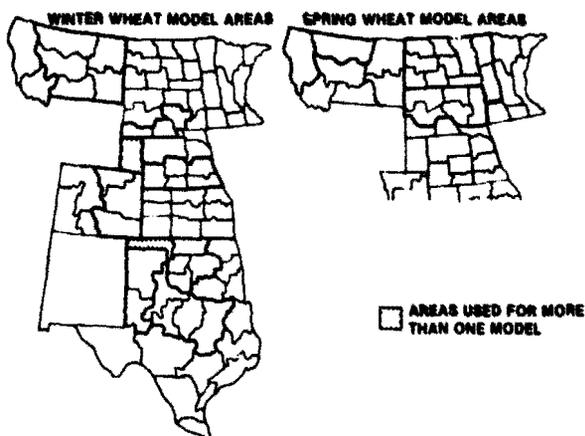


FIGURE 6.—Boundaries for weather data used in developing the yield models tested in Phase II.

TABLE V.—Phase II (1976 Crop Year) Results From LACIE Operational Yield Models Compared With SRS for Final Estimate

Area	Crop	SRS, bu/acre	LACIE, bu/acre	Error, bu/acre	RD ^a
Montana	SW	29.4	27.1	2.3	-8.5
North Dakota	SW	24.7	27.0	-2.3	8.5
Minnesota	SW	32.4	30.3	2.1	-6.9
South Dakota	SW	10.9	17.2	-6.3	36.6
Montana	WW	32.0	29.9	2.1	-7.0
South Dakota	WW	18.0	31.6	-13.6	43.0
Nebraska	WW	32.0	32.7	.7	2.1
Colorado	WW	21.5	19.6	1.9	-9.7
Kansas	WW	30.0	31.0	-1.0	3.2
Oklahoma	WW	24.0	22.6	1.4	-6.2
Texas	WW	22.0	18.7	3.3	-17.6
USGP	SW	25.3	26.2	-.9	3.4
USGP	WW	27.0	27.0	0	0
USGP	TW	26.4	26.7	-.3	1.1

^aRelative difference = ((LACIE - SRS) / LACIE) × 100 percent

PHASE III (1977 CROP YEAR) RESULTS

Model and Methodology Modifications

For Phase III, two additional yield models were developed to expand coverage to areas not previously modeled. Increasing wheat production in parts of Minnesota traditionally planted to other

TABLE VI.—Comparison of Phase II (1976 Crop Year) LACIE and SRS Yield Estimates by Month for the USGP

[In bushels per acre]

Month	Spring wheat		Winter wheat		Total wheat	
	SRS	LACIE	SRS	LACIE	SRS	LACIE
February			19.8	27.6		
March			19.8	27.0		
April			22.7	25.9		
May			24.9	25.3		
June			24.8	26.5		
July			26.4	26.7		
August	24.3	26.3	26.9	26.7	25.9	26.6
September	26.4	26.3	26.9	27.0	26.7	26.8
October	25.7	26.2	26.9	27.0	26.4	26.7
Final	25.3	26.2	27.0	27.0	26.4	26.7

crops resulted in a new model for that part of the state not previously covered by the Red River model. In addition, a model for south-central Texas and parts of the coastal areas was implemented. In each of these models, a linear trend from the period 1955-75 was used with no change in trend after 1975.

A modification to the historical testing of the models was implemented to make the historical tests more representative of the results that would have been obtained if the models had been in operation throughout the test period. On the assumption that an inflection point in the trend term would not be recognized in real time, the trend component of the model was continued for 2 years beyond the year that hindsight analysis had shown to be a breakpoint.

Historical Testing

The results of a 10-year (1967-76) period of using this "continued" trend procedure are shown in table

TABLE VII.—Phase II (1976 Crop Year) Results From LACIE Operational Yield Models for Foreign Indicator Regions Compared with FAS Estimates

<i>(a) U.S.S.R. winter wheat indicator region</i>				
<i>Period</i>	<i>FAS, qllha</i>	<i>LACIE, qllha</i>	<i>Error, qllha</i>	<i>RD^d</i>
Early season	24.0	25.7	-1.7	6.6
Midseason	24.7	25.3	-.6	2.4
Harvest	27.6	24.6	3.0	-12.3
<i>(b) U.S.S.R. spring wheat indicator region</i>				
<i>Period</i>	<i>FAS, qllha</i>	<i>LACIE, qllha</i>	<i>Error, qllha</i>	<i>RD^d</i>
Early season	10.0	10.7	-0.7	6.5
Midseason	10.9	10.6	.3	-2.8
Harvest	11.3	10.5	.8	-7.6
<i>(c) Canada spring wheat indicator region</i>				
<i>Period</i>	<i>FAS, bu/acre</i>	<i>LACIE, bu/acre</i>	<i>Error, bu/acre</i>	<i>RD^d</i>
Early season	29.6	27.7	1.9	-6.8
Midseason	29.6	27.8	1.8	6.5
Harvest	31.1	27.7	3.4	-13.3

^dRelative difference = $(LACIE - FAS) / FAS \times 100$ percent

VIII. The estimated yields for 1967, 1973, and 1974 are those affected by the modification in testing procedures. Other small changes from the results shown in table III stem from including the newly modeled regions in the aggregation. The reported RMSE's increased as a result of this procedure, reflecting a more realistic picture of the models' true predictive abilities. The test results also show that, based on the historical test, the Phase III yield models supported the 90/90 criterion.

Table IX shows that all the individual models supported the 90/90 criterion. The spring wheat models as a group tended to overestimate yield, with particular problems occurring in the North Dakota and Red River models. Figure 7 gives the results of a contingency test for the spring wheat models. The deviation of the actual yields from the model trend was compared with the relative model error. An overestimation of below-normal yields and an underestimation of above-normal yields were found. The χ^2 value was equal to 33.79 with 16 degrees of freedom and is significant at the 1-percent level. The modeled trends overall appear to be overestimates of the actual trend, contributing to the tendency toward a positive bias for the aggregated total spring wheat.

The winter wheat models performed well as a group. The models for the Badlands, Colorado, and Kansas showed the largest error rates. The Kansas

TABLE VIII.—Ten-Year Bootstrap Test for U.S. Phase III Models With Continued Trend

[In bushels per acre]

<i>Year</i>	<i>Total wheat</i>		<i>Spring wheat</i>		<i>Winter wheat</i>	
	<i>SRS</i>	<i>Model error^a</i>	<i>SRS</i>	<i>Model error^b</i>	<i>SRS</i>	<i>Model error^c</i>
1967	21.6	0.9	22.9	0.3	21.0	1.1
1968	26.0	-1.4	26.1	-1.9	25.9	-1.2
1969	28.4	1.0	28.4	2.2	28.4	.5
1970	28.2	-1.6	23.5	-1.0	30.4	-1.9
1971	30.8	-2.9	30.6	-1.7	30.9	-3.7
1972	29.3	-.2	28.5	2.2	29.7	-1.5
1973	30.8	-.2	27.7	.2	32.4	-.3
1974	23.8	4.6	20.8	6.6	25.5	3.4
1975	26.8	.5	25.7	.8	27.4	.3
1976	26.4	.7	25.3	2.0	27.1	-.1

^aMean error, 0.1 bu/acre, RMSE, 1.90 bu/acre

^bMean error, 1.0 bu/acre, RMSE, 2.56 bu/acre

^cMean error, -.04 bu/acre, RMSE, 1.84 bu/acre

model may be of particular concern because in recent years Kansas has accounted for nearly 40 percent of the winter wheat and 25 percent of the total wheat production in the USGP.

For comparative purposes, the acreage and production of the newly modeled regions were removed from the aggregation for 1976. The resulting estimated yield for the USGP in 1976 was 26.95 bushels per acre, which corresponds well with the 27.0 bushels per acre obtained during Phase II operations. This is important in that there was concern that the weather data used for model input during operations, which was derived from analyses of the synoptic scale weather network, was not comparable with the data used for model development and testing, which came from a much higher density climatic observation network. Figure 8 shows the results of a more detailed comparison between yield estimates calculated from high- and low-density meteorological data. On the basis of 2 years' data, it appears that the operational handling of meteorological data introduces no significant bias to the yield estimate.

TABLE IX.—Ten-Year Bootstrap Test (1967-76) for U.S. Phase III Yield Models Using Continued Trend With 90/90 Criterion Test by Model Region

Model	Crop	Mean error, bu/acre	RMSE, bu/acre	Support 90/90
Montana	SW	0.6	2.18	Yes
North Dakota	SW	-1.2	2.94	Yes
Red River	SW	-1.4	3.95	Yes
Minnesota	SW	-.6	3.81	Yes
South Dakota	SW	-.8	3.00	Yes
Montana	WW	.3	2.69	Yes
Badlands	WW	.1	4.61	Yes
Nebraska	WW	-.2	2.92	Yes
Colorado	WW	.8	3.42	Yes
Kansas	WW	.3	3.39	Yes
Oklahoma	WW	-.1	2.21	Yes
Panhandle	WW	.5	2.69	Yes
Texas Low Plains	WW	.6	2.74	Yes
Texas Edwards Plateau	WW	.8	2.88	Yes
Texas South Central	WW	-.8	2.69	Yes
Total	SW	-1.0	2.56	
Total	WW	.4	1.84	
Total	W	-.1	1.90	

Early in Phase III, a series of tests was conducted in which the ability of the models to predict yield before the at-harvest estimate was evaluated. The results of these trials for Kansas are typical and may be seen in table X. Clearly, as successive months of weather data are added to the model, the skill of the prediction increases. The variability associated with fitting the piecewise trend as each new year is added to the model is evident. Since the models are predicting weather-induced variations around the trend, the trend specification is a significant source of potential error.

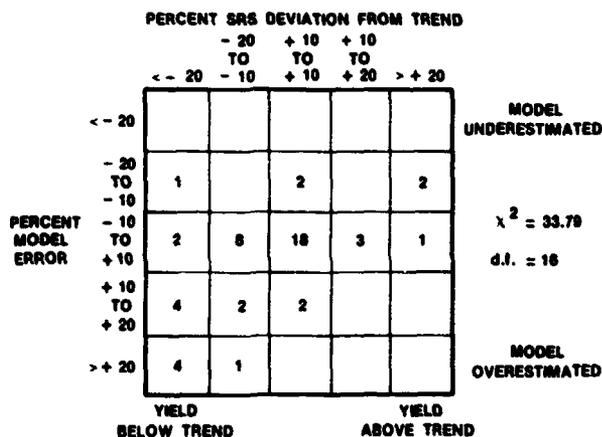


FIGURE 7.—Contingency table of model error and deviation of actual yield from trend for all spring wheat models.

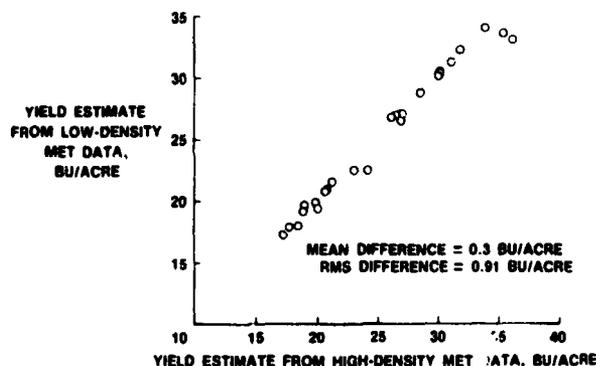


FIGURE 8.—Comparison of yield estimates resulting from high- and low-density input meteorological data for crop years 1976 and 1977 for each U.S. model.

Operational Testing

Table XI gives the results obtained by LACIE during the Phase III (1977) crop year. A comparison of the end-of-season yield estimates produced by LACIE and SRS by state shows sizable errors. Significant underestimates occurred in Montana and Minnesota for spring wheat. In winter wheat, notable underestimates also occurred in Texas and Oklahoma, with a corresponding overestimate in Kansas. The resulting aggregations also underestimated the yield for spring, winter, and total wheat. An examination of the historical tests shows that—in contrast with Phase II—only in 2 years were the aggregated yield estimates *worse* than during Phase III. Table XII shows the monthly progression of SRS and LACIE yield estimates.

During Phase III operations, the only foreign application of the LACIE yield models was in the U.S.S.R. Table XIII presents a comparison of the LACIE and FAS yield estimates on a month-by-month basis. LACIE yield estimates for total wheat yield were 11 percent below those supplied by FAS for the at-harvest estimate.

A complete assessment of the radically different performance of the LACIE yield models between Phase II and Phase III in the USGP is difficult. Phase II represented a relatively normal crop year. Although conditions were dry, the weather was fairly constant. Phase III began dry, and the LACIE yield models place emphasis on early-season moisture conditions. Plentiful but erratic precipitation followed throughout the growing season. The models were unable to reflect the adaptability of the crop, which took advantage of the erratic but improving conditions.

As a result of the ongoing evaluation of the LACIE yield models, a model-by-model revision was completed prior to operational use during the LACIE Transition Year. A complete statistical reevaluation of the weather and trend components of each model was completed, resulting in the first complete revision in yield models for LACIE operations. At this time, an evaluation of the revised models has not been completed. However, a consensus seems to have emerged that the results presented here represent the state of the art for a model of this spatial and temporal resolution.

TABLE X.—Results of Ten-Year Bootstrap Test for Phase III Kansas Winter Wheat Yield Model by Truncation

(In bushels per acre)

Year	SRS	CCEA truncation				
		Trend	Feb.	Mar.	May	June
1967	20.0	25.4	22.4	20.8	22.4	20.6
1968	26.0	24.5	23.3	22.3	24.0	24.4
1969	31.0	25.1	26.8	30.1	30.7	31.7
1970	33.0	26.9	26.9	29.1	29.3	30.0
1971	34.5	28.8	28.7	27.7	28.6	28.9
1972	33.5	30.7	29.9	28.6	29.6	29.6
1973	37.0	31.2	32.7	35.0	34.6	35.9
1974	27.5	32.1	33.4	33.6	32.2	32.8
1975	29.0	31.5	31.3	32.0	32.3	31.9
1976	30.0	31.2	29.0	29.2	30.2	30.3
Mean error		1.41	1.71	1.31	0.76	0.54
RMSE		4.54	4.17	3.89	3.35	3.11

TABLE XI.—Phase III (1977 Crop Year) Results From LACIE Operational Yield Models Compared With SRS for Final Estimates for the USGP Area

Area	Crop	SRS, bu/acre	LACIE, bu/acre	Error, bu/acre	RD ^a
Montana	SW	22.0	18.0	4.0	-22.2
North Dakota	SW	24.9	23.1	1.8	-7.8
Minnesota	SW	39.9	32.0	7.9	-24.7
South Dakota	SW	23.5	20.8	2.7	-13.0
Montana	WW	29.0	26.5	2.5	-9.4
South Dakota	WW	25.0	27.1	-2.1	7.7
Nebraska	WW	35.0	32.0	3.0	-9.4
Colorado	WW	22.0	22.5	-.5	2.2
Kansas	WW	23.5	28.8	-5.3	18.4
Oklahoma	WW	27.0	20.0	7.0	-35.0
Texas	WW	25.0	20.3	4.7	-23.2
USGP	SW	27.1	23.4	3.7	-15.8
USGP	WW	27.7	25.6	2.1	-8.2
USGP	TW	27.5	24.9	2.6	-10.4

^aRelative difference = ((LACIE - SRS) / LACIE) × 100 percent.

TABLE XII.—Comparison of Phase III (1977 Crop Year) LACIE and SRS Yield Estimates by Month for the USGP Area

(In bushels per acre)

Month	Spring wheat			Winter wheat			Total wheat		
	SRS	LACIE	Error	SRS	LACIE	Error	SRS	LACIE	Error
February				22.5	25.7				
May				28.2	25.5				
June				29.2	25.5				
July	27.0	24.8		28.4	25.6		28.0	25.4	
August	27.7	23.4		27.7	25.6		27.7	24.9	
September	26.9	23.6		27.8	25.5		27.5	24.9	
October	26.7	23.4		27.8	25.6		27.5	24.9	
Final	27.1	23.4	3.7	27.7	25.6	2.1	27.5	24.9	2.6

TABLE XIII.— Comparison of LACIE and FAS/U.S.S.R. Yield Estimates for Phase III (1977 Crop Year)

Month	Winter wheat			Spring wheat			Total wheat		
	FAS/U.S.S.R., ql/ha	LACIE, ql/ha	RD, percent	FAS/ U.S.S.R., ql/ha	LACIE, ql/ha	RD, percent	FAS/ U.S.S.R., ql/ha	LACIE, ql/ha	RD, percent
April		24.3							
May		24.1							
June		25.6							
July		25.9							
August	27.0	25.5	-5.9	11.0	9.0	-22.2	16.0	15.2	-5.3
September	28.8	25.6	-5.5	9.7	9.0	-7.8	16.1	14.7	-9.5
October	28.8	25.6	-5.5	9.7	8.8	-10.2	16.1	14.5	-11.0
Final	28.8	25.6	-5.5	9.7	8.8	-10.2	16.1	14.5	-11.0

CONCLUSIONS

The LACIE yield models which were developed, implemented, and tested during the three phases of this experiment represent the first generation of yield models designed for the large-area prediction of wheat yields. The models are capable of supporting the stated project goal of being within 10 percent of the actual wheat production 90 percent of the time.

The limitations of these models are inherent in their nature. The temporal resolution (1 month) limits their ability to handle the erratic weather occurring in critical situations. The assumption that the crop is always in the same growth stage during a given month limits the model's ability to respond to early or late crop development in a particular year.

This was particularly apparent in 1974 when planting in the spring wheat region was up to 1 month late. The relatively large spatial resolution of the individual models limits the capture of localized but important episodic events. However, the LACIE yield models have provided a valuable baseline understanding of the problems associated with predicting yields for large areas. In addition, these models provide a valuable benchmark against which to compare more sophisticated models which are designed to overcome current limitations and to provide a second generation of predictive yield models (see the paper by Stuff et al. entitled "Status of Yield Estimation Technology: A Review of Second-Generation Model Development and Evaluation" and the paper by Cate et al. entitled "The Law of the Minimum and an Application to Wheat Yield Estimation").

REFERENCES

1. Wheat Yield Models for the United States. LACIE-00431, Rev. A, NASA Johnson Space Center, June 1977.
2. YES Phase I Yield Feasibility Report. LACIE-00439, NASA Johnson Space Center, Jan. 1977.
3. The Effect of Flagging and Trend Adjustments to Wheat Yield Estimates With CCEA Great Plains Models. NOAA Center for Climatic and Environmental Assessment (Columbia, Mo.), 1978.
4. Yield-Weather Regression Models for the Canadian Prairies. LACIE-00432, Rev. A, NASA Johnson Space Center, June 1977.
5. Wheat Yield Models for the U.S.S.R. LACIE-00430, Rev. A, NASA Johnson Space Center, June 1977.

Accuracy and Performance of LACIE Crop Development Models

S. K. Woolley,^a V. S. Whitehead,^b R. G. Stuff,^b and W. E. Creab^b

INTRODUCTION

The intent of this paper is to describe the accuracy of the crop development model during each of the 3 years of LACIE. The estimated and observed crop development data were compared in order to establish a measure of confidence in the model and to identify consistent discrepancies that would adversely affect LACIE operations. Although the model provided reliable estimates for various wheat-growing regions of the world, it was found that there are still areas in need of further model improvement or development.

CROP CALENDAR MODEL

Crop development models, referred to in this report as crop calendar models, served two purposes in LACIE. First, they provided an estimate of the stage of development to the yield model builder. This is an important variable in the more recently developed yield models such as the Feyerherm model (ref. 1) and the Cate-Liebig model (see paper by Cate et al. entitled "The Law of the Minimum and an Application to Wheat Yield Estimation"). These model forms were developed to account for the change in response of wheat yield to environment as the plant progresses toward maturity. By providing growth stage estimates, the crop calendar model determined when the coefficients or response functions in the yield model changed from one set to another. Second, the crop calendar model output provided an indicator to the analyst-interpreter of what the appearance of the crop should be (i.e., at 2.3, the analyst-interpreter should begin to detect vegetation;

near 5.0, he should detect the vegetation turning from green to gold; etc.).

The accuracies required to support these users have been difficult to ascertain. Because the application of the crop development estimates refers to several fields or farms (sample segment for analysis; pseudozone for yield), it would not be practical to strive for better accuracy in time than the equivalent spatial scatter of the crop progression in the area of interest. The standard deviation of crop progression within the sample segments has not been previously determined and probably varies greatly with time and location and time of year, but a period of 5 to 7 days appears to be a reasonable goal for crop calendar accuracy.

A study was made to compare and evaluate the principal phenological crop calendar models available for spring wheat. The models evaluated were the growing-season degree-day, photothermal units, and Robertson's triquadratic model (ref. 2). From these tests, the Robertson model (ref. 3) was recommended and selected for use in LACIE applications.

The Robertson model predicts the rate of progression of wheat through its biological development. Daily maximum and minimum temperatures and day length are the input variables. Day length is calculated internally from the latitude and the date. The principal output of the crop calendar model is a daily increment of development (DID) through six physiological stages of growth.

Robertson's model consists of the product of quadratic expressions involving the three input variables (thus the term "triquadratic model"). For the LACIE application, a quadratic equation was used to calculate the DID within each of six physiological stages. The increments are accumulated from stage to stage. Because wheat responds differently to the environment during each physiological stage of growth, five different equations are required.

^aLockheed Electronics Company, Houston, Texas.
^bNASA Johnson Space Center, Houston, Texas.

The model was developed using only data from Canada for spring wheat. Terms and coefficients are the same for all locations. In 1976, Feyerherm (ref. 1) developed a scalar multiplier that was applied to the initial equations between emergence and heading and that reflected the effect of dormancy on winter wheat. With the Feyerherm multipliers, the model with Robertson's original coefficients produced fairly reliable estimates of the heading and ripening times for winter wheat. A more detailed explanation of the model appears in the paper by Whitehead and Phinney entitled "Growth Stage Estimation."

APPLICATIONS SUMMARY

Phase I Operations

The National Oceanic and Atmospheric Administration (NOAA) had the responsibility for design, implementation, and operation of the adjustable crop calendar (ACC) model. Because of limited NOAA resources, the National Aeronautics and Space Administration (NASA) assisted in the design and implementation of the model. In Phase I, the model was designed to be run at the crop-reporting-district (CRD) level on the U.S. Great Plains. NASA provided normal crop calendars for all CRD's in the U.S. Great Plains states (ref. 4). The model then brought together the data base and current meteorological data to generate updates. The model was not fully implemented until spring 1975. For winter wheat in each CRD in the U.S. Great Plains states where a winter wheat sample segment existed, the U.S. Department of Agriculture (USDA) supplied the Yield Estimation Subsystem (YES) with the actual date at which 50 percent of the crop had begun to joint. The model was started in spring wheat CRD's on the actual date when 50 percent of the crop had been planted. Daily maximum and minimum temperatures were selected for a representative first-order weather station in a particular CRD. If none existed, weighted values were used from the nearest neighboring stations. The temperature data used were in punched card form. The model was updated every 2 weeks in the batch mode on the IBM 370/168 computer facility at the University of Missouri by personnel from the NOAA Center for Climatic and Environmental Assessment (CCEA) at Columbia, Missouri, and was mailed to the NASA Johnson Space Center (JSC). Seven developmental stages

were implemented for the LACIE project for spring and winter wheat. These stages of development were identical to those used by Robertson, but the corresponding numbers selected by LACIE were one greater than those defined by Robertson. The stages of development and their corresponding numbers on the time scale were as follows.

- 1.0 Planted
- 2.0 Emerged
- 3.0 Jointed
- 4.0 Headed
- 5.0 Soft dough
- 6.0 Ripe
- 7.0 Harvest

The whole numbers indicate that 50 percent of the wheat in an area had reached that particular development stage. Thus, for stage 4.0, 50 percent of the wheat in the area had begun to head.

Phase II Operations

The NASA assumed the responsibility for design and implementation in the new winter and spring wheat regions. The winter wheat regions under study in Phase II were expanded to include the U.S.S.R. and the People's Republic of China (P.R.C.), in addition to the U.S. Great Plains. The corresponding spring wheat areas were Canada, the U.S.S.R., the P.R.C., and the U.S. Great Plains. NOAA retained the responsibility for the operation of the model, but the location for model operations was transferred from Columbia, Missouri, to the IBM 360/65 facility in Washington, D.C.

The model for winter wheat was started with a normal end-of-dormancy restart model (ref. 5) in the U.S. Great Plains. Corresponding end-of-dormancy dates and development stage numbers were obtained from climatic analogs and were transferred to crop calendar stations in the U.S.S.R. and the P.R.C. The model for spring wheat was started in all spring wheat countries using the planting model developed by Feyerherm (ref. 6). Crop calendar adjustments were provided for a weather station instead of for an area, such as the CRD usage in Phase I. These estimates were updated biweekly at first-order stations and transmitted to JSC via a time-sharing operation (TSO).

Toward the end of Phase II operations, the location of model operation was transferred from Washington, D.C., to the IBM 360/195 computer facility at Suitland, Maryland. The data continued to be transmitted to JSC via a TSO.

During the 1976 growing season, Southern Hemisphere calendars were initiated for 25 selected stations in Argentina, Australia, and Brazil. The transfer of crop calendar estimates was delayed several times because of model errors and meteorological data problems and was not completed until the summer of 1977.

Phase III Operations

Because of the lack of success of a winter wheat starter model, the ACC was started using normal fall planting dates in the U.S. Great Plains, the U.S.S.R., and the P.R.C. The model estimates were generated on the IBM 360/195 computer facility at Suitland, Maryland, and continued to be transmitted to JSC via a TSO. The estimates were updated biweekly and were modified by use of the Feyerherm scalar multiplier at each crop calendar station. The multipliers were applied to the ACC equations between emergence and heading to improve the model's accuracy. In addition to the multipliers, another control was introduced to the model to prohibit crop calendar advancement beyond stage 2.85 before January 1 to prevent the model from predicting jointing before spring green-up.

At the end of January 1977, a significant error in the LACIE crop calendar algorithm was identified. The problem stemmed from an ambiguity in Robertson's original paper on the crop calendar, which led to an error in the technique used in the LACIE model to eliminate "negative" growth. When temperatures were very low or day lengths short, the LACIE model erroneously allowed the development to continue during the emergence to jointing stage. As a result of the model's error, operations were temporarily suspended. After the ACC program was changed to incorporate the corrected algorithm, the winter wheat estimates were restarted from the fall planting dates and run through the winter. New estimates for the United States, the U.S.S.R., and the P.R.C. arrived at JSC during the middle of March and continued to be acquired on terminal via a TSO on the scheduled biweekly basis for the remainder of the growing period. The P.R.C. was soon dropped from further LACIE Phase III analysis by a project directive changing the scope of Phase III operations. Output of the adjustments continued to be provided in isoline map format.

The spring wheat starter model was once again used to begin operations in the spring wheat areas of

the United States, the U.S.S.R., and Canada. By the middle of July, it was obvious that the model estimates for the U.S.S.R. east of the Ural Mountains were running possibly 3 weeks ahead of the development stages as determined from the imagery acquired from Landsat. Agricultural practice in the U.S.S.R. is to delay planting in these "New Lands" areas until the end of May, regardless of weather conditions. Thus, the start dates derived from the spring wheat starter model proved to be several weeks early. To obtain revised estimates for this spring wheat region, the ACC model was operationally restarted using June 5 development stage estimates for 15 meteorological stations. The restart development stage estimates were agreed on jointly by the YES managers and the Classification and Mensuration Subsystem (CAMS) team of U.S.S.R. analysts. These were subjectively obtained from planting information contained in USDA Foreign Agricultural Service (FAS) reports from the region and from corresponding Landsat imagery. By the end of July, the revised crop calendar updates were processed and delivered to the users at JSC.

A second form of crop calendar output was made available by means of a gridded format at the segment basis, whereby the crop calendars produced by NOAA were extended to the LACIE sample segments (refs. 7 and 8).

PERFORMANCE RESULTS

Phase I

Approximately every 18 days, ground-truth data were recorded on forms at specified intensive test sites (ITS's) by personnel from the USDA Agricultural Stabilization and Conservation Service and mailed to JSC (fig. 1). Only data from six winter wheat and six spring wheat sites were available for Phase I analyses. The development stage of wheat at each ITS was obtained by converting the growth stage reported at the fields within the site to the LACIE version of the Robertson biometeorological time scale (BMTS) noted in table I and then simply averaging these numbers. Graphic comparisons of the differences between the ACC estimate, the historical calendar for that particular CRD, and the ITS ground truth were then made. Such a comparison is shown in figure 2 for the ITS in Finney County, Kansas.

LAND USE CODES		GROWTH STAGES	
100-SPRING WHEAT		01-NOT PLANTED	
200-BARLEY		02-PLANTED, NO EMERGENCE	
300-OATS		03-EMERGENCE	
400-WINTER WHEAT		04-TILLERING, PREBOOT, PRFBUD	
500-GRASSES/PASTURE		05-BOOTED OR BUDDED	
600-OTHER CROPS		06-BEGINNING TO HEAD OR FLOWER	
601-RAPESEED		07-FULLY HEADED OR FLOWERED	
602-RYE		08-BEGINNING TO RIPEN	
904-FLAX		09-RIPE MATURE	
607-CORN		10-HARVESTED	
617-SOYBEANS		11-DOES NOT APPLY	
618-COTTON			
700-SUMMER FALLOW			
900-UNKNOWN CROPS			

FIELD NO.	LAND USE CODE	GROWTH STAGE (CIRCLE ONE)
107	504	01 02 03 04 05 06 07 08 09 10 11
128	200	01 02 03 04 05 06 07 08 09 10 11
104	504	01 02 03 04 05 06 07 08 09 10 11
124	504	01 02 03 04 05 06 07 08 09 10 11

FIGURE 1.—Ground-truth periodic observation form.

The first step in the analysis was to determine the difference in days between the observed and the computed data. Tables of delta days were constructed using a plus sign to indicate that the model date was earlier than the observed date (i.e., model fast) and a minus sign to reflect a model date that

TABLE 1.—Robertson BMTS and Observed ITS Wheat Phenological Stages

Stage	BMTS	ITS	Description
Planted	1.0	01	Planted
		02	Planted, no emergence
Emergence	2.0	03	Emergence
Jointing	3.0	04	Tillering, prebooting, prebudding
		05	Booted or budded
Heading	4.0	06	Beginning to head or flower
		07	Fully headed or flowered
Soft dough	5.0	08	Beginning to ripen
Ripening	6.0	09	Ripe to mature
Harvest	7.0	10	Harvest

was later than the observed date (i.e., model slow). A similar comparison was made of the historical crop calendar and the development observed at the various ITS's in an effort to determine whether the model estimates provided more realistic information than the normals.

The bias at each stage was computed by the following expression:

$$BIAS_j = \frac{1}{n} \sum_{i=1}^n (ITS_i - ACC_j)$$

where n = number of ITS's

ITS_j = date when an ITS reported that 50 percent of the crop was at stage j , where $j = 3, 3.5, 4.0, 4.5, 5.0,$ and 6.0

ACC_j = date when the ACC model reached a particular stage j

Similarly, by substituting the historically averaged crop calendar for the ACC curve, one can use the expression to determine another bias, which becomes

$$BIAS_j = \frac{1}{n} \sum_{i=1}^n (ITS_i - HIST_j)$$

where $HIST_j$ = date that the crop normally reached a particular stage j (from historical average).

The standard deviation (SD) and the root mean

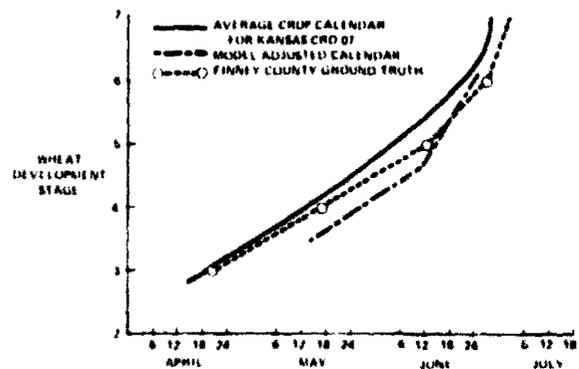


FIGURE 2.—Historical, observed, and predicted crop calendar stages for winter wheat in Finney County, Kansas, 1974-75.

square error (RMSE) were calculated for each data set by the following formula:

$$SD = \left(\frac{\sum_{j=1}^n [(ITS_j - ACC_j) - (\overline{ITS} - \overline{ACC})]^2}{n - 1} \right)^{\frac{1}{2}}$$

$$RMSE = \left[\frac{1}{n} \sum_{j=1}^n (ITS_j - ACC_j)^2 \right]^{\frac{1}{2}}$$

The computed bias could have been due to several factors, such as model errors, differences in the stage definitions from one locale to another, and observer errors. The values of the bias, SD, and RMSE for both model estimates versus ITS observed data and historical data versus ITS data are given in table II. The biases and SD's for winter and spring wheat were plotted for the crop development stages of heading and soft dough and are shown in figures 3 through 6. The crop calendar results at heading were chosen for closer examination because, at that particular development stage, very little ambiguity about its definition existed among ground observers. The analyses at soft dough were also scrutinized more closely because the general color-infrared characteristics of wheat change from red to orange shortly after reaching this stage and are readily distinguishable on Landsat imagery.

TABLE II.—Comparison of LACIE Phase I ACC and Historical CRD Calendars With Observed Development Stages in the U.S. 1974-75: Winter and 1975 Spring Wheat ITS's

ITS (county, state)	ACC vs. ITS, days						Historical vs. ITS, days						
	Jointing		Heading		Soft dough	Ripe	Jointing		Heading		Soft dough	Ripe	
	3.0	3.5	4.0	4.5	5.0	6.0	3.0	3.5	4.0	4.5	5.0	6.0	
<i>Winter wheat</i>													
Deaf Smith, Tex.	(a)	3	3	7	-3	5	10	15	18	14	12		
Oldham, Tex.	(a)	1	0	4	-3	-2	5	12	14	13	9		
Randall, Tex.		-4	-11	-8	0	0	6	9	11	10	8		
Morton, Kans.		-12	-9	-8	-2	3	19	10	1	5	11	3	
Finney, Kans.		-23	-13	-10	0	4	1	2	3	5	8	6	
Rice, Kans.	(a)	6	5	7	0	-16	-11	-4	2	9	7		
Bias, days			-3.8	-3.0	2.7	0.2	1.2	3.7	6.0	9.2	10.8	7.5	
SD, days			8.1	6.5	3.9	2.9	11.3	7.8	7.2	6.2	2.3	3.0	
RMSE, days			8.3	6.6	4.4	2.7	10.4	8.0	8.9	10.8	11.0	8.0	
<i>Spring wheat</i>													
Burke, N. Dak.		2	1	-1	1	3	8	23	18	12	9	8	11
Williams, N. Dak.		-13	-10	-5	-2	0	6	7	8	7	8	4	9
Hill, Mont.		5	2	1	2	2	-2	23	10	3	3	7	9
Liberty, Mont.		8	8	8	8	9	17	26	14	10	11	13	27
Toole, Mont.		-6	-8	-10	0	8	25	12	-1	-8	0	13	40
Polk, Minn.		-1	-3	-5	-4	0	3	15	6	0	5	10	14
Bias, days		-0.8	-1.7	-2.0	0.8	3.7	9.5	17.7	9.2	4.0	6.0	9.2	18.3
SD, days		7.7	6.7	6.2	4.1	3.9	9.9	7.5	6.6	7.4	4.1	3.6	12.6
RMSE, days		7.1	6.4	6.0	3.9	5.1	12.4	18.9	11.0	7.8	7.1	9.7	21.6

^aACC data not available

From the limited data set, it can be seen that the model development estimates were closer to those observed at the ITS's than were the historical (normal) curves. This statement held true at each development stage for both: spring and winter wheat. At soft dough, the model estimates were significantly better (at the 1-percent level) than the historical calendars for both winter and spring wheat.

The model for winter wheat averaged about 4 days behind the ITS's at heading, was 3 days ahead at soft dough, and then approached ground truth at ripe. A similar trend was noted for spring wheat, except that the model values were almost 10 days early at ripe because of the large differences at Liberty and Toole, Montana.

Much of the further testing and evaluation of the ACC originally scheduled for fall 1975 and winter 1976 was delayed because resources were diverted to the development of a winter wheat restart model and a spring wheat starter model and the support of testing, evaluation, and implementation of yield models.

Phase II

The performance of the ACC model was monitored for the United States, the U.S.S.R., Canada, and the P.R.C. This tracking effort enabled YES to evaluate the reliability of the ACC estimates as an in-house tool for quality control and served as an indicator to flag regions for which the estimates varied widely from those observed using ground-truth or Landsat imagery data.

The ground-truth ITS network was expanded in Phase II to include 26 ITS's in the United States and an additional 11 ITS's in Canada (figs. 7 and 8). The same analysis procedure used in Phase I was continued in Phase II. Comparisons were made throughout the growing season between ITS ground observations and both the crop calendar adjustments and the historical calendars. A sample plot of both the ACC estimate and the historical curve versus the development stages as reported in the Ellis, Kansas, ITS ground-truth report is given in figure 9. Tables III through V summarize these comparisons and present the model bias (in days), the SD's, and the RMSE's at various stages of development at these 37 winter and spring wheat sites in the United States and Canada. The sign convention remains the same as that used in Phase I.

From these data, it can be observed that for winter wheat, the model estimates were closer to the ITS data at jointing (-3 days compared to the historical

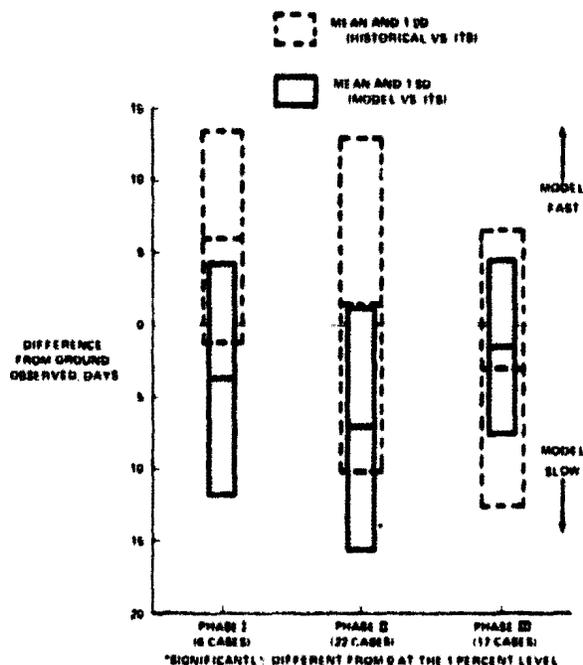


FIGURE 3.—Comparison of U.S. winter wheat ITS observations with CRD historical calendars and LACTE adjustments at heading.

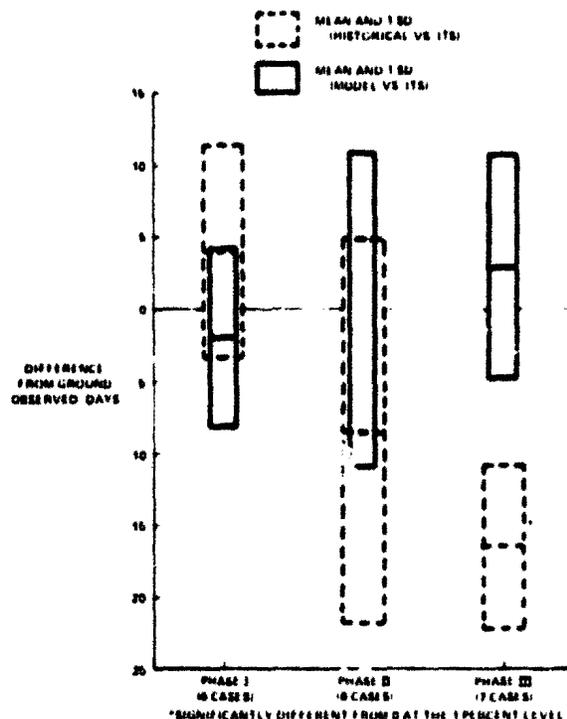


FIGURE 4.—Comparison of U.S. spring wheat ITS observations with CRD historical calendars and LACTE adjustments at heading.

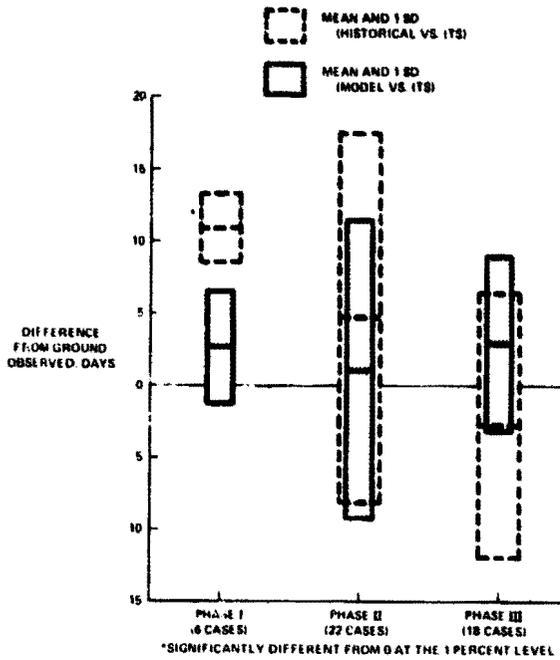


FIGURE 5.—Comparison of U.S. winter wheat ITS observations with CRD historical calendars and LACIE adjustments at soft dough.

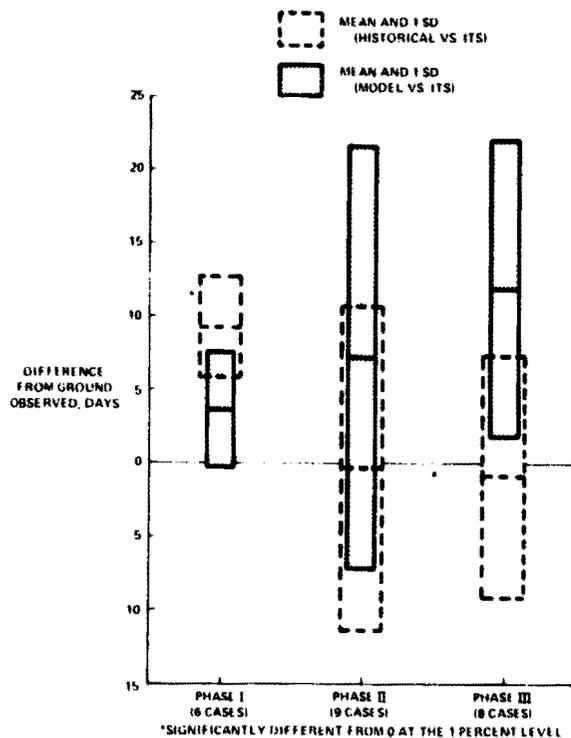


FIGURE 6.—Comparison of U.S. spring wheat ITS observations with CRD historical calendars and LACIE adjustments at soft dough.

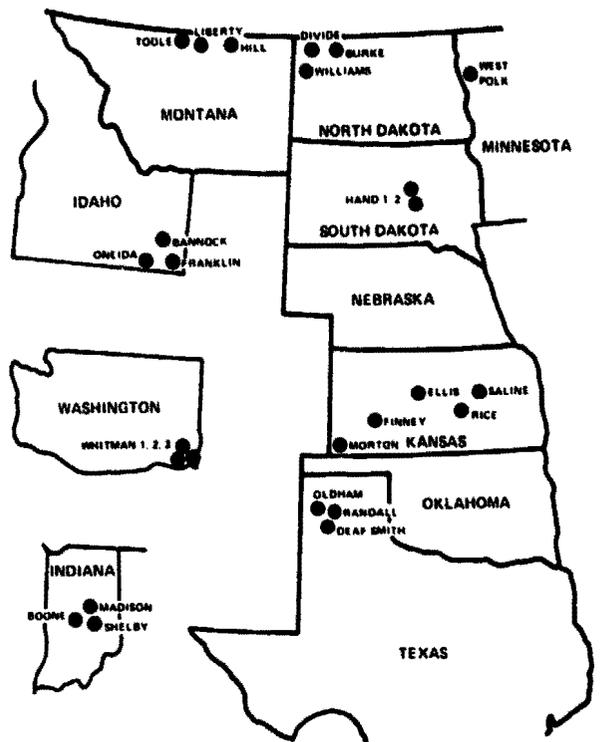


FIGURE 7.—The 26 U.S. intensive test sites.

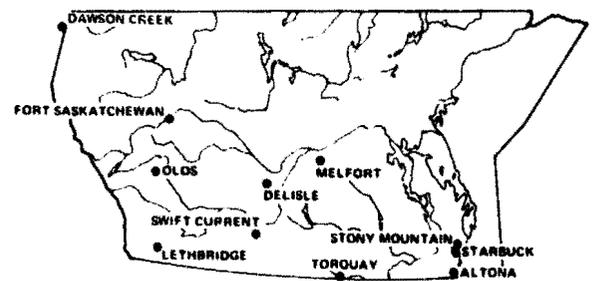


FIGURE 8.—The 11 Canadian intensive test sites.

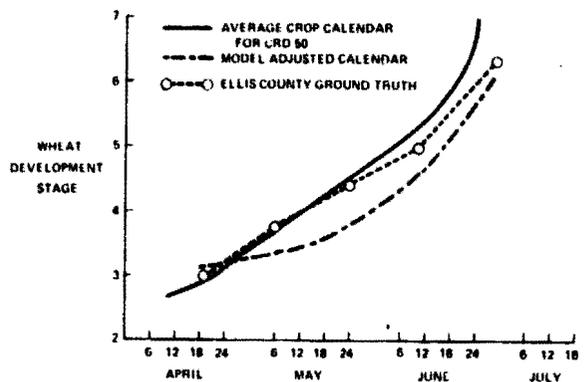


FIGURE 9.—Observed and predicted crop calendar stages for winter wheat in Ellis County, Kansas, 1975-76.

TABLE III.—Comparison of LACIE Phase II ACC and Historical CRD Calendars With Observed Development Stages in the 1975-76 U.S. Winter Wheat ITS's

ITS (county, state)	ACC vs. ITS, days						Historical vs. ITS, days					
	Jointing		Heading		Soft dough	Ripe	Jointing		Heading		Soft dough	Ripe
	3.0	3.5	4.0	4.5	5.0	6.0	3.0	3.5	4.0	4.5	5.0	6.0
Deaf Smith, Tex.	-4	-14	-18	-11	-4	-7	-10	0	7	15	17	8
Oldham, Tex.	14	0	-3	3	-1	-5	10	18	21	31	22	11
Randall, Tex.	-4	-10	-16	-10	-2	-7	-9	0	7	14	17	8
Finney, Kans.	-3	-8	-16	-7	-6	-9	-9	7	1	9	9	0
Rice, Kans.	(a)	(a)	-17	-12	-5	-3	(a)	(a)	-8	-1	5	2
Ellis, Kans.	(a)	-15	-17	-12	-7	-3	-2	-2	0	3	6	5
Saline, Kans.	-6	-16	-18	-14	-6	-3	-24	-15	-9	-2	4	2
Morton, Kans.	0	-1	-4	0	0	1	0	9	11	15	17	15
Boone, Ind.	1	-9	-11	-6	-8	-8	-17	-14	-9	-5	-12	-10
Madison, Ind.	0	-9	-11	-6	4	4	-17	-11	-7	-3	2	2
Shelby, Ind.	24	4	-9	-14	-13	-7	2	-3	-7	-12	-17	-10
Bannock, Idaho	8	3	-3	-8	0	-1	11	12	7	-4	-8	-20
Franklin, Idaho	(a)	6	-4	-11	6	4	(a)	15	8	-5	-3	-11
Oneida, Idaho	(a)	9	0	-5	-3	12	(a)	18	12	3	-8	-7
Whitman (1), Wash.	0	25	14	12	27	30	-3	36	26	18	28	20
Whitman (2), Wash.	-2	0	-8	2	16	11	-2	11	5	-2	13	2
Whitman (3), Wash.	0	8	6	7	27	29	-5	20	17	9	28	21
Hill, Mont. ^b	-18	0	-1	-6	5	7	-24	-7	-16	-16	-5	0
Liberty, Mont. ^b	-22	0	1	7	6	-4	-25	-6	-6	1	1	0
Toole, Mont. ^b	-24	-9	-14	-14	-7	-12	-23	-6	-15	-15	-4	1
Hand (1), S. Dak. ^b	-12	0	-6	-8	-3	-6	-20	2	-7	-6	-3	-13
Hand (2), S. Dak. ^b	-8	-10	-4	0	-4	-6	-17	-9	-6	-3	-6	-13
Bias, days	-3.1	-2.2	-7.2	-5.1	1.0	0.8	-9.7	3.6	1.5	2.0	4.7	0.6
SD, days	11.7	9.8	8.5	7.6	10.4	11.2	11.4	13.1	11.5	11.5	12.8	10.8
RMSE, days	11.8	9.8	11.0	9.0	10.3	11.0	14.7	13.3	11.4	11.4	11.3	10.6

^aITS data not available

^bAcquired as spring wheat ITS but contained numerous winter wheat fields

mean of -10 days) and at soft dough (1 day versus 5 days), whereas the historical crop calendars were significantly more accurate (at the 5-percent level) at heading (1 day versus -7 days). The winter wheat model estimates were approximately 2 to 2.5 weeks slow at heading in the Central and Southern U.S. Great Plains states but were generally less than 1 week off at ripe. At the other 11 ITS's outside Texas, Kansas, and Indiana, the magnitude of the model versus ITS dates at heading was usually less than 1 week. Generally, except for Whitman County, Washington, the model estimates were behind those observed at the ITS's. The Whitman County ITS's

received heavy rains, which apparently tended to slow the development rate of the wheat plant.

Tables IV and V show that the spring wheat model generally ran fast, except for Hand County, South Dakota; Dawson Creek, British Columbia; and Olds, Alberta. The U.S. and Canadian spring wheat ITS's were divided into separate groups for analysis. Figures 4, 6, 10, and 11 present the model's accuracy results at heading and at soft dough.

The spring wheat starter model was the mechanism by which the crop calendar model was begun in the spring wheat regions. Naturally, the accuracy of the planting dates to a great extent determined the

TABLE IV.—Comparison of LACIE Phase II ACC and Historical CRD Calendars With Observed Development Stages in the 1976 U.S. Spring Wheat ITS's

ITS (county, state)	ACC vs. ITS, days						Historical vs. ITS, days					
	Jointing		Heading		Soft dough	Ripe	Jointing		Heading		Soft dough	Ripe
	3.0	3.5	4.0	4.5	5.0	6.0	3.0	3.5	4.0	4.5	5.0	6.0
Hand (1), S. Dak.	-11	-11	-15	-12	-7	-19	-7	-17	-22	-16	-10	-15
Hand (2), S. Dak.	-13	-11	-15	-17	-20	-24	-9	-17	-21	-19	-18	-20
Burke, N. Dak.	9	18	15	14	24	22	8	12	10	9	12	7
Divide, N. Dak.	10	8	7	10	22	24	8	7	6	7	13	9
Williams, N. Dak.	9	10	3	3	13	8	9	6	0	-1	4	-5
Hill, Mont.	-13	-14	-5	0	11	12	-21	-27	-23	-16	-4	1
Liberty, Mont.	0	7	1	-2	1	14	-10	-4	-15	-13	-12	4
Toole, Mont.	17	13	8	12	17	25	10	4	-3	2	7	14
Polk, Minn.	(a)	(a)	(a)	-12	3	2	(a)	(a)	(a)	-11	5	5
Bias, days	1.0	2.5	-0.1	-0.4	7.1	7.1	-1.5	-4.5	-8.5	-6.4	-0.3	0
SD, days	12.0	12.5	10.8	11.3	14.3	17.9	11.7	14.2	13.3	10.8	11.1	11.3
RMSE, days	11.2	12.0	10.1	10.7	15.3	18.4	11.1	14.0	15.1	12.0	10.5	10.5

^aITS data not available

TABLE V.—Comparison of LACIE Phase II ACC and Historical CRD Calendars With Observed Development Stages in the 1976 Canadian Spring Wheat ITS's

ITS (town, province)	ACC vs. ITS, days						Historical vs. ITS, days					
	Jointing		Heading		Soft dough	Ripe	Jointing		Heading		Soft dough	Ripe
	3.0	3.5	4.0	4.5	5.0	6.0	3.0	3.5	4.0	4.5	5.0	6.0
Fort Saskatchewan, Alberta	14	5	-3	-7	-2	3	7	4	0	-8	-8	-9
Olds, Alberta	8	-5	-9	-14	-6	-12	6	8	-2	-10	-7	-11
Lethbridge, Alberta	12	8	6	6	12	15	10	11	11	12	19	18
Dawson Creek, British Columbia	-9	-9	-12	-17	-14	-12	-1	5	3	2	8	5
Stony Mountain, Manitoba	11	10	5	1	9	6	13	10	3	1	9	-1
Starbuck, Manitoba	-2	9	6	2	8	4	0	9	3	1	8	-1
Altona, Manitoba	-1	1	0	-1	-2	-2	2	4	1	0	-2	-4
Delisle, Saskatchewan	9	9	7	3	-1	(a)	12	14	14	13	12	(a)
Swift Current, Saskatchewan	-6	-3	-8	6	8	6	-9	-3	-5	8	9	7
Torquay, Saskatchewan	3	6	4	-1	2	-11	7	9	6	3	0	9
Melfort, Saskatchewan	11	10	5	1	9	6	-3	-3	-3	3	3	-1
Bias, days	3.6	2.6	-0.5	-1.5	1.8	0.3	4.0	6.2	2.8	2.3	4.6	1.2
SD, days	7.7	6.6	6.8	8.0	7.7	9.3	6.8	5.5	5.8	7.1	8.3	8.7
RMSE, days	8.2	6.8	6.5	7.7	7.6	8.8	7.6	8.1	6.2	7.2	9.2	8.4

^aITS data not available

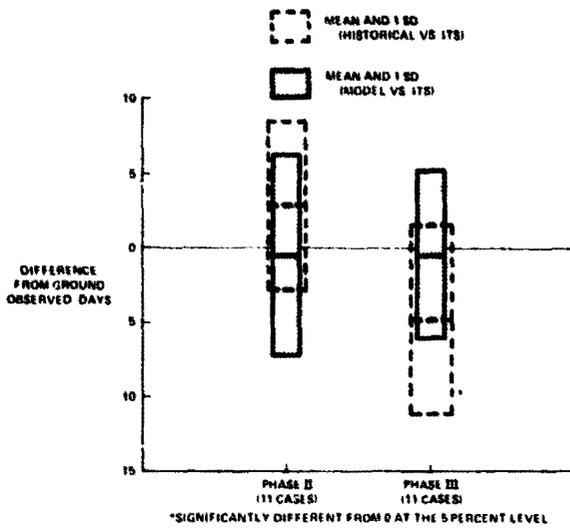


FIGURE 10.—Comparison of Canadian spring wheat ITS observations with CRD historical calendars and LACIE adjustments at heading.

reliability of the model's performance. For winter wheat, dormancy had a tendency to neutralize the effect of early or late plantings on the crop calendar model estimates because fields in the same area tend to emerge from dormancy at the same time. However, the application of incorrect spring wheat planting dates to the ACC model will lead to errors that tend to propagate throughout the remainder of the growing season. In the United States, the starter model predicted seeding in North Dakota about 10 to 15 days earlier than was reported at the three ITS's, whereas the starter model predicted planting in South Dakota as much as 15 days behind the ground-observed planting dates. The erroneous model-generated planting dates for these two states that were used to begin the ACC model thus introduced errors into the model estimates that did indeed remain through the ripe stage.

The severe drought at the two Hand County, South Dakota, ITS's forced the wheat to mature more rapidly and probably explains the model's lack of response to the real situation. Figure 12 illustrates some of the inadequacies associated with the LACIE use of the model; namely, that the model contains no moisture variable and the temperatures input to the model do not represent true canopy temperatures.

In the U.S. spring wheat region, the model performed better overall than the historical crop calendars at the heading stage, whereas the reverse was true at the soft dough and ripe stages.

In Canada, the model performed well as biases varied from -1.5 to 3.6 days over the development stages being examined; RMSE's ranged from 6.5 to 8.8. However, the historical crop calendars also proved to be fairly reliable, with biases ranging from 1.2 to 6.2 days and RMSE's from 6.2 to 9.2. For Phase II, no significant difference between the model and ITS estimates and the historical and ITS estimates was noted.

In the winter of 1976-77, a joint study was conducted by YES and CAMS personnel to assess the feasibility of using CAMS crop calendar evaluations from Landsat imagery to update and correct the YES-generated ACC's. Phase II imagery containing emergence and soft dough information for the state of Kansas was used for the test area. The results of the study indicated that in Kansas for that particular growing season, use of CAMS start information did not significantly improve the YES-generated calendars (ref. 9).

No ground-truth reports were available in the U.S.S.R. and the P.R.C., but the ACC estimates at sample segment locations were compared to the corresponding growth stage numbers as reported by the analysts on the CAMS evaluation forms. Some minor discrepancies were noted in the far southeastern winter wheat areas of the U.S.S.R., but the ACC model performed satisfactorily for the major

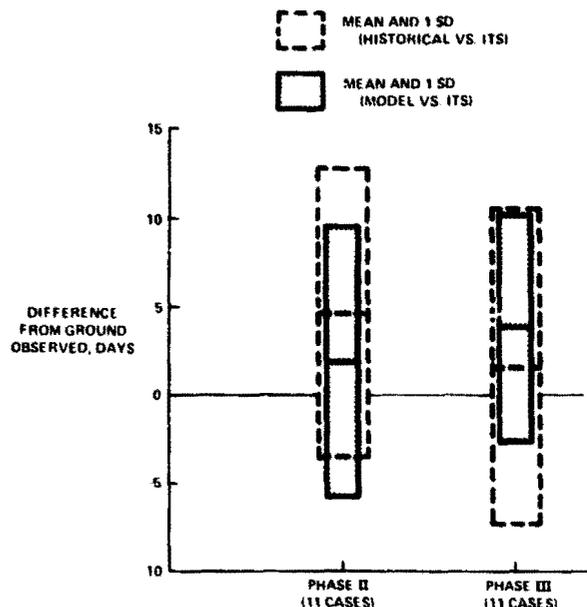


FIGURE 11.—Comparison of Canadian spring wheat ITS observations with CRD historical calendars and LACIE adjustments at soft dough.

wheat areas of the U.S.S.R. In the P.R.C., where wheat fields are often very small, it was extremely difficult to estimate development stage numbers from Landsat imagery. As a result, no major effort was focused on accuracy of the model in the P.R.C

Phase II

The tracking effort of comparing both the ACC model estimates and the historical values with those development stages reported at the ITS's was continued for Phase III. The ground truth consisted of data from 22 ITS's throughout the United States and the same 11 Canadian ITS's used in Phase II. The results of these comparisons for U.S. winter and spring wheat and Canadian spring wheat are given in tables VI to VIII and in figures 3 to 6, 10, and 11.

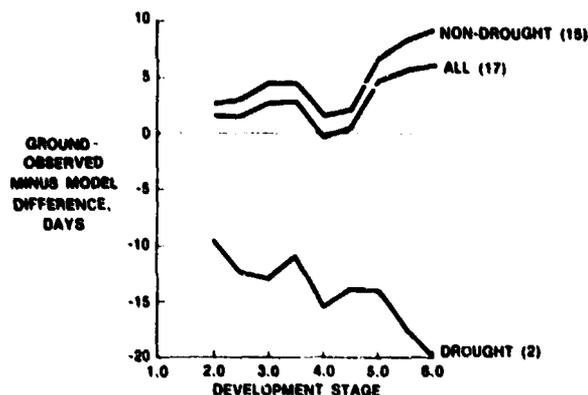


FIGURE 12.—Crop calendar error for Phase II spring wheat intensive test sites.

TABLE VI.—Comparison of LACIE Phase III ACC and Historical CRD Calendars With Observed Development Stages in the 1976 - 77 U.S. Winter Wheat ITS's

ITS (county, state)	ACC vs. ITS, days						Historical vs. ITS, days					
	Jointing		Heading		Soft dough	Ripe	Jointing		Heading		Soft dough	Ripe
	3.0	3.5	4.0	4.5	5.0	6.0	3.0	3.5	4.0	4.5	5.0	6.0
Oldham, Tex.	-4	17	17	9	9	8	-13	19	21	20	16	8
Randall, Tex.	3	7	5	4	8	8	-9	-2	3	11	13	5
Finney, Kans.	4	5	-3	3	8	-5	-23	4	3	2	9	3
Rice, Kans.	-12	0	-5	-14	0	7	-10	-4	-8	-10	-2	-3
Ellis, Kans.	-11	-3	-8	-15	1	-11	-27	-3	-6	-10	0	-4
Saline, Kans.	4	0	-3	-3	6	11	-10	-3	-4	-2	3	4
Morton, Kans.	2	0	1	0	5	8	-19	3	1	3	5	0
Boone, Ind.	10	9	2	0	2	5	-5	1	-4	-10	-13	5
Madison, Ind.	10	6	1	0	8	5	-5	0	-5	-10	-5	-8
Shelby, Ind.	10	-1	-3	-1	-4	2	-6	-8	-8	-9	-16	-11
Bannock, Idaho	15	3	0	-1	8	13	4	7	0	-11	-12	-19
Franklin, Idaho	(a)	(a)	(a)	11	14	(a)	(a)	(a)	8	-1	(a)	(a)
Oneida, Idaho	-11	-7	-7	-7	-5	-3	3	10	6	-2	-12	-23
Whitman (2), Wash.	-5	10	-3	-9	2	7	0	14	4	-8	-10	-16
Hill, Mont. ^b	3	-8	-9	-10	5	10	-13	-13	-23	-23	-9	1
Toole, Mont. ^b	-4	-8	-6	-9	-8	-8	-10	-11	-11	-12	-8	-5
Hand (1), S. Dak. ^b	17	5	-5	0	-6	-5	0	-7	-13	-3	-6	-13
Hand (2), S. Dak. ^b	17	7	0	-4	-3	-3	0	-1	-5	-12	-3	-10
Bias, days	2.8	2.5	-1.5	-2.6	2.8	2.9	-8.4	1.2	-2.9	-4.3	-2.8	-5.1
SD, days	9.6	6.8	6.1	7.2	6.1	7.5	8.8	8.5	9.5	10.2	9.2	9.2
RMSE, days	9.7	7.1	6.1	7.4	6.6	7.6	12.0	8.3	9.6	10.8	9.4	10.3

^aITS data not available

^bAcquired as spring wheat ITS but contained numerous winter wheat fields

TABLE VII.—Comparison of LACIE Phase III ACC and Historical CRD Calendars With Observed Development Stages in the 1977 U.S. Spring Wheat ITS's

ITS (county, state)	ACC vs. ITS, days						Historical vs. ITS, days					
	Jointing		Heading		Soft dough	Ripe	Jointing		Heading		Soft dough	Ripe
	3.0	3.5	4.0	4.5	5.0	6.0 ^a	3.0	3.5	4.0	4.5	5.0	6.0
Hand (1), S. Dak.	-10	-5	-2	-8	1	5	-11	-17	-20	-15	-8	-3
Hand (2), S. Dak.	-10	-8	-2	-3	-3	-3	-10	-19	-19	-14	-11	-10
Burke, N. Dak.	(a)	(a)	(a)	(a)	22	21	(a)	(a)	(a)	(a)	8	2
Williams, N. Dak.	0	5	2	4	12	10	-6	-4	-10	-9	-3	-8
Hill, Mont.	10	12	6	6	15	14	-10	-14	-22	-20	-9	-5
Liberty, Mont.	19	22	19	11	27	34	3	0	-7	-8	11	20
Toole, Mont.	2	0	-1	6	12	17	-6	-11	-17	-8	4	15
Polk, Minn.	-7	-5	-2	6	8	5	-19	-21	-20	-5	0	-3
Bias, days	0.6	3.0	2.9	3.1	11.8	12.9	-8.4	-12.3	-16.4	-11.3	-1.0	1.0
SD, days	10.9	10.8	7.7	6.4	10.0	11.4	6.7	7.8	5.7	5.2	8.2	10.9
RMSE, days	10.1	10.5	7.7	6.7	15.0	16.7	10.4	14.3	17.3	12.3	7.7	10.2

^aITS data not available.

TABLE VIII.—Comparison of LACIE Phase III ACC and Historical CRD Calendars With Observed Development Stages in the 1977 Canadian Spring Wheat ITS's

ITS (town, province)	ACC vs. ITS, days						Historical vs. ITS, days					
	Jointing		Heading		Soft dough	Ripe	Jointing		Heading		Soft dough	Ripe
	3.0	3.5	4.0	4.5	5.0	6.0	3.0	3.5	4.0	4.5	5.0	6.0
Fort Saskatchewan, Alberta	-1	0	-7	-11	4	-19	-9	-6	-6	-6	4	-4
Olds, Alberta	10	7	4	3	14	(a)	0	2	2	1	23	(a)
Lethbridge, Alberta	12	13	10	9	7	0	-2	0	-3	0	2	-10
Dawson Creek, British Columbia	-3	2	-3	-5	-6	4	0	4	5	7	8	5
Stony Mt., Manitoba	6	3	1	2	3	4	-3	-7	-9	-5	-2	-5
Starbuck, Manitoba	4	0	-3	-3	0	5	-4	-8	-12	-10	-6	-8
Altona, Manitoba	3	-1	-8	-9	-6	-6	0	-8	-14	-12	-11	-17
Delisle, Saskatchewan	11	5	0	10	8	10	7	5	2	14	13	14
Swift Current, Saskatchewan	9	5	-4	7	4	0	2	-4	-10	4	3	0
Torquay, Saskatchewan	7	3	-2	-2	1	6	0	-4	-8	-6	-3	-7
Melfort, Saskatchewan	9	9	7	6	12	-7	0	0	0	0	6	-3
Bias, days	6.1	4.2	-0.5	0.6	3.7	-0.3	-0.8	-2.4	-4.8	-1.2	3.4	-3.5
SD, days	4.9	4.2	5.6	7.2	6.4	8.4	4.0	4.8	6.4	7.7	9.3	8.5
RMSE, days	7.7	5.8	5.4	6.9	7.2	8.0	3.9	5.1	7.8	7.4	9.5	8.8

^aITS data not available.

From these results, it appears that increased accuracy was obtained in generating winter wheat estimates using the scalar multipliers. At heading, the average model estimates were less than 2 days behind the values observed at the winter wheat ITS's and the SD was about 6 days. At 14 of the 17 sites, it was noted that the magnitude of the ACC values differed from the ground-observed estimates by less than 1 week. At this same development stage, the corresponding historical calendars averaged about 3 days slower than the ITS's, with an SD of 9. Overall for winter wheat, the magnitude of the model versus ITS biases and SD's differed from the historical versus ITS by 2 days.

In the U.S. spring wheat region, the average ACC estimates were ahead of the ground-truth estimates for the entire development of the plant. The difference between the two estimates was smallest at jointing (0.6 day) and generally increased as the crop progressed toward maturity to a value of almost 13 days at ripe. For the historical versus ITS, the average historical values were approximately 8 days later than the ITS at jointing, regressed further at heading to 16 days behind, and then approached the ground-observed values at soft dough and ripe. At heading, the historical estimates were significantly different from 0 at the 1-percent level, whereas at soft dough, the model's ACC's were significantly different from 0 (at the 5-percent level).

During Phase III, the effects of the extended drought in the northern intermountain and western regions were still being felt in Montana and to a lesser degree in North Dakota. Planting was delayed at numerous fields in these two states, especially at the Liberty, Montana, ITS. The spring wheat starter model did not account for these deferred plantings and thus generated early planting dates for these states. An abundance of rain fell in July after the wheat had headed, and the moisture tended to slow the crop's actual development. The model did not respond to this slower development rate; thus, in Montana and North Dakota, it advanced still further ahead of the ITS values as the crop proceeded toward the ripe stage. At the same time, the combined historical calendars, which had averaged some 16 days behind ground truth at the heading stage, were within a day of the observed value at the ripe stage.

In Canadian spring wheat areas, the average model estimates were ahead of ground observations by 4 to 6 days at jointing and by 4 days at soft dough. There was little difference, on the average, between the observed and the predicted values at heading and

ripe. SD's varied between 4 days at midjointing to 8 days at ripe. The historical calendars also proved to be close to the ground-observed values, with average differences at the various stages ranging from 5 days behind (at heading) to 3 days ahead (at soft dough). Corresponding SD's varied from 4 days at jointing to 9 days at soft dough. At the 5-percent level, the normal calendar at the development stage of heading was significantly different from 0.

Variations in the wheat development observed within the Phase III ITS's were computed according to the following equation:

$$\text{Average SD} = \left[\frac{\sum (n_i - 1) s_i^2}{\sum n_i - 1} \right]^{\frac{1}{2}}$$

where n_i = number of fields within the i th ITS
 s_i = standard deviation of fields within the i th ITS

These SD's were computed at several development stages for the winter wheat ITS's, U.S. spring wheat ITS's, and their Canadian counterparts. These results are summarized in table IX. The deviations were smaller for the winter wheat sites than for the spring wheat sites, thus reflecting the neutralizing effect of dormancy on variations due to early and late plantings. For spring wheat, the variations of development within an ITS because of early and late planted fields generally continue to increase through the crop season. The approximate number of days associated with each stage's deviation is enclosed in parentheses. Thus, at heading, for example, the average SD within winter wheat ITS's was about 6 days, whereas for spring wheat ITS's, it was approximately 9 days.

Test of Applicability to Foreign Areas

Because the density of meteorological input data within a region affects both the reliability and the variability of the model's results, NOAA personnel performed studies to determine the percentage of meteorological stations in foreign areas for which 6-hourly observations were received regularly. It was found that observations from many of the stations

TABLE IX.—1976-77 Phase III Average Standard Deviations of Wheat Development Observed Within ITS's

ITS	Mean development stage ^a			
	Jointing	Heading	Soft dough	Ripe
	3.0	4.0	5.0	6.0
Winter wheat	—	0.24 (6)	0.19 (4)	0.29 (4)
U.S. spring wheat	0.13 (3)	.35 (9)	.22 (5)	.39 (6)
Canadian spring wheat	.28 (6)	.34 (9)	.36 (7)	.58 (9)

^aStandard deviation in units of a stage, approximate number of days associated with stage deviation given in parentheses.

previously selected for use in ACC operations were not received at the National Meteorological Center (NMC), whereas others were acquired only sporadically. NOAA's recommendation to use every available station in the vicinity of the wheat areas that reported regularly was followed for model operations.

As noted early in this paper, Robertson's crop calendar is a spring wheat model developed from Canadian data. In LACIE operations, the model was extended to estimate the development of winter wheat in various parts of the world. Feyerherm multipliers were computed for all winter wheat crop calendar stations in the United States, the U.S.S.R., the P.R.C., India, Argentina, Australia, and Brazil (ref. 10).

Studies were conducted to assess the applicability of the model and the scalar multipliers for the winter wheat grown in the southern latitudes, where the dwarf/semidwarf varieties are grown. Most of the dwarf/semidwarf varieties grown in the southern latitudes are not sensitive to day length, require no vernalization period, and may require warmer temperatures. The majority of the wheat grown in India is actually of a high-yield dwarf variety, and, except in far northern Kashmir, there is essentially no dormancy period. From the tests on India data, it was observed that the crop calendar did not advance because of the day-length factor in the jointing to heading equation. In India, winter wheat day lengths typically dip below 11 hours, which is the threshold value for crop advancement in this stage of the model. It was found that use of the multipliers in In-

dia produced a distorted crop calendar and that the calendar was not compressed enough to reflect an estimate of actual growing conditions. The model's estimates without the multipliers were better but still were totally inadequate. The results of the study clearly indicated that the ACC model was not valid for use in India (refs. 11 and 12).

A similar situation existed in the three Southern Hemisphere countries under study, where the wheat grown has no dormancy requirements for reproductive maturity. In essence, a spring-type wheat is grown during the winter months, with 150- to 200-day growing seasons. The ACC model was run for the 1977 crop year for Argentina, Australia, and Brazil without multipliers. No ground-truth data were available, but feedback from the CAMS analysts on their development stage estimates as determined from the Landsat imagery indicated that the ACC estimates for Australia were not reliable.

Very spotty and limited ground-truth data existed for the U.S.S.R. Feedback from the CAMS analysts was the primary means of determining the reliability of the model's estimates. Scalar multipliers were used in determining Phase III winter wheat predictions for the U.S.S.R. For the most part, no major discrepancies were observed between the model's results and the analyst's estimates. In spring wheat areas, as long as the planting or starter dates were realistic, the ACC performance was good. Problems did exist in the U.S.S.R.'s New Lands area east of the Ural Mountains during Phase III, but the discrepancy resulted from erroneous start dates generated by the spring wheat starter model.

CONCLUSIONS AND RECOMMENDATIONS

First, it should be restated that the ground-truth data set available for evaluating the accuracy of the ACC was very limited. At its peak, the network included reports from only 18 U.S. winter wheat sites, 9 U.S. spring wheat sites, and 11 Canadian spring wheat sites. Even with this sparse data set, certain trends were noticed during evaluation of the model's accuracy and certain conclusions were made.

For winter wheat, for each of the three growing seasons under study, the average model heading date varied between 2 and 7 days behind the date observed. A comparison of these results with those for soft dough shows that the model ran fast between the stages of heading and soft dough, where the average

model date occurred 1 to 3 days before the ground-observed date.

It was also found that, for winter wheat, the model's performance was best for Phase III (1976-77), the first time that the Feyerherm scalar multipliers were incorporated into the operational model.

It is also apparent that the model performed well for the Canadian spring wheat regions. This was expected since the model's coefficients were derived using phenological data for the Marquis variety of spring wheat grown in Canada.

The limited assessment of the model estimates at heading, soft dough, and ripe in the ITS's showed that, overall, the model's estimates provided more accurate information than that available from the historical normals and generally met the accuracy goal of being within 5 to 7 days of the ground observations.

The validity of planting dates and the lack of a moisture term in the model were observed to be important factors in the model's accuracy. Droughts in South Dakota in 1976 and in North Dakota and Montana early in the 1977 growing season were not reflected in the model's estimates. Also, some erroneous planting dates were generated by the spring wheat starter model during the growing seasons, and these errors tended to propagate throughout the seasons. From these results, it is concluded that an improvement in crop calendar capability is needed in the area of model startup dates. Some type of moisture term also needs to be incorporated into the model to account for periods of moisture excess or stress. A recent attempt has been made to incorporate a moisture variable in the Robertson model form. The only addition to the data was daily precipitation statistics. The results to date are encouraging. Details of both these efforts are given in the paper by Seeley et al. entitled "Prediction of Wheat Phenological Development: A State-of-the-Art Review."

Models that will adequately predict development stages for the dwarf and semidwarf varieties grown in the southern latitudes also need to be developed. As the technology progresses from a single-crop program, such as LACIE, to a multicrop effort, the need for accurate crop calendars to aid in the identification and separation of crops will be greatly increased. However, for some of the crops under consideration, reliable development stage data do not exist over enough stages and geographical areas to allow development of a Robertson-type model. Therefore, crop calendar models using remotely sensed data

from Landsat and meteorological satellites should be considered as prime data input and further developed and refined. For many of the crops, yield models will require as inputs the biological stage of development of the plant. This will be particularly true for the early warning techniques.

Early indications are that the remotely acquired data provide a new source of information on crop development independent of the meteorological inputs used now. This should be pursued not as a separate approach but as a combined approach (spectral with meteorological) to crop development estimation.

REFERENCES

1. Feyerherm, A. M.: Response of Winter and Spring Wheat Grain Yields to Meteorological Variation. Final Report, Contract NAS 9-14282, Kansas State University, Feb. 1977.
2. Stuff, R. G.: Preliminary Evaluation of Phenological Models for Spring Wheat. LEC-6134, Lockheed Electronics Co., Houston, Texas, June 1975.
3. Robertson, G. W.: A Biometeorological Time Scale for a Cereal Crop Involving Day and Night Temperatures and Photoperiod. Intern. J. Biometeorol., vol. 12, no. 3, 1968, pp. 191-223.
4. Wilcox, D. D. III; Champagne, G. J.; Baskett, R. L.; and Woolley, S. K.: A FORTRAN Implementation of the Robertson Phenological Model. LEC-5974, Lockheed Electronics Co., Houston, Texas, Apr. 1975.
5. Baskett, R. L.; Hildreth, W. W.; Wilcox, D. D.; and Woolley, S. K.: Large Area Crop Inventory Experiment (LACIE) Phase II Winter Wheat End-of-Dormancy Restart Model. LEC-8268, Lockheed Electronics Co., Houston, Tex., Apr. 1976.
6. Feyerherm, A. M.: Planting Date and Wheat Yield Models. Final Report for Period Feb. 1975 to Mar. 1977, Contract NAS 9-14533, Kansas State University, Sept. 1977.
7. Wilcox, D. D.: System Description of the LACIE/YES Gridded Crop Calendar Report Writer. LEC-10004 (Rev. A by M. L. Shelton), Lockheed Electronics Co., Houston, Tex., July 1977.
8. Wilcox, D. D.: Operating Procedures for the LACIE/YES Gridded Crop Calendar Report Writer. LEC-9955 (Rev. A by M. L. Shelton), Lockheed Electronics Co., Houston, Tex., July 1977.
9. Woolley, S. K.: Test Study To Verify the Assumption That CAMS Feedback Can Improve YES Generated Adjusted Crop Calendars. LEC Memo. Ref. 642-2296, Lockheed Electronics Co., Houston, Tex., Mar. 21, 1977.

10. Easley, J. W.: Scalar Multipliers for the Southern Hemisphere. LEC Memo., Ref. 642-2034, Lockheed Electronics Co., Houston, Tex., Nov. 16, 1976.
11. Easley, J. W.: Study of the Applicability of the Feyerherm Multiplier in India. LEC Memo. Ref. 642-944, Lockheed Electronics Co., Houston, Tex., Sept. 10, 1976.
12. Easley, J. W.: Scalar Multipliers for India. LEC Memo., Ref. 642-944, Lockheed Electronics Co., Houston, Tex., Sept. 24, 1976.

ACKNOWLEDGMENTS

The material presented in this paper represents the efforts of an aggressive team working the problem of crop calendars in LACIE. Contributing to the results but not listed as authors are J. W. Easley, R. L. Baskett, W. W. Hildreth, D. D. Wilcox, J. R. Zumbro, and D. E. Phinney, all of Lockheed Electronics Company in Houston.

Economic Evaluation: Concepts, Selected Studies, System Costs, and a Proposed Program

Frank H. Osterhoudt^a

INTRODUCTION

The LACIE has been a highly technical effort to estimate wheat production in major producing regions of the world by using satellite-derived data in combination with meteorological and historical data. The Accuracy Assessment Team, the Research, Test, and Evaluation Group, and the Information Evaluation Group within LACIE evaluated LACIE from a technical standpoint, often recommending new or modified technologies to improve overall performance. The specific questions addressed were "How good are the estimates?" and "What technology needs to be improved?" However, a crucial question has not been adequately addressed: "Are there sufficient benefits to justify implementation of a satellite-based crop information system?" In order to make this judgment, an economic evaluation is needed.

During the course of LACIE, an interagency Economic Evaluation Planning Team was established to develop and monitor an economic evaluation of LACIE. Drawing upon the team's experience and other benefit estimation studies, this paper first discusses a concept for valuing crop information, considering the more usual approaches, a recommended integrated approach, and problems of implementation. This is followed by a review of what has been done in the economic evaluation of LACIE-type information. The various studies of benefits are reviewed, and the costs of the existing and proposed systems are considered. Finally, a method and approach proposed for further studies is reported.

CONCEPT FOR VALUING INFORMATION

The prime benefit derived from improved infor-

mation is the ability to make improved decisions—those which affect buying, selling, investing, or setting government economic programs—with increased accuracy, in a more timely manner, or with more certainty.

Some decisions affected by crop information lie within the commodity markets and others lie outside. Market-related decisions are made by the domestic or international grain trade, U.S. producers and consumers, and foreign producers and consumers. Nonmarket decisions are reflected in government policy, the administration of government programs, and agreements with other nations. Some decisionmakers use crop production information directly, such as a federal official assessing the need to restrict wheat production during the next crop year. Some decisions and impacts are indirect; for example, consumers observe an increase in the price of bread when the price of wheat rises. Decisionmakers who use crop information also react to information concerning many other factors, including the state of the economy, money supply, trade policies, and pending legislation. All these factors are sources of uncertainty. Improved decisionmaking may require that one or all types of information be improved.

Crop production data have a number of properties that could be affected by a system incorporating satellite data collection and analysis. Accuracy is important, but it is conditioned by (1) when data are available in the crop season, (2) geographic location and detail, (3) comprehensiveness and continuity, and (4) the reliability of the estimates, including perceived objectivity.

Methods of Valuing Information

The methods of valuing information seem to fall into two general categories. The "global modeling approach" takes the form of simulations or

^aUSDA, Washington, D.C.

econometric models that estimate benefits derived from information by relevant sectors of the economy. Hayami and Peterson (refs. 1 and 2) used the measurement of consumer and producer surplus to assess the value of crop information. This has been elaborated by others, most recently by ECON, Inc. (refs. 3 to 7). Another global modeling approach uses decision analysis or decision theory to estimate the impact of information on the decision process and to assign a value to that impact. Decision analysis was developed by theorists such as Marschak (refs. 8 to 11), Howard (refs. 12 and 13), and Agnew (refs. 14 and 15). Although decision analysis usually reflects the value of information to an individual firm rather than to society, this technique has been used to estimate the aggregate value of information (refs. 15 and 16).

The second major method of valuing information is quite pragmatic, studying specific user groups. The Panel on Methodology for Statistical Priorities proposed this approach to estimate benefits attributable to data packages and programs (ref. 17). Savage (ref. 18) has summarized and criticized the Panel's proposal. Eisgruber¹ has warned that surveys of users, which frequently are a part of user studies, are notorious for their nearsightedness. Social scientists currently supporting the user approach include Hoos (refs. 19 and 20), Duncan (ref. 21), and Sharp (ref. 22). The pragmatic user-oriented approach may quantify estimates of benefits to specific user groups through the methods of the global modelers, but it recognizes the impossibility or impracticality of quantifying benefits to other user groups, such as researchers. For these, expert opinion is used to make qualitative assessments.

Integrated Methodology

Miller (refs. 23 and 24) has recommended an integrated methodology with one framework for market users of crop information and a second for nonmarket users (fig. 1). He emphasizes empirical estimates of information and decision models, especially for market users of crop information. For market users, the methodology comprises four models or components.

¹Personal correspondence from Ludwig Eisgruber, Oregon State University, to Forrest G. Hall, NASA Johnson Space Center, Jan. 20, 1978.

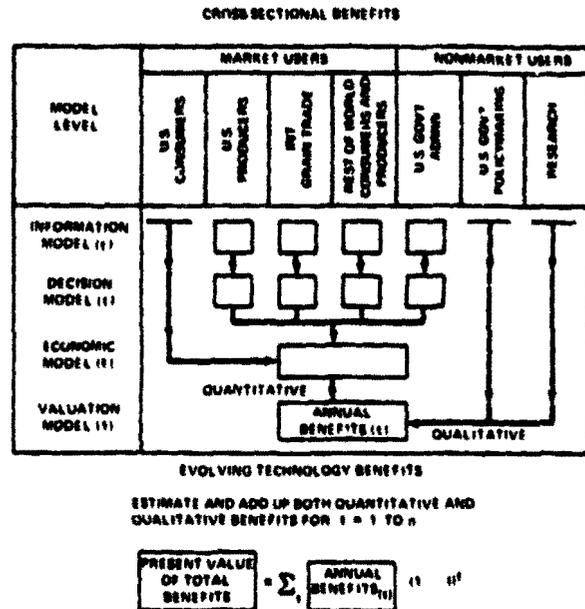


FIGURE 1.—Diagram of possible information benefits. (Distributional and structural benefits are not expressed in these summaries. When generated elsewhere, they may be accommodated herein.)

1. An information model relates the subset of information used in the decision process to the total set of information available and accommodates the fact that decisionmakers do not fully believe in or act upon specific forecast information. Other factors, such as alternative sources of information with their perceived costs and reliability, are also considered.

2. A decision model describes the decisionmaking process, considering both the information model and a system of economic rewards.

3. A traditional economic model, based on the assumption of profit maximization, represents behavioral characteristics involved in a specific kind of decision process.

4. An information valuation model estimates benefits in either of two ways. The first estimation procedure utilizes net social benefits from consumer and producer surplus. The alternative technique estimates the size of changes in user incomes, a methodology frequently used as a basis for benefit-cost ratios.

For the second group, nonmarket users, Miller proposes a basically qualitative assessment. The first subgroup of nonmarket users is concerned with the policymaking functions of the Federal Government and may represent the most important of all uses (refs. 17 and 25 to 28). Examples of frequently ad-

dressed policy issues are commodity trade negotiations and grain embargoes. Even if the presidential and executive decisionmaking process is unavailable, research into the uses of information for such decisions would provide insight for evaluation.

A second subgroup of nonmarket users is concerned with the administration of legislated federal programs. Estimation of the value of information for such activity is more feasible than for policymaking because budgets have been assigned and administrative latitude is relatively known.

The final major nonmarket user of information is the research community. Since the value of information as an input to research is highly correlated with the ultimate value of the research, the value of neither is readily predictable. The value of either information or research is ascribable only insofar as it affects buying, selling, investing, or human life.

As the benefits of global information generated from LACIE-developed technology are evaluated, elements of technology assessment emerge. The value of LACIE may be heavily influenced by the learning experience and the development of an advanced technology in the future. Thus, the assessment of the value of information should be generated for differing levels of technology at future time periods, with the present value assessed using an appropriate discounting procedure (fig. 1).

Problems of Implementation

Since there is no generally accepted methodology for estimating the value of information, appropriate methodology must be developed and tested. The relationships between information, decisions, and market structure are crucial to the development of this methodology (ref. 29). Future decisions concerning implementation of a satellite-based crop information system require better assessments of the need for information, which a satellite-based system can provide. In addition, expected investment and operational costs must be estimated. Economic analysis of the value of information can assist in making these decisions. However, the usefulness of the analyses is influenced by practical considerations such as budget restrictions and the "client's" acceptance of a new information source.

Thus, a major problem is measuring the quality of information and relating that quality to user requirements. There are several attributes of high-quality information. Authors have listed such factors as objec-

tivity, accuracy, reliability, adequacy, continuity, comprehensiveness, geographic detail, timeliness, availability, relevance, and believability (refs. 17 and 30). Some of these attributes are difficult or impossible to quantify, yet improvement in one characteristic could result in an increase in value. Because of the difficulty in quantification, most studies have limited their evaluations to measuring those characteristics of information quality that have the highest potential for generating economic value.

Other problems encountered in estimating the value of information include the following.

1. Extending the results of an analysis based on current system performance to a system using untested technology.

2. Anticipating political and economic conditions that change requirements for information and its value.

3. Assessing the effect of production changes on prices and market receipts, which in turn may make timely and accurate crop forecasts more valuable.

4. Extending results of analyses of information value that assume competitive conditions to situations dominated by government and large commercial organizations.

BENEFITS AND APPLICATIONS OF IMPROVED CROP CONDITION INFORMATION

The U.S. wheat crop is so large and the associated transactions are so great that modest improvements in the marketing system could have large aggregate benefits. For example, an increase of one cent per bushel, resulting from a price increase or an efficiency improvement, had an aggregate value to U.S. farmers of \$21 million in 1975 and nearly \$18 million in 1974. Exports in 1975 were 1.2 billion bushels; thus, one cent per bushel could have amounted to a difference in returns to the United States from the rest of the world of \$12 million. In 1975, the cost of moving wheat from the farms to the docks for export was \$873 million; small per-unit efficiencies in storing and shipping could have been large in total. Are there ways for satellite-based improvements in foreign wheat production information to affect decisions concerning planting, harvesting, buying, selling, or investing that would exceed the cost of the improvements? Although no final analyses have been made, there are several studies which add to our understanding of the question.

ECON Studies

Since 1972, ECON, Inc., of Princeton, New Jersey, has developed benefit-cost or economic evaluations of Earth resources satellite systems for NASA. In developing their estimates, ECON has made two key assumptions: (1) performance of at-harvest crop estimates is within 10 percent of true production 9 years of 10—the LACIE 90/90 criterion, and (2) the United States would be operating within a perfectly competitive international market. Under these assumptions, ECON has estimated that benefits to the United States of improved foreign production information would be in the neighborhood of \$300 million annually. They estimate that about \$240 million would be from improved wheat forecasts.

One of the ECON approaches was essentially descriptive or positive, used econometric techniques, and was sometimes known as ECON's production model (ref. 31). A second approach was partially normative and has been referred to as ECON's distribution model (refs. 4 and 32). Their present work is an extension of the distribution benefits model and is referred to as the integrated model (ref. 4). The integrated model is currently used to estimate both production and distribution benefits.

The integrated model, a stochastic dynamic decision model,² has its roots in the work of Hayami and Peterson (refs. 1 and 2). The model assumes perfect competition and imperfect foresight. It uses dynamic programming to solve for production, consumption, and trade-with-uncertain-information that would maximize producer and consumer surplus. The global activity of crop production and distribution is treated as a process, in which rates of planting, consumption, and exports are decision control variables. Estimates of supply in the two producing units, the United States and the rest of the world, are resultant state variables. The value of improved information is obtained by comparing producer and consumer surplus resulting from alternative information systems. ECON assumed that all participants operated in a perfectly competitive world wheat market and used production forecasts equivalent to those published by the Commonwealth Secretariat of the United Kingdom. A second analysis assumed that participants shifted completely to the more accurate

information represented by the LACIE 90/90 criterion.

According to ECON, the principal benefit to the United States from improved foreign crop information would be from selling larger quantities to the rest of the world in those months in which prices were higher. ECON's model maximizes returns to the global society by optimizing stocks, U.S. exports, and U.S. production. Inventories in the United States increase with improved information, while inventories in the rest of the world decrease.² Preliminary results show that at the end of the model's base year, buffer stocks or inventories would average 2.1 million metric tons (MMT) in the United States, 31.8 MMT in the rest of the world, and 1.2 MMT in transit. Under the improved information system, this would change to 13.3 MMT in the United States, 15.6 MMT in the rest of the world, and 1.0 MMT in transit at year's end. Note that total buffer stocks would decrease if improved information were available. Total annual exports would remain the same, but export revenues would be much higher under the improved system because export sales would occur at higher price levels. The higher export revenues would be somewhat offset by increased storage costs, including interest costs.

ECON shows that trade benefits from improved crop information sift down eventually to producers and consumers. Some benefits would result from adjustment in production. In addition to benefits to the United States, benefits to the rest of the world result from decreased inventory costs roughly equal to 10 percent of the U.S. benefits.

Futures Group Study

In the spring of 1976, the U.S. Department of Agriculture (USDA) contracted with the Futures Group, Inc., of Glastonbury, Connecticut, to study the use and usefulness of improved wheat information to the USDA (refs. 33 and 34). USDA officials were interviewed to determine how improvements in foreign wheat production information would affect program decisions. The use of improved information for broad policy direction or the need to provide improved information to the public were not pursued. The 30 respondents consisted of nearly equal numbers of program managers, analysts, and those performing both functions. Major findings of the study are discussed in the following paragraphs.

The principal uses of improved foreign crop infor-

²John Andrews, "A Stochastic Dynamic Decision Model of the Value of Improved Public Crop Information (Wheat)." Unpublished report, ECON, Inc., Princeton, New Jersey, 1977.

mation by the USDA would be in the management of export programs, especially PL-480, and in negotiations and adjustments within bilateral agreements. Improved information would also be used to support decisions concerning international wheat reserves or embargoes.

The need for wheat production data varies between programs. For some programs, current-season data cannot be obtained early enough to significantly affect program decisions. For other programs, improved wheat information would be important provided the improvements were significant. Under some circumstances, program decisions have been delayed until better crop data became available. Examples of commodity program decisions that could be affected by more up-to-date global information are summarized in table I.

Under supply and demand conditions current at the time of the study, programs relating to wheat production controls or income supports were not being applied. Therefore, foreign wheat production information was not needed to support most USDA program decisions.

The possibility of large worldwide wheat supplies and an associated weak export demand or of small supplies and a strong export demand could significantly affect the need for better information about world wheat production. The occurrence of short wheat supplies would be more disruptive and make good early production information especially useful.

Many variables, other than foreign production, are involved in projecting U.S. exports. These variables include prices, political attitudes and actions of foreign governments, grain carryover, livestock numbers and feed use, availability of foreign exchange, and transportation problems. For factors other than production, information usually can be improved only indirectly, if at all, with satellite-based data.

Department officials expressed the desire to have more accurate and timely information. They thought that if improved information were fed into the marketing system, the operation of the market would be improved. Furthermore, they anticipated the possible recurring need for agricultural price or supply programs.

TABLE I.—*Dates and Time Flexibility for USDA Wheat Programs Decisions*

<i>Decision</i>	<i>Legislated or required decision date</i>	<i>Is decision subject to change with new information?</i>
1. Wheat national program acreage (NPA)	By August 15 of previous crop year ^a	NPA may be adjusted based on new information
2. "Set-aside" acreage	By August 1 of previous crop year ^a	By clear precedent, could reduce set-aside required but not increase it
3. Price support level	No legal date; best done by planting time but seldom announced that early	Clear precedent not to reduce amount
4. Disposal of Commodity Credit Corporation (CCC) stocks	No legal date; cannot sell CCC inventory at less than 150 percent of current loan rate when reserve is in effect	Limited by desire to minimize effect on market
5. CCC aid for export sales	No legal date; made when there is U.S. buildup of price-depressing stocks	Full flexibility
6. Wheat available for PL-480	No legal date; announce plan by October 1	Full flexibility
7. Wheat reserve	No legal date; producer-held wheat reserves of 300 to 700 million bushels are required	Limited by policy created

^aExample: For the 1979 wheat crop, which is planted in the fall of 1978 and spring of 1979, the decision must be announced by August 15, 1978.

Overview of the U.S. Wheat Industry

An overview of the U.S. wheat industry made in mid-1977 by the Economics, Statistics, and Cooperatives Service (ESCS) provides some necessary background information concerning how crop information is used. These uses are the basis for its value (ref. 35). In itself, the study provides no expression of the value of information.

The study traces the flow of wheat between various sectors of the wheat industry. Accompanying the physical flows are decision flows. Decision flows may travel parallel to the physical flows, or separately, as in the case of the futures market. Both physical and decision flows are summarized in figure 2; the figure also indicates the number of participant locations in each sector. Researchers identified two sectors as key wheat-market information users. The first is the large integrated export firms. The second is the terminal markets. Information concerning a third sector, the Federal Government with its regulations and policies, is also critical to wheat decisions.

Timeliness of information was identified as of major importance for decisions based on wheat information. Decisions affected by wheat information include how much to plant, sell, feed, process, and consume; when to buy, sell, ship, and store; where to buy, sell, load, or ship; what to plant, what quality of product to buy or sell, and what transportation to use; how much to pay for land, where to build, and the size of facilities for storage and processing (ref. 35, p. 34). Deficiencies in information may lead to delay, waste, and other inefficiencies in the wheat industry and the general economy.

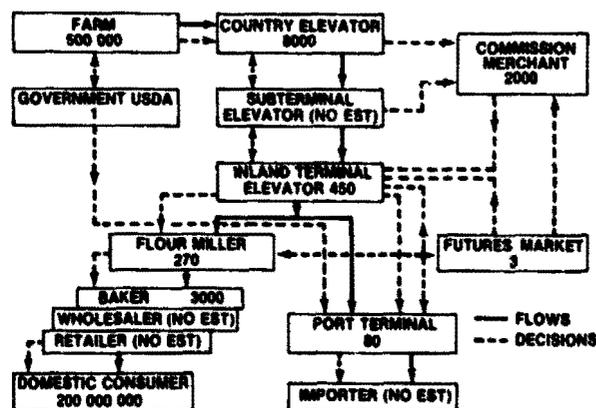


FIGURE 2.—Major wheat flow by sector, flow of merchandising decisions, and number of participant locations by sector.

Evaluation of the Oilseed and Products Program

During 1976, the information functions of the Foreign Agricultural Service (FAS) oilseed and products program were evaluated (ref. 36). Although conducted totally apart from LACIE, this study has clear implications concerning the value of improved wheat information. The actual extent of those implications must be judged in the light of the similarities of information use and needs of both oilseed and food grain information users and according to the nature of what is being compared.

A mail survey was made of the subscribers to the FAS oilseeds and products publications. Subscribers from government and international organizations were not included and are not considered below. The original population of about 1800 subscribers was categorized into three subpopulations: (1) private trade; (2) executives of firms; and (3) media, farm and trade associations, and educational institutions. About 42 percent of the subscribers were sent questionnaires; 69 percent of those returned usable data.

The composition of the subscriber list gives a rough index of where the interest lies for crop production information. Subscribers to the oilseeds and products publications may be categorized as follows.

Subscriber	Percent
Trade/broker	19
Manufacturing	16
Processing	10
Import/export	9
Consulting service	8
Media	8
Educational institutions	6
Farming	4
Banking/finance	3
Trade associations	3
Farm associations	2
Transportation	2
Storage/handling	1
Other	8

Two findings of the oilseed and products study are of particular interest when evaluating the need for wheat information. Crop production information was clearly the top-priority need of respondents. Subscribers were asked to designate their 3 highest priority information needs from a list of 11 topics. Seventy-nine percent of the private trade group ranked crop production as their highest priority information need, and 49 percent ranked information

on consumption as their second highest priority. Executives designated information on crop production as their highest priority (65 percent) and information on exports and imports as their second priority (41 percent).

Timeliness was the attribute of FAS information which received the most criticism. Although 73 percent of the private trade audience ranked all other attributes either good or excellent, 45 percent ranked timeliness as fair or poor. Among the executive group, all information attributes except timeliness were ranked good or excellent by 82 percent of the respondents. Timeliness was ranked fair or poor by 58 percent. It should be noted that the question of timeliness concerned all information from the program and did not distinguish crop production information separately.

The subgroup composed of the media, associations, and educational institutions acts as intermediaries in information transfer. They would not need foreign crop information for decisions in their own organizations but would use it in their reports to commodity decisionmakers. Sixty-one percent of this subgroup placed production as their most important information need; 34 percent classified timeliness of information in available reports as fair or good.

COSTS OF CROP INFORMATION SYSTEMS

The big question concerning a crop information system resting on LACIE-developed technology must be whether the benefits outweigh the costs sufficiently to warrant further investigation or use of the technology in an operational mode.

Two sets of cost estimates have been developed. Estimates of a satellite-based system were made in 1976 and updated in 1977. To give perspective, costs of the current system were also estimated in 1977.

Considerable caution must be exercised when considering these figures. The two are not directly comparable. The present USDA system primarily assembles, weighs, and disseminates information developed and paid for by others. On the other hand, the satellite-based system would be an entirely new data source that would provide additional information to the current system. The products of a satellite-based system would have different geographic comparability, different statistical properties, different timeliness features, perhaps different believability, and perhaps different uses. Thus, direct

comparisons of the costs of the two systems without simultaneous comparisons of product quality and associated benefits are likely to be misleading. An analysis of the uses and benefits of the present system and of a satellite-based system has yet to be made.

Costs of Satellite-Based Systems

The USDA/LACIE has developed cost estimates for satellite-based approaches that could provide better and faster information on important world crops. Two alternative systems were considered—one for a single crop and one for several crops. Each would produce repetitive area, yield, and production forecasts throughout the season in countries with 95 percent or more of total production; either could provide periodic updates for areas of current critical interest. Projections of the total USDA capital investment required range from about \$10 million for a single-crop system to about \$29 million for coverage of eight major crops (table II). These estimates are

TABLE II.—Projected Costs of USDA Remote-Sensing-Based Crop Assessment System

[Millions of dollars^a]

Item	Range of costs	
	One crop	Multicrop
Investments		
Hardware	5.6	19.0
Software	2.3	7.4
Data base	.1	.1
Conversion	.03	.03
Relocation expenses	1.0	2.0
Other	.8	.9
Total (cumulative)	8.9 to 10.8	19.4 to 30.0
Operations		
Personnel	2.4	4.6
Administrative	.7	1.3
ADP services	.5	1.2
Facilities	.2	.4
Research and development	.3	.4
Support services	.6	.7
Total (yearly average)	4.3 to 5.1	6.0 to 9.4

^aAll costs based on constant (1977) dollars with the exception of salaries, which are increased 5 percent annually in accord with past trend, cf. OMB Circular No. A-94, revised Mar. 27, 1972.

cumulative over 10 years. Estimated annual operating costs range from about \$5 million to about \$9 million. Other systems could be defined, with each exhibiting associated differences in output, costs, and benefits.

The cost estimates were made using a model that separates costs into capital investments and operating costs. The costs of LACIE research and development are not included; however, USDA costs associated with the application, development, and test phases following LACIE are included. The costs of satellite data collection and a ground processing system are included only as an annual payment for Landsat products.

Costs of the Present System

The present USDA foreign crop information system reports on crop production, trade, stocks, consumption, governmental policies in 110 countries, and prices (ref. 37). Total costs for FY 1977 for this system were approximately \$19 900 000 (table III). Of that amount, FAS and ESCS production estimates for all crops in the seven LACIE countries accounted for slightly less than \$700 000. Wheat estimates alone cost approximately \$165 000.

The principal sources of foreign crop information in USDA are the 98 agricultural attaches and assistant attaches assigned to foreign countries by FAS. Sixty-one percent of the cost of developing wheat production estimates for the seven LACIE countries

is derived from the FAS agricultural attache program. The remainder is divided among different FAS and ESCS analyst units in Washington and wheat team trips to the U.S.S.R.

Primary responsibility for estimating foreign crop production lies with the FAS. However, ESCS analysts play a major role in developing estimates for the U.S.S.R. and the People's Republic of China. Extra effort is expended by the USDA to estimate crop production in those countries that do not release their information as freely as others.

The FAS and the ESCS made some very tentative cost projections of the present USDA foreign crop information system. They estimated that the costs of crop forecasts for the seven LACIE countries under the current system would increase about \$83 000 by 1981 for all crops, of which \$17 000 would be for wheat. By 1986, additional increases of \$24 000 for all crops (\$3000 for wheat) were anticipated. Total costs of the current system for the seven countries for the 1977-86 period were expected to be \$8.9 million for all crops, of which \$2.1 million would be for wheat (table IV).

FUTURE PROGRAM OF ECONOMIC EVALUATION

An approach for economic evaluation has been recommended. The approach is pragmatic, is oriented toward information users, and utilizes

TABLE III.—Cost of Current USDA Foreign Crop Production Estimates^a

USDA division	Total budget ^b for all countries	Production estimates for 7 LACIE countries	
		Major crops	Wheat only
FAS attache program	\$11 811 000	\$503 400	\$100 400
FAS Washington analysis ^c	4 797 000	95 900	18 700
ESCS-FDCD ^d analysis	3 137 000	55 300	12 300
Other ^e	159 000	35 000	33 900
Total	\$19 904 000	\$689 600	\$165 300

^aCost estimates include overhead.

^bTotal cost of USDA foreign crop information system, which includes considerably more than production estimation; see text.

^cForeign Commodity Analysis program area.

^dEconomics Statistics, and Cooperatives Service — Foreign Demand and Competition Division.

^eFor spring wheat team and winter wheat team trips to the U.S.S.R., plus costs of Task Force on U.S.S.R. Grain Situation not already counted above. Total budget is for U.S.-U.S.S.R. Secretariat, which paid most costs of the wheat teams but none of the Task Force.

^f\$32 700 was for wheat teams traveling to the U.S.S.R.

TABLE IV.—Projected Costs of Inputs of Current USDA Foreign Crop Production Estimation, 7 Foreign LACIE Countries, FY 1977-86

(Thousands of dollars^a)

	1977	1981	1986	1977-86
Major crops				
FAS attache input	\$503.4	\$629.5	\$770.1	\$6382.6
FAS Washington analysis	95.9	111.4	135.5	1144.0
ESCS-FDCD input	55.3	100.0	122.0	950.6
Other	35.0	37.5	41.5	380.4
Total	\$689.6	\$878.4	\$1069.1	\$8857.6
Wheat only				
FAS attache input	\$100.4	\$125.8	\$152.2	\$1270.3
FAS Washington analysis	18.7	21.7	26.4	222.9
ESCS-FDCD input	12.3	22.3	27.1	211.4
Other	33.9	36.2	39.9	366.7
Total	\$165.3	\$206.0	\$245.6	\$2071.3

^aAll costs are based on constant (1977) dollars with the exception of salaries, which are increased 5 percent annually in accord with past trend, cf. OMB Circular No. A-94, revised Mar. 27, 1972.

proved economic methodology whenever possible (ref. 38). Five tasks have been specified to cover the key economic questions concerning the potential value of satellite-developed estimates. The LACIE experience in terms of expected performance was considered when tasks were formulated. The objectives of these five tasks are as follows.

1. Appraise the usefulness of improved global wheat production information to major user groups.

2. Refine and extend available models to develop quantitative estimates of the expected value to the United States from improved wheat production estimates.

3. Evaluate the relationship between the structure of the international grain trade and the existence of different levels of public foreign crop information.

4. Develop and quantify, where possible, the relationship between evolving technology and the quality of information derived from its application; especially examine how the planned and expected improvements in the Landsat observing system and LACIE-developed methodology will improve the quality of wheat production information.

5. Update cost projections for a USDA crop forecasting and condition assessment system based on LACIE-developed techniques.

As mentioned earlier, two general schools of thought seem to prevail in the economic evaluation of crop information: a global modeling approach and

a pragmatic user-oriented approach. Elements of each have been selected for the recommended program. Task 1 is designed to define the problem setting by analyzing individual user groups. Benefits would be assessed and measured without the necessarily restrictive assumptions of quantitative modeling. Both market and nonmarket users would be considered. Task 2 would quantitatively assess the benefits of improved information on foreign crop production estimates. No new, integrated theoretical modeling would be attempted; instead, known available models would be adapted to estimate benefits. The results of Tasks 1 and 2 will complement each other. It is recognized that the resources required to fully accomplish the five tasks are quite high. Tasks 1 and 2 would be given first priority.

Task 3 is concerned with structural impacts of information. The models available to estimate the value of information systems assume a perfectly competitive market structure, but the markets to be assessed appear oligopolistic (i.e., contain few actors). Furthermore, the effect of improved information may be to alter the distribution of income among countries or groups. Some of this effect is exacerbated by an oligopolistic market.

Task 4 recognizes that the technology to develop information is not static and that the effect of evolving technology on the information produced is not a simple correlation. This task directly considers the

role of an evolving technology in improving the quality of information. The answers to the questions addressed by Tasks 3 and 4 should expand the interpretation of Tasks 1 and 2.

Finally, the benefits of information derived from satellite-based crop estimates must be weighed against the cost of their production. Task 5 provides for updating the costs of a satellite-based global crop forecasting and crop condition assessment system.

SUMMARY AND CONCLUSIONS

Decisions need to be made concerning the extent to which USDA should incorporate satellite-based crop estimation techniques into their global crop information system. The only current study to estimate total benefits to the United States from satellite-based foreign wheat forecasts suggests benefits to be about \$240 million annually. Costs of a satellite-assisted system are estimated to include investments of about \$10 to \$29 million cumulative over a 10-year period, with an annual operating cost of about \$5 to \$9 million. This system would be a new source of information and would provide information not now available.

The present USDA foreign crop information system is essentially a crop intelligence system that covers about 110 countries. It generates some new data, but primarily it assembles, analyzes, and disseminates preexisting information. In addition to reporting crop production, the current system generates information on trade, stocks, consumption, policies, and prices. The total annual cost for the system is about \$20 million. Of the \$20 million, FAS and ESCS estimates show that it takes about \$0.69 million to make production estimates for major crops in the seven foreign LACIE countries; to estimate wheat only in the seven countries requires about \$0.17 million annually.

The principal USDA use of improved foreign crop information would be to support the management of export programs and export policy decisions. The occurrence of a short wheat supply, either domestic or foreign, would especially increase the importance of such improved wheat information.

A recent study of the FAS oilseed and products program has direct implications for LACIE-type wheat information. Of all the information FAS supplied in the oilseed and products publications, crop production was clearly ranked by subscribers as their

top-priority information need. Lack of timeliness was the attribute of FAS information they criticized the most. Presumably, LACIE's strengths would be in these two areas.

Economic evaluation of LACIE is ongoing. Both present and future studies will emphasize benefits of forecasts of foreign crops, wrestle with conceptual problems, and come up with descriptions and estimates that will help decisionmakers determine whether or not improvements in information from satellite-based technology are worth the costs.

ACKNOWLEDGMENTS

The author wishes to give credit and thanks to the various members of the LACIE Economic Evaluation Planning Team, especially Howard L. Hill, USDA Project Manager, Preston E. LaFerney, Theodore F. Moriak, and William G. Stround for their ideas and comments, and to Thomas A. Miller of the Economics, Statistics, and Cooperatives Service for his earlier contributions of concepts and methodology for valuing information. The author is solely responsible for the actual contents of this report.

REFERENCES

1. Hayami, Yujiro; and Peterson, Willis: Social Return to Public Information Services; Statistical Reporting of U.S. Farm Commodities. *American Econ. Rev.*, vol. 62, no. 1, Mar. 1972, pp. 119-130.
2. Hayami, Yujiro; and Peterson, Willis: Social Return to Public Information Services; Statistical Reporting of U.S. Farm Commodities. *American Econ. Rev.*, vol. 63, no. 5, Dec. 1973, pp. 1020-1021.
3. Andrews, John: A Distribution Benefits Model for Improved Information on Worldwide Crop Production. Report 76-104-1, ECON, Inc., Princeton, New Jersey, June 30, 1976.
4. Andrews, John; Heiss, Klaus; and Sand, Francis: Economic Benefits of Improved Information on Worldwide Crop Production. Report 76-243-1A, ECON, Inc., Princeton, New Jersey, Apr. 1977.
5. Heiss, Klaus P.: The Value of Improved Global Crop Information: An Empirical Approach to Landsat Benefits. Paper presented at the First Conference on the Economics of Remote Sensing Information Systems, San Jose State University, Jan. 19-21, 1977.

6. Sand, Francis: Application of Optimal Stochastic Control Theory to Long-Range Planning in Resource Management: The Value of Wheat Supply Information. Paper presented at the Eighth Annual Pittsburgh Conference on Modeling and Simulation, University of Pittsburgh, Apr. 21, 1977.
7. Sand, Francis: An Econometric Approach to the Measurement of Benefits of Information Systems Involving Remotely Sensed Data: Worldwide Wheat Survey Case Study. Paper presented at the First Conference on the Economics of Remote Sensing Information Systems, San Jose State University, Jan. 19-21, 1977.
8. Marschak, Jacob: Economics of Information Systems. *J. American Stat. Asso.*, vol. 66, pp. 333, Mar. 1971, pp. 192-219.
9. Marschak, Jacob: Economics of Inquiring, Communication, Deciding. *American Econ. Rev.*, vol. 58, no. 2, May 1968, pp. 1-18.
10. Marschak, Jacob: Efficient Choice of Information Services. Working Paper No. 136, Western Management Science Institute, University of California, Los Angeles, May 1968.
11. Marschak, Jacob: Optimal Systems for Information and Decision. *Economic Information, Decision, and Prediction—Selected Essays*, Vol. II, Ch. 32, D. Reidel Publishing Company, Boston, 1973, pp. 342-355.
12. Howard, Ronald A.: Social Decision Analysis. *Proceedings of the IEEE*, vol. 63, no. 3, Mar. 1975, pp. 359-371.
13. Howard, Ronald A.: The Foundations of Decision Analysis. *IEEE Transactions on Systems Science and Cybernetics*, vol. SSC-4, no. 3, Sept. 1968, pp. 211-219.
14. Agnew, Carson E.: Proximal Decision Analysis With Imperfect Information. *Management Science*, vol. 23, no. 3, Nov. 1976, pp. 275-279.
15. Agnew, Carson E.; and Anderson, Robert J., Jr.: The Economic Benefits of Improved Climate Forecasting. Final report submitted to NOAA by MATHTECH, Inc., Princeton, New Jersey, Mar. 29, 1977.
16. Bradford, David; and Kelejian, Harry: The Value of Information for Crop Forecasting With Bayesian Speculators: Theory and Empirical Results. *ECON, Inc.*, Princeton, New Jersey, 1977.
17. Setting Statistical Priorities. Panel on Methodology for Statistical Priorities, Committee on National Statistics, National Academy of Sciences, U.S. Department of Health, Education, and Welfare, 1976.
18. Savage, Richard: Setting Statistical Priorities. (Review of report by the same name by the National Academy of Sciences.) *Statistical Reporter*, Dec. 1976, pp. 77-81.
19. Hoos, Ida R.: Criteria for "Good" Futures Research. *Technological Forecasting and Social Change*, vol. 6, no. 2, 1974, pp. 113-131.
20. Hoos, Ida R.: Utilization and Assessment of Remote Sensing Data. *Special Study No. 5, Social Sciences Group, University of California, Berkeley*, 1976.
21. Duncan, Joseph W.: Developing Plans and Setting Priorities in Statistical Systems. *Statistical Reporter*, Aug. 1976, pp. 281-286.
22. Sharp, James M.: On the Value of Evaluative Studies: Reflections on the First Conference on the Economics of Remote Sensing Information Systems. Unpublished manuscript, Space Sciences Laboratory, University of California, Berkeley, 1977.
23. Miller, Thomas A.: Observations of the 1978 Economic Evaluation of LACIE. Document prepared as a planning guide for LACIE Economic Evaluation Team, USDA FAS, Aug. 1977.
24. Miller, Thomas A.: Value of Information—A Project Prospectus. Document prepared as a planning guide for Economic Research Service management, USDA, Sept. 1977.
25. Breimyer, Harold E.: Economics of the Product Markets of Agriculture. Iowa State University Press, Ames, Iowa, 1976.
26. Heidhues, Theodor: The Gains From Trade: An Applied Political Analysis. Paper presented at the Symposium on International Trade and Agriculture, Tucson, Arizona, Apr. 17-20, 1977.
27. Huddleston, Barbara: Commodity Trade Issues in International Negotiations. *International Food Policy Research Institute Occasional Paper No. 1*, Jan. 1977.
28. Mayer, Leo V.; and Ahalt, J. Dawson: Public Policy Demands and Statistical Measures of Agriculture. *American J. Agr. Econ.*, vol. 56, no. 5, Dec. 1974, pp. 983-988.
29. Riemenschneider, Charles H.: Economic Structure, Price Discovery Mechanisms, and the Informational Content and Nature of USDA Prices. Michigan State University Agricultural Economics Staff Paper No. 77-19, presented at the AAFA, USDA Agricultural and Rural Data Workshop, Washington, D.C., May 4-6, 1977.
30. Food Information Systems. Hearings before the Technology Assessment Board of the Office of Technology Assessment, Congress of the United States, 94th Congress, 1st and 2nd Sessions, Sept. 24 and 25 and Dec. 10, 1975, and Feb. 4, 1976.
31. United States Benefits of Improved Worldwide Wheat Crop Information From a Landsat System, Report No. 76-122-1B, *ECON, Inc.*, Princeton, New Jersey, Jan. 1976.

32. Bradford, David; and Kelejian, Harry: The Value of Improved (ERS) Information Based on Domestic Distribution Effects of U.S. Agriculture Crops. ECON, Inc., Princeton, New Jersey, 1974.
33. The Use and Value of Foreign Wheat Production Information to USDA Programs and Activities. The Futures Group, Inc., Glastonbury, Connecticut, July 1976.
34. Vermeer, James: A Summary of the Study: The Use and Value of Foreign Wheat Production Information to USDA Programs and Activities. USDA LACIE Project Office/FAS, Oct. 1976.
35. Heid, Walter G.; Niernberger, Floyd; and Schnake, L. D.: Overview of U.S. Wheat Industry With Emphasis on Decision Making and Information Systems. Revised by F. Niernberger and J. Shull. USDA Economic Research Service, Manhattan, Kansas, July 1977.
36. Oilseeds and Products Program Evaluation. USDA FAS, Washington, D.C., Nov. 1976.
37. Osterhoudt, Frank H.: Cost of the Current USDA Foreign Crop Production Estimation System. USDA LACIE Project Office/FAS, Jan. 1978.
38. LACIE Economic Evaluation: 1978 Program. Economic Evaluation Planning Team, USDA LACIE Project Office/FAS, Dec. 19, 1977.



Supporting Research and Technology

FOREWORD

The purpose of the Supporting Research program in LACIE is to provide the technology improvements required to make LACIE—and subsequent crop inventory experiments like LACIE—successful. It is an applied program motivated by problems that have surfaced in the LACIE large-scale experiments. Although the program has provided some solutions of the “quick fix” variety, most of the research effort is directed to problem solving on a time scale of 1 to 3 years. And, since this period spans the duration of LACIE, a large portion of the program is aimed at providing improvements for later remote-sensing applications of this type. It is proper, therefore, to view LACIE as a series of large-scale experiments in which existing technology has been evaluated and as a series of studies which can form the basis for improvements in future applications.

Many institutions have contributed and are still contributing to the Supporting Research program. A complete list of these institutions and of the major disciplines they represent is given in table I. Besides contributing by performing research, members of these institutions have also contributed in at least two other ways. First, they established a certain base technology which formed the foundation of many of the concepts used in the beginning of LACIE. Secondly, they participated in periodic reviews of LACIE. These reviews were meant to be critical examinations of the functioning of the experiment to ensure that sound methods were being used. To a large extent, this contribution has had a reciprocating effect, for it has provided the institutional participants with a good realization of the actual problems encountered in an experiment of this size. Thus, it has had an enriching effect on the overall quality of research.

The research that is discussed in this section is categorized by functional areas that relate to the major elements within LACIE. These areas are (1) machine-processing methods for segment crop area

TABLE I.—Participants in LACIE Supporting Research and Their Areas of Contribution

<i>Participant</i>	<i>Study area</i>
<i>Area estimation</i>	
Colorado State University Environmental Research Institute of Michigan (ERIM)	Canopy modeling Physics, sensors, modeling, pattern recognition
International Business Machines, Inc.	Statistical design, problem solving
Oregon State University Pan American University Purdue University— Laboratory for Applications of Remote Sensing (LARS)	Agricultural economics Canopy modeling Agriculture, pattern recognition
Rice University University of California at Berkeley University of Houston University of Missouri	Computation, mathematics Image and data interpretation, sampling Mathematics Agricultural economics
<i>Yield/crop calendar modeling and estimation</i>	
Clemson University Development Planning & Research Associates Earth Satellite Corporation Fort Lewis College Kansas State University	Crop physiology Crop modeling Yield modeling Crop modeling Crop physiology, yield modeling
NOAA CCEA ^a	Yield modeling, meteorological data
Prairie View A & M University University of Wisconsin USDA SEA ^b	Agriculture Yield modeling Yield and winterkill modeling
<i>Sampling/aggregation and production estimation</i>	
South Dakota State University Texas A & M University	Stratification Sampling, statistics, mathematics, agriculture
TRW, Inc. University of Texas at Dallas	Error modeling Statistics

^aNational Oceanic and Atmospheric Administration Center for Climatic and Environmental Assessment.

^bU.S. Department of Agriculture Science and Education Administration.

estimation, (2) manual image interpretation, (3) yield estimation and crop calendar modeling, (4) sampling and aggregation, and (5) field research.

Under category (1), two main themes—accuracy and efficiency—are motivating factors. In statistical terms, accuracy deals with the bias and variance of an estimator. Thus, methods that can estimate the correct crop acreage within a segment on the average (i.e., the expected value of the estimator is the true value) and methods that consistently (i.e., with low variance) produce good answers are sought. Although low bias is often a standard requirement for any estimator, in LACIE it came to the foreground early as a potential problem area. The reason is that LACIE began with ideas that were basically directed not toward making an inventory but rather toward land use mapping. Indeed, the concept of image classification is to identify land areas in accordance with some given generic category rather than to estimate the amount of material (e.g., crop acreage) in that category. The problem with these land use mapping methods in inventory applications is that generally the classification process makes occasional mistakes, and it can be shown that this classification error will result in an estimate that is different from the true answer no matter how many pixels are classified. In other words, classification errors will in general not average out in any given classification of a large area where the classification parameters are not locally adjusted. There are a number of approaches that use classification as one step and subsequently correct for errors, and there are methods that do not use classification at all. All these approaches lead to unbiased or asymptotically unbiased estimators. The performance (as measured, for example, by the variance in the results) of those that use classification improves as classification error is reduced. And, since classification has continued to play a dominant role in LACIE, improved classification methods have continued to be a research theme.

Efficiency, as used here, relates to processing methods that require a minimum of manual intervention. In terms of classification, this implies methods that call for only a minimal amount of machine classifier training. Throughout the course of LACIE, there has been a belief that machine processing can play a dominant role in area estimation. Yet, in the large-scale experiments, training samples need to be drawn from each segment that is classified. Hence, a large portion of the research effort has been directed toward solving this efficiency problem—a

problem that has come to be called the signature extension problem.

Manual image interpretation has continued to play a key role throughout LACIE since the use of ground observation data, especially over foreign areas, was ruled out. Even though improvements in machine-processing methods have been sought, the requirement to be able to adapt to anomalous situations implies that a certain amount of manual interpretation will be necessary. Interpretation of Landsat imagery indeed presents a challenging problem. Unlike high-resolution aircraft image data, for example, a Landsat image captures very little textural information related to an agricultural crop. Because texture is so important in human recognition processes, conventional photointerpretation methods thus have to be rethought. The central research goal has been formulating human decision-making processes that are predicated on an understanding of crop development properties and determining how those properties appear in Landsat data under a variety of environmental circumstances. The design of more informative displays, better use of ancillary data, and development of methods based on a question-and-answer scheme are some of the research projects that have contributed in this area.

Yield estimation and crop calendar development were originally approached from a regression modeling point of view. For example, in yield estimation, records of yields and associated weather variables were used that dated back as far as 40 years. Although moderately successful, the approaches had some serious drawbacks. First of all, the general approach was developed over the United States, where data records of this kind are available. This is not the case in many foreign areas. Hence, even if a satisfactory model were developed in the United States, the chance of obtaining the same data records with which to derive model coefficients in foreign areas is small. Moreover, the model developed with coefficients derived from U.S. data would probably not work well elsewhere. Secondly, the regression approaches worked well when average weather was experienced during the year. However, unusual weather caused unacceptably large yield prediction errors. Hence, in yield research, the thrust has been to build a yield model that is responsive to weather variation and that is "transportable"; i.e., a model that can be applied in many different areas besides the one for which it was developed.

The original crop calendar model used in LACIE

was a version of the Robertson spring wheat model in which crop stage development rates were predicted from daily maximum and minimum temperatures and day length. Although the model works well for spring wheat when a proper starting time is known, it will not predict well in the case of winter wheat because winter wheat, unlike spring wheat, enters a period of dormancy, a stage that is not accounted for in the Robertson model. The bulk of the research, then, has been to develop methods for estimating planting dates, in order to properly start the models, and for estimating the length of dormancy.

The topic of sampling and aggregation deals with the development of areal sampling strategies and with methods for combining crop area and yield estimates in order to estimate production, for example, at a country level. Here, the emphasis has been to increase the efficiency of the sampling; i.e., to be able to estimate crop area to a given variance with fewer samples. The original sampling strategies in LACIE were developed using historical information on crop proportions in a political division (e.g., counties in the United States); however, in foreign areas, such data are seldom available. Thus, the need to improve sampling has led naturally to proposals for using Landsat data to develop rough estimates of crop acreage on which to stratify or to look for correlates of crop acreage which can be used as stratification variables.

The field research program was conceived as an area of fundamental research into the character and

controlling factors of spectral radiation patterns of crops and soils. For 3 years, high-resolution spectral measurements, supported by intensive agronomic observations, were made in controlled experimental plots at an agriculture experiment station in Kansas and at one in North Dakota. At nearby test sites in commercial production areas and at a third experimental site in South Dakota, similar spectral measurements were taken by airplane- and helicopter-borne spectral sensors, also supported by agronomic observations. All data acquisition was scheduled to coincide as nearly as possible with Landsat overpasses. The field research data have been integrated into the LACIE research data base and have been applied to studies directed at current LACIE problems, some of the results of which are reported in this Supporting Research results section. The field research data will continue to be used in studies of crop and soil radiation patterns.

There are two types of papers contained in this section. One is a review paper, which is intended to integrate the research approaches taken so that the reader can understand the main concepts without having to read through considerable technical detail. The other papers provide that technical detail and are intended to appeal to readers who are perhaps interested in pursuing research topics of this type on their own. Still more detail is provided in the references cited. Many of these references are research reports that have been compiled during LACIE and are available to the reader on request.

Methods for Segment Wheat Area Estimation

R. P. Heydorn,^a M. C. Trichel,^a and J. D. Erickson^a

INTRODUCTION

The classification and mensuration approach used in the large-scale experiment studies within LACIE was designed to provide wheat proportion estimates for each segment.¹ Such proportions were provided by integrated manual and machine processing, which is discussed in the paper by Heydorn et al. entitled "Classification and Mensuration of LACIE Segments." Two fundamentally different designs were attempted in the 3 years of LACIE. The first design, which was implemented during the first 2 years, was one in which the analyst was required to sample a color-infrared (CIR) Landsat image and pick examples of wheatfields and of nonwheat areas. The Landsat reflectance values contained in these examples were then used to estimate classification parameters. These examples will subsequently be referred to as "training data" and the function of estimating classification parameters will be called "training the classifier." Once the classifier was trained, every picture element (pixel) in the segment was classified as wheat or nonwheat. The wheat acreage estimate for the segment was then simply the total number of pixels classified as wheat.

The large-scale experiment results indicated that this approach would not support LACIE goals.² There were a number of difficulties. Basically, the classification performance was erratic. Given enough time, an analyst could generally rework the segment

by alternately choosing training fields and reclassifying until, by qualitative judgment, results were satisfactory. In short, it was a very "arty" process. Because of the number of segments that needed to be processed in a given period of time, the analyst was limited in his exposure to any one segment. In the allotted time, it was very difficult to obtain an acceptable multitemporal classification; consequently, almost all classifications were of the unitemporal variety.³

The root of the problem, however, was probably the fact that an analyst was required to examine a CIR image and discern from observations of colors all the spectral variety required to train a complex statistically oriented classification algorithm. The mental transformation from color to statistical concept, such as the "number of normal distributions required to fit the data" or "appropriate sample sizes required to estimate distribution means and covariances" is indeed a difficult if not an impossible chore.

The difficulties with this design motivated the development of a second design called Procedure 1 (see the paper by Heydorn et al.) Procedure 1 relieved the analyst of nonlabeling functions and required only that he label accurately given pixels (called dots) randomly selected from the image. Through the use of a clustering mechanism, these dots, in part, served as training data. The clustering provided both the number of normal distributions to fit the data and the estimates of the classification parameters. In addition, Procedure 1 allowed the analyst to correct for classification errors without reclassifying a segment. This correction process, when viewed statistically, is simply a stratified area estimate in which the classification is treated as a

^aNASA Johnson Space Center, Houston, Texas.

¹The term "segment" refers to the 5- by 6-nautical-mile primary sampling unit as discussed in the paper by Feiveson et al. entitled "LACIE Sampling Design."

²The stated objective in LACIE was to estimate wheat production at a country level in compliance with the 90/90 criterion (see the plenary paper by MacDonald and Hall entitled "LACIE: An Experiment in Global Crop Forecasting"). The segment estimates therefore had to support that criterion.

³Multitemporal classification means that Landsat measurements for multiple passes are concatenated to effect a multivariate classification. Unitemporal classification means that just one pass of Landsat data is used.

stratification of the segment into potential wheat and nonwheat strata.

To implement Procedure 1, it was necessary to design a suitable clustering algorithm. However, a year's experience with the operation of Procedure 1 indicated that a clustering problem still existed and this precipitated further research. This time, somewhat more sophisticated approaches were taken which considered spatial as well as spectral properties of Landsat data.

Other research that was incorporated into the Procedure 1 design dealt with the transformation of Landsat data to enhance properties related to crop growth. Specifically, a transformation was developed for converting the Landsat channels to variables that quantified soil brightness, the green development of the crop canopy, and senescence or yellowing of the crop canopy. This transformation was used to develop numerical displays for the analyst (only the soil brightness and greenness variables were used) to allow him to track crop growth and relate it to a crop calendar. In addition, it allowed him to view spectral space and deduce clustering properties within the measurement.

Besides the problems arising from Phases I and II of LACIE which stimulated research leading to Procedure 1 and other developments, certain problems were known to exist before LACIE. One of these was that whenever an area estimate is obtained using an error-prone classification process, errors in classification can introduce bias into the estimate. This known problem motivated research into the development of area estimation methods that do not depend on classification or that attempt to remove the bias caused by classification. In fact, one of the concepts considered in the early research was later incorporated into the stratified area estimation concept used in Procedure 1. Another problem that was suspected to exist before LACIE and that was confirmed as a problem after the start of the experiment dealt with the efficient use of machine processing in such an application. Simply stated, the problem is one of being able to classify large areas accurately with only a minimal amount of training data. The concept came to be known as "signature extension." In LACIE, training data are required for each segment to be classified. An example of signature extension would be a situation in which only one of five segments is used to obtain training data (with the implication that the "signatures" of wheat and nonwheat are obtainable from that segment) and all five are subsequently classified. In fact, such an attempt was made

in the first year of LACIE and it resulted in failure. Several approaches to the problem were subsequently considered. The final approach, which was in the developmental stage at the end of the third year of LACIE, was based on statistical sampling concepts.

This paper is intended to be an overview of the major research that was conducted during the 3 years of LACIE to solve problems associated with segment wheat area estimation. Papers detailing the mathematical notions are referenced. The research topics that have been alluded to previously and that will be covered in the following sections are Proportion Estimation, Clustering, Feature Extraction, and Signature Extension.

PROPORTION ESTIMATION

Proportion estimation considers methods that estimate the areal proportion of a crop type in a segment. The central idea is to obtain estimators that are unbiased or at least are asymptotically unbiased. "Asymptotically unbiased" means that as the sample size gets larger, the bias of the estimates gets smaller.

If crop types could be uniquely identified from the Landsat spectral measurements, then unbiased estimation of crop proportions would be relatively simple. An approach would be to classify each pixel into its correct crop type, tabulate these classifications, and divide each tabulation by the total number of pixels to obtain the estimate. The problem is that crop types are not uniquely identifiable from Landsat imagery, or, if they are, an error-free identification method has not been found. Consequently, one must develop methods that account for these errors. Fortunately, statistical methods have been proposed to deal with this problem. Some of these methods were known before the start of LACIE; others were developed during the course of the project.

To obtain an intuitive feel for the nature of the problems and of possible solutions, consider the diagrammatic explanation illustrated in figure 1. In this figure, weighted wheat and nonwheat probability densities or likelihoods are illustrated. The weighting consists of the actual proportions of crop acreage in a segment. Imagine that a maximum-likelihood classification procedure is to be used to calculate the proportions of crop types in a segment. (A maximum-likelihood classifier is used in LACIE. For discussion, see the paper by Heydorn et al.) Such a classifier would partition the Landsat segment into

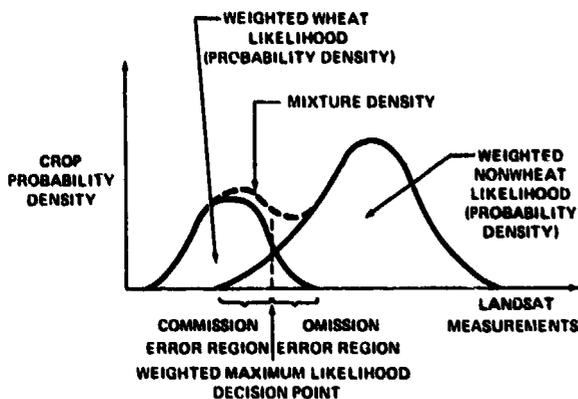


FIGURE 1.—Diagrammatic explanation of proportion estimation.

two groups. One group would contain all measurements for which the weighted likelihood of being wheat was greater than that of being nonwheat; in the other group, the reverse situation would hold. The point that defines this partition is labeled "weighted maximum-likelihood decision point" in the figure. If decisions are made in this way, errors of omission and commission will occur. If these omission and commission errors do not balance out, the final tabulation will be in error. This error is called a bias.

One approach to obtaining an unbiased estimate is to estimate the amount of error involved and in some fashion remove it from the final answer. Another approach is based on the observation that the mixture density, which is simply the probability density associated with each pixel of Landsat measurements without consideration of crop type,⁴ is (under appropriate assumptions) a unique mixture of the densities associated with each crop. When the crop type densities are known, the exact proportions in which they are mixed to form the mixture density can be computed.

A number of methods fall under the two approaches. The ones that appeared to fit the LACIE application best are mentioned in table I. Detailed description of each method can be found in the paper by Feiveson entitled "Estimating Crop Proportions From Remotely Sensed Data." To illustrate the ideas

⁴In more elementary terms, the mixture density can be viewed as that obtained by simply histogramming the Landsat measurements, where a histogram value (probability density) is the frequency with which a given pixel occurs in the scene.

involved, consider two of the methods mentioned in table I.

First consider the so-called "CDF mixture method." The basic model is

$$F = \alpha F_1 + (1 - \alpha)F_2$$

where F = marginal cumulative distribution function (CDF) corresponding to the mixture density⁵

F_1 = CDF corresponding to the wheat density

F_2 = CDF corresponding to the nonwheat density

α = proportion of wheat area in the segment

Given that F_1 and F_2 can be estimated through some sort of a sampling process, in which, for example, an analyst would pick and label examples of wheat and nonwheat pixels from a Landsat CIR image, an estimate of α , say $\hat{\alpha}$, would be the value of ξ that minimizes

$$\left| F(x) - \xi \hat{F}_1(x) - (1 - \xi) \hat{F}_2(x) \right|^2$$

subject to the constraint that $0 \leq \hat{\alpha} \leq 1$. Here,

$$\hat{F}_1(x) = [\hat{F}_1(x_1), \hat{F}_1(x_2), \dots, \hat{F}_1(x_n)]$$

$$\hat{F}_2(x) = [\hat{F}_2(x_1), \hat{F}_2(x_2), \dots, \hat{F}_2(x_n)]$$

where x_i = Landsat measurement for the i th channel, $i = 1, 2, \dots, n$. (The symbol "A" denotes "estimate of.")

In this approach, F can easily be determined to a high degree of accuracy by considering as many pixels in the segment as is deemed necessary, since crop-type labeling information is not required. On the other hand, since a knowledge of crop types is re-

⁵The marginal cumulative distribution function (or CDF) is the indefinite integral of the density function. That is, $F(x) = \int_{-\infty}^x f(y) dy$, where f is the density of one channel of Landsat measurements.

TABLE I.—Proportion Estimation Methods

Method	Description	Responsible institution
Inverting the confusion matrix	Estimate the omission/commission error matrix and use it to correct for bias	University of Texas at Dallas
Maximum-likelihood estimate of proportion	Assume normal component densities and maximize the likelihood of the mixture distribution with respect to mixing proportions	University of Texas at Dallas
Method of moments	Estimate the proportion of component moments in the mixture moments	Texas A. & M. University (TAMU)
CDF mixture method	Estimate the proportion of component marginal cumulative distribution functions (CDF's) in the mixture marginal CDF's	University of Texas at Dallas
"BIN" method	Same as CDF mixture method except density histograms used in place of CDF's	Lockheed Electronics Company (LEC)
Posterior probability	Treat classification as a small-grains/non-small-grains stratification and estimate small-grains proportion from a stratified random sample	NASA Johnson Space Center (JSC)

quired to estimate F_1 and F_2 , their estimates require considerably more effort. Consequently, the estimates of F_1 and F_2 are generally made up of considerably fewer samples than is the estimate of F . The variance of $\hat{\alpha}$ is dominated by the variance of F_1 and F_2 . Hence, in practice, $\hat{\alpha}$ is not expected to be exactly α , the true proportion.

Consider the method called "posterior probability" in table I. This method is the one used in Procedure 1. The basic idea here is to correct for classification error through a second sampling of the segment. More appropriately, however, the method can be considered as a two-stage sampling process. In the first stage, the analyst samples the segment to obtain a machine classification of wheat and nonwheat. The resulting classification map is treated as a stratification of the segment into a potential wheat stratum and a potential nonwheat stratum. The analyst then samples again and uses the sample and the stratification to complete a stratified area estimate. In the second sampling, the analyst can allocate his sample in proportion to stratum sizes. This procedure is called probability proportional to size (PPS) sampling.⁶ The proportion estimate takes the form

$$\hat{\beta} = \hat{\pi}_{11}\beta + \hat{\pi}_{10}(1 - \beta)$$

where $\pi_{11} = Pr(\text{analyst decides pixel is wheat} | \text{pixel classified wheat})$
 $\pi_{10} = Pr(\text{analyst decides pixel is wheat} | \text{pixel classified nonwheat})$
 $\beta = Pr(\text{pixel classified wheat})$

The estimators $\hat{\pi}_{11}$ and $\hat{\pi}_{10}$ are estimators of conditional probabilities. These conditional probabilities are often called posterior probabilities. The term β is the marginal probability of the classifier's calling a pixel wheat. It is obtained simply by tabulating the classification results.

Notice that the posterior probabilities, the $\hat{\pi}$'s, can be viewed as corrective terms to the machine estimate, β , of the true proportion. Although the posterior probabilities are not directly related to the omission and commission errors illustrated in figure 1, these are in a sense the inverse probabilities associated with those errors.

As with the other proportion estimation methods, the estimators associated with analyst interpretation, the $\hat{\pi}$'s, are costly to obtain and are generally obtained through only limited sampling. The $\hat{\beta}$ determined by this method, however, can be shown to be unbiased and has an associated variance that is lower

⁶This is one option available in Procedure 1. A second option called poststratified sampling, which was the one used in LACIE, is explained in the paper by Heydorn et al.

than the variance that could be obtained through the use of a second sampling without the benefit of machine classification.

Each of the methods listed in table I was evaluated on typical LACIE segment data. All labeling, however, was done from ground-truth information and not from analyst interpretation. The result showed that on Landsat data, no one method had a clear-cut advantage (see the paper by Feiveson entitled "Estimating Crop Proportions from Remotely Sensed Data"). The posterior probability approach, however, more easily fits into an approach in which some form of classification is used, as was the case in the first two phases of LACIE. Since there was a considerable body of knowledge about classification approaches, use of a classification approach was continued in the second LACIE design.

CLUSTERING

Clustering was used in LACIE as a means of automatically estimating parameters required for maximum-likelihood classification. In particular, the LACIE classification algorithm was based on the assumption that each crop type could be statistically modeled as a linear combination of normal distributions. By first clustering a segment, it is possible to associate each cluster with a crop type by analyst interpretation methods and thereby to associate a cluster with a crop-type distribution. The means and covariances obtained by using every sample in a given cluster can serve as estimates of the required crop-type means and covariances. Admittedly, these estimates have some undesirable properties;⁷ however, on the whole, the approach is feasible,⁸ as was demonstrated in the large-scale experiments with Procedure 1.

Before the design of Procedure 1, research was started to develop clustering algorithms that would

⁷Points from a cluster are in fact samples from a truncated distribution. This means that the estimates made on the basis of this truncated distribution of the mean and covariance of the untruncated distribution are likely to be biased.

⁸This statement is not intended to address the notion of using clustering alone as a means of classification. This concept is considered in the paper by Kauth and Richardson entitled "Signature Extension Methods in Crop Area Estimation."

function well in this application. The first approach was simply to take an available cluster algorithm⁹ which clustered point spectral data and refine it so that it would produce acceptable clusters without requiring continual adjustment of parameters. This approach was moderately successful, as judged by the results obtained with Procedure 1.

To develop an improved algorithm, two fundamentally different approaches were taken. One was simply to use point-clustering ideas but add additional information regarding the spatial properties of Landsat agricultural data. For the most part, this additional information was associated with the agricultural field structure in the scene. Algorithms of this type will be referred to as spatial clustering algorithms. The other approach is related to the concepts explained in the previous section on proportion estimation. That is, the mixture density is resolved into its component probability densities in the hope that these component densities can be associated with unique crop types.

Table II is a list of the clustering algorithms that were investigated in LACIE. Detailed explanations of each algorithm can be found in the references cited in that table. A synopsis of the main ideas follows.

ISOCLS

The Iterative Self-Organizing Clustering System (ISOCLS) is fundamentally a point-clustering algorithm in that each pixel of Landsat measurements is grouped with some other pixel without regard to the coordinate (spatial) location of that pixel in the image. A set of starting points (or seed points) is given and each pixel is assigned to the closest (in the l_1 or "city-block" metric) seed. This operation forms the initial set of clusters. Means and standard deviations are estimated for all clusters in this initial set. Next, a sequence of operations is performed in which clusters having too large a standard deviation in any given coordinate direction are split to form two smaller ones, clusters whose means are too close together are joined, and clusters that are too

⁹An algorithm called ISOCLS (Iterative Self-Organizing Clustering System) was available at the Johnson Space Center. It was the one used in Procedure 1.

TABLE II.—Clustering Algorithms

Clustering algorithm	Type of clustering	Responsible institution
ISOCLS ^a	Spectral	LEC
BCLUST ^b	Spectral	Environmental Research Institute of Michigan (ERIM)
ECHO ^c	Spatial	Laboratory for Applications of Remote Sensing
AMOEBA ^d	Spatial	TAMU
CLASSY ^e	Spectral	JSC (NRC) and LEC
UHMLE ^f	Spectral	University of Houston

^aE. P. Kan, "The JSC Clustering Program ISOCLS and Its Applications," LEC-0483, Lockheed Electronics Co. (Houston), July 1973.

^bR. J. Kauth and W. Richardson, "Signature Extension Methods in Crop Area Estimation," LACIE Symposium.

^cR. L. Kettig and D. A. Landgrebe, "Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects," IEEE Trans. Geoscience Electronics, vol. GE-14, no. 1, Jan. 1976, pp. 19-26.

^dJ. Bryant, "On the Clustering of Multidimensional Pictorial Data," LACIE Symposium.

^eR. K. Lenington and M. E. Rassbach, "CLASSY—An Adaptive Maximum Likelihood Clustering Algorithm," LACIE Symposium.

^fW. A. Coberly and C. L. Wiginton, "UHMLE—Program Description User Guide," Rep. 48, Dept. of Math., Univ. of Houston, Oct. 1975.

small are combined with larger ones. Following this operation, the mean of each cluster is recomputed. Pixels can be reassigned to the nearest mean and the process repeated as many times as desired. The algorithm can therefore compute the number of clusters in a given application.

BCLUST

The BCLUST or "blob-clustering" algorithm includes spatial information in the clustering process by augmenting each vector of Landsat channels with two location coordinates; i.e., the line and point numbers of the pixel. These augmented vectors are clustered into groups called "blobs" that are intended to be representative of agricultural fields. This "blobbing" is done by comparing a given augmented vector with the means of established blobs. The vector is assigned to the closest blob provided its distance is less than a given threshold. If the distance is greater than that, then the vector is designated as the mean of a new blob. After a new vector has been added to a blob, a new blob mean is computed using all vectors in that blob. When all vectors have been assigned to blobs, then a similar procedure is used to build

clusters from blobs. During this phase of the algorithm, clusters are formed by grouping the means of blobs. A blob mean is defined as the arithmetic average of the vectors for all the pixels in that blob. The spatial coordinates are deleted from the vectors during this calculation.

ECHO

Rather than operate on one pixel at a time as do the ISOCLS and BCLUST algorithms, the algorithm ECHO (Extraction and Classification of Homogeneous Objects) operates on four pixels at a time, using a "region-growing" concept. The hope is that regions (or clusters) will resemble agricultural fields. Groups of four pixels, or superpixels, are tested for homogeneity. Nonhomogeneous superpixels are judged to be on field boundaries and are initially excluded. The homogeneous superpixels are considered to be field interior pixels. Next, field superpixels are combined into regions. Two contiguous superpixels are grouped if they pass a statistical "goodness of fit" test (that is, a test is made to determine whether the pixels in two superpixels come from the same distribution). If two superpixels are grouped, then any subsequent comparisons are made between contiguous superpixels and the existing group or cluster. In this way, clusters grow until the test fails. Finally, the boundary pixels are assigned using maximum-likelihood classification.

AMOEBA

AMOEBA is also designed around spatial clustering principles. Here, however, a region-growing concept is not applied. Rather, a set of heuristically derived rules is first employed to single out boundary pixels and all points judged very near these boundaries. The complement of all such points makes up another set of pixels called "patches." The basic hope is that these patches closely resemble agricultural fields or at least homogeneous portions of fields. Pixels are then clustered by grouping with a set of patch means using the same comparative logic (nearest neighbor) as does the first group operation in ISOCLS. Clusters are then tested for purity using a misclustering criterion that considers a clustering error to have occurred if, for a given pair of pixels, (1) the pixels came from the same real class but are clustered differently or (2) the pixels came from

different real classes and are in the same cluster. The error given by (1) can be tested by assuming that all pixels in the same patch come from the same real class. This assumption is reasonable if a patch represents one agricultural field (assuming that a field can contain only one crop type). The examination of the error given by (2) is not as straightforward. Basically, the idea is that two pixels that are spectrally distant ought to belong to different real classes. After the misclustering criterion has been computed, the cluster with the most errors is eliminated. The pixels in the deleted cluster are combined with the existing clusters, again using a nearest neighbor logic. The misclustering criterion is again computed and the process is repeated. The final clustering is that which gives the lowest value for the misclustering criterion.

CLASSY

The CLASSY algorithm attempts to fit the total mixture density (where mixture density is as defined in the section on proportion estimation) of the segment pixels by normal distribution density functions. The number of distributions, the mean and covariance matrix of each, and the mixing proportions are estimated. Initially, the algorithm attempts to fit one normal distribution to the mixture density by using iterative maximum-likelihood estimation procedures to estimate the mean and covariance matrix. On the basis of higher order moments, a test is made to decide on the goodness of the fit. If the test fails, two normal distributions are tried. The parameters of the two new distributions are obtained by fitting to the first through the fourth moments of the original parent cluster. This splitting operation may be repeated. Periodically during the computations, the choice to split is reexamined, and, as a result, the parent cluster may be restored. When the choice is made to fit with more than one normal distribution, the mixing proportions (prior probabilities) are available, as they have been estimated using maximum-likelihood estimation methods. Thus, in an inventory application (like LACIE), CLASSY will estimate crop-type proportions provided a crop-type label is assigned (through analyst interpretation or otherwise) to each cluster.

UHMLE

The University of Houston Maximum-Likelihood

Estimation (UHMLE) algorithm is similar to CLASSY except that the number of component distributions to fit the mixture must be specified. No splitting or joining of the clusters is attempted. Maximum-likelihood iteration is used in estimating the proportions and the mean and covariance matrix of each cluster.

Evaluation

Extensive testing of the algorithms other than ISOCLS has not yet been done. (For details of tests done on ISOCLS, see the paper by Wheeler et al. entitled "An Evaluation of Procedure 1.") However, it is reasonable to believe that either the spatial variety or the distributional variety (e.g., CLASSY) would offer significant improvement just as a basic classifier of Landsat data. Since agricultural fields are very likely to contain the same crop type (at least, this is normally the case in the United States), the spatial algorithms should clear up a substantial amount of spectral confusion. The distributional variety is theoretically the ideal type of algorithm for an inventory application. Indeed, these algorithms can directly estimate the proportion of a crop type using principles similar to those discussed in the first section of this paper. It remains to be seen whether or not the assumptions related to this theory hold in real applications and whether or not the algorithms are sufficient to withstand violations of the assumptions.

FEATURE EXTRACTION

In the initial phases of LACIE, research in feature extraction topics was pursued along traditional lines found in the pattern recognition literature; that is, transformations of Landsat data were sought for converting the Landsat channel data to a new set of variables, called features, which preserved some desirable statistical property. These properties were expressed in terms of criterion functionals (e.g., Bhattacharyya coefficient or divergence) related to the probability of correct classification. In each case, a transformation was developed which would best preserve the probability of correct classification.

The motivation for this research was to reduce the number of Landsat measurements that must be considered in clustering and classification operations. In a multitemporal application, as many as 16 Landsat

channel measurements make up a pattern vector. Thus, a multitemporal clustering application can require a prohibitive amount of time if done on conventional computers. However, later in LACIE, parallel processor computing devices were purchased and, thus, the motivation for further research vanished.

Two accomplishments, however, stand out in this research. One was that a method was derived for computing the linear transformation that converts the Landsat measurement variables to a single variable in such a way as to best preserve the probability of correct classification. That is to say, if X is a vector of random variables of Landsat measurements, a vector b is found so that $Y = b'X$ is a new random variable for which the probability of correct classification obtained using Y in place of X is as close as possible to that obtained using X . The other accomplishment is related to the efficiency of computing features. Basically, a method was developed for computing the best linear transformation (in terms of a given criterion functional) by constructing the transformation from a set of Householder transformations. For details of these two accomplishments, see the paper by Decell and Guseman entitled "Linear Feature Selection With Applications."

In later phases of LACIE, a different concept was developed for deriving features. Rather than using statistical criteria for their definition, criteria related to crop growth were proposed. Here, the basic plan was not to minimize (or maximize) some functional, as was explained previously, but rather to develop transformations heuristically. Table III contains descriptions of the major transformations. Of those listed, the one that transforms the Landsat measurements to brightness, greenness, yellowness variables—the "tasseled-cap" transformation—has received the most attention.

Figure 2 illustrates the basic ideas involved. The figure shows that the Landsat spectral values in an agricultural scene, containing vegetation at different stages of development, will occupy a region (in the Landsat measurement space) that resembles (according to Kauth et al. in the paper entitled "Feature Extraction Applied to Agricultural Crops as Seen by Landsat") a tasseled cap. The "sweatband" part of the hat contains what is called the "plane of soils" and is essentially the region describing soil brightness. The brighter the soil, the larger the radius. As a crop develops and covers more of the soil with green vegetation, the spectral values move upward along

the greenness axis. Once the crop begins to yellow (or otherwise lose its green appearance), a component of motion along the yellowness axis is established.

The axes corresponding to this linear transformation were derived largely by empirical means through examination of Landsat data and data on soil color. While an underlying theory has as yet not been found, it appears that the transformation does bring out crop growth properties within Landsat data.

SIGNATURE EXTENSION

It would be desirable in an inventory application such as LACIE to automate the process as much as possible so that only a few individuals (analysts) could inventory an entire country. In LACIE, it was believed that this automation could be accomplished by computer analysis methods whereby an analyst could examine a small amount of Landsat data and thereby train a computer to recognize wheat over some large area.

Initially, the concept was implemented and tested in a very rudimentary way. Within a group of LACIE segments, one was selected for obtaining the training data, and then the training segment and its four nearest neighbors were classified on the basis of that training. It was soon discovered that regional effects, such as soil background or variations in cropping practices, as well as atmospheric effects modified spectral appearance to the point where such an approach failed.

The second major attempt sought to develop methods to remove the atmospheric and possibly the regional effects. Some success was achieved in developing a method that adjusted for atmospheric haze. However, when other perturbations not due to haze were present, as is almost always the case, these methods also failed. (See the paper by Minter entitled "Methods of Extending Crop Signatures From One Area to Another" for a detailed discussion of the approaches considered.)

One of these approaches is illustrated in figure 3. Essentially, the idea was to pair two segments and normalize one with respect to the other by a cluster matching process. The belief was that crop types would produce a unique cluster pattern and that if the same crop types existed in two segments, the cluster patterns would only be shifted, one with respect to the other, by the operation of some

TABLE III.—Transformations on Landsat Data^a That Enhance Crop Growth Characteristics^b

Transformation	Description	Comments
Tasseled-Cap Transformation	$\begin{pmatrix} B \\ G \\ Y \\ N \end{pmatrix} = \begin{pmatrix} 0.433 & 0.632 & 0.586 & 0.264 \\ -0.290 & -0.562 & .600 & .491 \\ -0.829 & .522 & -0.039 & .194 \\ .223 & .012 & -0.543 & .810 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$	<i>B</i> is intended to be a soil brightness variable; <i>G</i> is intended to be a variable that measures green biomass development (i.e., a vegetative index variable); <i>Y</i> is intended to be a variable that measures crop "yellowing"; <i>N</i> is called "non-such" since few or no crop development characteristics are measured by this variable; <i>X_i</i> is the <i>i</i> th Landsat channel measurement, <i>i</i> = 1, 2, 3, 4.
Transformed Vegetation Index (TVI)	$\sqrt{\frac{x_4 - x_2}{x_4 + x_2} + \frac{1}{2}}$	Channel 4 - channel 2 appears to be related to green development. The sum of these channels is used as a normalizing factor and the 1/2 is simply a constant to ensure that $\frac{x_4 - x_2}{x_4 + x_2} + \frac{1}{2} > 0$.
Transformed Vegetative Index (TVI6)	$\sqrt{\frac{x_3 - x_2}{x_3 + x_2} + \frac{1}{2}}$	TVI6 merely substitutes channel 3 for channel 4.
Differenced Vegetative Index (DVI)	$2.4x_4 - x_2$	DVI again measures a difference between channels 4 and 2. The constant 2.4 is intended to adjust the index for soil brightness; i.e., soils should give a DVI value of very near zero.
Ashburn Vegetative Index (AVI)	$2x_4 - x_2 \text{ if } 2x_4 - x_2 > 0$ $0 \text{ if } 2x_4 - x_2 < 0$	AVI is very similar to DVI without the soil line adjustment.
Ratioed Vegetative Index (RVI)	$\frac{x_2}{x_4}$	RVI has some properties similar to those of TVI. For soils, this may be roughly constant at a value of about 2.4.
Perpendicular Vegetative Index (PVI)	$\sqrt{(s_2 - x_2)^2 + (s_4 - x_4)^2}$ <p>where $s_2 = 0.851x_2 + 0.355x_4$</p> $s_4 = 0.355x_2 + 0.148x_4$	PVI is intended to measure the green development that occurs along the perpendicular to the soil brightness line. <i>S₂</i> and <i>S₄</i> are intended to measure the soil reflectance.
Perpendicular Vegetative Index (PVI6)	$\sqrt{(\tilde{s}_2 - x_2)^2 + (s_3 - x_3)^2}$ <p>where $\tilde{s}_2 = 0.498 + 0.543x_2 + 0.498x_3$</p> $s_3 = 2.734 + 0.498x_2 + 0.457x_3$	PVI6 is similar to PVI with channel 4 replaced by channel 3.

^aTransformation coefficients apply to Landsat-2 data.

^bSee A. J. Richardson and C. L. Wiegand, "Distinguishing Vegetation From Soil Background Information," *Photogram. Eng. & Remote Sens.*, vol. 43, no. 12, Dec. 1977, pp. 1541-1552.

unknown affine transformation.¹⁰ The transformation could be obtained by attempting to "overlay" the cluster pattern from one segment onto the next, as shown in figure 3. Several algorithms were developed for implementing this concept. The first was called MASC and subsequent refinements were called CROP-A, ROOSTER, and OSCAR.¹¹

Even though spectral crop-type observations are not stable from segment to segment, it may be that observations from a collection of segments maintain their discriminating properties. This concept is illustrated in figure 4. In that figure, the collection of

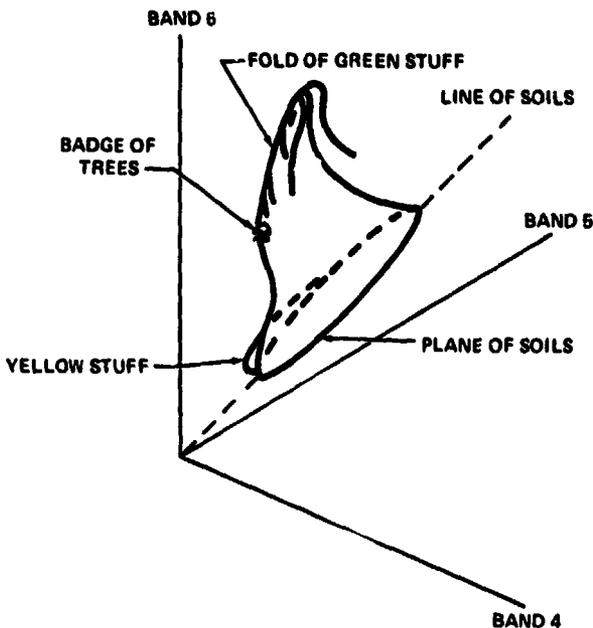


FIGURE 2.—The tasseled cap.

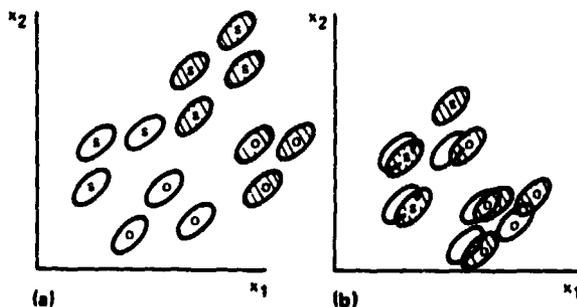


FIGURE 3.—One approach to signature extension based on a recognition-segment-to-training-segment transformation. (a) Original clusters from the training (open) and recognition (hatched) segments. (b) Recognition segment clusters transformed to conform to training segment clusters.

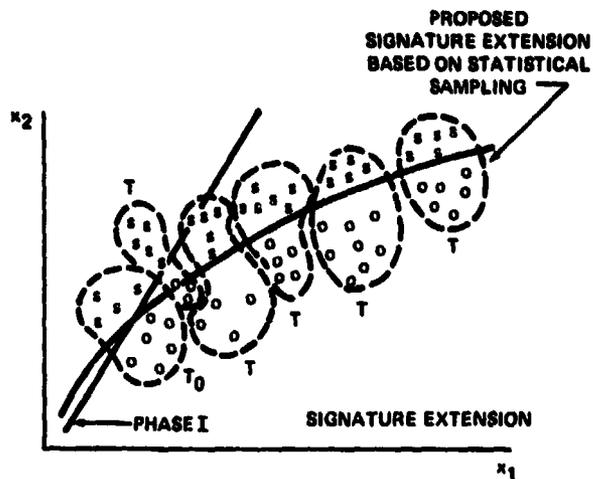


FIGURE 4.—Contrast between original and proposed signature extension concept.

points labeled T_0 is meant to represent the spectral values from one segment. It is seen that the straight line indeed separates wheat from nonwheat in that segment but not in all others. However, if data are taken from more than one segment, such as from the segments labeled T , then it becomes apparent that one curved line could separate the wheat from nonwheat in all the segments.

When dealing with LACIE segments, this concept would imply that one should be able to find a small number of segments which, when pooled, would serve as training segments for classifying the other segments. "Pooling" means that several training segments would be grouped and in effect this group treated as a larger segment in a training process. If, for example, a classification of an entire Landsat frame is desired, then it may be possible to find a very small portion of that frame that would serve as the training area. Basically, the approach is similar to the one used to classify a segment; i.e., a training

¹⁰An affine transformation is one that would rotate, contract or expand, and translate the Landsat measurement vectors. Thus, if X is a vector of Landsat measurements, $AX + b$ (where A is a matrix and b a vector) is the mapping resulting from the affine transformation $A(\cdot) + b$.

¹¹MASC (Multiplicative and Additive Signature Correction) and CROP-A (Cluster Regression Ordered on Principal-Axis) were developed by ERIM; ROOSTER (Rank Order Optimal Transformation Estimation Routine) and OSCAR (Optimal Signature Correction Algorithmic Routine) were developed by LEC.

sample is taken from that segment and used to classify the entire segment. The difference is that in the signature extension approach, much more variation in spectral values can be expected when the area of interest begins to expand from the size of a segment.

To cope with this variation in spectral values, at least two things can be done. One is first to partition the area to be estimated into strata in which the variation is expected to be small. Thus, one approach for constructing such strata would be to single out all environmentally static variables thought to affect Landsat reflectances, estimate their effects (using, for example, regression or analysis of variance approaches with field spectral mean as the dependent variable), and then establish areas in which changes in these variables are small. These areas are then the strata. This approach was in fact taken and is discussed in detail in the paper by Thomas et al. entitled "Development of Partitioning as an Aid to Spectral Signature Extension."

The so-called dynamic effects, such as those due to changes in atmospheric haze, should also be minimized. One such transformation, which uses only Landsat data, was developed to rescale the data to minimize haze distortion. The resulting algorithm, called XSTAR, applies an affine transformation, $A(\cdot) + \mathbf{b}$, where A is a matrix and \mathbf{b} a vector. The elements of A and \mathbf{b} are exponential functions of a variable that depends on the shift in yellowness (see the previous section on feature extraction). It was determined empirically that movement along the yellowness axis in a given segment is correlated with changes in atmospheric haze, and it is this observation that is exploited in the XSTAR transformation.

Given a stratification, the problem that remains is to select a small number of samples that can be used to obtain an estimate of the wheat area in a stratum. The approach (discussed in the paper by Kauth and Richardson entitled "Signature Extension Methods in Crop Area Estimation") is based on first clustering the segments into spectral groups and applying a probability proportional to size sampling to this grouping. In PPS sampling, a number of samples are allocated to each group in proportion to the size of the group. For the approach to be efficient, both spectral homogeneity and the random mix of wheat area within the agricultural area (wheat area homogeneity) should be present within a given stratum. Thus, the spectral stratifications discussed previously should be intersected with an area stratification. Such area stratifications are discussed

in the paper by Hallum and Basu entitled "Natural Sampling Strategy."

The overall approach (i.e., area stratification, haze correction, and segment selection) has yet to be demonstrated. Each of the three elements has been separately studied.

CONCLUSIONS

The Large Area Crop Inventory Experiment began with concepts that could be traced back to early mapping approaches. Here, the basic idea is to classify a given area as accurately as possible. The need to meet LACIE objectives with Landsat-type data motivated ideas related to making an inventory of an area without the use of classification (or dealing with classification error to remove estimator bias) while minimizing the number of manual interpretations required by a signature extension approach. Although many of these were not demonstrated in the large-scale experiment, significant progress was made in the supporting research program which can be applied to future designs.

As is generally the case, however, research leads to yet more research, and LACIE is no exception. It would appear that Landsat-1 and Landsat-2 data do not contain enough information to discriminate between crop types perfectly all the time and therefore a basic problem arises when no ground-truth data on crop types in an area are available. LACIE attempted to use analyst interpretation to supply this ground truth, but labeling errors resulted. Depending on the area and the time of year, these errors could be large. It would appear then that new approaches are needed to reduce labeling error. Perhaps better use of multi-year Landsat data, a more detailed understanding of the cropping practices in an area, better crop calendar prediction, and a better understanding of the limiting sources of error in Landsat data related to crop discrimination may provide the insight required to develop improved designs.

If analyst labeling errors cannot be reduced to sufficiently low levels, then methods considerably different from those used in Procedure 1 need to be considered. One such method might be centered around the use of a LIST¹² concept, in which the

¹²LIST is an acronym for Label Identification by Statistical Tabulation (see the paper by Pore and R. Abotteen entitled "A Programmed Labeling Approach to Image Interpretation").

analyst is required only to answer questions for which consistently accurate responses are possible. These responses would be used integrally in machine processes that also consider Landsat and meteorological data. Another approach might be to use some of the mixture distribution fitting methods discussed previously. In such methods, the analyst may only be required to select the appropriate component distributions to be used from a "bank" of possible distributions.

It may be that a satellite system that provided data with higher spatial and spectral resolution and that greatly increased the number of looks at a crop during its development cycle would significantly increase accuracy when coupled with the same basic processing concepts used at the beginning of LACIE. However, throughout LACIE, the desire has been to design the most accurate processing approach possible for a given sensor system. It is expected that future research will be motivated by the same desire.

Estimating Crop Proportions From Remotely Sensed Data

A. H. Fetteson^a

*The golden wheat doth flower forth throughout the LACIE segment
While eyes of Landsat scrutinize from infrared to green.
The imagery is analyzed and cluster maps are made
So that each pixel's classified by looking at its sheen.
The ones called wheat are counted up to tell how much is there,
But evil bias twists our aim—we know not how we fare!*

*What should we do? It would be nice to classify the pixels.
But lo! Our acreage estimate is subject to distortion;
Hence thus we must new methods try. For after all, in LACIE.
The goal we seek is nothing but to know the wheat proportion!*

INTRODUCTION

In LACIE, wheat acreage is estimated by the sample survey approach; i.e., the proportion of wheat is directly estimated for each of a number of 5- by 6-nautical-mile sample segments and then used in an aggregation to obtain a large area estimate. With a reasonable sample, this method stands or falls on the accuracy of the wheat proportion estimates for the individual segments.

The standard approach until recently has been to have analysts label data to train a maximum likelihood classifier, then classify every pixel in the segment and use the fraction classified as wheat for the wheat proportion estimate. Although it is intuitively appealing to be able to classify the pixels as wheat or nonwheat, such an enumeration is not required in LACIE—all that is needed is an estimate of the proportion of wheat in each segment. Because the classification/pixel-count method is theoretically biased even if all distributional assumptions are met, a search for alternative ways to estimate crop proportions has been initiated. In this paper, some of these methods are described and some initial test results on intensive site data are presented.

STATEMENT OF THE PROBLEM— THE PIXEL-COUNTING TECHNIQUE

Let $\{x_i\}_{i=1}^N$ be a set of p -dimensional measurement vectors corresponding to the N pixels in a LACIE segment. In the rest of this paper, it will be assumed that the vectors x_i are independent observations sampled from a mixture population with class distribution function (CDF)

$$F(x) = \sum_{j=1}^m \alpha_j F_j(x). \quad (1)$$

where α_j is the proportion of the segment in ground-cover class j , defined as

$$\alpha_j \geq 0; \quad \sum_j \alpha_j = 1.$$

and $F_j(x)$ is the corresponding CDF for the j th class.

^aNASA Johnson Space Center, Houston, Texas.

The most general problem is, "Given $\{x_i\}_{i=1}^N$, estimate $\{\alpha_j\}_{j=1}^m$ without any knowledge of m or $F_j(x)$." Since this problem cannot be solved without further assumptions, one of the following cases is usually assumed.

1. The number of classes m is known (or at least an upper bound is known) and $F_j(x)$ is from a known *identifiable* family, such as a set of distinct multivariate normal distributions.

2. The number of classes m is known, a set of observations from each distribution $F_j(x)$ is available (i.e., training data), and the functions $F_j(x)$ are assumed to be members of some identifiable family, usually known up to a set of parameters.

With respect to both cases, basically, a class of finite mixing distributions is *identifiable* if and only if

$$\sum_{i=1}^M \alpha_i F_i(x) = \sum_{j=1}^N \alpha'_j G_j(x)$$

for all values of x implies that $M = N$ and that, for each i ($1 \leq i \leq N$), there is some $j(i)$ ($1 \leq j(i) \leq N$) such that $\alpha_i = \alpha'_{j(i)}$ and $F_i(x) = G_{j(i)}(x)$. See Teicher (ref. 1) for more details.

The LACIE investigators are forced to work under the framework of case 2. The F_j values are not known in advance; they are estimated from training samples chosen by analysts. Since the designation of training data as being from a particular class is subject to error and since the training data set, even if designated correctly, represents only a relatively small sample from its parent population, the resulting estimates of $F_j(x)$ are not always reliable. However, until large signature banks are built up and models for adjusting distributions for haze, Sun angle, etc., are functional, there is no choice but to estimate the F_j values from training data. One of the rules of LACIE has been that no ground data can be used. If this restriction were relaxed, one could guarantee that training data were from the correct class; however, since LACIE was designed to operate in foreign areas where no ground truth is available, it would not be realistic to assume errorless labeling of training data.

The determination of m , the number of classes, is itself an interesting problem. Since the training data in a segment are insufficient to estimate the class distributions nonparametrically, it is assumed that F_j represents either a parametric family of multivariate

normal distributions or mixtures of multivariate normal distributions with known mixing weights. Which model is used depends on whether m is taken as the total number of ground-cover classes (including all subclasses of wheat and nonwheat) or is equal to 2 (wheat versus nonwheat).

In the first case, where m is taken to be the total number of ground-cover classes (i.e., the multiclass model), distribution of the j th class is assumed to be multivariate normal, with mean μ_j and covariance matrix Σ_j ; i.e., the density function for the j th class is given by

$$f_j(x) = \frac{1}{|\Sigma_j|^{1/2} (2\pi)^{p/2}} \exp \left\{ -\frac{1}{2}(x - \mu_j)^T \Sigma_j^{-1} (x - \mu_j) \right\} \quad (2)$$

in the second case ($m = 2$), there are only two "classes," wheat and nonwheat; however, since there is bound to be a large variety in the signatures of nonwheat (other crops, nonagriculture, etc.), it would be unrealistic to assume a multivariate normal distribution for all nonwheat. Instead, it is assumed that the nonwheat is itself a mixture of subclasses, each of which is multivariate normal.

A similar model is used for wheat to take into account the different varieties, growth stages, etc., of wheat found in a segment. As a consequence, the class density functions are given by

$$f_1(x) = \sum_{k=1}^{m_1} \alpha_{1k} f_{1k}(x) \quad (3a)$$

and

$$f_2(x) = \sum_{k=1}^{m_2} \alpha_{2k} f_{2k}(x). \quad (3b)$$

where $f_1(x)$ and $f_2(x)$ are the density functions for wheat and nonwheat, respectively; $f_k(x)$ is the density function for the k th subclass within the j th class; and $\{\alpha_{1k}\}$ and $\{\alpha_{2k}\}$ are the respective mixing weights.

Each approach has problems. In the first case ($m =$ number of subclasses), it is not obvious what m

should be. Present procedures employ a standard clustering algorithm in an attempt to define m and break up the data into subclasses. In the second method, m is obviously equal to 2, but clustering is still necessary to define subclasses needed for the estimation of f_{jk} . Furthermore, the coefficients α_{jk} are not known and, for lack of better information, are assumed equal within a main class; i.e., $\alpha_{jk} = m_j^{-1}$ ($j = 1, 2$).

Once classes and/or subclasses are defined, estimates of the density functions $\hat{f}_j(\mathbf{x})$ are obtained from training data, then all pixels in the segment are classified by the maximum likelihood classification rule; i.e., a pixel with measurement vector \mathbf{x} is classified as class j if

$$\hat{f}_j(\mathbf{x}) = \max_{j'} f_{j'}(\mathbf{x}).$$

The proportion of the pixels classified as "wheat" or as a subclass of "wheat" is then taken to be the wheat proportion estimate.

The accuracy of this or any proportion estimation method will depend to some extent on how well the class distributions are estimated. Some methods, however (including classification/pixel counting), are theoretically biased even if the f_j values are known. What is sought here are procedures that theoretically are relatively unbiased and fairly insensitive to errors in estimates of class distributions, so that reasonably accurate crop proportion estimates can still be made in the context of LACIE-type applications.

Bias of Pixel Counting

The present LACIE procedure of counting pixels classified as wheat will be called "pixel counting" or "PC." In this section, it will be shown that PC is biased even if the density functions $f_j(\mathbf{x})$ are known. To do this, consider the sample space \mathcal{X} of all possible measurement vectors \mathbf{x} . In maximum likelihood classification, with continuous density functions, \mathcal{X} is broken into disjoint (except for sets of measure zero) regions R_j such that

$$\bar{\mathbf{x}} \in R_j \text{ iff } f_j(\bar{\mathbf{x}}) = \max_{j'} f_{j'}(\bar{\mathbf{x}}).$$

Define conditional probabilities p_{ij} by

$$p_{ij} = \int_{R_i} dF_j(\mathbf{x}); \quad (4)$$

i.e., p_{ij} is the probability of an observation from class j being classified as class i ($i, j = 1, \dots, m$).

If α_i is the (unconditional) probability of classifying a pixel into class i , it follows that

$$\alpha_i = \sum_{j=1}^m p_{ij} \alpha_j,$$

or

$$\mathbf{e} = P\alpha, \quad (5)$$

where \mathbf{e} is the $m \times 1$ vector $(\alpha_1, \dots, \alpha_m)^T$ and P is the $m \times m$ matrix of the probabilities p_{ij} . It thus follows that, if α_w is the proportion of pixels classified as wheat, then its expectation is not α_w , the true proportion (where the subscript w denotes wheat), but is instead equal to

$$\sum_j p_{wj} \alpha_j,$$

even if the distribution for each class is known. In general, the bias vector of PC for all classes with known densities is equal to

$$\mathbf{e} - \alpha = (P - I)\alpha. \quad (6)$$

When the class densities are not known (as in LACIE) and estimates $\hat{f}_j(\mathbf{x})$ are used for $f_j(\mathbf{x})$, the situation is analogous; i.e., \mathcal{X} is split into regions \hat{R}_j such that

$$\hat{R}_j = \left\{ \mathbf{x} \mid \hat{f}_j(\mathbf{x}) = \max_{j'} \hat{f}_{j'}(\mathbf{x}) \right\}$$

and the expectation of $\hat{\theta}$ is given by

$$E(\hat{\theta}) = Q\alpha \quad (7)$$

where

$$Q = (q_{ij}) \\ = \int_{\hat{R}_i} dF_j(x); \quad (8)$$

i.e., q_{ij} is the probability of an observation from class j , falling in the i th (estimated) classification region \hat{R}_i . It follows that the bias of PC is given by

$$E(\hat{\theta} - \alpha) = (Q - I)\alpha \quad (9)$$

Unbiasing PC

If Q were known and nonsingular, one would unbiasedly estimate α by $Q^{-1}\hat{\theta}$; i.e., from equation (7), $E(Q^{-1}\hat{\theta}) = Q^{-1}(Q\alpha) = \alpha$.

It is interesting to note that knowledge of the matrix P is not necessary, or even sufficient, for unbiasing PC based on estimated densities. The advantage of knowing the densities and P lies in the variance of the corrected estimate. If $\hat{\theta}$ and \hat{u} are the PC proportions using the known and estimated densities, respectively, and if $\hat{\alpha}_k = P^{-1}\hat{\theta}$ and $\hat{\alpha}_u = Q^{-1}\hat{u}$, both of the estimates are unbiased with respective covariance matrices $P^{-1}V_{\bullet}(P^{-1})^T$ and $Q^{-1}V_u(Q^{-1})^T$, where V_{\bullet} and V_u are the covariance matrices of $\hat{\theta}$ and \hat{u} , respectively.

To obtain V_{\bullet} and V_u , note that $N\hat{\theta}$ is distributed multinomially $(\bullet_1, \dots, \bullet_m)$, where N is the number of classified pixels in the segment; hence, $V_{\bullet} = 1/N [D_{\bullet} - \bullet\bullet^T]$, where $D_{\bullet} = \text{diag}(\bullet_1, \dots, \bullet_m)$. Similarly, $V_u = 1/N [D_u - uu^T]$, where $D_u = \text{diag}(u_1, \dots, u_m)$ and $u = E(\hat{u}) = Q\alpha$; i.e., u_j is the probability of a random observation from the segment falling in the region \hat{R}_j .

Thus,

$$V(\hat{\theta}) = \frac{1}{N} P^{-1} [D_{\bullet} - \bullet\bullet^T] (P^{-1})^T \\ = \frac{1}{N} P^{-1} D_{\bullet} (P^{-1})^T - \frac{1}{N} \alpha\alpha^T \quad (10)$$

and

$$V(\hat{u}) = \frac{1}{N} Q^{-1} [D_u - uu^T] (Q^{-1})^T \\ = \frac{1}{N} Q^{-1} D_u (Q^{-1})^T - \frac{1}{N} \alpha\alpha^T \quad (11)$$

An examination of equations (10) and (11) shows that only the terms $P^{-1}D_{\bullet}(P^{-1})^T$ and $Q^{-1}D_u(Q^{-1})^T$ contribute to the difference between $V(\hat{\theta})$ and $V(\hat{u})$. The elements of D_u and D_{\bullet} , although not equal to each other, are clearly of the same average size since

$$\sum_{i=1}^m \alpha_i = \sum_{i=1}^m u_i \\ = 1.$$

As a consequence, $V(\hat{\theta})$ and $V(\hat{u})$ differ mainly because of the relative size of the elements of Q^{-1} compared to those of P^{-1} . By knowing the densities and hence the classification regions $\{R_j\}$ as opposed to estimated regions $\{\hat{R}_j\}$, one should obtain better classification accuracy and hence P should be "closer" to the identity matrix than Q . As a result, one would expect the elements of P^{-1} to be generally smaller than those of Q^{-1} .

Odell-Chhikara estimator.—Chhikara and Odell¹

¹R. S. Chhikara and P. L. Odell, "Acreage Estimates for Crops Using Remote Sensing Data." NASA Johnson Space Center Annual Technical Report (JSC-08971), University of Texas at Dallas, 1974.

proposed that PC could be made unbiased by estimating Q by \hat{Q} , say, then letting the corrected estimate of α be given by

$$\hat{\alpha}_{o-c} = (\hat{Q})^{-1} \hat{u}. \quad (12)$$

To investigate the properties of this estimator, a number of questions must be addressed. How is Q estimated? Is \hat{Q} always nonsingular? What about bias caused by the fact that $E(\hat{Q}^{-1})$ is not equal to Q^{-1} even if $E(\hat{Q}) = Q$? What if some elements of $\hat{\alpha}_{o-c}$ are negative or greater than 1 or the sum of all the elements is not 1? What is the variance of $\hat{\alpha}_{o-c}$, if it exists?

Estimation of Q : Odell and Chhikara suggested that Q be estimated by classifying training data, so that

$$\begin{aligned} \hat{Q} &= (\hat{q}_{ij}) \\ &= \frac{n_{ij}}{n_j}, \end{aligned} \quad (13)$$

where n_j is the number of training pixels labeled as class j by the analyst and n_{ij} is the number of those pixels which were classified as class i . A potential problem with this estimate is that the same data used for computing \hat{Q} are also used to train the classifier. As a result, \hat{Q} will be biased toward an identity matrix which represents perfect agreement between labeling and classification. Furthermore, errors are caused by mislabeling of training data; i.e., \hat{q}_{ij} estimates the probability of a pixel's being classified as class i given that it was labeled (by the analyst) as class j , not the probability of its being classified into class i given that it actually was from class j .

Another approach to the estimation of Q is to assume the data from class j are actually distributed $N(\hat{\mu}_j, \hat{\Sigma}_j)$, where $\hat{\mu}_j$ and $\hat{\Sigma}_j$ are means and covariances estimated from the j th set of training data, and to compute \hat{q}_{ij} by the Monte Carlo method; i.e., classifying randomly generated observations from $N(\hat{\mu}_j, \hat{\Sigma}_j)$ and using equation (13).² This procedure, of course, depends heavily on the normality assumption as well as the accuracy of $\hat{\mu}_j$ and $\hat{\Sigma}_j$, which is also subject to labeling error.

Existence of \hat{Q}^{-1} ; expectation and variance of $\hat{\alpha}_{o-c}$: If equation (12) is to have meaning, the nonsingularity of \hat{Q} must be established. Unfortunately, this is not always true when \hat{Q} is computed using equation (13) even if Q itself is nonsingular. The coefficients n_{ij} are distributed multinomially and thus have a nonzero probability of obtaining values such that \hat{Q} is singular; thus, strictly speaking, $E(\hat{\alpha}_{o-c})$ does not even exist in this case! In the second case (Monte Carlo), n_j can be taken to be arbitrarily large; hence, if the estimated class distributions are reasonably separated, \hat{Q} will not be singular. However, a theoretical problem in computing $E(\hat{\alpha}_{o-c})$ still exists.

Even if equation (13) is taken as being conditional on \hat{Q} being nonsingular and even if $E(\hat{Q}) = Q$, $\hat{\alpha}_{o-c}$ is still biased because, in general, $E(\hat{Q}^{-1} | \hat{Q} \text{ nonsingular})$ is not equal to Q^{-1} . Furthermore, the variance of $\hat{\alpha}_{o-c}$ can be quite large, resulting in estimates having elements which are negative or greater than unity.

Modified O-C estimator.—Because of the previously described shortcomings, it was decided to consider a modified estimate $\hat{\alpha}_M$, which is defined as the solution to the problem:

$$\text{Minimize } \left\| \hat{u} - \hat{Q} \hat{\alpha}_M \right\|^2 \quad (14)$$

subject to the elements of $\hat{\alpha}_M$ being nonnegative and summing to 1.

If \hat{Q} is nonsingular and the elements of $\hat{Q}^{-1} \hat{u}$ are nonnegative, then $\hat{\alpha}_M = \hat{\alpha}_{o-c}$. To show this relationship, it is necessary only to show that the elements of $\hat{Q}^{-1} \hat{u}$ sum to 1, since $\hat{\alpha}_M = \hat{\alpha}_{o-c} = \hat{Q}^{-1} \hat{u}$ makes equation (14) equal to zero; i.e., it is certainly a minimum. Since the columns of \hat{Q} sum to 1, then $\mathbf{e}^T \hat{Q} = \mathbf{e}^T$, where $\mathbf{e}^T = (1, 1, \dots, 1)$. Since $\mathbf{e}^T \hat{Q} = \mathbf{e}^T$, then $\mathbf{e}^T = \mathbf{e}^T \hat{Q}^{-1}$ and hence

$$\begin{aligned} \mathbf{e}^T (\hat{Q}^{-1} \hat{u}) &= (\mathbf{e}^T \hat{Q}^{-1}) \hat{u} \\ &= \mathbf{e}^T \hat{u} \\ &= 1 \end{aligned} \quad (15)$$

²In the m -class case. In the two-class case, the data are assumed to be distributed according to equation (3), where $\hat{\mu}_k$ is taken as $N(\hat{\mu}_k, \hat{\Sigma}_k)$ with $\hat{\mu}_k$ and $\hat{\Sigma}_k$ being estimated from the k th subclass within the j th class of training data.

since it is already known that the elements of $\hat{\mu}$ sum to 1.

Note that $\hat{\alpha}_M$ can be computed for either of the two methods of computing \hat{Q} . To distinguish between the two estimators in further references, $\hat{\alpha}_T$ will be the one solving equation (14) when \hat{Q} is computed from training data, and $\hat{\alpha}_{MC}$ will be the solution to equation (14) when \hat{Q} is computed by the Monte Carlo method.

Because of the convexity constraints in equation (14) and because $E(\hat{Q}^{-1})$ is not equal to Q^{-1} , $\hat{\alpha}_T$ and $\hat{\alpha}_{MC}$ are still biased. Whether they are less biased than the "uncorrected" estimate $\hat{\alpha}$ is at present an unanswerable question theoretically. Results of some numerical testing of these estimators will be given later.

Two-class versus m-class models.—Guseman and Walton³ have suggested that it would probably be more accurate to attempt unbiasing PC in the two-class case than in the m -class case, mainly because of the improved condition of the matrix \hat{Q} , e.g., if wheat and nonwheat were well separated but their respective subclasses were not, then the m -class "Q" matrix would be ill conditioned whereas the two-class "Q" matrix would be a near identity.

One problem with the two-class model is that, in general, Q cannot be unbiasedly estimated (even with perfect labeling of training data) unless the training data consist of a random sample within each major class. Until the advent of Procedure 1 (P-1) (see the paper by Heydorn entitled "Classification and Mensuration Approach of LACIE Segments"), training samples were not chosen at random either between major classes or within them, because it was thought that the resulting labeling accuracy would be excessively low; in the last 2 years of LACIE, however, random training samples became operationally available. The next section will include discussion of the unbiasing technique used in Procedure 1, which takes advantage of this randomness. With nonrandom sampling, however, \hat{Q} and hence proportion estimates of the form of equation (14) could easily have more bias, albeit less variance in the two-class case than in the m -class case.

Guseman and Walton give a procedure for combining the m -class and two-class estimators.³ It is

³L. F. Guseman and J. R. Walton, "Methods for Estimating Proportions of Convex Combinations of Normals Using Linear Feature Selection." Submitted to Communications Statistics—Theory, Methods, Part A.

claimed that, if the densities $f_{jk}(x)$ are known and the m -class estimator is unbiased, then a certain linear combination of the m -class and two-class estimators is also unbiased. For completeness, a brief review of the Guseman and Walton technique is given and how it may be generalized is shown.

Let β_{jk} be the proportion of the pixels from class j , subclass k , where $j = 1, 2$ and $k = 1, \dots, m_j$. Suppose one uses equation (14) to obtain estimates $\hat{\beta}_{jk}$ of β_{jk} where $m = m_1 + m_2$. If the estimates $\hat{\beta}_{jk}$ are unbiased, one can then unbiasedly estimate by

$$\alpha_j = \sum_k \beta_{jk} \hat{\alpha}_j = \sum_k \hat{\beta}_{jk} \hat{\alpha}_j.$$

(Presumably, in this situation, the estimates $\hat{\alpha}_j$, although unbiased, have a large variance and are hence undesirable as estimates by themselves.)

At the same time, let $\bar{f}_j(x)$ be an arbitrary weighted average of the given subclass densities $f_{jk}(x)$. (Guseman and Walton take \bar{f}_j to be

$$\frac{1}{m_j} \sum_{k=1}^{m_j} f_{jk}(x),$$

but this is more a convenience than a necessity.) One can then define a partition of the sample space into regions \bar{S}_1 and \bar{S}_2 such that

$$\bar{S}_j = \{x | \bar{f}_j(x) \geq \bar{f}_{3-j}(x)\}, \quad j = 1, 2.$$

Corresponding to the \bar{S}_j is a 2×2 "Q" matrix denoted by \bar{Q} , where

$$\begin{aligned} \bar{Q} &= (\bar{q}_{ij}) \\ &= \int_{\bar{S}_i} \bar{f}_j(x) dx. \end{aligned}$$

Suppose $\bar{d} = (\bar{d}_1, \bar{d}_2)^T$ is the vector of proportions of the observations lying in \bar{S}_1 and \bar{S}_2 , respectively. Then, one can let $\bar{\alpha} = \bar{Q}^{-1} \bar{d}$ be another estimate of

α , if \bar{Q}^{-1} exists. Note that $\bar{\alpha}$ is biased because the arbitrary mixing weights are not the true ones; i.e., $E(\bar{\alpha})$ is not equal to $\bar{Q}\alpha$, but is in fact equal to $\bar{Q}^*\alpha$, where

$$Q^* = q_{ij}^* \\ = \int_{S_i} \alpha_j^{-1} \sum_k \beta_{jk} f_{jk}(x) dx.$$

Since \bar{Q} is only 2×2 , however, one would expect $\bar{\alpha}$ to be a much more stable estimate than $\hat{\alpha}$.

Guseman and Walton suggest that $\bar{\alpha}$ and $\hat{\alpha}$ could be combined to obtain an estimate ("Estimator 3") of the form

$$\alpha^* = \hat{\alpha} \pm (\bar{\alpha} - \bar{Q}^{-1} \hat{Q}^* \hat{\alpha}) \\ = \hat{\alpha} \pm \bar{Q}^{-1} (\bar{\alpha} - \hat{Q}^* \hat{\alpha}), \quad (16)$$

where

$$\hat{Q}^* = (\hat{q}_{ij}^*) \\ = \int_{S_i} \hat{\alpha}_j^{-1} \sum_k \beta_{jk} f_{jk}(x) dx. \quad (17)$$

From equation (17), it is easily seen that $E(\hat{Q}^* \hat{\alpha}) = \bar{Q}^* \alpha = E(\bar{\alpha})$ and hence

$$E(\bar{\alpha} - \hat{Q}^* \hat{\alpha}) = 0 \\ \Rightarrow E(\alpha^*) = \alpha.$$

Note that any 2×2 matrix, H , could be substituted in equation (16) for $\pm \bar{Q}^{-1}$ and still leave α^* unbiased. An interesting question still unresolved is, "What should H be such that solutions of $\alpha^*(H) = \hat{\alpha} + H(\bar{\alpha} - \hat{Q}^* \hat{\alpha})$ have minimum mean-squared error?"

"Inverse" probabilities: the P-I approach.—Much of the difficulty in unbiased PC would be relieved if it

were possible to estimate Q^{-1} (if it exists) unbiasedly, instead of inverting an estimate of Q . Although no method for estimating Q^{-1} directly has been developed, it is possible to define a set of "inverse" probabilities which are estimable if a random subsample of correctly labeled training data is available. These probabilities can then be used to linearly unbiased \hat{u} without going through an inversion process, thus resulting in an improved estimate of α , guaranteed to be convex.

Suppose one had a set of estimated density functions $\{f_j(x)\}$ and corresponding classification regions $\{R_j\}$ ($j = 1, \dots, m$). If a pixel were chosen at random from the segment, one could ask what the probability is of the pixel's having come from class i given that it was classified as j . These "inverse" probabilities, h_{ij} , are, of course, related to the coefficients q_{ij} in equation (8) through the relation

$$h_{ij} = \frac{q_{ji} \alpha_i}{u_j}, \quad (18)$$

where q_{ij} , α_i , and u_j are as defined previously.

If one had a random sample of T observations and could correctly "label" them (i.e., tell what class each one came from), then an almost unbiased estimate of h_{ij} could be made by

$$\hat{h}_{ij} = \frac{t_{ij}}{T_j} \quad (T_j \neq 0), \quad (19)$$

where T_j is the number of observations classified as class j and t_{ij} is the number of observations classified as class j and labeled as class i .

Because the sample is random, T_j is a random variable (as opposed to n_j in eq. (13)); however, it is true that $E(\hat{h}_{ij} | T_j > 0) = h_{ij}$; hence, \hat{h}_{ij} is unbiased as long as T_j is not equal to zero.

If labeling were sufficiently accurate such that $\hat{H} = (\hat{h}_{ij})$ were a reasonable estimate of $H = (h_{ij})$, then PC could be corrected to give an estimate $\hat{\alpha}_j$, where

$$\hat{\alpha}_j = \hat{H} \hat{u}, \quad (20)$$

which is essentially what is done in LACIE, as described in the paper by Heydorn.

To show that $\hat{\alpha}_j$ is approximately unbiased under perfect labeling, consider $\hat{\alpha}_j^{(i)}$, the i th element of $\hat{\alpha}$, which is equal to

$$\sum \hat{h}_{ij} \hat{u}_j.$$

Its expectation is then equal to

$$\sum_j E(\hat{h}_{ij} \hat{u}_j) = \sum_j E\left(\frac{t_{ij}}{T_j} \hat{u}_j\right). \quad (21)$$

Remembering that \hat{u}_j is the proportion of the N pixels in the segment that were classified as class j and that T_j is the number of pixels classified as class j from the (random) training sample of size $T < N$, one can write $\hat{u}_j = (T_j + G_j)/N$, where G_j is the number of pixels *not in the training sample* that were classified as class j . Since T_j and G_j are from independent samples of respective sizes T and $(N - T)$, they are independent; hence,

$$\begin{aligned} E(\hat{u}_j | T_j) &= E\left(\frac{T_j + G_j}{N} \mid T_j\right) \\ &= \frac{T_j + (N - T)u_j}{N}. \end{aligned}$$

By a similar argument, given T_j , t_{ij} is independent of G_j . Also, if T_j is not equal to zero, $E(t_{ij} | T_j) = h_{ij} T_j$; hence, if T_j is not equal to zero,

$$\begin{aligned} E(t_{ij} \hat{u}_j | T_j) &= E(t_{ij} | T_j) E(\hat{u}_j | T_j) \\ &= h_{ij} T_j \left[\frac{T_j + (N - T)u_j}{N} \right] \\ \Rightarrow E\left(\frac{t_{ij} \hat{u}_j}{T_j} \mid T_j\right) &= h_{ij} \left[\frac{T_j}{N} + \frac{(N - T)}{N} u_j \right]. \end{aligned} \quad (22)$$

Since T_j is distributed binomially (T, u_j),

$$\begin{aligned} E(T_j | T_j \neq 0) &= Tu_j / [1 - (1 - u_j)^T] \\ &\approx Tu_j. \end{aligned} \quad (23)$$

Substituting into equation (22) yields the unconditional expectation

$$\begin{aligned} E\left(\frac{t_{ij} \hat{u}_j}{T_j}\right) &\approx h_{ij} \left[\frac{Tu_j}{N} + \frac{(N - T)u_j}{N} \right] \\ &= h_{ij} u_j \\ \Rightarrow E(\hat{\alpha}_j^{(i)}) &= \sum_j h_{ij} u_j \\ &= \sum_j q_{ji} \alpha_i \quad (\text{by eq. (18)}) \\ &= \alpha_i. \end{aligned} \quad (24)$$

Thus $\hat{\alpha}_j$ is unbiased under perfect labeling when all the T_j values are nonzero.

Note that, under the preceding assumptions, the raw estimates

$$\hat{\alpha}_R^{(i)} = T^{-1} \sum_j t_{ij}$$

are also unbiased estimates of α_i . It is shown in Heydorn's paper, however, that $\hat{\alpha}_j^{(i)}$ has lower variance than $\hat{\alpha}_R^{(i)}$; in fact, $\hat{\alpha}_j^{(i)}$ can be shown to be a poststratified estimate of α_i , where the strata are precisely the partition of the sample induced by the classifier.

Generalization of PC-Based Techniques

It may be noted that there is no requirement that the sample space \bar{X} be partitioned into regions through a classification rule. The discussion of unbiased PC applies for any partition of \bar{X} into disjoint regions $\{R_j\}_{j=1}^k$ where the phrase "is classified as class j " is replaced by " $\bar{x} \in R_j$ " where \bar{x} is an observation. For every partition $R = (R_1, \dots, R_k)$ of \bar{X} , there corresponds a $k \times m$ "Q" matrix, say

$$Q(R) = q_{ij}(R) \\ = \int_{R_i} dF_j(\mathbf{x}),$$

and the analogous relationship

$$\mathbf{u} = Q \alpha, \quad (25)$$

$k \times 1 \quad k \times m \quad m \times 1$

where $u_i = P\{\mathbf{x} \in R_i\}$, \bar{x} is a random observation from the mixture distribution in equation (1), and $\mathbf{u} = (u_1, \dots, u_k)^T$. Note also that the number of partitions k does not have to equal the number of classes m although, for all the elements of α to be estimable, one needs $k \geq m$. If $k > m$, one could use least-squares estimates of α to replace " Q^{-1} " in earlier discussions.

An example of a partition not based on classification would be one in which $u_i = 1/k$ for all values of i ; i.e., "statistically equivalent blocks." One desirable property of such a partition is that estimates of the type of equation (19) would be as stable as possible since, with a reasonable sample size, the expectation of all the T_j values would be safely removed from zero.

OTHER PROPORTION ESTIMATION METHODS

General Linear Functional Estimates

The second section of this paper included discussion of some proportion estimation methods which

were all similar in that (1) the sample space \bar{X} was split into disjoint regions and (2) the primary statistic used for estimation was the vector, $\hat{\mathbf{u}}$, of proportions of the pixels the measurements of which fell into each of the regions. The expectation of this vector was then shown to be a linear function of the target proportion vector α which could be solved in turn for an estimate of α as a function of $\hat{\mathbf{u}}$.

Suppose one defines indicator functions

$$g_i(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \in R_i \\ 0 & \text{otherwise,} \end{cases} \quad (26)$$

where R_i is the i th disjoint region with union \bar{X} , as in the preceding section. Then, using equation (1), equation (25) as given in the preceding section can be written

$$u_i = E(\hat{u}_i) \\ = \sum_{j=1}^m q_{ij} \alpha_j \\ = \sum_{j=1}^m \int_{R_i} dF_j(\mathbf{x}) \alpha_j \\ = \sum_{j=1}^m \alpha_j \int_{\bar{X}} g_i(\mathbf{x}) dF_j(\mathbf{x}) \\ = \int_{\bar{X}} g_i(\mathbf{x}) \sum_{j=1}^m \alpha_j dF_j(\mathbf{x}) \\ = E(g_i(\mathbf{x})) \quad (i = 1, \dots, k)$$

$$\Rightarrow E(g_i(\mathbf{x})) = \sum_{j=1}^m q_{ij} \alpha_j. \quad (27)$$

By writing equation (25) in the form of equation (27), it can easily be seen that such an equation holds for any set of functions, $\{g_j\}$, the expectations of which exist under the distribution of equation (1).

The coefficients q_{ij} are simply the conditional expectations of $g_i(\bar{\mathbf{x}})$ given that $\bar{\mathbf{x}}$ is a random observation from population j . Given enough linearly independent functions g_i , one can then use equation (27) to estimate α by replacing $E(g_i)$ with g_i . This class of estimators will be called *linear functional estimators* (LFE's).

Special Cases of LFE's

Method of moments.—To establish the *method of moments* estimator, let $g_i(\mathbf{x}) = x^{(i)}$ ($i = 1, \dots, p$) and $g_{p+i+j}(\mathbf{x}) = x^{(i)} x^{(j)}$ ($i = 1, \dots, p; j = 1, \dots, p$), where $x^{(v)}$ is the v th component of the p -dimensional observation vector \mathbf{x} . In this case, k is equal to $p + p(p + 1)/2$ and the expressions in equation (27) are safely overdetermined for any reasonable number of classes.

The left side of equation (27) is obtained by computing the first and second moments from the entire mixture distribution (i.e., the whole segment), whereas the coefficients q_{ij} on the right side of equation (27) are the conditional first and second moments, which can be computed from training data from each of the m classes. Then, as in equation (14), the system $g = Q\hat{\alpha}$ is solved for $\hat{\alpha}$ by constrained least squares, where $g = (g_1, \dots, g_k)^T$.

Marginal CDF.—Let $F_v(\mathbf{x})$ be the sample univariate CDF from the mixture distribution (eq. (1)) of $x^{(v)}$ taken over all the pixels in the segment. Then $F_{jv}(\mathbf{x})$ is simply the fraction of the observations the v th component of which is less than or equal to \mathbf{x} . Its expectation is

$$F_v(\mathbf{x}) = \sum_j \alpha_j F_{jv}(\mathbf{x}),$$

where $F_{jv}(\mathbf{x})$ is the marginal CDF of the v th component of the j th class distribution. For any set of arbitrary real numbers $\{x_{iv}\}_{i=1}^M$, the set of equations

$$F_v(x_{iv}) = \sum_{j=1}^m \alpha_j F_{jv}(x_{iv}) \quad \begin{matrix} \nu = 1, \dots, p \\ i = 1, \dots, M \end{matrix} \quad (28)$$

can be approximated by replacing F_v by \hat{F}_v and F_{jv} by \hat{F}_{jv} (the latter being computed from training data

from the j th class) and solved for α by constrained least squares. To avoid degeneracies, it is a good idea to spread x_{iv} over the distribution of $x^{(v)}$.

Specifically, x_{iv} can be taken as the (i/M) -th quantile of \hat{F}_v . Equation (28) is then replaced by

$$\frac{i}{M} = \sum_{j=1}^m \hat{\alpha}_j \left(\frac{n_{j\nu}}{N_{j\nu}} \right) \quad \begin{matrix} \nu = 1, \dots, p \\ i = 1, \dots, M \end{matrix} \quad (29)$$

where $n_{j\nu}$ is the number of observations from training sample j which had a ν th component less than or equal to x_{iv} and $N_{j\nu}$ is the total number of observations from training class j . (Note that there are pM equations and m unknowns.)

Marginal "BIN" estimator.—The CDF estimator can be slightly modified to produce what is termed the "BIN" estimator by replacing $F_v(x_{iv})$ with $F_v(x_{iv}) - F_v(x_{i-1,v})$, where x_{0v} is taken as $-\infty$. Similarly, $F_{jv}(x_{iv})$ is replaced by $F_{jv}(x_{iv}) - F_{jv}(x_{i-1,v})$. If the sets $\{x_{iv}\}$ are taken as quantiles again, then the left side of equation (29) becomes $1/M$ for all values of i and ν , whereas the coefficients $n_{j\nu}$ are taken as the number of observations in training sample j the ν th components of which lie in the interval $(x_{i-1,v}, x_{iv})$.

Density functions.—Another example of an LFE is that which uses estimated densities for g_i . Let $\hat{f}_j(\mathbf{x})$ be the estimated density function for the j th class. Then, one can let $g_i(\mathbf{x}) = \hat{f}_i(\mathbf{x})$ ($i = 1, \dots, m$). In this case, the system of equation (27) is square; i.e., $k = m$.

Maximum Likelihood Estimators

An altogether different approach to proportion estimation is that of maximum likelihood. Suppose class density functions $f_j(\mathbf{x})$ are known. If $\{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ is a random sample of N observations from equation (1), then the likelihood L of the sample is given by

$$L = \prod_{i=1}^N \sum_{j=1}^m \alpha_j f_j(\mathbf{x}_i). \quad (30)$$

Maximum likelihood estimates of α_j are obtained by

maximizing L (or $\log L$), subject to $\alpha_j \geq 0$ and sum $\alpha_j = 1$. Let

$$Q = \log L - \lambda \left(\sum_j \alpha_j - 1 \right).$$

Then

$$\frac{\partial Q}{\partial \alpha_j} = \sum_{i=1}^N \frac{f_j(x_i)}{f(x_i)} - \lambda \quad (31)$$

where

$$f(x_i) = \sum_v \alpha_v f_v(x_i)$$

and λ is a LaGrange multiplier. Setting $\partial Q / \partial \alpha_j = 0$ yields

$$\begin{aligned} \sum_{i=1}^N \frac{f_j(x_i)}{f(x_i)} &= \lambda \\ \Rightarrow \sum_{i=1}^N \frac{\alpha_j f_j(x_i)}{f(x_i)} &= \lambda \alpha_j. \end{aligned} \quad (32)$$

Summing equation (32) over j yields

$$\begin{aligned} \sum_i \frac{\sum_j \alpha_j f_j(x_i)}{f(x_i)} &= \lambda \left(\sum_j \alpha_j \right) \\ &= \lambda \\ \Rightarrow \sum_i \frac{f(x_i)}{f(x_i)} &= \lambda \\ \Rightarrow \lambda &= N. \end{aligned}$$

As a consequence, equation (31) can be rewritten as

$$\frac{1}{N} \sum_{i=1}^N \frac{\alpha_j f_j(x_i)}{f(x_i)} = \alpha_j, \quad (33)$$

which suggests a fixed-point iteration procedure for solving for α_j by successive substitution of trial values. Since any positive values of α_j substituted into the left side of equation (33) will result in positive values in the right side, such a procedure, if it converges, is guaranteed to produce a nonnegative estimate. To show that the solution to equation (33), if it exists, is a maximum, consider the second derivatives of Q :

$$\begin{aligned} h_{jk} &= \frac{\partial^2 Q}{\partial \alpha_j \partial \alpha_k} \\ &= - \sum_{i=1}^N \frac{f_j(x_i) f_k(x_i)}{f^2(x_i)}. \end{aligned} \quad (34)$$

Since the matrix $H = (h_{jk})$ is clearly seen to be a sum of seminegative definite rank 1 matrices, it is itself seminegative definite. Furthermore, if $N > m$, H is negative definite with probability 1. As a consequence, any solution to equation (33) is a maximum likelihood estimate of α .

If the functions f_j are not known, a similar scheme can be implemented using parametric estimates (usually Gaussian) for f_j . Details of this type of estimation, including some convergence theorems, may be found in reference 2.

PERFORMANCE OF PROPORTION ESTIMATORS

All the proportion estimators described in the two preceding sections are dependent on certain assumptions in order to be unbiased and to have a reasonably small variance. Every estimator requires a random sample from each of the m classes; therefore, nonrandomness or outright mislabeling could seriously affect the performance of the proportion estimator. In addition, some of the estimators have

distributional assumptions, usually normality; others, however, are nonparametric. In some cases, failure of an assumption to hold may only increase the variance of the estimate; in most, however, the estimator will become biased. As a consequence, no attempt is made to compare the merits of various estimators under ideal conditions (i.e., when all assumptions are met). What really matters in LACIE-type applications is the degree of insensitivity of the estimator to violations of its assumptions.

Because of the myriad of possible proportion estimators which can be constructed, it would be an impossible task to test them all, especially on real data. There has been very little comparative testing to date; however, Ziegler (ref. 3) did compare the performance of eight estimators on some LACIE intensive test site data. It appeared that the best results were obtained with the CDF or "BIN" estimators. The theoretical bias of raw pixel counting proved no worse than that caused by the failure of various assumptions in other estimators. The reader is referred to reference 3 for details.

In reference 4, Guseman did some limited testing comparing the m -class and two-class versions of equation (16) for the special case $H = Q^{-1}$. He found that the two-class case gave better results.

The LACIE Accuracy Assessment personnel did some comparison studies of pixel counting and Procedure 1 on eight LACIE segments. No meaningful differences in performance were noted. Again, despite being theoretically biased, PC did no worse than P-1, probably because of labeling errors in P-1.

Earlier studies by the Environmental Research Institute of Michigan (ERIM) (refs. 5 and 6) on moment-type estimators (first moments only) were also inconclusive. The only general conclusion possible at present is that poor to inadequate "training" data cause most of the proportion estimation methods to be indistinguishable. When future surveys are made with some ground truth available for training, the selection of an efficient proportion estimation method will become a much more important task than it is at present.

REFERENCES

1. Teicher, H.: Identifiability of Mixtures. *Ann. Math. Stat.*, 1961, pp. 244-248.
2. Peters, B. C.; and Walker, H. F.: An Iterative Procedure . . . (and other articles). Final Report (Univ. of Houston), NASA Contract NAS 9-12777, May 1, 1975 to May 30, 1976.
3. Ziegler, L. R.: An Evaluation of Algorithms for Estimating Wheat Proportions From Landsat Data. Technical Memorandum 642-2031, Lockheed Electronics Company, Houston, Tex., May 1977.
4. Guseman, L. F.: Applications of Feature Selection. Final Report, NASA Contract NAS 9-14689-4S, June 1, 1975 to May 31, 1976.
5. Nalepka, R. F.; and Hyde, P. D.: Estimating Crop Acreage From Space-Simulated Multispectral Scanner Data. NASA CR-ERIM 31650-148-T, Aug. 1973.
6. Horowitz, H. M.; Hyde, P. D.; and Richardson, W.: Improvements in Estimating Proportions of Objects From Multispectral Data. NASA CR-ERIM 190100-25-T, Apr. 1974.

ACKNOWLEDGMENTS

The author wishes to acknowledge the work of many contributors whose research has been included in this paper. In particular, H. O. Hartley, P. L. Odell, W. S. Coberly, J. P. Basu, and R. P. Heydorn have made many valuable inputs to the material presented herein. Of special note is the series of articles by University of Texas at Dallas personnel in their Annual Report for 1974-1975.⁴

⁴P. L. Odell, "Statistical Theory and Methodology for Remote Sensing Analysis With Special Emphasis on LACIE." NASA Johnson Space Center Annual Report (JSC-09703), University of Texas at Dallas, 1975.

Appendix

Categorization and Error Analysis of Proportion Estimators

An inspection of figure 1 will show how the proportion estimators discussed in this paper may be grouped. They are labeled "E1" to "E16" for easy reference. Under each estimator, symbols are shown which indicate the members of the given list of assumptions that must hold for the procedure to be feasible as a proportion estimator. In addition to this list in figure 1, all estimators assume, of course, that there is enough information in the data to actually separate the classes. (Such has not always been the case in LACIE!)

Note that there are three main categories of estimators shown. The first type, maximum likelihood, is heavily dependent on distributional assumption—usually that of normality. If the data were really normal for each of m known classes, maximum likelihood would probably give the best performance.

Unfortunately, the use of clustering to define classes is essentially in contradiction with the nor-

mality assumption, thus making E1 and E2 questionable unless clearly identifiable classes are known a priori. Since true means and covariances are generally unknown even if the distribution of the data is normal, E1 tends to be biased; however, it is not subject to wild variation, as is E2, when adequate training samples are not available.

The second major category of proportion estimators is that based on classification of individual pixels. Within this category are the "raw" (uncorrected) pixel-counting estimators E3 and E4; the " Q " matrix corrected estimators E5, E6, and E7; and "P-1" (E8), which is really in a category by itself.

All the estimators in this class depend to some extent on normality because they use estimated multivariate normal densities to define classification regions. Failure of the data to be normal, however, does not cause as much damage as it does under maximum likelihood; in fact, for E5 and E8, only the variance of the estimator, not its bias, is affected.

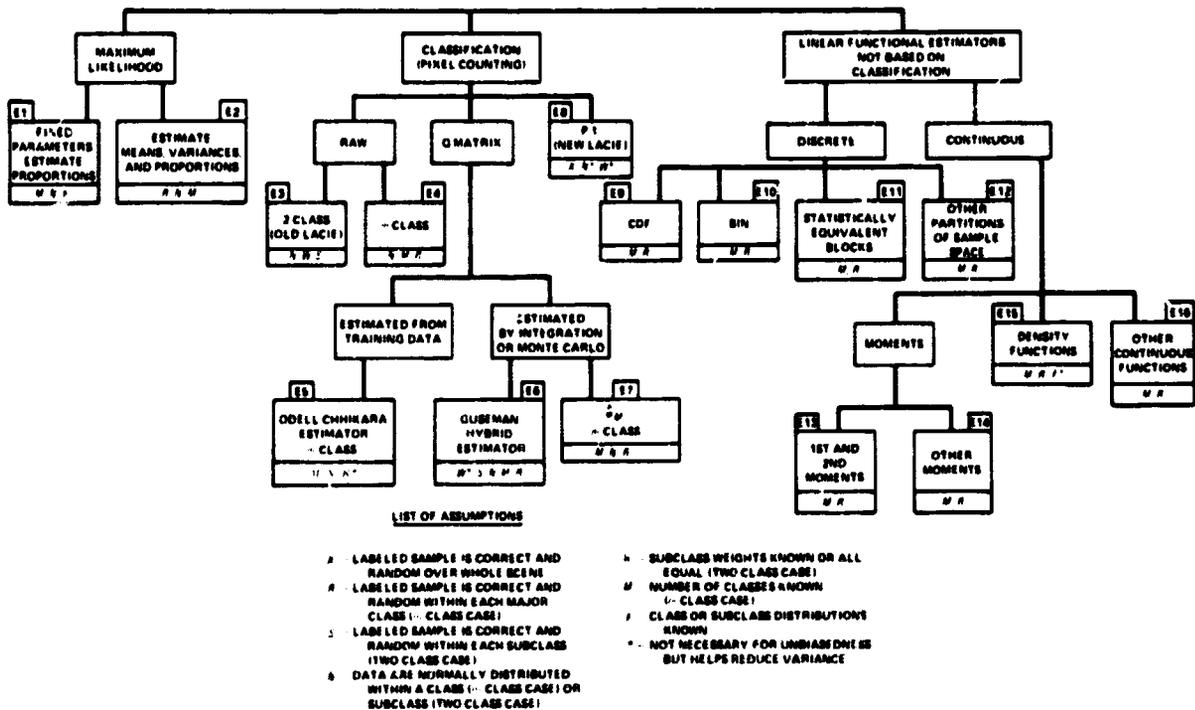


FIGURE 1.—Categorization of proportion estimation procedures.

In order to estimate the " Q " matrix unbiasedly, a good random training sample is required for each class. In E6 and E7, where the Q matrix is estimated by Monte Carlo or numerical integration (as opposed to E5, where it is estimated nonparametrically), distributional assumptions are more important. The trade-off of E7 against E5 is that fewer training observations are needed to estimate Q in E7 if the data are known to be normal, although serious errors may result if, in fact, they are not.

The Guseman estimator, E6, is a "hybrid" because it combines a two-class biased estimator with an m -class " Q " matrix procedure. Since Guseman's Q matrix is estimated by integration of normal densities, it, too, is dependent on the normality assumption.

Procedure 1, or E8, is unique in that the only real assumption for its unbiasedness is that samples selected at random throughout the *whole segment* can

be unbiasedly labeled and that all classes are represented. In E8, classification is only used as a stratification device; hence, the assumption of normality helps only through producing effective strata, not by reducing bias. Although E8 does not require as many assumptions as other procedures, the one it does require is the most stringent; i.e., the ability to label randomly sampled pixels from the mixture distribution in equation (1). Other procedures only require random observations *within* a class, not between classes.

The third major group of proportion estimators (LFE's) are essentially nonparametric. They require only knowledge of the number of classes and random samples from each class. They tend to have higher variances than some other methods, but they have the advantage of not being biased by the failure of distributional assumptions to hold.

On the Clustering of Multidimensional Pictorial Data

J. D. Bryant^a

ABSTRACT

A new approach to problems of clustering and classification of multidimensional pictorial data is presented. Proceeding logically from simple models and assumptions, the author describes the development of a clustering technique and program. Some tests of the program have been performed, and this work is reported. The techniques make use of information from the spatial domain.

INTRODUCTION

One application of remote sensing is the use of satellite- or aircraft-acquired multispectral scanner (MSS) data to conduct land usage inventories over large geographical areas. An essential part of a realistic program to conduct such an inventory is the application of cluster analysis to help the human analyst label remotely sensed data. Cluster analysis (clustering) lets the analyst label clear-cut cluster cases such as a large field (easily and accurately), while assuming that other members of the cluster are in the same class. Other members of the cluster then have the same label; since this category includes difficult-to-label cases, much tedious analysis is saved.

In this paper, several related ideas on using spatial relationships to aid clustering and classification are presented. Techniques for selecting "pure" picture elements (pixels) and "starting" cluster centers (for an iterative or *k*-means clustering program) and for assigning pixels to clusters (classification) are discussed. A new clustering technique that is both elegant and economical is proposed. Each technique is described in detail in an appendix; each has been implemented as a computer program. First tests of

the clustering program on four-pass Landsat data are described.

The origin of the innovations suggested here lies in the philosophy of the approach.¹ In dealing with such an ill-defined concept as clustering, it is essential that one identify the assumptions being made about the objects being clustered (the model) and study whether the desired results of clustering are justified on the basis of the model and the methodology being employed. It is essential that the reality of the model be verified independently of the methodology. It is not advocated that these philosophical issues be resolved here, but they are considered.

The methodology used here is derived in a new analytical framework for studying pictorial information. One job of the theorist is to proffer analytical frameworks (and models within). Another is to create sample theories consistent with the framework. The success of the ideas presented here has one inescapable consequence: current theory, based on mathematical statistics, needs to be critically examined. These early results definitely suggest that real data are inconsistent with the assumptions of current theory. What is lacking is a replacement theory. It would surely be unwise to wantonly abandon current theoretical work because a new methodology, which at first seems to contradict existing theory, has appeared. The theoretical problem is open.

A thread which starts in the preceding two paragraphs runs through all the ideas discussed here. It is the concept of reality. The reader must be cau-

^aDepartment of Mathematics, Texas A. & M. University, College Station, Texas.

¹To some extent, this work is a contribution to the "Philosophy of Clustering." It certainly seems clear that, when the approach taken here is carefully studied and combined with other work in understanding multidimensional pictorial data, a significant advance in the methodology of using spatial associations will result. (Meanwhile, it is realized there may be no such subject as the Philosophy of Clustering.)

tioned that the word "real" is used in the naive sense, principally to distinguish between what happens in the model and what has happened in experiments performed in the setting of the model. For example, in the model, the concept of a field is defined, whereas actual (real) fields are naively believed to exist. Of course, a thoughtful reader will reflect seriously on this situation and realize that naive realism cannot be justified.

CURRENT CLUSTERING TECHNIQUES

Of the many clustering techniques suggested in the literature, several on remotely sensed data seem to be effective. Iterative algorithms such as ISODATA (refs. 1 and 2) and CLASS (ref. 3) have been used successfully on four channels of aircraft scanner (C-1) data and on Landsat data (refs. 4 to 6). A similar technique is described by Wacker and Landgrebe (ref. 7). (For a discussion of the application of some of these techniques to remote sensing, see Duran and Odell (ref. 8, pp. 100 to 102). See also Anderberg (ref. 9, pp. 156 to 175) for a comprehensive comparison of variants of ISODATA in use at the time (1973).) Recent developments (in addition to CLASS), such as the gravitational clustering ideas of Ball (ref. 10) and later Wright (ref. 11), show some promise for clustering arbitrary metric data; their potential in remote sensing has not been fully evaluated.

All the techniques discussed previously are non-hierarchical clustering methods. The dominant idea in each of these techniques is to take some initial clustering of the data and rearrange the assignments (of data to clusters) to improve the partition. Even the simplest of these methods, the basic k -means program of MacQueen (ref. 12), requires two passes through the data with each pixel being classified on each pass. When programmed, most of the techniques will be structured with a starting procedure module (ref. 13): "INITIAL NUMBER OF CLUSTERS, CLUSTER CENTERS." As a module, it can be studied separately.

Although these methods include the most efficient clustering techniques known, the cost (in computer resources) of applying them to problems of remote sensing is still high. This point is discussed by Dubes and Jain (ref. 13) and by Wright (ref. 11). The cost is high because of the amount of data to be clustered (typically 23 000 pixels in one Landsat segment), the dimensionality of the data (four channels

for each temporal acquisition), and the nature of the clustering program. There are four obvious approaches to reducing the cost.

1. Reduce the dimensionality of the problem.
2. Select a small but representative subset of the data and cluster it.
3. Get a better starting partition.
4. Get a better clustering program.

Implementation of the first of these methods is variously called feature selection, factor analysis, or multidimensional scaling. An account of the mathematical-statistical trickery of feature selection can be found in Andrews (ref. 14). There is convincing evidence (refs. 15 and 16), for example, that each acquisition of Landsat agricultural data is at most two rather than four dimensional. However, even if the dimensionality is reduced to two for each acquisition (using a transformation such as the one developed by Kauth and Thomas (ref. 15)), high dimensionality remains because at least three acquisitions are required to separate the real classes present (in this case, crops). The second point, which might be called data selection, and the third are discussed next; the fourth point is discussed after a brief discussion of models and assumptions.

TECHNIQUES FOR SELECTING DATA

The easiest way to reduce the number of data units to k is to choose the first k units encountered. The obvious disadvantage for pictorial data is that some of the real classes present in the data may not be represented in the first k , even if k is large. Therefore, pixels are selected at random or in some fixed spatial pattern which spreads the selected points over the data. Consider, however, the following problems.

1. If the sample is sparse, a real class may not be represented. In particular, prior knowledge about the structure of the data (e.g., the number of classes) may be lost.
2. Some samples may come from classes which are of no interest so that processing resources are wasted.
3. Some (in the case of Landsat data, many) samples may be mixtures of real classes (being on a spatial boundary); iterative clustering programs such as ISODATA are likely to produce clusterings in which a mixture class captures a real class.

Implicit in this discussion are two ideas. The first is very simple—real classes exist. The second idea is

about the pictorial nature of the data—real classes are present in spatial associations (i.e., think fields). In the first part of this paper, a technique for sampling multidimensional pictorial data based on these ideas is presented; in appendix A, a computer program for implementing the technique is described.

Others have considered the problem of selecting representative pixels. For example, Hall et al. (ref. 17) delete pixels from a scene with fewer than four occurrences. A similar scheme has been developed by W. Coberly (private communication). High-frequency pixels are selected for clustering and factor analysis; however, small classes are still likely to be lost. Also, as the dimensionality of the data increases, the histograms become more scattered and the hashing program (by which one accumulates multidimensional histograms) becomes more complex and time consuming. The technique proposed here has the same purpose as histogram-based selection of high-frequency pixels—to select for analysis prototypes which are pure (i.e., not mixture) pixels and which represent each real class.

The third method mentioned for reducing the cost of an iterative clustering program is to start near a "solution" (so that fewer iterations are required). If this procedure can be managed with negligible added computational burden, computer time will be saved. The program described in appendix A for finding initial cluster centers is fast and automatic; the idea of the technique is introduced next.

FINDING FIELDS IN REMOTELY SENSED DATA

Let the multidimensional data vector in row i , column j , be denoted by d_{ij} . Let $|v|$ denote the euclidean length of a vector v (so that $|v - w|$ is the distance between v and w). Suppose there are r rows and c columns; pixels with row index 1 (r) or column index 1 (c) will be called border pixels. Others are said to be inside the scene. Each pixel d_{ij} inside has four nearest neighbors; i.e., left, right, above, and below. In Landsat data, these are probably the only neighbors that matter. In data with better resolution (such as most aircraft-acquired data), many more neighbors can be considered in forming spatial judgments.

An irresistibly interesting problem in pictorial pattern recognition is the boundary detection problem. The approach taken here to finding fields actually defines a set which is almost certain to contain the boundary; spatially connected (ref. 18) sets remain-

ing are called fields. Only the four nearest neighbors are considered in deciding connectedness. Since thin boundaries are not required, a simple one-dimensional gradient thresholding technique (with thresholds set automatically) is used to mark probable boundary points. The thresholds are set so that about one-third of the scene is boundary.

This technique, when tested on real Landsat agricultural data, has been observed to select fields which, on comparison with ground-truth maps, are found to contain representatives from each real class (crop type) and never to include two or more distinct real classes. This experimental evidence supports the following three assumptions.

Assumption 1: Real classes exist—Mark each pixel pair inside the scene which spectrally differs by more than a threshold from its left or above neighbor; set the threshold so that between one-fourth and one-half of the pixels are marked. The complement of the set of marked points is called the set of pure pixels, and the spatially connected components of this set are called fields.

Assumption 2: Each field contains exactly one real class.

Assumption 3: Each real class is presented in at least one field.

The program described in appendix A is based on these assumptions. Once the fields are formed, pixels are selected (roundrobin fashion) until an adequate supply for clustering is obtained. On the assumptions, it is known that

1. Each real class is represented.
2. Each pixel comes from a real class.

Also, real classes present only in small fields are not as likely to be captured by classes present in large fields in the clustering program, as is the case with random selection; this is true, since the roundrobin selection technique will select the same number from each field regardless of the field size.

The process of obtaining starting cluster centers (rather than pixels to cluster) is now outlined. Suppose s starting cluster centers are desired. Form fields as described previously and select 5 pixels from each field; call these pixels test pixels. Form the mean vectors of each field. (In one Landsat segment, there are typically 300 fields.) Classify test pixels by nearest spectral neighbor to the cluster centers and count the number of times a center attracts a test pixel. Eliminate all cluster centers to which no test pixels are assigned. (This step will automatically guarantee distinct cluster centers.) Now eliminate all centers with 1, 2, . . . assignments until s are obtained,

reclassifying test pixels which were assigned to an eliminated center.

THE CLASSIFICATION PROBLEM

Following this starting procedure, a clustering program (designed to cluster arbitrary metric data) will produce a clustering of the sampled data; on exit, the cluster centers will be known. The problem now becomes one of classification (of all pixels). How can one label each pixel in the scene with the correct cluster center? This question has been studied in some depth recently by Bauer et al. (ref. 19). They utilize five different classification algorithms.

1. Maximum likelihood per-point classifier
2. ECHO (refs. 20 to 22) classifier
3. Layered (ref. 23) classifier
4. Minimum distance to the means, per-point classifier
5. A parallelepiped per-point or "levels" classifier

In terms of accuracy, the minimum distance to the means (classifier 4) ranks better (but perhaps not significantly better) than maximum likelihood, and the cost is about one-third as much. All the other classifiers tested fall behind maximum likelihood and cost more than nearest neighbor classification. In the report, the authors express surprise that nearest neighbor classification is more accurate but point out that the training statistics were developed using an unnamed clustering algorithm which used minimum distance assignments; the means, variances, and correlation matrices were then formed assuming multivariate normal distribution for the training data. This explanation is not very convincing. In the opinion of this author, the reason that nearest neighbor assignments are superior to maximum likelihood lies in the basic failure of the assumptions, especially in sample sizes encountered here (i.e., the training classes are not Gaussian).

Another report by Richardson and Pentland (ref. 24) contains a comparison of maximum likelihood against 13 basic methods, 6 of which are actually spatial cleanup operations which follow a maximum likelihood (or, in principle, any) classification. None of these methods (except the maximum likelihood classifier) are mentioned previously; thus, it is difficult to compare the results. However, the "9-point rule" spatial operations classifier always slightly improved the classification accuracy. This

work is difficult to evaluate since the authors reduce the problem to an unnatural two-class problem from the beginning, presumably to give the 9-point-rule classifiers a better chance.

In all this work, the ECHO classifier is unique in its use of spatial information. An earlier technique related to ECHO was developed by Gupta and Wintz (ref. 25); they use hypothesis testing to grow "blobs" with similar gray levels and textures and develop a classification algorithm which performs well on aircraft-acquired MSS data. The main problem (other than computer time) with the application of these ideas to Landsat data really lies in the fact that textural information in a single pixel is meaningless and that 2- by 2-pixel areas are already too large. In addition, it is not credible that an estimate on textural information obtained from a 2- by 2-pixel area is sufficiently significant to warrant the action the program takes (growing a field). Setting the thresholds to prevent propagation of random fields leaves more points which must be handled on a per-point basis. Apparently, the thresholds (which must be set by the user) in the tests of ECHO reported were set to achieve at about two-thirds the cost of maximum likelihood classification with about a 1-percent loss in classification accuracy. This result would be quite impressive if it were not for the nearest neighbor classification, which achieves better accuracy in half the time.

Still, the use of spatial information in classification is a good idea. Van Rooy and Lynn (ref. 26) discuss the use of spatial information for improving the accuracy of a classification. They assume (without defining "field") that

1. The majority of fields contain many data points.
2. The initial classification was accurate.

These assumptions held for C-1 flight data. For Landsat data, the first assumption usually is not valid. What replaces it here results from a simple analysis of boundary pixels.

THE BOUNDARY PIXEL MODEL

Consider two adjacent real fields containing distinct real classes c and d . A pixel b on the real boundary between these two fields will be averaged by the remote-sensing hardware and appear as $b = \alpha c + (1 - \alpha)d$, $0 < \alpha < 1$. However, it is possible (and

likely, if c and d are well separated) that b will actually be spectrally nearer some other class (say class e) so that b is misclassified in class e by any per-point classification. Two surprisingly useful observations can be made from this simple model. The first is clear—apparent boundary classifications should be suspected. (Suspect a pixel which fails to be classified like at least two of its four nearest neighbors, and totally reject one which is unlike all four.) The second is based on looking at the spectral distance from b to c . Since $b - c = (1 - \alpha)(d - c)$, the euclidean distance from b to c is simply $(1 - \alpha)|d - c|$, and therefore the distance from b to the nearer of c or d is not greater than $1/2|d - c|$. This model of a boundary pixel thus leads to a cluster-dependent threshold. If, for each cluster center, a, z is the cluster center with $|a - z|$ largest, then $r(a) = 1/2|a - z|$ is a rejection threshold (for cluster a), and no pixel p should be assigned to a if $|a - p| > r(a)$.

The spatial-spectral classification technique can now be outlined; an associated program is described in appendix B. First determine the rejection threshold for each cluster. Recall that fields were found (in the spatial starting procedure/pixel selection step). Spectrally classify each field by nearest cluster. If the distance to the nearest cluster exceeds the rejection threshold for that cluster, increase the number of clusters by 1 (adding the field mean which was rejected as a new cluster center) and recompute the rejection threshold. When all field means have been classified, begin a spatial map of labels. Classify all unlabeled (and so nonfield point) pixels. Declassify any pixel which has less than two neighbors in the same class (considering only the four nearest neighbors). Examine the four neighbors of each unclassified pixel and find the class (of these four) to which the pixel is nearest but which it has not rejected. If one is found, label the pixel with this label. If none is found, restore the pixel's old label. Now perform a spatial cleanup operation. First, declassify a pixel with no neighbor (of the four nearest) in the same class; then, when three of the four neighbors of a declassified pixel are in the same class, transfer this class label to the pixel. The second operation is similar to what Van Rooy and Lynn (ref. 26) propose, with the following revised assumption.

Many fields have few members, but each field has at least two. It may be that the fuzzy set theory (ref. 27) can be applied here to make

more sense of these assumptions. (This work has not been done.)

THE PHILOSOPHY OF CLUSTERING PICTORIAL DATA

In the book "Patterns of Discovery" (ref. 28), Hanson examines the problem of how an analyst can analyze and construct hypotheses about the data. Although no explicit mention of cluster analysis is made, the book contains many references to the problems of spatial perception. In fact, the book opens with two microbiologists viewing the same image.

Imagine these two observing a Protozoan-Amoeba. One sees a non-celled animal. The first sees Amoeba in all its analogies with different types of single cells Within this class Amoeba is distinguished only by its independence. The other, however, sees Amoeba's homology not with single cells but with whole animals

Although the two view the same image, what they perceive as significant or relevant is not the same. Similar points are stressed by Anderberg (ref. 9, pp. 22 to 24). Polya (ref. 29, p. 110) makes the point simpler and with more generality. "Let us not neglect the obvious and let us note: two people presented with the same evidence may honestly disagree" and (p. 111) ". . . two persons presented with the same evidence and applying the same patterns of plausible inference may honestly disagree." These philosophical observations are important to one interested in clustering, for they bring into question the reality of the information in the data. A specific example in remote sensing (patterned after Hanson's) follows.

Three observers view film of a four-pass Landsat segment taken from a region of agricultural interest. One observer is interested in labelling "wheat vs. other"; that is, the temporal-spectral behavior of wheat and the conditions which prevailed in the place and time this data was acquired (as known to the observer) are used along with spatial associations to separate wheat from "other." The second is interested in yield estimation: although each individual pixel comprises over an acre of real area, this observer is attempting to understand the fine structure of the data so as to predict each pixel's

yield. It may be that differing amounts of soil moisture and variations in soil type or agricultural practices will affect yield, and this observer is attuned to perceive such differences.

A third observer,² interested in but ignorant of agriculture, perceives commonplace spatial features (fields, roads, clouds and cloud shadows and so on), and distinguishes between the passes as being "noisy" or "clean," the fields as being rectangular or round, wide or narrow. This observer sees the clustering problem more as one of filtering: it is desired to clean up the noisy fields, to enhance the fuzzy boundaries and somehow transform four passes to one pleasing and plausible image without losing much information. (Of course, this observer knows a precise mathematical definition of "information" which is probably unrelated to the needs of the first two.)

Let us adopt the third observer's orientation (despite his obvious unsuitability for solving the real problem), believing that a product (a clustering of the data) which pleases him will be useful to the others. His background (not merely temperament, as is pointed out by Polya) leads him to raise questions such as "Given a clustering of the data, which is the probability that a pixel is misclustered?" If this question is meaningful, then it seems that an objective function could be defined and that various clusterings of the data could be compared, with the best of those compared being selected. Unfortunately, the question is actually meaningless. Without external labels, the correctness of a classification of a single pixel has no meaning. (Although schemes exist to transform labels to clusters, a deeper question of the reality of the labels remains.)

Although the absolute real class into which a pixel should be clustered is unknown, there are samples from the same real class; recall the assumption that each field is assumed to contain exactly one real class, so that all pixels from the same field are in the same real class. Consider, therefore, a pair of pixels and a clustering of the data. There are four mutually exclusive possibilities.

1. The pair can be in the same real class and be clustered in the same cluster.

2. The pair can come from two different real classes and be clustered differently.

3. The pair can come from the same real class and be clustered differently.

4. The pair can come from different real classes and be clustered alike.

The last two cases represent errors. The probability of case 3 or 4 will be called the pair probability of misclassification (PPMC). Error case 3 can be estimated using samples from the same real class; a strategy for estimating case 4 will be proposed presently. Clearly, the estimate of the PPMC can be used as an objective function to choose one clustering among many.

AN INTERNALLY SUPERVISED CLUSTERING TECHNIQUE

The pair idea leads to a technique which has the internal structure of a pattern recognition algorithm in the sense of Kaminuma et al. (ref. 30). However, since paradigms are extracted from the data without external supervision, a user sees it as a clustering technique. Others have used internal parameter-free similarity to determine the number of clusters (ref. 31, for example). Furthermore, the use of pairs in measuring similarity between two clusterings was proposed by Rand (ref. 32), and the technique was used to compare clustering programs by Dubes and Jain (ref. 13, p. 260). The novelty of our approach to finding the clusters is that the pairs for supervision are selected from a "perfect" real clustering, and the technique tries to make the clustering equal real clustering.

The actual technique proceeds as follows. Once the fields are formed and labeled by the starting procedure or by some other means (ref. 33), sets of 5 pixels are drawn from each field which has at least 5 elements. The sets of 5 are called test sets. Each test set contains 10 unordered pairs from the same real class. To obtain samples from different real classes, take the family of test sets and rearrange the family (keeping test sets together) so that the values in channel 1 of the first element in each test set are in nondecreasing order. Let there be s test sets and let $n = s/4$. To obtain samples presumed to be from different real classes, select the five from test set k and one each from test set $k - n$ and $k + n$. (Details,

²It is unfortunate that, generally, remote-sensing software is developed by people with mathematical-statistical-engineering backgrounds. The present author (a typical third observer) is no exception.

such as what to do when one of $k \pm n$ test sets is not a valid index, are provided in appendix C.) This procedure furnishes 10 pairs, which probably³ are from different but not too different real classes.

For any given clustering, it is now easy to estimate the PPMC objective function. The selection of clusterings to evaluate is discussed next. Start by using the spatial-spectral starting module to obtain, at most, 200 starting cluster centers. Classify all test pixels and count, for any cluster center, the number of test set pixel pairs which were split plus the number of different field pixel pairs which were not split (these are errors). At the same time, evaluate the estimate on the PPMC. Eliminate the cluster which has the largest number of errors, reclassify test pixels which were assigned to that cluster, and continue until one cluster remains. The clustering with the smallest PPMC wins. (It is interesting that a clustering with but one cluster has a high PPMC, since samples from different real classes are all classed alike.)

A program (named AMOEBA) to implement this technique is described in appendix C.

RESULTS

The new clustering program AMOEBA was tested on four Landsat segments from the U.S. Great Plains. The data, which are described in table 1, were furnished by the NASA Johnson Space Center (JSC). Before entering the program, a transformation somewhat like the Kauth-Thomas transformation (ref. 15) is used to halve the dimensionality of the problem: the linear combinations $(c_1 + c_2 + c_3)/4$ and $(-c_2 + c_3 + c_4)/3$ are roughly the brightness and greenness; since channel 1 is used to sort test sets to obtain pairs from different real classes, the "best" acquisition (here taken to be pass 3) is transformed to channels 1 and 2.

In table II, the execution characteristics of AMOEBA are described. This version of the program includes performing all preprocessing and making one universal formatted image tape for display on NASA JSC hardware. The program is written in Fortran language and compiled using the Fortran H extended optimizing compiler (with OPT = 2). (Of

TABLE I.—Description of Four Test Data Sets

Identification	Acquisition dates ^a	Comments
1857	76073, 76109, 76154, 76191	Data exceptionally free of noise
1865	76073, 76109, 76164, 76190	Typical data
1854	76055, 76154, 76164, 76199	Passes 1, 2, and 3 noisy; typical scan line noise in pass 2
1861	76056, 76128, 76164, 76182	All passes noisy; much of the segment fallow or pasture

^aThe first two digits represent the year, the last three represent the Julian date.

³A more reliable way of selecting pairs from different real classes is needed. This "probably" situation is undoubtedly the weakest and least understood feature of the clustering technique.

TABLE II.—Execution Characteristics of AMOEBA Version 6

Characteristic	Value
Computer used	AMDAHL 470 V/6
Execution time per 117-line by 196-pixel segment, sec	$14.2 + 2.4 \times \text{no. of passes}$
Memory used, kbytes	$200 + 100 \times \text{no. of passes}$
Compilation time, sec	8

TABLE III.—AMOEB A Version 6, Two-Pass Data

Characteristic	Segment			
	1857	1865	1854	1861
Starting no. of clusters	360	337	341	338
No. of test pixels	1500	1430	1270	1285
Final no. of clusters	11	22	12	16
Cluster size				
Largest	5286	4844	7040	7093
Smallest	74	36	75	104
No. of unclassified pixels	11	5	22	11

course, the compilation needs to be done only once.)

In tables III and IV, typical results are shown. The four-pass tests used all the data from table I. Two-pass tests used passes 2 and 3. The results show one rather surprising characteristic: in three of the four tests, fewer clusters were found using four acquisitions than when using two. (Of course, the difference may not be significant.) In segment 1854, however, 37 clusters were found in four-pass data and only 12 in two-pass data.

The comparison of line-printer cluster maps with 1976 ground truth was generally very encouraging. As might be expected, the two-pass tests yielded much less accurate approximations (particularly in segment 1854, where the two acquisitions are only 10

TABLE IV.—AMOEBAs Version 6, Four-Pass Data^a

Characteristic	Segment			
	1857	1865	1854	1861
Starting no. of clusters	350	316	314	312
No. of test pixels	1450	1340	1260	1235
Final no. of clusters	9	21	37	14
Cluster size				
Largest	5939	4085	3806	2834
Smallest	155	125	77	479
No. of unclassified pixels	18	13	53	39

^aThe colors were "taken" from pass 3 and "enhanced"; the work to produce color displays will appear elsewhere.

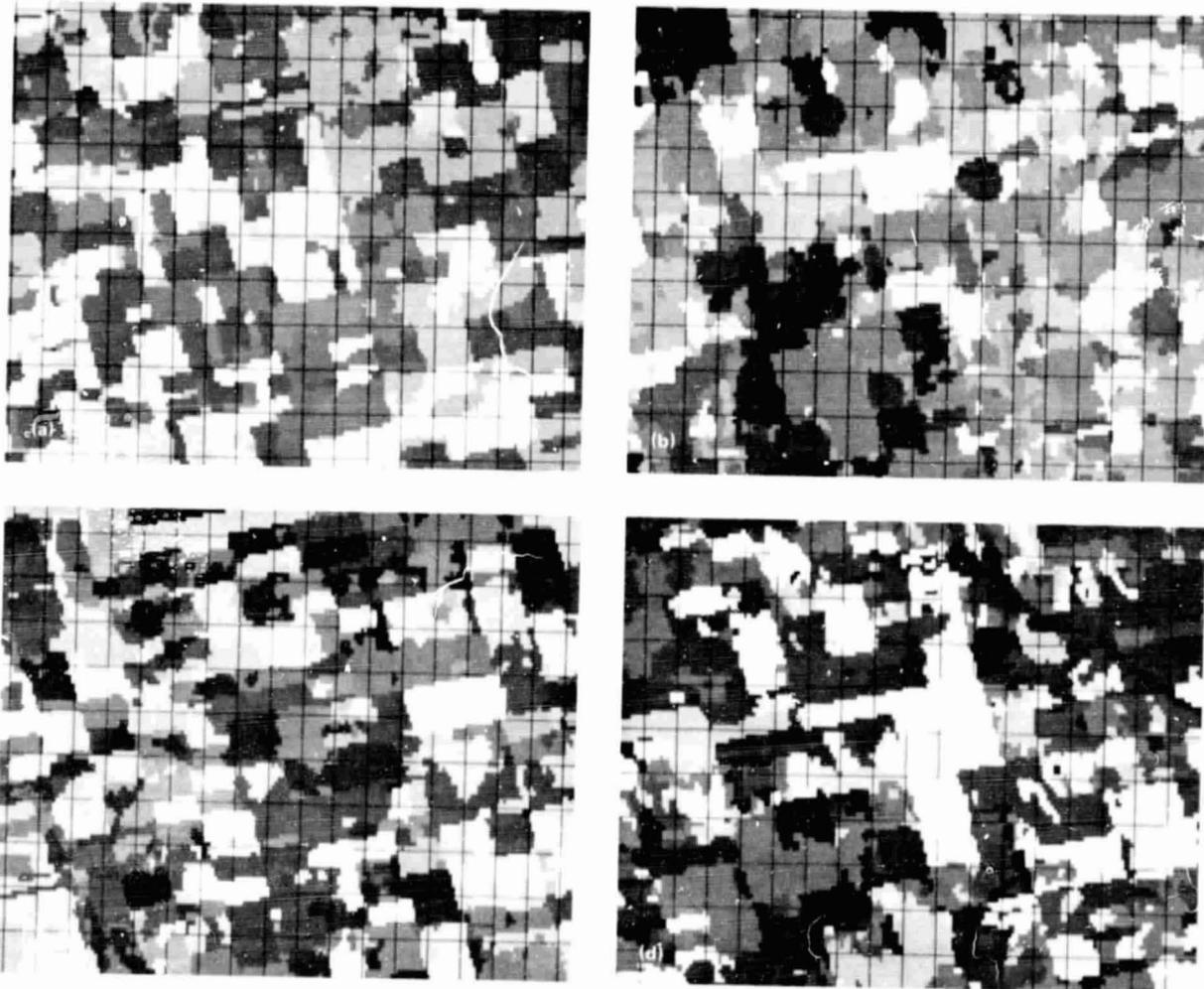


FIGURE 1.—Clustering of four-pass data. (a) Segment 1857. (b) Segment 1865. (c) Segment 1854. (d) Segment 1861.

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

days apart). In figure 1, typical film products of the output of the clustering are displayed.

SUMMARY

Two related aspects of using information from the spatial domain have been presented. The first concerns finding sets of pixels which are spatially and spectrally associated; these sets may be called "fields." Three assumptions on the relation between (unknown) real classes and fields are formalized. On these assumptions, one is able to automatically extract samples of the data known to represent each real class; in addition, popular field mean vectors are proposed as starting cluster centers for an iterative clustering program to process.

The second use of spatial information concerns the problem of classifying mixture pixels. A simple model for the boundary between two real fields leads to logic for detection of probable erroneous classifications. The fields found by the starting procedure summarized previously are classified as monoliths; and all other pixels are classified, yielding a map of labels. Any pixel not having at least two neighbors in the same class is suspected of being incorrectly classified; the boundary model furnishes a method for reclassifying such pixels.

Incidental to this work, it was noticed that samples of the data which probably come from different real classes could be obtained. First, the samples could be ordered from the fields on some (essentially arbitrary) one-dimensional attribute. Selected pairs spread out in this order can be believed to come from different real classes. Using these pairs, together with the samples from the same field (and thus real classes), one can evaluate an arbitrary clustering of the data by estimating the probability that a pair from the same real class is clustered differently plus the probability that a pair from different real classes is clustered alike. At the same time, clusters which are most involved in splitting pairs from the same real class or which have the least discriminatory ability for pairs from different real classes are identified and eliminated (thereby reducing the number of clusters). The clustering with the lowest pair-misclustering probability is selected.

Based on these ideas, the clustering program AMOEBA is described. Internally, the program is a pattern recognition program; but, from without, it appears to be an unsupervised clustering program.

The program is fast and automatic; thus, no choices (such as arbitrary thresholds to set or split/combine sequences) need be made. The difficult problem of finding the number of clusters is solved automatically. At the conclusion of the program, all points in the scene are classified; however, a provision is included for a "reject" classification of some points which, within the theoretical framework, cannot rationally be assigned to any cluster.

ACKNOWLEDGMENTS

L. F. Guseman, Jr., read portions of the manuscript, and many others attended presentations of these ideas in various forms. I found their criticism to be especially helpful. The friendly guidance of G. Dibrell, B. Hedges, and L. Swanson is noted. The research support provided by NASA over the past year (NASA contract NAS-9-14689-85) gave me sufficient free time to develop the basic ideas; it is a pleasure to acknowledge this assistance.

REFERENCES

1. Ball, G. H.; and Hall, J. D.: ISODATA, A Novel Method of Data Analysis and Pattern Classification. Tech. Rep., Stanford Research Institute (Menlo Park, Calif.), 1965.
2. Ball, G. H.; and Hall, J. D.: A Clustering Technique for Summarizing Multivariate Data. *Behavioral Sci.*, vol. 12, 1967, pp. 153-155.
3. Fromm, F. R.; and Northouse, R. A.: CLASS: A Non-parametric Clustering Algorithm. *Pattern Recognition*, vol. 8, 1976, pp. 107-114.
4. Kan, E. P. F.: ISODATA: Thresholds for Splitting Clusters. Tech. Rep. 640-TR-058, Lockheed Electronics Co., Inc. (Houston, Tex.), 1972.
5. Kan, E. P. F.; and Holley, W. A.: Experience With ISODATA. Tech. Memo. TM 642-354, Lockheed Electronics Co., Inc. (Houston, Tex.), 1972.
6. Kan, E. P. F.; and Holley, W. A.: More on Clustering Techniques With Final Recommendations on ISODATA. Tech. Rep. LEC 640-TR-112, Lockheed Electronics Co., Inc. (Houston, Tex.), 1972.
7. Wacker, A. G.; and Landgrebe, D. A.: The Minimum Distance Approach to Classification. Information Note 100771, Laboratory for Applications of Remote Sensing (LARS), Purdue University (W. Lafayette, Ind.), 1972.

8. Duran, B. S.; and Odell, P. L.: Cluster Analysis. Lecture Notes in Economics and Mathematical Systems, vol. 100. Springer-Verlag (New York), 1974.
9. Anderberg, M. R.: Cluster Analysis for Applications. Academic Press (New York), 1973.
10. Ball, G. H.: Classification Analysis. Tech. Note AD 716 482. Stanford Research Institute (Menlo Park, Calif.), 1970.
11. Wright, W. E.: Gravitational Clustering. Pattern Recognition, vol. 9, 1977, pp. 151-166.
12. MacQueen, J. B.: Some Methods for Classification and Analysis of Multivariate Observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, 1967, pp. 281-297.
13. Dubes, R.; and Jain, A. K.: Clustering Techniques: The User's Dilemma. Pattern Recognition, vol. 8, 1976, pp. 247-260.
14. Andrews, H. C.: Introduction to Mathematical Techniques in Pattern Recognition. Wiley-Interscience (New York), 1972.
15. Kauth, R. J.; and Thomas, G. S.: The Tasseled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. LARS Symposium, Machine Processing of Remotely Sensed Data, Purdue University (W. Lafayette, Ind.), 1976.
16. Wheeler, S. G.; Misra, P. N.; and Holmes, Q. A.: Linear Dimensionality of Landsat Agricultural Data. LARS Symposium, Machine Processing of Remotely Sensed Data, Purdue University (W. Lafayette, Ind.), 1976.
17. Hall, R. N.; McGuire, K. G.; and Bland, R. A.: Landsat Signature Development Program. LARS Symposium, Machine Processing of Remotely Sensed Data, Purdue University (W. Lafayette, Ind.), 1976.
18. Haralick, R. M.: Glossary and Index to Remotely Sensed Image Pattern Recognition Concepts. Pattern Recognition, vol. 5, 1973, pp. 391-402.
19. Bauer, M. E.; Silva, L. F.; Hoffer, R. M.; and Baumgardner, M. F.: Agricultural Scene Understanding. LARS Contract Rep. 112 677, LARS, Purdue University (W. Lafayette, Ind.), 1977.
20. Ketting, R. L.: Computer Classification of Remotely Sensed Multispectral Image Data by Extraction and Classification of Homogeneous Objects. Ph. D. Thesis, Purdue University (W. Lafayette, Ind.), 1975.
21. Swain, P. H.: Land Use Classification and Mapping by Machine-Assisted Analysis of Landsat Multispectral Scanner Data. LARS Information Note 111276, Purdue University (W. Lafayette, Ind.), 1976.
22. Kast, J. L.; and Davis, B. J.: Tests of Spectral/Spatial Classifier. LARS Contract Rep. 112717, LARS, Purdue University (W. Lafayette, Ind.), 1977.
23. Wu, C. L.; Landgrebe, D. A.; and Swain, P. H.: The Decision Tree Approach to Classification. LARS Information Note 090174, LARS, Purdue University (W. Lafayette, Ind.), 1974.
24. Richardson, W.; and Pentland, A. P.: Evaluation of Algorithms for Estimating Wheat Acreage From Multispectral Scanner Data. ERIM 109600-69-F, Environmental Research Institute of Michigan (Ann Arbor, Mich.), 1976.
25. Gupta, J. N.; and Wintz, P. A.: Multi-Image Modelling. TR-EE 74-24, School of Electrical Engineering, Purdue University (W. Lafayette, Ind.), 1974.
26. Van Rooy, D. L.; and Lynn, M. S.: Use of Spatial Information in Classification of Remotely Sensed Data. Institute for Computer Services and Applications, Rice University (Houston, Tex.), 1973.
27. Zadeh, L. A.: Fuzzy Sets. Information and Control, vol. 8, 1965, pp. 338-353.
28. Hanson, N. R.: Patterns of Discovery. Cambridge University Press (Cambridge), 1958.
29. Polya, G.: Mathematics and Plausible Reasoning, Vol. II, Patterns of Plausible Inference. Princeton University Press (Princeton, N.J.), 1954.
30. Kaminuma, T.; Takekawa, T.; and Watanabe, S.: Reduction of Clustering Problem to Pattern Recognition. Pattern Recognition, vol. 1, 1969, pp. 194-205.
31. Gitman, I.: A Parameter-Free Clustering Model. Pattern Recognition, vol. 4, 1972, pp. 307-315.
32. Rand, W. M.: Objective Criteria for the Evaluation of Clustering Methods. J. American Statist. Assoc., vol. 66, 1971, p. 846.
33. Kauth, R. J.; Pentland, A. P.; and Thomas, G. S.: BLOB, An Unsupervised Clustering Approach to Spatial Preprocessing of MSS Imagery. Eleventh International Symposium on Remote Sensing of Environment, Vol. 2, ERIM (Ann Arbor, Mich.), 1977, pp. 1309-1317.

Appendix A

A Spatial Starting Procedure

Notation: Suppose the data to be analyzed are in an array $d_{i,j}$ of vectors, $i = 1, \dots, r$ and $j = 1, \dots, c$. Also available is an array $l_{i,j}$ for saving labels.

Step 1. Cover the boundary.

a. Let $u = rc/2$, $m = rc/3$, and $l = rc/4$. Sum the squared distance $|d_{i,j} - d_{i,j+1}|^2$ over every fifth row and every fifth column and divide by the number summed, obtaining an estimate of the intrinsic variability n of the data. Set $h = 2n$; h is the horizontal tolerance.

b. Set $v = 13h/10$; v is the vertical tolerance. Initialize the array $l_{i,j}$ to all zeros. For column $j = 2, \dots, c$ and row $i = 2, \dots, r$, compare $|d_{i,j} - d_{i,j-1}|^2 > h$; if so, set $l_{i,j} = 1$ and $l_{i,j-1} = 1$. Then test $|d_{i,j} - d_{i-1,j}| > v$; if so, set $l_{i,j} = 1$ and $l_{i-1,j} = 1$. When all pixels have been processed in this way, pass through the inside of the array $l_{i,j}$ and fill in the holes. If $l_{i,j} = 1$, skip; otherwise, test the neighbor above and below, seeing if both are marked (i.e., = 1) and, if not, test the left and right neighbors. If either pair is marked, set $l_{i,j} = 1$. Count the number n of points with $l_{i,j} = 1$. If $u > n > l$, exit to Step 2; otherwise, replace h by $(h + h \cdot mn)/2$ and repeat Step b. Call pixels at (i,j) with $l_{i,j} = 0$ pure pixels.

Step 2. Label the fields and extract test sets.

Comment: In this step, connected sets of pure pixels called fields are marked in the array $l_{i,j}$. However, the maximum field size is limited to 50 pixels. At the same time as fields are being formed, 5 test pixels are extracted from each field which contains at least 5 and are stored in an array of test sets. The slightly obscure program to perform the field labeling is actually an efficient maze-solving algorithm.

a. Initialize a stack S of fifty 6-vectors by setting $S(D,p) = 2$ for $D = 1,2,3,4$ and $p = 1, \dots, 50$. Initialize a number-of-test-sets counter to $t = 0$ and a number-of-fields counter to $f = 0$. Set the row counter to $i = 2$.

b. Let the column counter be $j = 2$.

c. If $l_{i,j} \neq 0$, the pixel at (i,j) is not a pure pixel, so proceed to Step 2m. Otherwise, initialize a stack pointer p and a stack index x both to zero and set $l_0 = l_{i,j} = j$, and $d = 0$.

d. Set $l_{0,p} = 1$. If $p = 50$, go to Step 2i. Otherwise, set $p = p + 1$ and save the coordinates (l_0, j_0) in $S(5,p)$ and $S(6,p)$.

e. Search directions $D = 1$ (left), $D = 2$ (right), $D = 3$ (up), and $D = 4$ (down) except for direction d ; if the location searched is off the scene, go to Step 2f. If it is not, examine the label at the point. If the label is negative, go to Step 2k. If it is positive, go to Step 2f. Otherwise, change the label at the location to 1, set $S(D,p) = 0$, and go to Step 2g.

f. Set $S(D,p) = 1$.

g. Search the next direction (repeating Step 2e). When all directions have been searched, locate the pixel pointed to by the stack index x : the pointer is to location $(u + 1, v)$, where $u \equiv x \pmod{4}$ and $v = \lfloor x/4 \rfloor + 1$. The search is performed by incrementing x until $x \geq 200$ (Step 2i); $S(u + 1, v) = 0$ (Step 2h); $S(u + 1, v) = 1$ (next x) or $S(u + 1, v) = 2$ (Step 2i).

h. Set $x = x + 1$; set $d =$ the reflection of direction $u + 1$, and construct the address (l_0, j_0) , which is the stacked pure pixel location pointed to by $u + 1$ modifying $[S(5,v), S(6,v)]$. Go to Step 2d.

i. All pure pixels (up to 50) in the current component have been marked; they must now be labeled. If $f \leq -400$, exit to Step 2n. Otherwise, set $f = f - 1$ and set $L = f$.

j. Go through the $s = 1$ to p points stacked and set $l_{S(5,s), S(6,s)} = L$; at the same time, restore $S(D,s) = 2$ for $D = 1,2,3,4$. Proceed to Step 2l.

k. Set $L =$ the negative label found (in Step 2e). Go to Step 2j.

l. If $t > 1495$ or if $p < 5$, go to Step 2m. Otherwise, set $q = (p - 1)/4$, $k = 1 + nq$ ($n = 0, \dots, 4$), and save the set of 5 test pixels (one test set) by moving data in location $[S(5,k), S(6,k)]$ to the test set array. Set $t = t + 5$.

m. Set $j = j + 2$; if $j < c$, repeat Step 2c. Otherwise, set $j = 2$, $i = i + 2$. If $i < r$, repeat Step 2b; otherwise, exit to the next step.

n. The $t/5$ test sets of 5 pixels (from the same field) and $m = -L$ labeled fields have been formed. Note that $t \leq 1500$ and $m \leq 400$.

o. In another pass through the data and labels, form the sum, the count, and then the means of each field. Let the mean vectors be $\mathbf{M}_i, i = 1, \dots, m$.

Step 3. Select the cluster centers. Copy the m field means to cluster center vectors $\mathbf{C}_i, i = 1, \dots, m$, and initialize $k_i = 1, i = 1, \dots, m$. Set $m' = m$. Suppose that at most, n starting cluster centers are desired and that no duplicates are wanted.

a. Classify all t test pixels by nearest cluster center (euclidean distance); count the number t_i of times cluster center \mathbf{C}_i is the object of a classification. Eliminate any cluster with $t_i = 0$ (by setting $k_i = 0$), decrementing m' each time this happens. (Incidentally, this will eliminate duplicates in the family of cluster centers.) Set $e = 0$.

b. Replace e by $e + 1$ and set $i = 1$.

c. If $t_i \neq e$, skip to Step 3d; otherwise, set $t_i = 0$ and reassign test pixels previously classed in cluster \mathbf{C}_i , updating the counter for centers which receive

new classifications. Set $m' = m' - 1$ and exit to Step 3e if $m' \leq m$.

d. Set $i = i + 1$; if $i > t$, set $e = e + 1$ and $i = 1$. Repeat Step 3c.

e. The $m' \leq n$ starting clusters have been found. Rearrange them to be $\mathbf{C}_1, \dots, \mathbf{C}_n$ and set $n = m'$.

Usage: The following options apply.

Option 1. The user specifies the number N of representative (pure) pixels desired; Steps 1 and 2, parts a through n are executed. Then, pixels are picked roundrobin from the test sets until either N have been picked or all t have been picked. (Emit a warning if $N < t/5$.)

Option 2. The number n of starting cluster centers is specified; Steps 1 through 3 are performed.

Clearly, Options 1 and 2 can be combined. A third use of the starting procedure is mentioned in appendix C.

Appendix B

A Spatial Classification Program

In this program, it is assumed that the program in appendix A has been executed, giving fields (in a map of labels l_j) and field means (in an array $\mathbf{M}_i, i = 1, \dots, m$). Also, cluster centers $\mathbf{C}_i, i = 1, \dots, n$, are given (by an intervening clustering program). Should no fields be available, the performance of the program will be degraded but it will still work.

Step 1. Determine a rejection threshold and classify and label fields.

a. For $i = 1, \dots, n$, define $r_i = |\mathbf{C}_i - \mathbf{C}_j|^2/4$ where $|\mathbf{C}_i - \mathbf{C}_j| \geq |\mathbf{C}_i - \mathbf{C}_k|$ for $k = 1, \dots, n$.

b. Classify each field k by nearest neighbor to a cluster center \mathbf{C}_i ; if $|\mathbf{M}_k - \mathbf{C}_i|^2 > r_i$, increase the number n of clusters by 1, set $\mathbf{C}_n = \mathbf{M}_k$ and repeat Step 1a.

c. Go through the map of labels l_j and replace a field label (negative integer) by the cluster of the field k , a boundary label (1) by 0.

Step 2. Label the remainder of fields and check boundaries.

a. Classify all unlabeled pixels by nearest unre-

jected neighbor; transfer the classification to the label map.

b. Declassify any pixel which has less than two of its four nearest neighbors in the same class. (This process is order-dependent; it is recommended that the testing go down rows.) Declassify the pixel by replacing the label by its negative.

c. Proceed down columns and, for each declassified pixel, reclassify it by finding the nearest unrejected class of the four spatially nearest neighbors. If this attempt to reclassify the pixel fails, restore the original classification.

d. Again, go through the map of labels by columns and label with zero any pixel whose four nearest neighbors do not have the same label.

e. The preceding steps have been rather conservative. It may be desirable at this point to restore some classifications. One technique is to label an unlabeled pixel with the class of a majority of its neighbors. For most accurate maps, the three-of-four neighbors criterion is recommended.

Appendix C

AMOEBA: A Spatial Clustering Program

Step 0. Preprocess the data, perhaps reducing the dimensionality. Make sure that information most likely to separate clusters of interest is contained in channel 1 of the data vector.

Step 1. Set $M = 200$ and execute the program of appendix A. On exit, a map ℓ_j of labels is passed to Step 3. The n cluster centers C_i and t test pixels (in test sets of 5 each) are used here. Sort the test sets so that the value of the first test pixel in each test set in channel 1 is nondecreasing. (Keep the test sets together, of course.)

Step 2. Determine the number of clusters and what they are.

a. There are n starting cluster centers C_1, \dots, C_n . Let $k_i = 1, i = 1, \dots, n$. Classify all test pixels t_j by nearest cluster center; let L_j denote the label of the cluster center t_j that is nearest. Let $\text{min} = 100\,000, i = n$.

b. Set $n_s = 0, n_a = 0, j = 1$, and $s_h = 0, h = 1, \dots, n$.

c. Examine the classification of test set $\{t_j, \dots, t_{j+4}\}$. Let $j_1 = j + //4$; if $j_1 > t$, let $j_1 = j - //2$. Also, let $j_2 = j - //4$; if $j_2 < 1$, let $j_2 = j + //2$. Compare $L_{j_1}, \dots, L_{j_1+4}$ with L_{j_2} and L_{j_2+4} ; for $\mu = j_1, \dots, j_1 + 4$, if $v = L_\mu \neq L_{j_2}$, set $n_s = n_s + 1$ and $s_v = s_v + 10$; repeat for L_{j_2} . Then, compare all 10 distinct test set classification pairs: each time a pair is split (an event which involves two cluster centers), set $n_a = n_a + 1$ and (for each of the two clusters μ, v) set $s_\mu = s_\mu - 10, s_v = s_v - 10$. Finally, if a test set pair is (correctly) clustered in the same cluster μ , set $s_\mu = s_\mu + 1$. Set $j = j + 5$. If $j < t$, repeat Step 2c; otherwise, go to the

next step.

d. Find the cluster C_μ with $k_\mu \neq 0$ having s_μ minimum and set $k_\mu = 0$; save the index μ just eliminated. Compare $n_a - n_s$ with min : If $n_a - n_s < \text{min}$, set $\text{min} = n_a - n_s$ and set $i_s = i$. Set $i = i - 1$. If $i > 2$, repeat Step 2b; otherwise, continue with Step 2e.

e. The number of clusters is i_s (except for clusters added by Step 3); go through the cluster centers $C_i, i = 1, \dots, n$, and eliminate those with saved indices μ encountered before i_s were obtained. Rearrange the cluster centers for more efficient search.

Step 3. Set $n = i_s$; there are n cluster centers C_1, \dots, C_n . Recall the fields were labeled in ℓ_j . Execute the classification program in appendix B.

Step 4. A user-dependent step: various display products and statistical summaries can be performed now. The version described in this paper prints (on the line printer) a cluster map and a summary of counts of the number of elements in each cluster. Also, a universal image formatted magnetic tape is written.

Usage: Three options are available.

Option 1. In the preferred mode, the program finds the number of clusters.

Option 2. In a possibly suboptimal mode, the program seeks $\geq n$ clusters.

Option 3. In a probably suboptimal mode, the program seeks exactly n clusters.

These options are easily implemented and are incorporated in Version 6 of AMOEBA.

On Evaluating Clustering Procedures for Use in Classification*

M. D. Pore,^a T. E. Moritz,^a D. T. Register,^a S. S. Yao,^a and W. G. Eppler^b

ABSTRACT

Evaluation of clustering as a preprocessing step in classification is discussed. Special emphasis is given to the case in which a limited number of labeled samples are available for the evaluation. An estimated probability of correct classification and a variance of proportion estimate (a measure of cluster purity) are proposed. Three cluster-labeling techniques are described; two are presented in an application and one is theoretically developed to measure labeling errors on a per-cluster basis.

1. INTRODUCTION

In highly complex classification problems, a method of partitioning the samples into subpopulations or clusters and labeling each subpopulation is sometimes used. Each member of a subpopulation then assumes the label of that subpopulation. Errors may occur from two sources: (1) the labeling of the subpopulations may be inaccurate and (2) the partitioning into subpopulations may not be pure.¹ In the latter case, the question of the appropriateness of a subpopulation label must be considered. In this paper, the problem of evaluating clustering algorithms and their respective computer programs for use in this type of classification procedure is ad-

ressed. The major problem in cluster evaluation is the determination of a measure of excellence. However, in clustering for classification, the probability of correct classification (PCC) immediately is suggested as the ultimate measure of accuracy on training data. A means of implementing this criterion and a measure of cluster purity are discussed in section 2, and examples are presented in section 4. A procedure for cluster labeling that is based on cluster purity and sample size is presented in section 3.

Throughout the paper, a two-class classification problem is assumed; however, much of this development is readily applicable to the general classification problem.

2. CLUSTER EVALUATION CRITERIA

Clustering algorithms and their respective computer programs group data points according to characteristics of the respective points. For example, some algorithms group points that are numerically similar, whereas others weight numerical proximity with the spatial proximity of where the data point was observed. Regardless of the grouping philosophy employed by the algorithm, the question of cluster effectiveness in classification is valid and has been inadequately developed; in other words, criteria are needed for the comparison of clustering algorithms. Two such criteria are presented in the following paragraphs with a discussion of theoretical considerations concerning the merit of each. Although computer programs do not always represent an optimal implementation of an algorithm (i.e., defining an optimal implementation criterion is another problem), no distinction will be made between an algorithm and a program for implementing the algorithm. Hence, effectively, clustering programs will be compared. An evaluation of programs is

*This material was developed under NASA Contract NAS 9-15200 and prepared for the Earth Observations Division, NASA Johnson Space Center, Houston, Texas.

^aLockheed Electronics Company, Inc., Houston, Texas.

^bLockheed Missiles and Space Company, Palo Alto, California.

¹The word "pure" is used to imply that all elements of the cluster are samples from the same generic class of objects.

recommended for eliminating errors and inefficiencies caused by poor programming.

The PCC criterion for cluster evaluation in classification is theoretically optimal in that it is exactly the criterion which determines the accuracy of a classification. However, an environment will be assumed where there is an abundance of data for which the true classification is unknown, whereas the true classification is known for a relatively small subset of data. This creates a two-sided evaluation problem. If only the labeled data (the subset for which the true classification is known) are used for the evaluation process, then the clustering algorithm will not be operating in a typical environment (too few data points); on the other hand, if all the data points are used in clustering, the true cluster structure and, therefore, the true cluster label will be unknown. Either evaluation design introduces a source of procedural error into the evaluation. The problem with the first evaluation technique is that its error source is more fundamental; by not clustering in a typical environment, the procedure is systematically biased toward the peculiarities of the clustering algorithm operating on a small data set. The second evaluation technique will be addressed here, and the problem of error in cluster labeling will be dealt with as a problem in statistical estimation. As a general rule, experimental conditions must be identical to performance conditions for effective evaluations.

The problem of measuring classification accuracy is dependent on the errors in labeling clusters and the applicability of labels to mixed clusters.² To accomplish this labeling in an unbiased manner, three techniques have been developed. The first is to select the labeled sample nearest the cluster mean. (Either l_1 or l_2 metrics are usually used.) This technique is referred to as the nearest neighbor labeling (NNL) procedure and is favored because of its ease of automation. The second technique is to observe all labeled samples that fall in a cluster and follow a majority rule (MR) procedure. This technique requires two default procedures characterized by the following examples.

- a. In the event of a tie, the sample farthest from the sample mean will be omitted from consideration.
- b. If no labeled samples fall in a cluster, the

cluster is labeled by the NNL procedure.

This technique will presumably minimize the probability of mislabeling a cluster. This probability of mislabeling a cluster is a function of the number of labeled samples; hence, there is a trade-off in accuracy of cluster labeling and expense of labeling samples. This technique is more difficult to implement in an automated computer-oriented procedure and requires many more labeled samples to effect a measurable difference in labeling accuracy. The third technique consists of labeling samples sequentially within each cluster until a labeling confidence of pre-designated accuracy is achieved. This technique is developed in section 3.

A word of warning is required about labeling samples. If the labeled samples are not proportionally representative of the population of samples (with respect to classification categories), then labeling biases will influence cluster labels and give rise to errors other than those due to sampling variance. Therefore, randomized sampling schemes should be employed to select samples for labeling.

Although PCC is the most direct measure of accuracy in the clustering-classification procedure, it does not measure the cluster purity or the adaptability of the technique to relaxation of usual procedures, to bias sampling, to incrementing of cluster parameters, etc. As a pathological example of how PCC confounds cluster purity with cluster labeling accuracy, consider a two-category case where every cluster is labeled (by whatever method) as belonging to the first category. Then, given an equal number of samples from each category, the PCC is 0.5, regardless of whether the clusters are extremely mixed (i.e., each cluster has exactly 50 percent samples from the first category) or relatively pure (i.e., one-half the clusters have exactly 25 percent, say, from the first category, and the other one-half have exactly 75 percent from the first category).

Relatively pure clusters are thought to lend to the procedure a stability or low variance of error that cannot be achieved by proportional labeling of mixed clusters and also to lend credibility to the concept of clustering as an effective partitioning procedure before classification. The variance (VAR) of cluster proportion is proposed as a measure of cluster purity. Precise definitions of PCC and VAR follow. Let

- a. N_i denote the number of samples in cluster i
- b. M_i denote the number of labeled samples in cluster i
- c. P_i denote the proportion of labeled samples in cluster i which were labeled correctly

² "Mixed" clusters are clusters that are not pure; that is, all elements of the cluster are not samples from the same generic class.

Then, by definition,

$$PCC = \frac{\sum_i P_i N_i}{\sum_i N_i}$$

and

$$VAR = \sum_i \left(\frac{N_i}{\sum_j N_j} \right)^2 \left[\frac{P_i (1 - P_i)}{M_i - 1} \right]$$

The VAR is simply the variance of a proportion estimator stratified across clusters, where the proportion is that percentage of each cluster that is labeled correctly. This statistic is independent of the cluster-labeling procedure or the accuracy of the label and reflects cluster purity weighted by the number of labeled samples in the clusters. The weighting comes from the $M_i - 1$ term and magnifies the weight given to clusters with very few labeled samples, but this magnification has not been precisely analyzed.

3. ESTIMATING ERRORS FOR CLUSTER LABELING

3.1 Introduction

The following development is a fairly general Bayesian model for calculating the probability of correctly labeling a cluster by randomly selecting and labeling a subset of size n of that cluster. This model can be relaxed to samples "near" the cluster if it is assumed that those samples near the cluster make up a subpopulation of proportions identical to those within the cluster.

The purpose of a Bayesian development to cluster labeling is to apply prior experience (on similar data) as to frequencies of various cluster purities to current labeling. This prior information is necessary, as will be seen, to provide a probability confidence that the cluster is labeled correctly; or, conversely, it may provide thresholds on the proportions of observed categories to determine a necessary sample size for "confident" labeling in a sequential labeling procedure.

The model is given in sections 3.2 and 3.3. Section 3.4 gives the general example of a symmetric, quadratic prior density. Section 3.4, step a, contains four cases that demonstrate the generality of this example; and section 3.5 consists of specific solutions for $n = 1$ and $n = 2$. The reader may wish to substitute values of c from section 3.4, step a, into section 3.5 to see the effect of different a priori densities.

An itemized format is used in this section to facilitate later referencing.

3.2 Notation

Let θ_i denote the true proportion of category 1 in cluster i ; $0 \leq \theta_i \leq 1$. Since clusters are dealt with individually and identically, the subscript will be dropped; $\theta = \theta_i$. Further notation follows.

a. $n = M_i$ is the number of labeled samples in cluster i .

b. $x = x_i$ is the number of category 1 labeled samples in cluster i .

The cluster purity θ is treated as a random variable to reflect the fact that clusters assume particular purities with ascertainable frequencies. Also let $g(\theta)$ denote the generalized (possibly discrete) probability density function (p.d.f.) of θ . The p.d.f. g represents a priori information about the capability of the algorithm to generate "pure" clusters and will probably have to be estimated from empirical studies. The fact that different clustering algorithms produce clusters of different purities is reflected in the differing values of g for each algorithm.

3.3 Decision Rule Development

This section will establish a decision rule in its most general form for labeling clusters.

a. Assume each sample from cluster i is independent and is in category 1 with probability θ (a Bernoulli process).

b. It follows from the assumption in step a that x is binomially distributed with parameters θ and n .

$$x \sim f(x|\theta)$$

$$= \binom{n}{x} \theta^x (1 - \theta)^{n-x} I_{\{0,1,\dots,n\}}(x)$$

c. Interest is in the posterior p.d.f. of θ , the proportion or probability of category 1.

$$f(\theta|x) = \frac{g(\theta)f(x|\theta)}{f(x)}$$

d. The posterior probability that the cluster i is in category 2 is

$$\begin{aligned} p &= p \left[0 \leq \theta \leq \frac{1}{2} \right] \\ &= \int_0^{\frac{1}{2}} f(\theta|x) d\theta \\ &= \frac{1}{f(x)} \int_0^{\frac{1}{2}} f(x,\theta) d\theta \end{aligned}$$

and $1 - p$ is the probability that cluster i is in category 1.

e. Decision rules

1. If $x = n/2$, regardless of cluster label, $p = 1/2$.
2. If $x > n/2$, the cluster is labeled category 1 and p is the probability of commission (classifying category 2 into category 1) in cluster labeling and $1 - p$ is the probability of correctly labeling a category 1 cluster.
3. If $x < n/2$, the cluster is labeled category 2 and p is the probability of correctly labeling a category 2 cluster and $1 - p$ is the probability of omission (classifying category 1 into category 2) in cluster labeling.

3.4 General Example

Assume g is a symmetric, quadratic a priori density for θ (i.e., the mathematica! calculation is easier if this assumption is made); then, $g(0) = g(1)$ and $g(\theta) = a\theta^2 + b\theta + c$.

a. The fact that g is a p.d.f. implies

$$g(\theta) = (6c - 6)\theta^2 + (6 - 6c)\theta + c \text{ and } 0 \leq c \leq 3.$$

Now, four special cases are examined by varying the parameter c .

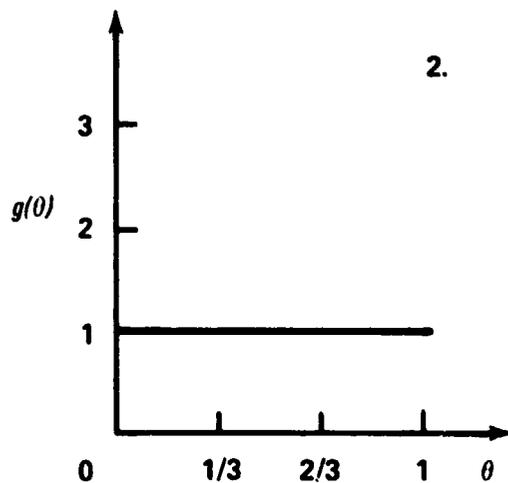
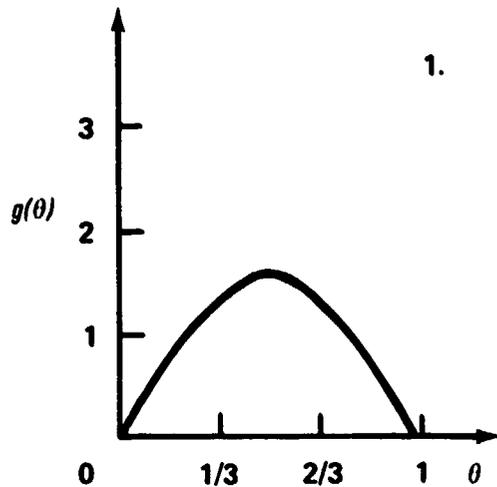
1. ($c = 0$) $\rightarrow g$ is concave downward; cluster algorithm gives mixed clusters only (i.e., $-$ means implied).

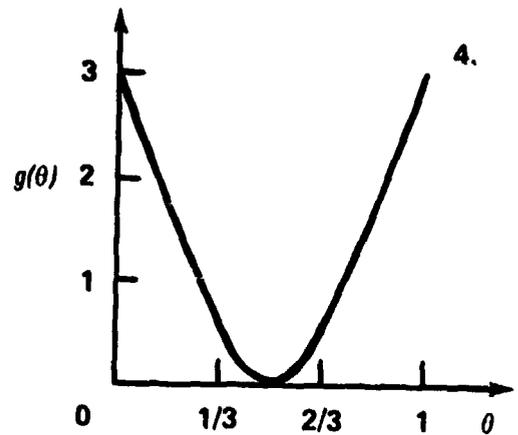
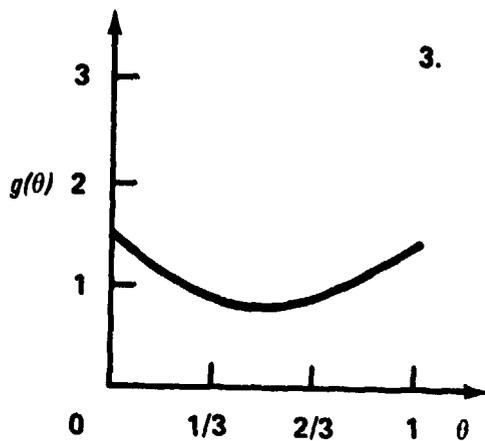
2. ($c = 1$) $\rightarrow g(\theta) = I_{[0,1]}(\theta)$; equi-ignorance principle.

3. ($c = 3/2$) $\rightarrow g(0) = 2g(1/2)$ and $g(\theta) = 3\theta^2 - 3\theta + 3/2$.

4. ($c = 3$) $\rightarrow g(1/2) = 0$ and $g(\theta) = 12\theta^2 - 12\theta + 3$.

These equations are illustrated as follows.





b. The unconditional density of x can be shown to be

$$f(x) = \frac{6(c-1)(x+1)(x+2) + 6(1-c)(x+1)(n+3) + c(n+2)(n+3)}{(n+1)(n+2)(n+3)}$$

c. The probability that the majority of the cluster is in category 2 is

$$p = \frac{4[6(1-c)(x+1)(n+1-x) + c(n+2)(n+3)] \binom{n-x}{s=0} \binom{n+1}{s} + \binom{n}{x} [6(c-1)(n+1)(n-2x)]}{2^{n+3} [6(1-c)x(n-x) + cn(n-1) + 6(n+1)]}$$

3.5 Specific Example

Recall that n is the number of labeling pixels and x is the number of category 1 labeled samples;
 $0 \leq x \leq n$.

a. Specific case $n = 1$

1. If $x = 0$, $p = \frac{c+11}{16}$.

2. If $x = 1$, $p = \frac{5-c}{16}$.

b. Specific case $n = 2$

1. If $x = 0$, $p = \frac{23c+117}{2^4(c+9)}$.

2. If $x = 1$, $p = 1/2$.

3. If $x = 2$, $p = \frac{27-7c}{2^4(c+9)}$.

3.6 An Additional Development

The posterior Bayes estimate of θ , the proportion of category 1 in the cluster, is

$$\hat{\theta} = E(\theta|x)$$

$$= \frac{1}{f(x)} \int_0^1 \theta g(\theta) f(x|\theta) d\theta$$

a. For the model given in steps a through d in section 3.3 and the general a priori assumption of step a in section 3.4, it follows that

$$\hat{\theta} = \left(\frac{x+1}{n+4} \right) \left[\frac{6(c-1)(x+2)(x+3) + 6(1-c)(x+2)(n+4) + c(n+3)(n+4)}{6(c-1)(x+1)(x+2) + 6(1-c)(x+1)(n+3) + c(n+2)(n+3)} \right]$$

b. Specific case $n = 1$

1. If $x = 0$, $\hat{\theta} = \frac{6-c}{15}$.
2. If $x = 1$, $\hat{\theta} = \frac{9+c}{15}$.

c. Specific case $n = 2$

1. If $x = 0$, $\hat{\theta} = \frac{6-c}{18+2c}$.
2. If $x = 1$, $\hat{\theta} = 1/2$.
3. If $x = 2$, $\hat{\theta} = \frac{12+3c}{18+2c}$.

In summary, a method of determining cluster labeling accuracy is developed that may be used after each labeling of sets of samples to sequentially deter-

mine whether sufficient confidence of labeling (by observing p) has been achieved to terminate and label a cluster (and establish n).

4. A CLUSTER PARAMETER EVALUATION EXAMPLE

The following example was part of a study performed for the Earth Observations Division of the NASA Johnson Space Center (JSC) and was used in the LACIE. The purpose of the study was to compare two sets of parameters in the clustering algorithm/program, Iterative Self-Organizing Clustering System (ISOCLS), used to partition 22 932 samples of picture elements (pixels) from Landsat multispectral scanner photographic images. The ISOCLS algorithm has been defined by Kan and Holley (ref. 1) and by Kan (ref. 2), and program

documentation is presented by Minter (ref. 3).

The two sets of parameters that were compared in ISOCLS actually constitute two clustering algorithms. One set of parameters constitutes "nearest neighbor clustering"; that is, 40 pixels (samples) were selected at random and labeled "seed pixels," and then each pixel was assigned to the seed nearest to it. This procedure generates 40 clusters, each with the label of its seed. This is a peculiar algorithm where the>NNL procedure and the MR labeling procedure will always yield the same labels; it is denoted the NN cluster parameter set. The other parameter set involves similar seeding but then goes into a complex (R -means like) set of splittings and

combinations of clusters. These parameters were recommended by Wylie and Bean³ and are referred to herein as the MPAD cluster parameter set. The numerical values used are given in table I. The data are 4 by 1 vectors of spectral values acquired on a particular date. Data were acquired on four different

dates and concatenated into supervectors of data. The number of channels or elements in a vector is thus a function of the number of dates used. The parameters listed in table I are briefly described as follows.

TABLE I.—MPAD and NN Cluster Parameter Sets

Parameter	Cluster parameter set for no. of channels—				
	MPAD				NN—any
	4	8	12	16	
CLUSTERS	60	60	60	60	60
THRESHOLD	8191	8191	8191	8191	8191
SEP	1.0	1.0	1.0	1.0	1.0
PERCENT	100	100	90	90	—
STDMAX	3.2	3.6	3.6	3.6	15.0
DLMIN	3.2	3.9	4.1	4.5	0
NMIN	50	50	50	50	20
ISTOP	8	8	8	8	0
SEQUEN	SC	SC	SC	SC	—
DOTFIL	(a)	(a)	(a)	(a)	(a)

^aRandomly selected starting dots.

- a. DOTFIL—self-generating or randomly selected starting vectors
- b. STDMAX—maximum standard deviation in a cluster before splitting occurs
- c. DLMIN—minimum distance between the means of two clusters needed to combine them
- d. ISTOP—maximum number of iterations in the initial splitting sequence
- e. SEQUEN—the final split/combine (SC) sequence
- f. NMIN—minimum number of pixels needed to form a cluster
- g. SEP—amount of separation between two new clusters after splitting occurs

- h. PERCENT—required percentage of stabilized clusters needed to stop the splitting sequence
- i. CLUSTERS—maximum number of clusters allowed per class
- j. THRESHOLD—the percentage of outlier observations to be deleted from consideration (Zero thresholding was used in this study.)

Each set of 22 932 samples (pixels) in a given area is referred to as a segment and covers a rectangular area of approximately 8.0 by 9.6 kilometers.

The four segments used and the four acquisition dates of each using the Julian calendar system are given in table II. The first two digits represent the year and the last three digits are the number of days into the year. For example, 76218 represents August 5, 1976, since that is the 218th day of 1976.

A question arose during the course of this study as to the merit of using pixels (seed samples) for which

³Alan D. Wylie and William C. Bean, "MPAD LACIE Clustering Parameter Study," NASA JSC Internal Note 76-FM-116, 1977.

TABLE II.— Acquisition Dates for Four LACIE Segments for Crop Year 1975-76

Location	Segment	Acquisition date			
		A	B	C	D
Morton County, Kans.	1961	75277	76164	76236	76254
Finney County, Kans.	1988	75295	76127	76164	76272
Stevens County, Kans.	1865	75349	76136	76172	76190
Randall County, Tex.	1978	75313	76074	76164	76218

the label was unsure. In particular, the categories used were those of small grains (agriculture crops) and non-small-grains. Pixels selected for labeling were sometimes found to be on agricultural field borders, making a "pure" label difficult. However, a "majority of the pixel" strategy was employed in labeling, and each clustering was performed with one of three types of labeling techniques.

a. Mixed—40 pixels were selected at random.

b. Pure border—border pixels that spanned fields of differing categories (small grains/non-small-grains) were replaced with pure pixels.

c. Pure—all border pixels were removed (even small grains/small grains borders) and replaced with pure pixels.

A ground-truth label was used in labeling to avoid confounding clustering purity with analyst-in-

terpreter errors. After the PCC was calculated for each test, the PCC's and the VAR's were averaged over the four sites. These results are given in tables III and IV. Several conclusions can be drawn from table III.

a. For MR labeling, pure-border-type labeling pixels are uniformly (across number of channels) better; i.e., produce higher PCC's.

b. For NNL and MPAD parameters, mixed labeling pixels are uniformly better.

c. For NNL and NN parameters, pure-border labeling pixels are virtually uniformly better; the one exception is virtually a tie.

The comparisons of VAR from table IV for these two clustering parameter sets do not yield such definite results. The NN parameter set yields lower VAR's for 4 channels uniformly (across pixel type).

TABLE III.— Mean PCC Values Averaged Over Sites

Channel	Labeling technique	Cluster parameter set for pixel type—					
		MPAD			NN		
		Mixed	Pure border	Pure	Mixed	Pure border	Pure
4	MR	85.32	86.57	86.52	87.17	88.04	87.96
	NNL	77.02	76.89	76.84	83.29	84.21	84.29
8	MR	89.38	91.13	91.02	89.78	91.69	91.26
	NNL	85.01	84.53	85.71	88.61	90.41	89.99
12	MR	90.31	90.40	90.31	91.64	92.24	91.82
	NNL	82.19	80.78	81.10	89.26	89.74	89.32
16	MR	94.24	95.17	95.17	91.82	91.76	91.36
	NNL	91.69	86.18	86.29	90.23	90.01	89.61

TABLE IV.— Mean Variance Values Averaged Over Sites

Channel	Cluster parameter set for pixel type—					
	MPAD			NN		
	Mixed	Pure border	Pure	Mixed	Pure border	Pure
4	7.63	7.17	7.20	6.89	6.40	6.50
8	4.32	3.64	3.65	5.84	4.48	4.53
12	4.39	4.39	4.47	4.80	4.56	4.60
16	1.95	1.56	1.56	5.48	5.48	5.52

whereas the MPAD parameter set yields lower VAR's for 8, 12, and 16 channels uniformly. The results for PCC (table III) and VAR (table IV) do not give identical conclusions; that is, one parameter set is not clearly superior to the other. For example, if clusters are labeled using the NNL procedure, NN parameters with pure-border-pixel labels are best (highest PCC), but the clusters generated are not as pure as with the MPAD parameters (and mixed pixel labels). An evaluation of clustering purposes and uses is called for in the trade-off of high PCC against low VAR.

Another example of the use of PCC is in the selection of other parameters, such as the number of seed samples. The preceding results were all based on the use of 40 starting dots. The effect on PCC of increasing the number of starting dots to 60 was tested, and the results are given in table V. Only pure-border pixels and the NNL of clusters were used to compare the MPAD and NN parameter sets.

TABLE V.— Comparison of PCC's Using 40 and 60 Starting Vectors^a

Set	Starting vectors	Channels			
		4	8	12	16
MPAD	40	76.9	84.5	80.8	86.2
	60	79.4	89.0	85.3	89.7
NN	40	84.2	90.4	89.7	90.0
	60	83.3	89.6	89.6	90.0

^aPure-border pixels were used with NNL of clusters averaged over four sites.

The values in table V indicate that the NN parameter set with 40 starting vectors produced the highest average PCC. It can be observed also that there is either little or no gain in PCC values when the number of channels is increased from 8 to 12 or 16, regardless of the parameter set used or the number of starting vectors.

5. SUMMARY

Two criteria, PCC and VAR, were presented as measures in cluster algorithm/program evaluation; and an example from the LACIE project at NASA JSC illustrates their use. The theoretical foundation for a system of cluster labeling as a function of cluster purity and size of labeled samples is developed, and an example for rather general assumptions is generated.

6. REFERENCES

1. Kan, E. P. F.; and Holley, W. A.: More on Clustering Techniques With Final Recommendations on ISODATA. Technical Memorandum 640-112, Lockheed Electronics Company, Inc., Houston, Texas, 1972.
2. Kan, E. P. F.: The Latest Version of ISOCAT(A)/ISOCLS. Technical Memorandum 642-570, Lockheed Electronics Company, Inc., Houston, Texas, 1972.
3. Minter, Ruth: ISOCLS Iterative Self-Organizing Clustering Program, CPD202. Lockheed Electronics Company, Inc., Houston, Texas, 1972.

CLASSY—An Adaptive Maximum Likelihood Clustering Algorithm*

R. K. Lenington^a and M. E. Rassbach^b

ABSTRACT

A new clustering method called CLASSY, which alternates maximum likelihood iterative techniques for estimating the parameters of a mixture distribution with an adaptive procedure for splitting, combining, and eliminating the resultant components of the mixture, has been developed. The adaptive procedure is based on maximizing the fit of a mixture of multivariate normal distributions to the observed data using its first through fourth central moments. The method generates estimates of the number of multivariate normal components in the mixture and the proportion, mean vector, and covariance matrix for each component.

This paper describes the mathematical model which is the basis for CLASSY and outlines the actual operation of the algorithm as currently implemented. Results of applying CLASSY to real and simulated Landsat data are presented and compared with results generated by the Iterative Self-Organizing Clustering System (ISOCLS) algorithm, a derivative of the ISODATA algorithm, on the same data sets.

INTRODUCTION

The Large Area Crop Inventory Experiment (LACIE) is dependent on clustering for the determination of spectral classes within a Landsat image of a sample segment (ref. 1). Currently, the Iterative

Self-Organizing Clustering System (ISOCLS) is used for this purpose (refs. 2 and 3). ISOCLS is basically a variation of the *k*-means or ISODATA algorithm of Ball and Hall (refs. 4 and 5). Although this algorithm may be interpreted as a simplified maximum likelihood procedure, it is fundamentally a heuristic algorithm for breaking a data set into fairly homogeneous compact clusters.

A new clustering algorithm called CLASSY, which approximates the mixture distribution of a given data set such as Landsat data with a linear combination of normal distributions, has been developed. CLASSY operates by interleaving maximum likelihood iterative estimation with an adaptive procedure for splitting, combining, and eliminating the resultant components of the mixture density (or clusters). The adaptive procedure is based on maximizing the fit of a mixture of multivariate normal distributions to the observed data using its first through fourth central moments. This procedure allows new components (or clusters) to be created if any existing one appears to be multimodal or otherwise nonnormal. CLASSY produces an estimate of the proportion, mean vector, and covariance matrix for each component in the multivariate normal mixture. It differs from standard maximum likelihood procedures in that it also generates an estimate of the number of components in the mixture.

The CLASSY algorithm is currently implemented on an IBM 370-148 computer. It is written in Fortran IV language and currently accepts as input Landsat imagery on magnetic tape. Both line printer and magnetic tape output are generated by the program.

The following section of this paper describes the mathematical model that is the basis for CLASSY and provides a brief description of the actual operation of the algorithm. The section entitled "Results" contains comparisons of the performances of CLASSY and ISOCLS on simulated data and on actual Landsat data used in LACIE. Finally, these results are evaluated and conclusions are developed.

*The current material for this paper was developed under NASA contract NAS 9-15200 and prepared for the Earth Observations Division, NASA Johnson Space Center, Houston, Texas. CLASSY was developed by M. E. Rassbach while he was a National Research Council postdoctoral fellow working at the Johnson Space Center.

^aLockheed Electronics Company, Houston, Texas.

^bElogic, Inc., Houston, Texas.

MATHEMATICAL DESCRIPTION

Assumptions and Problem Definition

The fundamental mathematical assumption underlying CLASSY is that the data may be usefully approximated by a mixture of multivariate normal densities. That is, if x is an observation vector and p is its probability density function, then

$$p(x|m, \pi_m) = \sum_{i=1}^m a_i p_i(x|\mu_i, \Sigma_i) \quad (1)$$

where a_i is the a priori probability of occurrence of class i ; $p_i(x|\mu_i, \Sigma_i)$ is the multivariate normal probability density function for class i with mean vector μ_i and covariance matrix Σ_i ; m is the total number of classes; π_m is the full set of parameters (i.e., $\{a_1, \dots, a_m, \mu_1, \dots, \mu_m, \Sigma_1, \dots, \Sigma_m\}$).

Given a set of statistically independent, unlabeled sample vectors $\{x_j\}$, the likelihood function may be formed in the following manner:

$$L(\{x_j\}|m, \pi_m) = \prod_{j=1}^N \left[\sum_{i=1}^m a_i p_i(x_j|\mu_i, \Sigma_i) \right] \quad (2)$$

where N is the total number of samples.

So far, the assumptions and equations parallel the usual maximum likelihood development. CLASSY makes the additional assumption that each value of the parameters m and π_m occurs with an a priori probability distribution $A(m, \pi_m)$. This Bayesian formulation of the problem is taken to avoid the degenerate situation of increasing the likelihood by generating more and more clusters with smaller and smaller values of a_i . The practical limit of this process is that each class will be associated with only one data point.

The objective of CLASSY, then, is to determine the discrete parameter m and the continuous

parameter vector π_m so as to maximize the following function:

$$L(\{x_j\}|m, \pi_m) = A(m, \pi_m) \prod_{j=1}^N \left[\sum_{i=1}^m a_i p_i(x_j|\mu_i, \Sigma_i) \right] \quad (3)$$

The values of m and π_m which maximize equation (3) specify a set of distributions that will be called clusters. Of course, $A(m, \pi_m)$ must be chosen so that it satisfies the normalization constraint

$$\sum_{m=1}^{\infty} \int A(m, \pi_m) d\pi_m = 1 \quad (4)$$

The upper limit on m is infinity since the possibility of generating an infinite number of clusters must be considered (in theory).

Typically, in the absence of other information, the a priori probabilities may be chosen as

$$A(m, \pi_m) = \begin{cases} \beta \prod_{i=1}^m C_i, & \pi_m \in R_m \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

where $C_i = C$ is a constant containing normalization factors over π_m space, β is an overall normalization constant, and R_m is a finite region of π_m space corresponding to allowable values for the parameters. Using this simple form for $A(m, \pi_m)$ in equation (4), the following is obtained.

$$\sum_{m=1}^{\infty} \int_{R_m} \beta C^m d\pi_m = \beta \sum_{m=1}^{\infty} \left(C \int_{R_1} d\pi_1 \right)^m \quad (6)$$

Now if

$$C = \gamma \left(\int_{R_1} d\pi_1 \right)^{-1}$$

where $\gamma < 1$, then the sum in equation (6) will converge and $\beta = 1 - \gamma$ provides the proper normalization. Thus, larger values of γ provide a priori bias in favor of more clusters, whereas smaller values provide bias in favor of fewer clusters.

In the current version of CLASSY, the authors have been using $\gamma = e^{-1}$ and approximating the R_1 integral of $d\pi_1$ by e^{2d} . This represents a crude approach to the problem of determining the form of $A(m, \varpi_m)$. However, in practice, the overall technique to be described in the next section has proven not to be sensitive to reasonable changes in the value of C .

With the form for $A(m, \varpi_m)$ assumed in equation (5), the function to be maximized becomes

$$L(\{x_j\}, m, \varpi_m) = \begin{cases} \left(\prod_{i=1}^m c_i \right) \prod_{i=1}^m \left[\sum_{k=1}^N \frac{a_i}{(c_i)^d |x_k - \mu_i|^d} \right] \\ \exp \left[\frac{(x_j - \mu_i)^T \Sigma_i^{-1} (x_j - \mu_i)}{2} \right] & \text{if } x_j \in R_m \\ 0, \text{ otherwise} \end{cases} \quad (7)$$

where d is the dimensionality of the observations x_j .

Solution Procedure

Many approaches may be taken to maximize equation (3). The approach chosen in CLASSY is to interleave maximum likelihood iteration (designed to maximize $L(\{x_j\}, m, \varpi_m)$ with respect to the continuous parameter vector ϖ_m) with a discrete split, join, and combine process (designed to maximize $L(\{x_j\}, m, \varpi_m)$ with respect to the discrete parameter m). Although the theoretical convergence properties of this procedure have not been examined, it is expected that, by alternating these two techniques, values of m and ϖ_m corresponding to at least a local maximum of $L(\{x_j\}, m, \varpi_m)$ will be determined.

Because the splitting and combining techniques operate around each existing cluster and the statistics for hypotheses concerning different numbers of clusters are maintained separately, it has been observed that the final local maximum will often be global.

Necessary conditions for a maximum of $L(\{x_j\}, m, \varpi_m)$ with respect to ϖ_m , assuming a fixed number of classes m , are well known (see Dunn and Hart (ref. 6) and Wolfe (ref. 7)) and are given by the following equations:

$$p(i|x_k, \varpi_m) = \frac{a_i p_i(x_k | \mu_i, \Sigma_i)}{\sum_{j=1}^m a_j p_j(x_k | \mu_j, \Sigma_j)} \quad (8)$$

$$a_i = \frac{1}{N} \sum_{k=1}^N p(i|x_k, \varpi_m) \quad (9)$$

$$\mu_i = \frac{\sum_{k=1}^N p(i|x_k, \varpi_m) x_k}{\sum_{k=1}^N p(i|x_k, \varpi_m)} \quad (10)$$

$$\Sigma_i = \frac{\sum_{k=1}^N p(i|x_k, \varpi_m) (x_k - \mu_i) (x_k - \mu_i)^T}{\sum_{k=1}^N p(i|x_k, \varpi_m)} \quad (11)$$

where $p(i|x_k, \varpi_m)$ is the posterior probability of class i , given the k th sample vector and the values of the parameters, and a_i , μ_i , and Σ_i , $i = 1, \dots, m$, are the elements of ϖ_m .

Numerous techniques have been proposed for obtaining a solution to this set of coupled, simultaneous nonlinear equations. Specific methods have been suggested by Quirein and Trichel (ref. 8), Day (ref. 9), Hasselblad (ref. 10), and Wolfe (ref. 7), among others. CLASSY uses direct functional iteration for equations (10) and (11); that is, use of estimates for

μ_i and Σ_i on the right side to produce improved estimates on the left side.

Estimates for the a priori class probabilities a_i are computed using an iteration scheme which has proved to converge more rapidly than simple functional iteration using equation (9). The scheme used is specified by the following equation, which is derived in the appendix.

$$a_i = \frac{a_i \sum_{p_i > q_i} \frac{p_i - q_i}{p}}{N - \sum_{p_i > q_i} \frac{q_i}{p} - \sum_{p_i < q_i} \frac{p_i}{p}} \quad (12)$$

where

$$p_i = p_i(\mathbf{x}_k | \mu_i, \Sigma_i)$$

$$p = \sum_{j=1}^m a_j p_j(\mathbf{x}_k | \mu_j, \Sigma_j)$$

$$q_i = \sum_{j \neq i} \left(\frac{a_j}{1 - a_j} \right) p_j(\mathbf{x}_k | \mu_j, \Sigma_j)$$

N = the total number of observations

This equation is used by substituting old values of a_i , μ_i , and Σ_i , $i = 1, \dots, m$, on the right to obtain an updated estimate for a_i on the left. The summations are taken over all values of \mathbf{x}_k such that $p_i > q_i$ or $p_i < q_i$.

Initially, each new data point \mathbf{x}_j is used to update the parameter values using equations (8) through (12). This procedure allows rapid evolution of the parameters as new data points are processed. A danger lies in the fact that the data are considered sequentially. If significant correlation is present in the data, updating the parameters with each new data point could theoretically cause the maximum likelihood equations to converge very slowly or to undergo cyclic drifts. This problem has been found to be particularly severe in Landsat data, which exhibit high correlation within fields. To reduce the effects of this correlation, the data are initially scrambled in

a random fashion. Using scrambled data and updating the parameter values with each new data point, the authors have observed that the number of samples (N) required for initial convergence is on the order of a few hundred, even for large data sets. Following initial convergence, the parameters are updated only after a complete pass has been made through the data. This second type of iteration allows a fine tuning of the parameter values and is not subject to problems related to data correlation. The conditions under which the second mode of parameter iteration is entered are discussed later in this section.

The same iteration scheme used to update the parameters is also used to accumulate third- and fourth-order central moments. That is, current values of the parameters are used with each new data point to form the new terms to be accumulated for estimating the moments. The fundamental equations for the estimates of the third- and fourth-order moments are generalizations of equations (10) and (11) and are given as

$$s_{kpq}^{(i)} = \frac{1}{W_i} \sum_{j=1}^N \bar{x}_{jk} \bar{x}_{jp} \bar{x}_{jq} p(i | \mathbf{x}_j, \pi_m) \quad (13)$$

and

$$k_{klpq}^{(i)} = \frac{1}{W_i} \sum_{j=1}^N \bar{x}_{jk} \bar{x}_{jl} \bar{x}_{jp} \bar{x}_{jq} p(i | \mathbf{x}_j, \pi_m) \quad (14)$$

where $\bar{x}_{jk} = (x_{jk} - \mu_{ik})$
 x_{jk} = the k th component of the j th sample vector
 μ_{ik} = the current estimate for the k th component of the mean vector of cluster i

and where

$$W_i = \sum_{j=1}^N p(i | \mathbf{x}_j, \pi_m) \quad (15)$$

The parameter W_i is defined as the weight for cluster i and may be considered as the number of points

assigned to a cluster on a fractional probabilistic basis; $\mathbf{S}^{(i)}$ is a three-dimensional "skewness" tensor, and $\mathbf{K}^{(i)}$ is a four-dimensional "kurtosis" tensor. To reduce the number of parameters to be estimated and stored, traces of these tensors are formed using the inverse of the estimated sample covariance matrix for cluster i (Σ_i) to obtain

$$s_k^{(i)} = \frac{1}{W} \sum_{j=1}^N \bar{x}_{jk} (\bar{x}_j^T \Sigma_i^{-1} \bar{x}_j) p(i | \mathbf{x}_j, \pi_m) \quad (16)$$

where $k = 1, 2, \dots, d$; and

$$K_{kl}^{(i)} = \frac{1}{W} \sum_{j=1}^N \bar{x}_{jk} \bar{x}_{jl} (\bar{x}_j^T \Sigma_i^{-1} \bar{x}_j) p(i | \mathbf{x}_j, \pi_m) \quad (17)$$

where $k, l = 1, 2, \dots, d$; and

$$\bar{x}_j^T = [\bar{x}_{j1} \dots \bar{x}_{jd}]$$

During the initial iteration mode, when parameter values are changing with each data point, the estimates for

$$\mathbf{s}^{(i)} = (s_1^{(i)} \dots s_d^{(i)})$$

and

$$\mathbf{K}^{(i)} = (K_{kl}^{(i)})$$

for each cluster i are only approximately correct. The second mode of iteration produces a more accurate estimate of these statistics. As shall be seen, the estimates of $\mathbf{S}^{(i)}$ and $\mathbf{K}^{(i)}$ are used in the maximization of the likelihood with respect to the discrete parameter m .

The optimization of $L(\{\mathbf{x}_j\}, m, \pi_m)$ with respect to the discrete parameter m takes the form of generating hypotheses concerning the number of clusters and

the subsequent testing of these hypotheses using a likelihood ratio test. At certain points in the process of maximum likelihood iteration, it is possible to generate a hypothesis concerning the fit of a given cluster to the data; namely, either that the data are better represented by two clusters rather than one (a split hypothesis) or that the data are better represented by combining the given cluster with another cluster (a join hypothesis). Each cluster is checked to determine whether either a split or a join hypothesis seems reasonable when the weight for that cluster as defined in equation (15) exceeds a threshold. At this same time, a portion of the old data, which have been accumulated using less accurate parameter values, is subtracted from the appropriate sum for each of the parameters given in equations (8) through (11). The weight threshold is initially set at 200 and increases each time it is exceeded. This procedure allows an initial fit to the major clusters in the data and a subsequent development of more detailed cluster structure.

The generation of a split hypothesis is governed by comparing scalar measures of multivariate skewness and kurtosis for each cluster to thresholds derived from the appropriate distribution for these measures computed under the assumption of a multivariate normal distribution. The scalar measures of multivariate skewness and kurtosis are contractions of the skewness vector $\mathbf{S}^{(i)}$ and the kurtosis matrix $\mathbf{K}^{(i)}$ with respect to the inverse of the estimated covariance matrix for cluster i , Σ_i^{-1} . These measures are given by

$$s_i^2 = \mathbf{S}^{(i)T} \Sigma_i^{-1} \mathbf{S}^{(i)} \quad (18)$$

$$k_i = \text{Tr}(\mathbf{K}^{(i)} \Sigma_i^{-1}) \quad (19)$$

$$(k_i^0)^2 = \text{Tr}(\mathbf{K}^{(i)} \Sigma_i^{-1} \mathbf{K}^{(i)} \Sigma_i^{-1}) - \frac{k_i^2}{d} \quad (20)$$

Here, k_i is the trace of the normalized kurtosis matrix for cluster i and $(k_i^0)^2$ is the trace-free component of the square of k_i .

If any one of these three statistics given by equations (18) to (20) exceeds its threshold value, the hypothesis is formed that the i th cluster may be split into two parts. The parameters for each of the two

new component clusters are estimated by minimizing the squared differences between the observed covariance matrix, the skewness vector, and the kurtosis matrix and the corresponding quantities for the mixture distribution composed of the two new normal distributions. The proportion and mean for the mixture composed of the subclusters are defined to be exactly equal to the corresponding quantities for the parent cluster. That is, if a_i and μ_i are the current estimates of proportion and mean for cluster i and a_{i_1} , a_{i_2} , μ_{i_1} , and μ_{i_2} are the corresponding initial values of the subcluster parameters, it is required that

$$a_i = a_{i_1} + a_{i_2} \quad (21)$$

and

$$\mu_i = \frac{a_{i_1} \mu_{i_1} + a_{i_2} \mu_{i_2}}{\hat{a}_{i_1} + \hat{a}_{i_2}} \quad (22)$$

Thus, the difference in subcluster proportions and the difference in the subcluster mean vectors are left as free parameters. The other free parameters are the independent elements of the two subcluster covariance matrices. Therefore, a total of

$$1 + d + 2 \left[\frac{d(d+1)}{2} \right] = (d+1)^2$$

parameters must be determined.

There are $[d(d+1)]/2$ equations, each of which matches the covariance matrix and kurtosis matrix parameters for the parent cluster to the corresponding parameters for the subcluster mixture. In addition, there are d equations matching the skewness vector parameters for the parent cluster and the subcluster mixture. This is a total of $d^2 + 2d$ equations. Thus, there is one more free parameter or unknown than there are equations and a unique solution is not possible.

The approach taken to obtaining a solution is to

minimize by means of a steepest descent algorithm a quadratic form that may be expressed as

$$\phi = \alpha_1 \|\Sigma_i - \Sigma_p\|^2 + \alpha_2 \|\mathbf{K}^{(i)} - \mathbf{K}_p\|^2 + \alpha_3 \|\mathbf{S}^{(i)} - \mathbf{S}_p\|^2 \quad (23)$$

where Σ_i , $\mathbf{K}^{(i)}$, and $\mathbf{S}^{(i)}$ are the current estimates of the covariance matrix, the kurtosis matrix, and the skewness vector, respectively, for cluster i ; Σ_p , \mathbf{K}_p , and \mathbf{S}_p are the corresponding "pooled" estimates from the mixture of the subclusters under the restrictions of equations (21) and (22); and α_1 , α_2 , and α_3 are arbitrary constants. The norms are the appropriate matrix and vector norms. That is, if M_i is one of the symmetric matrices in equation (23) and $V_i = \mathbf{S}^{(i)} - \mathbf{S}_p$, then

$$\|M_i\|^2 = \text{Tr} (M_i \Sigma_i^{-1} M_i \Sigma_i^{-1})$$

$$\|V_i\|^2 = V_i^T \Sigma_i^{-1} V_i$$

Minimization of equation (23) under the restrictions of equations (21) and (22) produces estimates for the proportions, mean vectors, and covariance matrices which define two new multivariate normal clusters. In the generation of a split hypothesis, the statistics defining the multivariate normal parent cluster are not discarded. When the maximum likelihood iteration cycle is begun again, it is performed for the previously existing clusters, including the parent cluster, and for the two new clusters, which may be thought of as subclusters of the parent cluster. Thus, as split and join hypotheses are generated, a hierarchical cluster structure or cluster tree evolves. Final decisions concerning the choice of a parent cluster or its subclusters to represent the data are made on the basis of likelihood ratio tests as will be described later.

The generation of a join hypothesis is the inverse of the split hypothesis generation procedure. That is, if the generation of a join hypothesis for two already existing clusters is deemed reasonable, then statistics for a new parent cluster are calculated from the multivariate normal mixture distribution defined by the two clusters to be joined. The new parent cluster

is inserted at the level of the clusters to be joined and the clusters to be joined are moved to the next lower level in the tree as subclusters of the new parent.

It should be noted that only clusters which have a common parent are eligible to be joined. The test for determining when a join hypothesis should be generated is designed to measure the degree of overlap between clusters having a common parent cluster. (All the clusters at the top level of the tree are assumed to have a common parent.) The overlap is checked by comparing the mean vectors and the diagonal elements of the covariance matrices for two clusters. A heuristic criterion is used to perform this check. This criterion is given by equation (24).

$$R_{ij} = \frac{(\mu_i - \mu_j)^T \left(\frac{W_i \Sigma_i^{-1} + W_j \Sigma_j^{-1}}{W_i + W_j} \right) (\mu_i - \mu_j) + A \sum_{k=1}^d (|\ln |a_{kk}| - \ln |a_{kk}|)^2}{B \left(\frac{W_i}{W_j} - \frac{W_j}{W_i} \right)^2 + 1} \quad (24)$$

where W_i is the current weight for cluster i and A and B are arbitrary constants (currently, $A = 0.3$ and $B = 0.18$).

The first term in the numerator is a weighted distance between the mean vectors of clusters i and j . The weighting is accomplished by an average inverse covariance matrix for clusters i and j . The second term in the numerator is a measure of the difference in the diagonal elements of the two covariance matrices. The diagonal elements rather than the full covariance matrices are used for computational simplicity. A more complete expression involving all covariance terms would be $\ln[\det \Sigma_i \Sigma_j^{-1}]$. The denominator is designed to discriminate against small clusters in the sense that R_{ij} will be artificially reduced if the weight of one cluster is small relative to the weight of the other cluster. This factor is designed to give large clusters an opportunity to absorb small clusters if such a join does not substantially affect the statistics of the larger cluster.

The R_{ij} criterion is computed for each cluster having the same parent as cluster i . If the cluster j for which R_{ij} is a minimum is less than an empirically set fixed threshold, then a join hypothesis for cluster i and j is generated.

Final decisions concerning the acceptance or rejection of split and join hypotheses are made in terms of likelihood ratio tests. If there are m_i subclusters for

a given parent cluster i , then the logarithm of the likelihood ratio of the subclusters to the parent is accumulated at the same time that maximum likelihood iteration is taking place. The form of this likelihood ratio is given by equation (25).

$$\begin{aligned} \ln \Lambda_i &= \ln \left\{ \frac{\beta C^{m_i} \prod_{j=1}^N \left[\sum_{k=1}^{m_i} a_{k_i} p(x_j | \mu_{k_i}, \Sigma_{k_i}) \right]}{\beta C \sum_{j=1}^N a_{j_i} p(x_j | \mu_i, \Sigma_i)} \right\} \\ &= (m_i - 1) \ln C + \sum_{j=1}^N \left\{ \ln \left[\sum_{k=1}^{m_i} a_{k_i} p(x_j | \mu_{k_i}, \Sigma_{k_i}) \right] \right. \\ &\quad \left. - \ln [a_{j_i} p(x_j | \mu_i, \Sigma_i)] \right\} \quad (25) \end{aligned}$$

where Λ_i is the likelihood ratio for cluster i , a_{k_i} , μ_{k_i} , and Σ_{k_i} are the current estimates of the parameters for cluster i ; and a_{k_i} and Σ_{k_i} are the corresponding subcluster parameters. This log likelihood ratio is tested against a threshold computed assuming that $2 \ln \Lambda_i$ is approximately distributed as an χ^2 random variable with degrees of freedom equal to $a + 1$. A one-tailed test is used, and the probability of a type I error is set at 0.01. If $2 \ln \Lambda_i$ exceeds the threshold set by the test, then the statistics for the parent cluster are eliminated and the subclusters take the place of the parent cluster.

It is also possible that $\ln \Lambda_i$ may become negative, even though in theory this should not occur. In practice, negative values may occur because of poor initial estimation of the subcluster parameters or lack of convergence in these estimates. To avoid the expense of maintaining poor subclusters, the subclusters are eliminated in favor of the parent cluster when $\ln \Lambda_i$ falls below a fixed negative threshold. This threshold is set to a large negative value to allow the subcluster statistics to converge if they are going to converge.

One other possibility in testing the likelihood ratio is that the subcluster statistics may actually converge so that the mixture distribution defined by the subcluster parameters reproduces or very nearly

reproduces the parent cluster distribution. In such cases, Λ_i will remain at a low value possibly slightly greater than or less than zero. If this occurs, it may be assumed that the parent cluster is the most economical description of the data and the subclusters may be eliminated. To test for this situation, another statistic based on the accumulated point probabilities under the parent and subcluster hypotheses is examined. Defining

$$p_{i_i}(x_j) = \sum_{k=1}^{m_i} a_{k_i} p(x_j | \mu_{k_i}, \Sigma_{k_i})$$

where a_{k_i} , μ_{k_i} , and Σ_{k_i} are the current estimates of the parameters for the subclusters of cluster i , the statistic computed is

$$E_i = \sum_{j=1}^N \left\{ \frac{p_i(x_j | \mu_i, \Sigma_i) - p_{i_i}(x_j)}{p_i(x_j | \mu_i, \Sigma_i) + p_{i_i}(x_j)} \right\}^2 \quad (26)$$

Equation (26) gives a crude measure of how much a parent cluster differs from the mixture of its subclasses. If E_i becomes smaller than a fixed empirically determined threshold and the log likelihood ratio is less than a fixed small positive value, then the subclusters are eliminated in favor of the parent cluster.

The one remaining test in the portion of the program that performs maximization with respect to the number of classes is a simple test on the proportion a_i of each cluster or subcluster. If this proportion falls below a threshold value, currently set to 0.01, then the cluster is eliminated. This test is used primarily in the interest of efficiency since very small clusters do not significantly affect the overall mixture distribution.

All the tests for the generation of hypothesized new clusters and for the elimination of clusters or subclusters occur at certain intervals during the process of maximum likelihood iteration and statistics accumulation; namely, when the weight for a given cluster has increased by a fixed amount or when a complete pass has been made through the data since the last tests were performed. After the tests have been made and any resultant restructuring of the cluster tree has taken place, E_i (given by eq. (26)), K_i ,

S_i , and Λ_i are reset. Thus, these statistics depend only on the data processed since the last testing of the cluster statistics for cluster i .

The present program cycles through the data a fixed number of times. (The number of passes through the data is controlled by an external parameter.) When the desired number of passes is complete, the program clusters the data by examining it point by point and assigning each data point to the cluster in the cluster tree for which the probability of occurrence of this data point is the greatest. This is the only time in the program that points are assigned to clusters. When all the points have been assigned, a cluster map showing the cluster symbol for each point is printed out. The program also prints out the final values for the parameters for each cluster in the cluster tree.

Figure 1 is a general flow diagram for the CLASSY program. This is not a detailed flow diagram for the program but merely serves to summarize the information given in this section in a convenient manner.

The initial values assumed at the beginning of the program are as follows.

$$\left. \begin{array}{l} m = 1 \\ a_1 = 1 \\ \mu_1 = \begin{bmatrix} 0.04 \\ \cdot \\ \cdot \\ \cdot \\ 0.04 \end{bmatrix} \\ \Sigma_1 = \begin{bmatrix} 10 & 0 \\ & \cdot \\ & \cdot \\ & \cdot \\ 0 & 10 \end{bmatrix} \end{array} \right\} \quad (27)$$

DATA, PROCEDURES, AND RESULTS

To evaluate the CLASSY clustering algorithm, it was applied to both real and simulated Landsat data. Performance measures were defined and calculated for each trial of the algorithm. The measures were compared with those derived from applying the ISOCLS algorithm to the same data.

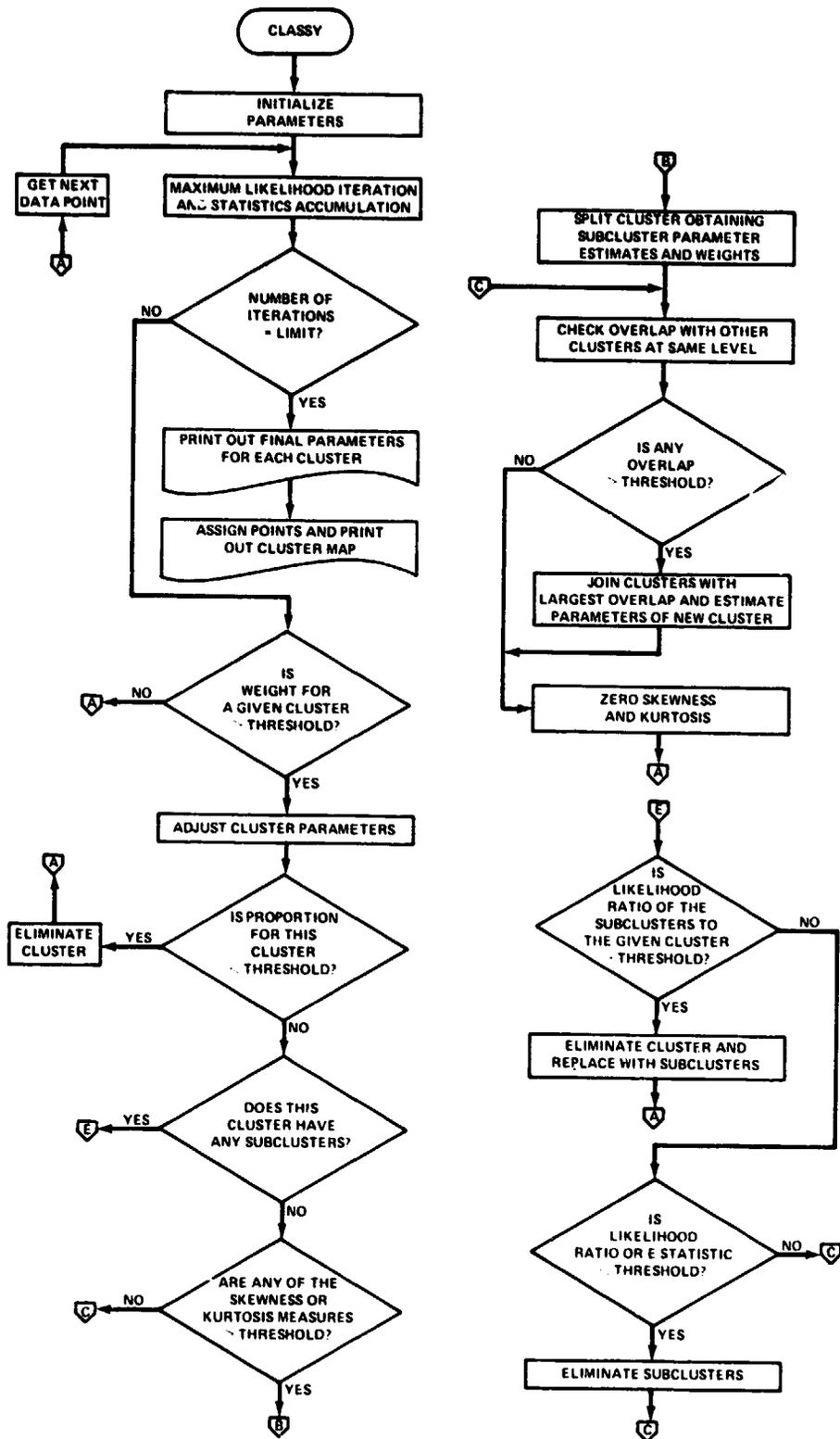


FIGURE 1.—Flow diagram for the CLASSY algorithm.

Data Sets

Two different data sets were used in the comparative evaluation of CLASSY and ISOCLS. The first was a set of Landsat acquisitions of four different LACIE segments. Each LACIE segment is 196 picture elements (pixels) per line by 117 lines and corresponds to a 5- by 6-nautical-mile area on the ground. The second data set was a group of four different simulated acquisitions of a simulated LACIE segment. Each of these data sets is described separately in the following paragraphs.

The four LACIE segments were selected on the basis of the availability of ground truth at regularly spaced pixels in the image and the provision of a representative sampling of LACIE segments in terms of field structure and the proportion of wheat present. Once the segments had been chosen, the acquisition that had the greatest separability, as measured by the Bhattacharyya distance, was selected. The Bhattacharyya distance was computed between wheat and nonwheat classes where the class statistics were obtained from ground-truth fields. The segment number and location, the acquisition date with the largest separability, and the ground-truth percentages of wheat and small grains for each segment are given in table 1.

TABLE 1.—Description of LACIE Sample Segments

Segment	Location	Acquisition date	Ground truth, percent wheat	Ground truth, percent small grains
1181	Kansas	Mar. 10, 1976	23.4	29.0
1988	Kansas	Nov. 8, 1975	33.0	33.0
1961	Kansas	July 18, 1976	8.2	8.2
1965	North Dakota	Aug. 8, 1976	41.6	47.0

The simulated data set consisted of four simulated Landsat acquisitions, each 196 pixels by 117 lines. This data set was generated by IBM for the Mission Planning and Analysis Division at the Johnson Space Center (ref. 11). Each "acquisition" was obtained first by specifying the mean vector and covariance matrix for 10 different classes. The class statistics for

each class were specified so as to simulate the LACIE data for two wheat classes (W_1 and W_2), two barley classes (B_1 and B_2), two classes of grass (G_1 and G_2), two stubble classes (S_1 and S_2), and two classes of fallow (F_1 and F_2). The statistics for these classes were actually obtained from Landsat data representing an agricultural area in Hill County, Montana. Once the statistics for a given class were specified, independent samples were generated from a four-dimensional multivariate normal distribution having those statistics. These samples were then placed in rectangular fields arranged over the simulated segment. This process was repeated for each class and for each of the four acquisitions. The arrangement of the simulated fields over the segment was the same for each acquisition. The pattern of the simulated fields is given in figure 2.

W_1	G_2	B_1	S_1	W_2	S_2	W_2	W_1	G_1	B_2
F_2	W_2	G_1	W_1	S_1	S_2	G_2	B_2	W_1	B_1
W_1	G_1	S_2	G_2	S_1	W_2	B_2	W_2	B_1	F_1
G_2	S_1	W_2	B_1	S_2	W_1	W_2	G_1	F_1	B_2
W_2	W_1	G_1	B_1	W_1	S_1	G_2	S_2	B_2	W_2

FIGURE 2.—Distribution of classes in simulated segment.

Evaluation Method and Procedures

CLASSY was evaluated using a comparative analysis method in which the clustering results of CLASSY were compared with those of ISOCLS using the ground truth as a reference. The evaluation procedure consisted of two steps.

1. The CLASSY and ISOCLS algorithms were applied to each segment in each data set. CLASSY was run for three complete iterations through all the data in each segment. ISOCLS was run in the nearest neighbor mode with 40 ground-truth pixels as start-

ing vectors. In this mode, ISOCLS merely assigns pixels to the nearest starting vector measured in terms of L-1 distance rather than operating iteratively. This mode was chosen for ISOCLS because this was the manner in which the algorithm was currently being used in the LACIE project.

2. The clusters in the line printer map produced by each algorithm were analyzed by first recording the cluster symbol and the corresponding ground-truth label (either wheat or nonwheat) for each pixel where ground truth was available. These results were tabulated, so that the number of ground-truth wheat pixels and ground-truth nonwheat pixels falling in each cluster was known. The clusters were then labeled wheat or nonwheat by majority rule.

A measure of the accuracy of each clustering algorithm in separating wheat from nonwheat (or a measure of the overall purity of the wheat and nonwheat clusters) was computed by estimating the probability of correct classification (PCC) for the labeled clusters. This probability is given by

$$PCC = \sum_{i=1}^{m_1} P(O_i|O) P(O) + \sum_{i=1}^{m_2} P(W_i|W) P(W) \quad (28)$$

where m_1 is the number of clusters labeled "other"; m_2 is the number of clusters labeled wheat; $P(O_i|O)$ is the probability that a pixel falls in the i th "other" cluster, given that it is other than wheat; $P(W_i|W)$ is the probability that a pixel falls in the i th wheat cluster, given that it is wheat; $P(W)$ is the a priori probability that a pixel is wheat; and $P(O)$ is the a priori probability that a pixel is other than wheat. Empirical proportions were used to estimate these probabilities and a priori values, resulting in the following estimate:

$$\hat{PCC} = \frac{1}{N_T} \sum_{i=1}^{m_1} N_{O_i|O} + \sum_{i=1}^{m_2} N_{W_i|W} \quad (29)$$

where N_T is the total number of ground-truth pixels, $N_{O_i|O}$ is the number of ground-truth "other" pixels falling in the i th "other" cluster, and $N_{W_i|W}$ is the number of ground-truth wheat pixels falling in the i th wheat cluster. It is noteworthy that, to obtain an accurate estimate of PCC using equation (29), it is necessary that several ground-truth pixels fall in each

cluster. Specifically, if there are clusters which have only one or two ground-truth grid-intersection pixels, the estimate of PCC will be biased on the high side.

As a part of the analysis, the proportion of wheat was also estimated for the labeled clusters and compared to the ground-truth value. The equation used for this estimate is

$$\hat{P}(W) = \frac{1}{N_T} \sum_{i=1}^{m_2} N_{W_i} \quad (30)$$

where N_{W_i} is the total number of ground-truth pixels (wheat and other) falling in the i th wheat cluster.

Estimates computed using equations (29) and (30) were obtained for each algorithm as applied to both the real and simulated data sets.

Results

The results of these computations are given in tables II through XI. Tables II, III, V, and VI compare CLASSY and ISOCLS results for the LACIE segments examined; the corresponding results for simulated segment data are given in tables VII through XI.

Table II compares the number of clusters and the PCC estimates for ISOCLS (\hat{PCC}_I) and for CLASSY (\hat{PCC}_C) as a result of clustering each of the four LACIE segments examined using both methods. The PCC estimates for CLASSY are, on the average, about 4 percentage points lower than those for ISOCLS. However, since the version of ISOCLS used

TABLE II.—Comparison of the Number of Clusters and the Estimated Probability of Correct Classification Using Single-Pass Segment Data

Segment	ISOCLS		CLASSY		$\hat{PCC}_C - \hat{PCC}_I$
	No. of clusters	\hat{PCC}_I	No. of clusters	\hat{PCC}_C	
1181	40	0.8410	7	0.8052	-0.0358
1988	40	.8070	8	.7661	-.0409
1961	40	.9236	11	.9028	-.0208
1965	40	.7419	9	.6774	-.0645
Average	40	.8284	8.75	.7875	-.0405

generates a factor of 4 to 6 times as many clusters as CLASSY, many of the ISOCLS clusters contain only one or two ground-truth grid-intersection points. As discussed in the preceding section, this means that the PCC estimates for ISOCLS will be biased high relative to CLASSY. In addition, each ISOCLS cluster typically contains one ground-truth point used as a starting vector for that cluster. Since the label of these starting vectors almost always agrees with the cluster label, this amounts to a further high bias in the PCC estimates for ISOCLS. In the light of this bias in favor of ISOCLS and the economy represented by the greatly reduced number of CLASSY clusters, CLASSY compares very favorably to ISOCLS.

The LACIE segments used in this study contained varying amounts of wheat. The ground-truth percentages of wheat $P(W)$ and small grains $P(SG)$ are given in table III. The estimate of the proportion of wheat computed using the ground-truth grid-intersection dots $\hat{P}_D(W)$ is also included. An estimate of the proportion of wheat in the whole scene determined from the clusters labeled wheat can be obtained using equation (30). The wheat proportion estimates resulting from applying this equation to the CLASSY results $\hat{P}_c(W)$ and ISOCLS results $\hat{P}_l(W)$ are also given in table III. Comparing these percentages to the ground-truth wheat proportions shows that, with the exception of segment 1965, the wheat proportion estimates are about 4 to 6 percent higher than the ground-truth wheat proportion values. These slightly high estimates may be due to the fact that, even though only wheat ground-truth dots were used to label clusters, labeled wheat clusters may reasonably be assumed to include some small grains. The last column in table III shows that the ISOCLS estimate was closer to the ground-truth

wheat proportion for two segments and the CLASSY estimate was closer for the other two segments.

The imagery for segment 1965 was examined in detail because the wheat proportion estimates for both CLASSY and ISOCLS deviated considerably from the ground truth and the PCC estimates for both algorithms were correspondingly low for this segment. This segment contained numerous small strip fields. Typically, small-field regions accentuate misregistration problems, and such appears to be the case for this segment. The misregistration of the ground-truth reference acquisition relative to the acquisition clustered reduced PCC values and distorted the proportion of wheat estimates for both algorithms.

To obtain an idea about the relative performance of CLASSY and ISOCLS when applied to multitemporal data, four-channel "green" images were formed for each segment by applying the Kauth (ref. 12) transformation to each of four acquisitions for a given segment and then selecting the green number from each acquisition. (It was necessary to reduce the 16-dimensional data to 4 dimensions since CLASSY is limited to 4 dimensions at the present time.) Table IV lists the four acquisitions used for each segment. The results of comparing the PCC values and the wheat proportion estimates for the two algorithms are given in tables V and VI, respectively. Comparing table V and table II shows that the PCC values for both algorithms remained about the same for segments 1181 and 1961 and that they increased significantly for segments 1988 and 1965. The average difference between the CLASSY and ISOCLS PCC values remained about 4 percent. However, the CLASSY PCC equaled the ISOCLS PCC for segment 1988, and the difference was very small for segment 1961. The last column of table VI

TABLE III.—Comparison of Wheat Proportion Estimates for Labeled Clusters Using Single-Pass Segment Data

Segment	Ground truth		Ground-truth dots $\hat{P}_D(W)$	ISOCLS $\hat{P}_l(W)$	CLASSY $\hat{P}_c(W)$	$D_l = \hat{P}_l(W) - \hat{P}(W)$	$D_c = \hat{P}_c(W) - \hat{P}(W)$	$ D_l - D_c $
	$P(W)$	$P(SG)$						
1181	0.234	0.290	0.333	0.287	0.303	0.053	0.069	-0.016
1988	.330	.330	.322	.397	.287	.067	-.043	.024
1961	.082	.082	.097	.042	.069	-.040	-.013	.027
1965	.416	.470	.516	.526	.645	.110	.229	-.119
Average	.266	.293	.317	.313	.326	.047	.061	-.021

TABLE IV.—Acquisitions Used in Creating Four-Channel Green Images

Segment	Acquisitions
1181	Mar. 10, 1976
	Apr. 16, 1976
	May 3, 1976
	July 14, 1976
1988	Oct. 20, 1975
	May 6, 1976
	June 12, 1976
	Sept. 28, 1976
1961	Aug. 15, 1975
	June 12, 1976
	Aug. 23, 1976
	Sept. 10, 1976
1965	May 11, 1976
	July 21, 1976
	Aug. 8, 1976
	Sept. 14, 1976

shows that, when the four-channel green images were used, the wheat proportion estimates from the CLASSY clusters were closer to the ground-truth values than were the ISOCLS estimates in every case.

Tables VII and VIII are analogous to tables II and III, except that they give the results for the single-pass simulated data. The column labeled maximum likelihood PCC (PCC_M) gives the overall PCC when using standard maximum likelihood classification where the statistics for each class were computed from fields in the simulated image given the class label for each field. Note that the PCC estimates for CLASSY were higher than those for ISOCLS in two of the four passes. In fact, on pass 2, where the separability was greatest, the PCC for CLASSY equaled

TABLE V.—Comparison of the Number of Clusters and the Estimated Probability of Correct Classification Using the Four-Channel Green Image Data

Segment	ISOCLS		CLASSY		$\frac{PCC_c - PCC_1}{PCC_1}$
	No. of clusters	PCC_1	No. of clusters	PCC_c	
1181	40	0.8667	4	0.8000	-0.0667
1988	40	.9357	16	.9357	0
1961	40	.9167	23	.9097	-.0070
1965	40	.8065	13	.7290	-.0775
Average	40	.8814	14	.8436	-.0378

the maximum likelihood PCC. On the average, the PCC for CLASSY was 1.4 percent higher than that for ISOCLS.

The proportion estimate computed from the labeled clusters is given in table VIII. Again, the estimate from CLASSY was closer to the true value in two of the four passes. However, the average individual ISOCLS estimate was about 2 percent closer to the true value.

The results for the simulated data using band 1 from each of the four passes are given in table IX. Band 1 was selected arbitrarily to assess the use of multitemporal data. Note that the PCC estimate for CLASSY was 1.0, meaning that none of the CLASSY clusters contained a mixture of wheat and nonwheat grid-intersection pixels.

Using the simulated data makes it possible to identify a cluster with a certain class in the data by determining which class contributes the majority of pixels to the cluster. After such an identification, the generating statistics for the class may be compared with the cluster statistics produced by CLASSY. Table X presents the results of such a comparison for

TABLE VI.—Comparison of Wheat Proportion Estimates for Labeled Clusters Using Four-Channel Green Image Data

Segment	Ground truth		ISOCLS $\hat{P}_1(W)$	CLASSY $\hat{P}_c(W)$	$D_1 = \hat{P}_1(W) - \hat{P}(W)$	$D_c = \hat{P}_c(W) - \hat{P}(W)$	$ D_1 - D_c $
	$P(W)$	$P(SG)$					
1181	0.234	0.230	0.292	0.241	0.058	0.007	0.051
1988	.330	.330	.316	.342	-.014	.012	.002
1961	.082	.082	.066	.069	-.016	-.013	.003
1965	.416	.470	.625	.565	.209	.149	.060
Average	.266	.293	.325	.304	.059	.039	.029

TABLE VII.—Comparison of the Number of Clusters and the Estimated Probability of Correct Classification Using Single-Pass Simulated Data

Pass	PCC _M	ISOCLS		CLASSY		PCC _M - PCC ₁	PCC _M - PCC _c	PCC _c - PCC ₁
		No. of clusters	PCC ₁	No. of clusters	PCC _c			
1	0.935	40	0.9139	5	0.9043	0.021	0.030	-0.0096
2	.986	40	.9713	5	.9857	.015	.000	.0144
3	.970	40	.9761	8	.9522	-.006	.018	-.0239
4	.928	40	.8852	7	.9187	.043	.009	.0335
Average	.955	40	.9366	6.25	.9402	.018	.014	.0144

TABLE VIII.—Comparison of the Wheat Proportion Estimates for Labeled Clusters Using Single-Pass Simulated Data

Pass	P(W)	P ₁ (W)	P _c (W)	D ₁ =	D _c =	D ₁ - D _c
				P ₁ (W) - P(W)	P _c (W) - P(W)	
1	0.3398	0.3301	0.2536	-0.0097	-0.0862	-0.0765
2	.3398	.3254	.3541	-.0144	.0143	.0001
3	.3398	.3636	.2917	.0238	-.0481	-.0243
4	.3398	.3254	.3349	-.0144	-.0049	.0095
Average	.3398	.3361	.3086	-.0147	-.0312	-.0228

the pass 2 simulated data, whereas table XI gives similar results for the clustering using band 1 from each of the four passes.

In the pass 2 CLASSY results, four of the five clusters could be clearly identified with one of the generating classes or distributions. A comparison of the mean vector and covariance matrices shows a remarkable correspondence between the CLASSY statistics and the generating statistics. Cluster 3 was about equally divided between grass 1 and grass 2.

TABLE IX.—Probability of Correct Classification Using Multipass Simulated Data

Data	ISOCLS		CLASSY		PCC _c - PCC ₁
	No. of clusters	PCC ₁	No. of clusters	PCC _c	
Band 1 from each of 4 passes	40	0.9809	7	1.0000	0.0191

Only the statistics for grass 1 are shown in the table. Similarly, cluster 2 was a mixture of stubble, fallow, and barley 2. The statistics for each of these classes are very similar for this pass. The statistics for stubble 1 are given as a representative example of that group of classes.

The data from band 1 of each of the four simulated passes had more separability; thus, CLASSY was able to distinguish more classes. The comparison of the generating statistics and the CLASSY statistics is presented in table XI. Only the variance terms from the multipass covariance matrix were available. Again, there is remarkable correspondence between the CLASSY statistics and the generating statistics.

CONCLUSIONS

The main conclusion of this paper is that the performance of the CLASSY clustering algorithm compares favorably with that of ISOCLS on both the real and simulated LACIE segment data. In terms of performance, these results were obtained despite the

TABLE X.—Comparison of Cluster Statistics for Pass 2 Simulated Data

Cluster number	Identification	Generating statistics					CLASSY statistics				
		Mean vector	Covariance matrix				Mean vector	Covariance matrix			
4	Wheat 1	20.36	0.91	1.21	0.34	-0.01	20.56	1.21	1.04	0.13	-0.19
		20.19	1.21	3.24	.24	-.65	20.67	1.04	2.87	-.10	-.95
		27.29	.34	.24	1.77	1.75	27.45	.13	-.10	1.84	1.76
		28.14	-.01	-.65	1.75	3.15	28.26	-.19	-.95	1.76	3.50
5	Wheat 2	18.55	0.82	0.69	-0.01	-0.47	18.76	1.08	0.80	-0.03	-0.50
		17.62	.69	1.11	-.48	-1.19	17.13	.80	1.54	-.47	-1.20
		26.35	-.01	-.48	1.23	1.41	26.36	-.03	-.47	1.46	1.50
		28.00	-.47	-1.19	1.41	3.25	27.97	-.50	-1.20	1.50	3.51
1	Barley 1	23.30	1.55	1.74	1.22	0.96	22.97	2.00	1.97	1.83	1.41
		25.80	1.74	3.16	1.52	1.12	25.45	1.97	3.59	2.36	1.77
		25.98	1.22	1.52	1.65	.91	25.27	1.83	2.36	2.92	1.84
		24.19	.96	1.12	.91	1.19	23.50	1.41	1.77	1.84	2.22
3	Grass 1 (grass 2)	20.83	1.31	2.07	0.54	0.11	20.71	1.15	1.48	0.55	0.22
		20.86	2.07	4.70	.91	-.29	20.54	1.48	4.10	1.01	.28
		23.37	.54	.91	1.10	.70	23.18	.55	1.01	1.40	.64
		22.50	.11	-.29	.70	1.23	22.52	.22	.28	.64	1.24
2	Stubble 1 (stubble 2, fallow, barley 2)	21.90	0.97	0.62	0.77	0.69	22.40	0.96	0.44	0.31	0.22
		23.64	.64	1.12	.70	.66	24.43	.44	1.17	.38	.29
		24.22	.77	.70	1.51	1.40	24.18	.31	.38	1.41	.92
		23.12	.69	.66	1.40	2.31	22.77	.22	.29	.92	1.68

fact that CLASSY reduces the number of clusters by a factor of 4 to 6 as compared to ISOCLS. This performance indicates that CLASSY is indeed approximating the empirical mixture density rather than just breaking up the data space into small homogeneous areas as does ISOCLS. This conclusion is further substantiated by noting the high degree of correspondence between the CLASSY cluster statistics and the generating statistics of classes in the simulated data. When data from band 1 of each of the 4 simulated acquisitions was clustered using CLASSY, 5 of the 10 classes were very accurately identified. The remaining classes, whose statistics were very close together, were broken into two reasonable groups. It appears, therefore, that the CLASSY algorithm may well provide a solution to the fundamental problem of clustering—the determination of the inherent number of classes in the data.

REFERENCES

1. MacDonald, R. B.; Hall, F. G.; and Erb, R. B.: The Use of LANDSAT Data in a Large Area Crop Inventory Experiment (LACIE). Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue Univ. (W. Lafayette, Ind.), 1975, pp. 1B-1 through 1B-23.
2. Kan, E. P. F.: The JSC Clustering Program ISOCLS and Its Applications. LEC-0483, Lockheed Electronics Co., Houston, Tex., July 1973.
3. Turley, R. C.: Algorithm Simulation, Test, and Evaluation Program (ASTEP)—Users Guide and Software Documentation. NASA Johnson Space Center, Houston, Tex., Apr. 1978, pp. 1-22.
4. Ball, G. H.; and Hall, D. J.: A Clustering Technique for Summarizing Multivariate Data. Behavioral Science, vol. 12, Mar. 1967, pp. 153-155.

TABLE XI.—Comparison of Cluster Statistics for Band 1 for Each of Four Passes of the Simulated Data

Cluster number	Identification	Generating statistics				CLASSY statistics					
		Mean vector	Covariance matrix			Mean vector	Covariance matrix				
5	Wheat 1	26.93	1.06	0.91	2.15	3.30	26.84	1.27	0.69	1.42	1.61
		20.36					20.27	.69	1.21	1.25	1.62
		17.39					17.22	1.42	1.25	2.32	2.65
		17.27					17.02	1.61	1.62	2.65	3.49
2	Wheat 2	25.79	1.03	0.82	0.47	1.76	25.90	1.22	0.94	0.78	0.98
		18.55					18.76	.94	1.23	.78	.87
		16.85					16.88	.78	.78	.85	.67
		18.12					17.97	.98	.87	.67	1.80
4	Barley 1	28.41	2.16	4.86	4.15	4.47	28.40	2.30	1.56	3.03	2.18
		23.30					22.71	1.56	1.81	2.69	2.17
		22.01					22.56	3.03	2.69	5.33	3.80
		17.01					17.44	2.18	2.17	3.86	3.58
3	Barley 2	28.23	1.33	0.77	1.88	1.61	28.40	1.63	-0.08	1.79	1.05
		22.78					22.71	-.08	.79	-.40	-.09
		22.37					22.56	1.79	-.40	2.54	1.23
		17.34					17.44	1.05	-.09	1.23	1.86
1	Grass 1 (grass 2, stubble 1)	25.67	1.81	1.31	1.80	1.62	25.82	2.69	0.87	1.76	2.17
		20.83					21.20	.87	1.39	.74	.98
		20.10					20.35	1.76	.74	1.71	1.65
		20.60					20.72	2.17	.98	1.65	2.43
6	Fallow 1	24.59	0.67	0.52	0.90	0.66	24.68	0.75	0.38	0.42	0.48
		22.48					22.45	.38	.72	.68	.09
		23.22					23.21	.42	.68	1.06	.04
		21.56					21.67	.48	.09	.04	.75
7	Stubble 2 (fallow 2)	24.33	1.17	0.67	0.74	1.04	24.34	1.31	0.38	-0.01	-0.14
		22.21					22.25	.38	.86	.09	-.15
		22.69					22.70	-.01	.09	1.01	.84
		28.63					28.63	-.14	-.15	.84	1.35

5. MacQueen, J.: Some Methods for Classification and Analysis of Multivariate Observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, L. M. LeCam and J. Neyman, eds., Univ. of California Press, 1967, pp. 281-297.

6. Duda, R. O.; and Hart, P. E.: Pattern Classification and Scene Analysis. John Wiley & Sons (New York), 1973.

7. Wolfe, J. H.: Pattern Clustering by Multivariate Mixture Analysis. Multivariate Behavioral Research, vol. 5, 1970, pp. 329-350.

8. Quirein, J. A.; and Trichel, M. C.: Acreage Estimation, Feature Selection, and Signature Extension Dependent Upon the Maximum Likelihood Decision Rule. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue Univ. (W. Lafayette, Ind.), 1975, pp. 2A-26 through 2A-39.

9. Day, N. E.: Estimating the Components of a Mixture of Normal Distributions. Biometrika, vol. 56, 1969, pp. 463-474.

10. Hasselblad, V.: Estimating the Parameters for a Mixture of Normal Distributions. *Technometrics*, vol. 8, 1966, pp. 431-466.

11. Oliver, R. E.: Description of Simulated LACIE Segments 1851-1854. IBM Information Systems Technical Report, Nov. 18, 1975.

12. Kauth, R. J.; and Thomas, G. S.: The Tasseled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue Univ. (W. Lafayette, Ind.), 1976, pp. 4B-4I through 4B-5I.

Appendix

The equation used to obtain iterative estimates of the a priori class probabilities or proportions, a_i , is derived beginning with equation (9), which is repeated here in a slightly more expanded form.

$$c_i = \frac{1}{N} \sum_{k=1}^N \frac{a_i p_i(x_k | \mu_k, \Sigma_i)}{p(x_k)} \quad (A1)$$

where

$$p(x_k) = \sum_{j=1}^m a_j p_j(x_k | \mu_j, \Sigma_j)$$

Since a_i does not depend on k , a_i may be canceled from both sides of the equation to obtain

$$1 = \frac{1}{N} \sum_{k=1}^N \frac{p_{ik}}{p_k} \quad (A2)$$

where, for convenience, the functional notation has been simplified.

Now,

$$0 = \frac{1}{N} \sum_{k=1}^N \frac{p_{ik} - p_k}{p_k} \quad (A3)$$

But

$$\begin{aligned} p_k &= a_i p_{ik} + \sum_{j=1, j \neq i}^m a_j p_{jk} \\ &= a_i p_{ik} + (1 - a_i) q_{ik} \end{aligned} \quad (A4)$$

Here, define

$$q_{ik} = \begin{cases} \sum_{j=1, j \neq i}^m \left(\frac{a_j}{1 - a_i} \right) p_{jk} \cdot a_i \neq 1 \\ 0, \text{ otherwise} \end{cases} \quad (A5)$$

So,

$$\begin{aligned} 0 &= \frac{1}{N} \sum_{k=1}^N \frac{p_{ik} - a_i p_{ik} - (1 - a_i) q_{ik}}{p_k} \\ &= \frac{1}{N} \sum_{k=1}^N \frac{(1 - a_i) p_{ik} - (1 - a_i) q_{ik}}{p_k} \end{aligned} \quad (A6)$$

$$0 = \sum_{k=1}^N \frac{p_{ik} - q_{ik}}{p_k} \quad (\text{A7})$$

assuming $a_i \neq 1$. Breaking this sum up into those terms which are positive and those which are negative results in

$$0 = \sum_{p_{ik} > q_{ik}} \frac{p_{ik} - q_{ik}}{p_k} + \sum_{p_{ik} < q_{ik}} \frac{p_{ik} - q_{ik}}{p_k} \quad (\text{A8})$$

Now, a_i is reintroduced as follows:

$$0 = \frac{1}{a_i} \left[a_i \sum_{p_{ik} > q_{ik}} \frac{p_{ik} - q_{ik}}{p_k} \right] + \frac{1}{(1 - a_i)} \left[(1 - a_i) \left(\sum_{p_{ik} < q_{ik}} \frac{p_{ik} - q_{ik}}{p_k} \right) \right] \quad (\text{A9})$$

If we now solve for the a_i 's which are outside the square brackets in terms of the a_i 's, p_i 's, and q_i 's inside the square brackets, the following is obtained.

$$a_i = \frac{a_i \sum_{p_{ik} > q_{ik}} \frac{p_{ik} - q_{ik}}{p_k}}{a_i \sum_{p_{ik} > q_{ik}} \frac{p_{ik} - q_{ik}}{p_k} + (1 - a_i) \sum_{p_{ik} < q_{ik}} \frac{p_{ik} - q_{ik}}{p_k}}$$

$$= \frac{a_i \sum_{p_{ik} > q_{ik}} \frac{p_{ik} - q_{ik}}{p_k}}{\sum_{p_{ik} < q_{ik}} \frac{p_{ik} - q_{ik}}{p_k} + a_i \sum_{p_{ik} > q_{ik}} \frac{p_{ik} - q_{ik}}{p_k}}$$

$$= \frac{a_i \sum_{p_{ik} > q_{ik}} \frac{p_{ik} - q_{ik}}{p_k}}{N - \sum_{p_{ik} > q_{ik}} \frac{q_{ik}}{p_k} - \sum_{p_{ik} < q_{ik}} \frac{p_{ik}}{p_k}}$$

(A10) This equation reduces exactly to equation (A1).

This is the iterative equation used to obtain proportion estimates in CLASSY.

Equation (A10) may also be put into a form illustrating the nature of the update term to obtain

$$a_i = a_i + \frac{a_i (1 - a_i) \sum_{k=1}^N \frac{p_{ik} - q_{ik}}{p_k}}{N - \sum_{p_{ik} > q_{ik}} \frac{q_{ik}}{p_k} - \sum_{p_{ik} < q_{ik}} \frac{p_{ik}}{p_k}} \quad (\text{A11})$$

This equation illustrates that direct functional iteration using equation (A10) amounts to adding a correction term given by

$$\frac{a_i (1 - a_i) \sum_{k=1}^N \frac{p_{ik} - q_{ik}}{p_k}}{N - \sum_{p_{ik} > q_{ik}} \frac{q_{ik}}{p_k} - \sum_{p_{ik} < q_{ik}} \frac{p_{ik}}{p_k}}$$

to the old value of a_i in order to obtain the new value of a_i .

As a way of comparing the iterative equation for proportion estimates used in CLASSY (eq. (A10)) to the standard maximum likelihood iterative equation (eq. (A1)), one may rework the standard equation so that the nature of the update term is apparent. Using equation (A6), one obtains

$$1 = 1 + \frac{(1 - a_i)}{N} \sum_{k=1}^N \frac{p_{ik} - q_{ik}}{p_k} \quad (\text{A12a})$$

or

$$a_i = a_i + \frac{a_i (1 - a_i)}{N} \sum_{k=1}^N \frac{p_{ik} - q_{ik}}{p_k} \quad (\text{A12b})$$

A comparison of equations (A11) and (A12) shows that the difference is in the term N versus

$$N = \sum_{p_{ik} > q_{ik}} \frac{q_{ik}}{p_k} - \sum_{p_{ik} < q_{ik}} \frac{p_{ik}}{p_k}$$

Thus, the iterative equation used in CLASSY (eq. (A11)) will amplify the correction for a_i if there are a

significant number of points such that $0 < p_{ik} < 1$ and $0 < q_i < 1$. This corresponds to the case where cluster i is a "mixed" cluster; that is, there is a significant amount of overlap between cluster i and other clusters. Since it is precisely these "mixed" clusters for which the standard iterative equation (eq. (A1) or (A12)) converges slowly, the iterative equation used for proportions in CLASSY (eq. (A10) or (A11)) should converge more readily.

Linear Feature Selection With Applications

H. P. Decell, Jr.,^a and L. F. Guseman, Jr.^b

INTRODUCTION

The Large Area Crop Inventory Experiment (LACIE) is concerned with the use of satellite-acquired (Landsat) multispectral scanner (MSS) data to conduct an inventory of some crop of economic interest such as wheat over a large geographical area. Such an inventory requires the development of accurate and efficient algorithms for data classification. The use of multitemporal measurements (several registered passes during the growing season) increases the dimension of the original measurement space (pattern space) and thereby increases the computational load in classification procedures. In this connection, the cost of using statistical pattern classification algorithms depends, to a large extent, on reducing the dimensionality of the problem by use of feature selection/combination techniques. These techniques are employed to find a subspace of reduced dimension (feature space) in which to perform classification while attempting to maintain the level of classification accuracy obtainable in the original measurement space. The most meaningful performance criterion that can be applied to a classification algorithm is the frequency with which it misclassifies observations; that is, the probability of misclassification. Consequently, one should attempt to select/combine features in such a way that the probability of misclassification in feature space is minimized.

In the sequel, several ways in which feature selection techniques have been used in LACIE are discussed. In all cases, the techniques require some a priori information and assumptions (e.g., number of

classes or form of conditional class density functions) about the structure of the data. In most cases, the classification procedure (e.g., Bayes' optimal) has been chosen in advance. Dimensionality reduction is then performed so as to (1) choose an optimal feature space in which to perform classification and (2) determine a transformation to apply to measurement vectors prior to classification. In all that follows, the transformations used for dimensionality reduction are linear; that is, the variables in feature space are always linear combinations of the original measurements.

As mentioned previously, the most meaningful performance criterion for a classification procedure is the probability of misclassification (denoted in the sequel by G). However, if the dimension of feature space (and therefore measurement space) is greater than one, then G is difficult to compute without additional class structure assumptions (e.g., equal covariance matrices). As a result, several numerically tractable criteria have been developed in conjunction with LACIE which provide some information concerning the behavior of G . These criteria are discussed in the next section. In a subsequent section, a compendium of recent results on linear feature selection techniques, most of which are available only in scattered NASA contract reports, is presented. The final section includes a discussion of the use of these techniques in LACIE, an outline of some of the investigations underway in the use of linear feature selection techniques, and a discussion of some related open questions.

MATHEMATICAL PRELIMINARIES

Let $\Pi_1, \Pi_2, \dots, \Pi_m$ be distinct classes (e.g., crops of interest) with known a priori probabilities $\alpha_1, \alpha_2, \dots, \alpha_m$, respectively. Let $\mathbf{x} = (x_1, x_2, \dots, x_n)^T \in R^n$ denote a feature vector of measurements (e.g., Landsat MSS data from either a single pass or several

^aUniversity of Houston, Houston, Texas. This author was supported in part by NASA Johnson Space Center under contract NAS 9-15000.

^bTexas A & M University, College Station, Texas. This author was supported in part by NASA Johnson Space Center under contract NAS 9-14689.

registered passes) taken from an arbitrary element of

$$\bigcup_{i=1}^m \Pi_i$$

Suppose that the measurement vectors for class Π_i are characterized by the n -dimensional multivariate normal density function

$$p_i(\mathbf{x}) = (2\pi)^{-\frac{n}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \mu_i)^T \Sigma_i^{-1} (\mathbf{x} - \mu_i) \right], 1 \leq i \leq m$$

It is assumed that the $n \times 1$ mean vector μ_i and the $n \times n$ covariance matrix Σ_i for each class Π_i are known (with Σ_i positive definite), $1 \leq i \leq m$. The symbol $|A|$ is used to denote the determinant of the matrix A . The n -dimensional probability of misclassification, denoted by G , of objects from

$$\bigcup_{i=1}^m \Pi_i$$

is given (refs. 1 and 2) by

$$G = 1 - \int_{R^n} \max_{1 \leq i \leq m} \alpha_i p_i(\mathbf{x}) d\mathbf{x} \\ = 1 - \sum_{i=1}^m \alpha_i \int_{R_i} p_i(\mathbf{x}) d\mathbf{x}$$

where the sets R_i , $1 \leq i \leq m$, called the *Bayes' decision regions*, are defined by

$$R_i = \left\{ \mathbf{x} \in R^n : \alpha_i p_i(\mathbf{x}) = \max_{1 \leq j \leq m} \alpha_j p_j(\mathbf{x}) \right\}, 1 \leq i \leq m$$

The resulting classification procedure, called the *Bayes' optimal classifier*, is defined as follows (ref. 1):

Assign an element to Π_i if its vector \mathbf{x} of measurements belongs to R_i , $1 \leq i \leq m$.

The *Bhattacharyya coefficient* for classes i and j ($1 \leq i, j \leq m$) is given (ref. 3) by

$$\rho(i, j) = (\alpha_i \alpha_j)^{\frac{1}{2}} \int_{R^n} \left\{ p_i(\mathbf{x}) p_j(\mathbf{x}) \right\}^{\frac{1}{2}} d\mathbf{x}$$

It has been shown that

$$G \leq \sum_{i=1}^{m-1} \sum_{j=i+1}^m \left\{ \alpha_i \alpha_j \right\}^{\frac{1}{2}} \int_{R^n} \left\{ p_i(\mathbf{x}) p_j(\mathbf{x}) \right\}^{\frac{1}{2}} d\mathbf{x} \equiv \rho$$

The quantity ρ is usually called the *Bhattacharyya distance* (or the *average Bhattacharyya distance*).

There have been various attempts to utilize certain functions of $\rho(i, j)$ and ρ to generate Bhattacharyya-related separability measures. For further variations on this theme, we refer the reader to the complete listing of references and in particular to reference 4.

The *divergence* (ref. 5) between classes i and j ($1 \leq i, j \leq m$) is given by

$$D(i, j) = \frac{1}{2} \text{tr} \left[(\Sigma_i - \Sigma_j) (\Sigma_i^{-1} - \Sigma_j^{-1}) \right] \\ + \frac{1}{2} \text{tr} \left[(\Sigma_i^{-1} + \Sigma_j^{-1}) (\mu_i - \mu_j)(\mu_i - \mu_j)^T \right]$$

and the *average interclass divergence* is given by

$$D = \sum_{i=1}^{m-1} \sum_{\substack{j=1 \\ j \neq i}}^m D(i, j)$$

or, equivalently, as shown in reference 6, by

$$D = \frac{1}{2} \text{tr} \left\{ \sum_{i=1}^m \Sigma_i^{-1} S_i \right\} - \frac{m(m-1)}{2} n$$

where

$$S_i = \sum_{j=1}^m \left(\Sigma_j + \delta_{ij} \delta_{ij}^T \right)$$

and

$$\delta_{ij} = \mu_i - \mu_j$$

As in the case of $\rho(i, j)$ and ρ , various functions of $D(i, j)$ and D have been proposed as class separability measures.

Kanal (ref. 4) provides an excellent exposition of such measures (e.g., Shannon entropy, Vajda's average conditional quadratic entropy, Devijver's Bayesian distance, Minkowsky measures of non-uniformity, Bhattacharyya bound, Chernoff bound, Kolmogorov variational distance, Lissack and Fu's generalization of the latter, Ito's approximating functions, and the Jeffreys-Matusita distance). This work contains 304 references and is perhaps the only comprehensive exposition of the subject through early 1974. A more recent nonparametric separability measure due to Bryant and Guseman (ref. 7) will be outlined at the end of this section.

Devijver (ref. 8) develops a bound on G called the *Bayesian distance*. He gives an excellent development of the concept and its relationship to the aforementioned separability measures. His results are quite general with regard to the class densities $p_i(\mathbf{x})$ and class a priori probabilities α_i , $1 \leq i \leq m$. The Bayesian distance is defined to be

$$H = \sum_{i=1}^m E \left\{ \frac{\alpha_i^2 p_i(\mathbf{x})^2}{p(\mathbf{x})^2} \right\}$$

where

$$p(\mathbf{x}) = \sum_{i=1}^m \alpha_i p_i(\mathbf{x})$$

The measure H satisfies the inequality

$$H \leq G \leq \frac{1}{m} + \frac{m-1}{m} \sqrt{\frac{mH-1}{m-1}} \leq \sqrt{H}$$

Following the philosophy discussed in the introduction, the intractable nature of the expression for G (i.e., although in many instances unnecessary, attention is being restricted to a finite family of normally distributed pattern classes) was one, if not the single, reason for developing more tractable pattern class separability measures. These measures could then be used in lieu of G to determine mappings from pattern space to feature space in which the classification of patterns is equivalent to (G is preserved) or "nearly equivalent to" classification of patterns in pattern space. Two fundamental questions that arise are these. First, what (if any) relation do the class separability measures bear to G ? Second, can one develop tractable algorithms based on the separability measures to determine the dimension-reducing mappings?

In connection with these questions, only linear mappings B of the measurement space R^n onto R^k for $k < n$ will be considered. This is equivalent to requiring that B be a $k \times n$ rank- k matrix. This class of mappings certainly includes those of the "feature subset selection" type since the selection of any k -feature subset (i.e., any k components of $\mathbf{x} \in R^n$) can be accomplished by selecting the appropriate $k \times n$ matrix B consisting of only 0's and 1's. The class of $k \times n$ rank- k matrices is more general in the sense that linear combinations of the features are permissible.

In all that follows, it will be assumed that B is a $k \times n$ rank- k matrix and that X is a normally distributed vector-valued random variable. An observation on X will be denoted by $\mathbf{x} = X(\omega)$. It is well known that if $X \sim N(\mu, \Sigma)$ then $Y = BX \sim N(B\mu, B\Sigma B^T)$.

The transformed measurements, $y = Bx$, for class Π_i , are normally distributed with density function

$$p_i(y, B) = (2\pi)^{-\frac{k}{2}} |B\Sigma_i B^T|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (y - B\mu_i)^T \cdot (B\Sigma_i B^T)^{-1} (y - B\mu_i) \right]$$

and the resulting probability of misclassification is given by

$$G(B) = 1 - \int_{R^k} \max_{1 \leq i < m} \alpha_i p_i(y, B) dy$$

$$= 1 - \sum_{i=1}^m \alpha_i \int_{R_i(B)} p_i(y, B) dy$$

where the transformed Bayes' decision regions are given by

$$R_i(B) = \left\{ y \in R^k : \alpha_i p_i(y, B) \right.$$

$$\left. = \max_{1 \leq j < m} \alpha_j p_j(y, B) \right\}, 1 \leq i \leq m$$

The *B-Bhattacharyya coefficient* for classes i and j is given by

$$\rho_B(i, j) \equiv \left\{ \alpha_i \alpha_j \right\}^{\frac{1}{2}} \int_{R^k} \left\{ p_i(y, B) p_j(y, B) \right\}^{\frac{1}{2}} dy$$

It has been shown by Decell and Quirein (ref. 6) that for each B

$$G \leq G(B) \leq \sum_{i=1}^{m-1} \sum_{j=i+1}^m \rho_B(i, j) \equiv \rho(B)$$

The quantity $\rho(B)$ is called the *B-Bhattacharyya distance* or the *B-average Bhattacharyya distance*.

In addition, it has been shown by Decell and Quirein (ref. 6) that $G = G(B)$ if and only if $\rho = \rho(B)$.

The *B-divergence* between classes i and j ($1 \leq i, j \leq m$) is

$$D_B(i, j) = \frac{1}{2} \text{tr} \left\{ \left[B\Sigma_i B^T - B\Sigma_j B^T \right] \cdot \left[(B\Sigma_i B^T)^{-1} - (B\Sigma_j B^T)^{-1} \right] \right\}$$

$$+ \frac{1}{2} \text{tr} \left\{ \left[(B\Sigma_i B^T)^{-1} - (B\Sigma_j B^T)^{-1} \right] \cdot (B\mu_i - B\mu_j) (B\mu_i - B\mu_j)^T \right\}$$

and the *B-interclass divergence* is

$$D(B) = \sum_{i=1}^{m-1} \sum_{\substack{j=1 \\ j \neq i}}^m D_B(i, j)$$

or, equivalently (ref. 6),

$$D(B) = \frac{1}{2} \text{tr} \left\{ \sum_{i=1}^m (B\Sigma_i B^T)^{-1} (B\Sigma_i B^T) \right\}$$

$$- \frac{m(m-1)}{2} k$$

where

$$S_i = \sum_{\substack{j=1 \\ j \neq i}}^m (\Sigma_j + \delta_{ij} \delta_{ij}^T)$$

and

$$\delta_{ij} = \mu_i - \mu_j$$

Although there is no explicit relationship between G and D (or $G(B)$ and $D(B)$), it was shown by Decell and Quirein (ref. 6) that $D = D(B)$ if and only if $G = G(B)$.

In the present setting and with the obvious general meaning of the definition, the B -Bayesian distance is defined to be

$$H(B) = \sum_{i=1}^m E \left\{ \frac{\alpha_i^2 p_i(y, B)^2}{p(y, B)^2} \right\}$$

where

$$p(y, B) = \sum_{i=1}^m \alpha_i p_i(y, B)$$

It has been shown in reference 9 that $G(B) = G$ if and only if $H(B) = H$. In this connection, the authors of this paper plan to extend the variational results of the next section to include Bayesian distance.

In the next section, related new results, some of which concern questions raised earlier, will be outlined, and the connection between linear feature combination and the classical concept of statistical sufficiency will be explained.

RECENT RESULTS IN LINEAR FEATURE SELECTION

In what follows, we will be concerned with finding an extreme value of some function Φ (of the reduction matrix B). For example, it may be desired to choose $\Phi(B) = G(B)$ and find \hat{B} such that

$$\Phi(\hat{B}) = \min_B G(B)$$

or, perhaps, to choose $\Phi(B) = D(B)$ and find \hat{B} such that

$$\Phi(\hat{B}) = \max_B D(B)$$

In seeking an extremum of Φ , it is natural to consider the differentiability of Φ with respect to the elements of B . In the sequel, use is made of the *Gateaux differential* of Φ at B with increment C , denoted by $\delta\Phi(B; C)$ and defined (if the limit exists) by

$$\delta\Phi(B; C) = \lim_{s \rightarrow 0} \frac{\Phi(B + sC) - \Phi(B)}{s}$$

where C is a $k \times n$ matrix. If, for a given $k \times n$ matrix B of rank k , the previously defined limit exists for each $k \times n$ matrix C , then Φ is said to be *Gateaux differentiable* at B . Similarly, when the limit exists, the following is defined.

$$\delta p_i(y, B; C) = \lim_{s \rightarrow 0} \frac{p_i(y, B + sC) - p_i(y, B)}{s}$$

where C is a $k \times n$ matrix. For an excellent discussion of Gateaux differentials, see reference 10.

Theoretical results related to minimizing $G(B)$ for two multivariate normal classes with equal a priori probabilities and a one-dimensional feature space were initially presented by Guseman and Walker¹ (ref. 11). The associated computational procedure was presented by Guseman and Walker (ref. 12).

The following results for the general case of m n -dimensional normal classes with arbitrary a priori probabilities and a one-dimensional feature space appear in reference 13.

¹L. F. Guseman, Jr., and H. F. Walker, "On Minimizing the Probability of Misclassification for Linear Feature Selection," JSC Internal Tech. Note JSC-08412, NASA Johnson Space Center, Houston, Tex., Aug. 1973.

Lemma. Let \mathbf{B} be a nonzero $1 \times n$ vector. Then (omitting subscripts)

$$\delta p_i(y, \mathbf{B}; \mathbf{C}) = -p(y, \mathbf{B}) \left[\frac{\mathbf{C} \Sigma \mathbf{B}^T}{\mathbf{B} \Sigma \mathbf{B}^T} - \frac{\mathbf{C} \mu}{\mathbf{B} \Sigma \mathbf{B}^T} (y - \mathbf{B} \mu) - \frac{\mathbf{C} \Sigma \mathbf{B}^T}{(\mathbf{B} \Sigma \mathbf{B}^T)^2} (y - \mathbf{B} \mu)^2 \right]$$

for each $1 \times n$ vector \mathbf{C} .

Theorem 1. Let \mathbf{B} be a nonzero $1 \times n$ vector for which $\alpha_i p_i(y, \mathbf{B}) \neq \alpha_j p_j(y, \mathbf{B})$ for $i \neq j$. Then, G is Gateaux differentiable at \mathbf{B} , and

$$\delta G(\mathbf{B}; \mathbf{C}) = -\sum_{i=1}^m \alpha_i \int_{R_i(\mathbf{B})} \delta p_i(y, \mathbf{B}; \mathbf{C}) dy$$

Theorem 2. Let \mathbf{B} be a nonzero $1 \times n$ vector at which G assumes a minimum. Then, G is Gateaux differentiable at \mathbf{B} .

By substituting the expression for $\delta p_i(y, \mathbf{B}; \mathbf{C})$ given by the lemma into the expression from theorem 1, and using integration by parts, the following result is obtained.

Theorem 3. Let \mathbf{B} be a nonzero $1 \times n$ vector for which $\alpha_i p_i(y, \mathbf{B}) \neq \alpha_j p_j(y, \mathbf{B})$ for $i \neq j$. Then, G is Gateaux differentiable at \mathbf{B} , and

$$\delta G(\mathbf{B}; \mathbf{C}) = \sum_{i=1}^m \alpha_i p_i(y, \mathbf{B}) \left[\frac{\mathbf{C} \Sigma_i \mathbf{B}^T}{\mathbf{B} \Sigma_i \mathbf{B}^T} (y - \mathbf{B} \mu_i) + \mathbf{C} \mu_i \right] \Big|_{R_i(\mathbf{B})}$$

where the notation $|_{R_i(\mathbf{B})}$ denotes the sum of the values of the function at the right endpoints of the intervals comprising $R_i(\mathbf{B})$ minus the sum of its values at the left endpoints.

If $\hat{\mathbf{B}}$ is a nonzero $1 \times n$ vector that minimizes

$G(\mathbf{B})$, then $\hat{\mathbf{B}}$ must satisfy the vector equation

$$\frac{\partial G(\mathbf{B})}{\partial \mathbf{B}} = \begin{pmatrix} \delta G(\mathbf{B}; \mathbf{C}_1) \\ \vdots \\ \delta G(\mathbf{B}; \mathbf{C}_n) \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

where \mathbf{C}_j , $1 \leq j \leq n$, is a $1 \times n$ vector with a 1 in the j th slot and 0's elsewhere. Using the preceding formula for $\partial G(\mathbf{B})/\partial \mathbf{B}$ resulting from theorem 3, a numerically tractable expression for the variation in the probability of misclassification G with respect to \mathbf{B} is obtained. The use of this expression in a computational procedure for obtaining a nonzero $1 \times n$ $\hat{\mathbf{B}}$ that minimizes G was developed by Guseman and Marion (ref. 14).

If $\hat{\mathbf{B}}$ is a nonzero $1 \times n$ vector that minimizes G , then the entries $p_{ij}(\hat{\mathbf{B}})$ in the error matrix $P(\hat{\mathbf{B}})$ for the optimal classification procedure determined by the regions $R_i(\hat{\mathbf{B}})$ can be readily computed from the expression

$$p_{ij}(\hat{\mathbf{B}}) = \int_{R_i(\hat{\mathbf{B}})} p_j(y, \hat{\mathbf{B}}) dy, \quad i, j = 1, 2, \dots, m$$

The linear feature selection procedure for minimizing $G(\mathbf{B})$ has been extended to the case where the density function for each class is a convex combination of multivariate normals. This extension allows for the design of a one-dimensional "class A—not class A" classification procedure that could be used (for example) to classify wheat(s) versus nonwheat(s). The associated computational procedure for this extension was developed by Guseman and Marion (ref. 15).

Decell and Quirein (ref. 6) develop explicit expressions for $\delta D(\mathbf{B}; \mathbf{C})$ and $\delta \rho(\mathbf{B}; \mathbf{C})$ in terms of \mathbf{B} and the known means and covariance matrices μ_i and Σ_i , $1 \leq i \leq m$. These expressions immediately provide $\partial(D(\mathbf{B}))/\partial \mathbf{B}$ and $\partial(\rho(\mathbf{B}))/\partial \mathbf{B}$ for use in a Davidon-Fletcher-Powell (ref. 16) iteration scheme for determination of an extremum value of $D(\mathbf{B})$ and $\rho(\mathbf{B})$, respectively.

The explicit expressions are

$$\frac{\partial(D(\mathbf{B}))}{\partial \mathbf{B}} = \frac{-2}{m^2 - m} \sum_{i=1}^m (\mathbf{B}\Sigma_i\mathbf{B}^T)^{-1} \cdot \left[(\mathbf{B}\Sigma_i\mathbf{B}^T) (\mathbf{B}\Sigma_i\mathbf{B}^T)^{-1} \mathbf{B}\Sigma_i - \mathbf{B}\Sigma_i \right]$$

and

$$\frac{\partial(\rho(\mathbf{B}))}{\partial \mathbf{B}} = \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{\partial \rho_{\mathbf{B}}(i,j)}{\partial \mathbf{B}}$$

where

$$\begin{aligned} \frac{\partial(\rho_{\mathbf{B}}(i,j))}{\partial \mathbf{B}} = & -\frac{1}{2} \left[\mathbf{B}(\Sigma_i + \Sigma_j)\mathbf{B}^T \right]^{-1} \left\{ (\mathbf{B}\delta_{ij}\delta_{ij}\mathbf{B}^T) \right. \\ & \cdot \left[\mathbf{B}(\Sigma_i + \Sigma_j)\mathbf{B}^T \right]^{-1} \mathbf{B}(\Sigma_i + \Sigma_j) \\ & \left. - \mathbf{B}\delta_{ij}\delta_{ij}\mathbf{B}^T \right\} + \left[\mathbf{B}(\Sigma_i + \Sigma_j)\mathbf{B}^T \right]^{-1} \\ & \cdot \mathbf{B}(\Sigma_i + \Sigma_j) - \frac{1}{2} \left[(\mathbf{B}\Sigma_i\mathbf{B}^T)^{-1} \mathbf{B}\Sigma_i \right. \\ & \left. + (\mathbf{B}\Sigma_j\mathbf{B}^T)^{-1} \mathbf{B}\Sigma_j \right] \end{aligned}$$

It is also shown in reference 6 that, in general, an absolute extremum of $G(\mathbf{B})$, $\rho(\mathbf{B})$, and $D(\mathbf{B})$ always exists. For any one of the given functions $G(\mathbf{B})$, $\rho(\mathbf{B})$, or $D(\mathbf{B})$, the absolute extremum is attained at $\mathbf{B} = (I_k | Z)U$, where I_k is the $k \times k$ identity and U is some $n \times n$ unitary matrix, thus parameterizing the aforementioned extreme problems on the compact group of unitary matrices. In reference 17, it is shown that the nature of the eigenvalues of U in no way provides any information about the extreme

values of $D(I_k | Z)U$. In reference 18, these results were refined in the sense that any extremal transformation can be expressed in the form $\mathbf{B} = (I_k | Z)H_p \dots H_1$, where $p \leq \min\{k, n - k\}$ and H_i is a Householder transformation $i = 1, \dots, p$. The latter result suggests constructing a sequence of transformations $(I_k | Z)H_1, (I_k | Z)H_2H_1, \dots$ such that the values of the class separability criterion (e.g., $G(\mathbf{B})$, $\rho(\mathbf{B})$, $D(\mathbf{B})$) evaluated at this sequence comprise a bounded monotone sequence of real numbers. The construction of the i th element of the sequence of transformations requires the solution of an n -dimensional optimization problem. Recall that $T(H)$, the Householder transformations (refs. 19 and 20) $H = I - 2\mathbf{x}\mathbf{x}^T$, $\mathbf{x} \in R^n$ with $\|\mathbf{x}\| = 1$, is a compact connected subset of the unitary matrices for which $H^T = H = H^{-1}$. Some of these results are outlined beginning with the definition (for a case, say, when it is desired to maximize Φ):

$$\Phi(I_k | Z) H_1 = \underset{H \in T(H)}{\text{1.u.b.}} \Phi(I_k | Z) H$$

Theorem 4. For each positive integer i , let the element H_i of $T(H)$ be chosen such that

$$\Phi(I_k | Z) H_i H_{i-1} \dots H_1 = \underset{H \in H_n}{\text{1.u.b.}} \Phi(I_k | Z) H H_{i-1} \dots H_1$$

then

- (1) $\Phi(I_k | Z) H_i \dots H_1 \leq \Phi(I_k | Z) H_{i+1} H_i \dots H_1$
- (2) $\Phi(I_k | Z) H_i \dots H_1 H \leq \Phi(I_k | Z) H_{i+1} H_i \dots H_1$ for every $H \in T(H)$
- (3) $\Phi(I_k | Z) H H_i \dots H_1 \leq \Phi(I_k | Z) H_{i+1} H_i \dots H_1$ for every $H \in T(H)$
- (4) $\Phi(I_k | Z) H_i \dots H_{i-p+1} H H_{i-p+1} \dots H_1 \leq \Phi(I_k | Z) H_{i+1} H_i \dots H_1$ for every $H \in T(H)$ and $p = 0, \dots, i - 2$

Theorem 5. The sequence

$$\left\{ \Phi(I_k | Z) H_i \dots H_1 \right\}_{i=1}^{\infty}$$

is bounded above and

$$\lim_{i \rightarrow \infty} \Phi(I_k|Z)_{H_i \dots H_1} = \text{l.u.b.} \left\{ \Phi(I_k|Z)_{H_i \dots H_1} \right\}$$

These theorems give rise to a sequential monotone procedure for possibly obtaining a Φ -extremal rank- k linear-combination matrix. At each stage in this procedure, the extremal problem is a function of only n variables. One can conjecture, under certain conditions, that the process should terminate in at most $\min\{k, n - k\}$ steps. The conjecture is clearly in line with the $\min\{k, n - k\}$ representation of the actual Φ -extremal solution. Certainly, the conjecture could further depend on any pathological behavior of Φ and Tally² constructs such a pathological failure point. Tally³ shows that the procedure actually converges to a Φ extremum provided Φ is $\mathcal{T}(H)$ -sloped. (See definition 1.) Some of these results will be outlined. Let U denote the set of unitary matrices and $\mathcal{T}(H)$ the Householder transformations.

Definition 1. Φ will be called $\mathcal{T}(H)$ -sloped provided that $U \in U$ and $\Phi(U) < \Phi_{\max}$ imply there exists some $H \in \mathcal{T}(H)$ (dependent on U) such that

$$\Phi(U) < \Phi(HU) \leq \Phi_{\max} = \text{l.u.b.} \Phi(U)$$

Definition 2. A sequence $\{U_i\}_{i=1}^{\infty}$ in U will be called Φ -convergent provided $\{\Phi(U_i)\}_{i=1}^{\infty}$ converges.

Definition 3. A sequence $\{U_i\}_{i=1}^{\infty}$ in U will be called a Φ -Householder sequence provided that $H \in H$ and i , an integer, imply

$$(1) \Phi(U_i) < \Phi(U_{i+1})$$

$$(2) \Phi(HU_i) < \Phi(U_{i+1})$$

²W. Tally, "On the Convergence of Optimal Linear Combination Procedures," *Comp. & Maths. with Appls.* (To be published).

³W. Tally, "A Convergence Criterion for Optimal Linear Combination Procedures," *Comp. & Maths. with Appls.* (To be published).

Proposition 1. Each Φ -Householder sequence $\{U_i\}_{i=1}^{\infty}$ is Φ -convergent and

$$\lim_i \Phi(U_i) = \Phi(U) = \text{l.u.b.} \Phi(U_i)$$

for some $U \in U$.

Proposition 2. Each Φ -Householder sequence converges to Φ_{\max} if and only if Φ is $\mathcal{T}(H)$ -sloped.

Proposition 3. If $\{U_i\}_{i=1}^{\infty}$ is a Φ -Householder sequence and Φ is $\mathcal{T}(H)$ -sloped, then exactly one of the following

- (1) $\{\Phi(U_i)\}_{i=1}^{\infty}$ is strictly monotonic (and convergent to Φ_{\max})
- (2) for some integer k ,

$$\text{l.u.b.} \Phi(HU_k) < \Phi(U_k)$$

(in which case, $\Phi(U_k) \equiv \Phi_{\max}$!)

These techniques have been applied to the functions $\Phi(\mathbb{B}) = D(\mathbb{B})$ and $\Phi(\mathbb{B}) = \rho(\mathbb{B})$, respectively, by Decell and Mayekar (ref. 21) and Decell and Marani (ref. 22) using C1 flight line data.

In each case, explicit expressions for $(\partial/\partial x)$ $\{D(I_k|Z)H\}$ and $(\partial/\partial x) \{\rho(I_k|Z)H\}$, where $H = I - 2xx^T, \|x\| = 1$, have been developed for m -pattern classes (equal proportions) and have been used sequentially, according to the aforementioned theorems, to calculate the extreme values and the unitary matrices (as products of elements of $\mathcal{T}(H)$) at which the extreme values occur. Some of the results are outlined in what follows.

Let

$$\Sigma_{ij} = \Sigma_i + \Sigma_j$$

$$J_{ij} = \Sigma_{ij}H(I_k|Z)^T \left[(I_k|Z)H\Sigma_{ij}H(I_k|Z)^T \right]^{-1}$$

$$K_{ij} = \Sigma_i H(I_k|Z)^T \left[(I_k|Z)H\Sigma_i H(I_k|Z)^T \right]^{-1}$$

and

$$L_{ij} = \Sigma_j H(I_k|Z)^T \left[(I_k|Z)H\Sigma_j H(I_k|Z)^T \right]^{-1}$$

Let

$$\hat{Q}_Y = \left[\mathbf{x}\mathbf{x}^T Q_Y(I_k|Z) - Q_Y(I_k|Z) \mathbf{x}\mathbf{x}^T \right]^T \\ - \left[\mathbf{x}\mathbf{x}^T Q_Y(I_k|Z) - Q_Y(I_k|Z) \mathbf{x}\mathbf{x}^T \right]$$

and let \hat{J}_Y , \hat{K}_Y and \hat{L}_Y be similarly defined by substituting, respectively, J_Y , K_Y and L_Y for Q_Y in the expression for \hat{Q}_Y , $l, j = 1, \dots, m$. The resulting expressions are

$$\frac{\partial}{\partial \mathbf{x}} \left[\rho(I_k|Z) H \right] = \frac{1}{m} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{\exp(F_{ij} + G_{ij})}{(\mathbf{x}^T \mathbf{x})^2} \\ \cdot (\hat{Q}_{ij} + 2\hat{J}_{ij} - \hat{K}_{ij} - \hat{L}_{ij}) \mathbf{x}$$

where

$$\delta_{ij} = (\mu_i - \mu_j) (\mu_i - \mu_j)^T$$

$$\text{tr}(\cdot) = \text{trace of } (\cdot) \text{ and } |\cdot| = \det(\cdot)$$

$$F_{ij} = -\frac{1}{4} \text{tr} \left\{ \left((I_k|Z) H \Sigma_{ij} H (I_k|Z)^T \right)^{-1} \right. \\ \left. \cdot (I_k|Z) H \delta_{ij} H (I_k|Z)^T \right\}$$

and

$$G_{ij} = -\frac{1}{2} \ln \left| (I_k|Z) H \Sigma_{ij} H (I_k|Z)^T \right| \\ + \frac{1}{4} \ln \left| (I_k|Z) H \Sigma_j H (I_k|Z)^T \right| \\ + \frac{1}{4} \ln \left| (I_k|Z) H \Sigma_i H (I_k|Z)^T \right| + \frac{k}{2} \ln 2$$

$$\frac{\partial}{\partial \mathbf{x}} \left[D(I_k|Z) H \right] = \frac{-2}{(\mathbf{x}^T \mathbf{x})^2} \\ \cdot \sum_{i=1}^m \left\{ (M_i - N_i)^T - (M_i - N_i) \right\} \mathbf{x}$$

where

$$M_i = \mathbf{x}\mathbf{x}^T Q_i(I_k|Z)$$

$$N_i = Q_i(I_k|Z) \mathbf{x}\mathbf{x}^T$$

$$Q_i = \left[\left\{ S_i \mathbf{B}^T - \Sigma_i \mathbf{B}^T (\mathbf{B} \Sigma_i \mathbf{B}^T)^{-1} (\mathbf{B} S_i \mathbf{B}^T) \right\} (\mathbf{B} \Sigma_i \mathbf{B}^T)^{-1} \right]$$

$$\mathbf{B} = (I_k|Z) (I - 2\mathbf{x}\mathbf{x}^T)$$

Peters et al.⁴ approach the problem of finding a minimum of $G(B)$ from the point of view of treating the mapping $B: R^n \rightarrow R^k$ (for some $k \leq n$) as a statistic, and they provide necessary and sufficient conditions that such a B be a sufficient statistic in the classical sense of Halmos and Savage (ref. 23), Lehmann and Scheffe (ref. 24), Bahadur (ref. 25), LeCam (ref. 26), and Kullback (ref. 5). Although their results are much more general than required for dealing with the dimensionality reduction problem for a finite number of normal populations, the application they provide for such families actually allows one to write down the optimal dimension-reducing $k \times n$ statistic B such that $G(B) = G$ (whenever such a B exists). Moreover, they also guarantee that there is no other B of smaller rank (i.e., of rank less than k) for which $G(B) = G$. Their application to the problem will simply be stated; the reader is referred to Peters et al.⁴ for the more general applications to exponential families (e.g., Wishart and normal multivariate sampling).

⁴B. C. Peters, Jr., R. Redner, and H. P. Decell, Jr., "Characterizations of Linear Sufficient Statistics," submitted to Sankhya, A, 1976.

Let $N(\mu_l, \Sigma_l)$, $l = 0, 1, \dots, m-1$, be an n -variate normal family with $\mu_0 = \Theta$ and $\Sigma_0 = I$, each member having density

$$p_l(x) = (2\pi)^{-\frac{n}{2}} |\Sigma_l|^{-\frac{1}{2}} \exp \left[-\frac{1}{2} (x - \mu_l)' T \Sigma_l^{-1} (x - \mu_l) \right]$$

The requirement $\mu_0 = \Theta$ and $\Sigma_0 = I$ imposes no loss of generality since there exists a nonsingular matrix M_0 for which $M_0 \Sigma_0 M_0^T = I$ and a change of coordinate system defined by the transformation $x \rightarrow M_0(x - \mu_0)$ allows one to recover the sufficient statistic in the original coordinate system.

Theorem 6. Let $\mu_0 = \Theta$, $\Sigma_0 = I$, and

$$M = \begin{bmatrix} \mu_1 & \mu_2 & \dots & \mu_{m-1} & \Sigma_1 & I & \Sigma_2 & I & \dots & \Sigma_{m-1} & I \end{bmatrix}$$

Matrix B is a linear sufficient statistic for the given finite n -variate normal family if and only if $\text{range}(B^T) = \text{range}(M)$. Moreover, $k = \text{rank } M$ is the smallest integer for which there exists a $k \times n$ sufficient statistic for the given family.

Again, note that theorem 6 completely determines the greatest dimensionality reduction possible such that $G(B) = G$. Moreover, as will be shown by example in what follows, there are any number of ways of finding a B such that $\text{range}(B^T) = \text{range}(M)$. In fact, the theorem states that if $\text{rank } M = n$, then there is no dimension-reducing sufficient statistic (i.e., $G(B) > G$ for every $k \times n$ matrix B for which $k < n$).

The following result due to Decell et al.⁵ provides one means of calculating (and determining the existence of) the aforementioned sufficient statistic B for which $G(B) = G$.

Theorem 7. Let Π_l be an n -variate normal population with a priori probability $\alpha_l > 0$, mean μ_l , and covariance Σ_l ; $l = 0, 1, \dots, m-1$ (with $\mu_0 = \Theta$, $\Sigma_0 = I$), and let

$$FG = M = \begin{bmatrix} \mu_1 & \mu_2 & \dots & \mu_{m-1} & \Sigma_1 & I & \Sigma_2 & I & \dots & \Sigma_{m-1} & I \end{bmatrix}$$

be a full-rank ($k \leq n$) decomposition of M . Then, the n -variate Bayes procedure assigns x to Π_l if and only if the k -variate Bayes procedure assigns $F^T x$ to Π_l . Moreover, k is the smallest integer for which there exists a $k \times n$ matrix T preserving the Bayes assignment of x and Tx to Π_l ; $l = 0, 1, \dots, m-1$.

These results completely characterize the nature of data compression for the Bayes classification procedure for normal classes in the sense that k is the dimension of greatest allowable data compression consistent with preserving Bayes population assignment. Moreover, the theorem provides an explicit expression for the compression matrix T depending only on the known population means and covariances. The statistic $T \equiv F^T$ given by the theorem is by no means unique (e.g., for any nonsingular $k \times k$ matrix A , $T \equiv AF^T$ will do). It is also true that there may be more efficient methods for calculating the statistic T (yet to be determined) than the method of full-rank decomposition of M .

It should be noted that the matrix M has an "excellent chance" of having rank equal to n . Even in the case of two populations ($m = 2$), there may well be n linearly independent columns among the $2(n+1)$ columns of M . Consequently, there do not exist an integer k and a $k \times n$ rank- k compression matrix T preserving the Bayes assignment of x and Tx .

Peters (ref. 27) treats the problem of determining sufficient statistics for mixtures of probability measures in a homogeneous family. The reader is referred to Teicher (refs. 28 and 29), Yakowitz (ref. 30), and Yakowitz and Spragins (ref. 31) for the treatment of homogeneous families.

When the linear feature selection techniques mentioned previously are used in a LACIE-type application, they are based on the assumptions that each class conditional density function is multivariate normal and that the associated parameters (μ_l, Σ_l , $1 \leq l \leq m$) are known or can be estimated. In some cases, either the normality assumptions may be violated or else the determination of the number of classes present and their associated parameters is not possible. The question then arises as to how, without having samples from each class, one might perform a dimensionality reduction without losing much of the "separation" present in measurement space. For ex-

⁵H. P. Decell, Jr., P. L. Odell, and W. A. Coberly, "Linear Dimension Reduction and Bayes Classification," J. American Statist. Assoc., submitted Jan. 1978.

ample, one might be interested in displaying a registered multipass Landsat data set on a three-color display device without a priori knowledge of class structure in the data.

Each of the previous linear feature selection techniques uses a statistical definition of the word "separation." The following procedure, due to Bryant and Guseman (ref. 7), makes no statistical assumptions about the data. In addition, no labeled subsets (training data) are required. In this sense, the linear feature selection technique outlined in the following is distribution free. Basically, the problem can be stated as follows.

Given distinct (prototype) vectors x_1, x_2, \dots, x_p in R^n and k ($1 \leq k < n$), determine a linear transformation $A: R^n \rightarrow R^k$ which minimizes

$$F(A) = \sum_{1 \leq i < j \leq p} (\|x_i - x_j\| - \|Ax_i - Ax_j\|)^2$$

where the norms $\|x_i - x_j\|$ and $\|Ax_i - Ax_j\|$ are the euclidean norms in R^n and R^k , respectively. Let $m = p(p-1)/2$ and let $\{z_i; 1 \leq i \leq m\}$ denote the m distinct differences of the prototypes x_j . If $A = (a_{ij})_{k \times n}$, $z_i = (z_{i1}, \dots, z_{im})^T$, and $a' = (a_{i1}, \dots, a_{im})^T$, then the gradient of F at A is given by

$$\frac{\partial F(A)}{\partial A} = AS - AT(A)$$

where S is the $n \times n$ matrix

$$S = \left(\sum_{i=1}^m z_{iq} z_{ir} \right)$$

and $T(A)$ is the $n \times n$ matrix

$$T(A) = \left(\sum_{i=1}^m \frac{\|z_i\|}{\|Az_i\|} z_{iq} z_{ir} \right)$$

Standard optimization techniques can be used to obtain \hat{A} , which minimizes F .

For a given data set (e.g., a Landsat sample segment), there are several ways to choose the pro-

tototype vectors $x_i, 1 \leq i \leq m$. For example, one might choose cluster centers from the output of a clustering algorithm.

CONCLUDING REMARKS

There are, of course, ad hoc feature selection procedures based on specific problem knowledge and empirical studies. An example of such a procedure is the transformation of Kauth and Thomas (ref. 32) used in the analysis of Landsat data. This transformation is based on an empirical data study and is described by an orthogonal coordinate change $U: R^4 \rightarrow R^4$. Application of the transform U to Landsat measurements simply produces a reduced feature space of dimension 2 (brightness-greenness). This is essentially accomplished at each Landsat measurement $X = (x_1, x_2, x_3, x_4)^T$ by the mapping:

$$\lambda \rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_{11} & \dots & u_{14} \\ \vdots & & \vdots \\ u_{41} & \dots & u_{44} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_2 \end{pmatrix}$$

The Kauth-Thomas transform has proven to be of value in LACIE applications (e.g., physical interpretation, dimensionality reduction, scatter plots). As one would expect, the Kauth-Thomas transform is not a sufficient statistic nor will it, in general, preserve Landsat Bayes class assignment in feature space.

Feature selection techniques are currently being studied as a tool for "optimum pass" selection problems in LACIE. The basic objective is to develop a technique for a priori selection (based on some separability criterion) of subsets of Landsat acquisitions for analysis to separate wheat from nonwheat when given an adequate sample of labeled wheat and nonwheat LACIE segment pixel data. There are preliminary results in this direction due to Guseman and Marion (ref. 33) using one-dimensional feature selection that minimizes $G(B)$.

In still another LACIE application, studies are being performed on parametric and nonparametric feature selection techniques that allow analyst/interpreters to better separate spring wheat from other small grains in a reduced feature space (e.g., brightness-greenness). In this connection, labeled wheat and other small-grains LACIE-segment pixel

data and ancillary data are being used to estimate the distribution functions for spring wheat and other spring small grains. Feature selection methods are being used to find a priori statistically optimum features and associated discriminant functions. These will be compared to the brightness and greenness features currently used at the NASA Johnson Space Center.

Methods for estimating class proportions, based on the linear feature selection procedure for minimizing $G(B)$, have been developed by Guseman and Walton⁶ (ref. 34). In both papers, the proportion estimation techniques rely on the fact that one can readily compute the error matrices associated with the optimal classifier produced by the linear feature selection procedure.

Other results of general related interest appear in Babu and Kalra (ref. 35), Kadota and Shepp (ref. 36), Marill and Green (ref. 37), Swain and King (ref. 38), Tou and Heydorn (ref. 39), Watanabe et al. (ref. 40), Wee (ref. 41), and Wheeler et al. (ref. 42).

REFERENCES

1. Anderson, T. W.: An Introduction to Multivariate Statistical Analysis. Wiley, 1958.
2. Andrews, H. C.: Introduction to Mathematical Techniques in Pattern Recognition. Wiley-Interscience, 1972.
3. Kailath, T.: The Divergence and Bhattacharyya Distance Measures in Signal Selection. IEEE Trans. Comm. Theory, Com.-15, no. 1, 1967, pp. 52-60.
4. Kanal, L.: Patterns in Pattern Recognition: 1968-1974. IEEE Trans. Pat. Recog., IT-20, 1974, pp. 697-722.
5. Kullback, S.: Information Theory and Statistics. Wiley, 1969.
6. Decell, H. P., Jr.; and Quirein, J. A.: An Iterative Approach to the Feature Selection Problem. IEEE Conference on Remotely Sensed Data, Cat. no. CH0834-2 GE, 1973, pp. 3B-01 to 3B-12.
7. Bryant, J.; and Guseman, L. F., Jr.: Distribution-Free Linear Feature Selection. Report 12, NASA Contract NAS-9-14689-5S, Dept. of Math., Texas A & M University, College Station, Tex., June 1977.
8. Devijver, P. A.: On a New Class of Bounds on Bayes Risk in Multihypothesis Pattern Recognition. IEEE Trans. on Computers, C-23, no. 1, 1974, pp. 70-80.
9. Guseman, L. F., Jr.; Peters, B. C., Jr.; and Swasdee, M.: Estimating Normal Mixture Parameters From the Distribution of a Reduced Feature Vector. Report 11, NASA Contract NAS-9-14689-4S, Dept. of Math., Texas A & M University, College Station, Tex., Nov. 1976.
10. Luenberger, D. G.: Optimization by Vector Space Methods. Wiley, 1969.
11. Guseman, L. F., Jr.; and Walker, H. F.: The Differentiability of the Probability of Misclassification as a Function of a Linear Feature Selection Matrix. Report 28, NASA Contract NAS-9-12777, Dept. of Math., University of Houston, Houston, Tex., 1973.
12. Guseman, L. F., Jr.; and Walker, H. F.: On Minimizing the Probability of Misclassification for Linear Feature Selection: A Computational Procedure. In The Search for Oil, D. B. Owens, ed., Marcel Dekker, 1975.
13. Guseman, L. F., Jr.; Peters, B. C., Jr.; and Walker, H. F.: On Minimizing the Probability of Misclassification for Linear Feature Selection. Ann. Statist., vol. 3, 1975, pp. 661-668.
14. Guseman, L. F., Jr.; and Marion, B. P.: LFSPMC: Linear Feature Selection Program Using the Probability of Misclassification. Report 3, NASA Contract NAS-9-13894, Dept. of Math., Texas A & M University, College Station, Tex., Jan. 1975.
15. Guseman, L. F., Jr.; and Marion, B. P.: LFSPMC (VERSION 2): Linear Feature Selection Program Using the Probability of Misclassification. Report 6, NASA Contract NAS-9-14689-4S, Dept. of Math., Texas A & M University, College Station, Tex., Mar. 1976.
16. Fletcher, R.; and Powell, M. J. D.: A Rapidly Convergent Descent Method for Minimization. Comput. J., vol. 6, 1963, pp. 163-168.
17. Brown, D. R.; and O'Malley, M. J.: The Role of Eigenvalues in Linear Feature Selection Theory. J. Comput. Appl. Math., vol. 3, 1977, pp. 243-250.
18. Decell, H. P., Jr.; and Smiley, W. G., III: Optimal Linear Combinations Using Householder Transformations. Commun. Statist.—Theor. Meth., A5, vol. 12, 1976, pp. 1153-1162.
19. Householder, A. S.: Unitary Triangularization of a Nonsymmetric Matrix. J. Assoc. Comput. Mach., vol. 5, 1958, pp. 339-342.
20. Householder, A. S.: The Theory of Matrices in Numerical Analysis. Blaisdell, 1964.
21. Decell, H. P., Jr.; and Mayekar, S. M.: Feature Combinations and the Divergence Criterion. Comp. & Maths. with Appls., vol. 3, 1977, pp. 71-76.

⁶L. F. Guseman, Jr., and J. R. Walton, "Methods for Estimating Proportions of Convex Combinations of Normals Using Linear Feature Selection," Commun. Statist.—Theor. Meth., submitted 1978.

22. Decell, H. P., Jr.; and Marani, S. K.: Feature Combinations and the Bhattacharyya Criterion. *Commun. Statist.—Theor. Meth.*, A5, vol. 12, 1976, pp. 1143-1152.
23. Halmos, P. R.; and Savage, L. J.: Application of the Randon-Nikodym Theorem to the Theory of Sufficient Statistics. *Ann. Math. Statist.*, vol. 20, 1949, pp. 225-241.
24. Lehmann, E. L.; and Scheffe, H.: Completeness, Similar Regions, and Unbiased Estimation. *Sankhya, A*, vol. 10, 1950, pp. 305-340.
25. Bahadur, R. R.: Sufficiency and Statistical Decision Function. *Ann. Math. Statist.*, vol. 25, 1954, pp. 423-463.
26. LeCam, L.: Sufficiency and Approximate Sufficiency. *Ann. Math. Statist.*, vol. 35, 1964, pp. 1419-1455.
27. Peters, B. C., Jr.: Sufficient Statistics for Mixtures of Measures in a Homogeneous Family. Report 64, NASA Contract NAS-9-15000, Dept. of Math., University of Houston, Houston, Tex., Mar. 1977.
28. Teicher, H.: Identifiability of Mixtures. *Ann. Math. Statist.*, vol. 32, 1961, pp. 244-248.
29. Teicher, H.: Identifiability of Finite Mixtures. *Ann. Math. Statist.*, vol. 34, 1963, pp. 1265-1269.
30. Yakowitz, S.: Consistent Estimator for the Identification of Finite Mixtures. *Ann. Math. Statist.*, vol. 40, 1969, pp. 1728-1735.
31. Yakowitz, S.; and Spragins, J.: On the Identifiability of Finite Mixtures. *Ann. Math. Statist.*, vol. 39, 1968, pp. 209-214.
32. Kauth, R. J.; and Thomas, G. S.: The Tasselled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. Symposium on Machine Processing of Remotely Sensed Data, Laboratory for Applications of Remote Sensing, Purdue University, W. Lafayette, Ind., June 29-July 1, 1976.
33. Guseman, L. F., Jr.; and Marion, B. P.: Optimal Selection of Passes. Report 2, NASA Contract NAS-9-13894-1S, Dept. of Math., Texas A & M University, College Station, Tex., Jan. 1976.
34. Guseman, L. F., Jr.; and Walton, J. R.: An Application of Linear Feature Selection to Estimation of Proportions. *Commun. Statist.—Theor. Meth.*, A6, vol. 7, 1977, pp. 611-617.
35. Babu, C. C.; and Kalra, S. N.: On Feature Extraction in Multiclass Pattern Recognition. *Int. J. Control*, vol. 15, 1972, pp. 595-601.
36. Kadota, T. T.; and Shepp, I. A.: On the Best Finite Set of Linear Observables for Discriminating Two Gaussian Signals. *IEEE Trans. Inform. Theory*, IT-13, 1967, pp. 278-284.
37. Marill, T.; and Green, D. M.: On the Effectiveness of Receivers in Recognition Systems. *IEEE Trans. Inform. Theory*, IT-9, 1963, pp. 11-17.
38. Swain, P. H.; and King, R. C.: Two Effective Feature Selection Criteria for Multispectral Remote Sensing. LARS Info. Note 042673, Laboratory for Application of Remote Sensing, Purdue University, W. Lafayette, Ind., Apr. 1973.
39. Tou, J.; and Heydorn, R.: Some Approaches to Optimum Feature Extraction. In *Computer and Information Sciences*, vol. 2, J. Tou, ed., Academic Press, 1967.
40. Watanabe, S.; et al.: Evaluation and Selection of Variables in Pattern Recognition. In *Computer and Information Sciences*, vol. 11, J. Tou, ed., Academic Press, 1967.
41. Wee, W. G.: A Survey of Pattern Recognition. Systems and Research Division, Honeywell, Inc., St. Paul, Minn. (undated).
42. Wheeler, S. G.; Misra, P. N.; and Holmes, Q. A.: Linear Dimensionality of LANDSAT Agricultural Data With Implications for Classification. Symposium on Machine Processing of Remotely Sensed Data, LARS, Purdue University, June 29-July 1, 1976.

Feature Extraction Applied to Agricultural Crops As Seen by Landsat

R. J. Kauth,^a P. F. Lambeck,^a W. Richardson,^a G. S. Thomas,^a and A. P. Pentland^a

INTRODUCTION AND SCOPE

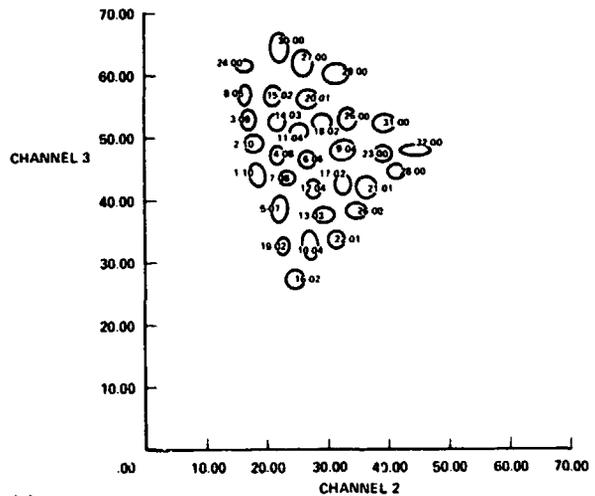
The physical interpretation of the spectral-temporal structure of Landsat data can be conveniently described in terms of a graphic descriptive model which has been named the Tasseled Cap. This model of Landsat data has been a rich source of development not only in crop-related feature extraction but also for data screening and for haze effects correction. The authors will first describe this model qualitatively and indicate its applications and then use it to analyze several feature extraction algorithms.

The Tasseled Cap

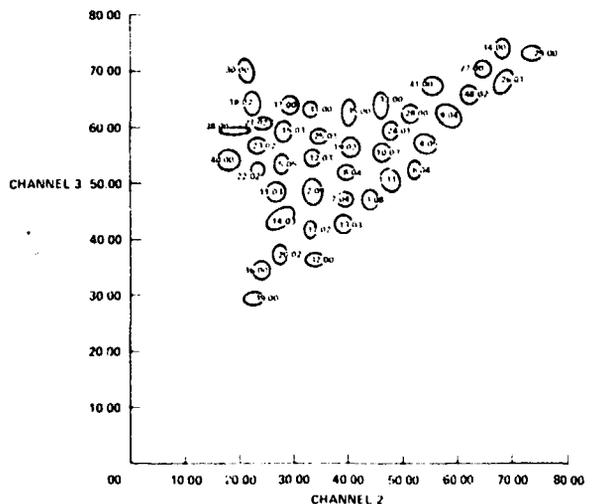
An examination of scatter plots of Landsat data from a number of different sites (fig. 1) discloses several distinct characteristics of the data. When plots from multispectral scanner (MSS) channels 2 and 3 are examined, a roughly triangular shape above the diagonal of the scatter plot can be seen. In a scene the size of a LACIE sample segment, this triangle is seldom "full" of data—usually some part of it is missing so that one might easily miss seeing the shape. However, if scatter plots from several segments are overlaid, as shown in figure 2, the overall pattern becomes visible. One of the authors, G. Thomas, hypothesized a physical explanation of these patterns as follows (ref. 1).

"In order to achieve a better understanding of just what is portrayed in the cluster patterns and why a general or 'complete' cluster pattern has the shape it does, ERIM's vegetative canopy model (ref. 2) was called into play. As it happened, the necessary model inputs from a certain type of vegetation, *Ionia* wheat (a variety grown in Michigan) were readily available.

^aEnvironmental Research Institute of Michigan (ERIM), Ann Arbor, Michigan.



(a)



(b)

FIGURE 1.—General cluster patterns in channels 2 and 3. (a) Livingston County, Illinois, July 16, 1973. (b) Fayette County, Illinois, June 11, 1973.

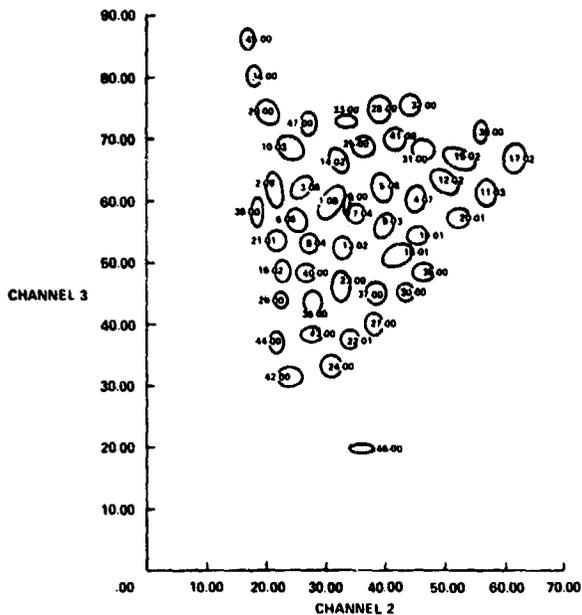


FIGURE 2.—Overlay of figure 1 cluster patterns of Fayette County, July 16, 1973.

And so, two soil reflectances were selected, one to simulate a darker, perhaps more organic or moist soil and the other to simulate a lighter colored, perhaps sandier or drier soil (for more information on the importance of soil moisture on soil reflectance, see Blanchard et al., 1974 (ref. 3), and Parks et al., 1974 (ref. 4)) and a construction made of the phenology of a sample of wheat, Ionia variety with two very different soil backgrounds (fig. 3). As may be seen, the soil background plays a dominant role in the bidirectional reflectance of a stand of Ionia wheat until the onset of plant maturity. If the bare soil points are connected by a line, hereafter called the bare soil line, the outline of the phenology of Ionia wheat is very similar to the outline of the 'complete' cluster pattern. It is not unreasonable to suspect, therefore, that location within a cluster pattern represents, to a degree, vegetative state of development as modified by such factors as soil reflectance, stress of various kinds, mixtures of vegetation and so on." (Soil reflectances are taken from reference 5.)

These same observations are expressed as an artistic conception in figure 4(a), showing a scatter plot of band 5 versus band 6. Band 5 is centered in the chlorophyll absorption band of green vegetation around 0.65 micrometer, whereas band 6 is centered on the cellulose reflectance peak of green vegetation around 0.75 micrometer. The signal from green vegetation is thus found to be small in band 5 and

large in band 6. Such a point is indicated in figure 4(a) by the designation "green stuff."

The main variability found in soils is their brightness. Hence, the signals from bare soils are distributed primarily along a line radiating from the origin.

Additional observations may be made. When looking at a scatter plot of band 4 versus band 5, one sees the data take a narrow cigar shape, as shown in figure 4(b). The physical explanation is that band 4 is centered on the cellulose reflectance around 0.55 micrometer but extends significantly into the chlorophyll absorption region so that bands 4 and 5 are highly correlated. (There are, however, differences between bands 4 and 5, which will be discussed later.)

When looking at figure 4(b), one imagines he is seeing the "edge" of a three-dimensional object embedded in four-dimensional Landsat space. A scatter plot of band 6 versus band 7 would show the same thing. Does this mean that the three-dimensional object is in fact nearly planar; i.e., two dimensional? Can the data be projected in such a way as to enable seeing its structure more clearly?

Such considerations led one of the authors¹ (ref. 6) to define axes of maximum variation in the Landsat data and to ascribe physical interpretation to these axes. First, a collection of points along the diagonal soil line was chosen from scatter plots and a

¹R. J. Kauth, "Soil Reflectance," NASA Johnson Space Center Memo TF3-75-5-190, Apr. 10, 1975.

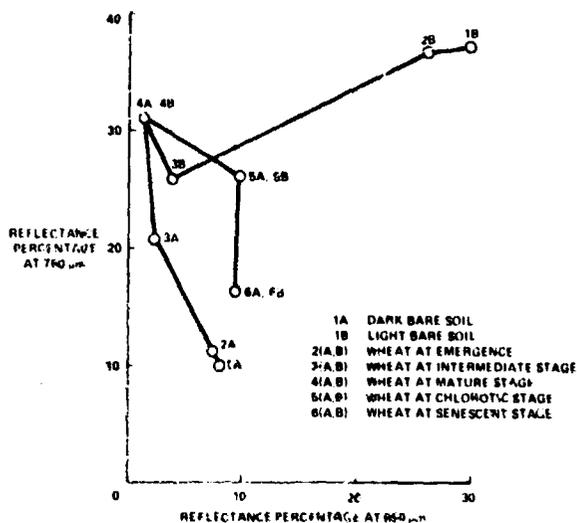


FIGURE 3.—Phenology for wheat (Ionia variety) based on canopy model.

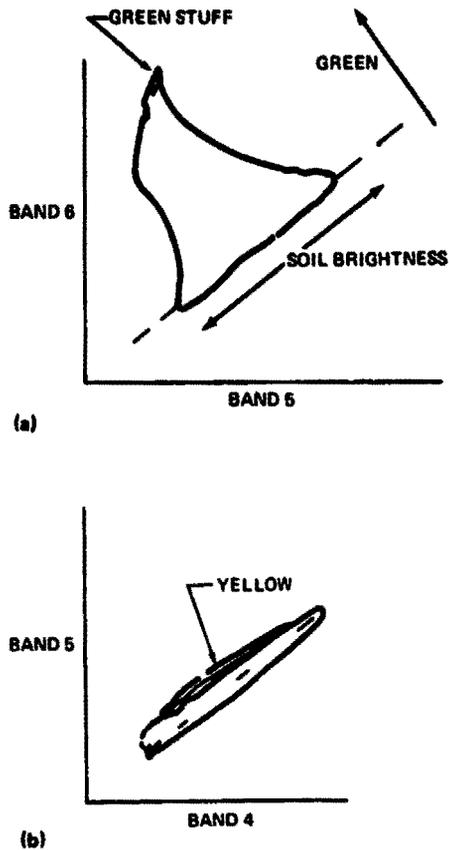


FIGURE 4.—Scatter plot comparisons of MSS band reflectance. (a) Bands 5 and 6 (fairly uncorrelated). (b) Bands 4 and 5 (highly correlated).

single axis was fitted to these (four dimensional) points. This axis was termed "brightness." Then, a point near the green peak was chosen and used with the Gram-Schmidt procedure (ref. 7) to produce a second axis that was normal to the first axis and that, together with the first, spanned the soil line and the point labeled "green stuff." This second axis was called "greenness."

Figure 5 expands the illustration to three dimensions. Soil sample points fall near a line and, further, fall predominantly in a planar region surrounding that line. Plants start out on bare soil and grow toward the region of green stuff. Among the variety of green plant canopies, some have large amounts of shadow, which shifts the observation directly toward the origin. Trees are an example of a green canopy with a fair amount of shadow, as shown by the "badge of trees." The variety of shadowing in various canopies creates a region called the green arm (some-

times called the green fold), which appears to be a line over which the Tasseled Cap is folded.

When a particular plant canopy has reached its appropriate stage of maximum green development, it may then turn yellow. Yellowing is often accompanied by darkening, either in the vegetation itself or as a result of drooping leaves which cause shadowing to increase. By yellowing, withering, or being cut down, every canopy eventually returns to the soil. These varieties of the return path in signal space are the "tassels" of the cap.

In Landsat data, the nearness of the yellow region to the side of the Tasseled Cap makes it almost indistinguishable. Hence, for most purposes, the agricultural information is substantially contained in a plane defined by unit vectors in the brightness and green-stuff directions but with a small amount of variation in the yellow-stuff direction perpendicular to the plane. A fourth direction, "non-such," is orthogonal to the other three and contains primarily noise variation. The unit vectors describing these four directions together form an orthogonal (rotation) matrix called the Tasseled Cap (or Kauth-Thomas) transform. Numerical values are given in appendix A.

Some further comments regarding the yellow-stuff direction in Landsat signal space are in order. Three physical effects cause signal variations with significant components in this one direction. The first is the effect for which the direction is named;

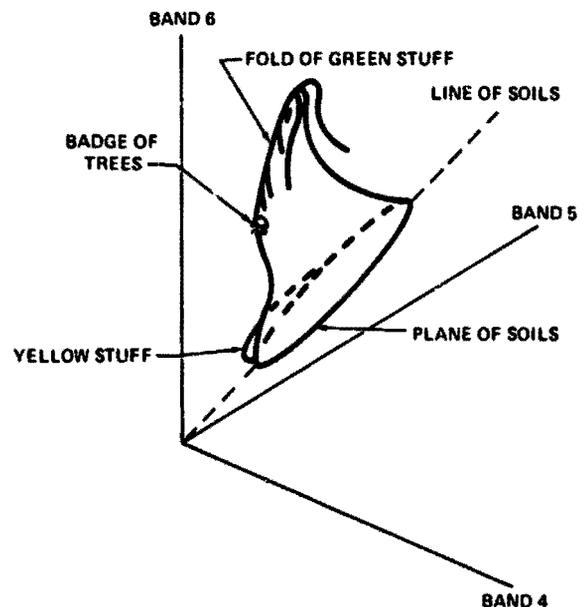


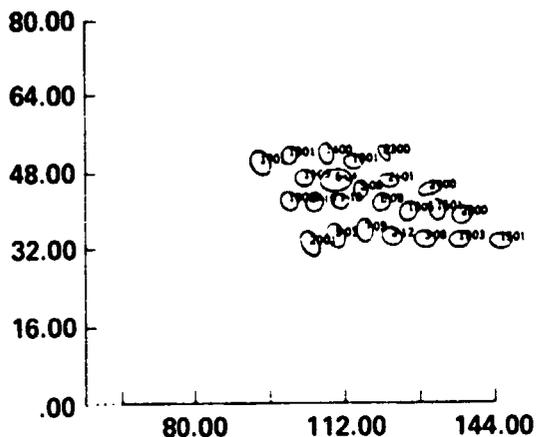
FIGURE 5.—The Tasseled Cap.

i.e., the yellowing of vegetative canopies. In this process, the chlorophyll absorption disappears and the reflectance in band 5 increases significantly. The reflectance of bands 4 and 6 also increases somewhat. Mixed in with the increased reflectance of the actual plant material is additional shadowing because of drooping leaves in some canopies or, alternatively, further decreases of shadowing due to laying out of the leaves or heads in others (see reference 6 for an example). However, these shadow effects cause changes in brightness and/or greenness, which are already established directions. The component that is not already represented by brightness and greenness is defined as the yellow direction. This direction is dominated by the difference between MSS bands 5 and 4.

The second effect is the difference between average and red soils. As seen by Landsat, the dominant direction of variation in soil reflectance is a scaling factor applied to all bands.¹ The next most important component is the difference caused by red soils. The existence of a second important soil reflectance component suggests the concept of the "plane of soils" mentioned previously in the discussion of figure 5. (It is fun to refer to crops planted upon and growing out of the plane of soils!) The direction of the second component of soil reflectance in Landsat data is nearly parallel to the yellow direction. Land-

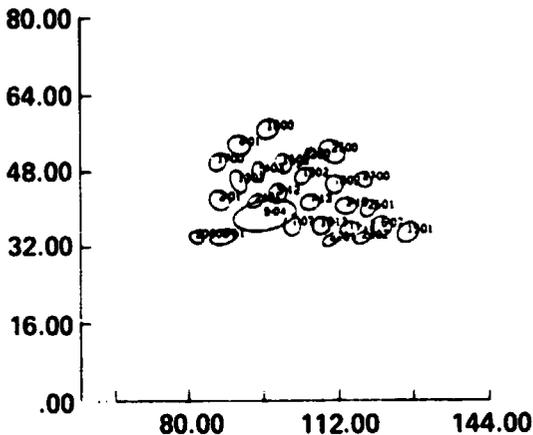
¹R. J. Kauth, "Soil Reflectance," NASA Johnson Space Center Memo TF3-75-5-190, Apr. 10, 1975.

GREENNESS



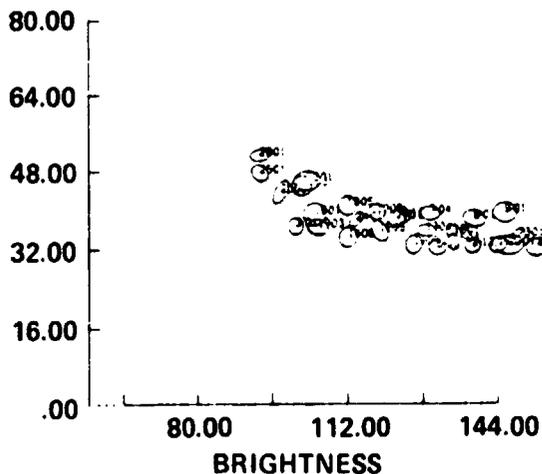
(a)

GREENNESS



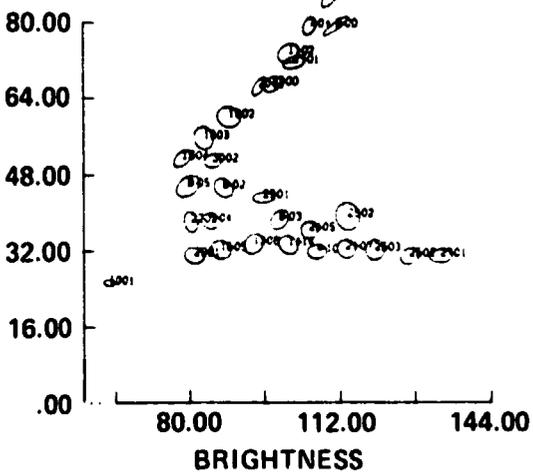
(b)

GREENNESS



(c)

GREENNESS



(d)

FIGURE 6.—Cluster scatter plots of several Kansas sample segments in transformed coordinates. (a) Grant County, May 27, 1974. (b) Haskell County, May 27, 1974. (c) Morton County, May 27, 1974. (d) Finney County, May 26, 1974.

sat sensors cannot distinguish between the visual colors red and yellow.

The third effect is due to haze over a scene, which causes changes in brightness and greenness and in the negative yellow direction. This effect will be discussed in more detail later.

The last two effects combine and make it difficult to observe canopy yellowing. Hence, attention is usually concentrated on developing crop signatures based on the greenness and brightness directions.

Figure 6 shows some typical cluster plots of Landsat data from 9- by 11-kilometer (5- by 6-nautical-mile) sample segments. In these figures, the abscissa is brightness and the ordinate is greenness or "green stuff." Notice that, in any particular scene, portions of the Tasseled Cap structure may be missing, depending on the crops planted and their state of development and on the types of soil. (The numbers on the ellipses are for identification only.)

Additional Structural Characteristics of the Data

Signals from vegetation and soil form the Tasseled Cap. Water, clouds, and cloud shadow also have their proper places in the four dimensions of Landsat signal space and can be described relative to the position of the Tasseled Cap and the points already defined. Haze and varying angles of solar illumination affect the data structure in systematic ways.

Figure 7 shows that clouds are located along the brightness axis but are shifted in the negative yellow direction as well. Figure 7(a) is an artistic conception of a scatter plot showing the data projected onto the brightness versus greenness plane, whereas figure 7(b) shows the data projected onto the yellow versus non-such plane.² Haze can be thought of as intermediate or thin clouds. When haze is present over a scene, the contrast in the scene itself is reduced while a portion of a cloudlike signal is added. The intermediate condition between no haze and completely hazy (i.e., cloudy) is sketched in figure 8 as a shaded region. It can be seen that, as the haze amount is in-

²Earlier, it was mentioned that non-such is the direction orthogonal to the other axes which "fills out" the four-dimensional space of the Landsat data. Variations in the non-such direction are mostly noise although some information regarding water and snow appears to be contained in this direction also.

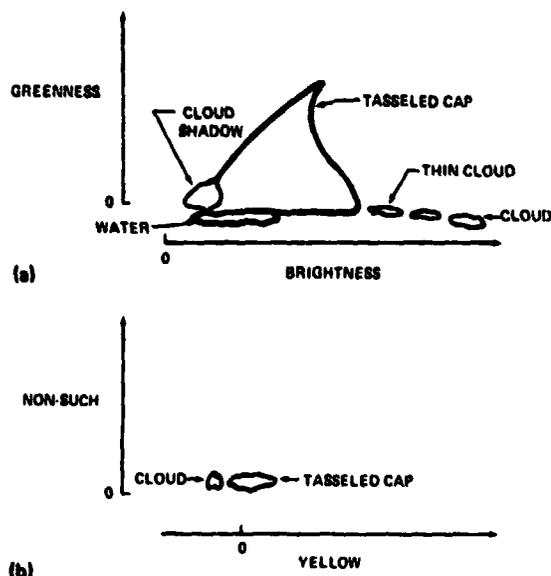


FIGURE 7.—Schematic diagram of the Tasseled Cap, showing the positions of water, clouds, and cloud shadows. (a) Projection on brightness/greenness plane. (b) Projection on yellow/non-such plane.

creased, the triangular shape of the Tasseled Cap shrinks toward the point of the clouds.

This diagram demonstrates why haze has a severe effect on signature extension capability. The entire data structure shifts out along the brightness axis and shrinks in the greenness direction; but worse, the

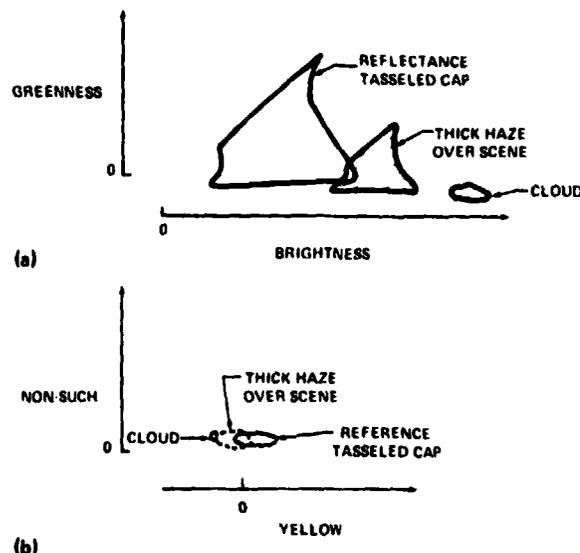


FIGURE 8.—Schematic diagram of Tasseled Cap as seen through haze. (a) Projection on brightness/greenness plane. (b) Projection on yellow/non-such plane.

shift in the negative yellow direction is sufficient to move the Tasseled Cap sidewise, completely off its haze-free image.

The angle of incidence of the Sun's illumination on a scene viewed by Landsat sensors also affects the data structure significantly. To a first order of correction, the radiance incident at the Landsat detector system changes directly with the cosine of the solar zenith angle at the Earth point viewed. (This theorem would be exactly true if the atmospheric haze and the ground both behaved as Lambertian reflectors.) To make the data gathered under different Sun zenith angles commensurate, the ERIM procedure is to correct all data to the standard zenith angle of 39° , which is typical of Landsat data over Kansas in April.

Up to this point, the four-dimensional structure of the Landsat data has been discussed qualitatively and geometrically, using the Tasseled Cap as a tool for visualizing the data. The major effects of haze and angle of solar illumination on the data structure have been indicated. Numerical values can be and are associated with this qualitative description (appendixes A and B). The immediate question is, how can this knowledge be exploited to gain useful information from the Landsat data?

PREPROCESSING/FEATURE EXTRACTION

The preprocessing and correction steps that take place before signatures can be extracted should be an integral part of feature extraction. The objectives of preprocessing and feature extraction may be several (ref. 8). Their purpose may be

1. To make the data more comprehensible by adjustment to standard conditions of observation
2. To eliminate or flag bad or noisy observations in the data
3. To make the data more comprehensible by extracting physical features or by projecting in such a way as to display the physical structure of the data
4. To compress the data, retaining most of the information and averaging over noise and redundancy
5. To make the distributions of the derived features fit some convenient model, such as the multivariate normal distribution

Several factors should be mentioned. First, all of the preceding objectives may not be met by the same preprocessing transformations; in fact, in some cases, they may be mutually exclusive. Sometimes, it

may be desirable to perform transformations on parallel paths effecting different objectives; for example, a linear transformation might be desirable for producing projections which can be examined by a researcher to gain insight, whereas a nonlinear projection might be used to cause the preprocessed data to fit a normal distribution.

Further, the term "preprocessing" may be somewhat misleading in that it appears to define a computer architecture in which first the preprocessing steps are performed, then the classification steps are performed, and then a proportion estimate is made. In fact, however, all the different conceptual steps constitute merely one functional relationship between the data and the desired output and could in practice be performed in one step. The preprocessing/feature extraction steps discussed here are conceptual and might be implemented in a variety of computer architectures.

Knowledge of the Landsat data structure and its physical effects permits the institution of two very useful preprocessing steps; namely, screening and correcting the data for the effects of solar illumination angle and haze.

Screening

Since the structure of the agricultural data is known, certain classes of materials can be identified immediately without special training. Thus, decision surfaces have been established to identify water, clouds, cloud shadow, dense haze, cloud shadow over water, and haze over water. In addition, picture elements (pixels) which are far outside the domain of any known material class are labeled "bad" pixels, and pixels without undesirable characteristics are labeled "good."

To perform the screening, a sequence of decisions is made for each pixel, such as

Is pixel bad? — Y → label bad

Is pixel cloud? — Y → label cloud

Each decision in turn depends on a sequence of tests. For example, on which side of certain planes in the Tasseled Cap transformed space does the pixel fall after being corrected for the major effect of solar zenith angle? If the pixel falls either above 16 counts or below -8 counts in the non-such direction, it is a bad pixel. The complete sequence of tests is given in appendix A.

Correcting

As shown in figure 8, the effect of haze is to shrink the size of the Tasseled Cap and to shift it in the negative yellow direction completely off its haze-free position in data space. This distortion can severely limit the capability to extend crop signatures from one site to another.

If the data points are projected first onto the two-dimensional brightness-greenness plane, the problem becomes less severe. Even then, however, there will be significant differences between hazy and haze-free signatures. The proper approach is to measure the amount of haze present and adjust the data back to some reference condition.

Fortunately, the measurement of the amount of haze is possible by measuring the average shift of the data in the negative yellow direction. The method is to measure the average yellow-stuff value of good pixels corrected for Sun zenith angle and compare this value to a reference value. The estimate is made simultaneously with the screening process. To actually correct the data, an atmospheric model which describes the effect of haze is needed. The entire procedure of making the haze measurement and applying the atmospheric model was developed by Lambeck and is called the XSTAR algorithm³ (ref. 9). The procedures for applying the XSTAR algorithm are given in appendix A.

Solar Zenith Angle Correction

Solar zenith angle effects are corrected implicitly during the screening process in order to correctly screen out bad pixels and to calculate the haze diagnostic for the XSTAR algorithm. The correction is applied to data during the process of haze correction, so that the XSTAR algorithm can in effect operate on a more standardized set of data. The form of the corrected data is

$$X' = \frac{\cos \theta_0}{\cos \theta} X$$

where θ is the Sun zenith angle, $\theta_0 = 39^\circ$, X is the data vector, and X' is the data vector corrected for Sun angle.

Satellite Calibration Effects

The authors have found it useful to incorporate data from both Landsat-1 and Landsat-2 passes, although Landsat-2 was the primary available satellite. The two satellite sensors have slightly different calibrations, which would hardly be noticeable except when machine processing is attempted. Hence, it was necessary to find a transformation which would convert the Landsat-1 data to be compatible with the Landsat-2 data. The procedure by which this was accomplished is detailed in appendix B. Briefly, it consists of comparing pairs of Landsat-1 and Landsat-2 observations on the same sample segments on successive (9-day separated) passes.

To actually perform the fit, it was necessary to account for the differing haze levels and differing Sun zenith angles for the two observations and, in effect, to assume a nonlinear model for each of the satellites (although a useful linear relationship between the two was found).

Episodic Events—Drought Estimation

Knowledge of the structure of the data allows one to estimate approximately the degree of green vegetative development over an area. This estimate has been used to monitor the status of drought conditions in the U.S. Great Plains (see the paper by Thompson and Wehmanen entitled "Application of Landsat Digital Data for Monitoring Drought").

Analyst Aids

The greenness and brightness values of pixels have been used to create analyst aids, assisting the analysts to view relevant aspects of the temporal growth pattern of a crop (see the paper by K. Abotteen and Bizzell entitled "The Classification and Mensuration Subsystem" and the paper by Heydorn entitled "Classification and Mensuration of LACIE Segments").

³P. F. Lambeck, "Implementation of the XSTAR Haze Correction Algorithm and Associated Preprocessing Steps for Landsat Data," ERIM Memo IS-PFL-1272, Mar. 18, 1977.

FEATURE EXTRACTION PHILOSOPHIES EXAMINED

Since feature extraction is so largely a Gestalt process, the features extracted naturally vary greatly from individual to individual and from one group of researchers to another. Historically, there have been many different attempts at feature extraction from Landsat data. Each of these can be analyzed in terms of the Tasseled Cap model structure discussed previously, even including such effects as the solar zenith angle and haze. The Tasseled Cap description is particularly appropriate for this analysis because the Kauth-Thomas transform is merely a rotation of the data—a way to peer in at the data structure from some unusual directions. The euclidean shape of the data structure (i.e., the distances between various observable features in the four-dimensional Landsat data space) is in no way disturbed. On the other hand, once the structure is understood, it is recognizable, even under some nonlinear transformations of the data. In the following discussion, two feature extraction schemes, the band-7-to-band-5 ratio and the Delta Classifier, are briefly analyzed.

Band Ratio Features

Throughout the history of remote sensing, various ratio schemes have been proposed and used with some success. From the physical understanding of the Landsat data, what can be said about these schemes? Conceptually, what could they accomplish?

For illustration, consider a simplified two-band case, say bands 5 and 7. Figure 9(a) shows the Tasseled Cap projected onto these bands. The dotted lines are the equiratio contours of R_{75} , the ratio of band 7 to band 5. Thus, if only R_{75} is retained, the spectral description of the scene is reduced from two dimensions to one dimension.

This ratio is extremely insensitive to changes in direct illumination on a scene, as illustrated in figure 9(b). An increased solar zenith angle results in a nearly proportional decrease in both bands 7 and 5. The ratio R_{75} remains constant for each observation, and the dotted lines pass through the same set of points relative to the overall structure of the Tasseled Cap.

How well does this ratio describe the crop development? According to Kanemasu (ref. 10), the

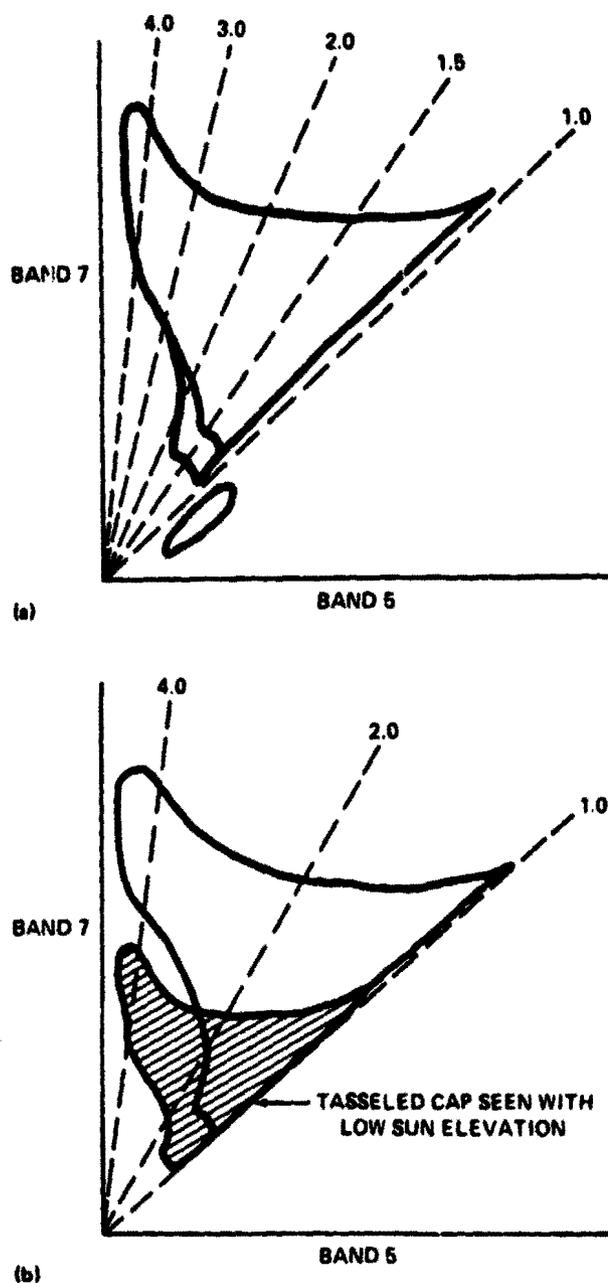


FIGURE 9.—The Tasseled Cap projected onto MSS bands 5 and 7. (a) Ratio of band 7 to band 5 with high Sun elevation. (b) Ratio of band 7 to band 5 with low Sun elevation (less illumination).

ratio of these bands is an increasing function of development through a major portion of the development cycle and is quite independent of solar illumination angle. (Data are not presented for the

period of time after full green development is reached.) Thus, it would appear that not much information is lost by retaining only the ratio of these two bands.

Are there other external effects to which the ratio R_{75} would be sensitive? The answer is yes. Figure 8 shows the effect of haze on the Tasseled Cap. As the amount of haze over the scene increases, contrast is reduced nearly proportionally in all bands so that again, as in the case of the varying solar zenith angle, the Tasseled Cap becomes smaller. But, at the same time, the haze scatters sunlight directly back to the observer, adding additional signals in all bands. The result is that the Tasseled Cap shrinks toward a point (called XSTAR) not located at the origin but located in the general neighborhood of cloud signals. Figure 8 shows this effect. The dotted lines in this case do not remain invariant in their position on the Tasseled Cap; and, in fact, the ratio R_{75} is greatly influenced by the amount of haze.

Evidently, the esthetically pleasing procedure is first to correct the data for the effects of haze and then to extract a ratio if desirable for describing crop development. Notice that the XSTAR algorithm (or any haze-correction algorithm which relies in part on measuring the yellow shift of the data structure) already contains a correction for the Sun zenith angle, which accounts by far for the largest part of the variation in illumination on the crops.

Why might it be desirable to take a ratio after having already applied the Sun-angle and haze correction? Because one might believe that most of the crop development information contained in the brightness-greenness plane is in fact implicit in a ratio, so that the remaining degrees of freedom that are discarded contain very little information. This is a conjecture which is not yet adequately tested. According to Malila (ref. 69, both spring wheat and barley are well described by the (Kauth-Thomas transform) greenness up to the time of heading; but, after heading, brightness changes become important to the description and are in fact the major source of possible discrimination between spring wheat and barley.

The Delta Classifier

The Delta Classifier extracts a development feature which combines aspects of both greenness and brightness and uses it in a decision logic based on

several acquisitions at different times during the growing season. The feature extraction occurs in three steps. In the first step:

$$f_1 = \frac{B1 - B2 + 32}{\Delta}$$

$$f_2 = \frac{B2 - B3 + 32}{\Delta}$$

$$f_3 = \frac{B3 - B4 + 32}{\Delta}$$

where $\Delta = B1 - B4 + 96$.

Notice that a fourth feature similarly defined would have the term $(B1 - B4)$ occurring in both numerator and denominator. Thus, the quantities f_1 , f_2 , f_3 , and Δ do not contain all the information in the original four-dimensional Landsat signal space. Notice that these features are somewhat independent of a constant scale factor correction, such as the cosine of the solar zenith angle, in all the Landsat bands, and they are exactly independent of an additive correction applied to all bands.

In the Delta Classifier, the quantity Δ is ignored and the features f_1 , f_2 , f_3 enter a second step of feature extraction. In the second step of feature extraction, the features are plotted on triangular graph paper. In this process, an additional degree of freedom is lost. The two new features are

$$XDEL = (f_1 - f_2) \frac{2}{\sqrt{3}}$$

$$YDEL = f_3$$

Figure 10(a) shows the relationship between the two feature sets extracted in steps 1 and 2. Figure 10(b) shows the major elements of the Tasseled Cap projected onto the XDEL, YDEL plane of the Delta Classifier. The elements shown are the origin, the mean of soils, the line of soils, the mean of the green arm, and the point XSTAR. Also shown are the points $32R_1$, $32R_2$, $32R_3$, and $32R_4$, where R_1, \dots, R_4 are the unit vectors for brightness, greenness, yellowness, and non-such, respectively.

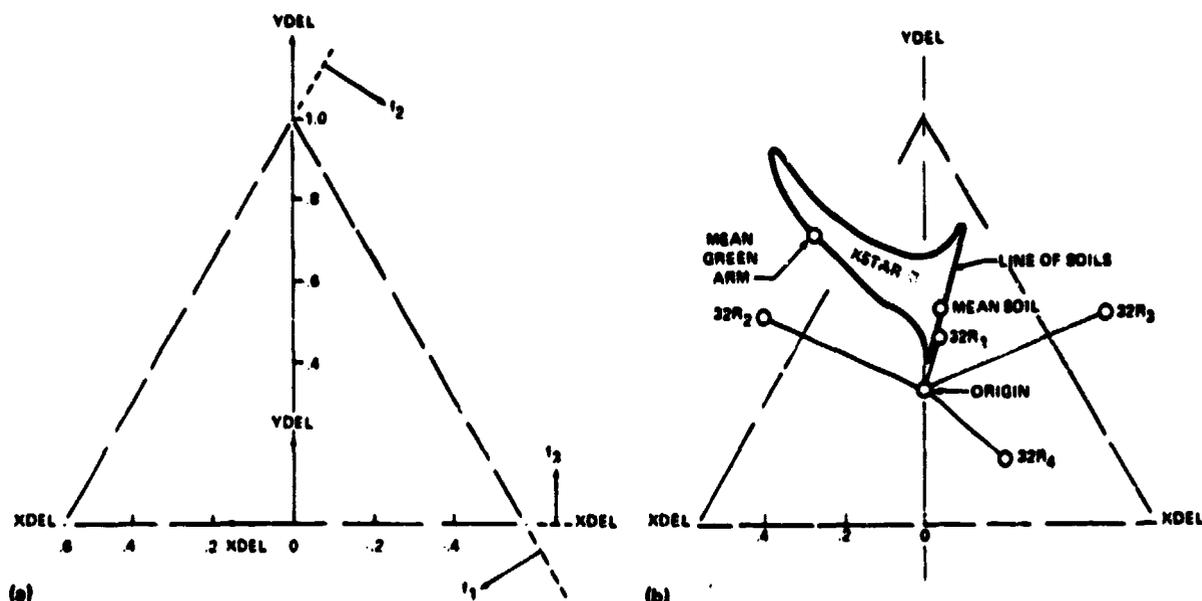


FIGURE 10.—The Delta Classifier. (a) Relationship between the two feature sets extracted in steps 1 and 2. (b) Major elements of the Tasseled Cap projected onto the XDEL, YDEL plane.

Figure 11 shows the effect of the two major external influences, solar illumination angle and haze, on the Tasseled Cap, as projected onto the Delta Classifier two-dimensional plot. The effect of a cosine ratio, $\cos(39^\circ)/\cos(\theta) = 0.66$, is shown by the dashed outline. The Tasseled Cap shrinks toward the origin. The effect of a haze amount (compared to the standard haze condition) having an optical depth of 0.42 to 0.55 micrometer is shown as a dash-dot outline. This is about as large an amount of haze as one could hope to correct successfully using the XSTAR algorithm. It has the effect of reducing contrast by approximately 33 percent, with the Tasseled Cap shrinking toward the point XSTAR, as discussed earlier.

Finally, in the third step of feature extraction, for the purpose of classifying winter wheat multitemporally during four successive phases of the growing season, only the XDEL feature is used. In general, this classification procedure is based on the assumption that, for identification as winter wheat, the crop will be emerged during the first biowindow (i.e., $XDEL(1) > 0$); will have significant green development during the second or third biowindow (i.e., $XDEL(2)$ or $XDEL(3) > XDEL(1)$); and will enter a stage of brightening and loss of greenness during the fourth biowindow (i.e., $XDEL(4) < 0$). (An increase in XDEL corresponds to an increase in greenness or a decrease in brightness.)

Figure 12 shows a typical trajectory of wheat during the four biowindows that passes the preceding tests. The following statements are of key importance in understanding the operation of the Delta Classifier with respect to the growth cycle of winter wheat.

1. During the long first biowindow, which encompasses both fall and spring emergence, the solar angle changes significantly. The requirement $XDEL > 0$ is essentially a ratio test similar to the band-7-to-band-5 ratio, as can be seen by comparing figure 12 with figure 9(a). Thus, the criterion line is substantially independent of the angle of solar illumination, as are the points near it, such as point 1 in figure 12.

2. The effect of haze on point 1 is to move it toward the point XSTAR, which means very little change in XDEL. This is true for any point just slightly above $XDEL = 0$. For points farther above it, the effect of haze is to reduce the value of XDEL but never to a value smaller than zero.

3. During the second and third biowindows, the XDEL values of winter wheat are normally large and, even when the haze level is substantial, still easily exceed the XDEL of point 1.

4. During the fourth biowindow, the XDEL value of point 4 is increased by the effect of haze. However, the typical "brightening" of the wheat at harvest is often sufficient to overcome haze effects, and point 4 still appears at a negative XDEL value.

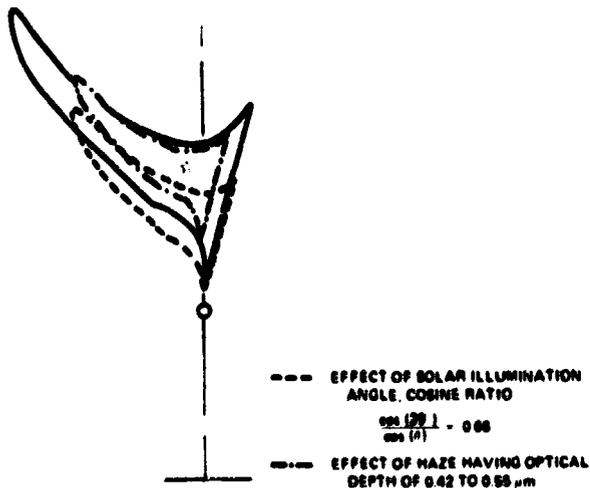


FIGURE 11.—Effect of solar illumination and haze on the Tasseled Cap as projected onto the XDEL, YDEL plane of the Delta Classifier.

In summary, a unique combination of events enables the Delta Classifier to correct for both illumination and haze effects at critical times during the wheat growth cycle, while failing to do so at other times.

SUMMARY AND RECOMMENDATIONS

The authors have summarized their knowledge of the spectral-temporal structure of agricultural scenes as viewed by Landsat and have shown how this knowledge is used to screen data, to correct for systematic external effects, and to obtain insight into the operation of various feature extraction algorithms. With respect to the problem of extracting features in the data which enhance ability to view the crop development, a number of investigation areas need to be pursued.

1. Optimal feature extraction with explicit attention to system noise should be considered. The Kauth-Thomas transform merely rotates the data structure so it can be viewed in different ways. The use of a single linear (greenness) feature in classification seems to work reasonably well in some cases, but there is no support for a contention that this is an optimal feature. Note that the noise in Landsat bands 4 to 6 is approximately proportional to the square root of the signal but is a constant in band 7. Transformations which make system noise an invariant function of feature value might be considered.

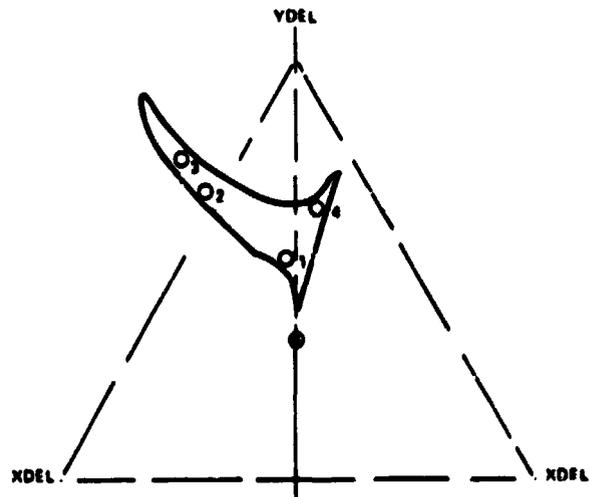


FIGURE 12.—Typical trajectory of wheat during the four blowdowns, as projected onto the XDEL, YDEL plane of the Delta Classifier.

2. Temporal-trajectory features need to be tried in which each feature is a measure of likeness between the pixel and one or the other crop development prototypes; e.g., likenesses to winter wheat, grassland, alfalfa, and corn. Essentially, these likeness features would constitute models of each major crop or confusion category (see, for example, reference 11).

3. These features (models) should be derived as conditional functions of some of the important ancillary conditions which could be observed, such as predicted crop yield for the crop in question or estimated planting date. The feature definitions could be based on a combination of crop modeling, field measurements data, and Landsat data. Landsat data must be the final basis on which the features are established. However, Landsat data are acquired irregularly; to create a general model of temporal development, the missing observations, in effect, must be estimated. To accomplish this estimate, some continuity condition which can be derived from field measurement data is required.

Appendix A

Coefficients of the Transformations

A number of transformations and data correction steps have been mentioned; each one implies certain coefficients or parameters. In this appendix, the procedures with numerical values for the automatic detection of garbled data, clouds, water, and cloud shadows; for the standardization of Landsat calibration and Sun-angle correction; and for implementing the XSTAR haze-correction algorithm are summarized.

The procedures documented herein were implemented at the Environmental Research Institute of Michigan (ERIM) in August 1977. The significant changes from previous procedures documented in ERIM Memorandum IS-PFL-1272 (Mar. 18, 1977) are as follows.

1. Revision of the calibration adjustment used for Landsat-2 full-frame data from computer-compatible tapes (CCT's) produced on or after July 16, 1975 (ref. 12) to fit the observed correspondence with Landsat-2 LACIE segment data (ERIM Memo IS-DR-1867, July 28, 1977)

2. Incorporation of the improved SCREEN procedure (refs. 9 and 13) for detecting garbled data, clouds, snow, cloud shadows, and water in Landsat multispectral scanner data

3. Reversal of the sign of the XSTAR haze parameter γ to correspond with reference 9

The steps for implementing the procedures are as follows.

Step 1.—Recalibrate the Landsat data as required.

Three distinctly different calibrations (counts versus radiance values) for Landsat data currently exist. These calibrations pertain to the following Landsat data products.

1. Landsat-1 data (LACIE segment and full-frame imagery)

2. Landsat-2 full-frame data from CCT's produced on or after July 16, 1975

3. Landsat-2 full-frame data from CCT's produced before July 16, 1975, and Landsat-2 LACIE segment data

The XSTAR haze-correction algorithm is specifically adjusted for Landsat-2 LACIE segment data. Other Landsat data can be recalibrated to simulate this same calibration by employing the appropriate multiplicative and additive transformation selected from the following.

Let x_i represent the Landsat signal in channel i . If recalibration is necessary, for each channel, set

$$x'_i = A_i x_i + B_i$$

Then, set

$$x_i = x'_i$$

Step 1a.—For Landsat-1 data (LACIE segment or full frame (app. B)), use

$$A = \begin{pmatrix} 1.04 \\ 1.00 \\ 1.09 \\ .82 \end{pmatrix} \quad B = \begin{pmatrix} -5.79 \\ 1.19 \\ -2.91 \\ 3.01 \end{pmatrix}$$

Step 1b.—For Landsat-2 full-frame data from CCT's produced on or after July 16, 1975 (ref. 12 and ERIM Memo IS-DR-1867), use

$$A = \begin{pmatrix} 1.275 \\ 1.141 \\ 1.098 \\ .948 \end{pmatrix} \quad B = \begin{pmatrix} -1.445 \\ -2.712 \\ -2.950 \\ .446 \end{pmatrix}$$

Step 1c.—For Landsat-2 LACIE segment data or for Landsat-2 full-frame data from CCT's produced before July 16, 1975, no recalibration is necessary.

Step 2.—Perform Sun-angle correction.

Let x_i represent the Landsat signal in channel i , following step 1. Let θ represent the solar zenith angle for the data acquisition. Then, for each channel of the acquisition, set

$$x'_i = \frac{\cos \theta_0}{\cos \theta} x_i \quad (\theta_0 = 39^\circ)$$

Then, set

$$x_i = x'_i$$

The data will now appear to have been acquired for the solar zenith angle θ_0 . The parameters of XSTAR have been adjusted for $\theta_0 = 39^\circ$.

Step 3.—Identify garbled data, clouds, water, and cloud shadows.

Let R be a rotation matrix (i.e., $RT = R^{-1}$, $\det(R) = 1$) defined by

$$R = \begin{pmatrix} 0.33231 & -0.28317 & -0.89952 & -0.01594 \\ .60316 & -.66006 & .42830 & .13068 \\ .67581 & .57735 & .07592 & -.45187 \\ .26278 & .38833 & -.04080 & .88232 \end{pmatrix}$$

The columns of R are unit vectors characterizing the axes of a rotated Landsat-2 data space akin to the Tasseled Cap data space but particularly oriented to suit the XSTAR algorithm (ref. 9).

Let x_i represent the Landsat signal in channel i , following step 2. Let i and j correspond to rows and columns of the R matrix. For each pixel of an acquisition, and for each value of j ($1 \leq j \leq 4$), calculate

$$z_j = \sum_{i=1}^4 R_{ij} x_i \quad (z = R^T x)$$

Step 3a.—A pixel is labeled as *garbled* data if any one of the following conditions applies.

$$z_4 > z_{4max} \quad (z_{4max} = 16)$$

$$z_4 < z_{4min} \quad (z_{4min} = -12)$$

$$z_3 - 0.09375z_1 > z_{3max} \quad (z_{3max} = -4)$$

$$z_3 + 0.18750z_1 < z_{3min} \quad (z_{3min} = 14)$$

$$z_2 + z_1/10 < z_{2min} \quad (z_{2min} = -20)$$

$$z_2 + z_1/1.8 > z_{5max} \quad (z_{5max} = 156)$$

$$z_2 - z_1/1.2 > z_{6max} \quad (z_{6max} = -8)$$

Step 3b.—A pixel is labeled as *cloud* if not labeled garbled and if both the following conditions apply.

$$z_1 > z_{Cmax} \quad (z_{Cmax} = 100)$$

$$z_3 + z_1/10 < z_{Cmin} \quad (z_{Cmin} = -7.5)$$

Step 3c.—A pixel is labeled as *diffuse cloud* (dense haze) if not labeled garbled or cloud and if both the following conditions apply.

$$z_1 > z_{Hmax} \quad (z_{Hmax} = 69)$$

$$z_3 + z_1/7 < z_{Hmin} \quad (z_{Hmin} = -3.25)$$

Sufficiently dense coverings of snow tend to be placed in the cloud or diffuse cloud category.

Step 3d.—A pixel is labeled as *water* if not labeled garbled, cloud, or diffuse cloud and if all the following conditions apply.

$$z_1 < z_{W1} \quad (z_{W1} = 75)$$

$$z_2 + z_1/16 < z_{W21} \quad (z_{W21} = -0.5)$$

$$z_4 < z_{W4} \quad (z_{W4} = 1.5)$$

$$z_2 + z_4 < z_{W24} \quad (z_{W24} = -4.5)$$

$$z_2 + z_1/2 + z_3 + 5 \times z_4 < z_{W2134} \quad (z_{W2134} = 10)$$

If a pixel is labeled water on the basis of satisfying the preceding conditions, a subcategory *cloud shadow over water* may be identified if, in addition, the following is applicable.

$$z_2 - 0.4 \times z_1 > z_{WS} \quad (z_{WS} = -12.2)$$

The subcategory cloud shadow over water is sometimes an artifact under the category cloud shadow (usually caused by striping effects in the data).

Step 3e.—A pixel is labeled as *cloud shadow* if not labeled garbled, cloud, diffuse cloud, or water and if

both the following conditions apply.

$$z_2 - 0.4 \times z_1 - 0.6 \times z_3 - 0.6 \times z_4 > z_{S2134} \quad (z_{S2134} = -0)$$

$$z_1 - 0.4 \times z_2 < z_{S12} \quad (z_{S12} = 37.75)$$

Step 4.—Compute XSTAR scene diagnostic signal value.

Let \hat{x}_i represent a scene diagnostic Landsat signal value in channel i for a single data acquisition. Set x_i equal to the average signal value of all pixels not labeled as garbled, cloud, diffuse cloud, water, or cloud shadow. (If especially subtle effects, such as nonuniform haze, are present in the scene, a bias will be introduced into the XSTAR haze diagnostic procedure that will lead to overcorrection or undercorrection of the data. Some types of nonagricultural data—as yet not studied—may also introduce a bias.)

Step 5.—Determine amount of change from reference haze condition (ref. 9).

Let \hat{x}_i be the Landsat scene diagnostic signal value in channel i . Let α_i and x_i^* be coefficients of the XSTAR algorithm for channel i . Let γ represent the change in optical thickness from the reference condition. Let Y^* be the yellow value characterizing the reference haze condition. Numerical values are as follows.

$$\alpha = \begin{pmatrix} 1.2680 \\ 1.0445 \\ .9142 \\ .7734 \end{pmatrix} \quad x^* = \begin{pmatrix} 61.9 \\ 66.2 \\ 83.2 \\ 33.9 \end{pmatrix}$$

$$Y^* = -11.2082$$

Using \hat{x}_i and Y^* , calculate the following.

$$a = \sum_{i=1}^4 \alpha_i^2 (\hat{x}_i - x_i^*) R_{i3}$$

$$b = \sum_{i=1}^4 \alpha_i (\hat{x}_i - x_i^*) R_{i3}$$

$$c = \left(\sum_{i=1}^4 \hat{x}_i R_{i3} \right) Y^*$$

$$\gamma = \frac{b}{a} \left\{ 1 - \sqrt{1 - \frac{2ac}{b^2}} \right\}$$

For extremely hazy conditions, the quantity under the radical in the equation for γ can be negative. In such cases, set the square root equal to zero; i.e., $\gamma = -(b/a)$. The solution given for γ is obtained (ref. 9) by approximating $e^{\alpha_i \gamma}$ in the XSTAR correction by a quadratic expression and then solving for γ such that the yellow value of x_i after the XSTAR correction will be Y^* .

As an alternative, in the event that the quadratic expression for $e^{\alpha_i \gamma}$ is inaccurate because of an extreme change from the reference haze condition, an iterative solution for γ is possible. The need for such a solution will be indicated by the presence of a negative quantity under the radical or by the obtaining of $|\gamma| > 0.5$. For the iterative solution, set

$$\gamma' = \gamma$$

$$\hat{x}'_i = e^{\alpha_i \gamma'} (\hat{x}_i - x_i^*) + x_i^*$$

$$\hat{x}_i = \hat{x}'_i$$

Then, repeat the solution for a , b , c , and γ using the new value of \hat{x}_i . Next, increase the new value of γ by γ' . If the quantity under the radical is again negative, or if $|\gamma - \gamma'| > 0.5$, the procedure might be repeated once more. If the iterations do not converge, discard the data acquisition as unusable or seek other remedies. Current experience indicates that the need for iterating should be rare.

Step 6.—Apply XSTAR correction.

Steps 1 to 5 can be accomplished with a single pass through the data. A second pass is required for step 6. Given a successful solution for γ in step 5, XSTAR may then be applied to correct each pixel of the acquisition as follows.

Let x_i represent the Landsat signal in channel i , following step 2. Then, for each pixel, calculate the following.

$$x'_i = e^{\alpha_i \gamma} (x_i - x_i^*) + x_i^*$$

$$x_i = x'_i$$

This correction may be applied to all pixels within the acquisition. (However, garbled data, clouds, or snow may convert to signal levels outside the normal dynamic range.)

To minimize roundoff or truncation errors, the data analyst should retain intermediate results from steps 1 to 6 in floating-point rather than integer format.

Appendix B

A Transformation to Make Landsat-1 and Landsat-2 MSS Data Compatible

INTRODUCTION AND DATA BASE

It was desired to estimate coefficients of a transformation which would convert Landsat-1 data to Landsat-2 data. In order to make the estimate, it was desirable to use the identical scene observed under identical conditions by both satellites. The nearest procedure in practice is to observe pairs of data over the same scenes separated by 9 days. An initial selection of about 25 such pairs was made, but natural attrition reduced this ultimately to 8 pairs. As will be shown, this is a quite minimal set and the fitting procedure should be repeated with a larger sample.

Pairs were eliminated from consideration if either member contained "patchiness" of cloud or haze, as evidenced in the histogram output of program SCREEN, or if the haze levels of the pair were thought to be markedly different from each other as evidenced by the yellow level of the mean of soils calculated by program SCREEN. The relative yellow level for each pass was judged against all non-"patchy" passes for the same satellite by plotting a separate yellow-level histogram over those passes for each satellite. The finally accepted pass pairs are listed in table B-I. Of the eight cases, four are cases in which the Landsat-2 pass precedes the Landsat-1 pass.

TABLE B-I.—Pass Pairs Used in Landsat-1 to Landsat-2 Fitting Procedure

Case	Segment	Landsat-1 (L1) date	Landsat-2 (L2) date	Sun zenith angle, θ_1 , deg	Sun zenith angle, θ_2 , deg
1	1020	155	164	37	32
2	1020	101	92	47	46
3	1154	153	162	37	31
4	1855	82	73	52	52
5	1855	82	91	52	46
6	1857	101	92	46	44
7	1861	101	92	46	45
8	1882	153	162	37	31

^aSame pass used twice.

The data used for fitting consisted of the four-band "mean of soils" and the four-band "mean of the green arm," both outputs of program SCREEN. These data and their averages are shown separately for each Landsat band in table B-II.

TABLE B-II.—Diagnostic Data Used in Fitting

Band	Case	Soil mean (YT)		Green arm mean (WT)		Average ^a	
		L1	L2	L1	L2	L1	L2
4	1	41.5	41.7	25.1	24.4	33.3	33.05
	2	40.2	36.8	28.0	22.0	34.1	29.4
	3	36.1	36.3	26.0	23.6	31.05	29.95
	4	34.2	30.3	26.9	22.1	30.55	26.2
	5	34.2	36.3	26.9	26.3	30.55	31.3
	6	39.7	37.6	28.7	25.0	34.2	31.3
	7	43.6	41.2	35.0	32.0	39.3	36.6
	8	34.9	35.1	25.6	24.6	30.25	29.85
5	1	40.3	47.1	16.9	25.4	30.1	36.25
	2	39.6	41.7	24.3	23.3	31.95	32.50
	3	31.3	38.0	16.6	20.3	23.95	29.15
	4	33.5	34.3	25.4	25.8	29.45	30.05
	5	33.5	41.9	25.4	29.7	29.45	35.80
	6	40.7	44.8	25.2	28.8	32.95	36.80
	7	46.3	49.7	36.8	40.7	41.55	45.20
	8	29.2	36.4	16.2	20.6	22.70	28.50
6	1	40.5	49.5	41.5	50.5	33.30	50.0
	2	38.8	41.7	40.7	40.5	39.75	41.1
	3	32.8	40.5	46.3	54.6	39.55	47.55
	4	32.3	34.3	32.5	33.9	32.4	34.1
	5	32.3	42.5	32.5	45.8	32.4	44.15
	6	38.8	45.1	44.3	46.1	41.55	45.6
	7	44.5	50.0	40.6	44.3	42.55	47.15
	8	34.3	41.0	43.5	59.6	38.9	50.30
7	1	17.1	20.5	22.0	25.3	19.55	22.90
	2	17.0	17.5	21.6	20.1	19.30	18.80
	3	14.0	16.1	25.8	26.7	19.9	21.40
	4	14.5	15.2	16.5	16.6	15.5	15.9
	5	14.5	18.0	16.5	22.3	15.5	20.15
	6	16.9	18.9	24.2	22.7	20.55	20.80
	7	19.8	20.9	19.1	19.3	19.45	20.1
	8	14.8	16.6	23.1	28.6	18.95	22.6

^aAverage = (YT + WT)/2.

MODELS

Several different models were used for fitting. In general, the models are of the form

$$x_2 = F(\alpha, \theta_1, \theta_2, x_1) + \Sigma$$

where x_2 is the Landsat-2 signal; x_1 is the Landsat-1 signal; θ_1 and θ_2 are the Sun zenith angles of Landsat-1 and Landsat-2 cases, respectively; α is a set of parameters which must be estimated; and Σ is an error.

Three specific models were tried, as follows.

a. Ratio model—assumes for each band

$$x_2 = A \frac{\cos \theta_2}{\cos \theta_1} x_1 + \Sigma$$

i.e., that there is no difference in signal offset between the two satellites and that the radiance returned from the scene is an inverse function of the cosine of the Sun zenith angle.

b. Offset model—assumes

$$\left(\frac{x_2}{\cos \theta_2} \right) = A \left(\frac{x_1}{\cos \theta_1} \right) + B$$

c. Three-parameter offset model—assumes

$$x_2 = A_2 A_1^{-1} \frac{\cos \theta_2}{\cos \theta_1} x_1 + B_2 - A_2 A_1^{-1} \frac{\cos \theta_2}{\cos \theta_1} B_1$$

where A_1 is the responsivity of Landsat-1, A_2 is the responsivity of Landsat-2, B_1 is the offset for Landsat-1, and B_2 is the offset for Landsat-2. Thus,

$$x_2 = A \frac{\cos \theta_2}{\cos \theta_1} x_1 + B_2 - A \frac{\cos \theta_2}{\cos \theta_1} B_1$$

Model "a" requires a fit to one parameter per band; model "b" requires two parameters per band; and model "c" requires three parameters per band. The residual error per band after fitting is shown in table B-III. Model "c" is considerably the best fit for band 7 and is slightly better for the other bands.

TABLE B-III.—Root Mean Squared Error of Three Models

Band	Model a	Model b	Model c
4	1.11	1.05	0.81
5	2.06	1.60	1.36
6	2.99	2.24	1.66
7	4.78	3.39	1.20

In general, one can write

$$z = a_0 + a_1 x + a_2 y$$

and identify

$$a_0 = B_2$$

$$z = x_2$$

$$a_1 = A$$

$$x = \frac{\cos \theta_2}{\cos \theta_1} x_1$$

$$a_2 = -AB_1$$

$$y = \frac{\cos \theta_2}{\cos \theta_1} x_1$$

In order to minimize the noise of observations of the soil and green arm points, the authors used their averages from table B-II. Thus, x_1 is the average L1 in table B-II. Using these tabulated values and regressing z on x and y gives the results shown in table B-IV.

Interest generally will be in converting Landsat-1 data to resemble Landsat-2 data at the same solar zenith angle. Therefore, the model will simplify to the form

$$x_2 = Ax_1 + B$$

where $B = B_2 - AB_1$; B is also given in table B-IV.

TABLE B-IV.—Regression Coefficients for Three-Parameter Model c

Band	$a_0 = B_2$	$a_1 = A$	$a_2 = -AB_1$	B_1	B
4	-19.48	1.04	13.69	-13.16	-5.79
5	-24.99	1.00	26.18	-26.18	1.19
6	-74.70	1.09	71.79	-65.86	-2.91
7	-21.53	.82	24.54	-29.93	3.01

REFERENCES

1. Henderson, R. G.; Thomas, G. S.; and Nalepka, R. F.: Methods of Extending Signatures and Training Without Ground Information. ERIM 109600-16-F (Ann Arbor, Mich.), May 1975.
2. Suits, G.: The Calculation of the Directional Reflectance of a Vegetative Canopy. Remote Sensing of Environment, vol. 2, 1972, pp. 117-125.
3. Blanchard, M. B.; Greeley, R.; and Goettelman, R.: Use of Visible, Near-Infrared, and Thermal Infrared Remote Sensing To Study Soil Moisture. Proceedings of the Ninth International Symposium on Remote Sensing of Environment (Ann Arbor, Mich.), Apr. 1974.
4. Parks, W. L.; Sewell, J. I.; Hilty, J. W.; and Rennie, J. C.: Utilizing ERTS Imagery To Detect Plant Diseases and Nutrient Deficiencies, Soil Types and Soil Moisture Levels. Contract NAS5-21873, NASA Goddard Space Flight Center, Mar. 1974.
5. Condit, H. R.: The Spectral Reflectance of American Soils. Photogrammetric Engineering, vol. 36, Sept. 1970, p. 955.
6. Malila, W. A.: Spectral Separability of Spring Wheat From Other Small Grains. Progress Report for Nov. 15, 1977, to March 14, 1978, on the Analysis of Scanner Data for Crop Inventories, ERIM 132400-2-L (Ann Arbor, Mich.), Mar. 1978.
7. Hildebrand, F. B.: Methods of Applied Mathematics. Prentice-Hall (New York), 1952.
8. Malila, W. A.; Crane, R. B.; and Turner, R. E.: Information Extraction Techniques for Multispectral Scanner Data. WRL 31650-74-T, Willow Run Laboratories, University of Michigan (Ann Arbor, Mich.), June 1972.
9. Lambeck, P. F.: Signature Extension Preprocessing for Landsat MSS Data. ERIM 122700-32-F (Ann Arbor, Mich.), Nov. 1977.
10. Kanemasu, E. T.: Seasonal Canopy Reflectance Patterns of Wheat, Sorghum, and Soybeans. Remote Sensing of Environment, vol. 3, 1974, pp. 43-47.
11. Abotteen, R. A.: Principal Component Greenness Transformation in Multitemporal Agricultural Landsat Data. Paper presented at the 12th International Symposium on Remote Sensing of Environment (Manila), Apr. 20-26, 1978.
12. Radiometric Calibration of Landsat 2 MSS Data. Landsat Newsletter No. 15, NASA Goddard Space Flight Center (Greenbelt, Md.), June 1, 1977.
13. Lambeck, P. F.: Signature Extension Preprocessing Algorithm Development. Contract Quarterly Progress Report, Task 2, Analysis of Scanner Data for Crop Inventory, May 15, 1977, to September 9, 1977, ERIM 122700-28-L (Ann Arbor, Mich.), Oct. 1977.

Compensation for Atmospheric Effects in Landsat Data

P. F. Lambeck^a and J. F. Potter^b

INTRODUCTION

It is well known that variations in Sun angle and in atmospheric aerosol and water-vapor levels change the spectral signatures collected by multispectral scanners (refs. 1 and 2). It has also been shown that the background reflectance changes signatures (ref. 3). These changes have a deleterious effect on classification accuracy. Hence, even before the beginning of LACIE, the remote-sensing community pursued the development of preprocessing techniques for removing or reducing the variations in multispectral data caused by such changes. In the initial design of LACIE and throughout its operation, it was anticipated that some of these techniques would be incorporated into LACIE procedures once they had been demonstrated to work within the necessary constraints of the experiment (refs. 4 to 6). Two of these techniques (Sun-angle correction and mean-level adjustment¹) were actually tried as components of alternative LACIE systems, but the results were not satisfactory. This failure could be attributed to inherent limitations in these techniques and to the difficulty in identifying sufficiently similar training and classification areas. This latter problem (partitioning, stratification, and sampling strategy for training) is discussed in a separate symposium paper on signature extension (see the paper by Kauth and Richardson entitled "Signature Extension Methods in Crop Area Estimation"). The following discussion documents some of the progress of the supporting research community and support contractors at the

NASA Johnson Space Center in developing preprocessing algorithms to support LACIE.

Most of the research efforts in developing preprocessing techniques to support LACIE were based on the assumption that changing observation conditions cause multiplicative and additive changes in each multispectral data channel (ref. 7). Consequently, two major options were available: (1) to develop data transformations (e.g., ratioing) which would tend to cancel out these effects and (2) to develop methods for estimating the appropriate multiplicative and additive factors and then to apply these directly to the data. For the most part, the latter option was taken. For Landsat data, four multiplicative and four additive factors needed to be determined.

Initial attempts to estimate the eight correction factors for Landsat data did not rely on any prior knowledge of how the multiplicative and additive factors might be interrelated but, instead, relied solely on statistical characteristics of the distributions of data to be preprocessed. This approach led to the development of cluster-matching algorithms (refs. 7 to 10) and distribution-matching algorithms (refs. 11 and 12) which attempted to extract significant statistical measures from appropriate subsets of the data distributions. A number of these algorithms were tested by Lockheed Electronics Company (LEC), and the results are given in reference 13 and in the symposium paper by Minter entitled "Methods of Extending Crop Signatures From One Area to Another." These approaches were found to produce unstable results at times and, therefore, they were considered unsuitable for use in LACIE.

Efforts were also undertaken to develop preprocessing algorithms using mathematical models to define interrelations between the required multiplicative and additive correction factors such that just a few statistical characteristics of a Landsat data distribution would be sufficient to drive the mathematical model and to calculate the preproc-

^aEnvironmental Research Institute of Michigan, Ann Arbor, Michigan.

^bLockheed Electronics Company, Houston, Texas.

¹Mean-level adjustment is a technique for normalizing LACIE Landsat data so that the mean values for different segments are equal.

essing corrections. The two most significant of these algorithms are the Atmospheric Correction (ATCOR) computer program developed by LEC (ref. 14) and the XSTAR haze correction algorithm developed at the Environmental Research Institute of Michigan (ERIM) (ref. 15).² These algorithms are discussed in the second and third sections of this paper, respectively, and both are discussed further in the fourth section.

ATCOR: AN ALGORITHM TO CORRECT LANDSAT DATA FOR THE EFFECTS OF HAZE, SUN ANGLE, AND BACKGROUND REFLECTANCE

Introduction

The radiance measured by the Landsat multispectral scanner (MSS) in a given channel I , where $I = 1, 2, 3, 4$, is determined primarily by four quantities.

1. The reflectance ρ_I of the target (i.e., the element of the Earth's surface in the field of view) in channel I (This quantity is actually a function of the wavelength λ but is assumed to be constant over the bandwidth of channel I .)
2. The solar zenith angle θ_0
3. The haze level τ_H in the atmosphere
4. The average reflectance $\bar{\rho}_I$ of the adjacent areas of the Earth's surface outside the field of view, assumed to be constant over the bandwidth of channel I

In this paper, the haze level τ_H is defined as the haze optical depth at wavelength 0.5 micrometer. (See reference 16 for discussion of optical depth and other concepts from radiative transfer theory.) The haze optical depth at other wavelengths λ is denoted by $\tau_H(\lambda)$. Normally, in the analysis of Landsat data, one wishes to classify certain objects on the Earth's surface on the basis of their reflectance ρ_I . These objects may be in the same Landsat image or in several different images separated in space and time. Variations in θ_0 , τ_H , and $\bar{\rho}_I$ within a scene or from one scene to another change the data and therefore reduce classification accuracy.

²P. F. Lambeck, "Revised Implementation of the XSTAR Haze Correction Algorithm and Associated Preprocessing Steps for Landsat Data," ERIM Memo IS-PFL-1916, Nov. 1977.

The ATCOR algorithm is a method for simulating the effects of such variations and correcting for them. Simulation and correction are really the same process since correction consists of simulating the MSS response for values of the Sun angle, haze level, and background reflectance that are different from the actual values. To simulate the effect of changes in θ_0 , τ_H , and $\bar{\rho}_I$, one must compute the MSS response as a function of these variables and of ρ_I .

An atmospheric model was developed and the Van de Hulst adding method (see the appendix) was used to compute the radiances at the MSS for a range of values of ρ_I , θ_0 , τ_H , and $\bar{\rho}_I$. This computation was done for all wavelengths in the MSS bands in steps of 0.01 micrometer, and the resulting radiances corresponding to each band were then multiplied by the MSS response function and integrated over wavelength to obtain the instrument response for that band. It was found that the Landsat gray-scale levels L_I could be written as

$$L_I = A_I(\bar{\rho}_I, \theta_0, \tau_H) \rho_I + B_I(\bar{\rho}_I, \theta_0, \tau_H)$$

where A_I and B_I are coefficients that are computed and tabulated for a full range of values for $\bar{\rho}_I$, θ_0 , and τ_H . Using this table, it is a simple matter to determine the effect on the Landsat data (i.e., L_I) of a change in any or all of these parameters.

These results allow one to make corrections for changes in θ_0 , τ_H , or $\bar{\rho}_I$ if values for these quantities are known for the segments to be corrected. Generally, θ_0 is known but τ_H and $\bar{\rho}_I$ are not known. However, if τ_H is known, $\bar{\rho}_I$ can be calculated using the table described in the preceding paragraph. The ATCOR computer program estimates τ_H using the method described later in this section, computes $\bar{\rho}_I$, and interpolates in the table of $A_I(\bar{\rho}_I, \theta_0, \tau_H)$ and $B_I(\bar{\rho}_I, \theta_0, \tau_H)$ to find the correction coefficients to make the desired correction.

The Atmospheric Model

The atmospheric model includes scattering by the molecular atmosphere and by haze. A factor which may be very important is scattering by cirrus clouds (ref. 17). This parameter could also be included in the model. However, the method used to determine the level of "haze" in the atmosphere (see the section

on the ATCOR program) in fact estimates the total effect of all aerosols in the atmosphere and cannot distinguish between haze and cirrus clouds. Therefore, it did not seem worth while to model the effect of cirrus clouds separately. ATCOR partly corrects for the effects of cirrus clouds since they contribute to the "haze" level τ_H . However, because the model assumes this contribution is from haze particles in the lower atmosphere, the correction is less than optimal.

It is assumed that the atmosphere consists of two homogeneous layers: a Rayleigh scattering molecular layer on top and a Mie scattering haze layer next to the Earth's surface. This two-layer model is expected to be a good approximation for the atmosphere since most of the haze is in the lower 1 kilometer of the Earth's atmosphere, whereas only about 11 percent of the molecular atmosphere is in this region. The two-layer model greatly simplifies the calculations. Water-vapor and other gaseous absorption is neglected, although it can be important in channel 4 of the Landsat data (ref. 2).

To define the atmospheric model, one must define the scattering diagrams (i.e., phase functions; ref. 16) and the optical depths for the two layers. These quantities completely define the scattering properties of the layers. They are well known for the Rayleigh case (ref. 16) and will not be discussed in detail here. For the haze layer, the scattering diagrams and optical depths were calculated from the Mie theory using a haze model by Reeser (ref. 18). The model is intended to represent a continental-type haze and assumes spherical particles with a size distribution given by

$$f(r) = 90 \quad 0.01 \text{ micrometer} \leq r \leq 0.1 \text{ micrometer}$$

$$= \frac{90}{10^4 r^4} \quad 0.1 \text{ micrometer} \leq r \leq 10.0 \text{ micrometers}$$

where r is the particle radius. This distribution corresponds to 100 particles/cm³. The real part of the index of refraction varied from 1.54 to 1.56 in the wavelength interval of interest, 0.4 micrometer $\leq \lambda \leq$ 1.1 micrometers. The imaginary part of the index was taken to be zero since absorption is neglected. Scattering diagrams for this model were computed for several wavelengths; the one for $\lambda = 0.8$ micrometer is shown in figure 1. The scattering diagram changes only a small amount with wavelength; therefore, the one shown in figure 1 was

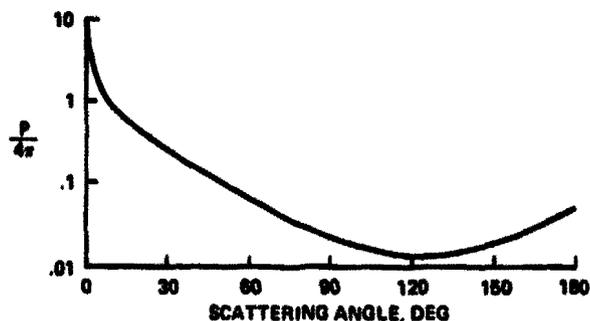


FIGURE 1.—Scattering diagram for haze at wavelength 0.8 micrometer.

used for all wavelengths. This procedure considerably reduces the computational effort involved.

The calculations described in this paper were made for haze levels of 0.0, 0.424, and 0.848. The variation of $\tau_H(\lambda)$ with wavelength for the cases $\tau_H = 0.424$ and $\tau_H = 0.848$ are shown in figure 2. The variation with λ of the Rayleigh optical depth $\tau_R(\lambda)$ is also shown in figure 2.

Radiance at the Sensor

To compute the MSS response for various values of ρ_p , θ_0 , τ_H , and ρ_p , one first calculates the corresponding radiance at the sensor. The method for doing this is rather complicated and is described in the appendix. There, it is shown (eq. (72)) that the radiance at the MSS can be written in the form

$$N(\bar{\rho}_p, \mu_0, \lambda) = \mu_0 F(\lambda) \left[a(\bar{\rho}_p, \mu_0, \lambda) \rho(\lambda) + b(\bar{\rho}_p, \mu_0, \lambda) \right] \quad (1)$$

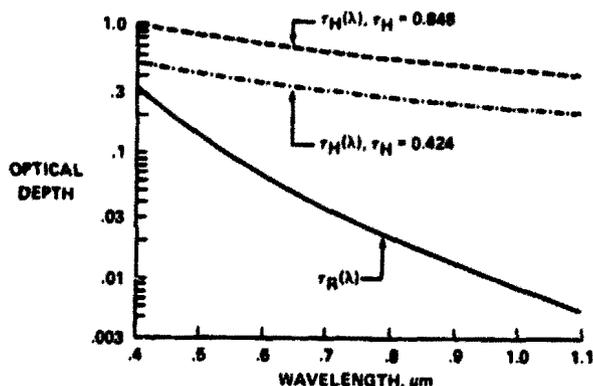


FIGURE 2.—Rayleigh and haze optical depths.

where $\mu_0 = \cos \theta_0$. Here, the dependence on the wavelength λ is explicitly indicated, but the dependence on the haze level τ_H is not. The quantity $F(\lambda)$ is the incident solar radiance at the top of the atmosphere divided by π . For each of the 3 values of τ_H , the coefficients $a(\bar{\rho}, \mu_0, \lambda)$ and $b(\bar{\rho}, \mu_0, \lambda)$ were computed for 71 values of λ (from 0.4 to 1.1 micrometers in units of 0.01 micrometer), 25 values of μ_0 (24 gauss points plus $\mu_0 = 1.0$), and 50 values of $\bar{\rho}$ (0.0 to 0.49 in units of 0.01).

The Landsat MSS Data

The Landsat MSS data (i.e., the MSS gray-scale levels L_I) are given by

$$L_I = \alpha_I N_I + \beta_I \quad (2)$$

where α_I and β_I are constants given in table I and N_I is the equivalent spectrally flat radiance defined by

$$N_I = \frac{\int N(\bar{\rho}, \rho, \mu_0, \lambda) S_I(\lambda) d\lambda}{\int S_I(\lambda) d\lambda} \quad (3)$$

Here, $S_I(\lambda)$ is the response function for band I of the MSS. In principle, $\bar{\rho}$ and ρ are functions of λ . If one assumes they are constant and equal to $\bar{\rho}_I$ and ρ_I across a given band, then

$$N_I = a_I(\bar{\rho}_I, \mu_0) \rho_I + b_I(\bar{\rho}_I, \mu_0) \quad (4)$$

where

$$a_I(\bar{\rho}_I, \mu_0) = \frac{\mu_0 \int a(\bar{\rho}_I, \mu_0, \lambda) F(\lambda) S_I(\lambda) d\lambda}{\int S_I(\lambda) d\lambda} \quad (5)$$

$$b_I(\bar{\rho}_I, \mu_0) = \frac{\mu_0 \int b(\bar{\rho}_I, \mu_0, \lambda) F(\lambda) S_I(\lambda) d\lambda}{\int S_I(\lambda) d\lambda} \quad (6)$$

Thus,

$$L_I = A_I(\bar{\rho}_I, \mu_0) \rho_I + B_I(\bar{\rho}_I, \mu_0) \quad (7)$$

where

$$A_I(\bar{\rho}_I, \mu_0) = \alpha_I a_I(\bar{\rho}_I, \mu_0) \quad (8)$$

$$B_I(\bar{\rho}_I, \mu_0) = \alpha_I b_I(\bar{\rho}_I, \mu_0) + \beta_I \quad (9)$$

To simplify the notation in the preceding analysis, the parameter τ_H was not explicitly indicated. However, in the rest of this paper, it will be indicated explicitly for the A_I and B_I coefficients; that is,

$$A_I(\bar{\rho}_I, \mu_0, \tau_H) = A_I(\bar{\rho}_I, \mu_0) \quad (10)$$

$$B_I(\bar{\rho}_I, \mu_0, \tau_H) = B_I(\bar{\rho}_I, \mu_0) \quad (11)$$

A complete set of $A_I(\bar{\rho}_I, \mu_0, \tau_H)$ and $B_I(\bar{\rho}_I, \mu_0, \tau_H)$ was computed using equations (5), (6), (8), and (9) for the range of values given previously for $\bar{\rho}_I$, μ_0 , and τ_H . Also, a complete set of the coefficients C_I given by

$$C_I(\bar{\rho}_I, \mu_0, \tau_H) = A_I(\bar{\rho}_I, \mu_0, \tau_H) \bar{\rho}_I + B_I(\bar{\rho}_I, \mu_0, \tau_H) \quad (12)$$

was computed. These were required for the ATCOR program, described in a following section.

Corrections for Changes in Sun Angle, Haze Level, and Background Reflectance

Assume that Landsat data are available for a segment corresponding to the values $\bar{\rho}_I$, μ_0 , and τ_H and it is desired to "correct" these data so that they correspond to some other set of "standard" values $\bar{\rho}'_I$, μ'_0 , and τ'_H for these parameters. With the first set of

TABLE 1.—Coefficients for Relating Landsat-2 Data to the Equivalent Spectrally Flat Radiance

<i>l</i>	α_l (a)	β_l
1	4980	-4.0
2	7471	-4.5
3	8699	-5.2
4	4961	-1.8

^aThe units of α_l are $\frac{\text{cm}^2 \cdot \text{W} \cdot \mu\text{m}}{\text{W}}$.

parameters, a target of reflectance ρ_l gives rise to a gray-scale level X_l given by

$$X_l = A_l(\bar{\rho}_l, \mu_0, \tau_H) \rho_l + B_l(\bar{\rho}_l, \mu_0, \tau_H) \quad (13)$$

With the second set of parameters, the same target would give rise to a gray-scale level X'_l given by

$$X'_l = A'_l(\bar{\rho}'_l, \mu'_0, \tau'_H) \rho_l + B'_l(\bar{\rho}'_l, \mu'_0, \tau'_H) \quad (14)$$

Eliminating ρ_l from equations (13) and (14), one obtains

$$X'_l = A_l X_l + B_l \quad (15)$$

where

$$A_l = \frac{A'_l(\bar{\rho}'_l, \mu'_0, \tau'_H)}{A_l(\bar{\rho}_l, \mu_0, \tau_H)} \quad (16)$$

$$B_l = B'_l(\bar{\rho}'_l, \mu'_0, \tau'_H) - A_l B_l(\bar{\rho}_l, \mu_0, \tau_H) \quad (17)$$

Thus, if the values of $\bar{\rho}_l$, μ_0 , and τ_H for a segment are known, the data can easily be corrected to correspond to any other values of these parameters. Normally, μ_0 is known but $\bar{\rho}_l$ and τ_H are not known;

and, in making the kind of corrections described in this paper, the most difficult task is to determine the values of $\bar{\rho}_l$ and τ_H .

The ATCOR program described in the next section was designed to provide approximate values for $\bar{\rho}_l$ and τ_H and to interpolate in the tables of the A_l and B_l coefficients to obtain the appropriate coefficients to correct the data.

The ATCOR Program

The ATCOR program is based on the assumption that it is possible to obtain a reasonable estimate for the reflectance of those portions of the Earth's surface that correspond to the darkest pixels in a given Landsat segment. The haze level can then be determined from the brightness of these pixels. This question is examined in detail in reference 19. For the present discussion, it will be assumed that a reasonable estimate for the Earth's reflectance corresponding to the "darkest pixels" can be made.

In the ATCOR program, band 1 is used to determine the haze level because the assumed haze model indicates that the effect of haze is greatest in this band. The set of darkest pixels is obtained by taking from each line of Landsat data the pixel that has the lowest value in band 1. An average minimum value, called $X_{1,\min}(\tau_H)$, is obtained by averaging the values of X_1 for these pixels. Also, the average value \bar{X}_1 for all the band 1 data in the segment is computed. It is assumed that the reflectance $\rho_{1,\min}$ corresponding to the darkest targets (i.e., corresponding to the value $X_{1,\min}(\tau_H)$) is known. Next, one calculates the average minimum values $X_{1,\min}(\tau_j)$, $j = 1, 2, 3$, corresponding to the same reflectance $\rho_{1,\min}$, and to the three haze levels for which coefficients were calculated (namely $\tau_1 = 0.0$, $\tau_2 = 0.424$, and $\tau_3 = 0.848$). They are obtained in the following manner. Using the table for $C_1(\bar{\rho}_l, \mu_0, \tau_j)$ generated by equation (12), an interpolation is performed to find the value $\bar{\rho}_{1,j}$ of $\bar{\rho}_l$ for which $C_1(\bar{\rho}_{1,j}, \mu_0, \tau_j) = \bar{X}_1$. This is done for $j = 1, 2$, and 3. Then, using the tables for A_l and B_l , an interpolation is performed to determine the coefficients $A_l(\bar{\rho}_{1,j}, \mu_0, \tau_j)$ and $B_l(\bar{\rho}_{1,j}, \mu_0, \tau_j)$. Finally, $X_{1,\min}(\tau_j)$ is determined from the equation

$$X_{1,\min}(\tau_j) = A_l(\bar{\rho}_{1,j}, \mu_0, \tau_j) \rho_{1,\min} + B_l(\bar{\rho}_{1,j}, \mu_0, \tau_j) \quad (18)$$

Using the three calculated values for $X_{j,\min}(\tau_H)$ corresponding to $J = 1, 2, 3$, the value of τ_H that gives the value obtained previously for $X_{j,\min}(\tau_H)$ is determined by interpolation. This value is the estimate of τ_H .

Once an estimate of τ_H has been obtained, an estimate of the value of β_j can be made. The first step is to calculate the average values for all the data in the segment for bands 2, 3, and 4 to provide \bar{X}_j for all four bands. Then, β_j is determined by interpolating to find the value of β_j for which $C(\beta_j, \mu_0, \tau_H) = \bar{X}_j$. Finally, the program interpolates in the tables for A_j and B_j to obtain $A_j(\beta_j, \mu_0, \tau_H)$ and $B_j(\beta_j, \mu_0, \tau_H)$, which are printed out and can then be used with equations (16) and (17) to make the desired corrections.

The ATCOR program was tested on a data set consisting of seven pairs of acquisitions over three sites in Kansas (ref. 13 and Minter's paper). Each pair consisted of two acquisitions, 1 day apart, of the same site. The objective was to determine whether ATCOR could correct for haze level differences when the target was the same. One acquisition was selected as the "training segment" and the other as the "recognition segment." The recognition segment was classified with the LARSYS classifier using

1. Local training
2. Signatures from the training segment corrected by ATCOR
3. Uncorrected signatures from the training segment

To correct the training segment signatures, both segments were processed by ATCOR to obtain the corresponding values of β_j , μ_0 , and τ_H , and then equations (16) and (17) were used to compute the A_j and B_j coefficients. These were then used to transform the training data. The results showed that ATCOR generally improved the classifications, by a substantial factor in some cases. Another test of ATCOR was performed by IBM.³ In this test, the training and recognition segments were not the same. ATCOR generally improved the results but only by a small amount. However, since there apparently were only small haze differences between the training and recognition segments in most cases, large improvements could not be expected. This test is further discussed in reference 20.

Although ATCOR was designed to correct for changes in haze level (τ_H), Sun angle (θ_0), and background reflectance (β_j), its principal application in LACIE to date has been to develop Sun-angle correc-

tion tables for use in the LACIE clustering algorithm.

XSTAR: AN ALGORITHM TO CORRECT LANDSAT DATA FOR THE EFFECTS OF HAZE AND SUN ANGLE

The XSTAR preprocessing algorithm is the result of a combination of physical intuition, empirical observation, and a formulation based on the ERIM radiative transfer model for an atmosphere with no absorption (ref. 21). The algorithm is derived as follows.

Letting primes denote quantities which characterize a standardized measurement condition, one first represents the optical thickness, τ'_i (for each MSS channel i), for this condition by

$$\tau'_i = \tau'_{Ri} + \alpha_i \gamma' \quad (19)$$

with τ'_{Ri} representing the Rayleigh optical thickness and with $\alpha_i \gamma'$ representing the aerosol optical thickness in each MSS channel such that γ' is a scalar parameter (independent of channel number) related to the amount of haze in the atmosphere and α_i is a corresponding function of channel number (or wavelength) which is assumed to be independent of the amount of atmospheric haze. For Landsat-2 data, by definition

$$\alpha = \begin{pmatrix} 1.2680 \\ 1.0445 \\ .9142 \\ .7734 \end{pmatrix} \quad (20)$$

for the Landsat spectral bands 1 through 4. The parameter γ' can then be seen to characterize the aerosol optical thickness (for the standardized condition) in a hypothetical spectral band for which $\alpha_i = 1$. The values for α_i were calculated from the esti-

³S. G. Wheeler, "Results of Signature Extension Experiment," IBM Memo 1874-RES-23-14, July 29, 1976. Also, see the paper by Minter.

mated Landsat in-band optical thickness for an atmosphere with a horizontal visual range of 23 kilometers (a relatively clear atmosphere).

Similarly, for an observed condition, the optical thickness τ_j is represented by

$$\tau_j = \tau_{RI} + \alpha_j(\gamma' + \gamma) \quad (21)$$

Since the Rayleigh optical thickness is independent of the amount of atmospheric haze ($\tau_{RI} = \tau'_{RI}$), one may write

$$\tau_j = \tau'_j + \alpha_j\gamma \quad (22)$$

The parameter γ then measures the change in optical thickness from the standardized condition.

Representing the MSS signals for the observed and standardized conditions by X_j and X'_j , respectively, and assuming that other variables in the radiative transfer equation are restricted so that the only significant variable is atmospheric optical thickness, the equation relating the signal X_j to its standardized value X'_j (ref. 15) becomes

$$X_j = e^{\alpha_j\gamma} X'_j + (1 - e^{\alpha_j\gamma}) X'_j + P(\alpha_j\gamma) \quad (23)$$

In general, the quantities X'_j and $P(\alpha_j\gamma)$ are both functions of the scanner calibration and of the illumination geometry, viewing geometry, optical thickness, and background albedo of the standardized condition (ref. 15). The polynomial function $P(\alpha_j\gamma)$ is also a function of $\alpha_j\gamma$, with its first term proportional to $(\alpha_j\gamma)^2$, and thus represents higher order effects of changes in optical thickness.

The XSTAR algorithm is based on the mathematical form of equation (23), excluding the higher order terms represented by $P(\alpha_j\gamma)$.

$$X_j = e^{\alpha_j\gamma} X'_j + (1 - e^{\alpha_j\gamma}) X'_j \quad (24)$$

Alternatively, one may write

$$(X'_j - X_j) = e^{\alpha_j\gamma} (X_j - X'_j) \quad (25)$$

From equation (25), it is apparent that the vector X^* specifies a point, or an origin, in the signal space relative to which the remainder of the signal space expands or contracts according to the effect of each multiplicative factor. The existence of the point established by X^* has led to the name XSTAR for the resulting preprocessing algorithm.

For Landsat-2 LACIE segment data (and for Landsat-2 full-frame data from computer-compatible tapes (CCT's) produced before July 16, 1975) which are acquired for a Sun zenith angle of 39° , by definition² (ref. 15)

$$X^* = \begin{pmatrix} 61.9 \\ 66.2 \\ 83.2 \\ 33.9 \end{pmatrix} \quad (26)$$

for the Landsat spectral bands 1 through 4. For other Sun zenith angles, a cosine Sun-angle correction must be applied to the data before applying the XSTAR correction; hence, equation (25) then becomes

$$X'_j = e^{\alpha_j\gamma} \frac{\mu'_0}{\mu_0} X_j + (1 - e^{\alpha_j\gamma}) X'_j \quad (27)$$

with $\mu'_0 = \cos 39^\circ$ and with μ_0 representing the cosine of the Sun zenith angle for the data acquisition to be corrected.

For Landsat-1 data (and for Landsat-2 full-frame data from CCT's produced on or after July 16, 1975), corrections to simulate the Landsat-2 LACIE segment data calibration first need to be applied before using equation (27) or any subsequent equations. These corrections are defined in ERIM memo IS-PFL-1916.

²P. F. Lambeck, "Revised Implementation of the XSTAR Haze Correction Algorithm and Associated Preprocessing Steps for Landsat Data," ERIM Memo IS-PFL-1916, Nov. 1977.

To apply the XSTAR preprocessing algorithm to Landsat data, one needs to determine the appropriate value for γ , which measures the amount of correction required. Fortunately, Landsat data distributions tend to lie within a two-dimensional hyperplane, when displayed in the four-dimensional Landsat data space, and this hyperplane shifts its position according to the effects of atmospheric haze. The direction in which this shifting is most easily discernible is specified by the unit vector \hat{y} (ref. 15).

$$\hat{y} \equiv \begin{pmatrix} -.89952 \\ .42830 \\ .07592 \\ -.04080 \end{pmatrix} \quad (28)$$

The \hat{y} direction is equivalent to the tasseled-cap "yellowness" direction (see the symposium paper by Kauth et al. entitled "Feature Extraction Applied to Agricultural Crops as Seen by Landsat"), and it measures the component of the shift of the data hyperplane which is perpendicular to the usual orientation of the plane. For the standardized condition, the average signal value, measured in the \hat{y} direction, is represented by Y^* , with

$$Y^* \equiv -11.2082 \text{ Landsat counts} \quad (29)$$

(This Y^* value has been chosen to represent a typical atmospheric condition, not necessarily a clear one.) Thus, one calculates the value for γ that will shift the mean signal value (\bar{X}_j) for the data acquisition to be corrected such that the corrected mean signal value, measured in the \hat{y} direction, will equal Y^* .

$$\sum_{I=1}^4 \left[e^{\alpha_I \gamma} \frac{\mu_0'}{\mu_0} \bar{X}_I + (1 - e^{\alpha_I \gamma}) X_I^* \right] \hat{y}_I = Y^* \quad (30)$$

If $e^{\alpha_I \gamma}$ is expanded as a series in ascending powers of $\alpha_I \gamma$ and if third order and higher order terms are ignored, one may estimate γ by calculating

$$a = \sum_{I=1}^4 \alpha_I^2 \left(\frac{\mu_0'}{\mu_0} \bar{X}_I - X_I^* \right) \hat{y}_I \quad (31)$$

$$b = \sum_{I=1}^4 \alpha_I \left(\frac{\mu_0'}{\mu_0} \bar{X}_I - X_I^* \right) \hat{y}_I \quad (32)$$

$$c = \sum_{I=1}^4 \left(\frac{\mu_0'}{\mu_0} \bar{X}_I \hat{y}_I \right) - Y^* \quad (33)$$

then setting

$$\gamma = -\frac{b}{a} \left\{ 1 - \sqrt{1 - \frac{2ac}{b^2}} \right\} \quad (34)$$

For extremely hazy conditions, the quantity under the radical in equation (34) can be negative. In such cases, the square root may be set equal to zero; i.e.,

$$\gamma = -\frac{b}{a} \quad (35)$$

The mean signal value (\bar{X}_j) used in equations (31) through (33) should be calculated from pixels that do not represent clouds, snow, nonuniform haze concentrations, cloud shadows, and water, so that the estimate for γ will not be biased. A fully automated technique for doing this (called SCREEN) has been developed and is documented in ERIM memo IS-PFL-1916.

The quantities \bar{X}_j and γ are calculated during one pass through the data. Equation (27) is then used to apply the correction during a second pass.

The XSTAR preprocessing algorithm is unique in its method for estimating relative changes in optical thickness (γ) in the absence of ground references or ground observations. The algorithm also retains the original form of the data after applying its preprocessing correction.

A test of XSTAR on 90 Landsat-2 consecutive-day data sets, representing a wide range of Sun zenith angles, scene characteristics, and atmospheric conditions, has indicated that XSTAR, compared to no preprocessing, doubled the number of consecutive-day data sets for which the day-to-day euclidean distance between the signal mean vectors was less than 3 Landsat counts (an estimated upper bound on acceptable performance). In all, one-half to two-thirds

of the data sets were brought within three Landsat counts of matching after applying XSTAR, and the remaining data sets (scenes more than 20-percent covered by clouds, cloud shadows, or snow) were in general significantly improved by XSTAR. These results are plotted in figure 3.

The XSTAR algorithm does not attempt to compensate for the effects of view angle, background albedo, atmospheric absorption, or inconsistencies in the calibration of the data. However, these effects appear to be of lesser consequence in Landsat data than the effects of haze and Sun angle for which XSTAR does apply a correction.

Since its development, the XSTAR algorithm has been tested only on Landsat agricultural data. Its performance characteristics on nonagricultural data are not yet known.

CONCLUDING COMMENTS

The ATCOR algorithm is based on a detailed atmospheric model and should give good results if the "haze diagnostic" part of the algorithm gives an accurate estimate of the haze level and if this haze level is reasonably constant across the image being corrected. The haze diagnostic should be accurate if the average reflectance of the darkest objects in the scene corresponds to the "standard value" assumed for this reflectance by the algorithm. When this correspondence is poor, the results can be unsatisfactory and this is undoubtedly the greatest source of error in applying ATCOR. However, the idea in ATCOR of tabulating preprocessing correction factors from an accurate mathematical model and then interpolating to estimate an appropriate correction for each scene appears to be a significant step. Future development efforts should concentrate on finding an improved haze diagnostic.

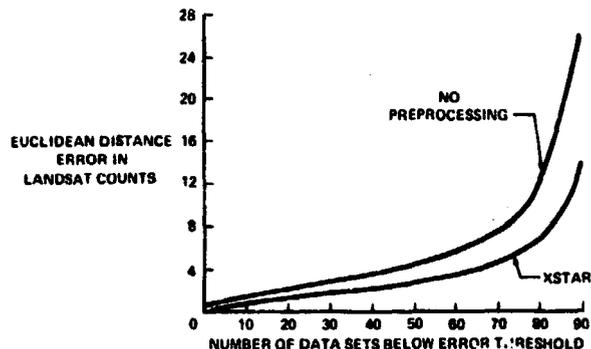


FIGURE 3.—Scene average euclidean distance error from XSTAR test on 90 consecutive-day LACIE acquisitions (after cosine correction for Sun angle).

Although the XSTAR preprocessing algorithm relies on a simplified atmospheric model, it uses a haze diagnostic (displacement of the data mean, measured in the ϑ direction) that is especially well suited to the requirements of the algorithm. This occurs because displacements of Landsat data measured in the ϑ direction correlate significantly with the multiplicative and additive changes caused by changing observation conditions. As a result, the XSTAR algorithm is capable of achieving a modest level of success with great consistency. Efforts are currently being made to use the XSTAR haze diagnostic with the ATCOR algorithm and thus combine the best features of both algorithms.

Neither the ATCOR algorithm nor the XSTAR algorithm provides an explicit compensation for the effects of changing Landsat view angle. Atmospheric models and practical experience both indicate that these effects are significant, even for the narrow range of view angles pertinent to Landsat data. Development efforts are currently underway at ERIM (ref. 15) to address this aspect of the preprocessing problem.

Appendix

Calculation of the Radiance at the MSS

REFLECTION AND TRANSMISSION MATRICES

In what follows, one will frequently be concerned with reflection and transmission matrices (R and T matrices), which describe the reflection and transmission properties of the plane-parallel scattering layers assumed to make up the atmosphere. These layers are assumed to be horizontally homogeneous and to extend to infinity in the horizontal direction. For a layer of optical depth τ_1 , the reflection and transmission matrices are defined by

$$R(\mu, \phi; \mu_0, \phi_0) = \frac{N(0, +\mu, \phi)}{\mu_0 F} \quad (36)$$

$$T(\mu, \phi; \mu_0, \phi_0) = T_{\text{diff.}}(\mu, \phi; \mu_0, \phi_0) + T_0(\mu, \phi; \mu_0, \phi_0) \quad (37)$$

where

$$T_{\text{diff.}}(\mu, \phi; \mu_0, \phi_0) = \frac{N_{\text{diff.}}(\tau_1, -\mu, \phi)}{\mu_0 F} \quad (38)$$

$$T_0(\mu, \phi; \mu_0, \phi_0) = \frac{N_0(\tau_1, -\mu, \phi)}{\mu_0 F} \quad (39)$$

Here, $N(\tau, \mu, \phi)$ is the radiance at optical depth τ in the direction specified by μ and ϕ , where μ ($0 < \mu \leq 1$) is the cosine of the zenith angle θ measured from the normal to the layer and ϕ is the corresponding azimuth angle. A minus sign in front of μ indicates the direction is downward. The optical depth τ is measured from the top of the layer downward; thus, $N(0, +\mu, \phi)$ is the upward-directed radiance at the top of the layer, and $N(\tau_1, -\mu, \phi)$ is the downward-directed radiance at the bottom of the layer. The symbols with subscript zero refer to the incident radiation. The incident beam has an irra-

diance πF through a unit area normal to itself. The subscript diff. refers to diffusely transmitted radiation; i.e., radiation that has been scattered at least once. The directly transmitted radiance N_0 is given by

$$N_0(\tau_1, -\mu, \phi) = \pi F e^{-\tau_1 / \mu_0} \delta(\mu - \mu_0) \delta(\phi - \phi_0) \quad (40)$$

where δ is the Dirac delta function. Note that upward-directed radiation is all diffuse, so the subscript diff. is omitted in this case.

THE ADDING METHOD

To compute the MSS response for various values of the parameters ρ_f , θ_0 , τ_H , and \bar{p}_f , one first computes the radiance at the MSS for these values of the parameters. This is done by computing the corresponding R matrix and using equation (36) to obtain the radiance.

The method used to compute the R matrix is the adding method originally proposed in an unpublished report by Van de Hulst.⁴ (See also reference 22 for comparison with other methods.) It allows one to take the R and T matrices for two separate layers of optical depths τ_1 and τ_2 and construct from them the R and T matrices for the layer of optical depth $\tau_1 + \tau_2$ consisting of the two layers, one on top of the other. A special case of the adding method occurs when the two layers are identical. It is then called the doubling method. In the calculations described in the following paragraphs, the doubling method is used to build up R and T matrices for the Rayleigh and aerosol layers that constitute the model atmosphere. The adding method is then used to combine these to obtain R and T matrices for the total atmosphere. Finally, the adding method is used to combine the atmospheric matrices and the R matrix

⁴H. C. Van de Hulst, "A New Look at Multiple Scattering," unpublished report, NASA Institute for Space Studies, Jan. 1963.

for the Earth's surface to obtain an R matrix that describes the reflectance of the overall Earth/atmosphere system. This matrix is somewhat different from the conventional R matrix since it describes a system that is not horizontally homogeneous.

The principle of the adding method is depicted in figure 4, which shows two scattering layers. It is assumed that the R and T matrices have been obtained for the two layers, and it is desired to obtain the R and T matrices for the two-layer system. In figure 4, the two layers are separated so that the upward and downward radiation field where they join can be indicated.

The R and T matrices for the top layer in figure 4 will be denoted R_T and T_T and those of the bottom layer R_B and T_B . It is understood that each matrix is a function of four angular variables, which are omitted to simplify the notation. In figure 4, a part, R_1 , of the incident flux is reflected by the top layer, and a part, D_1 , is transmitted by the top layer. Of the part described by D_1 , a part, U_1 , is reflected by the bottom layer, and a part, T_1 , is transmitted by the bottom layer. The process is continued as shown in the diagram. All the transmission matrices (T_T , T_B , T , and D) include both the diffusely and directly transmitted parts. The solution consists in determining R and T , the reflection and transmission matrices for the two layers taken together. The following relations can be read directly from the diagram.

$$\left. \begin{aligned} R_1 &= R_T & D_1 &= T_T \\ T_n &= T_B D_n & U_n &= R_B D_n \\ R_{n+1} &= T_T U_n & D_{n+1} &= R_T U_n \\ n &= 1, 2, \dots \end{aligned} \right\} \quad (41)$$

By substitution and addition, one obtains

$$\begin{aligned} D &= \left[1 + (R_T R_B) + (R_T R_B)^2 + \dots \right] T_T \\ &= [1 + S] T_T \end{aligned} \quad (42)$$

$$U = R_B D \quad (43)$$

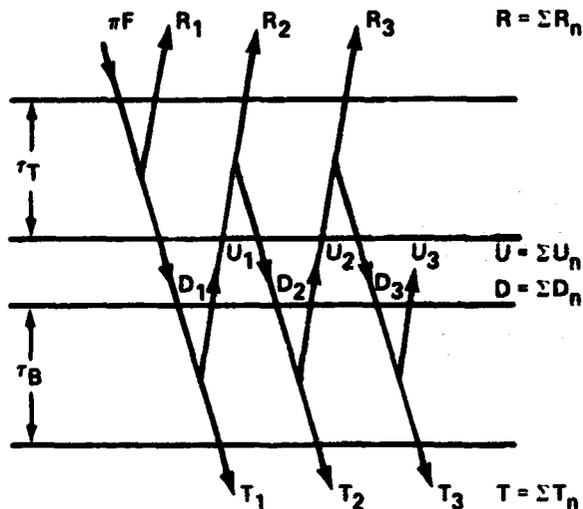


FIGURE 4.—Schematic representation of the adding method.

$$R = R_T + T_T R_B D \quad (44)$$

$$T = T_B D \quad (45)$$

The products in equations (41) through (45) stand for double integrals over the intermediate angles. For example, $U = R_B D$ stands for

$$\begin{aligned} U(\mu_1, \phi_1; \mu_2, \phi_2) &= \frac{1}{\pi} \int_0^1 \int_0^{2\pi} R_B(\mu_1, \phi_1; \mu', \phi') \\ &\quad \cdot D(\mu', \phi'; \mu_2, \phi_2) \mu' d\mu' d\phi' \end{aligned} \quad (46)$$

All other products are defined in a similar way.

Separating the directly transmitted and diffuse parts of T_T , T_B , D , and T , one obtains

$$D_{\text{diff.}} = T_{T,\text{diff.}} + S e^{-\tau T/\mu_0} + S T_{T,\text{diff.}} \quad (47)$$

$$U = R_B e^{-\tau T/\mu_0} + R_B D_{\text{diff.}} \quad (48)$$

$$R = R_T + e^{-\tau T/\mu} U + T_{T,\text{diff.}} U \quad (49)$$

$$T_{\text{diff.}} = e^{-\tau_B/\mu} D_{\text{diff.}} + T_{B,\text{diff.}} e^{-\tau_T/\mu_0} + T_{B,\text{diff.}} D_{\text{diff.}} \quad (50)$$

where the subscript diff. indicates the diffuse part of the corresponding matrix.

Since the bottom and top layers are homogeneous, the solutions are even functions of $\phi_0 - \phi$ and can be expanded in the form

$$R_B(\mu, \phi; \mu_0, \phi_0) = \sum_{m=0}^{N_B} R_B^{(m)}(\mu, \mu_0) \cos m(\phi_0 - \phi) \quad (51)$$

$$T_{B,\text{diff.}}(\mu, \phi; \mu_0, \phi_0) = \sum_{m=0}^{N_B} T_{B,\text{diff.}}^{(m)}(\mu, \mu_0) \cos m(\phi_0 - \phi) \quad (52)$$

with identical series for R_T and $R_{T,\text{diff.}}$, except that everywhere the subscripts are T instead of B . Most methods of solving the multiple scattering problem for a homogeneous layer, including the doubling method used in this paper, give solutions in this form. The number of components $N_B + 1$ in equations (51) and (52) is the number of components in the cosine expansion of the scattering diagram for the layer. (See equation (64) and the discussion following it.) Substituting equations (51) and (52) and the corresponding series for the bottom layer into equations (47) through (50), one obtains similar series for R and $T_{\text{diff.}}$ describing the two layers taken together.

$$R(\mu, \phi; \mu_0, \phi_0) = \sum_{m=0}^N R^{(m)}(\mu, \mu_0) \cos m(\phi_0 - \phi) \quad (53)$$

$$T_{\text{diff.}}(\mu, \phi; \mu_0, \phi_0) = \sum_{m=0}^N T_{\text{diff.}}^{(m)}(\mu, \mu_0) \cos m(\phi_0 - \phi) \quad (54)$$

where N is the larger of N_T and N_B . The coefficients $R^{(m)}(\mu, \mu_0)$ and $T_{\text{diff.}}^{(m)}(\mu, \mu_0)$ are given by the following equations.

$$Q_1^{(m)}(u, v) = (1 + \delta_{0,m}) \int_0^1 R_T^{(m)}(u, z) R_B^{(m)}(z, v) z dz \quad (55)$$

$$Q_{n+1}^{(m)}(u, v) = (1 + \delta_{0,m}) \int_0^1 Q_1^{(m)}(u, w) Q_n^{(m)}(w, v) w dw \quad (56)$$

$$S^{(m)}(u, v) = \sum_{n=1}^{\infty} Q_n^{(m)}(u, v) \quad (57)$$

$$D_{\text{diff.}}^{(m)}(u, \mu_0) = T_{T,\text{diff.}}^{(m)}(u, \mu_0) + S^{(m)}(u, \mu_0) e^{-\tau_T/\mu_0} + (1 + \delta_{0,m}) \int_0^1 S^{(m)}(u, v) T_{T,\text{diff.}}^{(m)}(v, \mu_0) v dv \quad (58)$$

$$U^{(m)}(z, \mu_0) = R_B^{(m)}(z, \mu_0) e^{-\tau_T/\mu_0} + (1 + \delta_{0,m}) \int_0^1 R_B^{(m)}(z, u) D_{\text{diff.}}^{(m)}(u, \mu_0) u du \quad (59)$$

$$R^{(m)}(\mu, \mu_0) = R_T^{(m)}(\mu, \mu_0) + e^{-\tau_T/\mu} U^{(m)}(\mu, \mu_0) + (1 + \delta_{0,m}) \int_0^1 T_{T,\text{diff.}}^{(m)}(\mu, z) U^{(m)}(z, \mu_0) z dz \quad (60)$$

$$T_{\text{diff.}}^{(m)}(\mu, \mu_0) = e^{-\tau_B/\mu} D_{\text{diff.}}^{(m)}(\mu, \mu_0) + T_{B,\text{diff.}}^{(m)}(\mu, \mu_0) e^{-\tau_T/\mu_0} + (1 + \delta_{0,m}) \int_0^1 T_{B,\text{diff.}}^{(m)}(\mu, u) D_{\text{diff.}}^{(m)}(u, \mu_0) u du \quad (61)$$

In equations (58) through (61), τ_T and τ_B are the optical depths of the top and bottom layers, respectively. Only the first few $Q_n^{(m)}(u, v)$ need to be calculated. As n increases, the series for $S^{(m)}(u, v)$ becomes a geometric series and the remaining terms can be approximated by a remainder term.

It will be assumed that the Landsat MSS is pointed vertically downward; i.e., that the look angle is 0.0° . This assumption greatly simplifies the multiple scattering calculations and seems justified since the maximum look angle is about 7° . With this assumption, the radiance at the sensor is independent of ϕ and ϕ_0 so only the $m = 0$ component in equations (51) through (61) needs to be computed.

THE DOUBLING METHOD

The doubling method is simply the adding method when the top and bottom layers are the same. By repeated doubling, one can obtain the solution for a thick homogeneous layer if one has the solution for a thin homogeneous layer. One begins with a layer of optical depth τ_1 and "adds" it to itself using the adding method to obtain solutions for a layer of depth $2\tau_1$. By repeating the procedure, one successively obtains solutions for depths $4\tau_1, 8\tau_1, 16\tau_1, \dots, 2^n\tau_1$ after n doublings.

Hansen (ref. 23) has shown that a good method for obtaining the initial layer of depth τ_1 is to take τ_1 small enough that only first-order scattering is important. One then has the solutions

$$R_B(\mu, \phi; \mu_0, \phi_0) = \frac{1}{4\mu\mu_0} \left(\frac{1}{\mu} + \frac{1}{\mu_0} \right)^{-1} \left\{ 1 - \exp \left[\tau_1 \left(\frac{1}{\mu} + \frac{1}{\mu_0} \right) \right] \right\} \cdot P(\mu, \phi; -\mu_0, \phi_0) \quad (62)$$

$$T_{B,\text{diff}}(\mu, \phi; \mu_0, \phi_0) = \frac{1}{4\mu\mu_0} \left(\frac{1}{\mu} + \frac{1}{\mu_0} \right)^{-1} \cdot \left[\exp \left(\frac{-\tau_1}{\mu_0} \right) - \exp \left(\frac{-\tau_1}{\mu} \right) \right] \cdot P(-\mu, \phi; -\mu_0, \phi_0) \quad (63)$$

where $P(\mu, \phi; \mu_0, \phi_0)$ is the scattering diagram describing scattering from the direction characterized by μ_0, ϕ_0 to that characterized by μ, ϕ . The convention regarding minus signs in front of μ_0 and μ was discussed previously (following eq. (39)). There are identical expressions for R_T and $T_{T,\text{diff}}$. In general, $\tau_1 = 2^{-25}$ is small enough for these solu-

tions to be sufficiently accurate. This is the method used in the doubling calculations described in this paper.

To perform the numerical calculations, one separates the azimuthal dependence by expanding the scattering diagram in a cosine series

$$P(\mu, \phi; -\mu_0, \phi_0) = \sum_{m=0}^N P^{(m)}(\mu, \mu_0) \cos m(\phi_0 - \phi) \quad (64)$$

where coefficients $P^{(m)}(\mu, \mu_0)$ are as given in reference 16, page 150, equation (87). One then has a similar expansion for R_B and $T_{B,\text{diff}}$.

$$R_B(\mu, \phi; \mu_0, \phi_0) = \sum_{m=0}^N R_B^{(m)}(\mu, \mu_0) \cos m(\phi_0 - \phi) \quad (65)$$

$$T_{B,\text{diff}}(\mu, \phi; \mu_0, \phi_0) = \sum_{m=0}^N T_{B,\text{diff}}^{(m)}(\mu, \mu_0) \cos m(\phi_0 - \phi) \quad (66)$$

Then, $R_T^{(m)}(\mu, \mu_0) = R_B^{(m)}(\mu, \mu_0)$ and $T_{T,\text{diff}}^{(m)}(\mu, \mu_0) = T_{B,\text{diff}}^{(m)}(\mu, \mu_0)$ are substituted into equations (55) through (61) to begin the doubling process. In the calculations described in this paper, only the $m = 0$ component was calculated, for the reasons given previously.

REFLECTION AND TRANSMISSION MATRICES FOR THE ATMOSPHERE

The reflection and transmission matrices that describe the total atmosphere are denoted $R_A(\mu, \mu_0, \lambda)$ and $T_A(\mu, \mu_0, \lambda)$, where the subscript A stands for "all." Similarly, $R_R(\mu, \mu_0, \lambda)$ and $T_R(\mu, \mu_0, \lambda)$ describe the upper Rayleigh scattering layer, and $R_H(\mu, \mu_0, \lambda)$ and $T_H(\mu, \mu_0, \lambda)$ describe the lower haze scattering layer. Here the superscript m has been omitted, but it is understood that all of these matrices correspond to $m = 0$. The parameter λ has been added to indicate the wavelength. The calculations are performed for 25 values of μ and μ_0 (namely, 24 gauss points plus the value 1.0) and for every value of λ from 0.4 to 1.1 micrometers in steps of 0.01 micrometer. They are

also performed for three values of τ_H ; namely, 0.0, 0.424, and 0.848. For a given value of τ_H , the calculations are made for all values of λ . Under the assumptions made previously, the only difference between the calculations for different values of λ is that the Rayleigh and haze optical depths are different, as shown in figure 2.

The simplest case is for $\tau_H = 0$. Then, $R_A(\mu, \mu_0, \lambda) = R_R(\mu, \mu_0, \lambda)$, and $T_A(\mu, \mu_0, \lambda) = T_R(\mu, \mu_0, \lambda)$. The doubling method was first used to obtain the R and T matrices corresponding to Rayleigh scattering layers of optical depth $2^{-24}, 2^{-23}, \dots, 2^{-11}$. This was done by starting with a Rayleigh scattering layer of optical depth 2^{-25} and doubling 36 times. The larger values of optical depth were not required for this paper but are routinely calculated by the doubling program. The calculation of $R_R(\mu, \mu_0, \lambda)$ and $T_R(\mu, \mu_0, \lambda)$ was begun with the largest value of λ ; namely, $\lambda = 1.1$ micrometers. The optical depth for the corresponding layer $\tau_R(1.1)$ was obtained (fig. 2) and the matrices $R_R(\mu, \mu_0, 1.1)$ and $T_R(\mu, \mu_0, 1.1)$, describing a Rayleigh scattering layer of this optical depth, were built up using the adding method to "add" certain previously calculated layers that were selected so that the sum of their optical depths was equal to $\tau_R(1.1)$. Next, $R_R(\mu, \mu_0, 1.09)$ and $T_R(\mu, \mu_0, 1.09)$, were calculated. Since $\tau_R(1.09)$ is larger than $\tau_R(1.1)$, this involved "adding" more Rayleigh scattering layers to the layer used to compute $R_R(\mu, \mu_0, 1.1)$ and $T_R(\mu, \mu_0, 1.1)$. This was done as before by using the adding method. This procedure was continued until the calculations had been made for all the selected values of λ . For the cases $\tau_H = 0.424$ and $\tau_H = 0.848$, the procedure was the same except that for each value of λ , the matrices $R_H(\mu, \mu_0, \lambda)$ and $T_H(\mu, \mu_0, \lambda)$ were built up (in the same way as $R_R(\mu, \mu_0, \lambda)$ and $T_R(\mu, \mu_0, \lambda)$) and the adding method was used to calculate $R_A(\mu, \mu_0, \lambda)$ and $T_A(\mu, \mu_0, \lambda)$ by "adding" the Rayleigh scattering layer on top of the haze layer.

RADIANCE AT THE MSS

This section describes the calculation of the R function associated with the Earth/atmosphere system from which the radiance at the MSS can be obtained using equation (36). It was assumed that, for a given wavelength, the Earth's surface is a Lambert reflector with a reflectance $\rho(\lambda)$ for the pixel in the field of view at a particular instant and an average reflectance $\bar{\rho}(\lambda)$ for the background (i.e., the area

around that pixel). If the whole surface of the Earth, including the pixel in the field of view, had a uniform reflectance $\bar{\rho}(\lambda)$, the desired R matrix could be obtained from equation (60). In this case, the top layer would be described by $R_T(\mu, \mu_0, \lambda)$ and $T_T(\mu, \mu_0, \lambda)$ and the bottom layer would be described by $R_B(\mu, \mu_0, \lambda) = \bar{\rho}(\lambda)$ and $T_B(\mu, \mu_0, \lambda) = 0$. Also, $\mu = 1$ since it was assumed that the MSS was pointed vertically downward. Under these conditions, the $m = 0$ component of equation (60) can be written

$$R(1, \mu_0, \lambda) = R_T(1, \mu_0, \lambda) + e^{-\tau_T(\lambda)} \bar{\rho}(\lambda) D'(\mu_0, \lambda) + 2 \int_0^1 T_{T, \text{diff.}}(1, z, \lambda) U(z, \mu_0, \lambda) z dz \quad (67)$$

where

$$\left. \begin{aligned} \tau_T(\lambda) &= \tau_H(\lambda) + \tau_R(\lambda) \\ D'(\mu_0, \lambda) &= e^{-\tau_T(\lambda)/\mu_0} + 2 \int_0^1 D_{\text{diff.}}(\mu, \mu_0, \lambda) \mu du \end{aligned} \right\} \quad (68)$$

Equation (67) is not exact because the top layer was assumed to be homogeneous in the derivation of equation (60) and the top layer is not homogeneous in equation (67). However, this difference should cause only a small error, which will be neglected in what follows. It is easy to reformulate the theory to treat an inhomogeneous top layer exactly; however, to date, no numerical results for this case have been obtained.

If one considers the three terms on the right-hand side of equation (67), it is clear that the second term represents the radiance that is directly transmitted to the MSS from the target in the field of view and the other two terms represent the path radiance. The first term represents a contribution from the atmosphere alone, and the third term represents a contribution to the path radiance from light that has been scattered by the Earth's surface. Thus, if the reflectance of the pixel in the field of view were changed from $\bar{\rho}(\lambda)$ to $\rho(\lambda)$, the main effect would be to change $\bar{\rho}(\lambda)$ to $\rho(\lambda)$ in the second term on the right-hand side of equation (67). The effect on the other terms and on $D'(\mu_0, \lambda)$ should be negligible. Thus, if the reflectance of the pixel in the field of

view is $\rho(\lambda)$ and the background reflectance is $\bar{\rho}(\lambda)$, the corresponding reflection matrix is given approximately by:

$$\begin{aligned} R(\bar{\rho}, \rho, \mu_0, \lambda) &= R(1, \mu_0, \lambda) \\ &+ e^{-\tau_T(\lambda)} D'(\mu_0, \lambda) [\rho(\lambda) - \bar{\rho}(\lambda)] \\ &= a(\bar{\rho}, \mu_0, \lambda) \rho(\lambda) + b(\bar{\rho}, \mu_0, \lambda) \end{aligned} \quad (69)$$

where

$$\begin{aligned} a(\bar{\rho}, \mu_0, \lambda) &= e^{-\tau_T(\lambda)} D'(\mu_0, \lambda) \\ &= e^{-\tau_T(\lambda)} U(1, \mu_0, \lambda) / \bar{\rho}(\lambda) \end{aligned} \quad (70)$$

$$b(\bar{\rho}, \mu_0, \lambda) = R(1, \mu_0, \lambda) - a(\bar{\rho}, \mu_0, \lambda) \bar{\rho}(\lambda) \quad (71)$$

In R , a , and b , the value 1 for μ has been dropped from the list of variables to simplify the notation. It should be noted that the function $R(\bar{\rho}, \rho, \mu_0, \lambda)$ has properties that are different from those associated with reflection functions as they are usually defined. However, for the present calculation, the important point is that the radiance at the sensor is given by

$$\begin{aligned} N(\bar{\rho}, \rho, \mu_0, \lambda) &= \mu_0 F(\lambda) R(\bar{\rho}, \rho, \mu_0, \lambda) \\ &= \mu_0 F(\lambda) [a(\bar{\rho}, \mu_0, \lambda) \rho(\lambda) + b(\bar{\rho}, \mu_0, \lambda)] \end{aligned} \quad (72)$$

where $F(\lambda)$ is the solar irradiance at the top of the atmosphere at wavelength λ .

For each of the 3 values of τ_H , the coefficients $a(\bar{\rho}, \mu_0, \lambda)$ and $b(\bar{\rho}, \mu_0, \lambda)$ were computed for the 71 values of λ , the 25 values of μ_0 , and 50 values of $\bar{\rho}$ ranging from 0.0 to 0.5 in units of 0.01. This was done by using the adding program in the usual way, with the top layer described by $R_T(\mu, \mu_0, \lambda)$ and $T_T(\mu, \mu_0, \lambda)$ and the bottom layer described by

$R_B(\mu, \mu_0, \lambda) = \bar{\rho}(\lambda)$ and $T_B(\mu, \mu_0, \lambda) = 0$. This produced the matrices $R(\mu, \mu_0, \lambda)$ and $U(\mu, \mu_0, \lambda)$, and the values of these for $\mu = 1$ were used to compute $a(\bar{\rho}, \mu_0, \lambda)$ and $b(\bar{\rho}, \mu_0, \lambda)$.

REFERENCES

1. Potter, J.; and Shelton, M.: Effect of Atmospheric Haze and Sun Angle on Automatic Classification of ERTS-1 Data. Proceedings of the 9th International Symposium on Remote Sensing of Environment, Apr. 15-19, 1974, pp. 865-874.
2. Pitts, D. E.; McAllum, W. E.; and Dillinger, A. E.: The Effect of Atmospheric Water Vapor on Automatic Classification of ERTS Data. Proceedings of the 9th International Symposium on Remote Sensing of Environment, Apr. 15-19, 1974, pp. 483-498.
3. Turner, R. E.; Malila, W. A.; and Nalepka, R. F.: Importance of Atmospheric Scattering in Remote Sensing. Proceedings of the 7th International Symposium on Remote Sensing of Environment, Willow Run Laboratories, University of Michigan (Ann Arbor), May 17-21, 1971, pp. 1651-1697.
4. Large Area Crop Inventory Experiment (LACIE), Phase I Evaluation Report. NASA Johnson Space Center, Houston, Tex., May 1976.
5. Large Area Crop Inventory Experiment (LACIE), Phase II Evaluation Report. NASA Johnson Space Center, Houston, Tex., July 1977.
6. Large Area Crop Inventory Experiment (LACIE), First Interim Phase III Evaluation Report. NASA Johnson Space Center, Houston, Tex., Jan. 1978.
7. Henderson, R. G.; Thomas, G. S.; and Nalepka, R. F.: Methods of Extending Signatures and Training Without Ground Information. ERIM 109600-16-F, Environmental Research Institute of Michigan (Ann Arbor), May 1975.
8. Finley, D. R.: A New Algorithm to Correct MSS Signatures for Atmospheric Effects. LEC-6246, Lockheed Electronics Co., Houston, Tex., June 1975.
9. Lambeck, P. F.; and Rice, D. P.: Signature Extension Using Transformed Cluster Statistics and Related Techniques. ERIM 122700-70-F, ERIM, Ann Arbor, Mich., May 1976.
10. Finley, D. R.; and Wehmanen, O. A.: Description of the OSCAR Algorithm for Implementation on the Prototype Signature Extension System. LEC-8774, Lockheed Electronics Co., Houston, Tex., June 1976.

11. Thadani, S. G.: The MLEST Algorithm. Proceedings of the 4th Symposium on Machine Processing of Remotely Sensed Data, Laboratory for Applications of Remote Sensing, Purdue University (W. Lafayette, Ind.), June 21-23, 1977.
12. Peters, B. C.; and Walker, H.: An Iterative Procedure for Obtaining Maximum Likelihood Estimates of the Parameters for a Mixture of Normal Distributions. Report 43, Dept. of Math., University of Houston, July 1975.
13. Abotteen, R.; Levy, S.; et al.: Performance Tests of Signature Extension Algorithms. Proceedings of the 11th International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan (Ann Arbor), Apr. 25-29, 1977, pp. 1523-1532. (Also available as rep. LEC-8668, Sept. 1976.)
14. Potter, J. F.: The Correction of Landsat Data for the Effects of Haze, Sun Angle, and Background Reflectance. Proceedings of the 4th Symposium on Machine Processing of Remotely Sensed Data, LARS, June 21-23, 1977.
15. Lambeck, P. F.: Signature Extension Preprocessing for Landsat MSS Data. ERIM 122700-32-F, ERIM, Ann Arbor, Mich., Nov. 1977.
16. Chandrasekhar, S.: Radiative Transfer. Dover Publications, Inc. (New York), 1960.
17. Pitts, D. E.; et al.: Temporal Variations in Water Vapor and Aerosol Optical Depth Determined by Remote Sensing. J. Appl. Meteor., vol. 16, 1977, pp. 1312-1321.
18. Reeser, W. K.: Aerosol Size Distribution Detection for PREPS Using Simple Processing Techniques. LEC 640-TR-091, Lockheed Electronics Co., Houston, Tex., Mar. 1972.
19. Potter, J. F.; and Menlowitz, M.: On the Determination of Haze Levels From Landsat Data. Proceedings of the 10th International Symposium on Remote Sensing of Environment, ERIM, Oct. 6-10, 1975, pp. 695-703.
20. Thadani, S. G.: Results of the MLEST Signature Extension Test. LEC-9700, Lockheed Electronics Co., Houston, Tex., Sept. 1977.
21. Malila, W. A.; Cranc, R. B.; and Turner, R. E.: Information Extraction Techniques for Multispectral Scanner Data. Report No. WRL 31650-74-T, Appendix I, Willow Run Laboratories, University of Michigan, Ann Arbor, June 1972.
22. Potter, J. F.: Scattering and Absorption in the Earth's Atmosphere. Proceedings of the 6th International Symposium on Remote Sensing of Environment, 1969, p. 415.
23. Hansen, J.: Radiative Transfer by Doubling Very Thin Layers. Astrophys. J., vol. 155, 1969, pp. 565-573.

Development of Partitioning as an Aid to Spectral Signature Extension*

R. W. Thomas,^a C. M. Hay,^a and J. C. Claydon^a

INTRODUCTION

The act of sorting sample segments into sets having similar spectral reflectance characteristics has come to be known as "spectral partitioning" in the LACIE context. As part of the LACIE supporting research effort, this technique has been investigated as a means of maximizing the efficiency of Landsat classification using signature extension procedures. Lambeck and Rice (ref. 1) suggested that performance of affine signature extension algorithms (incorporating corrections for both multiplicative and additive spectral differences) would be enhanced by attempting extensions between sample segments falling in the same spectral partitions. That is, within-group spectral variation should be more limited than the variation in the population as a whole, enabling more precise estimates of parameters used in signature extension algorithms (refs. 2 and 3). More recent results with multisegment classification (ref. 4) also suggest that partitioning may serve an important role in simultaneously classifying several segments using nonlocal signatures.

Lists of spectrally similar LACIE sample segments forming spectral partitions may be constructed with the use of three basic information types: static, seasonal, and pass-specific variables. All are directly or indirectly related to spectral signature behavior and can be used to define spatial domains over which crop-specific signatures should be extendable. For example, relatively slowly changing climatic and soil characteristics can be used to describe, for any area, growth potentials for specific crops. As a consequence, stratification of the landscape into static domains or strata within which crop development, and therefore spectral development,

should be similar is possible using these "static" variables. The static strata may in turn be subdivided or combined by use of seasonal variable information. For instance, departures from their long-term averages of accumulated growing-season precipitation, heat input, and other variables specific to the growing year can be employed to adjust stratum boundaries. Finally, Landsat pass-specific scanner, atmospheric, and spectral information can be used to further subdivide or refine spectral stratum boundaries.

Combination of static or seasonal spectral strata with pass-specific information produces a dynamic partitioning of the landscape and therefore of the LACIE sample segment population. The relative importance of static, seasonal, and real-time variables in defining useful spectral partitions has been the subject of significant debate during LACIE. This paper will address that question using a static spectral stratification as the initial partitioning device. After procedures for producing the static strata have been defined, the relative capability of those strata to account for variability in wheat signatures will be evaluated and contrasted with spectral variance explained by seasonal and Landsat pass-specific information. Much of this work has been reported more extensively by Hay and Thomas (refs. 3 and 5) and by Hay et al. (refs. 6 and 7).

STATIC SPECTRAL STRATIFICATION PROCEDURE

Overview

Viable static spectral stratification procedures depend on the use of parameters that influence long-term patterns of spectral signature and are easily measurable as well. These parameters are of two

*Work supported under NASA Contract NAS 9-14565.

^aUniversity of California at Berkeley.

basic types. The first set relates to physical site (field) conditions. Soil type, physiographic position, slope, and aspect represent such site characteristics. The second set may be described as crop growth drivers. These include (1) climatic variables and (2) long-term cultural practices, such as irrigation, fertilization, and mulching, which affect the amount and timing of available water and nutrients.

Hay and Thomas (ref. 3) describe a static spectral stratification technique employing subsets of both site and growth driver variables. Developed in support of the LACIE Supporting Research and Technology (SR&T) Program, their procedure uses a combination of broad climatic strata within which a finer mosaic of soil association and land use strata is set. Sample segments belonging to a given climatic/soil/land use stratum (or combination of these strata) are expected to have similar wheat signatures on the "average." On any given Landsat pass date, lists of matched training (signature source) and recognition (targets of signature extension) segments within strata must be adjusted for seasonal abnormalities or atmospheric effects.

Selection of Stratification Variables

The set of specific variables chosen by Hay and Thomas to produce static signature extension strata included (1) general soil type (soil association), (2) land use, (3) average long-term growing-season degree-days, and (4) average long-term growing-season precipitation. The rationale for selection of each of these variables was as follows.

Soil type and land use.—Excluding atmospheric effects, the spectral signature of a cropped field (i.e., wheat) is a composite signature made up of two general components. The first is the spectral reflectance of the soil background and the second is the spectral reflectance of the crop (vegetation) canopy. The amount that each component contributes to the composite signature is dependent on the percentage of canopy cover. The relative contribution of soil background to the composite signature on any given date is inversely related to the percentage of crop canopy present on that date.

The spectral reflectance of soils has been shown to be dependent primarily on surface moisture content, organic matter content, and particle size (refs. 8 and 9). The surface moisture content of a soil is a function of the moisture input (i.e., precipitation or irrigation water), the period of time since the last input,

the texture of the soil (particle size distribution), the organic content of the soil, and the quality of the site drainage. The general overall soil background tones within areas delineated on Landsat color composites in this study were positively correlated with available moisture capacity indices for the major soil series within the associations. The darker toned soils had higher available water capacity indices than the lighter toned soils, indicating that, on a given date, surface moisture content accounted for a significant amount of the variability in spectral reflectance of the soil background. Thus, soil association (general soil type) grouped by available water capacity indices and precipitation (moisture input) was chosen as a static signature extension stratification variable.

Within certain image biophases, there exist crops whose spectral signatures can be confused with those of wheat. The presence of a confusion crop within an area is dependent on the general land use/crop type distribution patterns within an area. In addition, crop canopy density and development can be affected by cropping practices such as irrigation. Irrigation availability also tends to diversify an agricultural environment, thereby increasing the probability of confusion crops. For this reason, land use was chosen as a static signature extension stratification variable.

Growing-season degree-days and precipitation.—Spectral reflectance from the vegetative canopy is a function of the crop canopy density and the phenological stage of crop development. Within a given region, crop phenological development, and therefore crop canopy, is greatly dependent on the climatic variables of temperature and precipitation. A wide review of the literature (including important references 10 to 13) indicates that the following climatic variables could be used for stratification.

1. Average growing-season degree-day¹ sums
2. Average growing-season precipitation
3. Average last date of spring frost
4. Average temperature and/or average minimum temperature for the coldest month of the year

¹Degree-day is a measure of daily accumulated temperature above a specified biological threshold temperature. The degree-day value for a given month S is defined (ref. 13) as the number of days in a given month n times the difference between the average temperature in that given month \bar{T} and a growth threshold temperature (commonly 40° F) below which wheat has been found not to accumulate significant biomass. Thus, the growing-season day-degree sum can be expressed as $S = \sum_j n_j \cdot (\bar{T}_j - 40^\circ \text{ F})$, where j is the month index.

Generally, the isolines of the temperature variables are very positively correlated with one another so that only one or two need be considered for use in stratification. Thus, average (long-term) growing-season precipitation and average growing-season day-degree sums were included as static stratification variables.

Stratification Procedure

Land use/soils.—The stratification procedure was developed over two LACIE test areas in the U.S. Great Plains: Kansas (winter wheat) and North Dakota (spring wheat). Basically, the stratification technique involved the following stepwise sequence. First, a single time of year was identified when soils and land use or cropping practice patterns of interest were most separable on Landsat color-infrared 9- by 9-inch format transparencies. One or two supplementary Landsat dates were selected to allow better identification of crop or soil moisture factors indicating different soil associations. Then, land use and soil association information was delineated on clear acetate overlaid on the full-frame 9- by 9-inch color-infrared Landsat transparencies. Soil association lines and land use lines were located according to interpretation of this imagery with reference to land use data published by the U.S. Department of Agriculture (USDA) Economics, Statistics, and Cooperatives Service (ESCS) and the USDA Soil Conservation Service (SCS). This soil association stratification was also referenced to cross-correlated county SCS soil data and regional soil association maps.

The land use classification system employed for stratification is given in the appendix. This system merges desirable features from the U.S. Geological Survey (USGS) Circular 671 system (ref. 14) and suggestions from the USDA system. The soil association classification was based on a cross-correlation of SCS county soil survey data, Aandahl's Great Plains soil distribution data,² and Landsat data.

Working initially within the counties for which detailed soil association map data were available, the map data were correlated to features observable on the Landsat transparency. Any needed boundary ad-

justments were made to bring the soil association map data into conformity with the more detailed spatial data inherent within the imagery. Correlations of soil associations across the crop reporting district (CRD) were based on

1. Commonality of soil association name
2. Image feature continuity across county boundaries
3. Similarity of descriptions of association and soil series within associations (The names of some soil series and associations changed when they crossed county boundaries.)

Correlated associations were given an alphanumeric code which contained the general soil group number from the legend code of Aandahl's map and a letter assigned to each subgroup determined from the detailed county soil survey data.

It was necessary to redefine some county soil associations in order to better serve the needs of a signature extension stratification. For example, some associations as originally defined within the county soil report contained a significant amount of variability in soil type, available moisture capacity, and land use as a result of intermingling of significantly different soil types. Within the constraints of the minimum mapping area of 10 to 15 square miles, it was possible to redefine some of the highly variable associations into more uniform associations with respect to general soil type, moisture capacity, and land use pattern. Care was taken, however, that any redefinition of soil associations did not violate the definitional concept of a soil association.

In areas for which detailed county soil association map data were not available, subdivision of Aandahl's general soil groups was accomplished through interpretation of the Landsat imagery. (This would be the technique employed over regions of the world where only very general soil maps may be available.) This procedure made use of such similarities as

1. Landform association (topographic site relationships)
2. Land use patterns
3. Soil background tone on given dates (correlatable with available moisture capacity and general soil type)
4. Proportional relationships among the preceding characteristics

Where possible, soil associations interpreted from Landsat imagery were correlated with soil associations from the county soil surveys. This was accomplished by analyses of the Landsat imagery for landform, land use, soil tone, and patterns similar to

²A. R. Aandahl, *Soils of the Great Plains Map, 1:2 500 000*, Lincoln, Nebr., 1972.

county-based descriptions and by the use of geological map data to assess parent material similarities.

In a small number of cases, soil association areas contained within their boundaries significantly different land use classes (excluding urban; i.e., intensive cropland versus rangeland). The question was then raised as to the validity of the soil association as found mapped within the county soil report. Analysis of the detailed soil series maps (more detailed than the soil association map) more often than not indicated that, in these cases, the imagery was probably more reliable for the placement of soil association boundaries than the county soil association map. Thus, land use delineation served as an iterative check on the soil association delineations.

Once the land use and soil associations had been delineated, they were combined and registered to 1:1 000 000-scale USGS base maps. Importantly, the land use and soil association classes, and consequently the combined land use/soil association classes, were defined with regard to the interrelated effects of several environmental factors (e.g., microclimate, soil characteristics, cropping practices) on wheat growth behavior and thus on spectral signature response. The assumption was that the spectral signature of wheat would tend to be similar within each such mapping unit type throughout the year, subject to the constraints of the other stratification variables (e.g., growing-season precipitation).

Combination of climatic strata with land use/soil strata.—Long-term (30-year) average growing-season degree-day and precipitation values were computed for all ground meteorological stations having complete temperature and rainfall data in the western two-thirds of Kansas and the entire state of North Dakota. These data, together with longitude and latitude coordinates for each station, were used by a computer software data handling package (MAPIT package available on the University of California at Berkeley (UCB) CDC-7600 computer) to generate isolines of degree-days and precipitation over each state. Isoline values were set to allow the production of meaningful climatic patterns while retaining as much resolution as possible. The resulting isolines were then smoothed by hand to remove meaningless boundary aberrations introduced by the interpolation algorithm.

Next, the degree-day and precipitation isolines were registered to each other to form climatic strata. Multidate Landsat full-frame imagery was then inspected in each state to determine gross patterns (soil moisture and crop development stage) of wheat

spectral response. Degree-day and/or precipitation intervals were then combined to form larger climatic strata if the gross pattern apparent on Landsat imagery suggested that larger regions were giving rise to similar wheat spectral response "on the average."

The resulting climatic strata were registered to the 1:1 000 000-scale map sheets on which the land use/soil association strata were overlaid.

Summary of Steps Used in Producing the Static Spectral Stratification

Step 1. A base date of Landsat imagery is selected from a period when soils and land use or cropping practices are most contrasted and most easily delineable.

Step 2. Soil associations are delineated on the base date color-infrared transparency, using available published soil data and interpretation of the Landsat imagery. The associations are then correlated across the CRD and ultimately across the entire area of interest.

Step 3. Land use or cropping practices are delineated on the base date color-infrared transparency, referencing the soil association delineations previously completed.

Step 4. The delineations from steps 2 and 3 are combined to produce one land-use/general-soil-type delineation.

Step 5. All remaining CRD's are processed in a similar manner. The resulting land-use/general-soil-type strata from each CRD are transferred to a 1:1 000 000-scale USGS base map and any boundary inconsistencies between CRD's are eliminated.

Step 6. Growing-season degree-day sums are calculated and plotted on the base coordinate system by the reporting meteorological station. Isolines are then determined by automatic interpolation and manual smoothing of the data.

Step 7. Growing-season precipitation is calculated and plotted on the base coordinate system by the reporting meteorological station. Isolines are then determined as in the case of degree-days.

Step 8. Climatic strata bounding isolines are selected by referencing the land use/soil association strata and several dates of Landsat imagery for consistent correlations of soil color-tone (soil moisture) and crop development stages with certain isolines.

Step 9. Climatic strata are registered to the 1:1 000 000-scale maps.

For a more complete description of the procedural steps, see references 3 and 5.

Figure 1 shows the climatic strata generated for the six western crop reporting districts in Kansas. Figure 2 shows both the climatic and the land use/soil association strata for North Dakota.

EVALUATION OF STATIC PARTITIONS RELATIVE TO VARIANCE CONTROL OBJECTIVES

Spectral Homogeneity

The static stratification developed as described in the preceding section was subsequently evaluated (ref. 7) in relation to its capability to group spectrally similar areas.

Approach.—The experimental procedure was composed of five basic parts. The first, *preprocessing*, standardized sample segments to a common Sun

elevation and haze condition. This was accomplished by implementation of XSTAR haze correction procedures (ref. 15) developed at the Environmental Research Institute of Michigan (ERIM). Preprocessing in this case provided a more stable measurement frame (Landsat or "Tasseled Cap" space) and thereby increased the ease with which real spectral differences could be identified and evaluated.

Each sample segment was partitioned according to land use/soil association strata as defined by the static stratification. Each segment partition was then individually *clustered* in a single-date mode by ISOCLAS (adapted from NASA Johnson Space Center (JSC)). The clustering process was limited to 10 iterations, a maximum band standard deviation of 3.2 Landsat counts within a cluster, and a distance between clusters of 3.2.

Resulting clusters by segment by UCB stratum were then stratified or *grouped* according to the percentage of wheat within the clusters. This was ac-

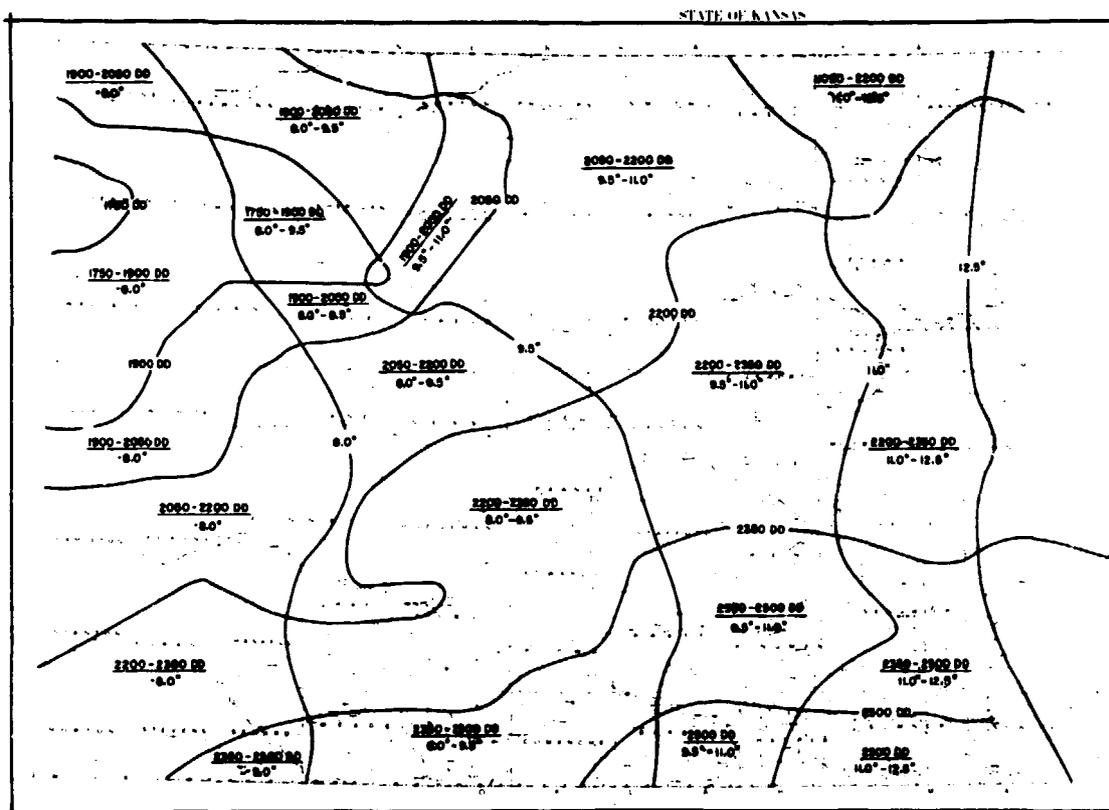


FIGURE 1.—Climatic strata for the state of Kansas used in the analysis of stratum spectral homogeneity. The ranges for long-term average growing-season degree-days and precipitation (reported in inches) are recorded in each stratum as numerator and denominator, respectively. When a range is not given, a range is assumed in the obvious plus or minus direction.

C-9

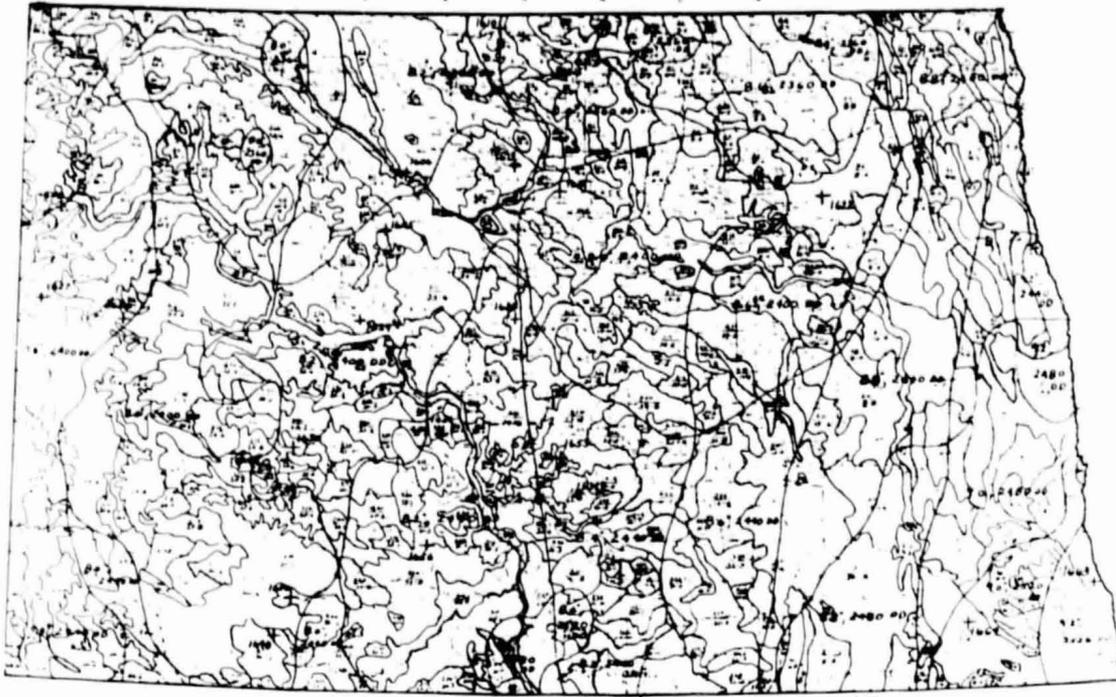


FIGURE 2.—Climatic and land use/soil association strata for the state of North Dakota used in the analysis of stratum spectral homogeneity. Long, gently curving lines delineate climatic strata. The finer resolution mosaic represents the land use/soil association stratification. Each land use/soil stratum is tagged by a fractional code: numerator for land use, denominator for soil type.

completed by comparing the cluster map with corresponding blind site ground data maps. To minimize the time required in this cluster grouping process, clusters were ordered (highest to lowest) by their $(2\times)$ band 7 to $(1\times)$ band 5 ratios of the cluster means on the Landsat pass date in question. This ratio was used as an indicator of vegetation and, depending on the date and state, of wheat versus other crop types. Using an interactive color television (TV) monitor system, clusters having the higher band 7 to 5 ratios were displayed and analyzed first, followed by clusters having lower ratios down to "a dry, stubble vegetation or soil line" (1.10). In this way, the multiple clusters occurring within fields could be "reconstructed" into field patterns and strongly correlated crop-type patterns on the initially blacked-out TV screen. The proportion of wheat in a given cluster could then be readily judged according to that cluster's distribution among fields. Four basic wheat percentage cluster groups were established: 75 to 100 percent, 50 to less than 75 percent, 25 to less than 50 percent, and 0 to less than 25 percent. Infor-

mation was also recorded regarding the cover-type makeup of the nonwheat portion of each cluster group.

A random sample of pixels was *labeled* from the cluster groups comprising 75- to 100-percent and 50- to less than 75-percent wheat in each stratum of each segment on each date. A random number generator, operated through the interactive color display system, minimized the time required for pixel selection. Ten to fifteen pixels in each of the two cluster groups were labeled as to crop type using the blind site ground data maps. This labeled pixel sample served three purposes: (1) it served as a check on the visual estimate of wheat percentage for each cluster group; (2) it provided the data employed in a Hotelling's T^2 test of wheat spectral difference between all possible pairs of land use and soil strata/climatic strata sampled; and (3) it provided the wheat pixel data used later in a spectral sensitivity analysis.

The fifth, and final, step in the analysis was to perform *pairwise spectral comparisons* of wheat signatures between all possible combinations of the

land use/soil strata and climatic strata sampled. These comparisons were made by applying Hotelling's T^2 test to the four-channel Landsat wheat signatures obtained from each pair of strata. Three sources of wheat signature data were evaluated separately: (1) the sample of pixels from the 75- to 100-percent wheat cluster group, (2) the sample of pixels from the 50- to less than 75-percent wheat cluster group, and (3) the combined sample of pixels from the 75- to 100-percent and 50- to less than 75-percent wheat cluster groups obtained in each stratum of each segment. Comparisons were limited to the same state and the same biostage.³

The result of the Hotelling test was a statistical significance or alpha value which gave the probability that the observed wheat signatures came from the same population. Alpha values of 0.05 (5 times in 100) or less were interpreted to mean that the null hypothesis (wheat signatures for the given pair of strata are similar) was to be rejected for the given pair of strata in question. By noting which pairs of strata did not cause rejection of the null hypothesis, sets of strata having statistically similar wheat signatures could be defined. Furthermore, it was assumed that nonrejection of the null hypothesis of spectral similarity implied a high probability of acceptable wheat classification performance. That is, if wheat spectral models (training statistics mean vector, covariance matrix) obtained from one portion of a set of spectrally similar strata were used to classify (using quadratic or linear discriminant functions) the remaining portion of that stratum set, an overall acceptable level of classification performance would be obtained. Acceptable as used here is defined in relation to the classification accuracy obtained by classifying on the basis of local stratum training statistics.

Data set.—In Kansas and North Dakota, two biophase periods were selected in which to apply the grouping and sensitivity analysis procedures just described. The first date in both states represented a wheat emergence condition. The second date corresponded approximately to a jointing or advanced jointing condition for the wheat crop. These time periods were selected on the basis of sensitivity analysis results reported in Hay et al. (ref. 6), which suggested that these stages were most difficult to

³For purposes of this analysis, a given biostage was considered to be extended over the several days (5-day period maximum) included in the data set.

characterize by static stratification variables. This analysis was therefore considered conservative in relation to the performance of the static stratification. Available sample segments were limited to those 1976 LACIE blind sites having ground data so as to minimize incorrect interpretation of results.

Results.—Within each state, all possible pairs of strata were tested for spectral similarity. For each stratum pair, Hotelling's T^2 test was applied separately to a sample of pixels from the 75- to 100-percent wheat cluster group (if this group was represented in both strata) and similarly to a sample of pixels from the 50- to less than 75-percent wheat cluster group. Pixel data from both cluster groups in each stratum were also pooled and tested against corresponding pooled data in other strata.

Results presented in table 1 for tests based on pooling cluster groups 1 and 2 (75- to 100-percent wheat and 50- to less than 75-percent wheat, respectively) show that *within a given climatic stratum*, the null hypothesis of spectral similarity between land use/soil strata was accepted 32 to 75 percent of the time. Significance levels used for rejection were $\alpha < 0.05$ and $\alpha < 0.01$. The acceptance rate between adjacent climatic strata (i.e., strata differing by one class of either long-term growing-season degree-days or precipitation (not both)) ran between 0 and 43 percent. Results for tests across climatic strata diagonally adjacent (differing by one class in both degree-days and precipitation) available from North Dakota gave acceptance rates of 50 percent (date 1) and 67 percent (date 2) for either significance level. In general, low rates of acceptance prevailed for land use/soil stratum pairs separated by more than one adjacent climatic stratum.

Based on the results, three basic patterns were evident for the two states and two dates involved.

1. The wheat signature population generally overlapped within a given climatic stratum. This pattern was more pronounced in the later as opposed to the earlier date.

2. Wheat signature overlap also occurred between horizontally, vertically, or diagonally adjacent climatic strata. The frequency of overlap was generally at a lower rate than within a given climatic stratum. It should be noted that no diagonally adjacent climatic stratum signature comparisons were available for Kansas. Given the somewhat larger areal extent of the climatic strata in Kansas relative to that in North Dakota (owing to the wider class values for degree-days and precipitation used in Kansas), the signature overlap rate between diagonally

TABLE I.—Hotelling's T^2 Test Results of Stratum Grouping Analysis

Cluster set	Frequency of null hypothesis acceptance, percent					
	Within same climatic stratum		Between adjacent climatic strata			
			Vertically or horizontally		Diagonally	
	$\alpha < 5\%$	$\alpha < 1\%$	$\alpha < 5\%$	$\alpha < 1\%$	$\alpha < 5\%$	$\alpha < 1\%$
^a Kansas date 1	32	42	19	43	— ^b	—
^c Kansas date 2	50	75	d_0	d_0	—	—
^e North Dakota date 1	33	50	74	24	50	50
^f North Dakota date 2	63	75	—	—	67	67

^a12 land use/soil strata distributed over 8 segments and 5 climatic strata.

^bDash indicates no stratum pairs available for test.

^c9 land use/soil strata distributed over 6 segments and 5 climatic strata.

^dBased on only 3 possible stratum matches available for test.

^e13 land use/soil strata distributed over 8 segments and 6 climatic strata.

^f9 land use/soil strata distributed over 5 segments and 5 climatic strata.

adjacent climatic strata in Kansas is expected to be lower than that obtained in North Dakota.

3. Wheat signatures rarely overlapped beyond an adjacent climatic stratum.

ERIM Evaluation of Static Spectral Stratification

Personnel of the ERIM also evaluated the UCB stratification as part of their work in the LACIE SR&T effort (ref. 16). Their approach was to perform all possible pairwise signature extensions among 23 segments (a total of 506 extensions) distributed across Kansas. Extensions were based on multitemporal Landsat data for biowindow 1 (Julian dates 291 to 90) and biowindow 2 (Julian dates 90 to 138) which had been Sun and haze corrected using the ERIM XSTAR algorithm. Resulting field mean classification accuracies (average percentage correct based on classification of mean spectral vectors by field) were then determined for each extension. These results in turn were summarized into within-stratum and between-stratum extensions and then evaluated using an analysis of variance.

Table II presents the outcome of the ERIM analysis of variance. The table shows that there was no significant difference in classification accuracy for within- versus between-stratum extensions in the case of land use strata. A similar comparison for soil association strata could not be made in that those strata divided the 23 segments into 23 different partitions. Significantly higher within- versus between-stratum classification performance occurred, however, in the cases of both degree-day and precipitation strata when evaluated separately. Finally, the best within- versus between-stratum classification performance (86.5 percent versus 66.6 percent) was obtained using the combined degree-day and precipitation climatic strata.

The ERIM analysis was based on an initial UCB stratification in which the individual climatic stratum areas were somewhat larger than those shown in figure 1. Nevertheless, the following interpretation of the results presented in table II remains valid. Namely, the climatic strata appear to be isolating general differences in crop development—differences related to the crop-development-driving nature of the climatic variables themselves. Wheat spectral differences did not appear to be highly correlated “on the average” with a relatively high-resolution land use static stratification.

ERIM Multisegment Classification Results

Recent work reported by Kauth and Richardson (ref. 4) also suggests the scale of spectral partitions should be on the order of that represented by

TABLE II.—XSTAR Field Mean Classification Results From ERIM Evaluation of UCB Static Spectral Stratification for Wheat in Kansas

Stratification	Within strata		Across strata		Significance of difference, α value
	No. extensions	Average XSTAR, percent correct	No. extensions	Average XSTAR, percent correct	
Land use	12	67.2	157	70.4	0.53
Degree-days	74	72.8	95	63.7	.01
Precipitation	41	82.4	128	66.2	.001
Degree-days plus precipitation	26	86.5	143	66.6	.001

climatic strata. They are in the process of developing a multitemporal, multisegment approach to crop proportion estimation in the LACIE context. The basic procedure is to (1) Sun and haze correct all Landsat passes for a given population of sample segments; (2) simultaneously cluster all segment data on the basis of both spectral (Tasseled Cap) and locational features; (3) cluster the spectral mean vector of each spatial "blob" (primarily field centers) that resulted from the initial clustering and assign each blob to one of the spectral strata that results from this second clustering; (4) inspect the population of blobs in each segment and select a minimum set of segments together containing adequate representation of all spectral strata in the entire population of segments; and (5) sample each spectral stratum to determine its crop composition and then expand these proportion estimates over all segments to obtain final crop proportion estimates and variances.

Application of this procedure to 17 LACIE Phase II segments spread across Kansas resulted in segment proportion estimates close to truth in a region covering the width of 3 to 4 climatic strata as shown in figure 1. Proportion estimates were significantly less accurate for most segments falling outside this region. This area of successful multisegment proportion estimation corresponds approximately to an area encompassed by a given climatic stratum plus the immediately adjoining climatic strata.

EVALUATION OF THE RELATIVE IMPORTANCE OF STATIC VERSUS SEASONAL AND PASS-SPECIFIC PARTITIONING VARIABLES IN SPECTRAL VARIABILITY

To gain insight into the underlying factors responsible for the results seen in the Hotelling's T^2 analysis and to obtain a measure of the relative importance of static and nonstatic spectral stratification variables, a spectral sensitivity analysis was performed on the data set described in the first portion of the preceding section.

Approach

The basic approach was to develop regression relationships relating spectral reflectance (dependent variable) to a set of static stratification, seasonal, and date-specific predictor variables. Matched spectral

response and predictor variable data were obtained for all pixels sampled in the analysis of stratum homogeneity.

The relative importance of each signature predictor variable listed in table III was expressed two ways. The first consisted of the percentage of total spectral variance (by band) explained by the addition of a given predictor variable to the regression equation. Variables were added in the same order as listed in table III, using a stepwise regression technique. The order—static, seasonal, date-specific—was chosen to most effectively identify the percentage of spectral variance accounted for by the static stratification variables before application of a signature extension algorithm. The R^2 (multiple correlation coefficient squared) increments, representing the percentage of variance added by each variable, were highly dependent on this ordering.

The second measure of signature predictor variable importance did not employ a prespecified order of entry into the regression. A forward selection regression procedure (as implemented by the Statistical Package for the Social Sciences) was used to order variables and tabulate the R^2 increments. Using this technique, the predictor variable having the highest simple correlation with the spectral band in question was entered into the regression first. The next variable entered was the one having the highest partial correlation with the spectral band after the effect of the first variable entered was removed from both the dependent and the independent variables. The third variable entered had the next highest partial correlation with the spectral response variable among all remaining predictor variables with the effects of the first two variables removed, and so on. Order of entry for a given variable among all bands for a given date provided the second measure of performance.

Results

Pixel data from both the 75- to 100-percent wheat and 50- to less than 75-percent wheat classes were pooled and regressed on corresponding static, seasonal, and Landsat pass-specific signature prediction variable data. Results for individual regressions on each Landsat band are presented in tables IV to VII. The tables are arranged first by state, then by date 1 or 2. Each table is then subdivided into results for ordered regression (part (a)) and regression without prior ordering (part (b)).

The most striking feature of the tables showing

TABLE III.—Signature Predictor Variables Used in the Kansas and North Dakota Wheat Spectral Sensitivity Analysis

Predictor variables	Measurement technique used for each field sampled
I. Static stratification variables (obtained from static strata map)	
A. Cultivated area percentage (CULTPCT)	Midpoint of cultivated area percentage range for the land use class covering the wheatfield.
B. Available soil water-holding capacity (AWC)	Average inches of water held per inch of soil at field capacity in the top 24 inches for the static strata soil association covering the wheatfield. These values are obtained from information available in county soil survey publications.
C. Long-term average growing-season degree-days (LTGSDD)	Midpoint of growing-season degree-day class covering the wheatfield. Degree-day classes obtained from 30-year average data by automatic and manual interpolation of ground meteorological station data for the period April through June in Kansas and June through August in North Dakota.
D. Long-term average growing-season precipitation (LTGSP)	Midpoint of growing-season precipitation class covering the wheatfield. Precipitation classes obtained from 30-year average data by automatic and manual interpolation of ground meteorological data for the period April through June in Kansas and June through August in North Dakota.
E. Long-term potential average available water in top 2 feet of soil ((24 × AWC) × LTGSP)	Multiply previously obtained values of AWC and LTGSP.
F. Long-term growing-season evapotranspiration (LTGSET)	Substitute 5-year average values for pan evaporation from nearest ground meteorological station making this measurement. Alternatively, empirical models using temperature and solar radiation may give satisfactory evapotranspiration estimates. Currently, only pan data are used here.
G. Long-term evapotranspiration stress on soil moisture reserve ((24 × AWC) × LTGSET)	Multiply previously obtained values AWC and LTGSET.
II. Seasonal variables (specific to 1975-76 growing season)	
A. Robertson biostage or bionumber—A numerical measure of crop development based on daily maximum and minimum temperature at selected meteorological stations in LACIE countries	Data obtained from Robertson biostage isoline maps reported for the Great Plains in the Weekly Meteorological Summaries produced in LACIE. The Robertson system divides the biological stages of wheat into seven development phases: (1) planting, (2) emergence, (3) jointing, (4) heading, (5) soft dough (turning greenish yellow to yellow), (6) hard dough, and (7) harvest. A Robertson number of 4.0 would mean that 50 percent of the crop is headed. Robertson numbers used in the sensitivity analysis were recorded to the nearest 0.1 of a development phase.
B. Growing-season degree-days accumulated to Landsat pass-date (SUMGSDD)	Calculated from temperature data supplied from nearest ground meteorological station having a physical/climatic setting most closely approximating the segment in which the wheatfield falls. Growing-season period: April through June (Kansas), May through August (North Dakota).
C. Growing-season precipitation accumulated to Landsat pass-date (SUMGSP)	Determined as in II.B. relative to precipitation data.

TABLE III.—Concluded

Predictor variables	Measurement technique used for each field sampled
D. Growing-season potential available soil water in top 2 feet of soil column ($24 \times \text{AWC} \times \text{SUMGSP}$)	Multiply previously obtained values for AWC and SUMGSP.
E. Growing-season evapotranspiration accumulated to Landsat pass-date (SUMGSET)	Substitute growing-season sum of pan evapotranspiration data from nearest station making this measurement.
F. Growing-season measure of available soil moisture in top 2 feet of soil [$(24 \times \text{AWC}) \times \text{SUMGSP}$] - SUMGSET = potential available soil water minus evapotranspiration loss)	Use values for AWC, SUMGSP, and SUMGSET obtained previously. Note that ground water table (a water source) is assumed not to be near the soil surface.
G. Average January 1976 temperature (JANTEMPT)	Determined from nearest meteorological station as in II.B.
H. Planting-season degree-days accumulated to Landsat pass-date (SUMPSDD)	Determined as in II.B. but for the period September through November (Kansas) and April (North Dakota).
I. Planting-season precipitation accumulated to Landsat pass-date (SUMPSP)	Determined as in II.B. relative to precipitation data in the period August through November (Kansas) and April (North Dakota).
III. Landsat date-specific variables	
A. Precipitation in the 4 days preceding Landsat pass-date (PPT4DA)	Determined as in II.B. relative to precipitation data.
B. $100 \times$ tangent of Landsat scan angle (SCANANG)	Departure measured along scan line of segment relative to an imaginary base line perpendicular to the scan direction and passing through the Landsat full-frame center point. Measurement based on full-frame center-point longitude and latitude coordinates given in Landsat Cumulative U.S. Standard Catalog and on sample segment coordinates supplied by JSC. The departure, reported in nautical miles, is defined as zero on the base line and increases positively to the east and negatively to the west. Then $\tan(\text{scan angle}) = \frac{\text{departure (n. mi.)}}{\text{mean satellite altitude (494 n. mi.)}}$
C. Landsat band 7 to band 5 ratio (RASR)—This ratio is one real-time indicator of biostage.	Obtain $(2 \times)$ band 7 to $(1 \times)$ band 5 ratio for the pixel.

results for Landsat bands with ordered regression is the significant importance of long-term growing-season degree-days and/or precipitation in accounting for the variation in spectral response. In this case, degree-days was the strongest for both Kansas dates and the second North Dakota date. Long-term growing-season precipitation accounted for the larger share of variance on the first North Dakota date. The other variable accounting for a substantial amount of spectral variance was cultivated area percentage.

This variable, obtained from the static stratification land use code, was significant in Landsat bands 6 and 7 on date 2 in both states and in all bands on date 1 in North Dakota. An evaluation of the cross variable correlation matrix suggests that the importance of the cultivated percentage was largely an artifact of the sample distribution in North Dakota. One other variable, available soil water-holding capacity (AWC), was expected to be significant in North Dakota. Unfortunately, AWC values could not be

TABLE IV.—Spectral Sensitivity Analysis of 454 Kansas Pixels Sampled January 18 to 20, 1976

<i>(a) Ordered regression</i>				
Variable	ΔR^2 value for Landsat band—			
	4	5	6	7
1. Cultivated percentage (CULTPCT)	0.01	—	0.02	—
2. Water-holding capacity (AWC)	.01	0.01	.01	0.01
3. Long-term growing-season degree-days (LTGSDD)	.17	.19	.43	.35
4. Long-term growing-season precipitation (LTGSP)	.04	.09	.04	.06
5. (24 × AWC) × LTGSP	.02	.03	.02	.04
6. Long-term growing-season evapotranspiration (LTGSET)	.01	—	—	.01
7. (24 × AWC) × LTGSET	.03	.03	.03	.04
8. Average January temperature (JANTEMPT)	.04	.06	.07	.05
9. Planting-season degree-days (PSDD)	—	—	—	—
10. Planting-season precipitation (PSP)	.02	.01	.01	.01
11. Scan angle	.02	.01	—	—
12. Band 7/band 5 ratio	.10	.18	.02	.17
Total R^2	.46	.61	.64	.73
Square root of mean square error	3.3	4.5	6.3	2.4
Total sum of squares	8.9×10^3	22.9×10^3	49.5×10^3	9.0×10^3

calculated for every land use/soil stratum; consequently, this variable (as well as composite variables using AWC) was omitted from the sensitivity analysis.

Although the reader is cautioned against putting much weight on the exact order of entry, the following observations were deemed significant in relation

TABLE IV.—Concluded

<i>(b) Regression without prior ordering</i>				
Variable	Order of entry for Landsat band— (a)			
	4	5	6	7
1. CULTPCT	4	4	8	7
2. AWC		3		
3. LTGSDD	5	5	6	6
4. LTGSP		11	5	4
5. (24 × AWC) × LTGSP	3	10	9	
6. LTGSET	8	7	3	2
7. (24 × AWC) × LTGSET	6	6	10	8
8. JANTEMPT	7	8	4	3
9. PSDD	1	1	1	1
	(0.24)	(0.28)	(0.51)	(0.43)
10. PSP	2	2	2	5
11. Scan angle	9	9	7	9

*Numbers in parentheses are ΔR^2 values

to the results of regression without prior ordering. In Kansas, variables entered first on date 1 were fall 1975 planting-season precipitation and degree-days. This was expected for a January 1976 pass date. Long-term growing-season degree-days, cultivated percentage, and scan angle were the first variables entered (i.e., having the highest correlation or partial correlation with the Landsat band values) into the regressions on the second date in Kansas. Precipitation in the 4 days preceding Landsat pass (4-day ppt.) and scan angle were entered consistently as the first and second variables for both dates in North Dakota. If present, 4-day precipitation can have an important impact on spectral signatures by wetting the soil or canopy surfaces. Other variables entered subsequently included either long-term or seasonal degree-day or precipitation variables.

**VIEWING THE EVIDENCE AS A WHOLE:
A CURRENT PERSPECTIVE ON
PARTITIONING**

Nature of the Spectral Surface

Using the results of the spectral sensitivity analysis as an aid, the following interpretation of results gained in the analysis of spectral homogeneity within and between strata is offered.

TABLE V.—Spectral Sensitivity Analysis of 162 Kansas Pixels Sampled May 4 to 7, 1976

(a) Ordered regression

Variable	ΔR^2 value for Landsat band—			
	4	5	6	7
1. CULTPCT	—	—	0.21	0.36
2. AWC	0.04	0.05	.01	.03
3. LTGSDD	.80	.78	.53	.30
4. LTGSP	.02	—	.01	.04
5. (24 × AWC) × LTGSP	.01	.02	.01	.01
Total R^2	.86	.84	.76	.74
Square root of mean square error	5.87	8.19	9.05	4.20
Total sum of squares	38.8×10^3	67.0×10^3	54.2×10^3	10.7×10^3

TABLE V.—Concluded

(b) Regression without prior ordering

Variable	Order of entry for Landsat band—			
	(a)			
	4	5	6	7
1. CULTPCT	2 (0.31)	2 (0.28)	5	6
2. AWC				5
3. LTGSDD	1 (0.43)	1 (0.51)	1 (0.71)	1 (0.65)
4. LTGSP	6			
5. (24 × AWC) × LTGSP	4	6	4	
6. LTGSET	5			
7. Robertson bionumber			6	2
8. JANTEMPT				4
9. Growing-season degree-days (GSDD)		4 5	2	
10. Scan angle	3	3	3	3

^aNumbers in parentheses are ΔR^2 values

1. The multivariate spectral surface for wheat appears to be relatively smooth, gradually changing over space. The spectral overlap encountered within and between climatic strata supports this notion.

2. Furthermore, the results of the sensitivity analysis indicate that this surface is strongly tied to degree-day and precipitation crop development variables. The spectral influences of long-term growing-season degree-days and, at times, long-term growing-season precipitation were found to be particularly significant. These, of course, were also the two variables used to define the climatic strata.

3. The sensitivity analysis also suggests that exceptions to interpretation 1 may be due largely to pass-specific precipitation differences, their interaction with soil type reflectances, and scan angle differences. Land use may also have an impact in situations where it is strongly correlated with soil type or with particular agricultural practices affecting plant canopy reflectance (e.g., irrigation, field size and shape).

4. Examination of the spectral surface on individual dates suggests that its average gradient changes throughout the crop year for wheat. That is, the region of spectral overlap giving adequate classification or proportion estimates will vary in size if analysis is performed on a single-date basis. In relative terms, this region may be of moderate size early in the crop year (soil background reflectance important), largest just before heading (accumulated weather/climatic influences dominant), and smallest during heading and ripening (local soil moisture/depth, crop practice influences apparent). A limited sensitivity analysis on several dates in Kansas for the 1975-76 crop year (ref. 6) suggested this pattern.

5. The ERIM multisegment results suggest that the effect of these changes in the shape of the spectral surface on classification performance may be controlled to some extent using multi-date classification. Further analysis is required.

Relative Role of Static Versus Real-Time Partitioning Variables in the Context of a Multitemporal, Multisegment Classification Approach to Signature Extension

At present, it appears that a multisegment clustering and classification approach similar to that described earlier (ref. 4) provides the most workable solution to the signature extension problem. The

TABLE VI.—Spectral Sensitivity Analysis of 192 North Dakota Pixels Sampled May 24 to 28, 1976

(a) Ordered regression

Variable	ΔR^2 value for Landsat band—			
	4	5	6	7
1. CULTPCT	0.27	0.29	0.23	0.17
2. LTGSDD	.02	.02	.01	.04
3. LTGSP	.31	.35	.13	.06
4. LTGSET	.02	.01	.03	.02
5. Robertson bionumber	.01	.01	.01	.01
6. GSDD	.01	.01	—	—
7. Growing-season precipitation (GSP)	.02	.01	.02	.03
8. Sum growing- season :vapo- transpiration (SUMGSET)	—	—	.03	.04
9. JANTEMPT	—	—	—	—
Total R^2	.61	.70	.46	.37
Square root of mean square error	3.7	5.0	7.5	3.7
Total sum of squares	6.3×10^3	15.5×10^3	18.7×10^3	3.8×10^3

following observations are addressed to the application of spectral partitioning in that context.

1. Multisegment clustering and classification should be possible using climatic strata, or, more generally, distance on a climate-related spectral surface, as a guide to segment grouping. In other words, it should be possible to use spectral training data (cover-type-specific Landsat or Tasseled Cap band means, variances, and covariances) obtained from a specially selected sample of LACIE segments to classify with acceptable accuracy the entire set of LACIE segments falling within climatic partitions. A cost savings over training and classifying each segment separately should result.

Spatial distribution of sample segments need not be dictated by spectral strata boundaries, but sample segments can and should be allocated among area, yield, or production strata to control sampling error in standard fashion. Training gains (i.e., number of segments used to develop spectral models for

TABLE VI.—Concluded

(b) Regression without prior ordering

Variable	Order of entry for Landsat band— (a)			
	4	5	6	7
1. CULTPCT	4	7	8	8
2. LTGSDD	5	4	3	3
3. LTGSP	8		5	5
4. Robertson bionumber		3		
5. GSDD	6	5		
6. GSP	3		6	6
7. SUMGSET	7	8		
8. JANTEMPT			4	4
9. PSDD				
10. PSP		6		7
11. 4-day precipi- tation	1	1	1	1
12. Scan angle	2	2	2	2

^aNumbers in parentheses are ΔR^2 values

classification versus total number of segments classified) should be largest with higher sampling intensities (number of sample segments classified per unit area). In this regard, it may be cost-effective to increase sample sizes somewhat over those required for normal estimate precision control in order to take advantage of a larger training gain. This action would ensure a higher probability of achieving assigned regional precision objectives and also enable the production of local crop proportion estimates of higher precision.

2. The center of any given spectral partition can be conveniently chosen to align with the center of a specific population of sample segments occupying a given area used for estimate summary (e.g., some reporting or aggregation unit). A method for defining the distance from the spectral partition center to the partition "boundary" remains a significant research question. Such a distance metric will undoubtedly be a function of crop development and in turn the growth-driving environmental variables. For the present, spectral partition size must be roughly defined in terms of climatic strata or combinations of climatic strata based (a) on degree-day and precipitation differences as they potentially affect growth and (b) on actual multisegment classification and proportion estimation performance.

TABLE VII.—Spectral Sensitivity Analysis of 157 North Dakota Pixels Sampled June 30 to July 2, 1976

<i>(a) Ordered regression</i>				
Variable	ΔR^2 value for Landsat band—			
	4	5	6	7
1. CULTPCT	0.04	0.03	0.10	0.10
2. LTGSDD	.02	—	.22	.16
3. LTGSP	.02	.01	—	.01
4. LTGSET	.07	.07	.03	.06
5. Robertson bionumber	.01	.03	.05	.07
Total R^2	.15	.13	.40	.40
Square root of mean square error	2.2	2.8	6.7	3.9
Total sum of squares	0.8×10^3	1.4×10^3	11.1×10^3	3.9×10^3

3. To achieve the successful multisegment classification described in observation 1 will, in all probability, require Sun angle and haze correction to a common standard. Although the XSTAR algorithm suffices for this purpose in the case of classification of Tasseled Cap bands (ERIM results), the authors have found that some question remains as to the proper algorithm to apply when classification is based on Landsat band combinations.

4. Although pass-specific precipitation, soil reflectance, and scan angle may generate spectral outliers, these should not generally pose significant problems to multisegment clustering within climatic strata. This is not to say, however, that recognition segments (segments into which signature is extended from others) having no adequate spectral analogs will not occur. Undoubtedly, they will. But, within many biostages or combinations of biostages, multisegment classification as described in observation 1 should be possible with at least some portion of the population of sample segments at hand. Further technical developments in scan angle correction and flagging of soil type conditions, etc., in which outliers will occur should serve to maximize successful use of the multisegment approach to signature extension within climatic strata.

TABLE VII.—Concluded

<i>(b) Regression without prior ordering</i>				
Variable	Order of entry for Landsat band— <i>(a)</i>			
	4	5	6	7
1. CULTPCT	4	4	4	4
2. LTGSDD	5			
3. LTGSP		3	5	
4. LTGSET		(0.11)		5
5. GSDD	3			
	(0.13)			
6. GSP			3	3
			(0.13)	(0.09)
7. 4-day precipitation	1	1	1	1
			(0.13)	(0.07)
8. Scan angle	2	2	2	2
			(0.12)	(0.12)

*Numbers in parentheses are ΔR^2 values

The general question of crop and environment interaction with spectral reflectance remains an area of major research concern. A complete, robust solution to the spectral partitioning problem must await the results of further work on signature prediction and modeling.

REFERENCES

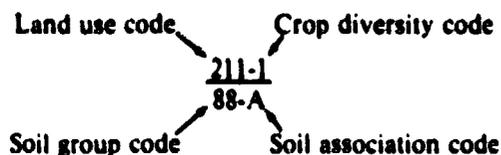
1. Lambeck, P. F.; and Rice, D. P.: Signature Extension Using Transformed Cluster Statistics and Related Techniques. Technical Report 109600-70-F, ERIM (Ann Arbor, Mich.), May 1976.
2. Landgrebe, D. A.; Bauer, M. E.; and Davis, B. J.: Development of Signature Extension Strata From Clustering Techniques. In Final Report, NASA Contract NAS 9-14016 (June 1, 1975, to May 31, 1976), Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Ind., 1976.
3. Hay, C. M.; and Thomas, R. W.: Application of Photointerpretative Techniques to Wheat Identification, Signature Extension, and Sampling Strategy. NASA Contract NAS 9-14565 (Principal Investigator R. N. Colwell). Space Sciences Laboratory, Series 17, Issue 33, UCB, May 1976.

4. Kauth, R. J.; and Richardson, W.: Procedure B: A Multisegment Training Selection and Proportion Estimation Procedure for Processing Landsat Agricultural Data. Technical Report 122700-31-F, ERIM (Ann Arbor, Mich.), Nov. 1977.
5. Hay, C. M.; and Thomas, R. W.: Development of Techniques for Producing Static Strata Maps and Development of Photo-interpretation Methods Based on Multitemporal Landsat Data. Six-Month Summary Progress Report, NASA Contract NAS 9-14565 (Principal Investigator R. N. Colwell). Space Sciences Laboratory, Series 18, Issue 2, UCB, May 1976.
6. Hay, C. M.; Thomas, R. W.; and Benson, A. S.: Development of Techniques for Producing Static Strata Maps and Development of Photointerpretation Methods Based on Multitemporal Landsat Data. Quarterly Progress Report, NASA Contract NAS 9-14565 (Principal Investigator R. N. Colwell). Space Sciences Laboratory, Series 18, Issue 60, UCB, May 1977.
7. Hay, C. M.; Thomas, R. W.; and Benson, A. S.: Development of Techniques for Producing Static Strata Maps and Development of Photointerpretation Methods Based on Multitemporal Landsat Data. Quarterly Progress Report, NASA Contract NAS 9-14565 (Principal Investigator R. N. Colwell). Space Sciences Laboratory, Series 18, issue 75, UCB, Aug. 1977.
8. Montgomery, O. L.; and Baumgardner, M. F.: The Effects of the Physical and Chemical Properties of Soil on the Spectral Reflectance of Soils. LARS Information Note 112674, Purdue University, West Lafayette, Ind., 1974.
9. Stockton, J. G.; Bauer, M. E.; Blair, B. O.; and Baumgardner, M. R.: The Use of ERTS-1 Multispectral Imagery for Crop Identification in a Semi-Arid Climate. LARS Information Note 040775, Purdue University, West Lafayette, Ind., 1971.
10. Pascale, A. J.; and Damarlo, E. A.: Agroclimatic Wheat Crop Types in the World. In Biometeorology, Proceedings of the Second International Bioclimatological Congress (London, Sept. 1960), Pergamon Press, 1962.
11. Peterson, R. F.: Wheat; Botany, Cultivation, and Utilization. Interscience Publishers, 1965.
12. Klages, K. H. W.: Ecological Crop Geography. The Macmillan Co., 1942.
13. Nuttonson, M. Y.: A Comparative Study of Lower and Upper Limits of Temperature in Measuring the Variability of Day-Degree Summations of Wheat, Barley, and Rye. American Institute of Crop Ecology (Washington, D.C.), 1956.
14. Anderson, J. R.; Hardy, E. H.; and Roach, J. T.: A Land-Use Classification System for Use With Remote-Sensor Data. Geological Survey Circular 671. U.S. Dept. of the Interior, 1972.
15. Lambeck, P. F.: Implementation of the XSTAR Haze Correction Algorithm and Associated Steps for Landsat Data. ERIM Memo IS-PFL-1272, Mar. 18, 1977, and revised by ERIM Memo IS-PFL-1916, Nov. 1, 1977.
16. Cicone, R. C.; Stinson, J. L.; and Balon, R. J.: An Evaluation of the Signature Extension Approach to Large Crop Inventories Utilizing Space Image Data. Technical Report 122700-33-F, ERIM (Ann Arbor, Mich.), Nov. 1977.

Appendix

Legend Code for Signature Extension Land Use/Soil Association Strata

The land use/soil association strata are annotated with a fractional code. The numerator is the land-use/crop-diversity designation and the denominator is the soil-group/soil-association designation.



LAND USE CLASSIFICATION CODE

- 100 Urban and built-up land
- 110 Residential, commercial, industrial, institutional, transportation, mixed, open, and other
- 120 Strip and clustered settlements
- 130 Resorts
- 200 Agricultural land (more than 15 percent of area is cultivated)
 - 211 Cropland and intensive pasture (more than 75 percent of the area is cultivated)
 - 212 Cropland and intensive pasture (more

than 50 percent but less than 75 percent of the area is cultivated)

- 213 Orchard and vineyards
- 220 Extensive agriculture (less than 50 percent of the area is cultivated)
- 300 Rangeland (less than 15 percent of the area is cultivated)
 - 310 Grassland range
 - 320 Woodland range
 - 330 Chaparral range
 - 340 Desert shrub range
- 400 Forest land
- 500 Water
- 600 Nonforested wetland
- 700 Barren land
- 800 Tundra
- 900 Permanent snow and icefields

CROP DIVERSITY CODE

- 1 Relatively high crop diversity
- 2 Medium crop diversity
- 3 Low crop diversity

Methods of Extending Crop Signatures From One Area to Another

T. C. Minter^a

INTRODUCTION

Background

The Large Area Crop Inventory Experiment (LACIE) is an attempt to establish the feasibility of inventorying the production of wheat on a world-wide basis by using Landsat data. A basic 5- by 6-nautical-mile sampling unit, called a segment, is employed. A wheat area estimate for a region is made by totaling estimates for subregions, where the wheat area estimate for a subregion is based on an estimate for each of the segments in the subregion. Wheat area estimates are made by extracting from a segment training data for wheat and nonwheat and then using the statistics for these training data to classify the segment pixel by pixel. This local training and classification procedure requires that training data in each segment be labeled by an analyst interpreter (AI).

Signature extension is an attempt to reduce the total amount of analyst work required in making a wheat area estimate for a region. The approach is to extract training statistics from one segment and use these statistics or signatures to classify several other segments; hence, the term "signature extension."

This paper summarizes much of the work accomplished by LACIE on signature extension in 1975 and 1976. Several significant advances in haze correction procedures are documented in this paper. All the work on signature extension described in this paper was accomplished by the Earth Observations Division of the NASA Johnson Space Center, Lockheed Electronics Company, IBM, and the supporting research institutions. All material presented

in this paper has been extracted from documents published by these groups.

It will be obvious from results presented in this paper that signature extension is a very difficult problem. The approach to signature extension described herein represents LACIE's understanding of the problem and its possible solution in 1975 and 1976. The lack of success of this signature extension approach led to the development of the multisegment training approach to signature extension described in the paper by Kauth and Richardson entitled "Signature Extension Methods in Crop Area Estimation."

Objective of Signature Extension

The objective of signature extension is to increase the spatial-temporal range over which a set of training statistics can be used to classify Landsat data without significant loss of recognition accuracy. Because of variations in measurement conditions when Landsat data are collected, the computer must be retrained on a regular basis. The crop signatures observed by Landsat are not constant in either time or space. The need to retrain the computer requires labeling of new examples of wheat and nonwheat, a process that is both costly and time consuming. A viable signature extension technology for LACIE would provide more timely and cost-effective classification over extensive land areas.

APPROACH USED IN SIGNATURE EXTENSION

The proposed approach to signature extension was to use the training samples developed by an

^aLockheed Electronics Company, Houston, Texas.

analyst in one segment (the training segment, or TSEG) to structure a classifier, which could then be used to identify wheat and estimate wheat proportions in the TSEG and in several nearby segments (called recognition segments, or RSEG's). It was acknowledged that the TSEG wheat and nonwheat signatures might not be representative of signatures in the RSEG. Any difference between TSEG and RSEG crop spectral signatures would result in poor classifier performance in the RSEG's, as measured in terms of probability of correct classification (PCC) and a higher variance of the wheat proportion estimate. The existence of a number of factors that might cause variations in the TSEG and RSEG crop signatures was postulated. These factors were divided into two main categories: dynamic factors and static sources of differences. Some of these are listed in table I.

To facilitate a successful extension of signatures from the TSEG to an RSEG, it was proposed that these sources of signature variation be accounted for and removed from the TSEG signatures before classifying the RSEG. Static sources of variation were to be removed by partitioning segments; i.e., by grouping together those segments from a large area which had similar characteristics of the sort listed as type B in table I. For more on partitioning, see the paper by Thomas et al. entitled "Development of Partitioning as an Aid to Spectral Signature Extension."

Differences in signatures from the TSEG and an RSEG which were caused by dynamic factors (i.e., atmospheric haze and Sun angle changes) were to be removed by mathematically modeling these effects and correcting the TSEG signatures accordingly on a pairwise basis. It is well known (see the paper by Lambeck and Potter entitled "Compensation for Atmospheric Effects in Landsat Data" and refs. 1 and 2) that signature changes caused by differences in atmospheric haze level and Sun angle can be mathematically modeled by an affine transformation of the form

$$y_k = a_k x_k + b_k \quad (1)$$

where x_k = a multispectral scanner measurement in the k th spectral band from the TSEG

y_k = the transformed equivalent of x_k in the RSEG

a_k = a multiplicative factor for the k th spectral band which is a function of the differences between TSEG and RSEG haze levels and Sun angles

b_k = an additive term for the k th spectral band which is a function of the differences between TSEG and RSEG haze levels and Sun angles

A number of algorithms were developed for estimating the coefficients of this affine transformation (eq. (1)) for Landsat multispectral scanner data. An important exception was the University of Houston Maximum Likelihood Estimation algorithm, which estimated the RSEG statistics directly. More will be said later about this approach. An important constraint imposed on these algorithms was that a_k and b_k had to be estimated without the aid of any training data in the recognition segment. If training data were made available in the RSEG, the signatures developed from these data would be preferable to corrected TSEG signatures in classifying the RSEG. But having an analyst develop these training data in the RSEG would defeat the purpose of signature extension.

Three approaches were taken in estimating the corrections to be applied to the TSEG signatures. These were

1. Cluster-matching algorithms
2. Distribution-matching algorithms
3. Atmospheric models

The cluster-matching algorithms used a clustering algorithm to identify the inherent spectral classes in the TSEG and RSEG data. Clusters corresponding to the same crop in the TSEG and the RSEG were then matched using various procedures. The coefficients a_k and b_k of the affine transformation (eq. (1)) could then be readily obtained. The algorithms using this approach were the Rank Order Optimal Signature Transformation Estimation Routine (ROOSTER) and the Optimal Signature Correction Algorithmic Routine (OSCAR). These algorithms are discussed in greater detail in a succeeding section.

The distribution-matching algorithms used maximum-likelihood estimation procedures to correct for differences between TSEG and RSEG probability density functions (pdf's). The University of Houston Maximum Likelihood Estimation (UHMLE) algorithm attempted to correct for these pdf differences without any assumptions as to the form the correction might take. The Maximum Likelihood Estimation of Signature Transformation (MLEST) algorithm assumed that differences between TSEG

and RSEG pdf's could be accounted for by an affine transformation of the form shown in equation (1). MLEST used a maximum-likelihood iteration approach to select a set of coefficients a_k, b_k that matched the TSEG and RSEG pdf's as closely as possible. MLEST and UHMLE are described in detail in succeeding sections.

The Atmospheric Correction (ATCOR) program employs an atmospheric model to predict the effect on Landsat data of changes in haze and Sun angle. Indicators of the haze level are derived from the data and processed through an atmospheric model to estimate the coefficients a_k, b_k of the affine transform (eq. (1)). This algorithm is discussed in another paper (Lambeck and Potter's) and will not be described in detail here.

Several experiments were conducted to evaluate the approaches listed previously. These experiments and their results are described in the fourth and fifth sections. Conclusions drawn from these experiments are presented in the final section.

A DESCRIPTION OF THE SIGNATURE CORRECTION PROCEDURES

In this section, several of the signature extension algorithms tested are described in detail. These include ROOSTER and its modification, OSCAR and its modification, MLEST, and UHMLE. ATCOR is described in the paper by Lambeck and Potter and will not be discussed here.

The following notation is used in the mathematical description of the algorithms discussed in this paper.

$\{x\}$	= set of samples from the training segment
$\{y\}$	= set of samples from the recognition segment
M_T	= number of subclasses in the training segment
M_R	= number of subclasses in the recognition segment
p	= dimensionality of samples
$\mu_i, i = 1, 2, \dots, M_T$	= subclass means in the training segment
$\Sigma_i, i = 1, 2, \dots, M_T$	= subclass covariance matrices in the training segment

$q_i, i = 1, 2, \dots, M_T$ = a priori probabilities of the training segment subclasses

$\mu'_i, i = 1, 2, \dots, M_R$ = subclass means in the recognition segment

$\Sigma'_i, i = 1, 2, \dots, M_R$ = subclass covariance matrices in the recognition segment

$q'_i, i = 1, 2, \dots, M_R$ = a priori probabilities of the recognition segment subclasses

Cluster-Matching Algorithms

Introduction.—In this section, two cluster-matching algorithms and their modifications are discussed. The basic theory is presented, then the algorithms are described in detail. The algorithms described are ROOSTER and OSCAR.

The theory of cluster-matching algorithms.—Given that an affine signature transformation is to be used to compensate for multiplicative and additive differences between two scenes, the values of the coefficients a_k and b_k for equation (1) must be estimated. For this purpose, one needs some effective way of comparing the data from the two scenes. One method for accomplishing this comparison is to compare cluster statistics for the scenes.

Consider two scenes where the same ground classes are present in the same proportions but the data values in scene 2 differ from those in scene 1 by a transformation of the form shown in equation (1); i.e.,

$$y_k = a_k x_k + b_k \quad (1)$$

In this case, the probability density function of the data in each scene should look the same; i.e., the same number of modes should be present in each scene, each with the same frequency, but the location of the modes will differ by the scale factor a_k and the displacement b_k in the k th Landsat band. Each scene is clustered separately to find these modes. Since the same classes are present in both scenes, the corresponding modes can be paired up. The parameters a_k and b_k can be estimated from the locations of these paired modes. In the algorithms reported on here, each mode is described by the mean of a cluster and a_k and b_k are estimated by the method of least

squares. In practice, several difficulties arise: first, the ground class labels for the modes of the data in the RSEG are unknown; second, the frequencies of the same ground class in the two scenes will usually differ; and third, some ground classes present in one scene may not be present in the other.

The first basic cluster-matching algorithm, called MASC (for Multiplicative and Additive Signature Correction) (ref. 2), was developed at the Environmental Research Institute of Michigan (ERIM) to test the cluster regression approach to determining the a_k and b_k coefficients. Although this algorithm achieved some occasional successes at signature extension, it did not include an adequate means for selecting only valid cluster pairs from the many potential cluster pairs.

The difficulty involved in identifying valid cluster matches between a pair of scenes may perhaps be partly appreciated by considering the problem of matching a set of 10 training scene clusters with a set of 10 recognition scene clusters. If one tries to examine all possible sets of 10 cluster pairs to find which is best, one finds that there are $10! (3\ 628\ 800)$ sets of pairs to be considered, assuming that there are no multiple pairings with the same cluster. If one happens to guess that only 8 valid pairs are possible, then the number of sets of pairs to be considered increases by a factor of $45/2$, to more than 80 million; i.e.,

$$\left[\binom{10}{8} \binom{10}{8} 8! \right]$$

Obviously, there are two basic difficulties to be dealt with in finding the valid cluster pairs from which to derive the required signature transformation. The first is to reduce to a practical number the sets of cluster pairs to be examined, and the second is to determine which among the remaining candidate sets of cluster pairs are most likely to be valid. ROOSTER and OSCAR take varying approaches to the solution of these two problems.

ROOSTER.—The Rank Order Optimal Transformation Estimation Routine (refs. 3 and 4) selects pairs based on channel ranks. Channel ranks have the important property of being invariant with respect to the affine haze/Sun angle correction (eq. (1)). Specifically, if

$$\mu_{ik} > \mu_{jk} \quad (2)$$

then

$$a_k \mu_{ik} + b_k > a_k \mu_{jk} + b_k \quad (3)$$

provided $a_k > 0$. Hence, corresponding RSEG clusters will manifest the same order relationship. Using this basic idea, the steps in the algorithm are as follows.

Step 1.—Cluster each segment. Let μ_{ik} be the mean of the i th cluster in the k th channel from the TSEG. Let μ'_{jk} be the mean of the j th cluster in the k th channel in the RSEG. Let M_T be the number of clusters in the TSEG and M_R be the number of clusters in the RSEG. The clusters are unlabeled; therefore, one does not know which, if any, of the μ_{ik} means corresponds to μ'_{jk} .

Step 2.—For all clusters, $i = 1, \dots, M_T$, compute the pseudorank vector \mathbf{u}_{ik} for the k th channel of the TSEG.

$$\mathbf{u}_{ik} = \frac{1}{M_T - 1} \sum_{\substack{w=1 \\ w \neq i}}^{M_T} G(\mu_{ik} - \mu_{wk}; t) \quad (4)$$

where $G(\mu_{ik} - \mu_{wk}; t)$ is defined as follows:

$$G(\mu_{ik} - \mu_{wk}; t) = \begin{cases} 1 & \text{if } \mu_{ik} - \mu_{wk} > t \\ 0 & \text{if } \mu_{ik} - \mu_{wk} < -t \end{cases} \quad (5)$$

and

$$G(\mu_{ik} - \mu_{wk}; t) = (\mu_{ik} - \mu_{wk} + t) / 2t \quad (6)$$

if

$$-t \leq \mu_{ik} - \mu_{wk} \leq t$$

The parameter t , an adjustment factor for determining pseudorank, is specified by the user and is intended to be small but positive. If $t = 0$, then \mathbf{u}_{ik} is the vector of ranks multiplied by a constant.

Step 3.—For all clusters, $j = 1, 2, \dots, M_R$, compute the pseudorank vector \mathbf{v}_{jk} for the k th channel of the RSEG.

$$\mathbf{v}_{jk} = \frac{1}{M_R - 1} \sum_{\substack{w=1 \\ w \neq j}}^{M_R} G(\mu'_{jk} - \mu'_{wk}; t) \quad (7)$$

where $G(\mu'_{jk} - \mu'_{wk}; t)$ is defined in the same manner as $G(\mu_{jk} - \mu_{wk}; t)$ in step 2.

Step 4.—Compute a measure of the similarity of ranking c_{ij} of the i th TSEG cluster to the j th RSEG cluster over $k = 1, 2, \dots, p$ channels.

$$c_{ij} = \sum_{k=1}^p |\mathbf{u}_{ik} - \mathbf{v}_{jk}|^Q \quad (8)$$

where Q is the power for the fit criterion and is a user-specified parameter; currently, $Q = 1$.

Step 5.—Rank c_{ij} in order of ascending values and select s of these where $s \leq \min(M_T, M_R)$. Pass over any c_{ij} for which i or j has been previously selected. Relabel corresponding pairs of cluster means $(\mu'_1, \mu_1), (\mu'_2, \mu_2), \dots, (\mu'_s, \mu_s)$. The user-specified parameter s is the maximum number of cluster pairs to be used in the regression for the coefficients of the affine transformation (eq. (1)).

Step 6.—Let $\{(H_w)\}$ be an r element subset of the set $\{1, 2, \dots, s\}$, where r is the minimum number of pairs to be used in the regression for the affine transformation. For each w and k , calculate α_{wk} and β_{wk} to minimize

$$l_{wk} = \sum_{i \in H_w} [\mu'_{ik} - (\alpha_{wk}\mu_{ik} + \beta_{wk})]^2 \quad (9)$$

The quantities α_{wk} and β_{wk} are the intercept and the slope for a simple least squares regression.

Step 7.—Select the w so that

$$w = \min \sum_{k=1}^p l_{wk}$$

The coefficients of the affine transformation (eq. (1)) are then

$$a_k = \alpha_{wk}, b_k = \beta_{wk} \quad (10)$$

In analyzing the procedure, one can see that if all (or nearly all) classes are present in both segments, the pseudorank vectors \mathbf{u}_i and \mathbf{v}_j should be nearly the same for both segments, and the proper matches can be obtained. Two possible difficulties are (1) the cluster means are subject to random variation that can cause rank reversals from one segment to the other and (2) some clusters may be found in only one segment.

The first difficulty is met by using the G functions to reduce the effect of small random variations and by forming pairings based on all channel ranks and thereby gaining the advantage of cumulative evidence. The problem of unmatched clusters is mitigated by the use of ranks. One unmatched cluster will cause a difference of at most $1/(M_R - 1)$ in a pseudorank value. Therefore, even in the presence of an unmatched cluster, the correspondence of matched vectors will be sufficient to produce proper pairings. Numerous unmatched clusters may present problems. There is, however, a reasonable chance that rank errors caused by unmatched pairs will average out to such a degree that mostly valid matches will be obtained. Moreover, the algorithm will deal with unmatched clusters more effectively than the alternative that assumes all clusters are matched.

The rationale for using regression to estimate a_k and b_k (eq. (1)) is obvious, if one can be sure that all matches are genuine. The proposed algorithm can work even if some matches are spurious. The use of spurious matches in selecting α and β will tend to produce a poor fit to the regression line. Thus, the proposed algorithm will tend to select valid pairs in determining the actual α and β values used to estimate a_k and b_k .

Modified ROOSTER.—Kauth and Thomas defined a set of four orthogonal physically interpretable coordinate axes that amount to a rotation of Landsat data (ref. 5). These directions correspond to (1) increasing soil brightness, (2) increasing vegetation greenness, (3) increasing vegetation yellowness, and (4) a direction called "non-such." In modified ROOSTER (ref. 6), the cluster mean vectors were projected onto the soil brightness axis

before applying the seven steps of the ROOSTER procedure described previously. This procedure had the effect of causing the clusters to be ranked on the basis of their brightness.

OSCAR.—The Optimal Signature Correction Algorithmic Routine (ref. 7) uses a "goodness of fit" function to evaluate candidate transformations. Ideally, an (a_k, b_k) transformation (eq. (1)) should transform most training segment cluster mean vectors so that each nearly equals one of the recognition segment cluster mean vectors. Therefore, a reasonable measure of goodness is how close each recognition segment cluster mean vector is to the closest transformed training segment mean vector.

In the algorithm ROOSTER, the goodness function used is the sum of squares of the differences for the best fitting cluster mean vectors. The ROOSTER algorithm deletes bad fits from consideration since it is unreasonable to expect all the signatures in one segment to have matches in the other segment. The deletion rule (i.e., the selection of a value for s) is somewhat arbitrary. The OSCAR algorithm sidesteps this deletion problem by weighting all potential matches using a negative exponential of their goodness of fit. The distance used to measure this goodness of fit also makes some use of the covariance structure of the clusters being matched, as in the Bhattacharyya distance. The rationale is developed as follows. First, define the following:

1. μ_i, μ_j' = training and recognition segment subclass/cluster mean vectors, respectively

2. Σ_i, Σ_j' = training and recognition segment cluster covariance matrices, respectively

An assumption is made that the recognition segment picture elements (pixels) y are multivariate normally distributed. Using the Mahalanobis distance as a measure, the average distance from a point in this distribution to the training segment cluster mean vector is determined by

$$D(y, \mu_i, \Sigma_i) = (y - \mu_i)^T \Sigma_i^{-1} (y - \mu_i) \quad (11)$$

where $D(y, \mu_i, \Sigma_i)$ is the Mahalanobis distance, and y is a point from the i th recognition segment subclass/cluster that is normally distributed with mean μ_j' and covariance Σ_j' .

The average value of this distance is

$$E[D(y, \mu_i, \Sigma_i)] = \int \dots \int D(y, \mu_i, \Sigma_i) N(y; \mu_j', \Sigma_j') dy \quad (12)$$

where $N(\cdot)$ denotes a normal density. Now, there exists a nonsingular matrix P such that

$$P^T \Sigma_i^{-1} P = H' \quad (13)$$

$$P^T \Sigma_j^{-1} P = I' \quad (14)$$

where H is a positive definite diagonal matrix. Next, a change of variables is made in equation (12) where the new variable w is defined to be

$$w = P^{-1}(y - \mu_j')$$

and then y is defined to be

$$y = \mu_j' + Pw \quad (15)$$

The Jacobian of y with respect to w is $J(y) = |P|$. From equation (14), it is evident that

$$|P| = \left(|\Sigma_j^{-1}| \right)^{\frac{1}{2}} \quad (16)$$

By substituting equations (13) and (14) in equation (12), one obtains

$$E[D(y, \mu_i, \Sigma_i)] = (2\pi)^{\frac{p}{2}} \int \dots \int (\mu_j' + Pw - \mu_i)^T \Sigma_i^{-1} \cdot (\mu_j' + Pw - \mu_i) e^{-\frac{1}{2} w^T w} dw \quad (17)$$

$$E \left[D(y, \mu_j, \Sigma_i) \right] = (\mu_j' - \mu_i)^T \Sigma_i^{-1} (\mu_j' - \mu_i) + 0 + (2\pi)^{-\frac{p}{2}} \cdot \int \dots \int w^T H w e^{-\frac{1}{2} w^T w} dw \quad (18)$$

$$E \left[D(y, \mu_j, \Sigma_i) \right] = (\mu_j' - \mu_i)^T \Sigma_i^{-1} (\mu_j' - \mu_i) + \text{trace}(H) \quad (19)$$

From equation (17), it follows that

$$pp^T = \Sigma_j' \quad (20)$$

Now,

$$\begin{aligned} \text{trace}(H) &= \text{trace}(p^T \Sigma_i^{-1} p) \\ &= \text{trace}(\Sigma_i^{-1} pp^T) \\ &= \text{trace}(\Sigma_i^{-1} \Sigma_j') \end{aligned} \quad (21)$$

Thus, by substituting, one obtains

$$\begin{aligned} D_{ij} &\equiv E \left[D(y, \mu_j, \Sigma_i) \right] \\ &= (\mu_j' - \mu_i)^T \Sigma_i^{-1} (\mu_j' - \mu_i) \\ &\quad + \text{trace}(\Sigma_i^{-1} \Sigma_j') \end{aligned} \quad (22)$$

On the basis of this result, the quantity D_{ij} is used as the average distance between RSEG cluster j and its corresponding TSEG cluster i with the usual estimates for μ and Σ . For any candidate transformation, one would like to have a good match for each

cluster. For a recognition segment cluster j , the following is defined.

$$G_j = \min_i D_{ij} \quad (23)$$

In computing the D_{ij} , the candidate transformation is applied to each μ_j . However, Σ_j remains unchanged.

To minimize the influence of unmatched clusters, it is useful to transform G_j to f_j , say where f_j ranges from 0 to 1 with 1 indicating a perfect fit; further, it is desirable that f_j diminish as the fit becomes poorer, with f_j approaching 0 rapidly as G_j becomes fairly large. The object is to minimize the influence of unmatched clusters. Set

$$f_j = \exp(-G_j/\sigma) \quad (24)$$

where σ is a user-supplied scaling parameter. The overall goodness measure is the sum of the f_j .

A detailed description of OSCAR appears in the appendix. In summary, the OSCAR method consists of four major steps.

Step 1.—Cluster the two segments and determine rank vectors for each cluster/subclass. The rank vectors are formed by computing the rank of each component cluster/subclass within its segment.

Step 2.—Compare all training segment/recognition segment pairs of rank vectors. Tag all those that are sufficiently close so that there is a fair probability that they belong to the same class. Such pairs will be called admissible.

Step 3.—For each nonoverlapping pair of admissible pairs, find the a_k, b_k transformation (eq. (1)) that fits the pairs of matching points. Pairs of pairs are nonoverlapping if all four clusters are unique. Determine the goodness measure for each resulting transformation that yields reasonable component values for the multiplicative factor a_k .

Step 4.—Rank the resulting transformations on the goodness measure and compute a weighted average of the best candidates. The weighting system is based on the ranks and goodness measures obtained for the candidate transformations.

Modified OSCAR.—The algorithm OSCAR chooses a pair of clusters in both the training and the

recognition segments. Using these four clusters, a channelwise linear transformation is computed and evaluated. The "best" transformation is chosen. Instead of using a pair of clusters from each segment, the modified OSCAR uses a single cluster and its projection onto the soil line. The steps in the procedure are as follows.

Step 1.—Rotate the means using Kauth rotation matrix R (ref. 5). (See the section on the modified ROOSTER procedure for an explanation of the Kauth transform.)

Step 2.—Find the minimum cluster-mean values in the rotated second and third channels.

Step 3.—Define the rotated projected vectors by replacing the second and third channels with the second and third channel minimums.

Step 4.—Rotate the projected vectors back into channel space using R^T .

Step 5.—For each pair of mean vectors (one in the training segment, one in the recognition segment, and both of which are greenness vectors), define the transformation that maps the training vector onto the projected recognition vector.

Step 6.—For each transformation, test whether it is close to a constant times the identity matrix.

Step 7.—For each transformation that passes the test in step 6, compute the OSCAR function, equation (22).

Step 8.—The transformation with the largest OSCAR function value is the desired transformation.

A detailed description of the modified OSCAR algorithm is given in reference 6.

The MLEST Algorithm

Introduction.—The MLEST algorithm (ref. 8) obtains maximum-likelihood estimates (MLE's) of the affine transformation that is assumed to relate the statistics of the training segment to those of the recognition segment. The MLEST algorithm is based on the following major assumptions.

1. The training and recognition segment samples are drawn from probability density functions that are mixtures of normally distributed subclasses.

2. The number of subclasses in the training segment is equal to the number of subclasses in the recognition segment. Training segment subclasses that do not exist in the recognition segment may be represented in the model by a priori probabilities of 0.

3. The training segment subclass statistics (i.e., means and covariances) are related to the recognition segment subclass statistics by a positive definite affine transformation.

Mathematical development.—In the mathematical development that follows, it is assumed that the number of training segment subclasses M_T equals the number of recognition segment subclasses $M_R = M_R$; therefore, $M_T = M_R = M$.

Let $p(x|i)$, $i = 1, 2, \dots, M$, represent the probability density functions for the training segment subclasses. Since the training segment subclasses are assumed to be normally distributed,

$$p(x|i) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1} (x-\mu_i)} \quad (25)$$

where $i = 1, 2, \dots, M$. The overall mixture density function for the training segment is given by

$$p(x) = \sum_{i=1}^M q_i p(x|i) \quad (26)$$

By assumption 3, the training segment subclass statistics are related to the recognition segment subclass statistics by a positive definite affine transformation. This transformation may be represented by the $(p \times p)$ real positive definite matrix A and the $(p \times 1)$ real vector B . It follows that the recognition segment subclass statistics (means and covariance matrices) are given by

$$\mu_i' = A\mu_i + B \quad (27)$$

and

$$\Sigma_i' = A\Sigma_i A^T; i = 1, 2, \dots, M \quad (28)$$

From equations (26) to (28), it follows that the mixture density function for samples from the recognition segment is given by

$$p(y) = \sum_{i=1}^M q_i p(y|i) \quad (29)$$

where

$$p(y|i) = \frac{1}{(2\pi)^{\frac{p}{2}} |A\Sigma_i A^T|^{\frac{1}{2}}} e^{-\frac{1}{2} (y - A\mu_i - \mathbf{B})^T (A\Sigma_i A^T)^{-1} (y - A\mu_i - \mathbf{B})} \quad (30)$$

and $i = 1, 2, \dots, M$. Next, suppose that one picks N statistically independent samples from the recognition segment y_1, y_2, \dots, y_N . Then, the likelihood function is given by

$$l(y_1, y_2, \dots, y_N) = \prod_{k=1}^N p(y_k) \quad (31)$$

The algebra is simplified considerably if one uses the logarithm of the likelihood function

$$L = \log_e l = \sum_{k=1}^N \log_e p(y_k) \quad (32)$$

It may be shown that the partial derivatives of L with respect to the matrix A , the vector \mathbf{B} , and the a priori probabilities q_i , $i = 1, 2, \dots, M$, are given, respectively, by

$$\frac{\partial L}{\partial A} = \left[\sum_{k=1}^N \sum_{i=1}^M p(i|y_k) (A\Sigma_i A^T)^{-1} \cdot (y_k - A\mu_i - \mathbf{B}) (y_k - \mathbf{B})^T - (N)I_p \right] (A^{-1})^T \quad (33)$$

$$\frac{\partial L}{\partial \mathbf{B}} = \sum_{k=1}^N \sum_{i=1}^M p(i|y_k) (A\Sigma_i A^T)^{-1} \cdot (y_k - A\mu_i - \mathbf{B}) \quad (34)$$

$$\frac{\partial L}{\partial q_i} = \sum_{k=1}^N p(i|y_k); i = 1, 2, \dots, M \quad (35)$$

subject to the constraints

$$q_i \geq 0; i = 1, 2, \dots, M \quad (36)$$

$$\sum_{i=1}^M q_i = 1 \quad (37)$$

where I_p is the $(p \times p)$ identity matrix and

$$p(i|y_k) = \frac{q_i p(y_k|i)}{p(y_k)} \quad (38)$$

The general MLEST algorithm obtains estimates of the $(p \times p)$ matrix A , the $(p \times 1)$ vector \mathbf{B} , and the a priori probabilities q_i , $i = 1, 2, \dots, M$, that maximize the logarithmic likelihood function L . Estimates obtained in this manner are called maximum-likelihood estimates.

In practice, the optimization indicated previously is performed by using the Davidon-Fletcher-Powell (DFP) constrained optimization program (ref. 9). The DFP program uses equation (32) for the likelihood function and equations (33) through (35) for its partial derivations to modify A , \mathbf{B} , and q_i , $i = 1, 2, \dots, M$, in a manner such that L is maximized. In many cases, the likelihood function (eq. (32)) is insensitive to q_i and therefore q_i can be set to a constant, $q_i = 1/M$, $i = 1, 2, \dots, M$, and not considered in estimating A and \mathbf{B} .

The UHMLE Algorithm

Introduction.—The UHMLE procedure (refs. 10 to 17) obtains estimates of the recognition segment subclass statistics (i.e., the means, the covariances,

and the a priori probabilities) by correcting the training segment subclass statistics for small differences between the training and recognition segment subclass statistics using an iterative maximum-likelihood correction procedure. The UHMLE algorithm is based on the following major assumptions.

1. The training and recognition segment samples are drawn from probability densities that are mixtures of normally distributed subclasses.

2. The number of subclasses in the training segment is equal to the number of subclasses in the recognition segment. Training segment subclasses that do not exist in the recognition segment may be represented in the model by a priori probabilities of 0.

3. Good initial estimates of the recognition segment subclass statistics may be obtained from the training segment.

4. The differences between training and recognition segment statistics at the subclass level are caused not only by haze and Sun angle differences but also by random differences between the two scenes (i.e., differences in crop growth stages, soil color, soil moisture, leaf area index, etc.) which cannot be modeled by an affine transformation. The parameters of each subclass require a correction that is independent of the corrections applied to the parameters of the other subclasses. This is considered to be an important feature of the UHMLE procedure.

Mathematical development.—Let $\{y_k\}$, $k = 1, \dots, N$, be an unlabeled sample of observations from the recognition segment. These samples are assumed to be drawn from a mixture of M_R populations in the RSEG, where each population is normally distributed. It is assumed that the number of training segment subclasses M_T equals the number of RSEG subclasses M_R ; that is, $M_T = M_R = M$.

Let $p(y|i)$, $i = 1, 2, \dots, M$, represent the probability density functions for the RSEG subclasses. Since the RSEG subclasses are assumed to be normally distributed,

$$p(y|i) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma'_i|^{\frac{1}{2}}} e^{-\frac{1}{2} (y - \mu'_i)^T \Sigma'_i^{-1} (y - \mu'_i)} \quad (39)$$

The overall mixture density function for the RSEG is given by

$$p(y) = \sum_{i=1}^M q'_i p(y|i) \quad (40)$$

The RSEG subclass statistics $\{q'_i, \mu'_i, \Sigma'_i\}$, $i = 1, \dots, M$, are unknown, but good initial estimates of the statistics $\{\hat{q}'_i, \hat{\mu}'_i, \hat{\Sigma}'_i\}$, $i = 1, \dots, M$, are assumed to be available from the TSEG. Therefore, using unlabeled independent samples y observation from the RSEG, $\{y_k\}$, $k = 1, \dots, N$, maximum-likelihood estimates of the subclass statistics $\{q'_i, \mu'_i, \Sigma'_i\}$, $i = 1, \dots, M$, may be obtained which locally maximize the log likelihood function

$$L = \sum_{k=1}^N \log_e p(y_k) \quad (41)$$

Clearly, L is a differentiable function of the parameters to be estimated. Equating to 0 the partial derivatives of L with respect to these parameters, one obtains, after a straightforward calculation, the following necessary conditions for a maximum-likelihood estimate for subclasses $i = 1, \dots, M$.

$$\hat{q}'_i = \frac{1}{N} \sum_{k=1}^N p(i|y_k) \quad (42)$$

$$\hat{\mu}'_i = \frac{\sum_{k=1}^N y_k p(i|y_k)}{\sum_{k=1}^N p(i|y_k)} \quad (43)$$

$$\hat{\Sigma}'_i = \frac{\sum_{k=1}^N p(i|y_k) (y_k - \hat{\mu}'_i)(y_k - \hat{\mu}'_i)^T}{\sum_{k=1}^N p(i|y_k)} \quad (44)$$

where

$$p(i|y_k) = \frac{\hat{q}_i^k p(y_k|i)}{\sum_{j=1}^m \hat{q}_j^k p(y_k|i)} \quad (45)$$

These are known as the likelihood equations.

An alternative set of likelihood equations proposed by Peters and Walker (ref. 10) for subclasses $i = 1, \dots, M$, is

$$\hat{q}_i^k = (1 - \epsilon)\hat{q}_i^k + \epsilon \frac{1}{N} \sum_{k=1}^N p(i|y_k) \quad (46)$$

$$\hat{\mu}_i^k = (1 - \epsilon)\hat{\mu}_i^k + \epsilon \frac{\sum_{k=1}^N y_k p(i|y_k)}{\sum_{k=1}^N p(i|y_k)} \quad (47)$$

$$\hat{\Sigma}_i^k = (1 - \epsilon)\hat{\Sigma}_i^k + \epsilon \frac{\sum_{k=1}^N p(i|y_k) (y_k - \hat{\mu}_i^k)(y_k - \hat{\mu}_i^k)^T}{\sum_{k=1}^N p(i|y_k)} \quad (48)$$

As shown by Peters and Walker (refs. 10 to 12, 14, 15, and 18), given any sufficiently small neighborhood of the true parameters and for $\epsilon \neq 0$, the probability is 1 that if N is sufficiently large, there is a unique solution of the likelihood equations in that neighborhood, and this solution is a maximum-likelihood estimate of the true parameters.

The likelihood equations, as written, suggest the following iterative procedure for obtaining a solution. Beginning with a set of starting values (obtained from the TSEG subclasses), obtain successive approximations to a solution by inserting the preceding approximations in the expression on the right-hand sides of equations (46) to (48). Category labels (i.e., wheat and nonwheat), attached to the TSEG subclass used for starting values, are carried through the successive approximations; i.e., if a starting mean vector $\hat{\mu}_i^k$ was computed using samples labeled wheat in the

TSEG, it will carry a wheat label through the successive approximations. On occasion, this successive approximation procedure may result in a TSEG wheat (or nonwheat) subclass being associated with a subclass of the opposite category in the RSEG. This process is called label switching. In those cases, an analyst may be required to correct for subclass mislabeling in the RSEG.

PERFORMANCE TESTS OF SIGNATURE EXTENSION ALGORITHMS ON SIMULATED DATA AND CONSECUTIVE-DAY DATA

Comparative tests were performed on seven signature extension algorithms to evaluate their effectiveness in correcting for changes in atmospheric haze and Sun angle in a Landsat scene (ref. 19). The evaluation criteria were classification accuracy and proportion estimation accuracy. The algorithms tested were the Maximum Likelihood Estimation of Signature Transformation, the University of Houston Maximum Likelihood Estimation, the Optimal Signature Correction Algorithmic Routine, modified OSCAR (MOD OSCAR), the Rank Order Optimal Signature Transformation Estimation Routine, modified ROOSTER (MOD R), and the Atmospheric Correction program.

The Data Sets

Two data sets were used—one consisting of simulated data described in the section immediately following and the other a set of acquisitions on consecutive days described in the succeeding section. The simulated data provided for a controlled experiment in which the transformations were known and in which the problems of nonnormal distributions and nonrepresentative statistics were avoided. The consecutive-day data set provided for a test of the capability of the algorithms to correct for atmospheric effects when effects caused by differences in the training and recognition segment ground scenes are eliminated. The algorithms ROOSTER, UHMLE, and MLEST were tested on the simulated data. All the algorithms were tested on the consecutive-day data set.

Simulated data.—The 1975 data base of the Earth Resources Interactive Processing System (ERIPS) contains four passes of four-channel simulated data for each of segments 429 and 432. Each segment has

117 lines and each line 196 pixels. The field coordinates reside in the ERIPS field data base. Five classes exist within each segment: wheat, barley, stubble, grass, and fallowed ground. Each class is divided into two subclasses.

The data were generated from means and covariance matrices determined from training fields in Hill County, Montana. An algorithm was used to generate multivariate normal data with the same statistics. This was done separately for the four passes of segment 429. Each pass of segment 432 was created from the distributions used in the corresponding pass of segment 429 by transforming them with an affine transformation so that the data corresponded to a different Sun angle. Segment 429 was chosen to be the TSEG and segment 432 the RSEG. All classifications were made using four channels corresponding to a particular pass. Each data set corresponds to one of the four passes: SIM1, SIM2, SIM3, and SIM4.

Consecutive-day data.—Seven sets of consecutive-day passes of Landsat-1 data from intensive test sites in Ellis, Finney, and Saline Counties, Kansas, were used in another test. The first data set is denoted F1709-8. (The F indicates Finney County; 1709-8 indicates the dates of the training and recognition passes, respectively; i.e., the training pass was made 1709 days and the recognition pass over the same segment occurred 1708 days after the launch of Landsat-1.) In all, four data sets from the Finney, two from the Saline, and one from the Ellis County test site were used.

Ground truth was available for all fields in all test sites. A subset was selected for training fields, and fields were grouped into subclasses with the aid of cluster maps. In general, the rectangular ground-observed areas were not oriented so that their sides were parallel to the scan lines in the Landsat-1 data. To facilitate the application of the various algorithms, a "signature extension area" was defined as the smallest rectangular area with sides parallel and perpendicular to the Landsat scan lines that included the ground-observed area in each case. For Finney County, this included the entire 9- by 11-kilometer segment (117 lines, 196 pixels) containing the ground-observed area. For Saline County, it included lines 26 to 91 and pixels 27 to 146; for Ellis, it included lines 24 to 109 and pixels 49 to 144.

Approach

The overall approach was to make signature extension runs using the algorithms being tested and to compare the results with local classification results with ground truth. Local classification results include the PCC results and the wheat proportion estimates obtained from classification of the recognition segment (or pass) with statistics generated from the same segment (or pass). The algorithms were to provide modified training statistics, which then were used to classify the recognition area. The UHMLE computes these modified statistics directly; all the other algorithms compute an affine transformation, which is then used to modify the training statistics.

The algorithms tested.—A short description of the algorithms and how they were operated in this test is provided here. Detailed descriptions of the algorithms are provided elsewhere in this paper. In the case of the consecutive-day data, the algorithms were usually run using the data from the signature extension area already defined. Exceptions will be noted.

MLEST: The MLEST technique uses an iterative gradient optimization procedure (the Davidon-Fletcher-Powell algorithm) to obtain maximum-likelihood estimates for the affine transformation assumed to relate the training and recognition statistics. The training subclass a priori probabilities and statistics are input to the program, which outputs the maximum-likelihood estimate of the affine transformation

UHMLE: The UHMLE takes subclass statistics from a TSEG and image data from an RSEG and computes maximum-likelihood estimates of subclass proportions and statistics for the RSEG. Two versions of UHMLE were used. The first, UH all, uses the ground-observed area as input data; when this version is used to obtain maximum-likelihood estimates of proportions generated internally by UHMLE, it is referred to as UH all MLE. The second, UH fields, uses only the training fields within the RSEG; when this version is used to obtain maximum-likelihood estimates of proportions generated internally by UHMLE, it is referred to as UH fields MLE. The second version was introduced to eliminate the effect of insufficient training. The statistics generated by UHMLE are used to classify the

ground-observed area in the RSEG.

OSCAR: The OSCAR considers every possible transformation defined by four cluster means: two in the TSEG and two in the RSEG. From these transformations, the algorithm selects those that are "best" able to match the training clusters with the recognition clusters. The amount of computation is kept to a manageable level (1) by rejecting pairings judged to be unreasonable on the basis of rankings and (2) by testing the remaining transformations, using each to transform all the training clusters, and calculating a goodness-of-fit measure based on the distance of the transformed training clusters from the recognition clusters. The five transformations giving the "best" fit are then averaged.

Modified OSCAR: The MOD OSCAR, in effect, defines a transformation for each pair of clusters — one in the RSEG and one in the TSEG. Each cluster is used with its projection onto the soil line¹ to define a transformation. The transformations are evaluated as in OSCAR, and the best transformation is output.

ROOSTER: To perform signature extension with ROOSTER (ref. 3), one first obtains a set of class means for the TSEG's and the RSEG's. These class means are obtained by clustering or by deriving class statistics from training fields.

The first step is to derive rank vectors corresponding to each of the class means. These rank vectors are obtained by computing for each channel the rank of each mean relative to the others for that segment. The rank vectors are used to match the classes (or clusters) in the training area with those in the recognition area. Then, a regression analysis is used to determine the affine transformation that best transforms the mean vectors from the training area into the corresponding mean vectors from the recognition area.

In this study, the ROOSTER was used in three different ways. The first, R(C), consisted of using clusters to define the class means for both segments; the second, R(S), used subclass means derived from training fields for both segments (it is expected to provide an estimate of how well ROOSTER would do if an ideal clustering algorithm were available);

¹The soil line is the "bottom of the tasseled cap" or that part of channel space containing bare soil (ref. 5).

and the third, R(S/C), used subclass statistics for the TSEG and clusters for the RSEG (this is an alternate way of using ROOSTER operationally, since subclass statistics are always available for the training area).

Modified ROOSTER: The MOD R is identical to ROOSTER except that the regression line is computed with the cluster means and the projections of the cluster means onto the soil line.

ATCOR: The ATCOR program is designed to correct for differences in haze level and Sun angle between the training and recognition data sets. The program processes each of these data sets separately. In each case, the input is the Landsat-1 data and the solar zenith angle. The ATCOR program determines the haze level from the brightness of certain dark targets in the scene and uses an atmospheric model to calculate a set of coefficients relating the Landsat data for that scene to the reflectance of the targets on the ground. The coefficients obtained from the training and recognition data sets are then used to compute the affine transformation to be applied to the training data to transform it to the observing conditions of the recognition segment.

REGRES: Rather than a signature extension algorithm, the REGRES program is a method for finding the optimum affine transformation to be applied to the statistics of the consecutive-day data. In each channel, a scatter plot is made of the second-day data versus the first-day data. A straight line is then fitted to the data which minimizes, in the least squares sense, the perpendicular distance from the points to the line. In principle, this line represents the best affine transformation for the training statistics.

Classification and evaluation.—After obtaining the modified statistics, the standard LACIE classification procedure was implemented on a Univac 1108 computer as part of the program EOD-LARSYS to classify the RSEG's. A two-class classifier was used with equal a priori probabilities for wheat and non-wheat. Within each class, the subclasses had equal a priori probabilities. A 1-percent chi-squared threshold was used at the subclass level to reject outliers. For the simulated data, entire areas were classified; for the consecutive-day data, the ground-observed areas were classified.

Classification accuracy: The classification accuracy was determined for wheat and nonwheat by

using the training fields previously defined as test fields. From these, the overall accuracy was computed.

$$\text{Overall accuracy} = q_w p(w/w) + q_\phi p(\phi/\phi) \quad (49)$$

where $p(w/w)$ is the wheat accuracy, $p(\phi/\phi)$ is the nonwheat accuracy, q_w is the wheat proportion in ground-observed area, and q_ϕ is the nonwheat proportion in ground-observed area. The proportions q_w and q_ϕ were known from ground truth. The wheat, nonwheat, and overall accuracies were compared with the results obtained from local classification.

Wheat proportions: The classification results yielded wheat proportions for the ground-observed areas. In addition, the UHMLE program yielded a maximum-likelihood estimate of the wheat proportions. These results were compared with the ground-observed proportions and the results obtained from local classification.

Results

The results of this processing are given in tables II through XII. Table II gives the a_k and b_k , $k = 1, \dots, 4$, coefficients determined for the consecutive-day data by those algorithms that produce an affine transformation. The algorithms are listed in the order in which they performed in the accuracy test.

On the basis of numerical calculations using an atmospheric model (ref. 1), certain constraints are expected to apply to the a_k and b_k coefficients corresponding to a change in the haze level. These should apply to the consecutive-day data if the haze levels present are uniform. Among these constraints, which apply to all channels, are the following.

1. If there is no difference in haze level between the TSEG and the RSEG, $a_k = 1.0$ and $b_k = 0.0$, $k = 1, \dots, 4$.

2. If the TSEG has more haze than the RSEG, $a_k > 1.0$ and $b_k < 0.0$, $k = 1, \dots, 4$.

3. If the TSEG has less haze than the RSEG, $a_k < 1.0$ and $b_k > 0.0$, $k = 1, \dots, 4$.

In many cases, the data in table II do not obey these rules. Examples can be found in the following anomalies.

1. $a_k > 1.0$ for some channels and $a_k < 1.0$ for others; e.g., R(S) for F1655-4.

2. $a_k > 1.0$ and $b_k > 0.0$; e.g., MLEST for F1673-2.

3. $a_k < 1.0$ and $b_k < 0.0$; e.g., R(C) for F1726-7, channel 2.

These failures to obey the constraints may be due in part to nonuniform haze levels in the data and to changes in the look angle across the scene.

Tables III to VI give the accuracy results for wheat and nonwheat using both data sets. The accuracy obtained with signature extension is expressed as a percentage difference from the local result; i.e.,

$$\text{Percentage difference} = \frac{\text{signature extension accuracy} - \text{local accuracy}}{\text{local accuracy}} \times 100 \quad (50)$$

Tables VII and VIII give similar results for overall accuracy.

Tables IX through XII give the differences for both data sets (1) between results obtained using signature extension and local classification and (2) between results obtained using signature extension and ground truth. The means and standard deviations were obtained using the absolute values of the numbers in the tables.

Analysis

This section reports a statistical analysis of the data in tables VII and XI. Data for the UHMLE algorithm were not included because of their large variances.

First, an analysis of variance was performed on the data in table VII. The purpose of an analysis of variance is to separate a response variable into component parts. In this way, the test for a particular factor will become more sensitive because variations due to other causes have been removed. In this experiment, two factors were present: signature extension algorithms and the seven consecutive-day acquisitions. The variation over the seven data sets could have been allocated to any of several different causes: ground scene variations from day to day, variations from one geographic location to another, or changes in the haze level from day to day.

The last alternative was chosen. Each pass was classified as either clear or hazy by visually inspecting the images of the data produced by ERIPS. The results are shown in table XIII. Three TSEG-RSEG

combinations occurred; namely, haze-clear, clear-haze, and clear-clear. It was assumed that each combination would produce different results (classification accuracy, etc.), thus the need for this factor in the analysis.

An interaction between the algorithms and haze combinations ($A \times H$) was also expected to be present; that is, one algorithm might have performed well for the clear-haze consecutive-day acquisitions and poorly for the haze-clear days, whereas the opposite results might have occurred for another algorithm.

The model for the experiment was

$$y_{ijk} = \mu + a_i + h_j + a_{ihj} + e_{ijk} \quad (51)$$

where y_{ijk} is the response variable, μ is the overall mean, a_i is the contribution of the i th algorithm, h_j is the contribution of the j th haze level, a_{ihj} is the contribution of algorithm i and haze level j to the interaction, and e_{ijk} is the error term for the k th observation for the i th algorithm and the j th haze level. In the analysis of variance for overall accuracy, y_{ijk} is the percentage accuracy difference; that is, the quantity given in table VII. The results of this analysis of variance are given in table XIV, where significant differences between the algorithms and between the haze conditions are apparent.

Table XV gives the average accuracy difference over the algorithms for each set of haze conditions. Because the analysis of variance indicated significant differences between haze conditions, one can infer from table XV that the presence of haze over the TSEG is significantly different from the other two conditions.

The results for the different haze conditions were plotted as a function of the algorithms (fig. 1). The condition with haze over the TSEG showed consistently better results than the other two conditions. A similar analysis was performed for the wheat proportion errors reported in table XI; R(S/C) was not included because of its large variance. The results showed a significant difference between the haze conditions but not between the algorithms (table XVI). Table XVII gives the average proportion difference over the algorithms for each haze condition, and figure 2 shows the performance of each algorithm for each of the three haze conditions. Here again, the haze-clear condition seemed to give the best results.

Conclusions

The results of these tests are summarized in table XVIII. The first two columns list the algorithms in the order in which they performed on the accuracy test for the simulated and consecutive-day data. The numbers given are the mean percentage differences (see tables VII and VIII). The minus signs indicate that the algorithm was less accurate than local classification. A statistical analysis was performed on the accuracy results for the consecutive-day data with the exception of data for the three versions of UHMLE (which were omitted because of large variances). The analysis indicated (1) that there were no significant differences among the algorithms and (2) that the results obtained when the TSEG appeared hazier than the RSEG were better than in the other two conditions observed; i.e., when both were clear or the RSEG was hazier.

The comparison of wheat proportion differences (between ground truth and proportions provided by algorithms and between local results and results from algorithms) in the last four columns of table XVIII shows the performance order of the algorithms to be nearly the same for simulated data but quite different for consecutive-day data. This was because local proportion estimates were quite different from ground-observed proportions for the consecutive-day data. These four columns of numbers are the means of the absolute values of the differences as given in tables IX to XII. A statistical analysis was performed on the consecutive-day data for wheat proportion differences from local results. Data from R(S/C) and the three versions of UHMLE were not used because of large variances. The results given in table XVI indicate no significant differences among the algorithms tested. Here again, the best results were obtained when the TSEG appeared hazier than the RSEG.

Finally, it must be mentioned that, because of time limitations, this test was performed using the currently available algorithms. Subsequently, it has been discovered that some of the algorithms show better performance when later versions are used. For example, the program UHMLE has a later version that uses a transformation of the form $(x + b)$ to get from the TSEG statistics an initial guess for the RSEG statistics. However, the results presented in this paper provide some indication of how the algorithms tested can be expected to work when they are applied to the signature extension problem.

GEOGRAPHICAL EXTENSION TESTS OF THE ATCOR, OSCAR, AND MLEST SIGNATURE EXTENSION ALGORITHMS

Geographical tests were performed on the ATCOR, OSCAR, and MLEST signature extension algorithms. The objectives of the tests were to evaluate these algorithms in a more realistic environment using LACIE image data and training field definitions provided by analyst interpreters. This is a more realistic test than the consecutive-day test and the simulated data test described in the preceding section. All three algorithms were tested on the same data set, but ATCOR and OSCAR were tested separately from MLEST.

Geographical Extension Data Set

The 1975-76 crop year operational data taken over Kansas, with associated labeled fields, were chosen for use in these tests. This set consisted of 28 LACIE training and recognition segment pairs, with approximately 7 segment pairs in each of the 4 wheat biowindows. Recognition segments and training segments were paired so as to minimize differences between the segments due to static factors in the scenes such as soil types, climate, and topography.

The Geographical Signature Extension Test of ATCOR and OSCAR

The objectives of the ATCOR and OSCAR test^{2,3} were as follows.

1. To compare the performance of the signature extension algorithms on the recognition segment to that attained using (a) local training signatures from the recognition segment and (b) untransformed signatures from the training segment
2. To relate performance to the biowindow in which the recognition and training segments were acquired
3. To evaluate the effects of clustering on the performance of the OSCAR algorithm

²S. G. Wheeler, "Signature Extension Experiment," IBM Memorandum IBM-RES-23-9, Contract NAS 9-14350, June 3, 1976.

³S. G. Wheeler, "Results of the Signature Extension Experiment," IBM Memorandum IBM-RES-23-1-1, Contract NAS 9-14350, July 29, 1976.

4. To evaluate the quality measures used in the algorithms as predictors of signature extension performance

Description of the experiment.—Five classification runs were performed on each TSEG and RSEG pair, as follows.

1. Local—using statistics from the recognition segment's training subclasses
2. Untransformed—using statistics from the training segment's training subclasses
3. OSCAR with clusters—using training segment statistics after applying an affine transformation derived by the OSCAR algorithm operating on cluster statistics in the training and recognition segments
4. ATCOR—using training segment statistics after applying an affine transformation derived by the ATCOR algorithm
5. OSCAR with subclasses—using training segment statistics after applying an affine transformation derived by the OSCAR algorithm operating on statistics from the AI-defined subclasses in the recognition and training segments

Response variables for the experiment were

1. Estimated percentage of wheat in the recognition segments at 0-percent and 1-percent thresholding
2. Observed classification accuracies in the test fields defined by the AI's in the recognition segment
3. Percentage of pixels thresholded in the recognition segment and in the test fields at the 1-percent thresholding level

The percentage of wheat in the segment was calculated by counting the number of pixels classified as wheat in the segment, subtracting the number of pixels classified as wheat within designated "other" fields, and then dividing by 22 932 (the total number of pixels in a LACIE segment). Thresholding in the segment was calculated in a similar manner. The OSCAR algorithm produces a quality factor that measures the closeness with which the transformed signatures match the recognition segment's cluster-based (or subclass-based) signatures. Also, the ATCOR algorithm produces an estimate of the haze levels in the recognition and training segments.

ATCOR and OSCAR geographical signature extension results.—Results of the ATCOR and OSCAR geographical signature extension tests are presented in the following paragraphs.

Percentage of wheat in the segment: The averages of the estimated values of the percentage of wheat in

the segments at 0-percent and 1-percent thresholding are given in table XIX. The same quantities are given, by biowindow, in tables XX to XXII. (Differences are caused by the fact that the category OSCAR with subclasses could not be run on one segment because of the small number of subclasses in the segment. Values in tables XX to XXII do not include that segment.) The standard errors (i.e., standard deviations of the average values) can be used to place confidence limits on the mean values but are not appropriate for testing for differences among the mean values. Significance tests, using the Fisher's F-factor, were performed, and appropriate standard errors that reflect the structure of the experiment were calculated.

Test statistics indicated that there were no significant differences among the estimates of the percentage of wheat in the segment by local classification or the four signature extension techniques, no significant differences among the estimated percentage of wheat in the different biowindows, and no significant interaction between the classification techniques and the biowindows. The lack of interaction suggests that, within the limits of experimental error, all the techniques work equally well, or poorly, in all biowindows.

The average values and standard deviations of the percentages of the pixels thresholded while classifying the recognition segment using statistics derived by the different techniques are given in table XXII. These tabulated values exhibit large differences among the techniques. However, an analysis of variance test indicated there were no significant differences among the classification techniques.

The previously described tests showed little difference between local classification in the recognition segment and the signature extension techniques except for classification of the RSEG with untransformed TSEG statistics, a technique that caused undue thresholding in the segment. Figures 3 through 6 are scatter plots of signature extension versus local estimates of percentage of wheat in the segments. (The solid line in these plots is $X = Y$ and the dashed line is the regression line $Y = f(X)$ given in the legend.) These plots show a large segment-to-segment variability between local and signature extension estimates and, consequently, a small (≈ 0.4) correlation between the estimates. Figure 7, a plot of ATCOR versus OSCAR estimates of percentage of wheat, shows a surprisingly high correlation between these estimates. This strong relationship implies that, despite their very different approaches toward

solving the signature extension problem, the techniques produce basically the same result and a large experiment would be required in order to choose one over the other.

Algorithm haze and quality factors: Table XXIII gives simple correlation coefficients relating the algorithm quality factors and ATCOR estimates of the differences between the haze levels in the recognition and training segments to the percentage of wheat estimates. As would be hoped, the thresholding produced by using untransformed training signatures is highly correlated with the change in haze measurements predicted by ATCOR. Figures 8 through 12 show plots of the significant factors in table XXIII versus the difference in haze levels between the recognition and training segments as predicted by ATCOR.

Since the performance of the cluster-based OSCAR algorithm does not correlate with any of the quality factors, these would not appear to be useful predictors of algorithm performance. The quality factor produced by using OSCAR with subclasses does correlate with some of the thresholding rates but these correlations could not generally be used in a signature extension situation.

Training field accuracies: Tables XXIV and XXV contain average training field accuracies for, respectively, the wheat and nonwheat training fields when classified without thresholding. F-tests in the analysis of variance with subsequent multiple comparison tests show significant differences between local classification and the signature extension techniques but no differences among the techniques. Similar results hold for the training field accuracies at 1-percent thresholding as shown in tables XXVI and XXVII. Each of these analysis of variance tests suggests that training field accuracies may be dependent upon the biowindow, but, as in the percentage of wheat estimates, there are no interactions between biowindows and algorithms.

Thresholding rates in the training fields, tables XXVIII and XXIX, show large differences between local classification and the signature extension techniques, particularly the untransformed technique. Both the average and the variability in the thresholding rates are much larger for the signature extension techniques than for local classification. There are, however, no significant differences among the results from ATCOR and the two OSCAR algorithms.

Conclusions from the ATCOR and OSCAR geographical signature extension tests: In terms of

average estimates of the percentage of wheat in the segment, there is little evidence for choosing local classification or one of the four signature extension techniques over the others, except for the untransformed approach, which shows a tendency to produce far too many thresholded pixels.

Local classification produces better training field accuracies than any of the signature extension algorithms, particularly untransformed. Differences between local classification and the algorithms may not be real since the observed differences probably can be explained by the known bias in local classification due to reclassifying training data. The untransformed approach, however, appears to do more poorly than the others.

There is little evidence in any of the data to suggest any difference between the performance of OSCAR with clusters and OSCAR with subclasses. It thus appears that clustering neither helps nor degrades OSCAR's performance appreciably.

Probably the most important and startling result of this analysis is the demonstrated high correlation between results from ATCOR and OSCAR. This shows that these, despite being completely different in concept, produce very similar results and only an exceptionally large experiment could establish a choice between them.

The Geographical Signature Extension Test of MLEST

The objectives of the MLEST test (refs. 20 and 21) were as follows.

1. To compare the performance of the MLEST signature extension algorithm on the recognition segment to that attained using (a) local training signatures from the recognition segment and (b) untransformed signatures from the training segment

2. To relate performance to the biowindow in which the recognition and training segments were acquired

Description of the experiment.—The experiment was performed using the same Kansas data set used in the OSCAR and ATCOR test. One of the 28 training/recognition segment pairs used in the OSCAR/ATCOR test had significant data dropout, which MLEST was not capable of processing. This TSEG/RSEG pair was dropped, and the MLEST test was performed on the remaining 27 TSEG/RSEG pairs.

The evaluation procedure for the 27 TSEG/RSEG pairs consisted of two major steps. In step 1, MLEST signature extension runs were made for each segment pair to determine MLE estimates for each A matrix and B vector. In all of these runs, the individual subclass a priori probabilities were assumed equal and held constant. The MLEST iteration was begun with A being the identity matrix and B being the null vector. Also, A was restricted to be diagonal in all runs. In step 2, the affine transformed training segment signatures were used to classify each recognition segment using the LACIE maximum-likelihood classifier. Classification accuracies were computed for wheat/nonwheat over recognition segment training fields. Overall classification accuracies were computed for each recognition segment using the formula

$$P_{\text{overall}} = 0.5 (P_{\text{CW}}) + 0.5 (P_{\text{CNW}}) \quad (52)$$

where P_{CW} is the wheat classification accuracy and P_{CNW} is the nonwheat classification accuracy.

The classification results were used to estimate wheat proportions at threshold values of 0 percent and 1 percent. The classification runs described previously were repeated using untransformed statistics from the training segment as well as statistics from the recognition segment training fields. Henceforth, the affine transformed classification results will be referred to as MLEST results; the untransformed training segment classification results will be referred to as UT results; and the recognition segment classification results will be referred to as LOCAL results.

MLEST geographical signature extension results.—The MLEST program converged normally for 23 of the 27 signature extension runs attempted. However, successful optimization iteration sequences could not be established for four segment pairs. Analysis of the data for these four segment pairs revealed that the recognition segment data were located relatively far from the modes of the corresponding initial estimates ($A = I$, $B = 0$) for the training segment mixture density functions in spectral space. This resulted in floating-point underflow problems in the likelihood function computations, which in turn caused the Davidon optimization iterations to abort. The MLEST program was rerun for these four segment

pairs using the following initial values for the affine transformation

$$A = I \quad (53)$$

$$B = \begin{pmatrix} \mu_1 & -\mu'_1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \mu_4 & -\mu'_4 \end{pmatrix} \quad (54)$$

where I = a 4×4 identity matrix

μ_i = mean value in channel i for the training segment, $i = 1, 2, 3, 4$

μ'_i = mean value in channel i for the recognition segment, $i = 1, 2, 3, 4$

In other words, a mean level adjustment (MLA) was used for the initial B vector. The reruns were successful, resulting in normal convergence for all four segment pairs.

Table XXX enumerates the classification accuracy results obtained with the geographical extension data set. Table XXXI lists, by biowindow, the average improvement in MLEST classification accuracy over UT accuracy and the average difference between MLEST classification accuracy and LOCAL classification accuracy. The average improvement and average difference percentages are defined as

$$\text{Average improvement} = \text{Avg}(P_{\text{MLEST}} - P_{\text{UT}}) \quad (55)$$

$$\text{Average difference} = \text{Avg}(P_{\text{LOCAL}} - P_{\text{MLEST}}) \quad (56)$$

where P_{MLEST} is the MLEST classification accuracy, P_{UT} is the UT classification accuracy, and P_{LOCAL} is the LOCAL classification accuracy.

Referring to tables XXX and XXXI, one can make the following observations.

1. The MLEST classification accuracies improved upon UT classification accuracies for a majority of the signature extension segment pairs. Improve-

ments in overall classification accuracy are indicated for 22 of the 27 segment pairs. Improvements in both the wheat and nonwheat classification accuracies are indicated for 14 of the 27 segment pairs.

2. The average improvement (table XXXI) in overall, wheat, or nonwheat classification accuracy is approximately 10 percent. The improvements in classification accuracy are particularly striking for segment pairs 1854/1025, 1882/1887, 1850/1887, 1880/1875, and 1883/1884. The improvements in wheat classification accuracy for these segment pairs range from approximately 23 percent for segment pair 1880/1887 to approximately 64 percent for segment pair 1882/1887.

3. The degradations ($P_{\text{UT}} - P_{\text{MLEST}}$) in classification accuracy resulting from the use of MLEST are relatively insignificant. The average degradation (five segment pairs) in overall accuracy is less than 2 percent. The average degradation in wheat classification accuracy (seven segment pairs) is less than 4 percent. The average degradation in nonwheat classification accuracy (nine segment pairs) is less than 3 percent.

4. The improvements in classification performance do appear to depend on the biowindow (table XXXI) in which the data were collected. Average improvements in classification accuracy are approximately 14 percent for biowindows 2 and 3, approximately 9.5 percent for biowindow 1, and approximately 4 percent for biowindow 4. These results are reinforced by the well-known fact that biowindows 2 and 3 provide maximum discrimination between wheat and nonwheat.

5. The MLEST classification accuracies fall short of the LOCAL accuracies. The average difference between MLEST and LOCAL accuracies is approximately 18 percent for the overall accuracies, approximately 21 percent for the wheat accuracies, and approximately 15 percent for the nonwheat accuracies (table XXXI). However, the LOCAL classification accuracies are biased estimates since they were estimated over the same fields that were used to train the classifier. Thus, the difference between the MLEST classification accuracy and the "true" LOCAL classification accuracy should be less than this observed difference of 15 to 20 percent.

6. MLA starting values for the B vector were used for segment pairs 1882/1887, 1880/1875, 1877/1875, and 1883/1884. Considerable improvement in MLEST classification performance was noted for these sites. The effect of the MLA starting values

was to place the initial mixture density function in the general neighborhood of the recognition segment data. It is conjectured that the use of MLA starting values for the remainder of the signature extension data set would have resulted in better MLEST classification performance.

Table XXXII lists the UT, MLEST, and LOCAL wheat proportion estimates. Table XXXIII lists mean absolute differences between MLEST wheat proportion estimates and LOCAL wheat proportion estimates and between UT wheat proportion estimates and LOCAL wheat proportion estimates. These mean absolute differences are averaged separately for each biowindow and collectively for the entire data set.

Referring to tables XXXII and XXXIII, one can make the following observations.

1. The MLEST proportion estimates are closer to the LOCAL proportion estimates than are the UT estimates in 14 of 27 segment pairs with 0-percent thresholding and 11 of 27 segment pairs with 1-percent thresholding.

2. The extent of improvement is erratic; however, the MLEST estimates are closer, on the average (table XXXIII) to the LOCAL estimates than are the UT estimates. The average absolute differences computed for each biowindow between MLEST and LOCAL and between UT and LOCAL indicate that the MLEST proportions represent improvements over UT proportions for biowindows 1 to 3. The MLEST proportions represent degradation with respect to UT proportions for biowindow 4. This finding is reinforced by the classification accuracy results presented earlier, which showed that the smallest improvement in classification accuracy using MLEST was in biowindow 4.

3. The averages of the UT, MLEST, and LOCAL wheat proportion estimates (all 27 sites) from table XXX at 0-percent thresholding are approximately equal (within 1 percent of each other). The variances of these estimates are also essentially equal. For a 1-percent threshold, the average MLEST and LOCAL estimates are approximately equal; however, the average UT estimate differs about 5 percent from these estimates.

4. The amount of thresholding with the MLEST classifications is significantly less than that obtained with the UT classifications. Drastic reductions in thresholding are indicated for segment pairs 1854/1025, 1168/1173, 1882/1887, 1880/1887, 1880/1875, and 1172/1181.

Conclusions from the MLEST geographical signature extension test.—On the basis of tests conducted thus far, the following conclusions can be made.

1. The use of the MLEST algorithm leads to improvements in classification accuracy.

2. The MLEST wheat proportion estimates are, on the average, closer to the LOCAL wheat proportion estimates than are the UT wheat proportion estimates.

3. In reference to the geographical extension results, the MLEST algorithm performs best on data from biowindows 1 to 3.

4. The use of the MLEST affine transformed training segment signatures for classification drastically reduces the percentage of pixels thresholded.

These results demonstrate the viability of MLEST as a signature extension algorithm. It is conjectured that the use of MLA starting vectors, physical constraints on A and B , and the iterative equations for the a priori probabilities would lead to improvements in the performance of the MLEST algorithm.

SUMMARY AND CONCLUSIONS

This paper has described the results of the effort to develop a technology for signature extension during LACIE Phases I and II (1975 and 1976). A number of haze and Sun angle correction procedures were developed and tested. These included the ROOSTER and OSCAR cluster-matching algorithms and their modifications, the MLEST and UHMLE maximum-likelihood estimation procedures, and the ATCOR procedure. All these algorithms were tested on simulated data and consecutive-day Landsat imagery. The ATCOR, OSCAR, and MLEST algorithms were also tested for their capability to geographically extend signatures using Landsat imagery.

Several conclusions can be drawn from these tests.

1. In general, the paired TSEG/RSEG segment approach to signature extension described in this paper was not successful. This conclusion is based on the poor geographical extension test results presented in the preceding section. The primary source of error appeared to be the lack of representative crop signatures in a single training segment for use in structuring a classifier to classify a recognition segment. This conclusion was reached by comparing the results from the consecutive-day signature extension tests with the geographical signature extension

results. The consecutive-day signature extension results were very encouraging in that they indicated that if haze was the primary source of signature variation, it could be successfully corrected for with little or no degradation in the classifier's performance. When other sources of signature variation were introduced in the geographical signature extensions, a significant degradation in classifier performance was noted which was not correctable using haze and Sun angle correction procedures. The lack of success noted in this paired segments approach to signature extension led to the development of the multisegment training approach to signature extension described in the paper by Kauth and Richardson.

2. The affine transformation appears to be an appropriate model for use in correcting Landsat imagery for uniform haze and Sun angle differences.

3. Of the algorithms tested, MLEST and ATCOR appear to offer the most promise for signature extension. MLEST outperformed all other algorithms on the simulated data, the consecutive-day data, and the geographical extension data set. In addition, it is supported by maximum-likelihood estimation theory.

In the consecutive-day test and the geographical extension test, ATCOR was shown to improve classification performance. ATCOR also has a theoretical foundation in its physical model explaining the interaction of light reflected by the surface of the Earth, the atmosphere, and the Landsat sensor.

REFERENCES

1. Potter, J. F.: The Correction of Landsat Data for the Effects of Haze, Sun Angle, and Background Reflectance. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Ind., June 21-23, 1977, p. 24.
2. Henderson, R. G.; Thomas, G. S.; and Malila, R. F.: Methods of Extending Signatures and Training Without Ground Information. Final Report, ERIM 109600-6-F, May 1975.
3. Finley, D. R.: A New Algorithm To Correct MSS Signatures for Atmospheric Effects. LEC-6246, Lockheed Electronics Co., Houston, Tex., June 1975.
4. Abotteen, R. A.: Test and Evaluation of Rank Order Optimal Signature Transformation Routine (ROOSTER). LEC-8067, Lockheed Electronics Co., Houston, Tex., Apr. 1976.
5. Kauth, R. J.; and Thomas, G. S.: The Tasseled Cap: A Graphic Description of the Spectral-Temporal Development of Agricultural Crops As Seen by Landsat. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Ind., June 1976, p. 48-41.
6. Wehmanen, O. A.: The Kauth Greenness Value With Application to Signature Extension and Drought Study. LEC-8680, Lockheed Electronics Co., Houston, Tex., Sept. 1976.
7. Finley, D. R.; and Wehmanen, O. A.: Descriptions of the OSCAR Algorithm for Implementation on the Prototype Signature Extension System. LEC-8774, Lockheed Electronics Co., Houston, Tex., June 1976.
8. Thadani, S. G.: A Status Report on the Maximum Likelihood Estimation of Signature Transformation (MLEST). LEC-8300, Lockheed Electronics Co., Houston, Tex., Sept. 1976.
9. Johnson, I. L.: The Davidon-Fletcher-Powell Penalty Function Method: A Generalized Iterative Technique for Solving Parameter Optimization Problems. NASA TN D-8251, May 1976.
10. Peters, B. C., Jr.; and Walker, H. F.: An Iterative Procedure for Obtaining Maximum-Likelihood Estimates of the Parameters for a Mixture of Normal Distributions. Report 43, Dept. of Mathematics, University of Houston, July 1975.
11. Walker, H. F.: On the Numerical Evaluation of the Maximum-Likelihood Estimate of Mixture Means. Report 44, Dept. of Mathematics, University of Houston, July 1975.
12. Peters, C.; and Walker, H. F.: Addendum to "An Iterative Procedure for Obtaining Maximum-Likelihood Estimates of the Parameters for a Mixture of Normal Distributions." Report 47, Dept. of Mathematics, University of Houston, Sept. 1975.
13. Coberly, W. A.; and Wiginton, C. L.: UHMLE — Program Description User Guide. Report 48, Dept. of Mathematics, University of Houston, Oct. 1975.
14. Peters, B. C., Jr.; and Walker, H. F.: The Numerical Evaluation of the Maximum-Likelihood Estimate of a Subset of Mixture Proportions, II. Report 50, Dept. of Mathematics, University of Houston, Feb. 1976.
15. Peters, B. C., Jr.; and Walker, H. F.: An Iterative Procedure for Obtaining Maximum-Likelihood Estimates of the Parameters for a Mixture of Normal Distributions. Report 51, Dept. of Mathematics, University of Houston, Feb. 1976.
16. Coberly, W. A.: User's Guide UHMLE/RTCC Version. Report 52, Dept. of Mathematics, University of Houston, Mar. 1976.

17. Coberly, W. A.: Program Documentation SIGEXT EOD-LARSYS Version of UHMLE. Report 53, Dept. of Mathematics, University of Houston, Mar. 1976.
18. Peters, B. C., Jr.; and Walker, H. F.: An Iterative Procedure for Obtaining Maximum-Likelihood Estimates of the Parameters of a Normal Distribution. *SIAM J. Appl. Math.*, vol. 35, no. 2, Sept. 1978, pp. 362-378.
19. Abotteen, R.; Levy, S.; et al.: Performance Tests of Signature Extension Algorithms. Proceedings of the 11th International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, Mich., Apr. 25-29, 1977, p. 1523.
20. Thadani, S. G.: The Maximum Likelihood Estimates of Signature Transformation (MLEST) Algorithm. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Ind., June 21-23, 1977, p. 66.
21. Thadani, S. G.: Results of the Maximum Likelihood Estimation of Signature Transformation (MLEST) Geographical Signature Extension Test. LEC-9700, Lockheed Electronics Co., Houston, Tex., Sept. 1977.

ACKNOWLEDGMENTS

The Earth Observations Division's development, testing, and evaluation of signature extension procedures were accomplished by personnel from NASA-EOD, Lockheed Electronics Company (LEC), IBM, and the supporting research institutions. The following personnel provided the major inputs to the study of signature extension: R. A. Abotteen, LEC; K. Baker, NASA; W. A. Coberly, University of Tulsa; D. R. Finley, LEC; F. Hall, NASA; S. Levy, LEC; M. Mendlowitz, LEC; T. Moritz, LEC; B. C. Peters, Jr., University of Houston; J. Potter, LEC; S. G. Thadani, LEC; H. F. Walker, University of Houston; O. Wehmanen, LEC; and S. G. Wheeler, IBM.

Appendix

Detailed Description of OSCAR

As described in the body of this paper, the OSCAR algorithm consists of four major steps: computation of pseudorank vectors, identification of admissible pairs, evaluation of candidate transformations, and computation of weighted average estimates. The following is a detailed description of these steps.

STEP 1. COMPUTATION OF PSEUDORANK VECTORS

Cluster each segment using a suitable clustering algorithm. Let $\mu_i, i = 1, 2, \dots, M_T$, be the cluster mean vectors for the training segment and $\mu_j, j = 1, 2, \dots, M_R$, be the cluster mean vectors for the recognition segment. We will use k to index components (channels) of μ_{ik} and μ_{jk} , where $k = 1, 2, \dots, p$.

Compute for each i and k

$$v_{ik} = \frac{1}{M_T - 1} \sum_{\substack{s=1 \\ s \neq i}}^{M_T} H(\mu_{ik} - \mu_{sk}; \tau) \quad (A1)$$

where

$$z = \mu_{ik} - \mu_{sk} \quad (A2)$$

$$H(z; \tau) = 0; z < -\tau \quad (A3)$$

$$H(z; \tau) = 1; z > \tau \quad (A4)$$

$$H(z; \tau) = (z + \tau)/(2\tau); -\tau \leq z \leq \tau$$

[0 < \tau < 5; currently, \tau = 3]

(A5)

Compute for each j and k

$$w_{jk} = \frac{1}{M_R - 1} \sum_{\substack{s=1 \\ s \neq j}}^{M_R} H(\mu_{jk} - \mu_{sk}; \tau) \quad (A6)$$

where the definitions are similar to those for equation (A1).

The H function is used to render the rank vectors robust with respect to slight random variations (in haze or in signatures) that could cause rank reversals. Differences greater than τ are counted as full ranks; differences less than τ are counted as partial ranks on a sliding scale from zero to one.

The pseudorank vectors are normalized to zero-to-one to enhance comparability between segments with different numbers of clusters. Thus, if i and j index similar classes, v_{ik} should be fairly close to w_{jk} for all k .

STEP 2: IDENTIFICATION OF ADMISSIBLE PAIRS

For each i, j pair, compute

$$c_{ij} = \sum_{k=1}^p |v_{ki} - w_{kj}|^p \quad (A7)$$

$$\left[\frac{1}{2} < p < 2; \text{currently, } p = 1 \right] \quad (A8)$$

For each i , find the β lowest c_{ij} values. Let M_{ih} denote the j index of the h th lowest c_{ij} .

List all i, j pairs such that either

- a. For some $h \leq \alpha$, there is an $M_{ih} = j$.
- b. For some $\alpha < h \leq \beta$, there is an $M_{ih} = j$ and $c_{ij} < \gamma$.

$$[0 < \alpha < \beta < 10, 0.8 < \gamma < 1.3;$$

currently, $\alpha = 4, \beta = 8, \gamma = 1]$

Pairs satisfying the above criterion are called admissible. Let admissible pairs be indexed by $g = 1, 2, \dots$. Let I_g be the i index and J_g be the j index for the g th pair.

If all clusters represent classes found in both segments, corresponding \mathbf{v}_{ik} and \mathbf{w}_{jk} vectors will be nearly equal. The vectors will differ somewhat because random variation can reverse ranks and because some clusters will be unmatched (i.e., found in one segment only). The approach is to test all pairs that appear to be promising.

STEP 3: EVALUATION OF CANDIDATE TRANSFORMATIONS

The basic structure is a double loop for all (g, q) pairs such that $1 \leq q \leq N(g)$. Thus, the basic loop consists of examining all pairs of admissible pairs. The steps are

- Set the subscript r equal to 1.
- Take the next (g, q) pair. If starting, $g = 1$ and $q = 2$. If all pairs have been examined, go to step 4.
- If either $I_g = I_q$ or $J_g = J_q$ ($|I_g - I_q| \leq 0.00001$ means $I_g = I_q$), go back to step b; otherwise, continue.
- Compute for each channel k

$$s_k = \frac{\mu'_{J_g k} - \mu'_{J_q k}}{\mu_{I_g k} - \mu_{I_q k}} \quad (\text{A9})$$

- If for all k , $\delta \leq s_k \leq \lambda$, then continue; if one or more of the inequalities are not satisfied, then go back to step b.

$$[0.3 \leq \delta \leq 0.6, 1.8 \leq \lambda \leq 3; \text{ currently, } \delta = 0.6, \lambda = 1.8] \quad (\text{A10})$$

- Set and compute for all k

$$c_{kr} = s_k \quad (\text{A11})$$

$$e_{kr} = \mu'_{J_g k} - c_{kr} \mu_{I_g k} \quad (\text{A12})$$

For each recognition cluster j , compute G_j as follows:

$$z_{ijk} = c_{kr} \mu_{ik} + e_{kr} - \mu'_{jk} \quad (\text{A13})$$

Form the vector \mathbf{z}_{ij}

$$\mathbf{z}_{ij} = [z_{ij1}, z_{ij2}, z_{ij3}, \dots, z_{ijp}]^T \quad (\text{A14})$$

then

$$D_{ij} = \mathbf{z}_{ij}^T \Sigma_i^{-1} \mathbf{z}_{ij} + \text{trace} (\Sigma_i^{-1} \Sigma_j') \quad (\text{A15})$$

and define

$$G_j = \min_{i=1, M_T} (D_{ij}) \quad (\text{A16})$$

g Compute

$$f_r = \sum_i e^{-G_i / \sigma}; [20 \leq \sigma \leq 80; \text{ currently, } \sigma = 50] \quad (\text{A17})$$

- Increment r ($r = r + 1$).

- Go to step b.

The algorithm forms a straight line between the means of each pair of clusters. If the resulting multiplicative factor is reasonable, the goodness criterion is evaluated and the transformation is stored. The rationale for the goodness criterion is given in the text.

STEP 4: COMPUTATION OF WEIGHTED AVERAGE ESTIMATE

Find the ψ highest f_r values. Let f_r be the r th highest value and let c_r and e_r be moved to correspond to f_r ; [$1 \leq \psi \leq 10$; currently, $\psi = 5$].

For all channels $k = 1, 2, \dots, p$, compute a_k , the multiplicative term in the affine haze/Sun angle correction equation (eq. (1)).

$$a_k = \frac{\sum_{r=1}^{\psi} c_{kr} \left(\frac{f_r}{f_1}\right)^r}{\sum_{r=1}^{\psi} \left(\frac{f_r}{f_1}\right)^r} \quad (\text{A18})$$

For all channels $k = 1, 2, \dots, p$, compute b_k , the additive term in the affine haze/Sun angle correction equation (eq. (1)).

$$b_k = \frac{\sum_{r=1}^{\psi} e_{kr} \left(\frac{f_r}{f_1}\right)^r}{\sum_{r=1}^{\psi} \left(\frac{f_r}{f_1}\right)^r} \quad (\text{A19})$$

The weighted average is used so that the final estimate will not be unduly influenced by a slight preference that could be caused by random variation. An inspection of equations (A18) and (A19) reveals that a_k and b_k (see eq. (1)) are formed by taking weighted averages of the best candidate transformations.

TABLE I.—Sources of Variation in Landsat Data Observed by the Multispectral Scanner

Type	Examples
A. Dynamic factors	Sun angle changes Atmospheric haze level changes
B. Static sources of differences	Soil color Principal crops Physiographic features Cropping practices Climate Field sizes and shapes

TABLE II.—Affine Signature Transformation Coefficients for Consecutive-Day Data

Date	R(S)		MLEST		OSCAR		REGRES		MODR		R(C)		MODOSCAR		ATCOR		R(SIC)	
	a _k	b _k																
F1709-8	1.20	-7.1	1.06	-2.32	1.12	-4.4	1.24	-9.1	1.03	-0.9	1.12	-5.1	1.06	-2.6	1.12	-6.6	1.05	-3.6
	1.18	-5.4	1.02	-9	1.08	-3.0	1.14	-5.2	1.04	-1.7	1.09	-4.5	1.05	-2.6	1.10	-5.4	.99	-2.2
	.99	2.2	1.05	-1.4	1.01	1.1	1.15	-6.6	1.03	.1	1.03	.5	1.05	-1.8	1.08	-5.8	1.23	-10.5
	.99	1.3	1.06	-7	.98	1.5	1.12	-2.1	1.03	.3	1.02	.9	1.05	-8	1.07	-2.3	1.23	-4.2
F1673-2	0.98	0.1	1.05	0.1	0.84	4.5	0.78	5.2	0.94	1.0	0.84	—	0.89	3.4	0.95	2.8	0.99	0.2
	.99	-9	1.06	.0	.90	1.8	.82	2.4	.94	.3	.88	1.5	.98	.5	.96	2.4	1.01	-7
	1.15	-8.6	1.10	.6	1.10	-8.8	1.3	-17.6	.98	-1.2	1.04	-5.6	.95	1.6	.96	2.4	1.15	-6.3
	1.16	-5.2	1.19	.8	1.06	-3.3	1.5	-13.3	1.00	-8	1.06	-3.3	.94	1.6	.97	1.0	1.16	-4.0
F1655-4	1.06	-0.1	1.36	-8.9	1.04	-0.1	1.05	-0.2	1.05	-0.2	1.00	1.6	1.07	-1.0	0.92	4.7	1.30	-8.1
	1.08	-1	1.24	-2.3	1.03	.5	1.04	1.0	1.06	.0	1.00	1.7	1.03	.6	.94	4.3	1.26	-5.9
	.92	6.5	1.12	-8	.98	3.1	.93	5.5	1.02	.8	.81	12.1	1.04	-8	.94	3.8	1.22	-6.2
	.95	2.2	1.35	-5.5	1.00	.9	1.05	.2	1.03	.3	.89	3.9	1.04	-6	.95	1.4	1.16	-1.7
F1726-7	0.91	0.0	0.94	-1.6	0.80	0.3	0.92	-0.9	0.88	0.4	0.89	0.1	0.86	2.0	0.94	3.3	1.00	-3.6
	.93	-6	.99	-3.5	.81	.5	.94	-1.2	.89	.6	.91	-8	.88	1.9	.95	2.8	1.07	-6.1
	.91	.6	1.22	-14.8	1.02	-6.0	1.06	-8.6	.90	.2	1.03	-6.4	.87	2.1	.96	2.5	1.33	-18.7
	.95	-6	1.17	-5.5	—	.5	.98	-1.6	.91	-3	.80	3.8	.87	1.3	.97	.9	1.17	-3.9
S1455-4	1.22	-4.9	0.98	0.4	0.89	2.1	0.94	1.0	0.91	1.3	0.88	2.2	0.91	2.1	0.98	0.8	2.12	-25.3
	1.18	-3.9	.98	.1	.93	.6	1.13	-3.0	.92	.6	.90	1.2	.92	1.5	.98	.6	1.58	-11.7
	1.08	-1.9	1.02	.0	.92	1.1	.99	-2	.97	-1.3	1.00	-2.3	.92	1.2	.98	.5	1.33	-7.0
	1.14	-1.4	.97	.5	.91	.6	.98	-2	.96	-8	.98	-1.3	.92	.3	1.00	.2	1.32	-3.5
S1725-4	0.95	3.5	0.99	1.7	1.01	1.2	1.01	1.4	1.05	-0.3	0.80	8.5	1.00	1.1	0.94	3.3	1.19	-4.9
	.98	2.9	1.03	.7	1.01	1.3	1.04	.9	1.06	-1.0	.81	8.1	1.06	-7	.95	2.8	1.22	-5.5
	.97	3.4	1.04	1.2	1.01	2.2	1.02	1.8	1.04	-3	.98	2.9	1.04	.1	.96	2.5	1.31	-4.4
	1.02	.5	1.04	.6	1.00	1.1	1.00	1.0	1.05	-6	.96	1.5	1.03	.2	.97	.9	1.32	-1.7
E1726-5	1.06	-0.5	1.03	0.1	0.99	2.1	1.04	0.6	0.96	3.5	0.95	3.7	0.96	2.6	0.92	4.7	1.20	-6.4
	1.06	-3	1.02	.0	.94	4.4	1.00	2.3	.95	4.4	.93	4.8	1.06	-1.5	.94	4.3	1.17	-5.9
	1.03	.3	1.01	.0	.99	3.0	.97	3.7	.98	3.1	.99	2.9	1.03	-6	.94	3.8	1.13	-4.2
	1.08	-7	1.02	.0	1.01	1.0	1.01	1.0	1.01	1.2	1.01	1.1	1.02	.5	.95	1.4	1.15	-2.0

TABLE III.—Wheat Accuracy for Simulated Data

Data	Local accuracy, percent	Percentage difference between local accuracy and that obtained with various algorithms (a)				
		R(S)	MLEST	US fields	R(C)	UT ^b
SIM1	84.4	-2.0	6.3	-1.0	-28.4	-100
SIM2	97.1	.0	1.2	-2.9	.0	-26.3
SIM3	94.8	.2	3.5	2.8	-12.8	-84.4
SIM4	87.9	-.1	6.5	-13.5	-1.7	-28.0
Mean	91.1	-.5	4.4	-28.4	-10.7	-59.7
SD ^c	5.9	1.0	2.5	48.2	13.1	38.1

^aA minus sign means the algorithm was less accurate than local classification.
^bUT = untransformed TSSO statistics.
^cStandard deviation.

TABLE IV.—Nonwheat Accuracy for Simulated Data

Data	Local accuracy, percent	Percentage difference between local accuracy and that obtained with various algorithms (a)				
		R(S)	MLEST	UH fields	R(C)	UT
SIM1	96.4	0.5	-6.3	-0.2	-30.0	-99.1
SIM2	99.1	.0	-.5	-.1	.0	-15.8
SIM3	97.7	.0	-1.1	-2.3	-2.9	-39.5
SIM4	94.3	-.1	-6.0	-2.4	-3.2	-3.2
Mean	96.9	.1	-3.5	-1.3	-9.0	-39.4
SD	2.0	.3	3.1	1.3	14.1	42.5

^aA minus sign means the algorithm was less accurate than local classification.

TABLE V.—Wheat Accuracy for Consecutive-Day Data

Data	Local accuracy, percent	Percentage difference between local accuracy and that obtained with various algorithms (a)											
		R(S)	MLEST	OSCAR	REGRES	MOD h	R(C)	MOD OSCAR	ATCOR	UH fields	UT	R(SIC)	UH all
F1709-8	96.7	0.5	1.1	0.7	1.2	0.7	1.1	1.3	1.2	-15.6	1.1	2.0	-21.9
F1673-2	97.3	-.8	.9	-3.9	-6.6	-2.6	-3.8	-2.9	-2.9	-3.9	-.6	.2	-2.2
F1655-4	93.5	-15.5	2.7	-12.8	-11.1	-12.7	-17.8	-14.5	-19.3	-9.6	-17.8	-1.8	-49.4
F1726-7	82.6	-3.3	-6.7	-2.3	-3.0	-5.9	-10.9	-1.9	-1.1	-27.4	-2.4	-8.4	-6.4
S1455-4	92.0	1.0	1.6	-4.0	-2.7	-7.1	-11.7	-6.1	-.7	-16.8	.0	-2.1	-42.6
S1725-4	79.7	7.3	2.0	4.6	7.0	-.9	8.8	1.8	-.6	1.8	-20.2	-13.6	21.3
E1726-5	92.6	-1.9	1.1	-.3	-2.3	-1.5	.4	-.4	-6.2	-50.9	-1.0	3.6	-49.2
Mean	90.4	-1.8	.4	-2.6	-2.5	-4.3	-4.8	-3.2	-4.2	-17.5	-5.8	-2.9	-21.5
SD	6.1	6.9	3.2	5.4	5.7	4.6	9.1	5.6	7.0	17.5	9.1	6.1	27.2

^aA minus sign means the algorithm was less accurate than local classification.

TABLE VI.—Nonwheat Accuracy for Consecutive-Day Data

Data	Local accuracy, percent	Percentage difference between local accuracy and that obtained with various algorithms (a)											
		R(S)	MLEST	OSCAR	REGRES	MOD R	R(C)	MOD OSCAR	ATCOR	UH fields	UT	R(SIC)	UH all
F1709-8	73.9	-8.5	-6.8	-10.3	-10.6	-11.2	-12.0	-11.5	-12.4	10.7	-12.0	-18.7	19.9
F1673-2	95.7	-2.3	-1.0	-3.0	-11.5	1.6	-9	.1	-5.7	-27.2	.3	-2.3	-30.8
F1655-4	95.4	.4	-3.4	1.4	.7	.4	-5	.6	1.4	-9	.6	-4.4	-4.0
F1726-7	79.1	3.7	3.9	5.8	7.6	-4	2.3	3.9	-7.5	10.6	-10.5	-6.6	-6.8
S1455-4	78.9	-2.3	-5.1	-2.8	-.5	3.0	7.6	3.0	1.3	-4.6	.0	-5.8	-8.0
S1725-4	93.5	-6.5	-3.5	-7.7	-8.4	-6.1	-14.7	-5.7	-9.7	-11.8	-7.0	-8.1	-23.5
E1726-5	45.2	-5.1	-17.3	-9.3	-5.3	-2.2	-11.1	-25.2	3.1	86.1	-28.5	-31.4	61.3
Mean	80.2	-2.9	-4.7	-3.7	-2.0	-2.1	-4.2	-5.0	-4.2	9.0	-8.2	-11.0	1.2
SD	17.9	4.2	6.5	5.9	6.1	5.0	8.4	10.4	6.1	36.5	10.4	10.4	31.0

^aA minus sign means the algorithm was less accurate than local classification.

TABLE VII.—Overall Accuracy for Consecutive-Day Data

Data	Local accuracy, percent	Percentage difference between local accuracy and that obtained with various algorithms (a)											
		R(S)	MLEST	OSCAR	REGRES	MOD R	R(C)	MOD OSCAR	ATCOR	UH fields	UT	R(SIC)	UH all
F1709-8	79.5	-5.8	-4.4	-7.0	-7.1	-7.6	-8.1	-7.8	-8.5	2.7	-8.2	-12.5	7.3
F1673-2	96.1	-2.0	-.5	-3.2	-10.2	.5	-1.7	-.7	-5.0	-21.3	.1	-1.7	-23.7
F1655-4	94.9	-3.3	-1.8	-2.1	-2.1	-2.7	-4.7	-3.0	-3.6	-3.1	-3.8	-3.8	-15.0
F1726-7	80.0	1.9	1.7	3.8	4.9	-1.9	-1.1	2.4	-5.9	.9	-8.5	-7.1	-6.8
S1455-4	86.5	-.2	-.9	-3.5	-1.8	-3.2	-4.4	-2.5	.1	-12.1	.0	-3.5	-29.5
S1725-4	85.4	1.1	-.5	-.9	.0	-3.2	-1.9	-5.0	-4.7	-4.3	-14.1	-11.0	.9
E1726-5	66.2	-3.2	-6.0	-3.8	-3.5	-1.8	-4.1	-9.8	-2.7	1.4	-11.5	-9.8	-7.3
Mean	84.1	-1.6	-1.8	-2.4	-2.8	-2.8	-3.7	-3.8	-4.3	-5.1	-6.6	-7.1	-10.6
SD	10.2	2.7	2.6	3.3	4.9	2.5	2.4	4.2	2.7	8.7	5.5	4.2	13.1

^aA minus sign means the algorithm was less accurate than local classification.

TABLE VIII.—Overall Accuracy for Simulated Data

Data	Local accuracy, percent	Percentage difference between local accuracy and that obtained with various algorithms (a)				
		R(S)	MLEST	UH fields	R(C)	UT
SIM1	93.5	0.0	-3.5	-21.7	-29.6	-99.3
SIM2	98.6	.0	.0	-.7	.0	-18.3
SIM3	97.0	.1	.0	-1.0	-5.2	-50.0
SIM4	92.8	-.1	-3.2	-5.0	-2.9	-8.8
Mean	95.5	.0	-1.7	-7.1	-9.4	-44.1
SD	2.8	.1	1.9	9.9	13.6	40.8

^aA minus sign means the algorithm was less accurate than local classification.

TABLE IX.—Wheat Proportions for Simulated Data as Determined Using Local Results

Data	Local	Signature extension proportion minus local proportion					
		R(S)	R(C)	MLEST	UH fields MLE	UH fields	UT
SIM1	24.3	-0.5	-1.9	0.6	-16.8	-22.3	-24.3
SIM2	24.7	.0	.0	.6	-1.0	-1.1	-3.2
SIM3	24.9	.1	-3.1	1.5	1.7	1.6	-20.0
SIM4	24.2	.0	-.2	6.8	-.3	-.6	-6.3
Mean absolute values	—	.2	1.3	2.4	5.0	6.4	13.5
SD	—	.2	1.5	3.0	7.9	10.6	10.3

TABLE X.—Wheat Proportions for Simulated Data as Determined Using Ground Truth

Data	Ground truth	Signature extension proportion minus ground-truth proportion					
		R(S)	R(C)	MLEST	UH fields MLE	UH fields	UT
SIM1	23.9	-0.1	-1.5	1.0	-16.4	-21.9	-23.9
SIM2	23.9	.8	.8	1.4	-.2	-.3	-2.4
SIM3	23.9	1.1	-2.1	2.5	2.7	2.6	-19.0
SIM4	23.9	.3	.1	7.1	-.3	-.3	-6.0
Mean absolute values	—	.6	1.1	3.0	4.9	6.3	12.8
SD	—	.5	.9	2.8	7.8	10.5	10.3

TABLE XI.—Wheat Proportions for Consecutive-Day Data as Determined Using Local Results

Data	Local accuracy, percent	Signature extension proportion minus local proportion												
		R(S)	REGRES	OSCAR	MODR	UT	MLEST	ATCOR	MOD OSCAR	R(SIC)	R(C)	UH all	UH fields	UH all MLE
F1709-8	35.4	5.9	7.5	9.0	9.4	9.8	7.8	10.7	10.3	16.5	10.3	-8.9	-8.5	-5.2
F1673-2	28.9	3.1	-2.0	.6	-2.0	1.9	5.0	-1.6	-1.9	4.4	-1.5	20.7	13.3	22.5
F1655-4	27.7	-2.2	-2.2	-2.1	-1.6	-4.7	3.9	-4.9	-2.8	2.8	-3.8	1.6	4.6	5.8
F1726-7	28.8	-1.0	-2.0	-1.4	-1.4	3.1	-4	2.6	-1.4	1.9	-2.6	-2.4	-13.7	-2.8
S1455-4	53.7	.5	-3.3	-3.5	-8.6	-2.0	4.8	-3.0	-7.6	-3	-13.3	-11.4	-10.6	-8.3
S1725-4	35.3	4.6	5.0	5.3	2.9	-8	2.6	4.4	3.6	-9	9.6	19.6	5.5	20.7
E1726-5	61.9	1.8	1.4	3.3	.9	6.6	4.9	-3.2	6.0	9.8	3.6	-25.1	-36.1	-29.7
Mean absolute values	—	2.7	3.3	3.6	3.8	4.1	4.2	4.3	4.8	5.2	6.4	12.8	13.2	13.6
SD	—	1.9	2.2	2.9	3.6	3.2	2.3	3.0	3.3	5.9	4.6	9.2	10.7	10.6

TABLE XII.—Wheat Proportions for Consecutive-Day Data as Determined Using Ground Truth

Data	Ground truth	Signature extension proportion minus ground-truth proportion												
		R(S)	REGRES	OSCAR	MODR	UT	MLEST	ATCOR	MOD OSCAR	R(SIC)	R(C)	UH all	UH fields	UH all MLE
F1709-8	24.6	16.7	18.3	19.8	20.2	20.6	18.6	21.5	21.1	27.3	21.1	1.9	2.3	5.6
F1673-2	24.6	7.4	2.3	4.9	2.3	6.2	9.3	2.7	2.4	8.7	2.8	25.0	17.6	26.8
F1655-4	24.6	.9	.9	1.0	1.5	-1.6	7.0	-1.8	.3	5.9	-.7	4.7	7.7	8.9
F1726-7	24.6	3.2	2.2	2.8	2.8	7.3	3.8	6.8	2.8	6.1	1.6	1.8	-9.5	1.4
S1455-4	58.3	-4.1	-7.9	-8.1	-13.2	-6.6	.2	-7.6	-12.2	-4.9	-17.9	-16.0	-15.2	-12.9
S1725-4	58.3	-18.4	-18.0	-17.7	-20.1	-23.8	-20.4	-18.7	-19.4	-23.9	-13.4	-3.4	-17.5	-2.3
E1726-5	44.2	19.5	19.1	21.0	18.6	24.3	22.6	14.5	23.7	27.5	21.3	-7.4	-18.4	-12.0
Mean absolute values	—	10.0	9.8	10.8	11.2	12.9	11.7	10.5	11.7	14.9	11.3	8.6	12.6	10.0
SD	—	7.9	8.4	8.5	8.8	9.6	8.8	7.8	9.9	10.7	9.3	8.7	6.2	8.7

TABLE XIII.—Haze Conditions on Consecutive-Day Data as Determined by Inspection of Images

<i>Data</i>	<i>TSEG</i>	<i>RSEG</i>
F1709-8	Clear	Clear
F1673-2	Haze	Clear
F1655-4	Clear	Clear
F1726-7	Haze	Clear
S1455-4	Clear	Clear
S1725-4	Clear	Haze
E1726-5	Clear	Haze

TABLE XIV.—Analysis of Variance for Overall Accuracy

<i>Source</i>	<i>Degrees of freedom</i>	<i>Sum of squares</i>	<i>Mean square</i>	<i>F-factor</i>	<i>Significance</i>
Algorithm (A)	9	217.61	24.18	2.02	6 or 7 percent
Haze (H)	2	113.69	56.85	4.74	5 percent
A × H	18	205.32	11.41	.95	NS ^a
Error	40	480.17	12.00		
Total	69	1 016.79			

^aNot significant.

TABLE XV.—Accuracy Percentage for the Three Different Haze Conditions

<i>Haze condition</i>		<i>Percent accuracy difference</i>
<i>TSEG</i>	<i>RSEG</i>	
Haze	Clear	- 1.71
Clear	Clear	- 4.26
Clear	Haze	- 4.82

TABLE XVI.—Analysis of Variance for Wheat Proportions

<i>Source</i>	<i>Degrees of freedom</i>	<i>Sum of squares</i>	<i>Mean square</i>	<i>F-factor</i>	<i>Significance</i>
Algorithm (A)	8	59.77	7.47	X	NS
Haze (H)	2	156.27	78.14	9.42	1 percent
A × H	16	45.21	2.83	X	NS
Error	36	312.95	8.69		
Total	62	574.20			

TABLE XVII.—Average Proportion Differences for the Three Different Haze Conditions

<i>Haze condition</i>		<i>Proportion difference</i>
<i>TSEG</i>	<i>RSEG</i>	
Haze	Clear	2.0
Clear	Clear	5.8
Clear	Haze	3.9

TABLE XVIII.—Summary of Test Results

Simulated data		Consecutive-day data	
<i>Percentage difference between local accuracy and that obtained with various algorithms</i>			
R(S)	0.0	R(S)	-1.6
MLEST	-1.7	MLEST	-1.8
UH fields	-7.1	OSCAR	-2.4
R(C)	-9.4	REGRES	-2.8
UT	-44.1	MOD R	-2.8
		R(C)	-3.7
		MOD OSCAR	-3.8
		ATCOR	-4.3
		UH fields	-5.1
		UT	-6.6
		R(S/C)	-7.1
		UH all	-10.6
<i>Wheat proportion difference from local</i>			
R(S)	0.2	R(S)	2.7
R(C)	1.3	REGRES	3.3
MLEST	2.4	OSCAR	3.6
UH fields MLE	5.0	MOD R	3.8
UH fields	6.4	UT	4.1
UT	13.5	MLEST	4.2
		ATCOR	4.3
		MOD OSCAR	4.8
		R(S/C)	5.2
		R(C)	6.4
		UH all	12.8
		UH fields	13.2
		UH all MLE	13.6
<i>Wheat proportion difference from ground truth</i>			
R(S)	0.8	UH all	8.6
R(C)	1.1	REGRES	9.8
MLEST	3.0	UH all MLE	10.0
UH fields MLE	4.9	R(S)	10.0
UH fields	6.3	ATCOR	10.5
UT	12.8	OSCAR	10.8
		MOD R	11.2
		R(C)	11.3
		MLEST	11.7
		MOD OSCAR	11.7
		UH fields	12.6
		UT	12.9
		R(S/C)	14.9

TABLE XIX.—Estimated Average Percentage of Wheat and Thresholding in All Recognition Segments Used in the Study

Classification technique	Percentage of wheat in segment					
	0-percent thresholding		1-percent thresholding		Percent thresholding at 1 percent	
	Estimate	SE ^a	Estimate	SE	Estimate	SE
Local	30.9	2.8	29.9	2.6	3.4	0.64
Untransformed	31.6	2.7	27.9	3.0	13.8	4.2
OSCAR with clusters	34.2	2.8	32.8	2.7	4.2	.89
ATCOR	33.7	2.7	31.7	2.7	6.9	1.8
OSCAR with subclasses	32.2	2.8	31.5	2.8	3.8	.7

^aStandard error.

TABLE XX.—Average Values and Standard Deviations of the Estimated Percentage of Wheat at 0-Percent Thresholding

Classification technique	Biowindow				Average
	1	2	3	4	
<i>Average values</i>					
Local	33.4	25.0	31.6	36.7	31.5
Untransformed	35.0	31.2	27.2	35.1	31.9
OSCAR with clusters	37.4	33.2	36.6	32.2	34.6
ATCOR	36.8	29.3	35.0	36.0	34.1
OSCAR with subclasses	33.7	32.2	30.9	32.5	32.2
Average	35.2	30.2	32.2	34.5	32.9
<i>Observed standard deviations</i>					
Local	15.8	7.8	15.0	17.4	
Untransformed	18.8	17.1	9.9	13.0	
OSCAR with clusters	17.7	18.5	12.5	12.9	
ATCOR	17.4	17.8	9.6	13.4	
OSCAR with subclasses	15.4	15.8	8.5	19.5	
<i>Number of segments</i>					
All techniques	5	7	7	7	(26) ^a

^aParentheses indicate total.

TABLE XXI.—Average Values and Standard Deviations of the Estimated Percentage of Wheat at 1-Percent Thresholding

Classification technique	Blowindow				Average
	1	2	3	4	
<i>Average values</i>					
Local	33.0	23.8	30.6	35.3	30.5
Untransformed	31.1	26.4	22.7	34.2	28.4
OSCAR with clusters	36.3	31.6	34.7	31.5	33.3
ATCOR	33.8	27.8	32.5	35.1	32.2
OSCAR with subclasses	32.6	32.7	29.3	31.8	31.5
Average	33.4	28.4	30.0	33.6	31.2
<i>Observed standard deviations</i>					
Local	15.6	7.1	14.3	16.0	
Untransformed	21.0	20.4	8.2	13.1	
OSCAR with clusters	17.7	18.3	11.3	12.8	
ATCOR	18.4	17.8	8.8	13.5	
OSCAR with subclasses	15.5	16.2	7.3	19.2	
<i>Number of segments</i>					
All techniques	5	7	7	7	(26) ^a

^aParentheses indicate total.

TABLE XXII.—Average Values and Standard Deviations of Percentage of Thresholding in the Segments at 1-Percent Thresholding

Classification technique	Blowindow				Average
	1	2	3	4	
<i>Average values</i>					
Local	1.5	5.7	2.1	3.8	3.4
Untransformed	14.0	21.3	14.3	4.3	13.4
OSCAR with clusters	3.4	3.9	4.5	4.5	4.1
ATCOR	7.5	4.6	10.0	3.6	6.3
OSCAR with subclasses	3.8	3.2	4.0	4.3	3.8
Average	6.1	7.7	7.0	4.1	6.2
<i>Observed standard deviations</i>					
Local	1.1	3.2	1.3	4.8	
Untransformed	25.6	31.5	20.7	2.0	
OSCAR with clusters	5.1	5.9	4.8	4.2	
ATCOR	12.0	3.7	14.5	1.5	
OSCAR with subclasses	5.4	4.1	4.3	2.3	
<i>Number of segments</i>					
All techniques	5	7	7	7	(26) ^a

^aParentheses indicate total.

TABLE XXIII.—Correlation Coefficients

Estimate	Change in haze as predicted by ATCOR	OSCAR quality factors	
		Cluster-based	Subclass-based
<i>Percentage of wheat in the segment at 1 percent</i>			
Local	-0.05	0.06	-0.42
Untransformed	.35	.25	-.14
OSCAR with clusters	.16	.14	-.13
ATCOR	^a .47	.23	-.04
OSCAR with subclasses	.07	.14	-.01
<i>Differences between signature extension and local percentage of wheat at 1 percent</i>			
Untransformed	0.37	0.19	0.21
OSCAR with clusters	.19	.07	.24
ATCOR	^a .48	.16	.34
OSCAR with subclasses	.09	.10	.34
<i>Thresholding rates at 1 percent</i>			
Local	-0.17	0.09	0.02
Untransformed	^a -.83	-.34	-.16
OSCAR with clusters	-.22	-.19	^a -.61
ATCOR	^a -.50	-.14	^a -.62
OSCAR with subclasses	-.30	-.07	^a -.76
<i>Difference between signature extension and local thresholding rates</i>			
Untransformed	^a -0.81	-0.35	-0.16
OSCAR with clusters	-.08	-.21	-.51
ATCOR	-.40	-.16	^a -.57
OSCAR with subclasses	-.12	-.12	^a -.59

^aSignificant at the 0.01 level.

TABLE XXIV.—Average Wheat Training Field Accuracies and Standard Deviations at 0-Percent Thresholding^a

Classification technique	Blowdown				Average
	1	2	3	4	
<i>Average values</i>					
Local	85.8	81.2	70.9	87.6	81.2
Untransformed	74.9	57.4	33.0	57.0	54.9
OSCAR with clusters	80.0	52.5	43.8	50.4	55.7
ATCOR	80.6	52.5	38.5	58.7	56.6
OSCAR with subclasses	72.7	57.4	34.1	58.8	55.2
Average	78.8	60.2	44.0	62.5	60.7
<i>Observed standard deviations</i>					
Local	14.5	14.8	6.9	8.7	
Untransformed	19.6	34.9	19.0	33.4	
OSCAR with clusters	19.6	40.5	25.9	32.3	
ATCOR	20.1	33.6	22.3	33.1	
OSCAR with subclasses	25.2	31.7	16.8	30.9	
<i>Number of segments</i>					
All techniques	6	8	7	7	(28) ^b

^aAll segments used in the study.

^bParentheses indicate total.

TABLE XXV.—Average Nonwheat Training Field Accuracies and Standard Deviations at 0-Percent Thresholding^a

Classification technique	Blowdown				Average
	1	2	3	4	
<i>Average values</i>					
Local	92.0	88.7	86.7	95.7	90.7
Untransformed	84.0	79.4	78.5	72.9	78.5
OSCAR with clusters	83.3	78.5	67.9	73.9	75.7
ATCOR	84.2	82.9	67.1	73.7	76.9
OSCAR with subclasses	84.8	78.3	71.2	83.1	79.1
Average	85.7	81.6	74.3	79.8	80.2
<i>Observed standard deviations</i>					
Local	5.7	5.2	7.2	3.4	
Untransformed	17.0	14.0	11.7	15.9	
OSCAR with clusters	16.4	16.9	14.5	20.4	
ATCOR	15.9	14.6	20.1	16.7	
OSCAR with subclasses	15.2	14.1	10.9	11.6	
<i>Number of segments</i>					
All techniques	6	8	7	7	(28) ^b

^aAll segments used in the study.
^bParentheses indicate total.

TABLE XXVI.—Average Wheat Training Field Accuracies and Standard Deviations at 1-Percent Thresholding^a

Classification technique	Blowdown				Average
	1	2	3	4	
<i>Average values</i>					
Local	85.5	80.8	70.4	87.2	80.8
Untransformed	68.2	50.9	26.7	56.3	49.9
OSCAR with clusters	78.0	49.4	37.2	49.8	52.6
ATCOR	76.8	49.4	34.1	58.1	53.6
OSCAR with subclasses	71.3	54.0	29.1	58.4	52.6
Average	76.0	56.9	39.5	62.0	57.9
<i>Observed standard deviations</i>					
Local	14.3	14.8	7.1	8.5	
Untransformed	27.9	36.0	17.3	33.2	
OSCAR with clusters	20.8	37.2	20.0	32.1	
ATCOR	24.7	31.8	18.2	32.8	
OSCAR with subclasses	26.2	29.0	15.0	30.8	
<i>Number of segments</i>					
All techniques	6	8	7	7	(28) ^b

^aAll segments used in the study.
^bParentheses indicate total.

TABLE XXVII.—Average Nonwheat Training Field Accuracies and Standard Deviations at 1-Percent Thresholding^a

Classification technique	Blowndow				Average
	1	2	3	4	
<i>Average values</i>					
Local	91.8	86.5	86.2	94.8	89.6
Untransformed	74.5	62.3	67.3	66.1	67.1
OSCAR with clusters	81.1	73.6	64.4	68.4	71.6
ATCOR	80.8	76.9	60.0	68.2	71.3
OSCAR with subclasses	82.8	73.0	67.1	76.9	74.6
Average	82.2	74.4	69.0	74.9	74.9
<i>Observed standard deviations</i>					
Local	5.8	9.4	7.5	3.3	
Untransformed	23.2	19.1	20.0	20.0	
OSCAR with clusters	16.6	17.4	14.6	23.4	
ATCOR	16.5	12.5	20.3	20.2	
OSCAR with subclasses	15.5	10.3	13.6	11.7	
<i>Number of segments</i>					
All techniques	6	8	7	7	(28) ^b

^aAll segments used in the study.
^bParentheses indicate total.

TABLE XXVIII.—Thresholding in Wheat Training Fields and Standard Deviations at 1-Percent Thresholding^a

Classification technique	Blowndow				Average
	1	2	3	4	
<i>Average values</i>					
Local	0.42	0.49	0.54	0.52	0.50
Untransformed	11.7	29.8	16.4	2.7	15.8
OSCAR with clusters	3.6	5.4	7.9	2.3	4.9
ATCOR	7.5	7.5	8.0	2.4	6.3
OSCAR with subclasses	3.7	5.0	8.4	2.0	4.8
Average	5.4	9.6	8.3	2.0	6.5
<i>Observed standard deviations</i>					
Local	0.38	0.28	0.51	0.28	
Untransformed	20.9	42.2	25.6	1.9	
OSCAR with clusters	4.9	6.3	13.2	2.2	
ATCOR	13.2	9.2	10.5	1.5	
OSCAR with subclasses	5.4	8.7	18.3	1.5	
<i>Number of segments</i>					
All techniques	6	8	7	7	(28) ^b

^aAll segments used in the study.
^bParentheses indicate total.

TABLE XXIX.—Thresholding in Nonwheat Training Fields and Standard Deviations at 1-Percent Thresholding^a

Classification technique	Blowdown				Average
	1	2	3	4	
<i>Average values</i>					
Local	0.31	0.53	0.57	1.0	0.62
Untransformed	10.7	21.3	13.0	9.7	14.0
OSCAR with clusters	2.7	7.8	4.2	8.7	6.0
ATCOR	4.6	6.7	9.6	8.5	7.4
OSCAR with subclasses	2.5	4.8	5.1	7.1	4.9
Average	4.1	8.2	6.5	7.0	6.6
<i>Observed standard deviations</i>					
Local	0.22	0.32	0.72	1.5	
Untransformed	20.3	27.6	18.9	8.2	
OSCAR with clusters	3.3	11.5	3.6	7.9	
ATCOR	6.8	8.9	13.1	7.7	
OSCAR with subclasses	3.3	9.7	5.6	5.1	
<i>Number of segments</i>					
All techniques	6	8	7	7	(28) ^b

^aAll segments used in the study.

^bParentheses indicate total.

TABLE XXX.—Classification Accuracy Results for the Geographical Extension Data Set

Segment pair	Wheat accuracy, percent			Nonwheat accuracy, percent			Overall accuracy, percent		
	UT	MLEST	LOCAL	UT	MLEST	LOCAL	UT	MLEST	LOCAL
<i>Blowdown 1</i>									
1854/1025	16.17	53.36	77.98	42.68	85.69	83.95	29.43	69.52	80.97
1031/1025	73.76	91.37	98.82	95.15	98.94	96.47	84.45	95.16	97.64
1176/1170	91.71	93.36	91.04	57.61	62.20	83.73	74.66	77.78	87.39
1889/1033	94.55	94.29	90.65	58.94	56.81	78.17	76.74	75.55	84.41
1169/1033	68.44	79.11	90.96	89.95	82.89	92.44	79.20	81.00	91.70
1168/1173	62.22	50.20	97.46	43.51	73.21	95.33	52.86	61.70	90.40
1174/1033	80.46	90.11	98.75	98.33	97.89	98.18	89.40	94.00	98.47
<i>Blowdown 2</i>									
1883/1881	94.13	95.24	87.30	65.54	70.79	89.60	79.83	83.02	88.45
1864/1025	70.89	73.39	93.07	83.22	81.20	89.71	77.06	77.29	91.39
1882/1887	0	64.29	87.30	19.63	66.00	89.60	9.81	65.15	88.45
1893/1891	53.60	53.60	82.53	68.84	70.55	84.13	61.22	62.07	83.33
1153/1875	74.12	82.43	80.48	62.69	56.02	70.80	68.51	69.22	75.64
1890/1887	.28	23.20	94.20	58.39	91.41	88.40	29.33	57.30	91.30
1178/1180	52.80	54.44	89.95	52.26	63.50	75.12	52.53	58.97	82.54
<i>Blowdown 3</i>									
1854/1852	50.97	47.15	80.87	84.45	85.11	79.90	67.71	66.13	80.38
1877/1875	28.41	41.47	67.97	71.99	87.38	77.92	50.20	64.43	72.95
1880/1875	22.65	74.59	75.14	40.65	74.08	93.32	31.65	74.33	84.23
1163/1165	84.21	76.61	88.30	30.22	67.14	61.46	57.22	71.87	74.88
1178/1165	66.82	75.93	91.82	52.16	52.83	89.82	59.49	64.38	90.82
1172/1181	16.61	42.08	64.04	43.49	45.15	75.90	30.05	43.61	69.97
<i>Blowdown 4</i>									
1859/1861	82.21	83.50	93.83	56.64	69.13	87.99	69.43	76.32	90.51
1032/1861	56.76	69.23	86.49	74.45	73.68	92.24	65.60	71.46	89.37
1031/1027	6.38	6.15	89.48	44.04	46.95	87.38	25.21	26.55	88.43
1892/1885	53.97	52.98	97.02	78.06	76.18	97.53	66.01	64.58	97.28
1883/1884	35.48	62.32	98.71	34.92	47.09	99.47	35.20	54.71	99.09
1888/1879	92.31	89.35	95.46	96.14	95.64	93.77	94.22	92.50	94.62
1176/1177	92.32	94.09	89.34	70.43	64.50	86.89	81.38	79.29	88.11

CA

**TABLE XXXI.—MLEST Classification Performance
Versus Blowdown for the Geographical
Extension Data Set**

Criterion	Blowdown				Overall average
	1	2	3	4	
<i>Average improvement, percent</i>					
Overall accuracy	9.71	13.53	14.74	4.05	10.35
Wheat accuracy	9.21	14.40	14.69	5.46	10.80
Nonwheat accuracy	10.21	12.70	14.79	2.64	9.91
<i>Average difference, percent</i>					
Overall accuracy	11.75	18.30	14.75	26.06	17.82
Wheat accuracy	13.41	24.05	18.39	27.53	20.94
Nonwheat accuracy	10.09	12.55	11.33	24.59	14.76

TABLE XXXII.—Proportion Estimation Results for the Geographical Extension Data Set

Segment pair	Wheat proportions, percent						Pixels thresholded, percent		
	Before thresholding			After thresholding			UT	MLEST	LOCAL
	UT	MLEST	LOCAL	UT	MLEST	LOCAL			
<i>Blowdown 1</i>									
1854/1025	27.0	38.7	47.8	10.1	37.4	47.4	59.5	4.0	0.7
1031/1025	4.7	6.5	10.1	4.4	6.3	9.8	8.8	2.5	1.3
1176/1170	61.0	58.9	44.7	59.1	58.2	44.4	7.4	2.6	1.7
1889/1033	41.6	43.6	32.8	41.3	43.4	32.7	.9	.5	.4
1169/1033	13.0	19.3	15.2	12.9	19.1	14.8	.8	1.3	2.0
1168/1173	21.8	14.1	15.2	15.4	9.8	14.6	23.7	9.6	3.0
1174/1033	38.8	49.4	49.7	37.7	48.9	48.7	1.4	1.0	3.0
<i>Blowdown 2</i>									
1882/1881	56.9	50.7	34.0	54.2	49.0	33.0	4.0	2.7	5.1
1864/1025	29.1	28.2	36.0	23.8	23.5	32.5	8.9	9.6	6.5
1882/1887	50.5	48.1	34.0	.1	47.9	33.0	86.9	1.3	5.1
1893/1891	16.9	19.2	45.1	16.3	18.6	41.8	1.2	.8	3.4
1153/1875	44.7	52.3	45.0	44.5	52.1	44.5	2.5	1.5	1.5
1880/1887	26.0	18.7	19.0	6.0	17.9	18.0	41.4	5.0	8.3
1178/1180	40.3	44.0	29.3	36.6	43.7	29.2	6.5	.8	2.3
<i>Blowdown 3</i>									
1854/1852	33.5	30.3	48.2	31.7	29.5	47.9	3.2	2.1	0.7
1877/1875	29.3	18.0	54.4	27.0	17.3	51.0	4.5	3.4	4.1
1880/1875	33.3	31.9	26.5	24.8	29.7	25.6	39.3	6.1	3.0
1163/1165	65.5	32.3	40.8	65.4	32.2	40.4	1.0	1.4	1.4
1178/1165	43.6	45.2	32.8	43.5	45.1	31.1	.4	.5	2.9
1172/1181	34.0	52.8	53.7	17.8	51.2	42.7	47.6	2.1	1.3
<i>Blowdown 4</i>									
1859/1861	33.6	41.5	31.2	32.5	41.0	30.3	6.8	1.4	2.5
1032/1861	39.9	45.5	39.8	39.4	44.7	39.6	2.3	2.4	.4
1031/1027	10.4	10.9	19.4	9.3	9.2	19.1	6.9	4.4	2.8
1892/1885	34.4	34.6	30.4	32.4	33.2	29.0	3.8	2.8	1.7
1883/1884	48.1	53.0	62.5	46.9	52.0	56.9	2.9	2.5	14.1
1888/1879	52.2	48.1	62.3	50.9	46.4	58.7	5.1	5.7	5.4
1176/1177	47.4	53.6	40.7	46.7	52.8	40.4	2.5	2.3	1.9

TABLE XXXIII.—Mean Wheat Proportion Estimate Differences Versus Blowdown for the Geographical Extension Data Set

Blowdown	0-percent thresholding		1-percent thresholding	
	$ Q_{LOCAL} - Q_{UT} $	$ Q_{LOCAL} - Q_{MLEST} $	$ Q_{LOCAL} - Q_{UT} $	$ Q_{LOCAL} - Q_{MLEST} $
1	10.14	6.17	11.39	6.76
2	13.26	12.4	15.39	12.19
3	16.97	13.98	18.88	13.32
4	6.67	9.33	5.67	8.50
Overall	11.57	10.25	12.61	10.08

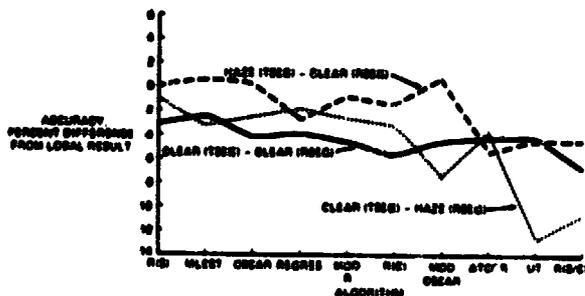


FIGURE 1.—Base-by-algorithm interaction showing overall accuracy difference for the three base conditions.

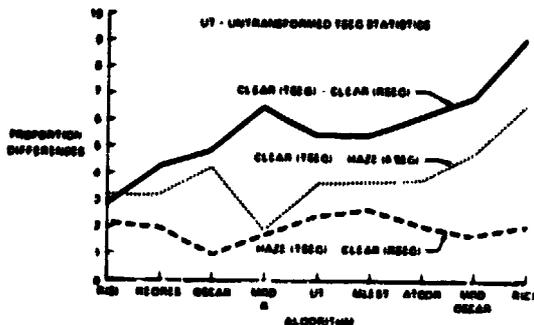


FIGURE 2.—Base-by-algorithm interaction showing proportion differences for the three base conditions.

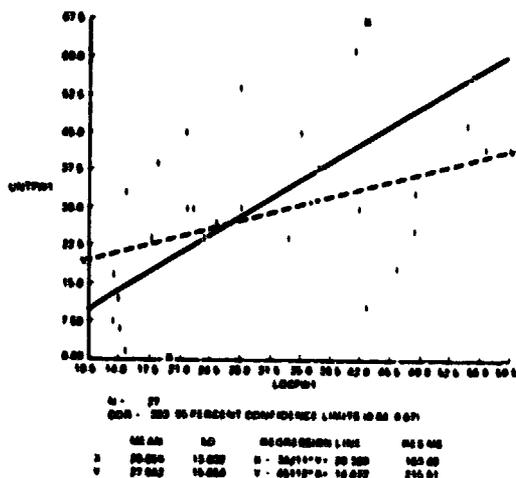
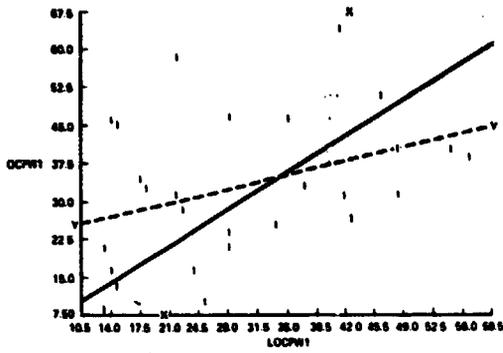


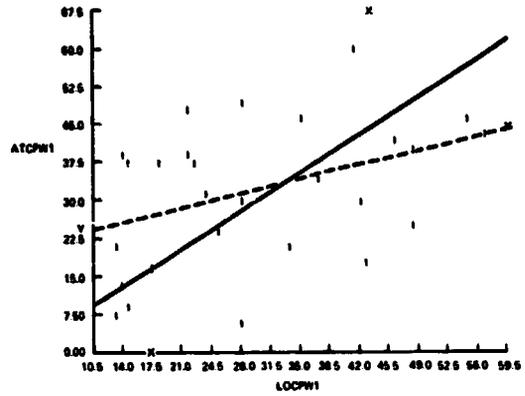
FIGURE 3.—Untransformed versus local estimate of percentage wheat in the segment at 1-percent thresholding.



N = 27
COR = 363 95 PERCENT CONFIDENCE LIMITS (0 03, 0 68)

	MEAN	SD	REGRESSION LINE	RES MS
X	29.889	12.632	$X = 36777 \cdot Y + 16.043$	167.76
Y	22.827	14.228	$Y = 37927 \cdot X + 21.500$	163.01

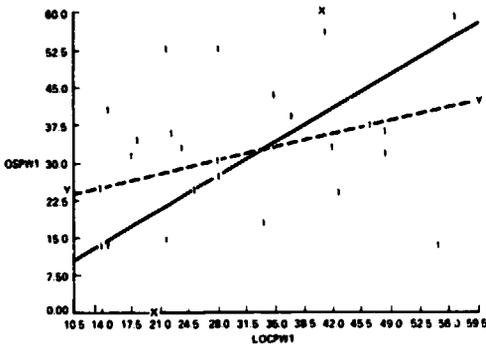
FIGURE 4.—OSCAR with clusters versus local percentage of wheat in the segment at 1-percent thresholding.



N = 27
COR = 401 95 PERCENT CONFIDENCE LIMITS (0 02, 0 68)

	MEAN	SD	REGRESSION LINE	RES MS
X	29.889	12.632	$X = 28895 \cdot Y + 17.515$	162.12
Y	31.726	14.070	$Y = 41433 \cdot X + 19.365$	172.70

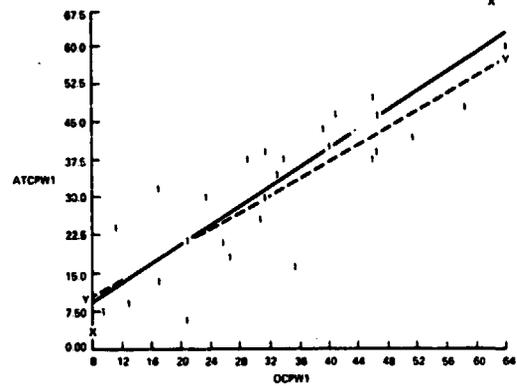
FIGURE 5.—ATCOR versus local percentage of wheat in the segment at 1-percent thresholding.



N = 28
COR = 326 95 PERCENT CONFIDENCE LIMITS (0 05, 0 67)

	MEAN	SD	REGRESSION LINE	RES MS
X	30.488	12.486	$X = 31606 \cdot Y + 20.526$	168.48
Y	31.520	14.289	$Y = 35430 \cdot X + 20.718$	168.87

FIGURE 6.—OSCAR with subclasses versus local percentage of wheat in the segment at 1-percent thresholding.



N = 27
COR = 843 95 PERCENT CONFIDENCE LIMITS (0 48, 0 94)

	MEAN	SD	REGRESSION LINE	RES MS
X	32.827	14.228	$X = 86279 \cdot Y + 5.7565$	61.062
Y	31.736	14.070	$Y = 83288 \cdot X + 4.3950$	59.628

FIGURE 7.—ATCOR versus OSCAR with clusters percentage of wheat at 1-percent thresholding.

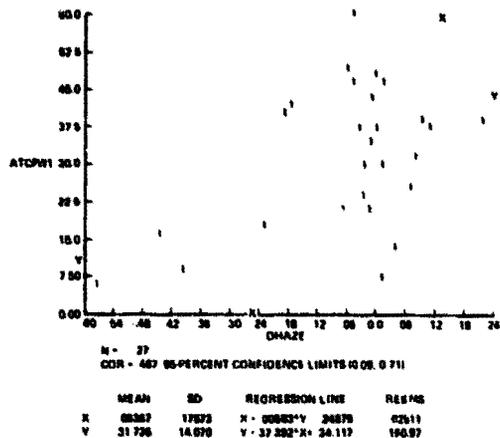


FIGURE 8.—ATCOR percentage of wheat in the segment at 1-percent thresholding versus change in haze levels.

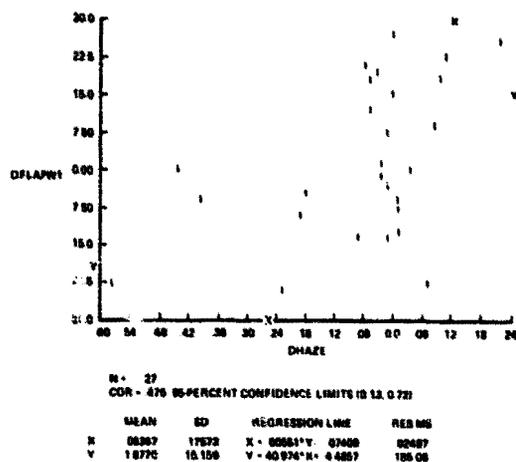


FIGURE 9.—Differences between ATCOR and local percentage of wheat in the segment at 1-percent thresholding versus change in haze levels.

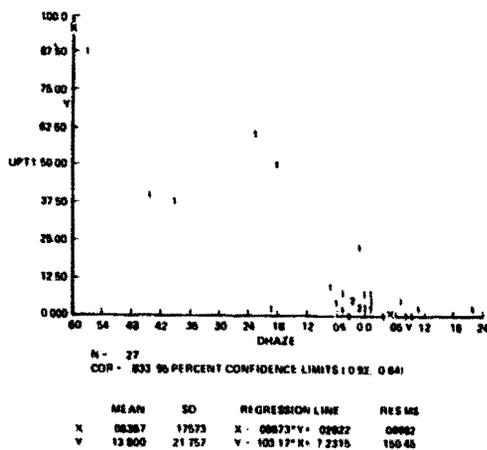
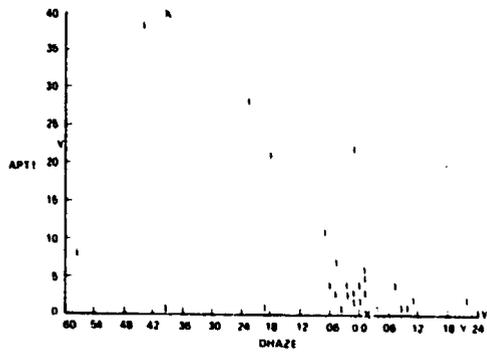


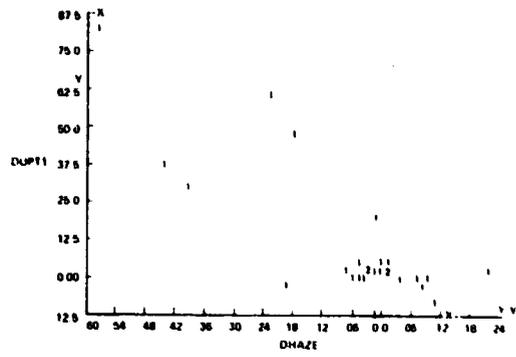
FIGURE 10.—Untransformed thresholding rate at 1-percent thresholding versus change in haze levels.



N = 27
 COR = 900 95 PERCENT CONFIDENCE LIMITS (0 73 0 141)

	MEAN	SD	REGRESSION LINE	RES MS
X	08.367	175.73	X - 00830*Y + 663E 6	02408
Y	8 9203	9 4826	Y - 26 896*X + 5.2080	08 663

FIGURE 11.—ATCOR thresholding rates at 1-percent thresholding versus change in haze levels.



N = 27
 COR = 807 95 PERCENT CONFIDENCE LIMITS (0 91 0 680)

	MEAN	SD	REGRESSION LINE	RES MS
X	08.367	175.73	X - 00851*Y + 00406	01121
Y	10 400	21 780	Y - 99 998*X + 4 0333	172 18

FIGURE 12.—Differences between untransformed and local thresholding rates at 1-percent thresholding versus change in haze levels.

Signature Extension Methods in Crop Area Estimation

R. J. Kauth^a and W. Richardson^a

INTRODUCTION

An encompassing rationale for crop area estimation over large regions using remotely sensed data has been developed over a period of several years as a result of the stimulation of LACIE and various supporting research institutions. Currently, most elements of the general idea have been implemented at the Environmental Research Institute of Michigan (ERIM) in a multispectral scanner (MSS) processing system called Procedure B. This paper describes this general idea, shows how various research efforts have contributed to its development, and indicates the elements of the rationale that are implemented in Procedure B.

During the last decade, remote-sensing specialists have witnessed a convergence of several disciplines in the development of MSS data processing techniques. The first techniques used in the classification of MSS data were methods of multivariate pattern recognition using a Gaussian signature model (ref. 1). Over this period of time, emphasis has been placed on trying to understand the underlying physical reasons for the structure of remotely sensed multispectral data (refs. 2 through 9), on compensating for interfering external effects (ref. 10), and on trying to apply MSS "signatures" obtained by training in one area to a wider region (ref. 11). Finally, there has been a series of statistically based attempts to estimate areal portions or acreages based directly upon the signatures and the multispectral data (i.e., without creating classification maps) (ref. 12).

The first viable attempt to combine a traditional pattern recognition approach and a statistically based stratified sampling approach was performed for wheat versus nonwheat at the NASA Johnson Space

Center (JSC) during the LACIE program (ref. 13). The resulting technique is called Procedure 1 (ref. 14). In this procedure, a classification map is produced by a clustering technique based on labeled samples. Next, the areal proportion estimate represented by that classification map is "debiased" by using additional labeled samples to estimate the percentage of correct classifications in each of the two mapped classes. In fact, this "bias correction" step in Procedure 1 is precisely equivalent to stratified sampling from the two classes or "strata" created by the operation of the classifier. The very existence of Procedure 1 is forcing a fundamental change in the research community's understanding of concepts that have been taken for granted: classifiers, signatures, and estimation of acreage (or yield).

During this same period, a sequence of improvements in understanding the physical structure of multispectral data has led to the development of effective automated procedures for screening Landsat multispectral data (to identify garbled data or pixels that are cloud, cloud shadow, water, etc.), for correcting for some of the significant external effects (varying solar zenith angle and varying amounts of haze over the scene), for extracting the most significant spectral features ("tasseled cap") from Landsat data, and for extracting spatial features (pseudofields) from the data. These improvements in preprocessing as well as the state-of-the-art stratified sampling aspects of Procedure 1 have been incorporated in Procedure B.

Differences between Procedure B and Procedure 1 are that Procedure B is both a multisegment and a multistratum procedure. Multisegment means that Procedure B uses data from several LACIE-sized segments together and makes a proportion estimate for the entire group of segments as well as for the individual segments. Multistratum means that, in the process of clustering data features, Procedure B produces multiple classes or strata rather than just two

^aEnvironmental Research Institute of Michigan, Ann Arbor, Michigan.

strata (as in Procedure 1), and performs stratified sampling on each of these multiple strata in order to make a proportion estimate.

In the following sections, a detailed description of Procedure B is given, the test results to date for the components and for the overall performance of Procedure B are presented, and the conclusions that can be drawn from these tests are discussed. In the final section, the overall rationale of signature extension for crop area estimation is summarized.

DESCRIPTION OF PROCEDURE B

This section contains a description of Procedure B. Some of the more detailed aspects are documented by reference to other papers in this LACIE symposium.

Broad Description

The objective is to develop improved techniques for estimating the amount of an agricultural crop present in a large geographical region or partition. Specifically, the goal is to define a training and classification procedure (or more generally, a labeling and estimation procedure) that will allow valid estimates for the region to be made on the basis of training information obtained from a few segments. Thus, the data to be processed are of two types: (1) a small amount of labeled training data from some segments in the region and (2) a large amount of unlabeled data from these and other segments.

Procedure B is a specific technique of proportion estimation that tells an analyst which scene elements to label and then uses those labels in an unbiased way to produce a proportion estimate for the scene. The extent to which Procedure B is used to perform signature extension is discussed later.

The fundamental concept of proportion estimation in Procedure B is similar to the concept used in Procedure 1 (ref. 14); namely, a stratified sampling technique performed in the spectral feature domain. There are, however, major differences in concept between Procedure B and Procedure 1. The steps in Procedure B are as follows.

1. **Data normalization:** There is a preprocessing stage in which data are screened and corrected for effects of satellite calibration, Sun angle, and haze over the scene.

2. **Feature extraction:** Spectral and spatial features are extracted from the multispectral image data, and these features are augmented with ancillary information such as weather and crop calendar data.

3. **Stratification of the feature space:** An unsupervised clustering algorithm divides the feature space into domains or strata. The number of strata produced is larger than the two (wheat versus non-wheat) of Procedure 1. Typically, for single segments, the number of strata produced is around 40.

4. **Multisegment:** The stratification may be performed for feature vectors encompassing several sample segments. Thus, spectrally similar features from several sample segments may be assigned to the same stratum. It is in this sense that Procedure B performs signature extension.

5. **Sample selection:** Certain numbers of samples are allocated to each stratum. Feature vectors are randomly drawn from each stratum and then are identified using analyst (or ground truth) labels in a production (or research) version of the procedure.

6. **Proportion estimation:** From the identified samples, proportion estimates are made for each stratum, for the entire group of segments, and for each segment in the group.

7. **Performance monitoring:** Each component of the procedure (in the research version) is monitored for performance according to criteria of unbiasedness and low variance.

Procedure B and some of its differences from Procedure 1 are discussed in more detail in the following subsections: Preprocessing and Feature Extraction, Stratification Procedure, Allocation of Samples, and Proportion Estimation.

Preprocessing and Feature Extraction

The objectives of the preprocessing and feature extraction steps may be several (ref. 6).

1. To make the data more comprehensible by adjusting all of them to standard conditions of observation

2. To eliminate or flag bad or noisy observations in the data

3. To make the data more comprehensible by extracting physically meaningful features or projecting the data in such a way as to display their physical structure

4. To compress the data, retaining most of the in-

formation and averaging out noise and redundancy

5. To make the distributions of the derived features fit some convenient model such as the multivariate normal distribution (This step is not in the current implementation of Procedure B.)

To define features to be extracted, the major characteristics of the data must be kept in mind. Remotely sensed data have three main attributes: spectral, spatial, and temporal. The spectral profile for each picture element (pixel) is provided by the MSS. The spatial characteristics include a pixel scan line and point number and the position of the LACIE segment in the region. The temporal characteristics include the changes associated with the passage of time during the growing season.

The problem of estimating and classifying in a wide region is complicated by several sources of variation in the data.

1. Systematic external effects, such as haze, viewing angle, Sun zenith angle, and scanner calibration

2. Effects upon particular crops of ancillary variables (such as moisture, growing degree days, and crop calendar) which are observable for an entire segment

3. Random noise due to scanner noise, to within-site variation of an ancillary variable whose site average is known, or, finally, to variation in underlying ancillary conditions which are not being currently observed but which are significant in their effects

Regarding random noise, the noise properties exhibited by MSS data are generally highly correlated spatially. A simple example is provided by the fact that the within-field variance of the MSS signals is less than the between-field variance of multiple examples of the same crop; hence, the choice of a reasonable number of pixels from a single field within a segment may constitute insufficient training for that segment. Similarly, the choice of a single segment and the fields within that segment as a training data base for a group of segments may constitute insufficient training for that group of segments. For this reason, development of a training procedure has proceeded on the assumption that multiple segments are needed for training to represent the variability of data present in a group of segments.

Some of the data variation caused by external effects can be removed by correcting for known external effects, so that all data are transformed to a standard reference condition. In Procedure B, the data are screened so that garbled data and pixels con-

taining clouds, cloud shadows, and water are detected and flagged, and a haze diagnostic is computed over the array of good data points (see the paper by Kauth et al. entitled "Feature Extraction Applied to Agricultural Crops as Seen by Landsat"). Next, the data are corrected for differences in satellite calibration (i.e., all Landsat-1 data are modified to simulate Landsat-2 data), for Sun zenith angle (i.e., all data are made to look as though they were gathered with a Sun zenith angle of 39°), and haze (all data are transformed to a standard haze condition) (see the paper by Lambeck and Potter entitled "Atmospheric Effects Compensation for LACIE Data").

The noise variation can be lessened and operating efficiency greatly increased by adopting methods of data compression that preserve useful information while averaging out noise. In Procedure B, the data are compressed in two ways, spectrally and spatially.

Spectral compression is accomplished by a linear transformation of the four Landsat bands through a matrix rotation called the tasseled-cap transform, or the Kauth-Thomas transform (see the paper by Kauth et al.). Most of the significant information regarding agricultural scenes has been found to lie in the plane defined by the first two components of the resulting transformed data (ref. 15); hence, these components are retained and the last two components are discarded, resulting in a data compression by a factor of 2.

Spatial averaging is accomplished by grouping together pixels that are near to each other and spectrally similar. These fieldlike groups are referred to as "blobs" (ref. 16). The spectral average of the group of pixels, the number of pixels in the group, and the average spatial position of the group are retained as data features describing that group (blob). As a result of "blobbing," the data are further compressed by a factor of 30.

The final step in feature definition is to associate with each blob certain ancillary data that vary from segment to segment, such as view angle, crop calendar, available soil moisture, and latitude and longitude. The idea is that, to the extent possible, the physical factors that affect the MSS data should be associated with those data, whereas an arbitrary factor, the segment identification, should be ignored. In Procedure B, the ancillary data features are treated as equal to the other features, and the net result is to perform a "soft" geographic partitioning to be discussed further in later sections.

Stratification Procedure

The words "stratification," "partitions," "groups," etc., have been heavily used during the course of LACIE, particularly in discussions of signature extension. At this point, about all one can do is to define and use these terms in a consistent way. In general, to stratify is to divide the space of observation into mutually exclusive regions (called strata) preparatory to making some kind of estimate separately within each of these regions. This definition includes diverse concepts within its scope, as the following examples show.

The first step in the application of Procedure B ought to be to limit the range of application to a large region of moderately constant ancillary data conditions. The region might be one-third to one-half the size of Kansas and might contain from 10 to 60 LACIE sample segments. A number of different approaches to such large-scale geographic partitioning have been developed (see the paper by Thomas et al. entitled "Development of Partitioning as an Aid to Spectral Signature Extension" and the paper by Hallum and Basu entitled "Natural Sampling Strategy"). In general, these partitions can be thought of as strata defined on the space of the *ancillary* variables.

In Procedure I, a wheat-nonwheat classifier separates the data into two classes. Here, the classes produced can be regarded as strata defined on the space of the *spectral* variables.

In Procedure B, the strata are domains defined on the feature space, which includes both spectral-spatial features and ancillary features. Hence, there is a spectral stratification similar to Procedure I and at the same time a "soft" geographic partitioning. Furthermore, the strata are not recombined into two classes as in Procedure I but are left as multiple strata.

An unsupervised clustering technique is used to group the blobs into strata. The algorithm, called BCLUST (i.e., blob-clustering), consists of the following steps.

1. The blobs from a number of segments are ordered randomly. Omitted from the list are the so-called small blobs, which are blobs that have no interior pixels. (An interior pixel is one that faces pixels from the same blob on all four sides.) The small blobs (usually stringy boundary areas between fields) are omitted because they are difficult to label and subject to registration errors. The blob algorithm parameters are set so that only a small proportion of

the pixels in the segment are contained in small blobs.

2. The big blobs are clustered using the spectral mean vectors and the ancillary variables jointly. The spectral means are computed from only the interior pixels because these pixels give a purer representation of the crop present in the blob. The distance measure used in the clustering is

$$d_i = \sum_{j=1, nchan} w_j (x_j - \bar{x}_{ji})^2 \quad (1)$$

where x_1, \dots, x_{nchan} is the data vector, $\bar{x}_{1j}, \dots, \bar{x}_{nchan,j}$ is the mean vector of cluster i and $nchan$ is the total number of spectral and ancillary variables.

3. For a given channel j , the weight w_j is constant; i.e., not varying from cluster to cluster. The means \bar{x}_{ji} are updated for clusters as new points are added. The clustering parameters are chosen to produce between 40 and 100 groups of blobs. These groups of blobs are termed B-clusters.

Training Selection Procedure

In the current configuration of Procedure B, the term "training selection" is a misnomer. Nevertheless, a discussion of how this terminology came into being is instructive.

In the initial development of Procedure B, it was intended that signature extension would be accomplished through a process of training a classifier on a subset of the segments in a large region. At that time, it was believed that it might be necessary to train on several segments simply to acquire robust signatures; the question being addressed then was this: Without reference to ground truth, which segments and which feature vectors (blobs) within the segments should be chosen for training? The problem was to choose training segments and blobs within those segments such that there was sufficient training for each class, even though the information about class membership was not available at the time the choice of segments was made. Faced with this problem, it at least seemed reasonable to choose training blobs representative of the distribution of unlabeled feature vectors.

The blob-clustering algorithm described in the previous section was originally developed to provide

a convenient representation of the empirical frequency function of the feature vectors in the entire region and within each segment. The procedure for training segment selection considered many possible combinations of segment choices and, for each one, computed a value function according to certain heuristically derived rules. The rules were based on the idea that training segments should be chosen that, in combination, would provide several blobs to be labeled within the domain of every important B-cluster. The details of the segment selection and blob selection algorithm that accomplished this type of selection are given in reference 17.

Having made the training selections, it was originally intended to create signatures, classify all the blobs into two classes, and finally correct for bias over all the segments by drawing and labeling additional samples, as in Procedure 1. However, in order to obtain an early check on the performance of the segment selection program alone, from a set of labeled feature vectors on hand, it was decided to go directly to proportion estimation by making estimates for each B-cluster and aggregating those estimates. It now appears that, in the absence of good prior signature information, the labeling effort is more efficiently used if samples are drawn from strata, labeled, and used directly in proportion estimation (refs. 18 and 19).

The term "training selection" is a misnomer because, in its current configuration, Procedure B has no training phase. Samples drawn are only for the purpose of making a stratified sample estimate (bias correction), not for training. The segment selection and blob selection algorithms have been modified to optimize sampling efficiency in view of this revised purpose.

If the best efficiency is found by just sampling, one may ask what has happened to the concept of training? How much prior information is needed to make training useful? These issues certainly deserve further discussion by the entire community.

Allocation of Samples

The Procedure B selection algorithm has three steps. First, a *choice of segments* is made that is best with respect to a certain value function. Second, the *number of blobs* to be labeled in each chosen segment represented in each B-cluster is specified. Finally, the *particular blobs* to be labeled are chosen at random

from those available in each chosen segment represented in each B-cluster.

Segment selection.—The objective of the segment selection algorithm is to provide a supply of blobs from which proportional allocation can be made, as described in the next section. The algorithm for segment selection is as follows.

1. Choose the number S of segments to be selected.
2. For each possible set of S segments, compute a value function,

$$\text{value} = \sum_{\text{all clusters } i} N_i H_i \quad (2)$$

where N_i is the number of pixels in the i th B-cluster and H_i is the "hit function" measuring how well the choice of segment intersects the i th B-cluster. The hit function is calculated as follows.

1. Count a hit for a segment only if the segment has at least LB blobs in the i th B-cluster (typically, use $LB = 2$).

2. Count W_1 for the first segment hit, W_2 for the second segment hit, and so on; H_i is the sum of the W 's for B-cluster i (typically, $W = 50, 10, 2, 0, 0 \dots$).

The reason for introducing the parameter LB is that it would not be desirable to pick a segment which had only one, perhaps spurious, example in it to represent a B-cluster. The choice of a relatively large value for W_1 is to ensure that the segments chosen will first of all do a good job of getting at least one blob for training in almost every B-cluster with emphasis on the large B-clusters. But when these objectives are mostly satisfied, then the weights W_2 , W_3 , and so on, ensure that the larger B-clusters will be represented by more than one segment if possible.

Determination of numbers of blobs to be labeled.—In the case where single pixels are sampled and where there is no prior information about the true proportion P_i in each stratum, it can be shown that the minimum variance allocation of a fixed total number of pixels is to allocate them in proportion to the number of pixels in each stratum. Procedure B applies this idea in determining the numbers of blobs to be labeled in each B-cluster and chosen segment.

First, it is decided how many blobs are a reasonable number to be labeled. This decision involves the weighing of costs and is outside the scope of this report. This number of chosen blobs is then divided

among B-clusters in proportion to the number of pixels in each.

The number of blobs in each B-cluster is then split up among the chosen segments represented in that cluster. The split is proportional to the number of blobs in each "cell" (intersection of the B-cluster and a segment). The purpose of this proportional allocation within B-clusters is to make the estimation formula approximately unbiased (see the following section) and to approximate a minimum variance allocation.

Choice of blobs to be labeled.—After the number of training blobs allocated to each B-cluster/segment cell has been determined, Procedure B is used to make the actual choice at random from all the blobs in each cell. The method of random selection that is used ensures that the crop proportion of the pixels in the sampled blobs is an unbiased estimate of the crop proportion of the pixels in the cell. Rather than a simple random sample of blobs, which results in a biased estimate, the method is to choose the first blob with probability proportional to size and the others (if any) with equal probability (ref. 20).

Proportion Estimation

Proportion estimates can be made either for individual sample segments or for an entire group of segments to which Procedure B has been applied. The proportion estimation formula for a group of segments is

$$\hat{p} = \frac{\sum_{i=1}^M N_i \hat{p}_i}{\sum_{i=1}^M N_i} \quad (3)$$

where \hat{p}_i is the estimated proportion of wheat in the i th B-cluster, M is the number of B-clusters, and N_i is the total number of pixels in the i th B-cluster.

$$\hat{p}_i = \sum_j \frac{N_{ij}}{N_i} \hat{p}_{ij} \quad (4)$$

where \hat{p}_{ij} is the estimated proportion of wheat in the

(i,j) th cell (i.e., the intersection of B-cluster i and segment j) and N_{ij} is the number of pixels in the (i,j) th cell. Finally,

$$\hat{p}_{ij} = \frac{\sum_k N_{ijk} P_{ijk}}{\sum_k N_{ijk}} \quad (5)$$

where P_{ijk} is the proportion of wheat in the k th sample blob in the (i,j) th cell and N_{ijk} is the number of pixels in the k th sample blob in the (i,j) th cell.

The formula for the proportion estimate of a single segment (say the j th segment) on the basis of the multisegment B-cluster is

$$\hat{q}_j = \sum_{i=1}^M \frac{N_{ij}}{N_j} \hat{p}_i \quad (6)$$

where M is the number of B-clusters and N_j is the number of pixels in the j th segment

$$N_j = \sum_{i=1}^M N_{ij} \quad (7)$$

If these estimates are aggregated over the entire group of segments, \hat{p} is obtained again; i.e.,

$$\hat{p} = \sum_{j=1}^G \frac{N_j}{N} \hat{q}_j \quad (8)$$

where G is the number of segments.

Inserting equation (6) into equation (8) and reversing the order of summation, the following is obtained.

$$\hat{p} = \sum_{i=1}^M \left(\sum_{j=1}^G \frac{N_{ij}}{N} \right) \hat{p}_i \quad (9)$$

which is equivalent to equation (2).

Thus, the same proportion estimate for a group of segments is obtained whether proportion estimates are made on single segments and then aggregated or whether a single estimate is made directly for the entire group of segments. These single-segment estimates based on multisegment B-clusters are not the same estimates as those obtained by running Procedure B on single segments, because the latter are based on single-segment B-clusters.

Bias of the estimator.—In this section, it is first shown that the proportion estimator is an unbiased estimate of the crop proportion measured on the pixels of the big blobs. (Recall that a big blob is a blob that has at least one interior pixel and that an interior pixel is one that faces pixels from the same blob on all four sides.) The blobs were drawn from each B-cluster/segment cell in such a way that the crop proportion in the sample blobs is an unbiased estimate of the crop proportion in the cell.

The cell proportions are combined into a cluster proportion by the same weighted average operation that combines the cell estimates into a cluster estimate. Therefore, the cluster estimate is an unbiased estimate of the cluster proportion. Again, the cluster proportions are combined into an overall proportion by the same weighted average operation that combines cluster estimates into an overall estimate.

Thus, the overall estimate is an unbiased estimate of the overall proportion. This conclusion applies only to the pixels in the big blobs. Leaving out the small blobs introduces a bias.

In the research version of Procedure B, the wheat proportion of each selected blob is measured exactly from a pixel-by-pixel specification of the ground truth and thus has no bias. This proportion would not be easy to estimate in practice because the pixels on the edge of a blob are likely to be on or near a field boundary and hence subject to multitemporal registration error and mixed spectral response. Therefore, two other more realistic estimates of the wheat proportion of a blob have been considered. One is to estimate the proportion from the ground truth of the blob interior (i.e., the interior pixels). The other is to label the blob purely wheat or not wheat on the basis of the percentage of wheat in the blob interior. The two estimates are very much alike because the tests on Kansas and North Dakota segments show that the blob interiors are generally pure.

The bias in the proportion estimate for a segment when using either of these two practical methods of

estimating the wheat proportion in a blob is

$$\beta = \beta_B + \frac{N_S}{N}(P_B - P_S) \quad (10)$$

where N is the number of pixels in the segment, N_S is the number of pixels in small blobs, P_B is the true proportion of wheat on the big blob pixels, P_S is the true proportion of wheat on the small blob pixels, and β_B is the bias in the estimate of P_B .

If the whole blob is labeled, then β_B is zero, as we have seen, and the bias is $\beta = N_S/N(P_B - P_S)$. In 13 Kansas segments tested, the bias in the wheat percentage estimate was found to be as indicated in table I.

To determine the variance of the proportion estimate, let

$$N = \sum_{i=1}^M N_i \quad (11)$$

$$N_i = \sum_{k \in L(i)} B_{ik} \quad (12)$$

where N is the number of pixels in the group of segments and N_i is the number of pixels in the i th B-cluster. From equation (3) and using the standard formula, the variance of \hat{P} is

$$\text{Var } \hat{P} = \sum_{i=1}^M \left(\frac{N_i}{N} \right)^2 \text{Var } \hat{P}_i \quad (13)$$

TABLE I.—Wheat Percentage Estimate Bias in 13 Kansas Test Segments

Identification method	Range of bias	Average absolute bias
All pixels in the blob	−3.1 to 1.2	1.2
Blob interiors	−3.8 to 1.0	1.3
Pure blob interiors	−4.8 to 0.6	1.4

The variance of \hat{P}_i is more difficult to analyze. If, for example, n_i statistically independent single pixels were drawn from the i th B-cluster and

$$\hat{P}_i = \frac{n_{wi}}{n_i} \quad (14)$$

was set, where n_{wi} is the number of pixels found to be wheat, then the binomial variance of \hat{P}_i would be

$$\text{Var}(\hat{P}_i) = \frac{P_i(1 - P_i)}{n_i} \quad (15)$$

If n_i is replaced by the number of blobs (e.g., b_i) drawn for training from within the i th B-cluster, then, because each blob represents a number of pixels,

$$\text{Var}(\hat{P}_i) < \frac{P_i(1 - P_i)}{b_i} \quad (16)$$

Therefore, an upper bound on the variance of \hat{P} is given by

$$\text{Var} \hat{P} < \sum_{i=1}^M \left(\frac{N_i}{N} \right)^2 \frac{P_i(1 - P_i)}{b_i} \quad (17)$$

If a random sample is made from each B-cluster/segment cell, rather than just from each B-cluster, and then equation (4) is used to estimate P_p , a two-way stratified sampling is being executed over B-clusters and over segments. In this case, an upper bound on the variance of the estimate for a group of segments is

$$\text{Var} \hat{P} < \sum_{i=1}^M \sum_{j=1}^G \left(\frac{N_{ij}}{N} \right)^2 \frac{P_{ij}(1 - P_{ij})}{b_{ij}} \quad (18)$$

As will be shown later, the variance of two-way stratification is always smaller than the variance of one-way stratification.

Variance due to segment selection.—So far, the variance of the proportion estimates due to random sampling within each B-cluster (eq. (17)) or within each B-cluster/segment cell (eq. (18)) has been discussed. In the application of Procedure B to signature extension, samples are identified from only a subset of the segments. If the wheat proportion in B-clusters is constant over all segments, then it does not matter which segments are sampled from. In practice, however, it has been observed that the wheat proportions within B-clusters are not homogeneous, especially when Procedure B is applied to a very large region, such as the entire state of Kansas. This observation implies that, even when the segments are chosen to represent the B-clusters, the particular choice of sample segments to be used for labeling affects the estimates \hat{P}_i . Thus, the process of randomly choosing subsets of segments may introduce an added random variation, whereas the process of systematically choosing segments may introduce a systematic variation, or bias. These questions have been addressed empirically and the results are included within the next section.

COMPONENT PERFORMANCE TESTS

In performance testing of Procedure B, two classes of tests are distinguished: (1) component performance tests, which are discussed in this section, and (2) overall performance testing, which is discussed in the following section.

The purpose of component testing is to determine optimal parameter settings for each component of Procedure B. The approach is to measure component performance as a function of the parameter values and then adjust the parameter values to achieve the best performance.

The performance measures used are bias, variance, and variance reduction factor (RV). Not all these measures are appropriate for every component. Following are the major components of Procedure B being examined.

1. Spectral/spatial stratification (BLOB)
2. Spectral/ancillary data stratification (BCLUST)
3. Training segment selection
4. Training blob selection
5. Proportion estimation algorithm

The appropriate measure of performance for the first two components is the RV. The performance measures for the last three components are the bias and the variance of the overall proportion estimate.

Parameter selection for each of these components will be described in reverse order. For the last four components, the data base consisted of nine segments in Kansas for which ground truth was known. The percentage wheat for each blob was estimated by comparing blob maps with ground-truth photographs. The tests of the spectral-spatial stratification, which required a more accurately specified ground truth, were based on a pixel-by-pixel ground-truth tape. This tape included the original nine segments and four more.

Proportion Estimation Algorithm Performance Tests

Parameters.—Although the algorithm does not have parameters to be optimized, the concern was that the estimator be unbiased with respect to the source of labels (i.e., the LACIE analyst or ground truth).

Experimental procedure.—Procedure B was run for all the nine ground-truthed segments and every available blob was allocated for training. (There were about 3000 in all.) The Procedure B estimate of the total percentage of wheat in the group of nine segments was compared with that measured by planimetry at JSC.

Discussion.—Initially, the mathematical form of the estimator did not account for the size (population) of each blob in estimating the proportion of wheat in a B-cluster. The resulting proportion estimate was found to be positively biased. This situation has been corrected and the difference between the ground-truth proportion and the Procedure B wheat proportion estimate is about 0.3 percent. Using about 3000 blobs, the proportion estimate for the 9 sample segments is 21.89 percent wheat. This number will be used as the standard of comparison in further work with this set of nine segments.

Blob Selection Performance Tests

Parameters.—The parameters that can be varied are the total number of blobs selected for labeling and the number of segments in which labeling is performed. For any given number of segments to be labeled, Procedure B chooses a specific subset; hence, the choice of a subset is not a parameter to be optimized.

In this section, the aspect to be studied is the variance introduced by random selection of blobs.

Experimental procedure.—For each setting of the parameters, 15 wheat proportion estimates were made. Each of these estimates was based on a different random drawing of blobs according to the constraints of Procedure B. Following are the parameter settings chosen for testing.

1. Number of blobs selected for labeling: 300, 600, 1500, 3000.

2. Number of sample segments in which labeling was performed: 1, 3, 6, 9

There are 300 to 400 blobs suitable for labeling in each sample segment. Hence, certain combinations are not possible; for example, it is not possible to draw 1500 blobs from only three segments.

Figures 1 through 3 give the results of this test. Figure 1 shows results when 300 blobs, drawn from 9, 6, 3, and 1 sample segments, were chosen for labeling. The ordinate is the resulting wheat proportion estimate for the entire group of nine segments. Each dot represents a proportion estimate obtained from a particular random draw of 300 blobs. The abscissa indicates the number of sample segments from which these 300 blobs were drawn. In this figure, the dots have been spread out slightly along the abscissa in order to make individual points more distinguishable.

The standard deviation of each group of data is shown by the error bars. The important thing to notice is that this standard deviation is about the same when blobs to be labeled are selected from nine, six, or three segments, but this standard deviation is notably smaller for the one-segment case. This is because 300 blobs drawn from 1 segment nearly exhaust the available blobs. Hence, the possible selections are fewer and the variance of the resulting proportion estimates is smaller.

Figure 2 shows the proportion estimates resulting when 600 blobs are selected from 9, 6, and 3 sample segments, whereas figure 3 shows the result of selecting 1500 blobs from 9 and 6 segments. In general, the previously mentioned trend is preserved. The case of 3000 blobs and 9 segments leads to a single point (no variance).

These results apply to the case where no ancillary variables were used. As will be discussed later, the use of an appropriate ancillary variable in the BCLUST algorithm improves the RV. This entire experiment was repeated using the choice of an ancillary variable, November soil moisture, which was

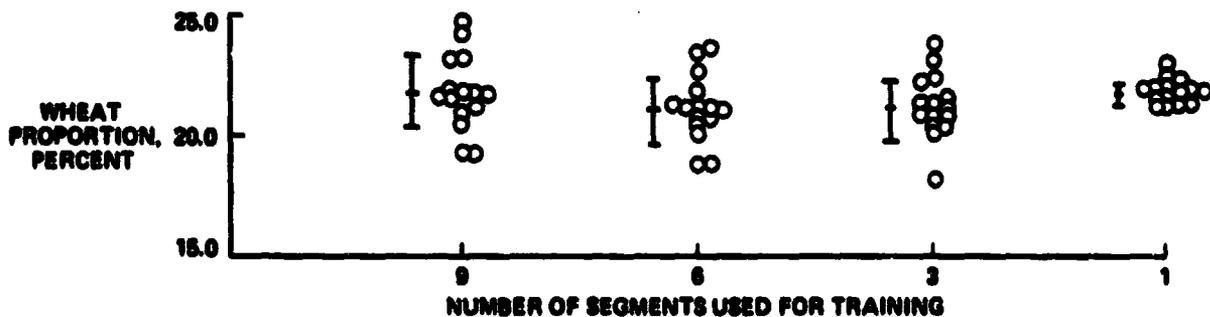


FIGURE 1.—Wheat proportion estimates for 9 segments with approximately 300 blobs selected for training and with no ancillary variable (true proportion = 21.89 percent).

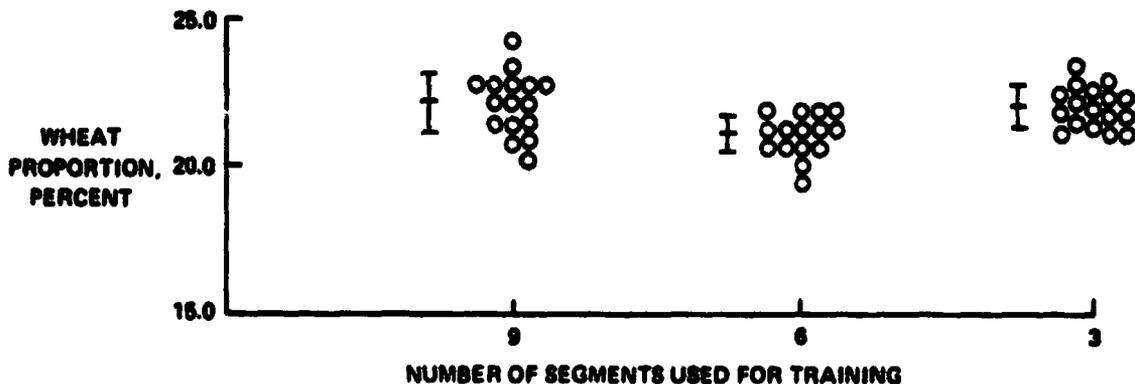


FIGURE 2.—Wheat proportion estimates for 9 segments with approximately 600 blobs selected for training and with no ancillary variable (true proportion = 21.89 percent).

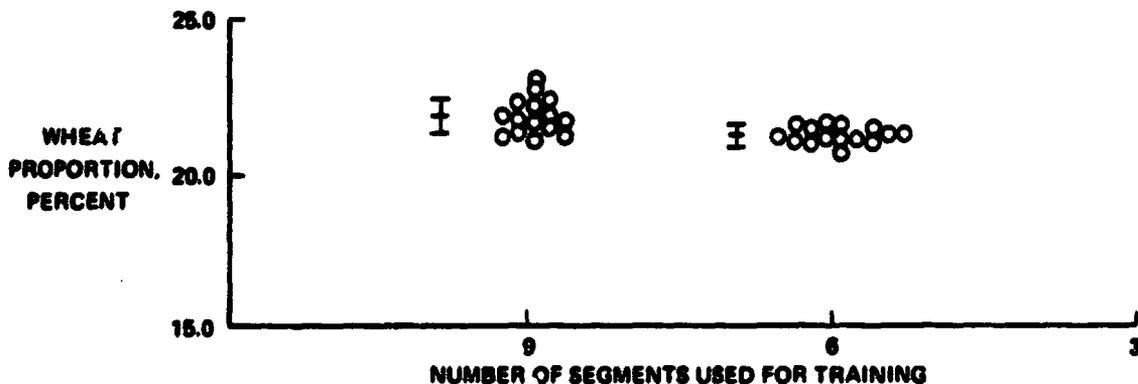


FIGURE 3.—Wheat proportion estimates for 9 segments with approximately 1500 blobs selected for training and with no ancillary variable (true proportion = 21.89 percent).

determined (in other tests to be described) to yield the best RV. These results and also the results of figures 1 through 3 are given in table II. This table will be referred to later in the discussion on constructing a model to explain the results.

Segment Selection Performance Tests

Parameters.—The segment selection program has several heuristically developed internal parameters that might be adjusted; however, these parameters

TABLE II.—Summary of Random Replications of a Proportion Estimation Procedure as a Function of the Number of Segments Used for Labeling, the Number of Blobs Labeled, and the Ancillary Variable Treatment

No. of segments used for labeling	No. of blobs labeled	Wheat proportion estimates without using an ancillary variable			Wheat proportion estimates with November soil moisture used as an ancillary variable		
		Mean, percent	Variance, percent	SD, ^a percent	Mean, percent	Variance, percent	SD, percent
9	3000	21.89	0.0090	0.095	21.80	0.0104	0.102
	1500	21.88	.2401	.490	21.87	.1832	.428
	600	22.23	.9006	.949	21.66	1.0506	1.025
	300	21.74	2.1287	1.459	22.47	2.3104	1.520
6	1500	21.34	.0590	.243	22.19	.1376	.371
	600	21.11	.4382	.662	21.95	.9565	.978
	300	21.15	1.9488	1.396	21.48	2.2134	1.521
3	600	22.00	.4597	.678	19.03	.2520	.502
	300	21.18	1.7742	1.332	18.98	1.1535	1.074
1	300	21.77	.2043	.452	37.51	.5852	.765

^aStandard deviation.

have not been optimized. What is perhaps more important is to measure the variance introduced by the process of selecting a subset of segments and to see whether the use of a systematic routine in any way reduces that variance or introduces a bias.

Experimental procedure.—There are 84 possible combinations of 9 segments taken 3 at a time and the same number of combinations taken 6 at a time. Of these possible combinations, 15 combinations of 3 and of 6 were chosen for labeling. The selection was done in two ways: for one set, a random selection of 15 combinations was made; for the second set, the best 15 combinations of segments as defined by the value function (eq. (2)) were selected. Procedure B was run for both of these sets with no ancillary variable, then the experiment was repeated with November soil moisture used as an ancillary variable. In each case, all of the blobs available from the selected three or six sample segments were used for training, so that the only source of variability would be the selection of segments.

Results.—The results of this experiment, expressed in percentage of wheat, are given in table III. Table IV shows these same results expressed in terms of their deviation from the "true" value of the wheat proportion, 21.89 percent.

In no case was the mean of 15 combinations significantly biased. Table III shows that segment selection introduces a variance into the estimate of per-

centage of wheat that amounts to a standard deviation of 1.5 to 5.3.

Is systematic selection of segments really helpful? An answer to this question is found in the comparison between the variance of the systematic and random selections. For a choice of three segments out of nine, systematic selection had just as large a variance as random selection, but for a choice of six segments, systematic selection had a smaller variance. The improvement is significant at the 0.01 level for the no-ancillary-variable case.

BCLUST Performance Tests

The spectral stratification algorithm BCLUST has a number of internal parameters, as mentioned earlier. Preliminary tests were run to establish a reasonable value for the growth parameter KTAU. This parameter was chosen such that the BCLUST tended to equalize the size of the B-clusters that were produced. The parameters that were evaluated in the main test series were the spectral variables used, the ancillary variables used, and the relative weight applied to each variable.

Performance measures.—The RV was the performance measure used in these tests. This measure has the advantages that it can be computed directly (if one has detailed ground truth) without making final

TABLE III.—Results Achieved With Procedure B With and Without the Use of an Ancillary Variable Using Random Selection and the Highest Ranked Selections of Sample Segments for Labeling

[True mean = 21.89 percent]

No. of segments used for labeling	Wheat proportion estimates without using an ancillary variable			Wheat proportion estimates with November soil moisture used as an ancillary variable		
	Mean, percent	Variance, percent	SD, percent	Mean, percent	Variance, percent	SD, percent
<i>Random selection</i>						
3	23.13	27.98	5.29	20.11	18.81	4.34
6	22.11	9.14	3.02	22.16	6.27	2.50
<i>Systematic selection</i>						
3	21.74	33.68	5.80	19.46	20.60	4.54
6	21.11	2.34	1.53	22.82	4.19	2.05

TABLE IV.—Procedure B Results Shown in Table III, Expressed as Deviations From True

No. of segments used for labeling	Estimates without using an ancillary variable			Estimates with November soil moisture used as an ancillary variable		
	Error of mean	SD of mean	Significant at 0.05?	Error of mean	SD of mean	Significant at 0.05?
<i>Random selection</i>						
3	1.23	1.37	No	-1.79	1.12	No
6	.21	.78	No	.26	.61	No
<i>Systematic selection</i>						
3	-0.16	1.5	No	-2.44	1.17	No
6	-.79	.40	No	.92	.53	No

proportion estimates and that it relates directly to the justification for doing any machine processing of the data. This point merits further discussion.

A wheat proportion estimate, $\hat{P}_a = n_w/n$, can be made directly by a LACIE analyst labeling n randomly chosen pixels, where n_w is the number labeled wheat. The variance of this estimate is

$$\text{Var } \hat{P}_a = \frac{P(1 - P)}{n} \quad (19)$$

where P is the true proportion of wheat in the area for which the estimate is made.

If the area is first stratified and then n_i samples are randomly drawn from the i th stratum, the proportion can also be estimated as

$$\hat{P} = \frac{1}{N} \sum_{i=1}^M N_i \hat{P}_i \quad (20)$$

where N_i = the number of pixels in the i th stratum
 N = the total number of pixels
 M = the total number of strata
 P_i = n_{wi}/n_i
 n_i = the number of samples drawn from the i th stratum
 n_{wi} = the number of samples in the i th stratum found to be wheat

The variance of this estimate is given in equation (17).

The larger strata will tend to contribute more significantly to the variance because of the squared factor $(N_i/N)^2$. This tendency can be offset by allocating more samples to the larger strata, and in fact it can be shown that, if there is no prior knowledge of P_i , the best allocation (i.e., least variance) is proportional to the sizes of the strata,

$$n_i = n \frac{N_i}{N} \quad (21)$$

With this allocation, equation (17) becomes

$$\text{Var } \hat{P} = \sum_{i=1}^M \frac{N_i}{N} \frac{P_i(1 - P_i)}{n} \quad (22)$$

The RV due to stratification is defined as

$$\begin{aligned} \text{RV} &= \frac{\text{Var } \hat{P}}{\text{Var } \hat{P}_a} \\ &= \frac{\sum_{i=1}^M \frac{N_i}{N} P_i (1 - P_i)}{P(1 - P)} \quad (23) \end{aligned}$$

The RV is not a function of the number of samples drawn. It can be interpreted either as the ratio between variances of stratified and unstratified estimates (as it has been defined) or equivalently as a measure of the number of samples required to achieve a certain variance of the resulting proportion estimate. The RV is thus better if it is smaller. Small

values of RV are obtained by making P_i or $1 - P_i$ very small; i.e., the RV is a measure of the purity of strata.

In defining the RV, some assumptions are made that are not completely fulfilled in Procedure B. For example, the RV is defined as a pixel-by-pixel measure, whereas the sampling in Procedure B is performed using blobs. Also, the assumption of proportional allocation defined in equation (21) is not strictly enforced and, therefore, the RV is a lower bound on the effective reduction of variance. Nevertheless, it is believed that the RV should correlate very well with the overall variance of Procedure B estimates and is therefore a useful and valid measure for optimization of BCLUST.

Experimental procedure.—BCLUST was run using the nine segments for which ground-truth proportion of blobs was known. Thus, the total number of pixels and the number of pixels known to be wheat could be calculated for each B-cluster, so that the RV could be calculated.

Six spectral variables (brightness and greenness in the first three LACIE biowindows) were first considered. An optimal set of weights (w_j in eq. (1)) was established among these six spectral variables. Next, the RV was computed using only the greenness variables with the previously determined optimal weights. Finally, maintaining the same relative weights among the spectral variables, various ancillary variables were added and their weights relative to the spectral variables were increased until best performance (i.e., minimum RV) was reached. (This best ancillary variable combination was used in the tests of blob selection and segment selection.)

The number of strata created affects the value of the RV. In fact, it can be shown (ref. 16) that the RV either decreases or stays constant whenever the number of strata is increased by splitting an existing stratum. The top curve in figure 4 illustrates this fact. (The reader should ignore the lower curves for the time being. Two-way RV will be discussed at the end of this subsection.) Because the curve seems to level out reasonably well for the 9 segments at about 90 B-clusters, it was decided to compare RV values using this number of B-clusters. The number of B-clusters can be controlled by a parameter, τ , internal to BCLUST.

Results.—The optimal weights for the six spectral variables were found by a search of six-dimensional space, starting at a point where the weights were in inverse proportion to the ranges of the values of the

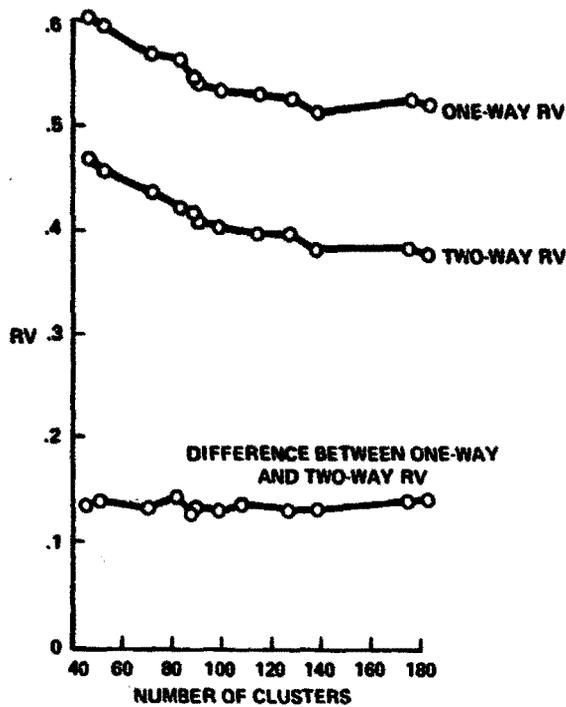


FIGURE 4.—RV due to clustering as a function of the number of clusters using six spectral variables.

variables. For each setting of the weights, the RV for 90 B-clusters was calculated. The search pattern was to follow the path of steepest descent to a setting of the weights that would result in the lowest RV. It happened that the optimal setting was found at the starting point.

Using weights determined in this manner, the RV for 90 B-clusters was also computed for the case where the three greenness variables alone were used by BCLUST. The comparisons are shown in table V. It was judged that the improvement due to retaining the three brightness variables was substantial. Thus, all six spectral variables were used in subsequent tests.

TABLE V.—RV for Three Greenness Variables Compared to RV for Three Greenness and Three Brightness Variables

RV	Three greenness variables	Six spectral variables
One-way	0.622	0.539
Two-way	.430	.406

The two-way RV shown in figure 4 will now be discussed. (See also eq. (18) and related discussion.) Two-way RV depends on a two-way stratification, by B-cluster and by segment. Two-way variance reduction can be achieved by using all the sample segments for labeling and allocating the samples drawn not only to B-clusters in proportion to their size but also to each segment in proportion to the population of the B-cluster within that segment. In fact, Procedure B allocates samples in this way when all the segments are used. It appears that roughly a 20-percent sampling advantage might be obtained by sampling for two-way reduction of variance, at the expense of using all the segments for training rather than a subset of them.

A number of different ancillary variables were considered, as well as combinations of ancillary variables. The details of each are not critical to this discussion. In general, the ancillary information falls into two categories: ancillary information derived from sources external to Landsat data, such as available soil moisture on certain dates, or the latitude and longitude; and ancillary information derived from the Landsat data itself, such as the green-index median, which is derived for these segments from the green-index numbers computed by Wehmanen (see the paper by Thompson and Wehmanen entitled "Application of Landsat Digital Data for Monitoring Drought"). This and the mean of the green arm are diagnostic features derived from each LACIE sample segment. They are believed to be a measure of the general vigor of vegetation within the segment.

With each combination, the procedure was the same. BCLUST was run with six spectral variables plus the ancillary variable with a certain weight. Then the value of τ was adjusted until there were 90 B-clusters. The value of τ is thus an indirect measure of the weight applied to the ancillary variable. Then the weight was changed, and the procedure was repeated. Figure 5 shows the RV versus τ for the case of November soil moisture. This happens to be the case that resulted in the lowest minimum value of the one-way RV. The pattern is fairly typical, however. As the weight of the ancillary variable increases, the one-way RV decreases to a minimum and then increases again. The explanation of the pattern follows.

Initially, with no weight on the ancillary variable, many B-clusters extend across all the segments. In some of the segments, a particular B-cluster may contain mostly wheat blobs; in some other segments, it may contain mostly nonwheat blobs. Thus, that B-cluster is "mixed" in the sense that P_i is neither very

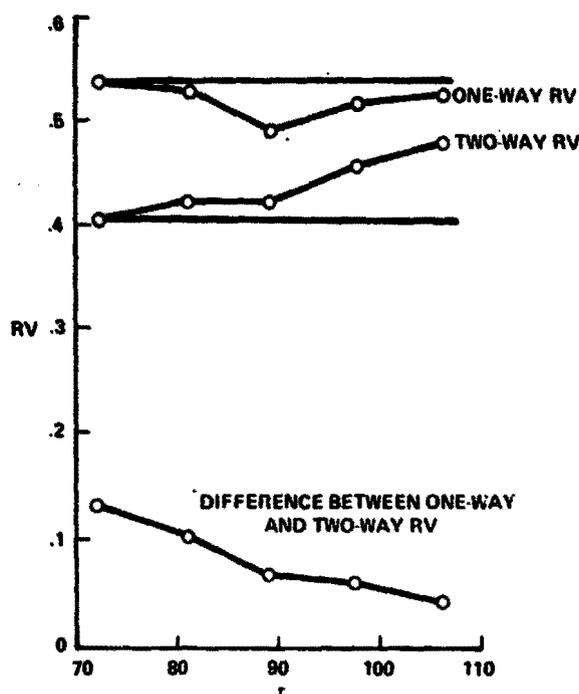


FIGURE 5.—RV due to clustering November 1 soil moisture along with the spectral variables; r measures the weight put on the ancillary variable ($r = 72$ corresponds to zero weight).

near zero nor very near one. The resulting RV is 0.539 averaged for all B-clusters. As the weight on the ancillary variable is increased, blobs from segments with differing values of the ancillary variable are forced into different B-clusters, even though they may have nearly the same value of spectral variables. B-clusters no longer can extend across all segments. The typical B-cluster is now split into two parts, and, if the signature of wheat depends on the ancillary variable, it is likely that each of the two parts will be purer than the original cluster, so the RV will decrease.

This trend cannot continue indefinitely. As the weight on the ancillary variable is increased, the effective weight on the spectral variables is decreased. In effect, the clustering procedure becomes incapable of distinguishing very well on the basis of spectral information. As the ancillary variable weight becomes too large, each segment contains its own little group of B-clusters that do not cross over to any other sample segment. In the case of 9 segments and 90 B-clusters, the average number of segments per B-cluster will be 10. The likelihood is small that many of these 10 B-clusters will be nearly pure.

The effect of the ancillary variable at its optimum weight is to create a sort of "soft" partitioning of the segments; i.e., a subset of the segments is found to intersect a certain subset of the B-clusters. The ancillary variable prevents an attempt at "signature extension" when it is, in fact, not feasible to extend signatures.

Table VI lists a few of the ancillary variables that have been tried and the weight at which they were found to have the best effect. (A weight of 1 means that the weight was inversely proportional to the range of the variable.)

BLOB Performance Tests

The spectral/spatial processing algorithm BLOB is the final step in feature extraction in Procedure B. BLOB creates pseudofields by grouping together pixels that are both spectrally and spatially near each other. The mean vector of the pixels assigned to each blob is then calculated and this value becomes the spectral feature which is used in BCLUST. BCLUST then performs a further grouping of blobs (and therefore of pixels) into B-clusters.

Parameters.—The BLOB algorithm contains a number of parameters. The algorithm is designed to add two more channels, line number and point number, to the multispectral data channels. Line and point are spatial variables. A standard clustering algorithm is then used to add pixels to existing clusters and to create new clusters when the pixel is not close enough to any existing cluster.

TABLE VI.—Ancillary Variables Considered and Performance Obtained With Each

Variable	One-way RV	Weight
No ancillary variables	0.539	—
November soil moisture	.492	0.5
Crop calendar, April	.502	.5
Green median:		
1st biowindow	.511	1.18
2nd biowindow	.500	.5
3rd biowindow	.497	2.0
Green arm mean:		
1st biowindow	.522	1.03
2nd biowindow	.519	2.0
3rd biowindow	.494	.48

The distance function used in the clustering is determined as follows. Each new data point (spectral: $x = (x_1, x_2, \dots, x_n)$; spatial: $x = (\text{line, point})$) is tested for admission into each existing blob by computing

$$\text{distance} = \frac{(x_1 - M_1)^2}{\text{VARI}} + \dots + \frac{(x_n - M_n)^2}{\text{VARN}} + \frac{(l - M_l)^2}{\text{VARL}} + \frac{(p - M_p)^2}{\text{VARP}} \quad (24)$$

where $(M_1, \dots, M_n, M_l, M_p)$ is the blob mean vector and $\text{VARI}, \dots, \text{VARN}, \text{VARL}, \text{VARP}$ are weights expressed as variances.

The pixel x is added to the existing blob with the smallest distance unless this minimum distance is greater than a parameter τ in which case the pixel becomes the seed point of a new cluster.

In this task, six spectral variables have been used, brightness and greenness from three biophases. The next problem is to specify the parameters of the BLOB algorithm, namely VARI through VAR6 , VARL , VARP , and TAU .

Another parameter of Procedure B that relates to blobs is the question of whether to use only big blobs in the operations from BCLUST on. Figure 6 is a gray map of the blobs in segment 1865. The big blobs are represented by printed characters and the small blobs by blanks. Although there are many small blobs, they do not contain many pixels. The small blobs mostly represent boundaries between fields.

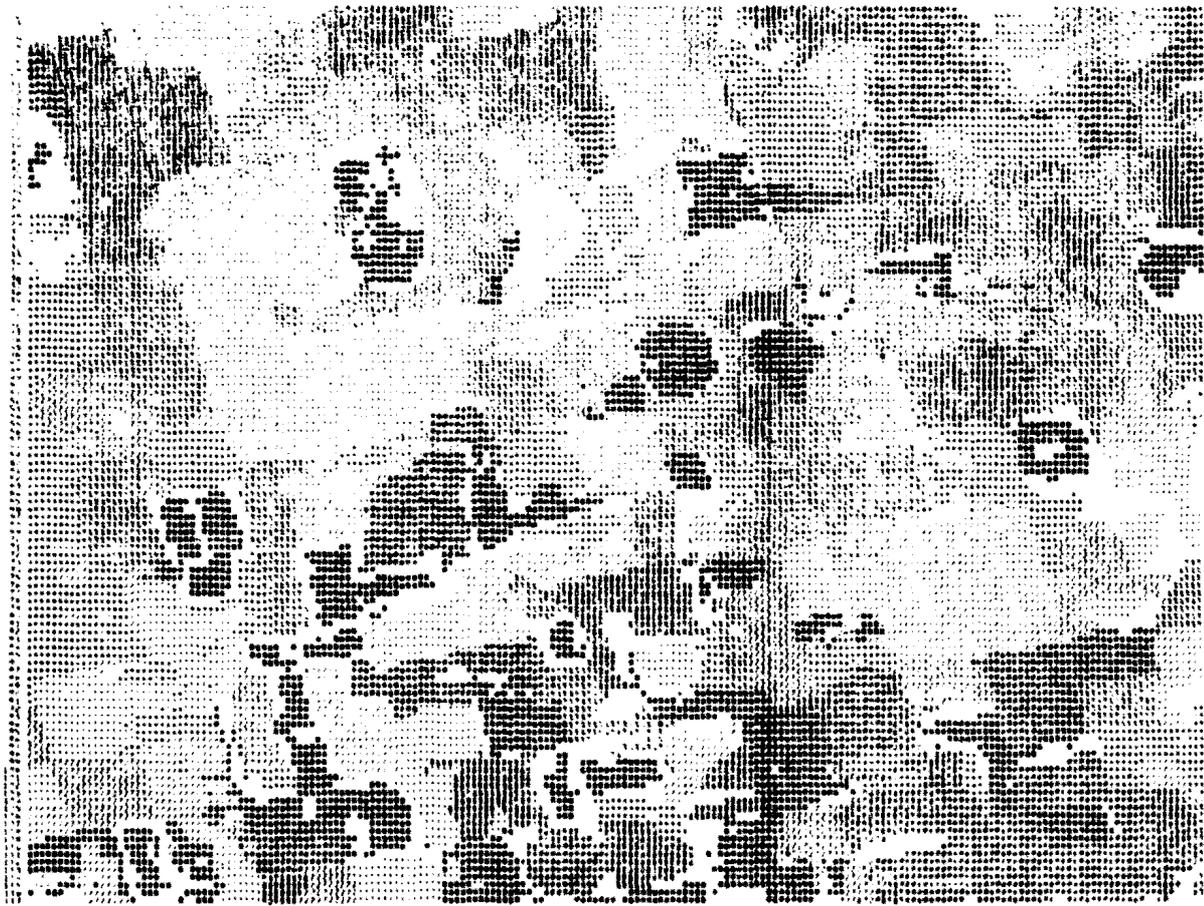


FIGURE 6.—Gray map of big blobs in segment 1865, Kansas.

The previously reported component tests of Procedure B have been run using big blobs only.

Performance measures.—The blob RV is analogous to that for B-clusters:

$$RV = \frac{\sum_{\text{all blobs}} \frac{N_i P_i (1 - P_i)}{N P_i (1 - P_i)}}{P(1 - P)} \quad (25)$$

where P_i is the proportion of wheat in blob i , N_i is the number of pixels in blob i , P is the overall wheat proportion in the segment, and N is the number of pixels in the segment. The blob RV is a measure of blob purity: the purer the blob, the smaller the RV. This measure also provides a lower bound for the B-cluster RV. No matter how well the blobs are combined into B-clusters, it is impossible to improve the RV found by blobbing alone, nor can the blob RV be any better than the pixel RV, which is

$$RV = \frac{\sum_{\text{all pixels}} p_j (1 - p_j)}{NP(1 - P)} \quad (26)$$

where p_j is the proportion of wheat in pixel j .

The usefulness of the blob RV and the pixel RV as performance measures was made possible by the recent provision of highly accurate pixel-by-pixel ground truth. The other performance measures are the bias and average absolute error of wheat estimates based on classes of blobs and pixels, such as big blobs, small blobs, and blob interiors.

Experiments and results.—The parameters of the BLOB algorithm were determined as follows. The spectral weights VAR1, . . . , VAR6 were set by referring to the previous work on finding the optimal spectral weights for grouping blobs into B-clusters. A search pattern in six-dimensional space indicated that for B-clustering, the best spectral weights, expressed as variances, are proportional to the effective ranges of the variables. Using the same proportion for blobbing weights, the spectral weights were determined relative to each other.

As for the spatial weights, the proportion of the line variance, VARL, was set to the point variance, VARP, so that the line standard deviation represents the same geographic distance as the point standard

deviation. This proportion is not 1 to 1 because points are sampled more frequently than lines.

The next step was to determine the balance between the set of spectral variances and the pair of spatial variances by holding the spectral variances constant and comparing the blob RV's for three sets of spatial variances, small, medium, and large. (If all the parameters are increased by the same proportion, the algorithm remains exactly the same. That is why holding the spectral variances constant and varying the spatial variances and τ is permissible.)

Figure 7 shows the result of computing the blob RV's for eight Kansas segments and three settings of the spatial variances. The lower setting corresponds to more emphasis on the spatial; the higher setting, to more emphasis on the spectral. The gentle trends indicate an optimal setting somewhere between the lower and middle settings.

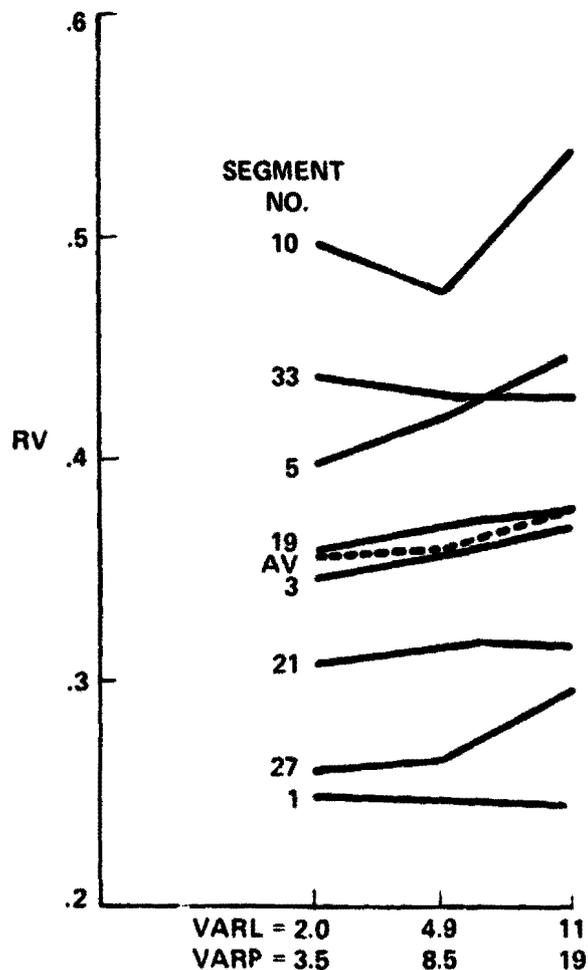


FIGURE 7.—Blob RV for three sets of spatial weights.

The problem with choosing τ is that when τ is decreased, the number of blobs is increased and the RV is generally decreased. However, decreasing τ too much reduces the number of interior pixels that are depended on for B-clustering. Also, the bias resulting from omitting small blobs is likely to be larger. A value of τ was chosen so that the number of big blobs would roughly correspond to the number of fields, and the τ value was kept constant for all the segments.

The final parameters are given in table VII. They apply to brightness and greenness variables in biophases 1, 2, and 3. If a subset of these spectral variables (e.g.; size m) were used, the appropriate weights would be obtained by keeping the corresponding spectral weights the same, multiplying VARL and VARP by $6/m$, and multiplying τ by $m/6$.

When these parameters were used for blobbing 13 segments in Kansas, big blob pixels averaged 81 percent and ranged from 65 to 93 percent, interior pixels averaged 30 percent and ranged from 15 to 43 percent, and edge pixels (i.e., those that are in big blobs but are not interior) stayed nearly constant at 50 percent. The number of big blobs was remarkably constant (from 342 to 489), whereas the total number of blobs ranged from 590 to 1991.

To determine limits of the efficiency of Procedure B, RV factors were computed for pixels, blobs, and B-clusters. The results are given in table VIII. Remembering that RV measures purity and that a low score is good, it is demonstrated that the pixel RV is a lower bound for the blob RV, which in turn is a lower bound for the B-cluster RV.

Noticeable features of table VIII are the considerable variability from segment to segment and

TABLE VII.—*Blob Parameters Appropriate for Using the Brightness and Greenness Transformed Channels in the First Three Biophases*

Description	Name	Weight
Brightness, biophase 1	VAR1	25.0
Greenness, biophase 1	VAR2	5.3
Brightness, biophase 2	VAR3	14.0
Greenness, biophase 2	VAR4	9.0
Brightness, biophase 3	VAR5	18.4
Greenness, biophase 3	VAR6	9.0
Line variance	VARL	3.46
Point variance	VARP	6.0
Distance limit	τ	23.2

TABLE VIII.—*RV's for Blob Tests on 13 Kansas Segments^a*

Segment	Pixels	Blobs	B-clusters	Blob Interiors
1020	0.09	0.24	0.25	0.04
1035	.11	.36	.66	.19
1041	.11	.41	.74	.21
1165	.16	.49	.86	.20
1851	.11	.36	.48	.16
1852	.08	.30	.48	.14
1861	.08	.26	.43	.09
1865	.08	.26	.56	.09
1886	.15	.43	.61	.17
1163	.18	.52	.71	.28
1167	.15	.39	.63	.18
1860	.13	.41	.53	.15
1887	.12	.39	.59	.17
Average	.12	.37	.58	.16

^aThe lower the RV, the purer the groups (perfect purity scores 0; perfect homogeneity scores 1).

the high correlation (0.80) between the 13 B-cluster RV's and the limit set by blobbing (i.e., the blob RV).

At this point, it is not clear why good (i.e., small) RV's are obtained for some segments and poor (i.e., large) RV's are obtained for other segments. The effects of multitemporal misregistration and of smaller field sizes are being examined.

The RV score for blob interiors is very good, showing that the blob operation is doing its job in the sense that, although there may be some confusion at the edges of the blobs, the interiors of the blobs are quite pure.

The average RV for B-clusters, 0.58, implies that the stratified estimate (i.e., the Procedure B estimate) would require only 58 percent of the identifications needed for an unstratified probability sample of pixels in the segment to achieve the same variance. This column is a misleading indication of the value of Procedure B because Procedure B samples blobs rather than pixels. Sampling blobs rather than pixels would be expected to reduce the variance because the proportion of wheat in a *blob* randomly chosen from a B-cluster would be expected to have less variance than the proportion of wheat in a *pixel* randomly chosen from a B-cluster.

The difference between the average B-cluster RV of 0.58 given in table VIII and the two-way RV of 0.406 given in table V results from the different types

of ground truth used as a standard. The table V result was based on wheat percentages of blobs estimated from ground-truth photographs and maps of blob interiors, whereas the table VIII results are based on a pixel-by-pixel ground-truth tape. Because the blob interiors are pure, the estimated wheat percentages based on them tended to be closer to 0 or 100 than was actually the case. This resulted in much smaller blob RV's, which in turn produced smaller B-cluster RV's.

The difference between the average B-cluster RV and the two-way RV is not a reflection of the difference between the single-segment and multisegment use of Procedure B. Using the estimated ground truth, the average single-segment B-cluster RV was 0.418 and the composite RV, obtained by collecting the single-segment strata into one big pool of strata, was 0.377. In short, the comparison of single-segment with multisegment use of Procedure B, as judged by the earlier, estimated ground truth, was a standoff.

The purity of blob interiors offers some hope of solving a practical difficulty with sampling. The current wheat estimates are all based on samples provided by the judgments of analyst interpreters (AI's). If these AI's are asked to identify a pixel at random, the chance is 70 percent that it will come from an edge or a small blob and is therefore likely to be on or near a field boundary. What with multitem-

poral registration errors and mixed spectral responses, it would seem a formidable, if not an impossible, task to identify such a pixel. But if asked to identify a relatively pure blob interior, the AI would have a much better chance to respond accurately.

Table IX, a table of different methods of estimating wheat percentage, gives empirical information about the accuracy of such a procedure. The "all pixels" column is the percentage of wheat computed from every known pixel in the segment. Those figures, and indeed all the others except those in the adjoining column, are based on the pixel-by-pixel ground-truth data recently computed at ERIM. The third column is a measurement of the percentage of wheat in the scene made by planimetry at JSC a couple of years ago. The average absolute difference between these two columns is 0.8 percent, showing that even the most careful measurements from high-resolution photography are subject to an error of about 1 percent. The discrepancy in segment 1865, which is caused by the failure of the photography to cover the top quarter of the segment, is left out of the calculation.

The next three columns are the percentage of wheat computed on various subsets of pixels. The big blob pixel estimate is quite close to the measured truth, whereas the small blob pixel and the interior pixel estimates are erratic.

The estimate made from small blob pixels has a

TABLE IX.—Various Estimates of Wheat Proportion

Segment	All pixels	JSC wheat, percent	Big blobs	Small blobs	Interior pixels	Extrapolated from interior	Extrapolated from pure interior
1020	26.1	25.3	24.0	43.2	21.1	23.9	24.1
1035	17.7	17.5	17.5	18.3	13.9	17.7	17.4
1041	14.4	14.4	14.3	15.3	14.5	14.7	14.4
1165	7.1	6.5	6.2	8.9	7.3	6.2	6.6
1851	22.8	21.9	20.4	33.6	16.6	19.9	19.6
1852	23.4	22.3	24.6	15.6	26.0	24.4	24.0
1861	34.9	34.4	34.4	42.5	28.3	34.2	33.9
1865	28.5	20.4	26.6	34.5	23.6	26.6	26.6
1886	29.7	28.9	29.9	28.4	29.6	29.8	29.9
1163	9.3	8.7	8.0	13.7	5.9	8.2	7.8
1167	10.1	8.0	7.0	15.7	5.7	5.3	5.3
1860	26.1	24.8	26.2	25.1	24.8	26.4	26.6
1887	11.4	10.9	10.2	17.8	7.0	10.2	10.0
Average bias		-0.8	-0.9	3.9	-2.9	-1.0	-1.2
Average absolute error		.8	1.2	5.5	3.3	1.3	1.4

bias that averages 4 percent and ranges from -8 to 17 percent. The bias in the estimate from the small blob pixels has to be balanced by a bias in the opposite direction on the big blob pixels. The big blob bias is smaller because there are more pixels in the big blobs. The estimate made from interior pixels has a bias that averages -3 percent and ranges from -6.5 to 2.5 percent.

The interior pixel estimate was computed by simply totaling the wheat percentages for the interior pixels and dividing by the number of interior pixels.

Another estimate based on interiors is to assign to all pixels in a big blob the proportion of wheat found in the blob interior. This is the "extrapolated from interior" column. This estimate is less biased and less erratic than the simple-minded interior pixel estimate. In fact, its bias and average absolute error are close to the error between the two planimetry measurements. This estimate, moreover, is more easily achieved by the AI because it is based only on interiors of blobs.

A more realistic estimate yet is obtained on the assumption that in the relatively pure interiors, the AI will identify either 100 percent wheat or 0 percent wheat. The extrapolation from pure interiors is such an estimate. Either 100 percent or 0 percent is extrapolated to all the pixels of the blob, and then the percentage of wheat of all pixels is obtained. This estimate also gives very good results and has the advantage of representing a practical sampling procedure.

OVERALL PERFORMANCE OF PROCEDURE B

As nearly as can be determined from the tests run so far, Procedure B is not importantly biased with respect to the source of labeling information. Thus, the primary overall performance measure should be the variance of the proportion estimate as a function of the cost of labeling. This cost has two main components: the cost of a single label and the cost of looking at a segment. The variance also appears to have two main components, the variance $V_{\text{blob}}(b)$ due to sampling blobs and the variance $V_{\text{seg}}(s)$ due to sampling segments. Let s be the number of segments selected and b be the number of blobs selected for labeling. Let C_{seg} and C_{blob} be the cost per segment

and the cost per blob, respectively. The total variance is

$$V = V_{\text{seg}}(s) + V_{\text{blob}}(b) \quad (27)$$

and the total cost is

$$C = sC_{\text{seg}} + bC_{\text{blob}} \quad (28)$$

For a fixed total variance, the minimum cost must be found. This can be done iteratively providing data on $V_{\text{seg}}(s)$ and $V_{\text{blob}}(b)$ are available.

A reasonable model for $V_{\text{blob}}(b)$ is that the variance follows the variance of the hypergeometric distribution, corresponding to sampling without replacement. For a single stratum, the variance of the hypergeometric distribution of the estimated proportion \hat{P}_a of wheat would be

$$\text{Var } \hat{P}_a = \frac{P(1-P)}{b} \left(\frac{B-b}{B-1} \right) \quad (29)$$

where P is the overall proportion of wheat, B is the total number of blobs in the stratum, and b is the number of sample blobs drawn.

For multiple strata, $i = 1, \dots, M$, the variance of the hypergeometric distribution is

$$\text{Var } \hat{P} = \sum_{i=1}^M \left(\frac{N_i}{N} \right)^2 \frac{P_i(1-P_i)}{b_i} \left(\frac{B_i - b_i}{B_i - 1} \right) \quad (30)$$

where B_i , b_i , and P_i refer to the i th stratum; N_i is the number of pixels in the i th stratum; and N is the number of pixels overall. As long as N_i is 1 and as long as b_i is allocated in direct proportion to B_i , the last bracket will be a constant fraction equal to $(B-b)/B$, and can be taken outside the summation. Thus,

$$\begin{aligned} V_{\text{blob}}(b) &= \text{Var } P \\ &= \frac{B-b}{B} \sum_{i=1}^M \left(\frac{N_i}{N} \right)^2 \frac{P_i(1-P_i)}{b_i} \end{aligned} \quad (31)$$

If it is assumed that the b_i are chosen in proportion to the number of pixels in each stratum,

$$V_{\text{blob}}(b) = \left(\frac{B-b}{B}\right) \sum_{i=1}^M \frac{N_i P_i (1-P_i)}{N b} \quad (32)$$

The summation is similar in form to the RV factor computed in equation (22) except that blob sampling (which may be slightly more efficient than pixel sampling) is involved and the proportional sampling rule is only approximately enforced. Allowing a constant factor K to account for these differences, it is assumed that the summation is equal to

$$\frac{1}{b} K P (1-P) (RV)$$

Thus,

$$V_{\text{blob}}(b) = \frac{B-b}{Bb} K P (1-P) (RV) \quad (33)$$

Currently, a model of this type is being constructed and fit to the data.

Similar but more complicated considerations are involved in creating a model for $V_{\text{seg}}(s)$. In general, the behavior of the model must be that the variance goes to zero as $s \rightarrow S$, as with the hypergeometric distribution. However, it also may be true that the variance due to segment sampling approaches zero for a fixed ratio of s/S as the number of segments increases in some large region. When larger numbers are considered, the distance between segments decreases and the representativeness of that fraction chosen should increase. The nine segments examined do not provide sufficient data to resolve this issue. Tests with additional segments are in progress.

CONCLUSIONS

It is evident that all the answers about Procedure B are not yet in. Before attempting to draw conclusions, some of the questions should be reviewed.

The two important considerations for the entire man/machine system for area estimation are accuracy and efficiency. For the machine processing

preceding and following labeling, these considerations may be expressed as bias and variance with respect to the labeling source. For the labeling portion of the system, the accuracy of the analyst labels and the ease of working with the image products in conjunction with the machine may be considered.

On the basis of the tests to date, Procedure B is not importantly biased with respect to analyst labels. In fact, it appears reasonable to allow some small bias to make the analyst's job more convenient by asking the analyst to label a blob interior, which is almost always pure, rather than a random dot, which is often a boundary or an edge pixel.

The efficiency of Procedure B may be expressed in terms of the variance reduction factor. For single segments, the RV is around 0.6. In terms of a classifier performance, this corresponds to a classification accuracy of about 0.84. For multiple segments, the RV is higher, indicating a loss in purity of some strata as they are extended over wider regions. With respect to two-way reduction of variance, the RV factor is again about 0.6. In order to exhibit either single-segment or two-way RV, samples must be drawn from every segment.

The purpose of the multisegment mode of operation is to reduce costs by allowing analysts to label in only a subset of segments. However, this is a type of block sampling that introduces an additional source of variance. It is still an open question whether gains can be made by sampling from only a subset of segments. The fact that multisegment Procedure B achieves better results with an ancillary variable than without is a hopeful sign that further gains are possible.

Another consideration of efficiency is the question of whether it is better to blob and then B-cluster or to B-cluster without any spatial processing. Tests are being conducted to determine this. Even if it turns out that blobbing reduces the RV only slightly, there are considerable benefits from data compression and from providing the analysts with pure field interiors to work with.

A LOOK TO THE FUTURE

Procedure B has a modular structure in which proposed improvements can be evaluated objectively. In the category of preprocessing, improvements may be made by developing individual detector calibration procedures, by implementing a spatially varying haze

effect correction, and by including the effects of view angle and background albedo. The Landsat-3 return beam vidicon (RBV) may provide additional useful information on the atmosphere.

In the category of feature extraction, spectral features other than brightness and greenness and spatial features other than blobs should be evaluated (ref. 21). The effects of the master data processor for Landsat-3 on spatial feature extraction should be examined. Again, the RBV may be of assistance in defining improved spatial features.

The largest gains will probably be made from an improvement in stratification and an improvement in ease and accuracy of labeling. In stratification, the problem is to improve the purity of strata while at the same time reducing the number of strata. Possibilities include the use of prior information and the improvement of the clustering method itself. For example, certain spectral classes are most unlikely to be wheat and could be eliminated a priori, thus reducing the portion of the data that must be sampled. Signature data from previous years or signature models based on field measurements may be used to isolate a subset of data likely to be wheat, leaving the more difficult cases to the analyst.

REFERENCES

1. Malila, W. A.: Investigation of Spectrum-Matching Techniques for Remote Sensing in Agriculture. Report No. 1674-10-F, Willow Run Laboratories of the Institute of Science and Technology, University of Michigan, Ann Arbor, Dec. 1968.
2. Horvath, R.; Braithwaite, J.; and Polcyn, F.: Effects of Atmospheric Path on Airborne Multispectral Sensors. Report No. 1674-5-T, Willow Run Laboratories, University of Michigan, Ann Arbor, Jan. 1969.
3. Elterman, L.: Vertical Attenuation Model with Eight Surface Meteorological Ranges, 2 to 13 Kilometers. Report No. AFCRL-70-0200, Air Force Cambridge Research Laboratories, Cambridge, Mar. 1970.
4. Wagner, T.: Automatic Processing and Analysis of Soils and Soil Conditions. Report No. 2760-2-F, Willow Run Laboratories, University of Michigan, Ann Arbor, July 1970.
5. Condit, H. R.: The Spectral Reflectance of American Soils. *Photogrammetric Engineering*, vol. 36, Sept. 1970, p. 955.
6. Malila, W. A.; Crane, R. B.; and Turner, R. E.: Information Extraction Techniques for Multispectral Scanner Data. Report No. 31650-74-T, Willow Run Laboratories, University of Michigan, Ann Arbor, June 1972.
7. Suits, G. H.: The Calculation of the Directional Reflectance of a Vegetative Canopy. *Remote Sensing of Environment*, vol. 2, 1972, pp. 117-125.
8. Kauth, R. J.; and Thomas, G. S.: The Tasseled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, LARS, Purdue University, W. Lafayette, Ind. IEEE Catalog No. 76 CH 1103-1 MPRSD, June 29-July 2, 1976, pp. 4B-41 through 4B-51.
9. Ahern, F. J.; Goodenough, D. G.; Jain, S. C.; Rao, V. R.; and Rochon, C.: Use of Clear Lakes as Standard Reflectors for Atmospheric Measurements. Proceedings of the 11th International Symposium on Remote Sensing of Environment, ERIM, Ann Arbor, Mich., Apr. 25-29, 1977, vol. 1, pp. 731-755.
10. Lambeck, P. F.: Signature Extension Preprocessing for Landsat MSS Data. ERIM 122700-32-F, ERIM, Ann Arbor, Nov. 1977.
11. MacDonald, R. B.; et al.: The Use of Landsat Data in a Large Area Crop Inventory Experiment (LACIE). Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, LARS, Purdue University, W. Lafayette, Ind. IEEE Catalog No. 75 CH 1009-0-C, June 3-5, 1975, pp. 1B-1 through 1B-23.
12. Peters, C.; and Coberly, W. A.: The Numerical Evaluation of the Maximum Likelihood Estimate of Mixture Proportions. Annual Report, University of Texas at Dallas, May 1975.
13. Transition Year Classification and Mensuration Subsystem (CAMS) Detailed Analysis Procedures. LACIE-00723, NASA Johnson Space Center, Houston, Tex., Mar. 1978.
14. Abotteen, K. M.; and Sumner, C. A.: Procedural Design Test. Technical Memorandum, LEC 10016, Lockheed Electronics Co., Houston, Tex., Mar. 1977.
15. Wheeler, S. G.; Misra, P. N.; and Holmes, Q. A.: Linear Dimensionality of Landsat Data With Implications for Classification. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, LARS, Purdue University, W. Lafayette, Ind. IEEE Catalog No. 76 CH 1103-1 MPRSD, June 3-5, 1976, p. 2A-1.
16. Kauth, R. J.; Penland, A. P.; and Thomas, G. S.: BLOB: An Unsupervised Clustering Approach to Spatial Preprocessing of MSS Imagery. Proceedings of the 11th International Symposium on Remote Sensing of Environment, ERIM, Ann Arbor, Mich., Apr. 25-29, 1977, vol. 2, pp. 1309-1317.

17. Kauth, R. J.; and Richardson, W.: Procedure B: A Multisegment Training and Proportion Estimation Procedure for Processing Landsat Agricultural Data. Technical Report 122700-31-F, ERIM, Ann Arbor, Mich., Nov. 1977.
18. Holatzynski, W.: Some Results Concerning Stratified Sampling. ERIM Internal Memo IS-WH-2267, ERIM, Ann Arbor, Mich., June 27, 1978.
19. Kauth, R.; and Richardson, W.: Classifiers and Stratified Sampling. ERIM Internal Memo IS-RJK-2257, ERIM, Ann Arbor, Mich., June 29, 1978.
20. Cochran, W. G.: Sampling Techniques. Second ed. Wiley, 1963, pp. 177-178.
21. Bryant, J.: AMOEBA: A Spatial Clustering Algorithm for Multitemporal Landsat Data. Report 15, Department of Mathematics, Texas A & M University, College Station, Aug. 1977.

An Evaluation of Procedure 1

S. G. Wheeler,^a R. P. Heydorn,^b P. N. Misra,^a W. Lee, Jr.,^a and R. T. Smart^a

INTRODUCTION

LACIE Procedure 1 has undergone continuous testing and evaluation, starting with analytical and experimental studies before it was implemented in the Earth Resources Interactive Processing System (ERIPS) software and continuing to the present with performance evaluations using blind-site data. In this paper, some of the evaluation studies and their results are described, some of the strengths and weaknesses of the procedure are indicated, and some areas for possible improvement are identified.

In concept, Procedure 1 is simple and straightforward. A detailed description of this procedure is given in another paper ("Classification and Measurement of LACIE Segments" by Heydorn et al.) but, for completeness, a short introduction is included here. The steps required to estimate the proportion of a segment devoted to small-grain production are the following.

1. Select a sample of pixels and label them as either small grains or other. These selected, labeled pixels are called type 1 dots.
2. Employing multispectral scanner (MSS) intensity values from some or all of the type 1 dots as cluster starting vectors, cluster the pixels in the segment.
3. Assign a label to each cluster using an automatic labeling technique based on the labels and intensity values of the type 1 dots.
4. Classify the pixels in the segment as either small grains or other using a LACIE mixture density classifier with cluster statistics serving as subclass mean vectors and dispersion matrices.
5. Calculate a classification-based proportion estimate by counting the number of pixels assigned a small-grain label by the classifier.
6. Select and label a second sample of pixels from

the segment. These pixels are called type 2 dots.

7. Generate an estimate of the segment's small-grain proportion by adjusting the classification-based proportion estimate for the classification errors observed in the type 2 dots.

Further details, in the form of equations, will be given in the section entitled "Analytical Results."

This set of steps has been greatly simplified. In practice, for example, one may be interested in more than two classes of crops, and there may be data problems such as cloud interference or bit drops. For simplicity, the simple two-class case without data problems will be assumed. Generalizations of the analytical expressions to be presented are readily available, and the experimentation to be reported did, in fact, take into account cloud interference and other data problems.

A number of questions arise as to the exact way each of the steps should be performed. For example: How many type 1 dots should be selected and what fraction of these should be reserved strictly for cluster labeling? What clustering algorithm should be used? How should clusters be labeled? How many type 2 dots should be selected and how should they be labeled? Also, what is the effect of using analyst interpreter (AI) labels for the type 1 and type 2 dots rather than generally unavailable, error-free ground-truth labels? The aim of the evaluation studies reported in this paper has been to help answer these questions by estimating the effects of the different factors on the sampling properties of the segment proportion estimators. A few results stem from analytical expressions; others, from experimentation guided by the analytical results.

ANALYTICAL RESULTS

Much of Procedure 1 is too complicated to allow its sampling properties to be fully described by tractable analytical expressions. Classification and clustering, for example, have been studied for a long

^aIBM, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

time by many people and yet their properties have been successfully evaluated in only the simplest, single-channel situations. Thus, Procedure 1's analytical results are limited to a few expressions which describe the sampling properties of the proportion estimators in terms of the results of a fixed, given classification of a segment.¹ These expressions are very important to LACIE because they identify the features of the labeling, clustering, and classification steps to which the proportion estimators are most sensitive.

Classification-Based Proportion Estimation

To start the discussion, some notation will be established. Assume that the N pixels in a segment have been classified and that N_1 have been assigned a small-grain classification label. Then, the ratio

$$\pi_s(C) = N_1/N \quad (1)$$

is called the classification-based estimator of the small-grain proportion of the segment. It is readily shown that $\pi_s(C)$ is a generally biased estimator with bias

$$b(C) = (1 - p_s)\theta_{so} - p_s\theta_{os} \quad (2)$$

where p_s is the true proportion of small grains in the segment, θ_{os} is the probability of erroneously classifying a small-grain pixel as other, and θ_{so} is the probability of erroneously classifying an other pixel as small grain. The bias in the classification-based estimator thus depends entirely on the error probabilities and, since all the pixels in the segment are assumed to be classified, there is no within-segment error variance. One purpose of Procedure 1 has been to remove, or at least decrease, the bias of the proportion estimates. As will be seen, this decrease in bias is purchased at the cost of adding a random error, with variance, to the estimates.

¹In particular, statements regarding expected values and variances of proportion estimators are conditional on the results of a fixed classification rule.

Analyst-Based Proportion Estimation

At this point, analyst labeling errors will be discussed. Results from this discussion will be used in developing the bias and variance of the Procedure 1 proportion estimator. Consider a situation in which an analyst "classifies" a subset of the pixels in a segment by assigning a label to each of a set of randomly chosen pixels. If the analyst labels a total of n pixels and assigns a small-grain label to n_1 of them, then the analyst-based estimate of the segment's small-grain proportion is

$$\pi_s(A) = n_1/n \quad (3)$$

Unless the analyst is using ground truth to label the pixels, this, also, is a biased estimator with bias

$$b(A) = (1 - p_s)A_{so} - p_sA_{os} \quad (4)$$

where A_{os} is the probability that the analyst labels a small-grain pixel as other and A_{so} is the probability the analyst labels an other pixel as small grain. Since the analyst has labeled a sample of the pixels in the segment, this estimator has a variance. Assuming that pixel labels are independent,² this variance is approximately

$$\alpha_s = p_s(1 - A_{os}) + (1 - p_s)A_{so} \quad (5)$$

where

$$V[\pi_s(A)] = \frac{\alpha_s(1 - \alpha_s)}{n}$$

²Since agricultural crops are grown in fields, pixels labels are clustered rather than independent. However, assuming a very small sample constrained to be well spread throughout the scene, independence is probably not a bad assumption.

is the probability that the analyst will assign a small-grain label to any randomly chosen pixel. The approximation is accurate in this case since the analyst would usually label a very small fraction of the 22 932 pixels in a segment.

Procedure 1 Proportion Estimation

Procedure 1 attempts to correct for the errors in the classification estimator $\pi_s(C)$ by using estimates of the misclassification probabilities θ_{os} and θ_{so} to adjust $\pi_s(C)$. According to the description of Procedure 1 in the introduction, classification in Procedure 1 depends on a sample of type 1 dots used first to guide a clustering run and then to label the resulting clusters. To estimate the classification error rates, a second sample of pixels, the type 2 dots, are selected at random, labeled by the analyst, and used to evaluate the classifier and to estimate the small-grain (SG) proportion. Table I shows the results of such a step. Here, m_{ij} is the number of the m type 2 dots that were classified into the i th class and assigned a j th class label by the analyst. A plus subscript denotes summation over that subscript; e.g., $m_{1+} = m_{11} + m_{12}$ is the number of type 2 dots that were assigned a small-grain label by the classifier. The analyst, of course, does not know the classification result when the labels are assigned. In fact, the type 2 dots are selected and labeled together with the type 1 dots before the classification is performed. This labeling is done for convenience since the variance of the proportion estimators can be reduced by controlling the numbers m_{1+} and m_{2+} . This point will be addressed later.

The classification bias, equation (2), can be put into another form by defining two alternate measures of classification error. Let λ_{os} be the probability that a pixel classified as small grain is, in fact,

an other pixel, and let λ_{so} be the probability that a pixel classified as other is, in fact, a small-grain pixel. Then, since $p_s \theta_{os} = (1 - \pi) \lambda_{so}$, equation (2) leads to the expression

$$b(C) = \pi_s(C) \lambda_{os} - [1 - \pi_s(C)] \lambda_{so} \quad (6a)$$

or alternately

$$\begin{aligned} p_s &= \pi_s(C) (1 - \lambda_{os}) \\ &= [1 - \pi_s(C) \lambda_{so}] \end{aligned} \quad (6b)$$

If the error probabilities λ_{os} and λ_{so} were known, this relation could be used to generate an unbiased proportion estimator from $\pi_s(C)$. The error probabilities are not known, but they can be estimated using the type 2 data results as

$$\hat{\lambda}_{os} = m_{12}/m_{1+}$$

and

$$\hat{\lambda}_{so} = m_{21}/m_{2+}$$

Using these, the Procedure 1 proportion estimator is

$$\hat{p}_s = \frac{m_{11}}{m_{1+}} \pi_s(C) + \frac{m_{21}}{m_{2+}} [1 - \pi_s(C)] \quad (7)$$

TABLE I.—Results of Comparing the Classification and Analyst Labels for the Type 2 Dots

Classification label	Analyst label		Total
	SG	Other	
SG	m_{11}	m_{12}	m_{1+}
Other	m_{21}	m_{22}	m_{2+}
Total	m_{+1}	m_{+2}	$m_{++} = m$

It is relatively easy to show that the Procedure 1 proportion estimator \hat{p}_s is a maximum-likelihood estimator of the segment's small-grain proportion under the assumptions that (1) $\pi_s(C)$ is not a random variable, (2) the type 2 dots are a multinomial sample from the segment, and (3) the analyst labels are error free. If the analyst labels have errors, \hat{p}_s is still a maximum-likelihood estimator, but it estimates the proportion of the segment the analyst would have labeled as small grains rather than the true proportion of small grains in the segment. The

fact that \hat{p}_s is a maximum-likelihood estimator under the reasonable multinomial assumption implies that this proportion estimator has the usual large-sample optimality associated with maximum-likelihood estimators and probably cannot be improved upon easily as long as information is limited to the classification results and the type 2 dot labels. Also, the estimator is asymptotically normal, with asymptotic mean and variance given by standard maximum-likelihood theory.

Since $E(m_{11}/m_{1+}) \doteq 1 - A_{os}$ and $E(m_{21}/m_{2+}) \doteq A_{so}$, the bias in \hat{p}_s is virtually the same as the bias in $\pi_s(A)$ given by equation (4). That is,

$$b(\hat{p}_s) \doteq b(A) = (1 - p_s)A_{so} - p_s A_{os}$$

The variance of \hat{p}_s can be derived by noting that it is a standard form of a regression estimator used in sample survey theory. That is, if one attaches random variables X_i and Y_i to the i th type 2 dot where

$$X_i = \begin{cases} 1, & \text{if the dot is classified as small grain} \\ 0, & \text{otherwise} \end{cases}$$

and

$$Y_i = \begin{cases} 1, & \text{if the dot has a small-grain analyst label} \\ 0, & \text{otherwise} \end{cases}$$

then \hat{p}_s is equivalent to a regression estimator of the mean of the Y_i 's and can be expressed as

$$\hat{p}_s = \hat{a} + \hat{b}[\pi_s(C) - \bar{X}] \quad (8)$$

where \hat{a} and \hat{b} are least squares estimators calculated from the (X_i, Y_i) pairs. Using this and following reference 1, it is seen that the variance of the Pro-

cedure 1 estimator is closely approximated by

$$\begin{aligned} V(\hat{p}_s) &= (1 - \rho^2) V[\pi_s(A)] \\ &= (1 - \rho^2) \alpha_s (1 - \alpha_s) / m \end{aligned} \quad (9)$$

where $V[\pi_s(A)]$ is the variance of the analyst-based proportion estimator. The ρ in equation (9) is the correlation between the random variables X_i and Y_i and is equal to

$$\rho = \frac{P(S,S) - \pi_s \alpha_s}{\sqrt{\pi_s (1 - \pi_s) \alpha_s (1 - \alpha_s)}} \quad (10)$$

where $P(S,S)$ is the probability that a randomly chosen pixel is both classified and labeled by the analyst as a small-grain pixel.³ If the type 2 dots were selected after classification so that the number of dots classified as small grains m_{1+} were equal to the product $m\pi_s(C)$, these would be exact equations for the bias and variance. However, since the type 2 dots are selected before classification, m_{1+} is a random variable. The error in the approximations given in equations (4) and (9) for the bias and variance of \hat{p}_s is caused by the probabilities that m_{1+} can take the values 0 and m . If m and p_s are moderately large, these will be very small probabilities and therefore equations (4) and (9) will give very good approximations. The true bias can be either larger or smaller than that given in equation (4), but true variance always exceeds that given by equation (9). In cases for which both the true variance and $V(\hat{p}_s)$ have been calculated, $V(\hat{p}_s)$ was found to underestimate by about 2 to 4 percent, with a maximum error of about 10 percent of the true variance. Equation (9), then, should be sufficient for the purpose of evaluating Procedure 1.

Since both the analyst-based and the Procedure 1 proportion estimators have the same expected value, it is useful to consider the ratio of their variances as a measure of their relative efficiencies. Remembering

³ $1 - \rho^2$ is equivalent to the variance reduction coefficient R discussed in the paper by Heydorn et al.

that Procedure 1 uses two sets of dots, the variance of an analyst-based estimator using these dots is

$$v_1 = \frac{\alpha_s (1 - \alpha_s)}{m_1 + m_2}$$

whereas the variance of the Procedure 1 estimator is

$$v_2 = (1 - \rho^2) \frac{\alpha_s (1 - \alpha_s)}{m_2}$$

where m_1 and m_2 are the numbers of type 1 and type 2 dots.⁴ The efficiency of Procedure 1 with respect to the analyst estimate is

$$E = \frac{v_1}{v_2} = \frac{m_2}{m_1 + m_2} \frac{1}{1 - \rho^2} = \frac{m_2}{m_1 + m_2} \frac{1}{R} \quad (11)$$

If $R < [m_2/(m_1 + m_2)]$, the variance of the Procedure 1 estimator is less than that of an analyst-based estimator and there has been a gain in sampling efficiency due to the clustering and classification processing in Procedure 1. Otherwise, there is a loss of efficiency and a better estimate would result by skipping the machine processing. It should be noted that the efficiency of Procedure 1 would traditionally be defined as the reciprocal of that in equation (11). This definition, however, is well established in the LACIE community and will be adopted here. In the sequel, R will be referred to interchangeably as efficiency or variance ratio.

Note that altering the clustering and classification schemes of Procedure 1 will not affect the bias of its proportion estimators since this depends only on the analyst labeling errors. Only the variance would be

⁴In the implementation of Procedure 1, boundary dots (i.e., pixels on or "near" the boundary of a field) were not used as type 1 dots. In this discussion, however, it is assumed that all pixels are used and that the same selection and labeling rules apply to both type 1 and type 2 dots.

affected by changing the clustering or classification and this only through $R \approx 1 - \rho^2$. The value R , then, should provide the most sensitive measure of the effect of these changes. In practice, one has the classification estimator $\pi_s(C)$ and the type 2 dot results as shown in table I from which to estimate R or, equivalently, ρ . As seen in equation (10), ρ is a function of the proportion of the segment which the analyst would have labeled as small grains and the classification errors with respect to the analyst labels. Two heuristic estimators of R have been proposed and studied. One of these uses equation (10) with $P(S,S)$ and α_s replaced by their estimators based on the type 2 dots; i.e.,

$$\hat{P}(S,S) = m_{11}/m$$

and

$$\hat{\alpha}_s = m_{+1}/m$$

The term π_s comes from the classification results. This first estimator of ρ has the form

$$\hat{\rho}_1 = \frac{m_{11} - m_{+1}\pi_s}{\sqrt{\pi_s(1 - \pi_s)m_{+1}m_{+2}}} \quad (12)$$

The second estimator of ρ was developed by assuming that, since it is a correlation coefficient, there is very little information to be gained by using the $N - m$ pixels which are not included as type 2 dots. Consequently, the standard product moment estimator of the correlation between the X_i 's and Y_i 's from the type 2 dots is used to estimate ρ . Because of the special 0,1 nature of these variables, this reduces to a particularly simple function of the entries in table I; i.e.,

$$\hat{\rho}_2 = r = \frac{m_{11}m_{22} - m_{21}m_{12}}{\sqrt{m_{1+}m_{2+}m_{+1}m_{+2}}} \quad (13)$$

In either case, the estimator of R is $1 - \hat{\rho}^2$.

Table II includes the results of a small Monte Carlo experiment to compare these two estimators.

TABLE II.—Comparison of the Sampling Properties of Two Estimators of the Correlation Between Analyst Labels and Classification Labels

Proportion small grains	Classification error rates		Errors between true and estimated correlations					
			Product moment (r)			Mixed ($\hat{\rho}_1$)		
			θ_{os}	θ_{so}	Av	SD	MSE	Av
0.3	0.6	0.2	-0.02	0.13	0.02	-0.03	0.17	0.04
.3	.5	.2	.02	.14	.02	.09	.21	.05
.3	.4	.2	.02	.15	.02	.03	.24	.05
.3	.3	.2	-.03	.11	.01	-.03	.16	.03
.5	.6	.2	0	.13	.02	.02	.25	.06
.5	.5	.2	.01	.6	.03	.05	.27	.07
.5	.4	.2	.02	.11	.01	.08	.19	.04
.05	.5	.2	.001	.19	.01	-.003	.21	.02

For this experiment, sets of 22 932 labels were generated for simulated segments using binomial distributions appropriate to the segment wheat proportions and classification errors given in the table. The experiment was replicated 20 times for each set of conditions with the resulting observed averages, standard deviations (SD's), and mean squared errors (MSE's) of the differences between the true and the estimated correlations. In every case, the product moment estimator performed better than the estimator that uses the classified small-grain proportion. The observed mean squared error was at least twice as large in every case for the mixed as compared to the product moment estimator. For this reason, all R values reported in this paper are calculated as $1 - r^2$, where r is the product moment correlation given by equation (13). As a point of interest, r is a standard measure of association for contingency tables. It is asymptotically normal, is unbiased, and has asymptotic variance (ref. 2)

$$V(r)\sigma_r^2 = \frac{1}{m} \left\{ 1 - \rho^2 + \left(\rho + \frac{1}{2}\rho^3 \right) \cdot \frac{(P_{1+} - P_{2+})(P_{+1} - P_{+2})}{\sqrt{P_{1+}P_{2+}P_{+1}P_{+2}}} + \frac{3}{4}\rho^2 \left[\frac{(P_{1+} - P_{2+})^2}{P_{1+}P_{2+}} + \frac{(P_{+1} - P_{+2})^2}{P_{+2}P_{+2}} \right] \right\}$$

where P_{ij} is the probability associated with cell i,j in table I. Since the expected value of $\hat{R} = 1 - r^2$ is

$$E\hat{R} = 1 - \left[(Er)^2 + V(r) \right] = 1 - \rho^2 - V(r)$$

\hat{R} underestimates R on the average. It has the expected value

$$E\hat{R} = \left(1 - \frac{1}{m} \right) R - \frac{1}{m} f(\rho, P)$$

where $f(\rho, P)$ is small with respect to m . Thus, at least approximately, \hat{R} underestimates R by a factor of $1 - (1/m)$. Nearly all the R values reported in this paper were calculated using 60 to 140 dots; thus, the attenuation factor can be ignored.

The remainder of this paper consists of results from experimental studies of Procedure 1. These are given in terms of proportion estimation errors, probabilities of correct classification, and the variance ratio R , which measures the return from machine processing.

EXPERIMENTAL RESULTS

A number of experimental studies have been conducted to evaluate the performance of Procedure 1.

This paper will be limited to a discussion of the results from three of the experiments and an evaluation of LACIE Procedure 1 proportion estimates for some blind-site segments.

The type 1 and type 2 dot labels used in the first two experiments were derived from ground-truth information; therefore, the corresponding Procedure 1 proportion estimates will be unbiased. The purpose of these experiments was to gain an understanding of Procedure 1 by validating the analytical expressions for means, variances, and sampling efficiencies and by studying effects of modifying some of the Procedure 1 parameters. Ground-truth labels were used in these studies because they removed a source of variation—the analyst labeling errors—from the experiments. Both analyst labels and ground-truth labels were used in the remaining experiment and in evaluating the LACIE estimates. Their results will include comparisons between results using ground-truth and analyst labels.

Experiment 1

The first experiment was designed to validate the analytical expressions and to measure effects of modifying the numbers of cluster starting vectors, cluster labeling vectors, and channels used in processing the imagery. Each of these three factors was used at two different levels in a 2^3 factorial experiment over the four segments identified in table III. This table contains the segment numbers and locations, the ground-truth small-grain (in this case, wheat) proportions, and a variance factor which is N times the variance of an analyst-based proportion estimator calculated from N accurate pixel labels. The factor levels used in this experiment are

1. Number of cluster starting vectors = (20,40)

2. Number of additional labeling vectors used with the cluster starting vectors to label cluster = (0,20)

3. Number of channels used in processing the imagery = (4,8)

All clustering runs were performed using the ERIPS Iterative Self-Organizing Clustering System (sometimes called ISOCLS; herein called Iterative). The cluster parameters were taken from the results of a study on clustering made by the Mission Planning and Analysis Division (MPAD) of the NASA Johnson Space Center.⁵ Classifications were performed using the standard LACIE mixture density algorithm with cluster statistics used as subclass parameters and prior probabilities set proportional to cluster population sizes.

The first step in evaluating the experimental data was to validate the expressions for the mean values and variances of the Procedure 1 proportion estimators. As anticipated, test results showed that these expressions were valid and useful. Other results from this small experiment were rather disappointing. The experiment produced evidence that, though Procedure 1 has a potential for generating estimates that are very efficient when compared to hand counting by an analyst, it did not do so for most of the segments at most of the factor levels. Perhaps not surprisingly, given the small number of segments, none of the three experimental factors had a significant effect on the Procedure 1 proportion estimators. More surprising was an almost complete lack of consistency in the effects from segment to segment. For example, when processing segment 1591 using 20 starting and labeling vectors, the R value was increased from 0.76 to 0.999 with the addition of a second acquisition. Under the same conditions, the R value decreased from 0.82 to 0.26 for segment 1988.

The distribution of the observed R values and averages of the R values for combinations of the experimental factors are shown in figure 1.⁶ The me-

TABLE III.—Segments Used in Experiment 1

Segment		Ground-truth wheat proportion	N × variance of analyst-based wheat proportion
Number	Location		
1965	North Dakota	0.42	0.24
1884	Kansas	.37	.23
1591	Nebraska	.05	.05
1988	Kansas	.33	.22

⁵A. D. Wylie and W. C. Beam, "MPAD LACIE Clustering Study." JSC Internal Note 76-FM-116, NASA Johnson Space Center, 1977.

⁶The "box" plot in figure 1 is a stylized histogram showing, from top to bottom, the maximum, 75th percentile, median, 25th percentile, and minimum of the data. The box shape is merely to aid the eye and its width has no meaning in this paper.

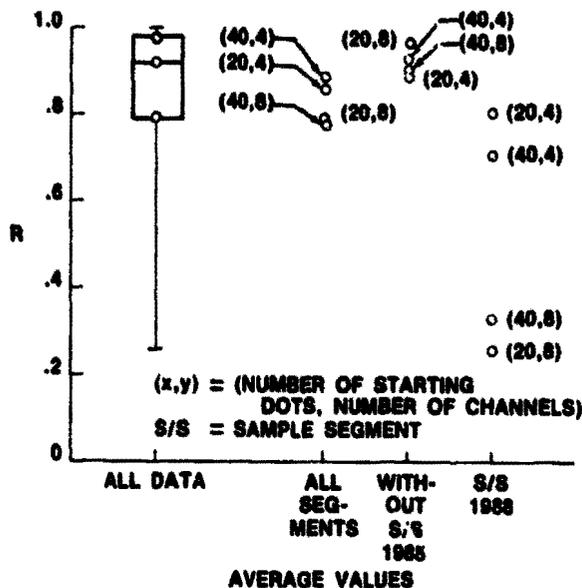


FIGURE 1.—Procedure 1 sampling efficiencies (R) from experiment 1.

dian R is about 0.92, with 75 percent of the values falling above 0.78. All of the small R values are from segment 1988 with eight-channel imagery.

Figure 2 illustrates the relationships between the measured sampling efficiencies and the variances of the proportion estimators. This plot shows the variances of two different analyst-based proportion estimators and two proportion estimators based on Procedure 1. The difference between the analyst-based estimators is that one assumes that only the type 2 dot labels are used to generate the estimator, whereas the other assumes that an additional 20 labels are available which would have been used as type 1 dots in Procedure 1. The two curves for Procedure 1 correspond to the best (0.259) and worst (0.823) observed R values from this segment. Since analyst-based estimates would be calculated using all available pixel labels, only the best case for segment 1988 shows a gain in sampling efficiency due to machine processing, but it is a very significant gain.

The results of this experiment indicate a weakness in the Iterative clustering algorithm since the R values should improve as more information is provided through additional starting vectors and acquisitions. The results may also imply that the acquisition selection technique is faulty. This implication may be true but it is not pertinent since the clustering

algorithm should be capable of extracting some information from every acquisition or at least it should not give a degraded performance with additional acquisitions. Note in this regard that, excepting clouds, etc., which were accounted for in this experiment, all acquisitions contain some information; poorly selected acquisitions do not contain poor data, merely a suboptimal choice of data.

To better evaluate the clustering algorithm, the cluster assignments of the labeled pixels were tabulated for each clustering run. An example of such a tabulation is given in table IV, which shows cluster assignments for the labeled pixels in segment 1988 when clustered using 20 starting vectors and 4 channels of data. Judging by the pixel assignments, there were at most four pure clusters (i.e., clusters 7, 14, 17, and 18) and none of these was pure wheat. Also, to point out a worst case, cluster 12 was apparently half wheat and half other. Problems with the clustering labeling algorithm also were indicated since cluster 4, which almost certainly contained between 79 percent and 99 percent class other, changed label from other to wheat when 20 additional labeling dots were provided. Results from the remaining segments were similar to those shown in table IV.

Since other experiments were larger than this one, discussion of the proportion estimates from this experiment will be omitted. The main conclusion drawn from this experiment was that Procedure 1 has potential to produce significant sampling efficiencies and, consequently, good proportion estimates, but it does require improvement if it is to attain this goal. In particular, performance of the ERIPS Iterative clustering algorithm appears poor despite the years of use and of study devoted to its development. One result of this finding was that a

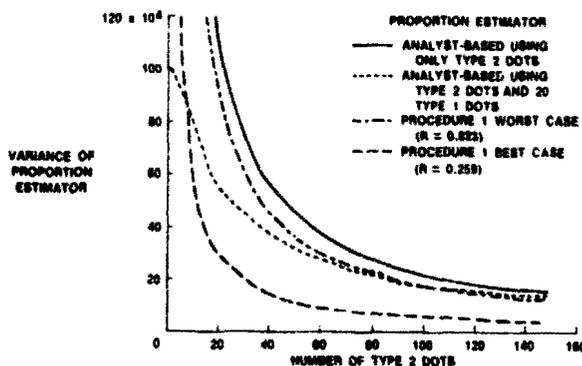


FIGURE 2.—Effect of Procedure 1 efficiencies on variances of proportion estimators for sample segment 1988.

nearest neighbor clustering algorithm was considered and subsequently adopted for LACIE. The next experiment to be discussed includes a comparison of the performances of Procedure 1 when used with these two clustering algorithms.

Experiment 2

The second experiment was designed to provide a detailed evaluation of Procedure 1 when ground-truth labels are available. A set of 13 Phase II segments was processed using numerous variations on the steps of Procedure 1. These segments are identified in table V, which also gives their ground-truth small-grain proportions, estimates of their small-grain proportions based on counting labels from (approximately) 209 pixels, and a run type. The seven segments with run type "sequential" were processed using, in turn, one, two, three, and four acquisitions of data. The remaining six "T-4" segments were processed using only four acquisitions. The acquisitions were selected by an automated selection algorithm developed for Procedure 1.

All segments were processed using both the iterative and nearest neighbor (NN) clustering algorithms, and proportion estimates were generated both with and without the classification step of Procedure 1. The latter without-classification process used the cluster label for each pixel in a cluster as that pixel's classification. The with-classification process used the standard Procedure 1 mixture density classification algorithm.

Proportion estimates were also produced with and without the "bias correction" step of Procedure 1. The without-bias-correction estimate is merely the proportion of the segment assigned a small-grain label by the classification or, as appropriate, clustering algorithm. Bias correction was done using ALL, 100, 75, 50, 25, and 10 type 2 dots, where "ALL" is interpreted as all labeled pixels not used as type 1 dots. This variation of the number of type 2 dots is not particularly interesting since its effects are completely explained by the analytical expressions given earlier. It did, however, provide an added opportunity to validate these expressions.

Three different schemes were used for selecting the pixels to be labeled and used as type 1 or type 2 dots. In each, the pixels were selected from the 209 grid dots situated at the intersections of the grid lines found on photographic products of LACIE imagery.

Under the first scheme, all grid pixels not covered by clouds or heavy haze and not located within "designated other" areas were available for labeling. Under the second scheme, only field-center pixels and pure

TABLE IV.—Representation of Grid Points in Clusters, Segment 1988, 4 Channels, 40 Starting Dots

Cluster number	Total number of pixels	Number of labeled grid points	Number of grid points labeled wheat	Cluster label
1	1326	9	1	O
2	1191	12	7	W
3	3196	29	3	O
4	2349	25	3	O ^a W ^b
5	2285	26	3	O
6	2159	20	1	O
7	2525	16	-	O
8	233	2	1	W
9	1013	9	6	W
10	1759	11	8	O
11	781	7	4	W
12	756	6	3	W
13	690	6	4	O
14	58	3	-	O
15	1217	12	9	W
16	456	4	1	O
17	777	9	-	O
18	161	3	-	O

^a40 labeling pixels.

^b60 labeling pixels.

TABLE V.—Segments Used in Experiment 2

Segment Number	Location	Percent SG			Run type
		Ground truth	Dots	Phase II estimate	
1003	Colorado	19.8	25.82	25.5	Sequential
1090	Colorado	32.8	37.50	29.1	T-4
1961	Kansas	8.2	10.0	8.0	Sequential
1988	Kansas	33.0	32.73	28.5	Sequential
1865	Kansas	23.4	24.39	12.0	Sequential
1178	Kansas	15.5	18.18	16.0	Sequential
1574	Nebraska	8.2	16.32	15.9	T-4
1624	North Dakota	53.89	57.89	46.8	T-4
1967	North Dakota	34.5	36.76	30.0	T-4
1046	Oklahoma	23.1	20.00	14.6	Sequential
1238	Oklahoma	11.99	10.05	0	T-4
1978	Texas	48.4	44.44	17.0	Sequential
1084	Texas	16.09	22.08	.4	T-4

boundary pixels (i.e., pixels lying on boundaries between fields of the same class) were available for labeling. Under the third scheme, only field-center pixels were labeled.

The data analysis discussion starts with figure 3, which contains box plots of observed errors in the Procedure 1 estimates of the small-grain proportions. These estimates used either Iterative or nearest-neighbor clustering with 4, 8, 12, or 16 channels of data. The 4-, 8-, and 12-channel data consisted of observations from the 7 sequential segments and the 16-channel data consisted of observations from all 13 of the segments. Using the interquartile ranges (i.e., length of the box portion of the plots as a measure of variability⁷), the following can be seen.

1. Excepting the surprisingly small variance of the 12-channel Iterative clustering data, the variances decrease gradually as the number of channels increases, with a large decrease at 16 channels.

2. With four and eight channels, the observed variances are slightly smaller for nearest neighbor than for Iterative clustering. With 16 channels, the variances are virtually identical. The 12-channel Iterative clustering case again stands out as anomalous.

Also, given the number of observations in each box, there is no indication that the true medians of these distributions differ from zero. This result was anticipated since ground-truth labels were used throughout.

Figure 4 shows the observed Procedure 1 sampling efficiencies (R values) associated with the proportion estimators used in figure 3. The R values have a skewed distribution with most of the observations falling well above 0.9. This, of course, implies a very small gain in sampling efficiency for the machine processing as opposed to an analyst-based hand counting estimator. In fact, there is a loss of efficiency if the total number of type 1 and type 2 dots is taken into account. The situation improves somewhat, though not sufficiently, with 16 channels, where the median of the R values drops to about 0.75. Procedure 1 again shows some potential for producing good sampling efficiencies since, excepting the four-channel cases, there are some low values of R .

⁷The distribution of the interquartile range depends on both the number of observations and the true distribution of the data. With normally distributed data, its expected value tends to about $4/3\sigma$ as the number of observations increases.

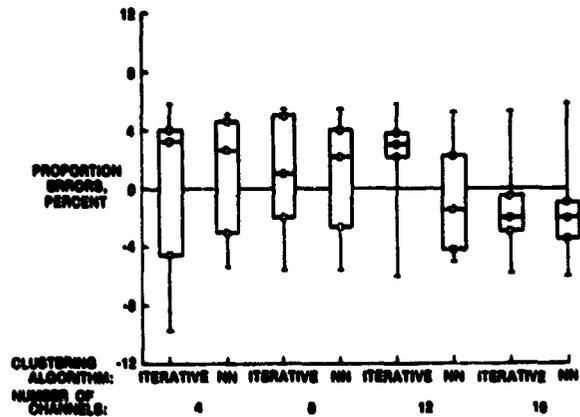


FIGURE 3.—Comparison of Procedure 1 proportion errors using iterative and nearest neighbor clustering (All-pixel labeling).

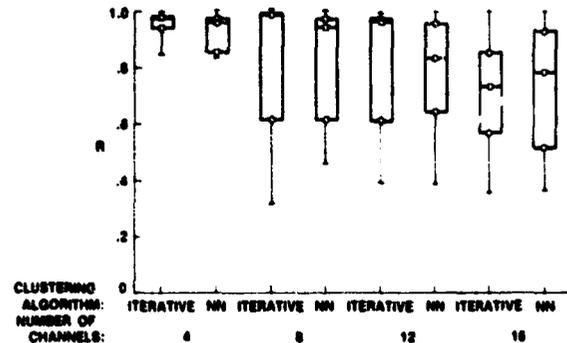


FIGURE 4.—Observed Procedure 1 sampling efficiencies (R) for iterative and nearest neighbor clustering (All-pixel labeling).

Figure 5 gives a scatter plot of the sampling efficiencies from Iterative and nearest neighbor clustering with 16-channel data. This plot includes data from both the All-pixel and field-center-pixel labeling schemes. As expected from the box plots, the All-pixel labeling data fall very close to a 45° line. The field-center-pixel labeling values do also, except for three values that are much higher for nearest neighbor than for Iterative clustering.

Although not explicitly shown to be so in the equations in the section entitled "Analytical Results," the sampling efficiencies are functions of the accuracy with which pixels are classified. Figures 6 and 7 show, respectively, the proportions of the labeled small-grain and other pixels that were correctly classified. In every case, the labeled other pixels were classified very accurately and the median accuracy increased slightly with the number of channels. The median classification accuracies of the

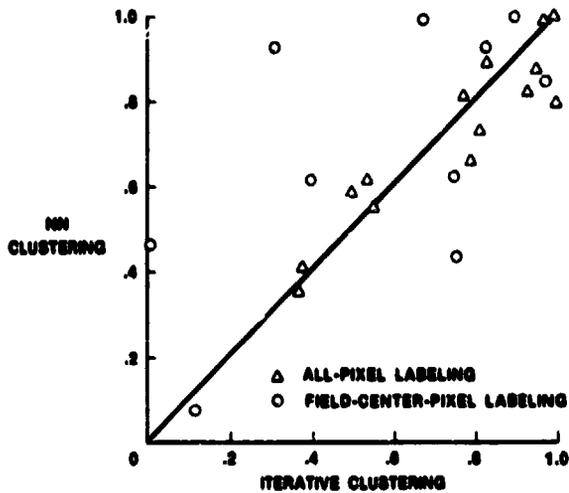


FIGURE 5.—Scatter plot of sampling efficiency (R) for nearest neighbor versus iterative clustering (16-channel data).

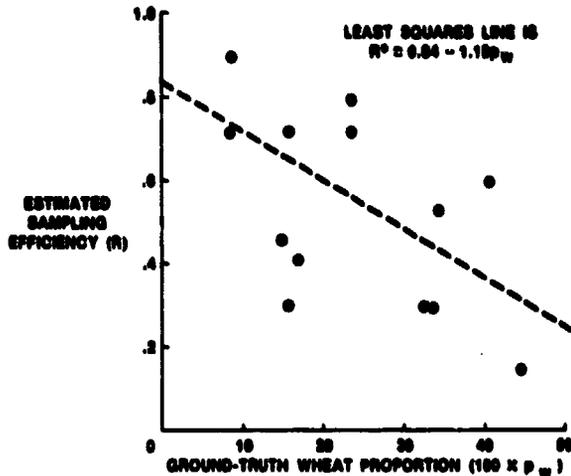


FIGURE 8.—Plot of the estimated sampling efficiencies (R) versus ground-truth wheat proportions.

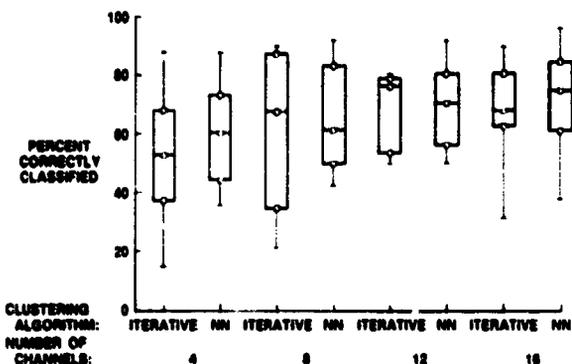


FIGURE 6.—Observed grid dot classification accuracies for all labeled small-grain pixels.

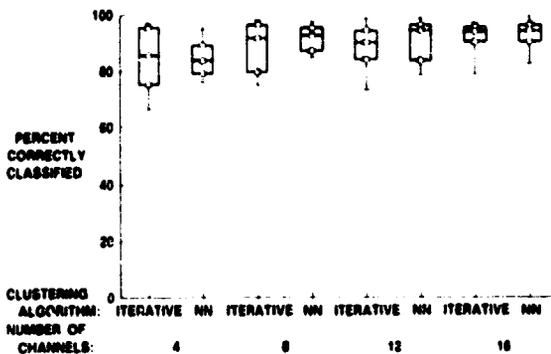


FIGURE 7.—Observed grid dot classification accuracies for all labeled other pixels.

small-grain pixels are low, starting between 0.5 and 0.6 for 4-channel data and increasing to about 0.7 with 16-channel data. This would seem to be the prime weakness of Procedure 1 as presently implemented—it does not classify small-grain pixels very accurately. It will be demonstrated subsequently that this weakness can be traced to the clustering algorithm.

Figures 8 and 9 show, respectively, plots of the estimated sampling efficiency R and probability of correct classification of wheat $\hat{P}(W|W)$ versus the true wheat proportion in these segments. The strong

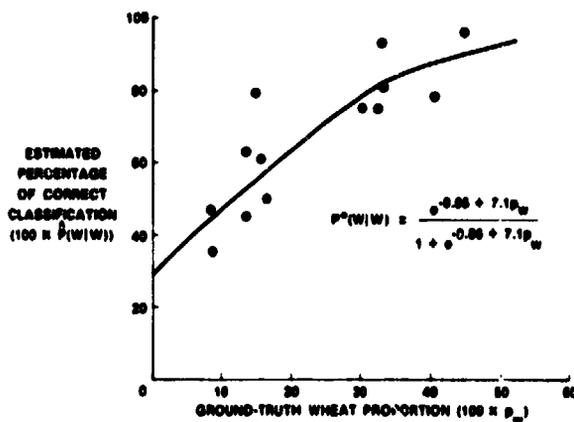


FIGURE 9.—Plot of $\hat{P}(W|W)$ versus ground-truth wheat proportion.

relationship of R and $\hat{P}(W|W)$ to the segment's wheat, or more generally small-grain, proportion is readily seen in these plots. As noted, the sampling efficiency R can be approximated by the least squares line

$$P^*(W|W) = \frac{e^{-0.85+7.1p_w}}{1 + e^{-0.85+7.1p_w}}$$

and the classification probability by the relationship

$$R^* = 0.84 - 1.18p_w$$

The main import of these figures is that the performance of Procedure 1 depends on the true proportion of small grains in a segment; there is improved performance with larger small-grain proportions. Experiments to detect the effects of changing the parameters of Procedure 1 should, therefore, use segments with a full range of small-grain proportions.

Figures 10 and 11 show two strong points of Procedure 1. Figure 10 contains box plots comparing the distributions of observed proportion errors with and without the bias correction step of Procedure 1. For each set of channels, the bias correction significantly decreases the incidence of large proportion errors and, as a result, the bias-corrected proportion estimates have a much smaller error variance than the others. The bias correction step of Procedure 1, then,

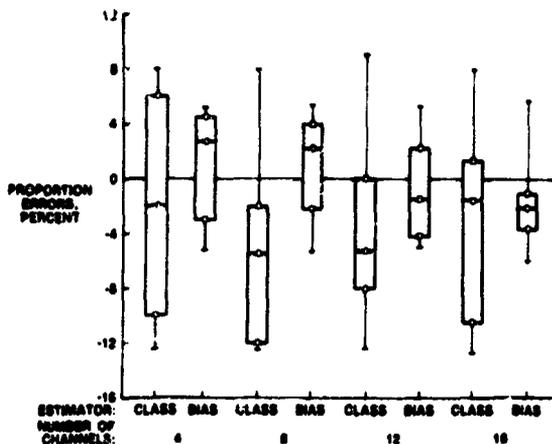


FIGURE 10.—Comparison of classification-based (CLASS) and bias-corrected (BIAS) proportion errors (nearest neighbor clustering, All-pixel labeling).

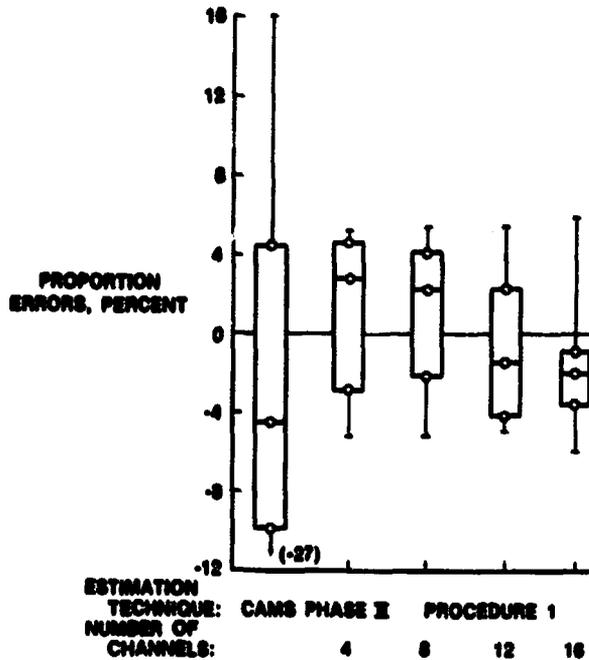


FIGURE 11.—Comparison of Phase II and Procedure 1 proportion estimation errors (nearest neighbor clustering, All-pixel labeling).

contributes to the efficiency of the Procedure 1 estimators.

Figure 11 compares errors in proportion estimates produced by the Classification and Mensuration Subsystem (CAMS) during Phase II of LACIE with similar errors in Procedure 1 estimates made using 4, 8, 12, and 16 channels of data. It is readily apparent that Procedure 1 represents a significant improvement even in the four-channel case, where sampling efficiencies were shown to be poor. Note that, since the Procedure 1 estimates are based on ground-truth labels whereas the CAMS estimates are based on analyst labels, the comparison in figure 11 is biased in favor of Procedure 1. It will be shown later that this bias does not appear to be great enough to invalidate the conclusion that Procedure 1 is an improvement over the procedure used in Phase II.

Figure 12 allows a comparison of the sampling efficiencies resulting from two pixel-labeling schemes. This figure gives box plots of the efficiencies resulting from the All-pixel and field-center-pixel labeling schemes for labeling the type 1 and type 2 dots. No large differences between the two schemes are seen, though the medians of the field-center-pixel efficiencies tend to be slightly lower (better) than those of the All-pixel efficiencies. As a result,

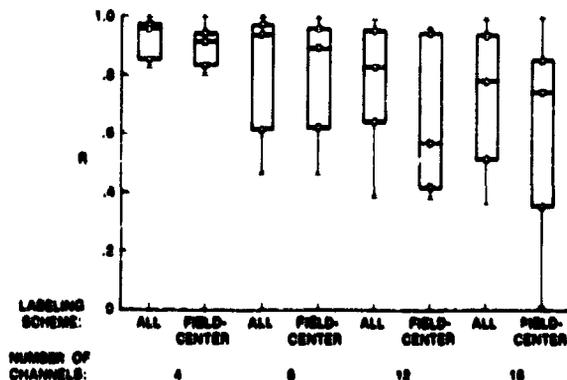


FIGURE 12.—Comparison of distributions of sampling efficiencies (R) with All-pixel and field-center-pixel labeling (nearest neighbor clustering).

LACIE has adopted a procedure in which only pure pixels are labeled for use as type 1 dots. Because a bias might result if boundary pixels were not considered in the bias correction step, all available type 2 pixels are labeled.

The final factor to be considered in the discussion of this experiment is the effect of the classification step in Procedure 1. The steps of Procedure 1 include a clustering run followed by classification to attach a small-grain or other label to each pixel in the segment. Since the clusters are also labeled as small grains or other, an alternate classification could easily be accomplished by attaching the cluster label to each pixel in a cluster. The number of pixels clustered into the small-grain class could then be counted and used to calculate a cluster-based proportion estimate. Bias correction would then be based on comparison of analyst and cluster labels for type 2 dots. The result would be a cluster-based Procedure 1 proportion estimator. Because of the use of the type 2 dots for bias correction, this estimator would have the same expected value as does the standard Procedure 1 estimator. The classification process requires computer resources and can be justified only if it produces improved sampling efficiencies with attendant decreases in the variances of the proportion estimators. The remainder of this discussion of experiment 2 will be devoted to a comparison of the classification-based and cluster-based Procedure 1 sampling efficiencies.

First, a small diversion. All R values presented thus far in the discussion of experiment 2 were calculated without use of classification results from the type 1 dots. All other labeled pixels were used in the calculation. The type 1 dots were excluded

because they are used in labeling the clusters and, consequently, will be more accurately classified than the remaining pixels. Their inclusion would cause a biased, optimistic estimate of the R values. Figure 13 illustrates this bias using 16-channel classification results. Every R calculated using the type 1 dots is smaller than the corresponding R calculated from all labeled pixels. It would be erroneous to accept an estimate of R based on all labeled pixels. On the other hand, there is a high correlation between the two estimators of R . If the purpose is to compare alternate approaches to Procedure 1, either estimator of R should provide a valid comparison.

The point of the preceding discussion is that the computer listings that contain the results of this experiment do not isolate the type 1 dots when tabulating cluster results. Consequently, all R values calculated for a cluster-based Procedure 1 proportion estimator used results from both the type 1 and the type 2 dots. The remaining figures will show these biased estimates of the sampling efficiencies. Even the classification-based results, then, will not be identical to those in previous figures.

Figures 14 and 15 provide a comparison of the sampling efficiencies from the cluster-based and classification-based schemes using the iterative and nearest neighbor clustering algorithms. There do not appear to be any significant differences between the R values from the two schemes, although the cluster-based values may tend to be slightly lower than the others. Figures 16, 17, and 18 are scatter plots of R values for a cluster-based procedure versus a

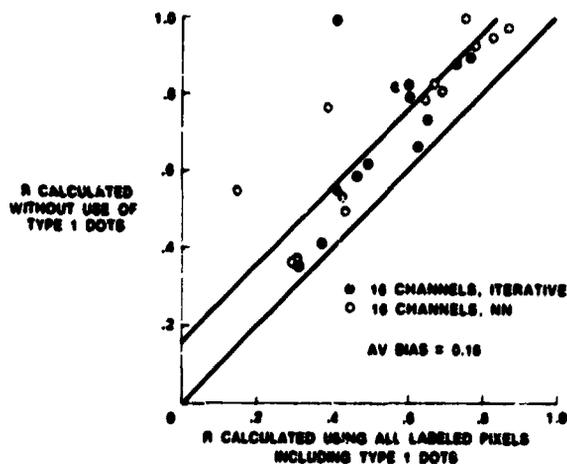


FIGURE 13.—Illustration of bias introduced when type 1 dots are used in estimating the sampling efficiencies (R) (16-channel data, All-pixel labeling).

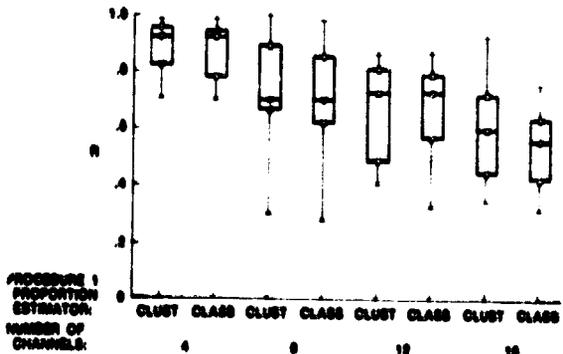


FIGURE 14.—Comparison of sampling efficiencies (R) for cluster-based (CLUST) and classification-based (CLASS) Procedure 1 proportion estimators (iterative clustering).

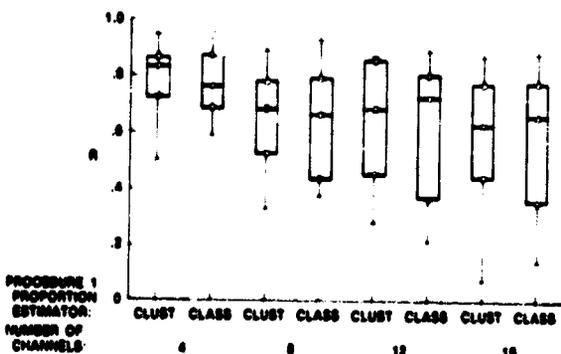


FIGURE 15.—Comparison of sampling efficiencies (R) for cluster-based (CLUST) and classification-based (CLASS) proportion estimators (nearest neighbor clustering).

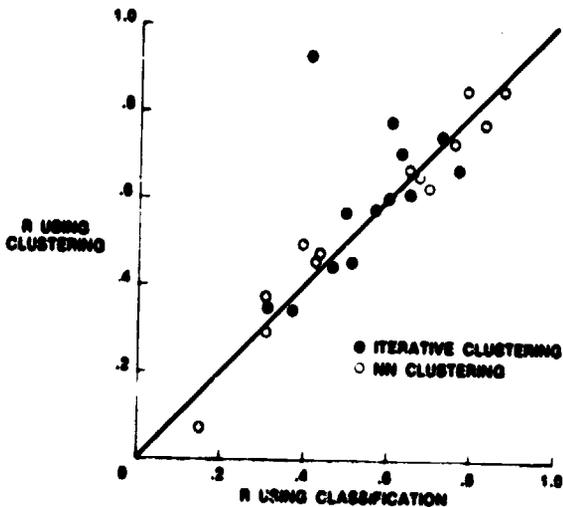


FIGURE 16.—Comparison of classification and clustering R values (All-pixel labeling, 16-channel data).

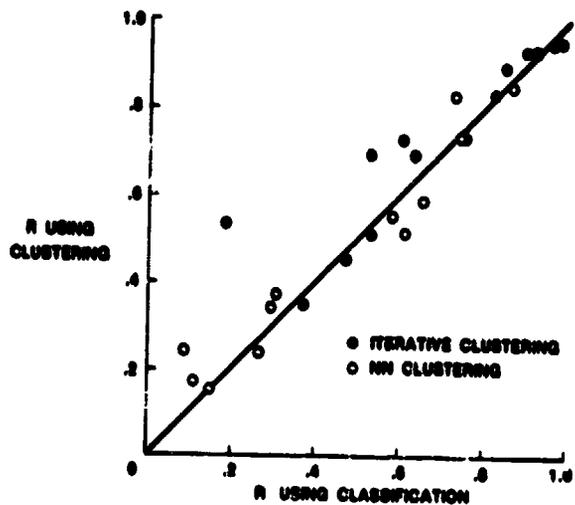


FIGURE 17.—Comparison of classification and clustering R values (pure-pixel labeling, 16-channel data).

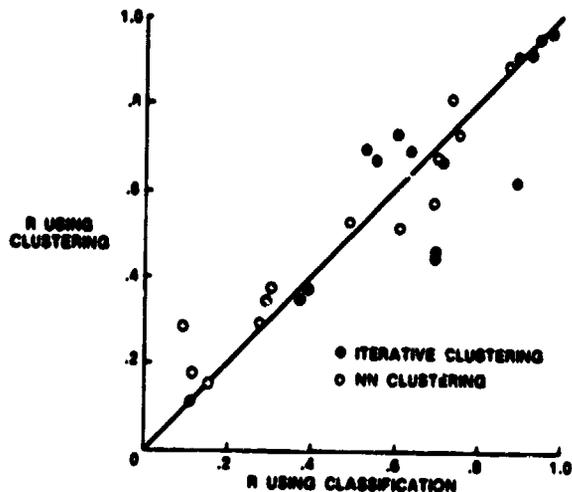


FIGURE 18.—Comparison of clustering and classification R values (field-center-pixel labeling, 16-channel data).

classification-based procedure using 16-channel data and the 3 pixel-labeling schemes. Again, there are no significant differences between the two sets of values and they appear very highly correlated in all plots. Slightly more than half of the cluster-based R values are higher than the corresponding classification-based values. This implies a slight, though hardly significant, advantage for the classification-based procedure. Finally, figures 19 and 20 show observed proportions of the labeled pixels that are correctly classified if the cluster labels are used as classification labels for the pixels in the clusters. A com-

parison of these figures with figures 6 and 7 will show that the two classification schemes produce virtually identical classification accuracies for the non-small-grain pixels. The medians of the percentage of the small-grain pixels that are correctly classified tend to be slightly higher for the classification results than for the cluster results. The differences between the two, however, are certainly not statistically significant. Thus, there does not appear to be a real gain in sampling efficiency due to the classification step in Procedure 1. Also, there is no degradation of results due to this step and, in particular, the poor classification accuracy results for the small-grain pixels are seen to stem directly from the cluster assignments.

Conclusions From Experiment 2

Results from experiment 2 reinforce and add to the conclusion from experiment 1. The main conclusions are as follows.

1. Procedure 1 produced much better proportion estimates than did the CAMS Phase II procedure. This observation must temporarily be tempered by the fact that the CAMS results are based on analyst labels, whereas the Procedure 1 results are based on ground-truth labels. Later results will demonstrate that use of the different labels does not invalidate this conclusion.

2. Procedure 1 was shown to produce very good sampling efficiencies for some segments under some conditions but, in general, is not more efficient than a hand counting procedure using both the type 1 and the type 2 dots.

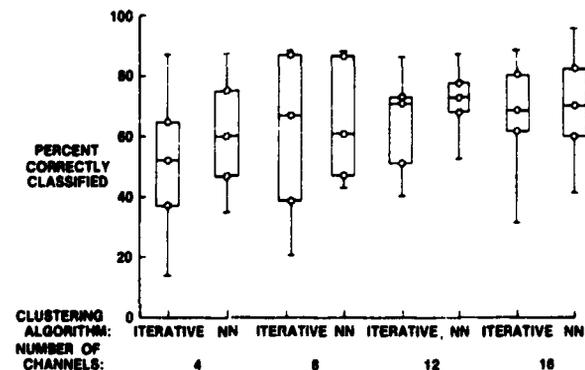


FIGURE 19.—Observed grid dot cluster classification accuracies for labeled small-grain pixels.

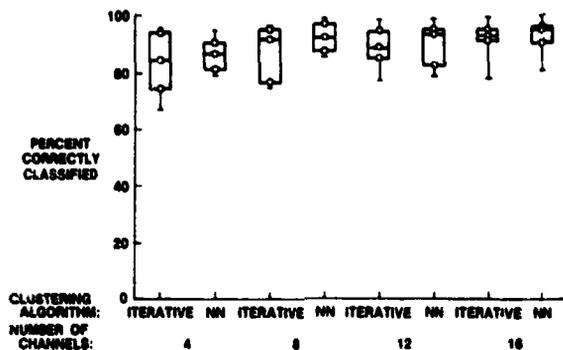


FIGURE 20.—Observed grid dot cluster classification accuracies for labeled other pixels.

3. The main problem with Procedure 1 is the high incidence of misclassification of small-grain pixels. This problem is caused by the clustering procedures, which do not adequately separate the small-grain pixels from the other pixels.

4. The nearest neighbor clustering algorithm appears to be slightly better than the Iterative algorithm in terms of its effectiveness in Procedure 1. This result is somewhat surprising since Iterative is a complicated algorithm whereas nearest neighbor uses an extremely simple, almost naive approach. The simplicity of the nearest neighbor algorithm holds hope for improvement in clustering.

5. For the present implementation of Procedure 1, the classification step following clustering does not appear to affect the sampling properties of the proportion estimators. This conclusion may bear further study with additional segments and, in particular, must be reconsidered if new clustering algorithms are proposed for Procedure 1.

6. Results improved as the number of acquisitions increased from one to four.

7. Finally, the bias correction step of Procedure 1 was shown to significantly improve the variances of the proportion estimators. Both the bias-corrected and uncorrected proportion estimators appear to be unbiased when ground-truth labels are used.

Experiment 3

The third experiment employed the nine segments identified in table VI to evaluate the effects on Procedure 1's performance of using analyst labels. This experiment used both the nearest neighbor and the Iterative clustering algorithms, and all segments

C-10

were processed using one, two, three, and four acquisitions. As in the earlier study, the nearest neighbor algorithm produced slightly better results than did the Iterative. Only nearest neighbor clustering results will be discussed here.

Figure 21 shows the distributions of observed differences between the Procedure 1 proportion estimates and the ground-truth small-grain proportions for these segments. It is readily seen that the estimates based on analyst labels have a significant negative bias and a very large variance. The ground-truth label results are unbiased, but they also have a large variance when compared with the ground-truth results from experiment 2.

Figure 22 shows the distributions of the sampling efficiency measures for these segments. Once again, there are some very good efficiencies; i.e., R values less than 0.6. Most of the efficiencies are poor, with the majority of the R values falling well above 0.8. The improvement in efficiency resulting from the use of more acquisitions seen in experiment 2 is not seen in these results.

One possible explanation for the poor results from this experiment as compared with those from experiment 2 is the presence of the four small-field segments in North Dakota. The locations of their proportion errors in the box plots in figure 21 are shown by the small circles in the boxes. The spread of these circles indicates that much of the variability is due to these segments. Unfortunately, there are not enough data to allow the questions raised by this experiment to be addressed.

The main conclusion from this study is that a larger study should be conducted using small-field segments with both ground-truth and analyst labels. Another conclusion of this study is that the analyst labeling errors caused a large negative bias in the pro-

portion estimates and introduced a large source of variability. Also, the analyst-based proportion estimates do not appear to change in distribution as the number of channels is increased. Because of the small number of segments and the mixture of small- and large-field segments, these conclusions must be tentative, pending a more complete study.

LACIE Operations Data⁸

The last set of data to be discussed contains actual LACIE proportion estimates for a set of segments with associated ground-truth labels. To provide a

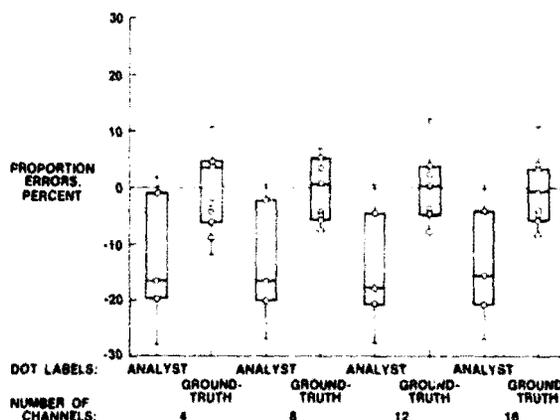


FIGURE 21.—Proportion errors from ground-truth and analyst labels.

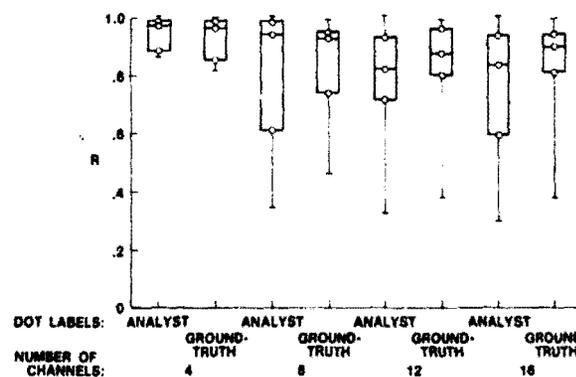


FIGURE 22.—Comparison of sampling efficiencies (R) for analyst and ground-truth labels with Procedure 1.

TABLE VI.—Segments Used in Experiment 3

Number	Segment Location	Ground-truth percentage	
		Wheat	Small grains
1660	North Dakota	26.6	34.6
1651	North Dakota	20.86	27.93
1642	North Dakota	39.2	57.6
1614	North Dakota	26.96	42.76
1003	Colorado	16.6	19.8
1046	Oklahoma	23.1	23.1
1961	Kansas	8.2	8.2
1988	Kansas	33.0	33.0
1865	Kansas	23.2	23.4

⁸The data evaluated in this subsection were provided by K. Havens of Lockheed Electronics Company. Much insight was gained from her evaluations of these data.

basis for comparison, these segments were processed by Procedure 1 using both ground-truth and analyst labels for the type 1 and type 2 dots. The data set is identified in table VII. Note that the segments were processed using two, three, or four acquisitions of imagery.

Figure 23 shows the observed errors between the estimated and the ground-truth small-grain proportions for these segments. As in experiment 3, there is a serious negative bias in the proportion estimates based on the analyst labels. This bias is caused by analyst labeling errors and indicates that the analysts tend to underestimate the small-grain proportions. This observation has been substantiated in studies of the analyst labeling errors. The estimates based on ground-truth labels are unbiased and have a slightly smaller variance than do the analyst-based estimates. There is a tendency for the variances from both label sets to decrease as the number of acquisitions is increased. This result is particularly encouraging in view of the fact that different segments were processed using the differing numbers of acquisitions.

Figures 24 and 25 show the distributions of the observed variance ratios and a comparison between the variance ratios from the analyst and ground-truth label processing. Only a few of these R values are lower than 0.6, indicating a general loss of sampling efficiency for Procedure 1 compared with analyst hand counting. (The value 0.6 is used since most of the segments were processed with about 40 type 1 and 60 type 2 dots leading to a break-even value of 0.6 for the R values.) There is no strong indication of an advantage in the R values from the ground-truth labeling over the analyst labeling results. This result implies that the increased variances noted in the proportion estimates based on analyst labels are caused mainly by segment-to-segment variation in the biases due to analyst labeling errors.

The main conclusion from this study is that analyst labels are biased and tend to underestimate the amount of small grain in the segments. If this bias source could be corrected (e.g., by providing additional information or training to the AI's), the Procedure 1 proportion estimators based on analyst labels would be very competitive with estimators based on ground-truth labels. This experiment again shows that Procedure 1 is not yet attaining the sampling efficiency required to be competitive with an analyst-based count estimator. This conclusion again indicates that Procedure 1 requires an improved clustering algorithm since problems can be traced to misclassification of small-grain pixels.

TABLE VII.—Segments Used in Evaluation of Procedure 1 Using Operations Data

Segment (a)	Location (county, state)	Number of acquisitions
1005 (W)	Cheyenne, Colo.	4
1032 (W)	Wichita, Kans.	4
1033 (W)	Clark, Kans.	2
1853 (W)	Ness, Kans.	3
1861 (W)	Kearny, Kans.	4
1512 (S)	Clay, Minn.	2
1520 (S)	Big Stone, Minn.	3
1544 (S)	Sheridan, Mont.	2
1739 (M)	Teton, Mont.	4
1582 (W)	Hayes, Nebr.	4
1604 (S)	Renville, N. Dak.	2
1606 (S)	Ward, N. Dak.	2
1648 (S)	Bowman, N. Dak.	2
1661 (S)	McIntosh, N. Dak.	2
1902 (S)	McKenzie, N. Dak.	2
1231 (W)	Jackson, Okla.	3
1242 (W)	Canadian, Okla.	4
1367 (W)	Major, Okla.	3
1677 (S)	Spink, S. Dak.	4
1690 (S)	Kingsbury, S. Dak.	3
1803 (W)	Shannon, S. Dak.	4
1805 (M)	Gregory, S. Dak.	4
1056 (W)	Moore, Tex.	3
1059 (W)	Ochiltree, Tex.	4
1060 (W)	Sherman, Tex.	2

^aW = winter wheat; S = spring wheat; M = mixed wheat.

A Postmortem on Figure 11

Figure 11 shows that the Phase II CAMS estimates of the small-grain proportions in the segments used in experiment 2 are much more variable than similar estimates produced by Procedure 1. In the discussion of Procedure 1, the fact that analyst labels were used in producing the CAMS results and ground-truth labels were used in producing the Procedure 1 results was identified as a possible source of this large difference in variability. However, the results from the operational data and experiment 3 indicate that the errors in analyst-based Procedure 1 proportion estimators have a standard deviation only about 1.4 times as large as the standard deviation of errors resulting from ground-truth labeling. If this is true, the results in figure 11 show a very real improvement for Procedure 1 over Phase II CAMS, even in the worst case in which a single acquisition is processed by Procedure 1.

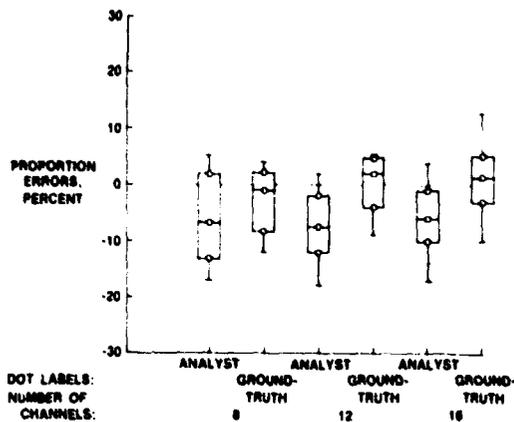


FIGURE 23.—Observed proportion errors using the operations data.

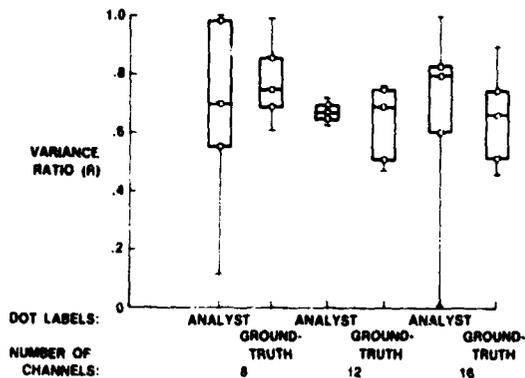


FIGURE 24.—Observed variance ratios using the operations data.

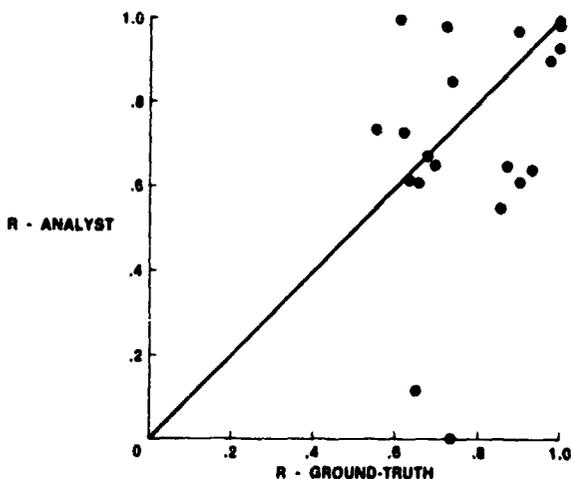


FIGURE 25.—Scatter plot of variance ratios from analyst labels versus variance ratios from ground-truth labels.

CONCLUSIONS

The main conclusion reached from these studies is that Procedure 1 needs an improved clustering algorithm if it is to attain a sampling efficiency competitive with simple analyst-based count estimates of proportions. There are new clustering routines available (e.g., AMOEBA, BCLUST, and ECHO) which use spatial as well as spectral information and, as a consequence, may be substantially better than the ERIPS algorithms. There is considerable room for improvement since the nearest neighbor algorithm, which produces better proportion estimators than does the iterative algorithm, is too simple to be optimal.

The results show that Procedure 1 has a potential for greatly increasing the sampling efficiencies since it did so for some segments. Also, Procedure 1 is a substantial improvement over the Phase II technology.

Currently, the primary effect of using analyst labels is that they introduce a negative bias in the proportion estimators. The Procedure 1 proportion estimators based on analyst labels have about a 40-percent larger standard deviation than do the estimators based on ground-truth labels. This 40-percent increase is due mainly to variations in the bias caused by analyst labeling errors rather than to increased R values.

Finally, the results of experiment 2, and other results not included in this study, show that the bias correction step in Procedure 1 produces a substantial decrease in the variance of the proportion estimator though it does not greatly affect the bias. Also, the classification following clustering does not appear to affect the properties of the estimators and could be eliminated.

REFERENCES

1. Kendall, M. G.; and Stuart, A.: *The Advanced Theory of Statistics*, Vol. 3. Hafner, 1968, pp. 218-219.
2. Bishop, Y. M. M.; Fienberg, S. E.; and Holland, P. W.: *Discrete Multivariate Analysis: Theory and Practice*. MIT Press, 1977, p. 381.

The Vegetative Index Number and Crop Identification

P. Ashburn^a

INTRODUCTION

Considerable research on the use of the vegetation index number (green number) (ref. 1) conducted by LACIE and other government agencies (ref. 2) revealed that the green-number approach to drought and yield monitoring has been successful. However, only a few studies (ref. 3) have examined the green-number approach for crop identification.

In this study, a vegetative index number of numerical value was calculated from the digital values of the Landsat system. The objective was to provide some measure of green growing vegetation. The purpose of this paper, as a pilot project, is to investigate the usefulness of the green numbers for schemes in crop identification and acreage estimation and to compare a new vegetation index number, the Ashburn Vegetation Index (AVI), with the Kauth-Thomas Vegetation Index (KVI) for crop identification schemes. Comparisons between the AVI and the KVI are given in table I. Table II shows the results of wheat acreage estimation using LACIE Procedure 1 (P-1) and the AVI for the eight LACIE sample segments used in this study. The process used by LACIE for crop acreage estimates for the 1976-77 crop season, P-1, will be described in another part of this paper. Visual results of the AVI may be found in the figures.

DESCRIPTION OF THE ALGORITHMS

Ashburn Vegetation Index

The AVI algorithm is 2 times band 7 minus band 5 with all the resulting negative values mapped to zero.

$$AVI = 2(7) - 5, \text{ AVI } 0 \text{ then set AVI} = 0 \quad (1)$$

The AVI is a straight linear equation (2 times band 7

minus band 5) of the Landsat data. This algorithm operates on the principle that as growing plants turn green, the chlorophyll in the leaves absorbs the red, thereby lowering the digital value in band 5; the green leaves increase the infrared reflectance, thereby increasing the digital value in band 7. Therefore, the greater the spread the higher the green number. When the values of these two bands are equal to or less than zero, the AVI indicates no growing vegetation for that pixel. The negative values are mapped to zero. Therefore, zero equals no vegetation and any positive value equals growing vegetation.

Kauth-Thomas Vegetation Index

The Kauth-Thomas (1976) Vegetation Index (KVI) is based on transformation of the four multispectral scanner (MSS) bands of the Landsat. They include a green vegetation index which is equal to $0.290MSS4 - 0.562MSS5 + 0.600MSS6 + 0.491MSS7$ and a soil brightness index which is equal to $0.433MSS4 + 0.632MSS5 + 0.586MSS6 + 0.264MSS7$. The green number, however, is derived from the following transformation.

$$y^i = Ax^i \quad (2)$$

where

$y^i = \begin{bmatrix} y_1^i \\ y_2^i \\ y_3^i \\ y_4^i \end{bmatrix}$ a vector representing the Landsat-1 version of the Kauth-Thomas transformation of $x^i(6)$; the superscript is the pixel number and the subscript is the Landsat channel.

0.4326	0.6325	0.5837	0.2641	Soil brightness
-.2897	-.5620	.5995	.4907	Greenness
-.8242	.5329	-.0502	.1850	Yellowness
.2229	.0125	-.5431	.8094	None such

^aUSDA Foreign Agricultural Service, Houston, Texas.

TABLE I.—A Comparison of the AVI and the KVI for Four Dates Over the Randall County Intensive Test Site

Pixel	Nov. 2, 1976 76307		Feb. 19, 1977 77050		June 6, 1977 77158		July 12, 1977 77194		AL ^a (d)	CL ^b (d)	GTL ^c (d)
	AVI	KVI	AVI	KVI	AVI	KVI	AVI	KVI			
Line 10											
10	24	23	8	10	14	16	0	4		W	
20	0	6	2	3	7	10	2	9	N	N	
30	0	4	3	5	0	4	0	3		N	
40	0	4	0	0	1	10	0	5	N	N	N
50	0	3	0	1	1	9	0	7	DO	N	N
60	0	4	0	1	0	8	0	4	DO	N	N
70	10	11	0	0	5	14	3	11	N	N	N
80	30	26	8	8	6	12	0	3		W	W
90	0	5	0	3	0	5	18	22		N	N
100	2	5	1	2	0	2	35	33		N	N
110	5	9	0	4	0	1	0	3	N	N	N
120	0	5	0	5	0	2	0	1		N	
130	0	2	0	2	0	2	0	4		N	
140	0	5	1	4	0	4	27	27		W	N
150	2	5	0	3	18	18	8	12		N	
160	0	2	0	2	0	5	0	3		N	N
170	0	3	4	7	0	3	3	11	W	W	W/B
180	0	2	2	3	15	23	24	23	N	N	N
190	5	4	4	5	14	17	20	19		N	N
Line 20											
10	0	1	0	5	16	21	17	23	N	N	N
20	8	4	0	4	0	11	2	8		W	N
30	0	3	0	1	1	8	0	5		N	N
40	2	5	0	0	1	11	0	7	DO	N	N
50	0	2	0	1	0	3	1	12	DO	W	N
60	8	8	0	1	0	10	54	44	N	N	N
70	11	12	4	6	10	14	0	4	W	N	W
80	11	9	2	3	12	17	0	10		N	W
90	0	3	0	0	0	8	43	34		N	N
100	2	6	0	4	0	3	21	24		N	N
110	4	6	0	3	0	6	5	11	N	N	N
120	0	3	0	3	14	20	5	9		N	W
130	3	6	0	4	1	11	45	38			
140	9	12	8	8	5	15	0	6		N	N
150	3	8	0	3	28	24	1	9		W	W
160	0	1	0	2	0	8	1	7	N	N	N
170	1	5	3	6	16	21	19	22		N	
180	0	1	0	2	21	26	22	23		N	N
190	0	2	0	3	1	12	9	16		N	

^aAL = analyst label from P.I

^bCL = classification label from P.I

^cGTL = ground-truth label

^dW = wheat, N = nonwheat, DO = designated other, W/B = wheat boundary, ITS = Intensive Test Site, N/B = nonwheat boundary, and TH = threshold

TABLE I.—Continued

Pixel	Nov. 2, 1976 76307		Feb. 19, 1977 77050		June 6, 1977 77158		July 12, 1977 77194		AL	CL	GTL	
	AVI	KVI	AVI	KVI	AVI	KVI	AVI	KVI				
<i>Line 30</i>												
10	2	7	0	2	4	12	2	10	N	N	N	
20	1	6	0	5	8	15	3	11		NN	NN	
30	2	5	0	4	0	3	0	0	N	NN	NN	
40	0	3	0	1	0	7	13	18		NN	N/B	
50	0	3	0	2	0	11	0	4	DO	NN	NN	
60	2	5	4	5	0	2	0	6	W	WW	WW	
70	0	3	3	4	0	3	0	3		W	NN	
80	6	10	0	1	0	3	0	8		TH	W	
90	10	10	0	4	0	8	17	18	N	NN	NN	
100	2	5	0	1	6	12	0	2	N	NN	W/B	
110	0	0	0	2	15	19	0	4		NN	W	
120	20	19	2	7	2	8	0	0		NN	W	
130	35	29	4	6	4	12	0	3		W	W	
140	1	3	0	6	8	11	10	11		TH	NN	
150	0	0	0	0	0	6	0	5	DO	NN	NN	
160	0	2	0	1	0	10	2	8	DO	NN	NN	
170	1	5	0	2	0	6	31	38	N	NN	NN	
180	0	0	0	3	24	29	39	38	N	NN	NN	
190	0	4	0	3	24	26	29	27	N	NN	NN	
<i>Line 40</i>												
10	0	6	0	5	0	11	0	8	N	N	N	
20	6	7	0	1	0	7	25	24		NN	NN	
30	4	8	8	7	17	17	0	8	W	WW	WW	
40	4	9	0	5	0	4	0	3	N	NN	NN	
50	0	6	0	3	0	2	0	1		NN	NN	
60	12	12	11	10	13	15	2	9	W	WW	WW	
70	12	12	4	4	0	7	3	8		W	W	
80	2	8	0	8	0	1	0	2		NN	NN	
90	8	10	0	3	0	2	31	26	N	NN	NN	
100	4	8	1	3	0	-1	0	2	N	NN	NN	
110	0	4	0	2	17	17	0	2		NN	W	
120	0	4	0	1	0	7	51	41		NN	NN	
130	0	2	0	4	4	11	51	41		NN	NN	
140	3	8	3	8	0	7	0	4		NN	NN	
150	0	2	0	1	0	10	6	10	DO	TH	NN	
160	3	2	0	0	0	8	0	11	DO	NN	NN	
170	0	6	0	1	33	34	43	36	N	N	N	
180	0	3	0	2	3	12	1	10	N			
190	0	1	0	0	3	14	6	14	N			

TABLE I.—Continued

Pixel	Nov. 2, 1976 76307		Feb. 19, 1977 77050		June 6, 1977 77158		July 12, 1977 77194		AL	CL	GTL
	AVI	KVI	AVI	KVI	AVI	KVI	AVI	KVI			
Line 50											
10	0	3	0	1	4	13	15	17			
20	14	14	1	3	4	14	0	2		W	W
30	3	6	0	1	0	10	33	30		N	N
40	3	6	0	0	0	6	2	9	N	W	N
50	3	6	0	3	0	5	46	40		N	N
60	14	12	10	9	5	12	2	10	W	W	W
70	20	18	8	10	12	14	2	9	W	W	W
80	0	4	0	2	0	3	3	8		N	N
90	7	10	0	0	0	5	39	37	N	N	N
100	0	5	0	3	0	4	0	2	N	N	N
110	7	6	0	6	0	3	0	0		N	N
120	5	7	0	1	0	3	0	2		N	N
130	5	7	0	2	0	6	31	28		W	N/B
140	5	5	0	4	0	1	0	7	N	N	N
150	0	7	0	5	0	1	28	26		N	N
160	0	1	0	2	0	1	39	36		N	N
170	0	2	0	1	0	10	0	1		N	N
180	9	9	8	10	8	18	0	3	W	W	W
190	0	1	0	0	0	7	28	26	N	N	N
Line 50											
10	0	4	0	2	0	10	5	10	DO	TH	N
20	0	2	0	3	6	14	2	10		N	W
30	0	4	0	2	0	9	14	15	N	N	W/B
40	0	2	0	2	6	15	37	33		N	N
50	0	0	0	-2	2	14	22	26	N	N	N
60	2	4	0	1	0	7	1	6	N	N	W/B
70	0	1	0	2	X	12	2	6		N	W
80	0	5	0	4	X	4	0	4	N	N	N
90	14	14	0	11	16	17	0	6		W	W
100	8	12	X	4	2	12	0	4		N	W/B
110	0	4	X	3	0	13	0	6		N	W/B
120	0	4	0	4	14	18	18	19	DO	N	N
130	1	6	0	3	0	3	0	5	DO	N	N
140	0	3	3	4	0	4	21	19		N	N
150	0	4	0	3	9	14	0	6		W	W
160	0	5	2	4	4	12	0	10	W	N	W
170	0	1	0	0	0	8	5	12		N	N/B
180	0	5	0	2	0	7	0	5		N	N
190	0	1	0	0	0	9	37	34		N	N

TABLE I.—Continued

Pixel	Nov. 2, 1976 76307		Feb. 19, 1977 77050		June 6, 1977 77158		July 12, 1977 77194		AL	CL	GTL		
	AVI	KVI	AVI	KVI	AVI	KVI	AVI	KVI					
<i>Line 70</i>													
10	1	5	0	4	2	10	4	9	DO	N	N	}	ITS
20	0	1	2	3	0	7	1	11	N	N	W		
30	0	5	0	4	8	11	9	15		W	W/B		
40	18	19	1	2	0	4	1	12	N	N	N		
50	14	15	1	6	0	8	11	14		W	W/B		
60	2	2	0	1	0	3	0	2		N	W		
70	0	2	0	2	0	5	2	2	N	N	W/B		
80	0	6	0	5	0	4	0	6	N	N	N		
90	6	8	0	2	0	2	56	44	N	N	N		
100	2	7	0	2	0	10	0	4		N	W/B		
110	0	5	3	5	0	7	0	3	W	W	W		
120	2	3	2	3	0	8	0	9	N	N	W		
130	0	2	0	2	0	4	31	27	N	N	W		
140	6	8	0	2	0	11	64	50		N	N		
150	2	7	0	3	3	9	0	1		N	N		
160	1	5	0	3	0	3	0	2		N			
170	0	5	0	2	0	5	0	8	DO	N			
180	0	7	0	3	0	4	0	6	W	W	N/B		
190	12	13	8	8	8	12	0	6	N	N	W		
<i>Line 80</i>													
10	2	5	0	4	5	13	2	11	DO	N	N	}	ITS
20	6	9	0	4	7	16	15	16	DO	N	N		
30	0	4	1	3	0	11	3	14		W	W		
40	6	9	0	2	0	9	0	8		N	W		
50	4	7	0	1	0	4	33	28		N	W		
60	0	6	0	4	0	3	18	22	N	N	N		
70	3	8	0	4	0	5	9	15	N	N	N		
80	0	3	0	0	0	4	67	53	N	N	N		
90	0	2	0	3	0	1	57	48	N	N	N		
100	3	6	0	3	0	1	0	5		N	N		
110	0	3	1	4	0	8	0	1		W	W		
120	0	3	1	4	3	14	2	14		W	W		
130	0	5	2	4	0	10	0	10		W	W		
140	2	5	0	2	17	21	18	19		N	W		
150	0	3	0	4	8	15	16	20	N	N	N		
160	0	4	0	1	0	6	2	12	N	N	N		
170	0	3	0	1	0	11	4	12		N	N		
180	0	2	0	0	0	10	6	13	DO	N			
190	0	2	0	0	0	9	9	13	DO	N			

TABLE I.—Continued

Pixel	Nov. 2, 1976 76307		Feb. 19, 1977 77050		June 6, 1977 77154		July 12, 1977 77194		AL	CL	GTL	
	AVI	KVI	AVI	KVI	AVI	KVI	AVI	KVI				
<i>Line 90</i>												
10	2	6	6	8	14	16	5	13		W	W	W
20	0	2	4	6	16	18	0	0	W	N	W	W
30	0	3	2	5	0	12	0	4		W	W	W
40	5	7	0	4	0	6	0	10	N	N	W	W
50	4	9	4	5	10	11	5	16	W	W	W	W
60	2	20	4	5	13	10	0	3	W	W	W	W
70	9	10	5	8	10	11	0	3	W	W	W	W
80	4	8	0	3	0	3	0	0	N	N	N	N
90	0	5	0	3	0	2	54	45		N	N	N
100	2	7	7	7	12	15	6	7		W	W	W
110	0	3	0	5	0	8	0	5		W	W	W
120	0	4	0	1	0	10	0	5		N	W	W
130	2	7	6	8	0	8	1	9	W	W	W	W
140	5	9	3	5	0	7	0	1		N	W	W
150	0	6	0	5	0	4	0	9		N	W	W
160	0	2	0	0	0	8	5	9	DO	N	N	N
170	0	2	0	1	0	10	4	11	DO	N	N	N
180	0	-1	0	-1	0	9	4	6	DO	N	N	N
190	0	0	0	2	0	9	0	5	DO	N	N	N
<i>Line 100</i>												
10	0	3	0	3	0	4	0	5	N	N	N	N
20	11	13	0	2	1	7	48	37		N	N	N
30	3	5	0	2	4	10	44	44	N	N	N	N
40	0	3	0	4	0	8	3	13	DO	N	W	W
50	0	3	0	1	0	7	0	9	DO	W	W	W
60	8	11	4	6	4	10	0	5		W	W	W
70	4	7	5	7	8	11	0	8	W	W	W	W
80	6	8	3	6	6	11	0	6	W	W	W	W
90	6	11	5	9	0	1	0	6		W	W	W
100	3	8	3	5	0	3	0	4	W	W	W	W
110	0	4	0	0	4	13	0	8		N	W	W
120	0	0	0	1	0	6	0	2		N	W	W
130	0	1	1	2	7	12	12	16	N	N	W	W
140	2	8	6	6	0	7	0	1	W	W	W	W
150	0	2	3	4	0	6	0	2		W	W	W
160	0	6	0	1	0	3	0	7		N	W	W
170	0	-1	0	0	0	4	18	19	DO	N	W	W
180	0	-1	0	1	0	6	5	9	DO	N	W	W
190	0	1	0	2	0	11	5	12	DO	W	W	W

ITS

TABLE I.—Concluded

Pixel	Nov. 2, 1976 7630*		Feb. 19, 1977 77050		June 6, 1977 77158		July 12, 1977 77194		AI	CI	GIL
	AVI	KVI	AVI	KVI	AVI	KVI	AVI	KVI			
Line 110											
10	16	13	8	9	11	12	0	5	W	W	W
20	0	1	0	2	0	0	0	11	N	N	N
30	0	4	0	5	0	6	2	9		N	N
40	0	3	0	2	7	12	3	12	DO	N	N
50	0	0	0	1	0	5	0	1	DO	N	N
60	2	4	0	2	0	1	44	38	N	N	N
70	2	9	0	5	0	1	0	4	N	N	N
80	0	5	0	5	0	1	0	7	N	N	N
90	0	6	2	6	0	5	1	11	W	W	W
100	0	2	2	4	0	5	0	4	W	W	W
110	0	3	0	3	0	7	0	3		W	W
120	03	8	0	5	0	3	3	13		W	N
130	0	4	0	1	0	-1	0	12		N	N
140	0	1	0	1	0	3	0	9		N	N
150	0	7	0	3	0	1	12	15	N	N	N
160	0	3	0	5	0	10	0	7	W	W	W
170	0	0	0	2	5	12	3	14		N	W
180	7	8	16	15	2	8	0	8	N	N	W
190											

Each vector is inspected automatically, and any vector having values unreasonable for agricultural data is discarded using the following procedure.

A pixel y^i is accepted as good only if

1. x_1 is less than 12 and $12x_1 - 34x_2$ is more than 108.

2. y_1 is less than 15 or more than 120.

3. y_2 is less than 8 or more than 30.

4. y_3 is less than 6.

5. y_4 is less than 10 or more than 35.

The greenness level m of the soil line then is estimated by 1 percent of the minimum greenness value y_2^i for acceptable pixels.

Then the green number g^i is computed for each pixel by

$$g^i = y_2^i - m \quad (3)$$

TABLE II.—Crop Identification Results

Segment	Julian date	AVI, percent (c)	P-I, percent (a)	Ground truth, percent (a)
1978	77050	33.3 (WW)	33.1	Unknown
	77158	26.0 (WW)	24.2	26.3 (WW)
1964	77084	43.8 (WW)	33.6 (WW)	41.02 (WW)
1971	77149	34.2 (WW)	30.0 (WW)	31.96 (WW)
	77202	7.3 (SG)	7.8 (SW)	7.17 (SG)
1988	77104	21.0 (WW)	34.9 (WW)	42.03 (WW)
	77175	36.5 (WW)	36.5 (WW)	—
	77211	—	42.8 (WW)	—
3159	77182	27.0 (SW)	26.9 (SW)	b25 (SW)
		47.0 (SG)		b50 (SG)
3186	77185	31.0 (SG)	15.7 (SW)	Unknown
8402	77187	25.0 (SW)	28.2 (SW)	Unknown
3192	77182	9.1 (SW)	9.0 (SW)	b8 (SW)

*WW = winter wheat, SW = spring wheat, and SG = spring grain.

^bApproximate results based on partial ground truth.

PROCEDURES

Procedure 1

Very briefly, P-1 is the technique used by the LACIE for crop identification and area measurement during the 1976-77 crop season. This procedure was duplicated on the General Electric IMAGE-100 (I-100) interactive computer system. It includes the selection of a number of picture elements (pixels) from a LACIE 5- by 6-nautical-mile sample segment. These pixels are randomly selected from a preset sequence of 209 pixels located at the intersection of a grid that is placed over the sample segments given in the figures. The pixels chosen are adjacent to the upper left of each intersection of the grid. From a random selection of these 209 pixels, two groups of pixels are identified and labeled. Type 1 includes a selection and labeling as spring wheat (SW), winter wheat (WW), nonwheat (N), or designated other (DO) of 40 or more pixels that are used to cluster and classify the rest of the sample segment.

The label DO in P-1 represents any area that has been removed from the scene. It could be alfalfa, barley, corn, water, roads, bare soil, etc. Each of the 40 single-pixel fields is labeled by the analyst. The statistics from these fields are then used for starting the clustering and classification procedures for the segment. Type 2 includes 40 or more additional pixels that are labeled and used to provide a bias correction for the Type 1 classification results. This procedure provided good classification results but required approximately 3.5 hours per segment. These results are shown in table II. One of the pixel identification aids is the KVI green number. This number is shown for each of the 209 pixels in table I. One use of this number is to determine whether a field has growing vegetation. This becomes highly significant during the very early stages of crop growth. For additional details on P-1, see reference 4.

AVI Change Detection

The AVI is run on segment data that are housed on the I-100 disk. It is calculated by using a tape load procedure housed in the consolidated tape read program. This program loads, in 5 minutes and 10 seconds, the AVI data to channel 3, band 4 to channel 1, band 5 to channel 2, and band 7 to channel 4 of the I-100 system. Channels 1, 2, and 4 provide a regular

color-infrared image. Three hardwired programs in the I-100 are then used. The first sets the I-100 perimeters so that all 256 gray-level values are used to develop AVI histograms. The second is a single-cell extension program that alarms all the pixels in the segment. The third is the multichannel histogram program. This task takes approximately 30 seconds. The histogram of the AVI is isolated on channel 3 and all values above zero are alarmed on the screen. The results are then assigned to one of the eight theme tracks of the I-100. That theme is measured and the percent of scene calculated for the final area measurement. This process requires approximately 7 minutes. This same process is done for each of the acquisition dates with each result assigned to a different theme in the I-100. These different themes can then be added, subtracted, or a logical AND/OR performed.

This allows the analyst to subtract the AVI mask, which represents the native vegetation, from a later AVI, which includes areas of native vegetation. The analyst can also see areas where native vegetation has been removed by using a logical AND/OR program in the I-100. Level thresholding the histogram can also separate low, high, or other densities within the AVI. The AVI levels range from 0 to 67 for this study. In theory, the upper band limit could be 128 or 2 times the value of band 7.

PILOT TEST

The LACIE sample segments used in this test were taken from operational segments used in LACIE Phase III. These segments were randomly distributed among the four U.S. Department of Agriculture (USDA) commodity analysts and processed on the I-100. The segments were worked using P-1 to provide an operational wheat estimate. They were taken from intensive test sites (ITS's) in the United States, from ITS and blind sites in Canada, and from 50 segments from Kokchetav Oblast, U.S.S.R. Blind sites are LACIE sample segments that have total ground-truth identification. They are called blind sites because the analyst does not know that ground truth is being taken over the site.

At the time of this writing, ground truth was available only for portions of the U.S. ITS's. Consequently, the results were compared to the ground truth where available and to P-1 results where ground truth was not available before the AVI was tested; however, the results were not compared until

all testing of the AVI was completed. These results, shown in table II, will be discussed later in this paper.

By happenstance, these segments provided differing conditions which helped identify various ways the AVI could be used for crop identification. Among these were a straight acceptance of the results of the algorithm, segments where the AVI results were identified as native vegetation on one date and subtracted from the AVI results of a second date, segments where growing vegetation was identified by the AVI and a histogram produced to isolate the crop in question, segments where the crops were identified and P-I used to separate one crop from the other, and instances where limited information was obtained. These results are described in a subsequent section of this paper.

Randall County, Texas, ITS 1978, has been used to show the digital values from the AVI/KVI (table I). This site has both irrigated and dryland winter wheat fields. It experienced a severe drought during the winter months of 1976-77.

Landsat imagery of the ITS for Julian dates 76290, 76307, 77032, 77050, 77158, and 77194 during the 1976-77 crop season was available. Of these, 76307, 77050, 77158, and 77194 (fig. 1) are used to compare the AVI with the KVI. The results are shown in table I. The first two numbers of the Julian dates listed above are for the year, and the last three numbers are for the day of the year.

Table I is set up so that each part of the table represents one line of the same segment. Every tenth pixel was sampled. There are 209 samples for each of

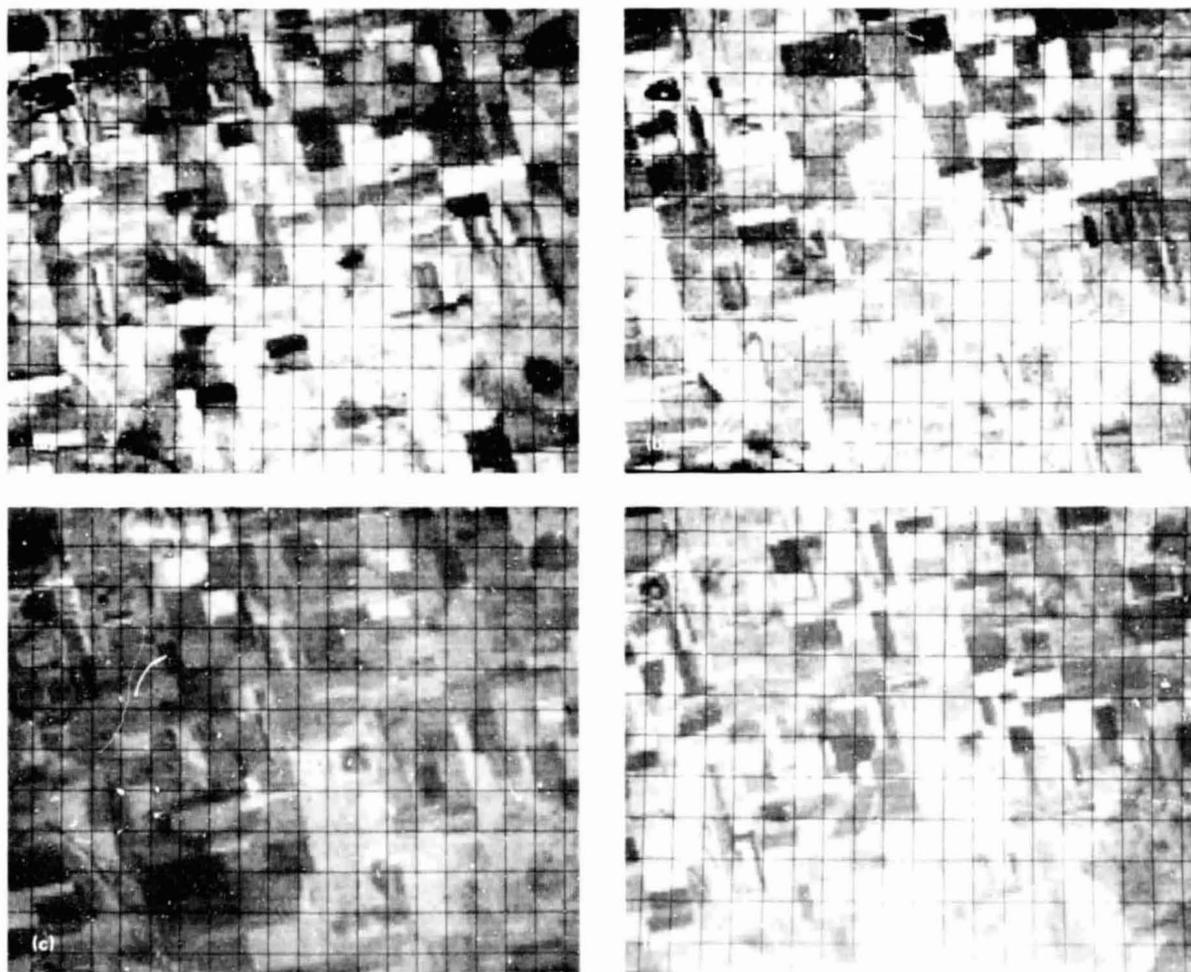


FIGURE 1.—Randall County, Texas; LACIE sample segment 1978. (a) Day 76307, November 2, 1976. (b) Day 77050, February 19, 1977. (c) Day 77158, June 6, 1977. (d) Day 77194, July 12, 1977.

REPRODUCIBILITY OF THE
ORIGINAL PAGE IS POOR

the four dates used. These are the same 209 single-pixel fields used in P-1. The last three columns are labels from P-1 and the ground truth for the ITS. The analyst labels (AL) are used as labels for starting statistics in the clustering and classification of the scene. The classification label (CL) is the result of the P-1 interactive cluster and a mixture density classification. The ground-truth label (GTL) is true only for the pixels within the brackets; the other labels in this column are interpreted from signatures in the ITS. Some of the GTL's are border pixels which, because of registration problems, may be in a field on one acquisition and out of the field on another acquisition. These pixel labels are followed by a "B."

The dryland wheat, because of dry weather and hot temperatures, was ripe and ready for harvest by June 6; however, the irrigated wheat was just getting ripe by July 12. A separation of time between dryland and irrigated wheat harvest is not unusual, but 3 weeks to a month separation is not common. The green numbers, however, reflect this separation.

Photographs of segment 1978 for the four dates used in table I are shown in figure 1. These photographs can be used to follow any of the single-pixel fields used to show what the AVI/KVI represent and are especially useful for comparing differences in the AVI/KVI. Line 40/pixel 40 is a good example of these differences. On November 2, 1976, this milo field was ripe and harvest had just begun. It is possible that this field was still partially green. Both the AVI at 4 and the KVI at 9 indicate some green vegetation. On February 19, 1977, it is obvious that the field has been harvested and only milo stubble remains. The AVI at zero indicates no green in the field, but the KVI at 5 is recording green. In this case, as well as in others in the series, the KVI is recording a number over white/yellow colors of stubble fields. The KVI drops from 9 to 3 in this series. The AVI drops from 4 to 0 and remains at 0 throughout the series.

The AVI/KVI, however, are highly consistent in the low to middle ranges where green vegetation is evident. This is shown on line 40/pixel 60. An irrigated wheat field has an AVI series of 12, 11, 13, and 2; the KVI has a 12, 10, 15, and 9 for the same pixel.

Photographs (from a 35-millimeter camera) of the I-100 television screen of the results of the AVI for segment 1978, February 19, 1977, are shown in figure 2. This acquisition date was chosen because confusion from native grasses, weeds, or other crops is at a



FIGURE 2.—Randall County, Texas; LACIE sample segment 1978 with AVI results, February 19, 1977 (orange = AVI/wheat).

minimum. The orange color is the AVI identification of wheat. The theme tracks of the I-100 are used to hold successive dates of AVI results. The P-1 results are shown in figure 3.

No ground truth was available for the February 19 date. Segment ground truth was first calculated by Accuracy Assessment for the June 6, 1977, date. Accuracy Assessment is a group within LACIE that takes the ground-truth information and assigns the ground-truth label to each of the single-pixel fields. In ITS segments, where only partial ground truth is collected, they use the imagery and available ground truth and expand these signatures to the single-pixel fields that do not have ground-truth labels.

The AVI results for the 77158 date identified both spring and summer crops. These crops were thresholded out of the scene by subtracting the AVI theme of 77050 from 77158. The results are shown in table II.

ITS 1964, Ellis County, Kansas, provides a case where only one winter date, 77084, was required to



FIGURE 3.—Randall County, Texas; LACIE sample segment 1978 with P-1 results, February 19, 1977 (pink = P-1/wheat).

obtain an estimate. There was no attempt to remove native vegetation from this date or to run additional dates. An estimate for this date was obtained in 6 minutes. Also, no effort was made to determine whether the fields identified were harvested for grain.

ITS 1971, Hill County, Montana, was used to identify a segment that contained both winter and spring wheat. The AVI for day 77149 was run and a histogram was produced. It was obvious, by viewing the results on the television screen, that some spring crops were just becoming visible. By taking out the lower five levels, the winter wheat and a small amount of native vegetation were separated from the spring crops. The results of the AVI for 77149 were subtracted from the AVI of 77184 and 77202. The final results are shown in table II.

ITS 1978, Finney County, Kansas, is a case where minimal results were obtained by using the AVI. Drought conditions during the winter and a lack of adequate segment coverage were the prime reasons for a failure to adequately identify winter wheat. By 77109, the spring rains had not come to this area, and there was no coverage between 77104 and 77175. Therefore, no good green period was available for the AVI to estimate. By 77175, the crop was already ripe and being harvested. An attempt was made to add the theme tracks from 77104 to a thresholded theme from 77175. The results of this attempt are shown in table II.

Ground truth for the Canadian blind sites had not been received at this writing, so the results of the AVI are compared to P-1 results. These segments show special uses of the AVI for obtaining an estimate.

Saskatchewan blind site 3159 provided a case where the AVI histogram provided a good spring wheat estimate. The spring grains were first identified by using an upper threshold to remove the native vegetation from the scene. A lower threshold was then used to separate the older spring wheat from the younger small grains in the segment. These threshold levels were identified by the analyst while viewing the color monitor. He systematically removed AVI levels until the native vegetation area and the younger small grains area were identified. These results are shown in table II.

Saskatchewan blind site 3186 provided a case where a multitemporal AVI was insufficient. In this case, the native vegetation was identified with the AVI on 77150. This theme was saved and subtracted



FIGURE 4.—Saskatchewan, Canada; LACIE sample segment 3186 with AVI results. Yellow = AVI of 77150; red plus yellow plus purple = AVI of 77185; red plus purple = difference of 77150 and 77185; and red = P-1 classification of wheat.

from the AVI results from 77185. No histogram separation could be found that would isolate the spring wheat. So P-1 was used over the results of the AVI to separate the spring grains. The results are shown in table II and figure 4.

Saskatchewan blind site 3192 is shown in figure 5. The yellow in this scene is the result of the AVI for 77145. This native vegetation mask was subtracted from the AVI on 77182. The results are shown in table II.

Kokchetav, U.S.S.R., sample segment 8402, is shown in figure 6. The AVI for this segment was run for acquisitions 77150 and 77187 and the results are presented in figure 7. The orange color was from 77150 and the orange plus the yellow was from 77187. All the orange was identified as native vegetation and all the yellow was identified as low-density wheat. The blue is the result of the P-1 classification for wheat.



FIGURE 5.—Saskatchewan, Canada; LACIE sample segment 3192 with AVI results from day 77145 (yellow = native vegetation (DO mask)).

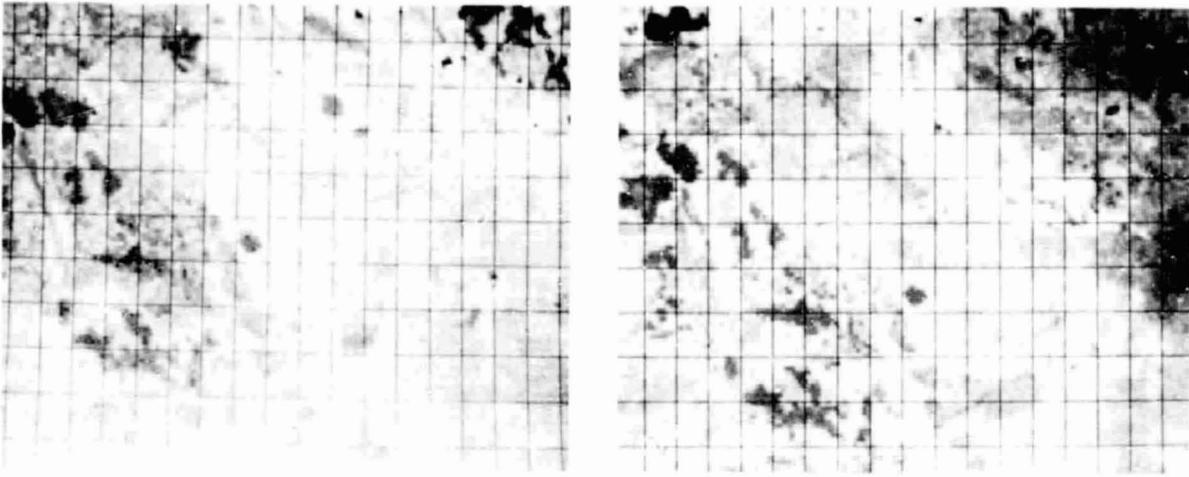


FIGURE 6.—Kokchetav, U.S.S.R.; LACIE sample segment 8402. (a) Day 77150, May 29, 1977. (b) Day 77187, July 5, 1977.

DISCUSSION

This study has demonstrated, on the basis of a limited number of ground-truth cases, that a change-detection green-number approach to crop identification and acreage measurement could be successful. This has been demonstrated over sample segments

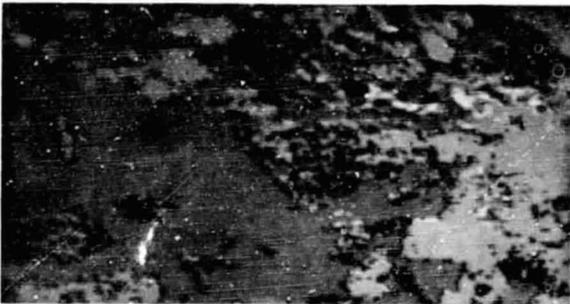
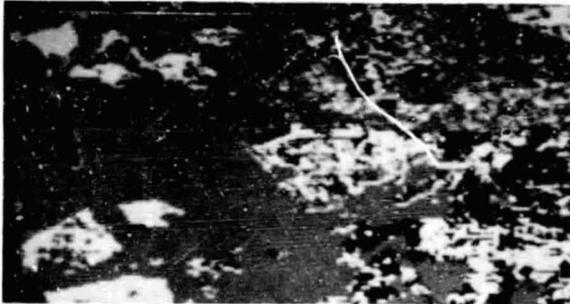


FIGURE 7.—Kokchetav, U.S.S.R.; LACIE sample segment 8402 with AVI and P-I results. Orange = AVI of 77150 = native vegetation (DO mask); orange plus yellow = AVI of 77187; yellow = difference = spring wheat; blue = P-I spring wheat.

in three countries (the United States, Canada, and the U.S.S.R.). Success has been demonstrated in the extremely fast calculation of the AVI; whereas P-I requires 210 minutes on the average, the AVI requires 6 to 12 minutes to obtain a comparable wheat estimate. Time, accuracy, and ease of use are considerations in the selection of crop identification methods.

Two significant results show that the AVI identifies growing vegetation very near the soil line; and that all positive values record growing vegetation. A project now underway will determine how much vegetation is required to first produce a positive value for the AVI and the KVI. Preliminary results of this project indicate AVI identification at the two- to three-leaf stage.

The AVI was preferable to the KVI for very early season identification of growing vegetation. Since the KVI records a positive green number over wheat straw and milo stubble, the soil line is difficult to obtain. The AVI, however, registers a positive value only when green vegetation is present. The AVI and the KVI, however, are remarkably close when growing vegetation is evident.

Since the AVI change-detection system is based on identification and measurement of growing vegetation, it is reasonable to assume that major crop types that have different growing seasons can be identified and measured. This is shown in the identification of native vegetation that was subtracted from later AVI results, and in the separation of winter and spring grains by using the AVI change-detection system.

Since there are large areas of interest where native vegetation is identifiable by the AVI before any crop is detectable, the AVI can be used to locate native vegetation areas (DO areas in P-1). The stored areas (masks) can then be subtracted from later acquisitions. Since the AVI for any acquisition only detects growing vegetation which also includes the crop or crops of interest, the subtraction of a vegetation DO mask leaves only the crop or crops of interest. These can then be identified and/or measured. All other DO areas such as roads, water bodies, cities, and bare soil are automatically removed from the scene by the AVI.

It is also reasonable to assume that information based on the AVI change-detection system can be derived from full-frame data. This is cost effective because the new interactive computers that use array or parallel pipe systems can calculate linear equations exceedingly fast. It is estimated that the PDP 11-70 with a parallel pipe system can calculate the AVI on full-frame data in less than 6 minutes. This makes feasible a total inventory of a state or country to check and/or replace sample segment aggregations.

This study has identified a procedure that could be tested in an operational system. This would provide additional information on the AVI usage, its relative strengths and weaknesses, and its cost effectiveness. Additional studies should be conducted using the AVI for stress measurement, soil moisture identification, direct yield or yield modification through stress factors, and bare soil for early-season estimates.

CONCLUSION

Based on the limited number of samples used in this study, the AVI change-detection system appears to be a promising procedure for crop identification. It was found effective in identification of crops where the crop was the only growing vegetation. It was found effective in identification of native vegetation in the spring wheat regions of the U.S.S.R. and Canada. It was found effective over areas where timely acquisitions allowed for the development of native vegetation masks which were subtracted from later AVI results to provide a good crop estimate. It was found effective when the various crops were at a growing stage that could be separated by density levels of the histogram.

The AVI was not effective when a green-phase acquisition was not received. It was marginally effective when crops in the scene could not be separated by green-number histograms. However, there was some advantage in knowing where the crops were so that other classification methods could be used. In these instances, P-1 was used and found to yield accurate results. Timeliness and accuracy are key factors in the selection of methods for data analysis. The AVI and P-1 were found to be equally accurate in this study. However, the time differential between crop identification on the I-100 Hybrid System and the AVI for equally accurate results highly favored the AVI. This suggests the consideration of AVI in a large test to determine its suitability for inclusion in an operational system. Such a combination of procedures could enhance the timeliness and cost effectiveness of analysis with no sacrifice of accuracy.

REFERENCES

1. Thompson, D. R.; and Wehmanen, O. A.: The Use of Landsat Digital Data To Detect and Monitor Vegetation Water Deficiencies. NASA Johnson Space Center, Apr. 1977.
2. Soil, Water, and Vegetation Conditions in South Texas. USDA ARS, Weslaco, Texas, June 1977.
3. Ashburn, Pat: The World Wheat Situation and a Color Composite Method for Monitoring Winter Wheat on ERTS Imagery. Cornell University, 1974.
4. IMAGE-100/Hybrid System Procedure 1 Requirements. LACIE-C00202, NASA Johnson Space Center, Jan. 1977.

Manual Landsat Data Analysis for Crop Type Identification

C. M. Hay^a

INTRODUCTION

An important component of the measurement procedures in LACIE has been the manual identification of crop type by human analysts. This paper will briefly describe the process of manual analysis for crop identification, the problems encountered in LACIE that were associated with the manual crop identification measurement procedures, and the research undertaken in cooperation with LACIE operations by the supporting research community to effect solutions to or greater understanding of the manual analysis problems.

HISTORY OF MANUAL INTERPRETATION IN LACIE

LACIE Phases I and II

Throughout LACIE Phases I and II (1975 and 1976), the analyst performed two main tasks. The first task was to outline representative areas (fields) for all spectral classes in a segment on the basis of their appearance on the Landsat image product. The spectral statistics generated from these areas were used as training for maximum likelihood classification. The second task was to label the crop type (wheat/nonwheat) in the selected training areas. This process of first selecting representative training areas and then labeling the crop type in the areas comprised what is called the "Fields Procedure." An analyst took approximately 12 hours to process a segment by the Fields Procedure and to evaluate and possibly rework the results. Half of this time was spent selecting and recording training areas; only one-eighth of the time was spent actually labeling the areas as to crop type.

^aUniversity of California at Berkeley.

LACIE Phase III

By contrast with the procedure of Phases I and II, a procedure was developed and implemented in LACIE Phase III (1977) which incorporated clustering for spectral class definition and training statistics generation. This procedure is called Procedure I. The analyst was freed from the time-consuming task of spectral class definition and could now concentrate solely on crop type labeling. A new within-segment sampling strategy involving randomly selected dots (pixels) was another innovation of Procedure I. The analyst had only to label sample dots as to crop type, thus reducing his segment processing time to approximately 3 or 4 hours. In Phase III, therefore, the analyst had only one main analysis task—crop type identification.

CROP TYPE IDENTIFICATION—THE ANALYSIS PROCESS

In simple terms, the interpretation process (also called labeling) consists of two main components: (1) feature detection and physical characteristics determination, and (2) feature evaluation, including identification and condition assessment. While these processes may occur simultaneously and iteratively, they can be treated separately to facilitate understanding. *Feature detection* is the action of discriminating a unique landscape feature (a field in the LACIE case) on the basis of spectral, spatial, and temporal characteristics observable in Landsat multitemporal-spectral data. *Feature evaluation* is the process of assessing available data by analytical means and then synthesizing the pertinent data to conclude the feature's identity and condition. Feature identification is the action of assigning a name (e.g., wheat, nonwheat) to the detected feature. Correct feature identification cannot properly proceed unless feature detection has first occurred.

Feature detection, however, does not ensure feature identification.

Feature Detection and Characteristics Determination

In agricultural environments, the *features* that an analyst wishes to detect are cropped fields. The *feature characteristics* that an analyst must determine are (1) the size and shape, the type of boundary elements, and the spatial relationships of similarly and dissimilarly appearing features (spatial characteristics); (2) the development patterns throughout the growing season for the field (temporal characteristics); and (3) the spectral response in specific time periods corresponding to given crop type biostages (spectral characteristics). Of these three characteristics, the second is the most significant to the analyst for the detection and identification of any crop type. The other two characteristics are necessary when significant temporal overlap exists between one crop type and some confusion crops, when key acquisitions are missing or of poor quality, or when ambiguity exists in the data. Obviously, the probability of correctly identifying a crop in a given field will be low if a spectral response indicating vegetation canopy is never detected during the growing season or is not detected at particularly significant vegetation biophases specific to given crop types.

Feature Characteristics Evaluation for Crop Identification

While site-specific Landsat data allow an analyst to detect a feature and determine its temporal, spectral, and spatial characteristics, ancillary data and a priori knowledge from outside the Landsat data are necessary for the analyst to identify and label a detected feature; that is, nowhere will one find the words "wheat field" written across a field as observed on Landsat data. A priori knowledge and ancillary data supply information about what crops are grown in a region, the rate and timing of crop-specific canopy development, cropping and cultivation practices employed in a region or specific to a given crop type, the characteristic appearances of given features on Landsat data, etc.

A priori knowledge is gained from training and past experience. Ancillary data consist of data that

are additional to the site- and date-specific Landsat spectral data. Ancillary data necessary for crop type identification are (1) crop calendar information, including average-normal and year-specific data; (2) historical crop proportions for several recent years; (3) regional cropping practice information, such as crop rotation sequence, cultivation practices, and irrigation practices; and (4) occurrence of meteorological events affecting crop development and crop spectral response. For a more complete description of the manual analysis process, see this author's symposium paper entitled "Manual Interpretation of Landsat Data."

PROBLEMS ENCOUNTERED IN LACIE WITH MANUAL CROP IDENTIFICATION

In Phases I and II of LACIE, it was found that, in some regions, the analysts' interpretation error was beyond the tolerance limits. Several problem areas associated with each of the two main interpretation components—feature detection and feature evaluation—were identified. Solutions to these problems were addressed in LACIE through cooperation between the research community and LACIE operations personnel. The problems identified as operative in LACIE can be grouped into three main areas related to the manual analysis process and measurement procedures. These are (1) problems associated with feature detection and characteristics determination (the first component of the manual analysis process), (2) problems associated with feature evaluation (the second component of the manual analysis process), and (3) problems associated with labeling procedures (measurement mechanics).

Problems Associated With Feature Detection and Characteristics Determination

Image product deficiencies.—As stated earlier, Landsat data are used for feature detection and physical characteristics determination. Thus, the ability of Landsat data products to represent spatial and spectral data clearly and accurately to the analyst is of great concern. During LACIE Phases I and II, the primary Landsat data product available to analysts was the color-infrared (CIR) image Product 1. This image product is a color composite of the data from three of the Landsat spectral bands. The three spectral bands selected to produce this color com-

posite are the green band (0.5 to 0.6 micrometer, multispectral scanner (MSS) band 4) assigned to a blue color gun, the red band (0.6 to 0.7 micrometer, MSS band 5) assigned to a green color gun, and an infrared band (0.8 to 1.1 micrometers, MSS band 7) assigned to a red color gun. The resultant color-composite image was designed to simulate the type of image secured from conventional color-infrared photographic imagery because analysts were most familiar with that type of image product. The remaining Landsat infrared band (0.7 to 0.8 micrometer, MSS band 6) is normally excluded from the composite because of the three-color limitation of light additive systems.

The Product 1 image is an effective format for the extraction of spatial information, such as feature size, shape, relationship to neighboring features, and distribution throughout the area. However, Product 1 can provide only gross, relative spectral information about a feature. While gross, relative spectral information is often sufficient for crop type identification where multitemporal analysis procedures are used, numerous situations were encountered in LACIE Phases I and II where Product 1 did not sufficiently represent the Landsat spectral data to allow correct crop type labeling.

Frequently, in situations where there was a sparse canopy, as early in the growing season, or where there was abundant vegetative cover in general, as in humid regions, or where close confusion crops were present with the crop of interest, Product 1 either did not represent the vegetated field in the manner normally expected (some shade of red or pink) or did not show subtle spectral differences between features that actually were present. These problems caused the analyst to (1) "misinterpret" vegetated fields as nonvegetated fields on the early-season acquisitions and (2) fail to detect spectral characteristic differences between close confusion crops.

Temporal sampling rate deficiencies.—Another crop identification problem related to feature characteristics determination is insufficient temporal sampling. As was stated above, the temporal-spectral pattern throughout the growing season is the most significant feature characteristic for crop identification. If this pattern is not adequately determined, there is a greater probability of confusion among crop types. Two causes of insufficient temporal sampling which lead to inadequate temporal-spectral pattern determination are (1) missing Landsat acquisitions due to cloud cover or other reasons and (2) periodicity of Landsat overpasses. Temporal-spectral pattern

changes that occur with a frequency of less than 18 days are unlikely to be consistently observed since Landsat passes over a particular site every 18 days. A problem created by this periodicity is confounded even more significantly when acquisitions are lost because of cloud cover or other cases of non-response.

Spatial resolution deficiencies.—Features below the resolution limit of the Landsat sensors (approximately 1 acre) cannot be detected. Thus, correct crop identification with Landsat-1 and Landsat-2 for fields smaller than 1 acre is impossible and for fields of up to approximately 10 acres is improbable. The improbability of correctly identifying 5- to 10-acre fields is a function of misregistration between acquisitions and boundary (mixed) pixel problems. It is necessary to determine fairly accurately the spectral changes of a field over time. If data points representing a given ground location cannot be overlaid from one acquisition to another with a fair degree of precision, an accurate temporal-spectral pattern, and thus crop type, cannot be determined.

Of the feature detection and characteristics determination problems just discussed, the most significant problem was deficient Landsat data products. A further discussion of the factors involved and some solutions implemented in Phase III relative to the deficient data products will be presented later in this paper.

Problems Associated With Feature Evaluation

Most of the remaining sources of error associated with manual crop type labeling in LACIE were a function of insufficient a priori knowledge or ancillary data or of nonoptimum labeling procedures.

Insufficient a priori knowledge and ancillary data.—One deficiency in a priori knowledge which had an effect on analysts' labeling accuracy, particularly in the early phases of LACIE, was the lack of adequate information concerning the *variability* in the temporal-spectral patterns of wheat, small grains, and other crop types. A related deficiency was the lack of adequate crop type temporal-spectral separability information. No specific information about the temporal-spectral patterns of crop types other than wheat was available to the analysts. These deficiencies resulted in omission errors for wheat and small grains. Incorrectly, analysts assumed less variability in wheat temporal-spectral patterns than

was actually present. Thus, analysts' labels were conservative with respect to wheat. As the analysts' experience in LACIE increased, they gained a better appreciation for the true wheat temporal-spectral pattern variability. However, additional variability information was definitely needed in abnormal situations, such as the occurrence of drought or winterkill or other episodal events. Similarly, without specific information about the temporal-spectral patterns of crops other than wheat, analysts could not "doublecheck" their identifications by working the problem in reverse. That is, in addition to considering the question, "Is this pixel wheat?" the analyst could have asked, (1) "What crop type is this pixel?" and (2) "What crop types are definitely not represented by this pixel?" Elimination of candidate crop types from consideration in the analysis often forces the analyst to go back, reconsider his initial analysis, and change his initial answer. However, since the analyst did not have the necessary data and temporal-spectral variability information for crop types other than wheat, he could not doublecheck his initial answer. The result was that some wheat was mislabeled or omitted.

Another problem that resulted in inconsistent labels among different analysts (analyst variability) was related to the differing amounts and types of a priori knowledge that individual analysts possessed. Each analyst had his own unique background and set of interpretation experiences upon which to draw. This meant that the quality and quantity of a priori knowledge was highly variable among the analysts. Before the start of operational interpretation in LACIE, the analysts had all undergone an extensive 2-week training course which was intended to help standardize their background and experience. After the start of LACIE operational interpretation, however, it was found that the variability among analysts was still greater than desired. Steps taken to help remedy this situation are addressed later in this paper.

Nonoptimum labeling procedure.—A large number of labeling errors traced to the analyst consisted of labels affixed to misregistered and boundary (mixed) pixels. Misregistered pixels are those that jump back and forth between one field and an adjacent one on successive acquisitions. Boundary pixels are mixtures of the signatures from more than one field. In LACIE, the analyst had to affix a definite crop type label (wheat or nonwheat) to a pixel, including the boundary and misregistered pixels. To do this, he specified a reference acquisition on which he labeled

the pixel. He "guaranteed" the pixel label for that reference acquisition only and not for any other acquisitions. This led to analyst-credited "mislabeled" when the pixel label was not appropriate for the majority of the segment's machine-processed acquisitions that were subsequently checked in accuracy assessment.

This problem was not significantly addressed while the LACIE experiment was being run. However, current opinion is that the problem can probably be lessened or completely eliminated by screening misregistered and boundary pixels and then labeling them in a different manner, since the analyst has no problem recognizing and describing these pixels as misregistered or boundary pixels. The problem is due to the lack in the current procedure of an adequate labeling option to affix to these pixels.

MAJOR RESEARCH EFFORTS TO IMPROVE ANALYST LABELS

Landsat Image Products Improved and Expanded

One of the deficiencies of Product 1 was due to the mapping function used to transform the Landsat digital data to image format. Each Landsat spectral band was scaled and biased separately to enhance overall image contrast. This was desirable for optimum extraction of the spatial information. However, this data mapping procedure altered the relationships between spectral bands such that the spectral information was very definitely distorted. Thus, fields with sparse vegetative canopy often failed to display the expected "red" tones.

An auxiliary image product, called Product 3 or the Kraus Product, was developed to restore proper spectral band proportions. On Product 3 (fig. 1), sparse canopy is represented in pale or dull red colors. This is more in line with analyst expectations of characteristic CIR image appearance for this vegetative condition. Product 3, however, exhibits a loss of image contrast and brightness, which causes analyst fatigue when interpreted for sustained periods. Thus, Product 3 was used only as an auxiliary to Product 1. A description of the manner in which Product 1 and Product 3 were produced is contained in the paper by Juday entitled "Colorimetric Consideration of Transparencies for a Typical LACIE Scene."

Another factor contributing to Product 1 limitations was that equal differences in digital spectral

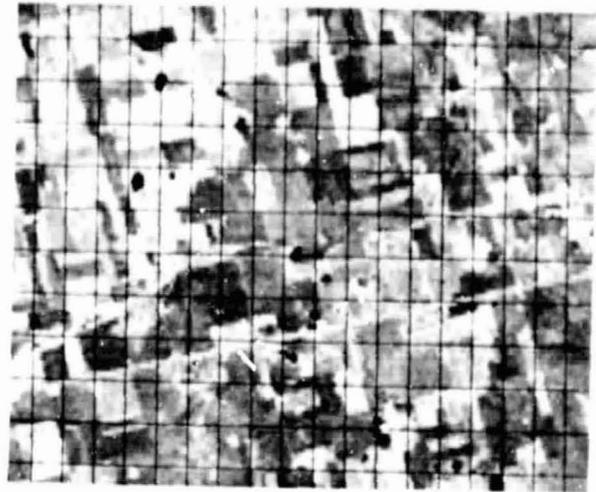
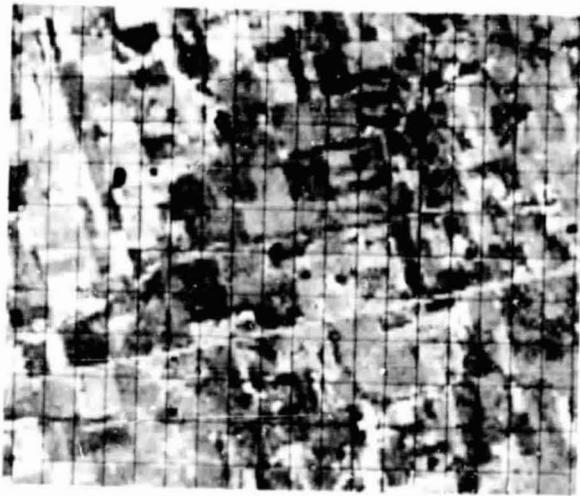


FIGURE 1.—Example of standard image Product 1 and auxiliary Product 3 for segment 1640, May 20, 1977. (a) Product 1. (b) Product 3.

responses were not seen as equally contrasting color differences. Thus, the analyst could not see some of the significant spectral differences between features on Product 1. Figure 2 clearly demonstrates this problem. In the quantization of Landsat data by the production film converter (PFC), 16 Landsat values are grouped as 1 PFC value. Landsat MSS band 7 (infrared) data are plotted against MSS band 5 (red) data and the data points in the resultant scatter plot have been assigned the color that those pixels would possess in the normal Product 1. As can be seen in figure 2, many pixels spanning significant regions in the Landsat spectral data space are assigned the exact same image color, and thus spectral differences are not detected. A more thorough discussion of this very interesting image display problem is found in the paper by Juday.

Numeric and Graphic Data Products

To offset the limitations of image products in adequately portraying the Landsat spectral data, several numeric and graphic data products were developed for LACIE analysts. These products were made available to analysts after Procedure 1 implementation in Phase III. The first step in constructing any of these products was to apply the Kauth-Thomas Tasseled Cap Transformation to the Landsat spectral data (see the paper by Kauth et al. entitled "Feature Extraction Applied to Agricultural Crops as Seen by Landsat"). This affine transformation allows a more

orthogonal view of the "hyperplane" in which Landsat spectral data fall than is possible within Landsat coordinated space. The Tasseled Cap Transformation defines a new set of coordinate axes referred to as TCH-1 (brightness), TCH-2 (greenness), TCH-3 (yellow stuff), and TCH-4 (non-such). Results of work at the Environmental Research Institute of Michigan (ERIM) indicate that TCH-1 corresponds to soil and scene brightness, TCH-2 corresponds to "green stuff" such as green vegetation within a scene, and TCH-3 appears to correspond to "yellow or dry vegetation" within a scene. TCH-4, which contains little data and has no specific correlation to ground conditions, is called "non-such" and is not presently used.

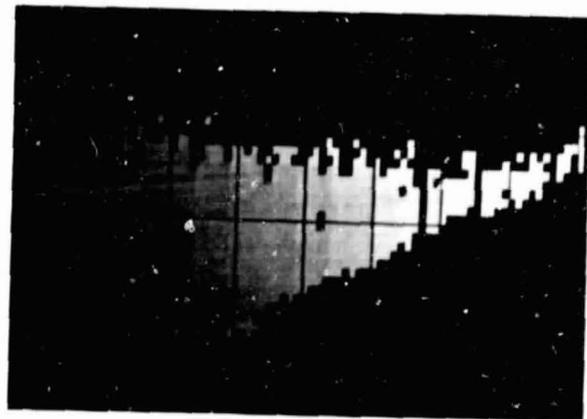


FIGURE 2.—Spectral scatter plot.

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

The transformed Landsat data were presented to the analyst in several different numeric and graphic formats. One such product was a scatter plot that allowed the analyst to compare the greenness (TCH-2 vertical axis) of a pixel in relation to its brightness (TCH-1 horizontal axis). Figure 3 is a scatter plot of TCH-2 versus TCH-1 for an acquisition of a segment in North Dakota where barley is turning and wheat is still green. Scatter plots were produced for each acquisition processed by the automatic classifier, and the data were sampled by means of a 10-point by 10-line grid placed over the registered Landsat data. This produced a sample of 209 pixels from the scene. From this set of 209 pixels were drawn the analyst starting and labeling (type 1) dots and the bias correction (type 2) dots. The subset of labeled sample pixels was displayed with the corresponding analyst labels. This allowed the analyst to check quickly the consistency of his dot labels (see the paper by Heydorn et al. entitled "Classification and Mensuration of LACIE Segments" for a discussion of within-segment sampling procedures).

Another "spectral aid" developed for Phase III LACIE was the trajectory plot (fig. 4). Again, TCH-2 (greenness) versus TCH-1 (brightness) plots were used. However, each trajectory plot was the multitemporal history of just 1 of the 209 dots; that is, the TCH-2 versus TCH-1 values for the given pixel for all multitemporal acquisitions (currently limited to four) were presented in one trajectory plot. The points on the plot were labeled in proper temporal sequence, and the analyst evaluated the dynamic change in spectral response through time for the pixel. The temporal change in spectral response of a crop type is a very significant identifying characteristic. The analyst was able to compare

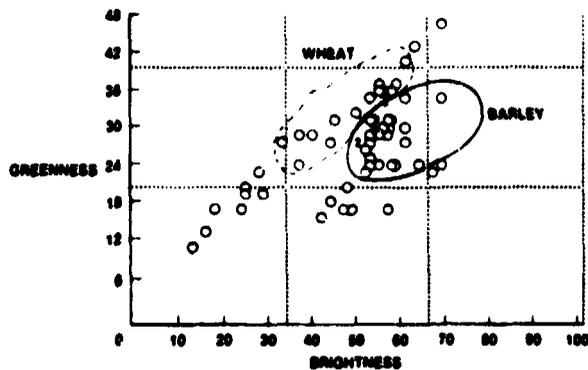


FIGURE 3.—Example of analyst spectral aid scatter plot.

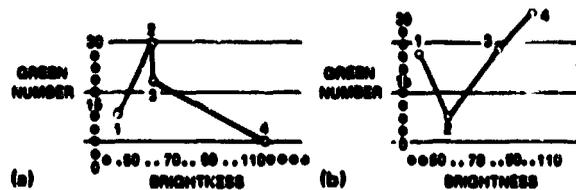


FIGURE 4.—Example of trajectory plots. (a) Spring wheat. (b) Nonwheat.

the trajectory plots of sample pixels to the trajectory plots of various crop types contained in keys.

The third and final numeric product (fig. 5) made available to the analysts in Phase III was a simple listing of the values of green number and brightness for each of the 209 sample pixels for each acquisition processed by the classifier.

The numeric and graphic data products described were available only after the segment had been machine processed. The analyst, therefore, used these products to check the consistency of his dot labels after machine processing of the segment. For example, the analyst could change bias correction dot labels (type 2 dots) after processing; such a change could have a beneficial effect on the segment wheat estimate. Thus, the spectral aids affected the quality and consistency of bias correction dot labels more directly than the starting and cluster labeling dots (type 1 dots). A fuller discussion of these products is presented in the paper by Abotteen entitled "Image and Numerical Display Aids for Manual Interpretation."

In LACIE Phase III, NASA and Lockheed Electronics Company did a study in which the spectral aids, particularly the scatter plots, were used to separate spring wheat from the other spring small grains in North Dakota, most of which were barley. As has been established in crop separability analysis, barley matures and turns golden sooner than spring wheat. On optimally timed acquisitions around the wheat soft-dough and barley turning biostage, barley will appear less "green" (lower TCH-2 green number)

DOT NO	GRID INTER SECTIONS	ANALYST LABEL	CLASSIFIED	ACQUIS 1		ACQUIS 2		ACQUIS 3		ACQUIS 4	
				G	B	G	B	G	B	G	B
7	10 20	S	S	13	54	32	66	24	66	12	84
61	40 60	N	N	36	88	30	77	31	73	11	88
90	70 100	N	N	11	43	21	64	27	62	13	81
122	70 90	S	S	12	43	30	62	30	67	6	110

FIGURE 5.—Example list of values for green number and brightness of individual pixels by acquisition classification.

and more bright (higher TCH-1 brightness value) than spring wheat. In a scatter plot for this optimally timed acquisition, barley will fall below and to the right of spring wheat (fig. 3). Thus, the spectral aids can greatly enhance analysis of subtle spectral differences. Note, however, the emphasis placed on optimally timed acquisitions.

A Priori Knowledge Expanded and Standardized

Interpretation keys.—After the start of LACIE operations, it was apparent that some additional analyst training and standardization was still needed. In response to this need, selective interpretation keys were compiled and made available to the analyst (see the paper by Baron et al. entitled "Analyst Interpretation Keys"). The keys were intended to help incorporate the analysts' experiences and standardize a priori knowledge concerning wheat and Landsat data analysis. It was hoped that the keys would help minimize variability in crop identification (labeling). Furthermore, the keys aided and hastened the training of new analysts who joined the project after the initial pool of analysts was selected and trained.

The interpretation keys were compiled in two volumes. Volume I contained introductory material with information concerning the general analysis of Landsat imagery and ancillary data for the identification of wheat. Volume II consisted of Landsat imagery showing examples of wheat development within specific geographic regions. In this way, regional eccentricities or problems that affected the wheat temporal-spectral response pattern could be efficiently presented to the analyst.

Crop separability studies.—When analysts began processing segments from spring wheat areas, it was quickly determined that the labeling of spring wheat versus other spring small grains was very unreliable. The analysts had no guidelines to help them separate the spring wheat from all other spring small grains. This problem of small-grains separability was most significant in the spring wheat segments because significant proportions of other spring small grains were grown along with the wheat. This was not the usual case in the winter wheat situation. Since LACIE was originally intended to be a wheat inventory system as opposed to a small-grains inventory system, research was undertaken to determine whether there were consistent temporal or spectral characteristics that could be identified and measured

from Landsat to enable the analyst to make the needed distinctions.

This work was carried out primarily at ERIM, the Laboratory for Applications of Remote Sensing (LARS), and the NASA Johnson Space Center (JSC). The results of this research were not available until near the end of the LACIE experiment and thus did not affect LACIE operational procedures. However, JSC in-house tests were conducted into procedural analysis changes that required the analyst to distinguish spring wheat from all other small grains present. The direct wheat evaluations were encouraging and discouraging at the same time. It was found that spring wheat and spring barley did differ from each other in their temporal characteristics. Barley fairly consistently matured faster and sooner than wheat. Thus, barley would start to turn before wheat did. The discouraging aspect, however, was that a Landsat acquisition was needed at this critical barley-turning/wheat soft-dough stage. It was found that acquisitions at this critical time were often missing or improperly timed. So, while spring wheat and barley could theoretically be separated from each other, they could not be separated consistently from one segment to the next because of the need for the often-missing critically timed barley-turning acquisition.

Procedural Modifications

The most significant procedural modification to be developed in LACIE was Procedure 1, which relieved the analyst of the responsibility for spectral class definition. This was accomplished by the use of clustering. Also in Procedure 1, the analyst labeled individual, randomly selected pixels from a systematic grid instead of labeling fields that he had previously delineated. Procedure 1 reduced analyst-segment interaction time by 65 percent, thus increasing throughput and decreasing turnaround time. A more complete description and discussion of Procedure 1 can be found in the paper by Heydorn et al.

One innovation incorporated in Procedure 1 was a change in the use of the classifier output. In the old Fields Procedure, the wheat acreage from the classifier was treated in the traditional remote-sensing manner as the final estimate for the segment. However, Procedure 1 recognized that there was bias in the classification and therefore used the classifier output as the stratification to be used in a stratified sampling scheme. Analyst labels for bias correction

dots (type 2 dots) were now used in connection with the stratification produced by the classifier to produce the final estimate for the segment.

Currently in Procedure 1, all pixels are clustered and assigned labels according to a nearest neighbor rule on the basis of a limited number of analyst labels (type 1 dots). All pixels are then processed through a maximum likelihood classifier for assignment to a wheat or nonwheat stratum. One supporting research study has been undertaken to evaluate alternative methods of producing the stratification used in Procedure 1. An alternative procedure developed by the University of California at Berkeley (UCB) for producing the crop type stratification is called the Delta Function Stratification Procedure.¹ This procedure utilizes an indicator of the temporal pattern of the vegetation; this indicator is produced by ratioing Landsat MSS band 7 (infrared) with Landsat MSS band 5 (red) to assign clusters to a crop type stratum. The strata produced from this procedure (usually five or six) are then "bias corrected" according to standard Procedure 1 methods. Advantages of the alternative stratification procedure are (1) a potential decrease in analyst segment-handling time due to a decrease in the number of pixels requiring labels (2) more accurate labeling of clusters, (3) a decrease in the amount of computer processing time required per segment achieved by eliminating the maximum likelihood processing step, and (4) the capability to extend the procedure to crops other than wheat. Tests of this alternative stratification procedure indicate that it produces results that are comparable to and not statistically significantly different from current Procedure 1 results. The alternative stratification procedure is undergoing further evaluation and has not yet been evaluated in a test on the scale of LACIE operations.

Automatic Crop Labeling—The Future

The foregoing discussion has addressed the manual labeling of crop types from Landsat data. The potential of automating these crop-labeling pro-

cedures will now be briefly addressed. There is a study in progress within the supporting research community which has as its objective the development of an automated or computer-aided labeling procedure. The motivation for such a study is to further decrease variability in dot labels due to differences between individual analysts. In addition, it is desirable to have an estimate of the reliability of each dot label; i.e., the probability that a given label is correct.

The automatic labeling procedure being developed and tested is called LIST (Label Identification from Statistical Tabulation) (see the paper by Pore and Abotteen entitled "A Programmed Labeling Approach to Image Interpretation"). Three types of information are input to the procedure: (1) spatial information provided by manual analysis of Landsat data; (2) Landsat spectral information which is automatically sampled; and (3) ancillary information compiled from meteorological data and other ancillary data types described earlier. The questions an analyst must answer for input to LIST and the automated questions are presented in table I. Presently, the procedure is "trained" on an area for which ground data are available. Relative weights for the input variables (answers to input questions) are determined by statistical analysis of these ground-observed areas. The "trained" LIST procedure is then applied to areas without benefit of ground data. Initial test results (table II) are comparable with results from analyst-labeled dots. Boundary and misregistered pixels were screened from the test so the test results are for "pure" pixels only. As the decision logic for specific crop identification becomes better defined, it too may be automated. Automated or partially automated crop-labeling procedures can enhance operational crop inventory systems in that manual analysis inputs can be minimized. The analyst can be freed from repetitive analysis tasks, and procedures can be more nearly standardized to reduce measurement procedure variability.

SUMMARY

In the simplest terms, manual identification of crop type consists of two components. The first is feature detection and physical characteristics determination. A feature of interest in LACIE is an agricultural field. The second component is feature identification or labeling. The data utilized for feature detection are Landsat data. The data neces-

¹C. M. Hay et al., "Development of Techniques for Producing Static Strata Maps and Development of Photointerpretation Methods Based on Multitemporal Landsat Data," Annual Report, NASA Contract #AS9-14565 (R. N. Colwell, Principal Investigator), Space Sciences Laboratory, Series 19, Issue 1, University of California at Berkeley, Dec. 1977.

TABLE I.—LIST Questions

Question	Response
<i>Analyst-interpreter questions</i>	
1. Is pixel clearly in non-agricultural area?	(1) Yes. Stop. (Blank) Agricultural area or indeterminate
2. Is pixel registered with regard to other dates (i.e., in the same category on all four dates)?	(1) No. Stop. (Blank) Yes or indeterminate
3. Is pixel a mixed pixel (part of more than one field or boundary)?	(1) Yes. Stop. (Blank) No or indeterminate
4. Is this an anomalous pixel (not representative of most of the other pixels within the field)?	(1) Yes. Stop. (Blank) No or indeterminate
5. PFC vegetation canopy indication is _____. (Use all available imagery film types.)	(0) No vegetation canopy (1) Low-density green vegetation canopy (2) Medium-density green vegetation canopy (3) High-density green vegetation canopy (4) Senescing (turning) vegetation canopy (5) Harvested canopy (stubble)
<i>Automated questions</i>	
1. Robertson biostages for winter and spring wheat, respectively	_____
2. Green number of pixel (corrected to 60° incidence)	_____
3. Is green number in the small-grains range?	_____
4. Brightness number of pixel	_____
5. Winter and spring principal component greenness (PCG) statistics, respectively	_____
<i>Automated analyst-interpreter keys</i>	
1. Is the vegetation indication of the pixel (using all available product types) valid for the Robertson biostage of wheat for the acquisition?	_____
2. Does the pixel follow a small-grains vegetation canopy development pattern?	_____

sary for feature identification are a priori knowledge and ancillary data. That is, only feature existence and information about the physical characteristics of the feature can be extracted from Landsat. The information that allows the correlation of feature characteristics to a specific type of feature (e.g., a wheat field) comes from a priori knowledge and ancillary data. Landsat data contain significant information which, in conjunction with ancillary data, can allow quite sophisticated analyses to be performed. The format in which Landsat data are presented, however, does significantly affect the usefulness of the information. Two types of information are extracted from the Landsat data—spatial information and temporal-spectral information. Originally, the sole format for Landsat data in LACIE was an image called Product 1. While this format was optimal for the extraction of spatial information, it was not optimal for the extraction of precise spectral information. Indeed, Product 1 distorted the spectral data and led to labeling problems for analysts. In answer to the Landsat spectral data format problem, the research community, working closely with LACIE operations personnel, developed numeric and graphic formats for analyst “spectral aids” which were more optimal for the extraction of spectral information from Landsat data.

Crop type identification is possible because of relatively unique temporal-spectral patterns associated with timing differences of certain phenological

TABLE II.—LIST Test Results^a

Labeling procedure	Omission error, percent	Commission error, percent
<i>Winter small-grains sites</i>		
Analyst	18	13
LIST	17	15
<i>Spring small-grains sites</i>		
Analyst	50	29
LIST	53	39

^aFour training and four test segments for each site.

growth stages of different crop types. Adequate sampling through time of the spectral response of given fields is necessary to identify crop types reliably. Thus, missing Landsat acquisitions because of cloud cover have frequently had a significant impact on crop-labeling accuracies in LACIE. Also, optimal timing of the acquisitions is critical to the separation of closely related crops such as wheat and barley. Temporal differences between closely related crops are subtle and are observable only within limited time periods. Since only one Landsat was used, tem-

poral sampling was limited to 18-day intervals; this periodicity did not allow for consistent, reliable separations between closely related wheat and barley.

Automated and partially automated crop-labeling procedures were developed, and initial testing demonstrates a significant potential for such procedures. Automated procedures offer decreased variability in crop type labels, reduced manual analysis requirements in operational crop inventories, and increased measurement reliability information about crop type labels.

Original photography may be purchased from
EROS Data Center

Sioux Falls, SD

57198

LACIE Analyst Interpretation Keys

J. G. Baron,^a R. W. Payne,^a and W. F. Palmer^a

INTRODUCTION

The Analyst Interpretation (AI) Keys were prepared within the Large Area Crop Inventory Experiment (LACIE) for incorporation into the Classification and Mensuration Subsystem (CAMS) Detailed Analysis Procedures. They were developed and tested during Phase II of LACIE (1976) and implemented during Phase III (1977).

The Classification and Mensuration Subsystem of LACIE was responsible for using Landsat data to determine the proportion of wheat in each sample segment required by the Crop Assessment Subsystem (CAS). Analysts within CAMS used Landsat color-composite images, crop calendars, ancillary data such as historical statistics, and computer-generated spectral plots and cluster maps to identify a subset of the picture elements (pixels). This identification formed the basis for a machine classification of the entire segment and calculation of the proportion of wheat. Identification of the pixels was accomplished by comparing their characteristics with known signatures (physical/cultural features or patterns of features which allow a particular crop type to be recognized on imagery). These signatures of wheat and other crops are described and documented in the Analyst Interpretation Keys.

Objectives of the Keys

The major objectives of the AI Keys were to improve accuracy and efficiency by minimizing variance in crop identification, to disseminate analyst experience gained in LACIE, to accelerate the training of new analysts or those new to specific geographic areas, to maintain an interpretation information base, and to provide a documentation format that would enable easy updating. Another important

objective was to provide the analyst with a better understanding of the expected ranges in color variation of signatures for individual biostages and of temporal sequences of Landsat signatures. Since signatures within these images are affected by image processing and environmental conditions as well as changing crop conditions, absolute color matching is not usually a reliable means of identification. However, relative similarities or differences of signatures and temporal sequences are useful and this usefulness is being increased by such new technology as haze correction.

Since crop discrimination in LACIE is sometimes dependent on the somewhat subjective interpretation of data ancillary to actual multigate satellite imagery, the construction of detailed decision logic has been an elusive but important objective. General interpretation decision logic (fig. 1) without documented signature variability and temporal signature sequences for small grains has been used and improved throughout LACIE. In this context, the AI Keys initially expanded and illustrated the logic related to the use of LACIE Product 1 (the primary imagery product—color-composite imagery of bands 1, 2, and 4) and treated ancillary data (other than crop calendars) as supplementary information for decisions.

Background

The first key used in LACIE was the "Wheat Identification Aid for Image Interpreters" (ref. 1). It was developed in June 1974 using the limited amount of data (Landsat and ground observations) available at that time. These data were from two intensive test sites (ITS's): Hill County, Montana, and Swift Current, Saskatchewan. An appendix containing five additional sites was added in 1975. This document proved to be very beneficial in the training of analysts new to the CAMS environment.

^aLockheed Electronics Company, Houston, Texas.

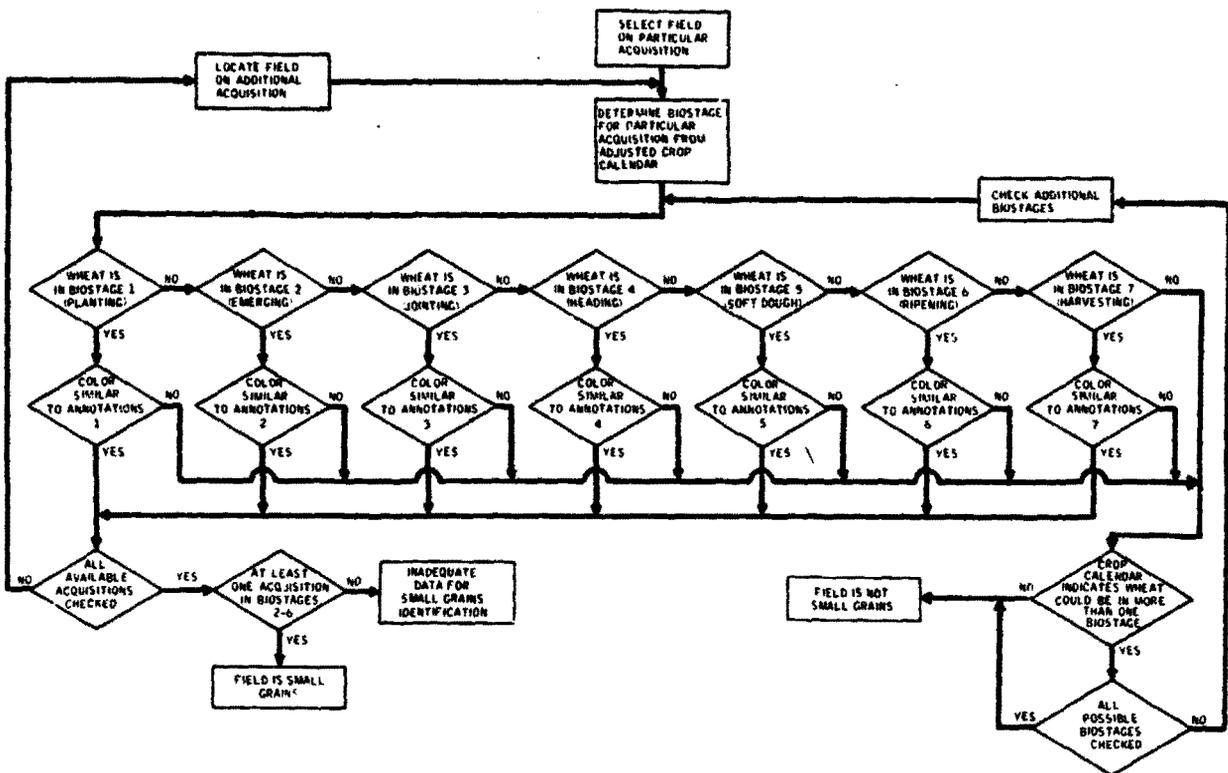


FIGURE 1.—Decision logic for small-grains identification.

In August 1975, the Acreage Estimation Technical Review Team¹ suggested that additional keys be developed to improve the accuracy and consistency of interpretation. This recommendation was subsequently endorsed by LACIE project management and the LACIE Keys Design and Planning Working Group² was organized in January 1976.

It was apparent to the LACIE Keys Design and Planning Working Group that the interpretation methodology that had evolved within CAMS and the combination of imagery, ancillary data, and ground observations already found beneficial in LACIE would be the basis of a timely, specific key. A review

¹The Review Item Disposition (RID) was initiated for the team by William Anderson of the Earth Resources Observation Systems (EROS) Data Center, Sioux Falls, South Dakota.

²The LACIE Keys Design and Planning Working Group consisted of personnel from Lockheed Electronics Company (LEC) and Lockheed Missiles and Space Company (LMSC) and two independent consultants, Robin Welch (Texas A & M University) and Joseph Clifton (U.S. Department of Agriculture, retired).

of interpretation keys in general use had provided little insight into a design that would be applicable in LACIE.

The first step the LACIE Keys Design and Planning Working Group took was to examine factors that would affect the identification of design requirements. These factors and their implications for the design were stated as follows.

Experimental factors

1. The Data Acquisition System does not produce unambiguous signatures.
2. Correct identifications have required ancillary data for evaluation of spatial and temporal signature variability.

Design implications

1. Training analysts on nominal signatures alone will not assure accuracy or consistency.
2. Correlation of temporal image clues, crop calendars, and other data is required for acceptable accuracy and consistency.

The design requirements resulting from the analysis were then identified as the following.

1. Both nominal signatures and signature variability must be recognized and understood.

2. Explicit logic is needed for best use of crop calendars to identify nominal signatures.

3. Ancillary data should be designed to cue analysts to sources of possible signature variability.

4. Reference imagery should be provided for segments and sites that are most representative of specific crops and that show relatively homogeneous signatures.

5. Reference imagery on all biostages, indexed to crop calendar, should be provided for reference sites.

6. Ground truth, where available, should be used for verification of the keys.

7. The selective key approach should be used.

With respect to item 7, interpretation keys can be categorized into two basic types: elimination keys and selective keys. An example of the elimination key is the dichotomous key in which, at each step in the analysis process, the various categories of objects are divided into two groups, on the basis of some characteristic that should be visible on the imagery. Each succeeding step subdivides the remaining group until the object of interest is correctly identified (ref. 2). This type of key usually works best when the interpretation task is straightforward. Because of the broad range of variability in wheat signatures and the sometimes confusing nature of the low-resolution Landsat imagery, this type of key was not practical for LACIE.

A selective key usually consists of illustrations and descriptions of the objects of interest and is designed for comparative analysis. The analyst selects from the key the example that most clearly represents the ground cover or the object to be identified. This type of key is well suited to crop identification, and it was determined to be better for the difficult interpretation task in LACIE.

It was planned to develop (1) a basic or introductory volume (Analyst Interpretation Keys, Volume I—Image Analysis Guide for Wheat/Small Grains Inventories) to provide nominal wheat signatures occurring in Landsat imagery and the signature variability caused by environmental influences and regional agricultural practices and (2) supplementary volumes containing regional selective keys for the United States and Canada, the U.S.S.R., and other countries to present annotated Landsat imagery on all biophases for selected reference sites, descriptions of wheat cropping areas, and structured interpretation logic based on crop calendars. Two important factors relating to the design of the keys were project decisions (1) to engage in a large ground-truth collection program throughout the U.S. Great Plains (blind

sites) and (2) to develop methods to "partition" this same region into homogeneous areas where certain factors relating to wheat acreage and yield would be grouped.

Ground observations were being routinely collected at 29 ITS's (United States and Canada), and these data plus the 1976 blind-site ground observations would be available to verify Landsat signatures in the initial version of the AI Keys. Along with these ground data, 2 years of Landsat imagery would be "in house" for construction of the document. Adequate ground truth and a sequence of good Landsat acquisitions over each test site to be considered as a "reference" site were of highest importance.

The initial partitioning of U.S. and Canadian wheat-growing regions by project personnel was scheduled for completion in 1976, and since this coincided with the keys development schedule, this version was incorporated into the keys design. (Later studies have resulted in slightly different delineations of homogeneous U.S. Great Plains regions.)

Additionally, the design of the keys was influenced by the following considerations.

1. The distribution of keys within a country should be adequate to document the major geographic differences in crop signatures.

2. The number of keys within a region would result from balancing the objectives of documenting smaller variations in signatures and retaining a convenient size for the AI Keys.

3. Fields used to illustrate the signatures of crops should be dated in terms of crop development instead of calendar date to minimize reflectance differences due to the intrascene range of planting dates and the geographic differences in average planting date.

4. Imagery used in the AI Keys must closely approximate operational imagery in resolution, color balance, and scale.

For the best use of project resources, the following items were considered to be desirable.

1. The existing data collection and processing systems would be used, if possible, to ensure extensive and continuous data flow for development and updating purposes.

2. The keys development would be closely tied to CAMS operations so that technically accepted products and procedures could be incorporated, testing could be conducted with analysts in a realistic operational environment, and retraining of analysts during the implementation phase would be minimized.

The design from this working group was pre-

sent to and accepted by the Image Analysis Keys Team³ at the March/April 1976 Project Review. Suggestions made by this team at this time and at subsequent reviews were incorporated into the development of the keys. Volume I and Volume II (Analyst Interpretation Keys—United States and Canadian Great Plains Regional Keys) were developed, tested, and used for the first time during LACIE Phase III (1977-78 crop year).

ANALYST INTERPRETATION KEYS DESCRIPTION

Detailed descriptions of the "Image Analysis Guide for Wheat/Small Grains Inventories," Volume I of the AI Keys, and Volume II, "United States and Canadian Great Plains Regional Keys," are presented in the following two sections.

Volume I, "Image Analysis Guide for Wheat/Small Grains Inventories"

Volume I is a synoptical key and basic information source for agricultural interpretation with an emphasis on small-grains identification. This volume is designed to help the analyst to recognize not only the normal spectral signatures of wheat cultivation and phenology but also the range of variations from the normal that occur and to understand the reasons for their occurrence.⁴ Volume I is divided into five major sections; the content of each is discussed herein.

Section 1, entitled "Introduction," gives the organization and use of the keys, the system objectives of LACIE, and the image interpretation objectives of LACIE. The principal wheat-growing regions of the world are described and the necessity of a global crop inventory using remote sensing is discussed.

³The Image Analysis Keys Team was composed of the following: J. G. Baron, R. W. Payne, and W. F. Palmer (LEC); W. E. Hensley and L. C. Wade (NASA Johnson Space Center); W. Draeger and W. Anderson (EROS Data Center); J. Lunch (Central Intelligence Agency); R. I. Welch (Texas A & M University); and W. Williamson (LMSC).

⁴This volume includes or is based on information compiled or developed by the Earth Satellite Corporation under a previous contract. This information includes a substantial portion of the textual material in Section 2 of Volume I.

"The Landsat Data-Acquisition System and MSS Image Products," Section 2, presents a detailed discussion of color, color-infrared (CIR), and black-and-white photography. Particular emphasis is placed on the image characteristics of CIR film and its application to vegetation analysis. The techniques used to produce simulated CIR imagery using the data acquired by the Landsat multispectral scanner (MSS) are discussed. An example of spectral reflectance data for winter wheat gathered by the LACIE Field Measurements Project is depicted. The spectral bandwidths imaged by the MSS are annotated on the winter wheat spectral curves for comparative purposes (fig. 2). In addition, aerial photographic views of several small-grain fields, both in color and CIR, are included to illustrate the signature responses to be expected from the two types of imagery. A comprehensive overview of Landsat operations and of the MSS and its output products is presented. Schematics of the overall Landsat system, the MSS system, and the Landsat groundtracks for a typical day are included to provide background data for the reader. Additional diagrams in this section illustrate seasonal changes in solar elevation and the relationship of the U.S. winter wheat belt to the various solar elevations. The processing sources of image tonal variations in both aerial photography and Landsat imagery are also discussed in this section.

Section 3, "Identification of Wheat/Small Grains Cultivation on Landsat Imagery," addresses the problems of detection, recognition, and identification of crop cultivation on the imagery produced from the Landsat MSS data. This is accomplished by a comprehensive description of the photophenology

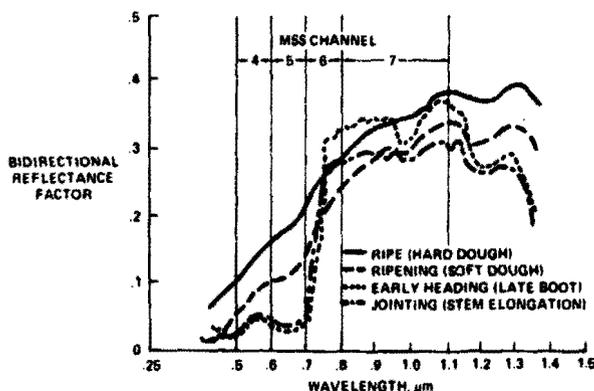
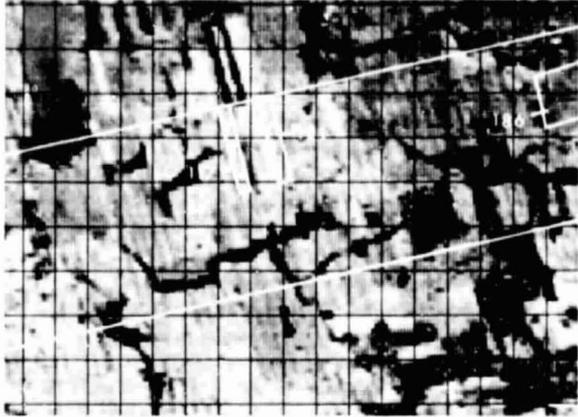
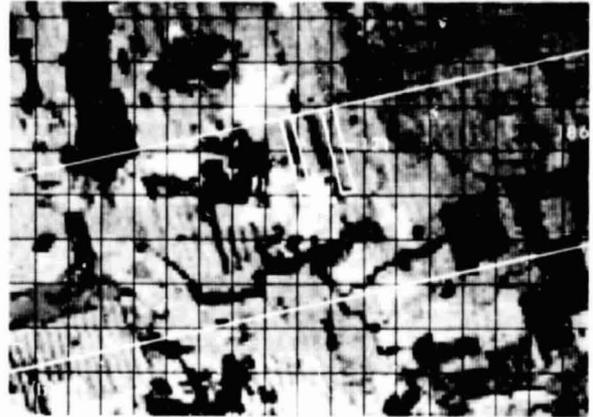


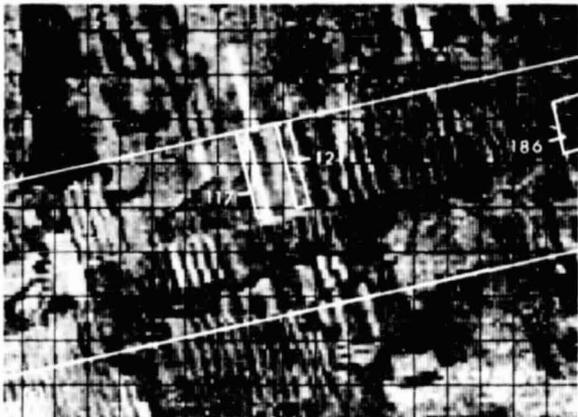
FIGURE 2.—Typical reflectance spectra for dryland Eagle variety winter wheat in various stages (Garden City, Kansas; 1975-76 growing year).



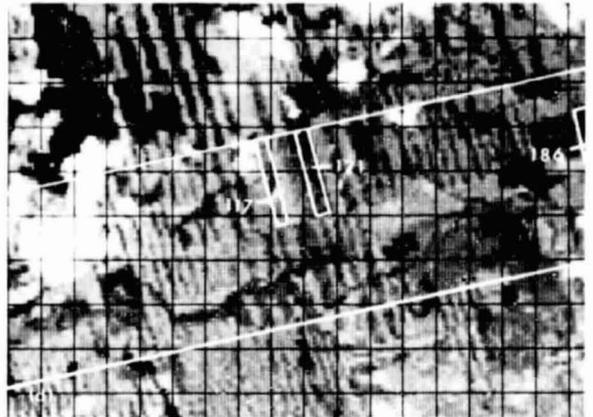
117 - Winter wheat, light pink, tillering, 20 cm (8 in.) high
 121 - Barley, light green, emergence, 5 cm (2 in.) high
 186 - Spring wheat, medium green, emergence, 5 cm (2 in.) high



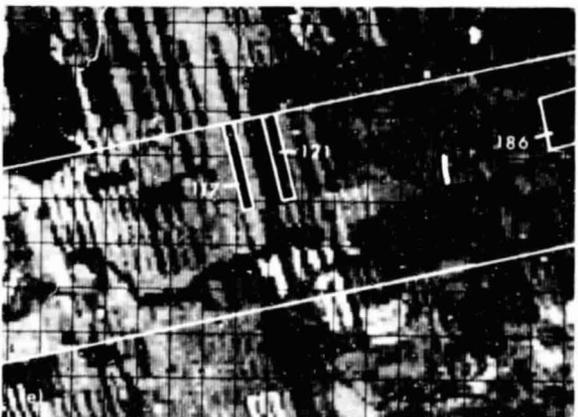
117 - Winter wheat, pink, booting, 41 cm (16 in.) high
 121 - Barley, light pink, tillering, 20 cm (8 in.) high
 186 - Spring wheat, light pink/green, tillering, 10 cm (4 in.) high



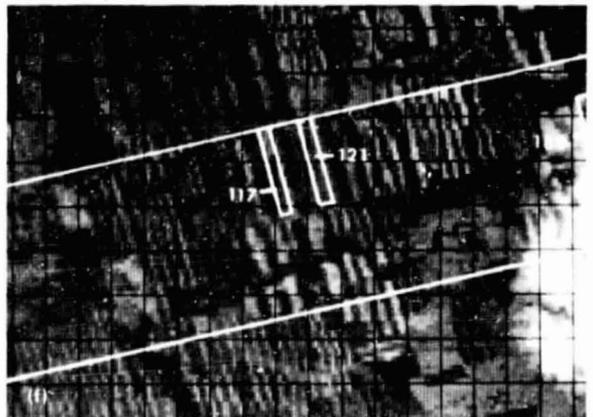
117 - Winter wheat, fully headed, 76 cm (30 in.) high
 121 - Barley, orange, beginning to head, 46 cm (18 in.) high
 186 - Spring wheat, orange, booting, 25 cm (10 in.) high



117 - Winter wheat, light green, ripening, 81 cm (32 in.) high
 121 - Barley, orange, fully headed, 71 cm (28 in.) high
 186 - Spring wheat, orange/olive, fully-headed, 81 cm (32 in.) high



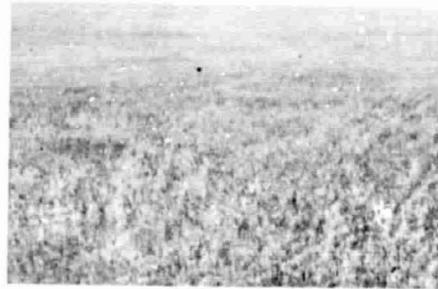
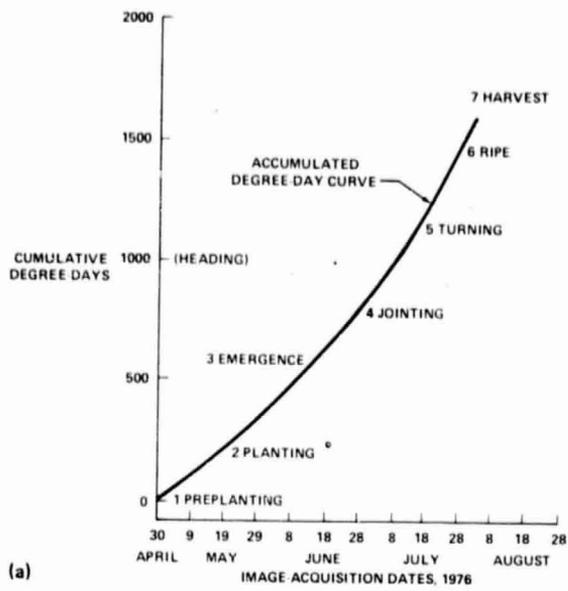
117 - Winter wheat, white, windrowed
 121 - Barley, pink, ripening
 186 - Spring wheat, beginning to ripen



117 - Winter wheat, dark green, fallow
 121 - Barley, medium green, fallow
 186 - Spring wheat, cloud covered

FIGURE 3.—Sample segment images of the Toole County, Montana, intensive test site on six successive dates. (a) June 10, 1975. (b) June 28, 1975. (c) July 16, 1975. (d) August 3, 1975. (e) August 21, 1975. (f) October 14, 1975.

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR



4. Jointing



1. Preplanting



5. Turning



2. Planting (pre-emergence)



6. Ripe

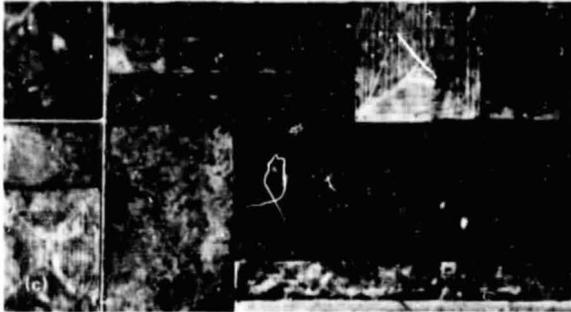


(b) 3. Emergence

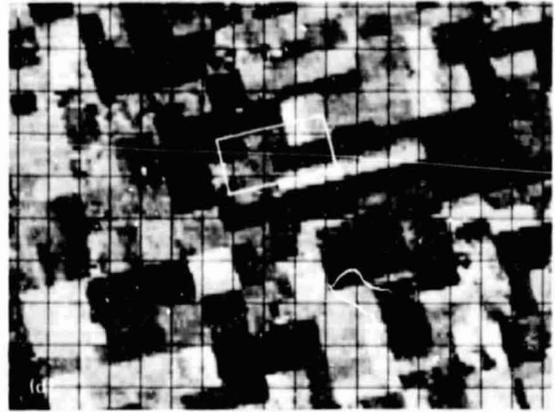


7. Windrowed (harvest)

Harvest (wheat)



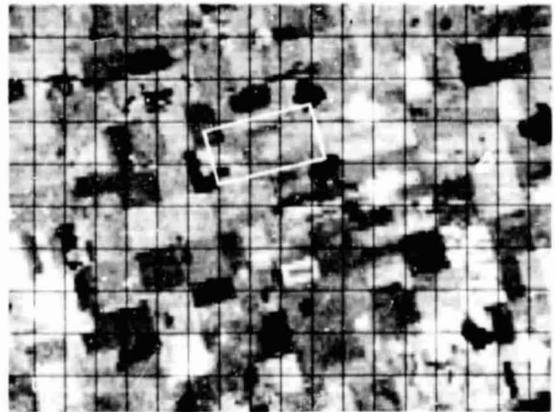
Harvest (wheat)



Jointing (spring wheat)



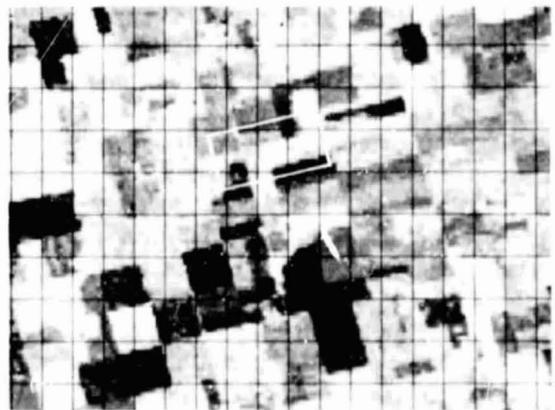
Jointing (spring wheat)



Preemergence (spring wheat)



Preemergence (spring wheat)



←
FIGURE 4.—Multistage wheat photophenology that relates image-acquisition date to the corresponding mean wheat-growth stage and the resultant photographic and image appearance. In this illustration, a late-June image-acquisition date has been extrapolated through the degree-day growth curve into the photophenological example and then into the corresponding aerial and spatial signatures (Hand County, South Dakota). (a) Wheat phenology. (The numbers along the degree-day growth curve correspond to the growth stages shown in figure 4(b).) (b) Wheat photophenology at ground level. (c) Color-infrared vertical photograph, August 11, 1976. (d) Sample segment image, August 23, 1976. (e) Color-infrared vertical photograph, June 20, 1976. (f) Sample segment image, June 30, 1976. (g) Color-infrared vertical photograph, May 5, 1976. (h) Sample segment image, May 7, 1976.

of small grains and by illustrating the application of multitemporal analysis to the detection and identification of the various biostages of wheat and other small grains.

The phenological development of spring and winter wheat is illustrated by a color schematic which depicts the color, height, and general appearance of the wheat plant during the different biostages. Another figure in Section 3 illustrates the concept of multitemporal acquisitions during the various biostages. The concept of multitemporal analysis is further demonstrated by the presentation of six LACIE sample segment images acquired at different times during the year over Toole County, Montana (fig. 3). Spring wheat, winter wheat, and barley fields are annotated on each of the six Landsat CIR images. These data are correlated with ground-level photographs of the three crop types. The ground photography was collected during the corresponding Landsat overpasses.

Section 3 also contains a detailed description of the crop calendar and its role in the identification of small grains and other crops. A LACIE crop calendar for Finney County, Kansas, is shown to illustrate the use of the calendar.

In Section 3, the general techniques for identifying small-grains cultivation on Landsat imagery are described step by step with imagery examples. Figure 4 is a typical example of the imagery. In addition to the annotated Landsat imagery, aerial and ground photographs of different agricultural crops are included to illustrate the varying levels of information contained in each type of imagery. Sequential imagery collected over the LACIE intensive test sites in Williams County (North Dakota), Divide County (North Dakota), Hand County (South Dakota), and Toole County (Montana) is used to illustrate the appearance and signatures of wheat/small grains and other types of ground cover encountered in the Northern Great Plains. Ground-truth maps for the ITS's are included in the sequence. General observations on some similarities and dissimilarities in signatures are made to familiarize the analyst with the different types of patterns and signatures he may encounter. These are also illustrated with imagery examples.

The selective key approach for Landsat agricultural analysis is demonstrated in Section 3, and the use of the LACIE regional keys contained in Volume II, Parts I and II, is described in detail. A regional key from Volume II is included later in this paper for illustration purposes. The decision logic

diagram shown in figure 1 is included to show the decision path followed by the analyst identifying wheat/small grains.

Section 4, "Environmental Effects on Wheat/Small Grains Signatures," provides imagery examples and explanations of the environmental factors that may alter the appearance of small-grains signatures. Specific factors discussed are soil moisture variation, planting date variation, physiographic variations, and the effects of drought, flooding, atmosphere, snow, and wind. A multistage view of the various environmental effects is provided by ground-level photography and aircraft CIR photography, which are correlated to the Landsat imagery examples. Figures 5 to-7 are typical of the data included in this section.

To facilitate the use of the LACIE Keys, photographic examples of common agricultural operations associated with the development and harvest of wheat and other small grains are presented in Section 5, "Common Agricultural Practices for Wheat and Associated Crops." The description and examples in this section deal with the preparation of agricultural land, planting and harvesting operations, and associated cultivation practices. The illustrations and definitions are drawn largely from North America, but they are not necessarily specific to the United States and Canada. Other examples are included to make this section applicable to the various wheat-growing areas of the world.

Cropping and associated practices illustrated in Section 4 include planting operations, crop rotations, fallowing, minimum tillage, irrigation systems, windrowing, and harvest practices (one-stage and two-stage harvesting). Imagery examples of the U.S.S.R. (fig. 8), the People's Republic of China (PRC), India (fig. 9), Australia, Brazil, and Argentina are also provided to illustrate field sizes and shapes typical of these countries. Ground photographs of different types of equipment used in the operations described previously are included. Aerial photographic examples are also provided to illustrate cropping details which, in some instances, are not discernible in the Landsat imagery (fig. 10).

Volume II, "United States and Canadian Great Plains Regional Keys"

Volume II is an operational Analyst Interpretation Key for use in identifying small-grains fields. Pixels within these fields are used as inputs to train a

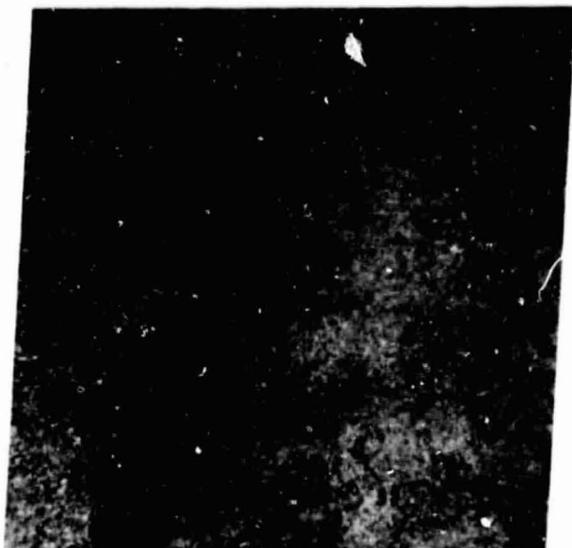


FIGURE 5.—Landsat full-frame image, showing surface-darkening effect from a light rainfall of brief duration (with overlay of total precipitation for March 19-20, 1973).



FIGURE 7.—Sample segment image containing the area shown in figure 6.



FIGURE 6.—Color-infrared high-altitude photograph showing ripe wheat (W), harvested wheat (W/H), windrowed wheat (W/W), and flow (F) fields on glaciated land with large kettleholes (K).



FIGURE 8.—Sample segment image of Khar'kov Oblast in the Ukraine region, U.S.S.R. (November 6, 1975).

pattern-recognition algorithm so that all small grains within a 5- by 6-nautical-mile sample segment can be classified.

Early in LACIE Phase I (1975), the loss of data because of cloud cover and acquisition problems was recognized as the most crucial factor affecting the analyst's ability to label small grains correctly. An analyst might process a segment with a good acquisition history, where the temporal development of small grains was readily apparent, and then, several days later, process a nearby segment with only one or

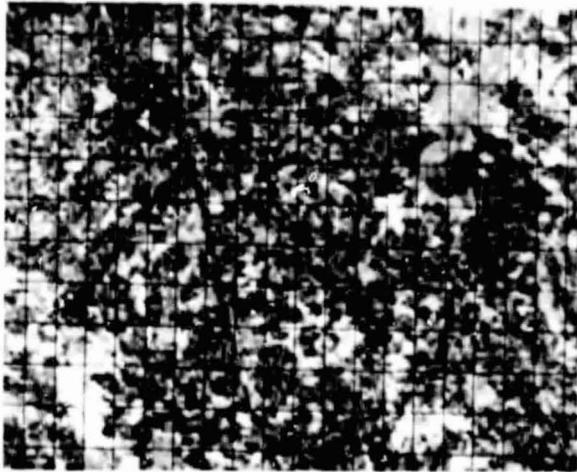


FIGURE 9.—Sample segment image of area near Meerut, Uttar Pradesh, in northern India (March 1, 1975).

two acquisitions and a potentially confusing situation. Some of the more experienced analysts soon began keeping a few "key" segments on hand for use in extending signatures to nearby areas. This practice proved to be very effective for their own reference as well as for training some of the less experienced analysts. Volume II is an attempt to formalize this approach by the systematic selection of appropriate reference segments and the documentation of the procedure for their use.

Regional partitions.—The North American Great Plains was divided into 41 regional partitions based on soils, climate, land use, topography, cropping practices, and the homogeneity of small-grain signatures. (See references 3 to 5 for source materials

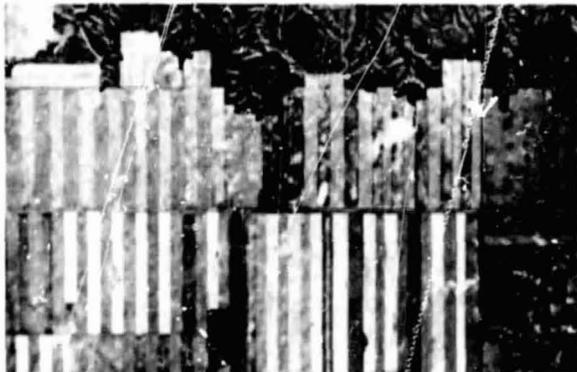


FIGURE 10.—Color-infrared medium-altitude photograph of Richland County, Montana, showing fallow/cropland strips oriented perpendicular to the prevailing wind.

used to compile the partitions.⁵) Thirty-seven of the regions had adequate data to enable the selection of a reference segment for inclusion in the AI Keys. The sample segments selected as reference segments are intended to be representative of the remaining segments within the partitions and, therefore, provide a key for identifying small grains throughout the region.

The regional partitions along with reference segments are listed in table I.

A reduced example of the material available for each partition is shown in figure 11. The actual keys are 11 by 17 inches and are bound in looseleaf binders. Volume II is in two separate binders for ease of handling. Partitions 1 through 19 are contained in Part I of Volume II and Partitions 20 through 39 are in Part II. Both print and transparency copies are available for use by the analysts.

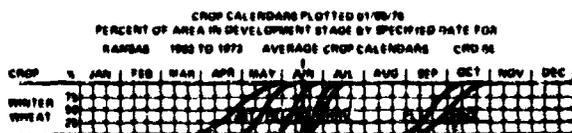
Use of Volume II.—The complete procedure used by analysts to estimate the proportion of small grains in LANDSAT sample segments can be found in the CAMS Detailed Analysis Procedures (ref. 6). One of the tasks performed by the analyst is the labeling of individual pixels as either small grains or non-small-grains. The application of Volume II to the performance of this task is described herein.

The specific steps in using the key are illustrated in the following example. An analyst is assigned segment 1854, which has an acquisition (fig. 12) collected 75/168 (168th day of 1975). After preparing and mounting the imagery according to standard procedures, the analyst should proceed as follows.

1. Establish the principal biostage from the appropriate crop calendar—The analyst should consult the crop calendar adjustment that is closest to the date of acquisition (75/168) and read the biostage for the location of segment 1854. He should adjust the crop calendar in accordance with instructions in the CAMS Detailed Analysis Procedures. In this case, the biostage is 5.3. To apply the key, this acquisition should be defined as biostage 5 and the value 5.3 should be used in step 2 to determine other possible biostages for wheat.

2. Determine other possible biostages from the average crop calendar—Using the average crop calendar from his packet, the analyst should plot the value 5.3 on the 50-percent horizontal line. He should construct a vertical line through this point as shown in the following diagram.

⁵L. D. Pearl, "AI Keys Partitioning Procedure," presentation to NASA management, May 1976.

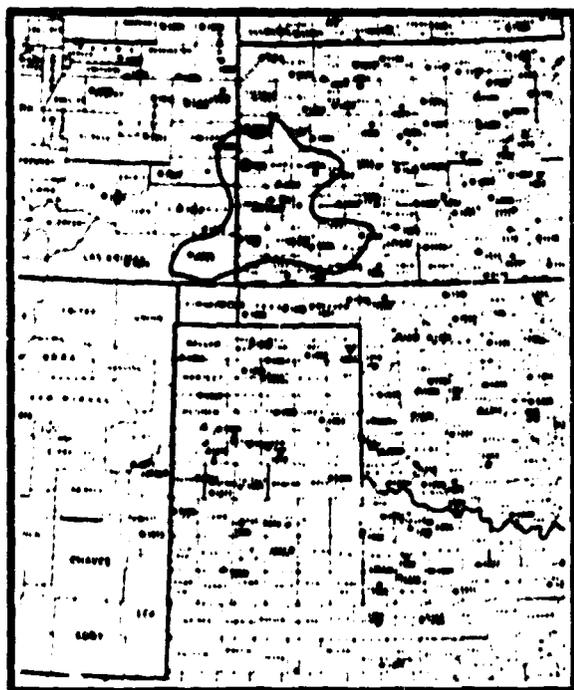


the analyst should conclude that some wheat may be in biostage 4 (heading), some in biostage 6 (ripe), and some in biostage 7 (harvested). Note that steps 1 and 2 have also been accomplished for the first acquisition (75/115—fig. 13) and the results are posted beneath the imagery.

Because the vertical line in this case crosses biostages 5 and 6 and because the accepted tolerance is 10 days,

PARTITION 10 - SOUTHERN CENTRAL HIGH TABLELAND

Segment 1025
Grealey, Kansas



ELEVATION, TOPOGRAPHY, AND SOILS

- Elevations range from 900 to 1200 meters (3000 to 4000 feet).
- This partition consists of smooth loess-mantled tableland slopes that are level (0 to 16 percent slope). Steep slopes border the Arkansas River Valley.
- Most soils in this partition are brown to nearly black, fine-silty and clayey in texture.

CLIMATE

- The average annual precipitation is 38 to 51 centimeters (15 to 20 inches), fluctuating widely from year to year.
- The average annual temperature is 283 to 287 K (50° to 57° F).
- The average freeze-free period is 170 to 185 days.
- Drought, wind, and low temperatures are major hazards.

CROPPING PRACTICES

- In all the Kansas counties, a crop rotation of summer fallow/wheat/sorghum is common. Because of water requirements, wheat does not often directly follow sorghum.
- In Baca County, Colorado, the possible cropping systems are sorghum/fallow/wheat; wheat/fallow; sorghum/fallow; and continuous sorghum.
- Alternate strips of crop and fallow across the direction of prevailing winds are used to help control erosion. The width of the strips varies with erodibility of the soil and size of machines used in the farming operation. Sandy soils require narrower strips than heavier soils.

LOCATION

- This partition is in southwestern Kansas and a small part of southeastern Colorado.

LAND USE

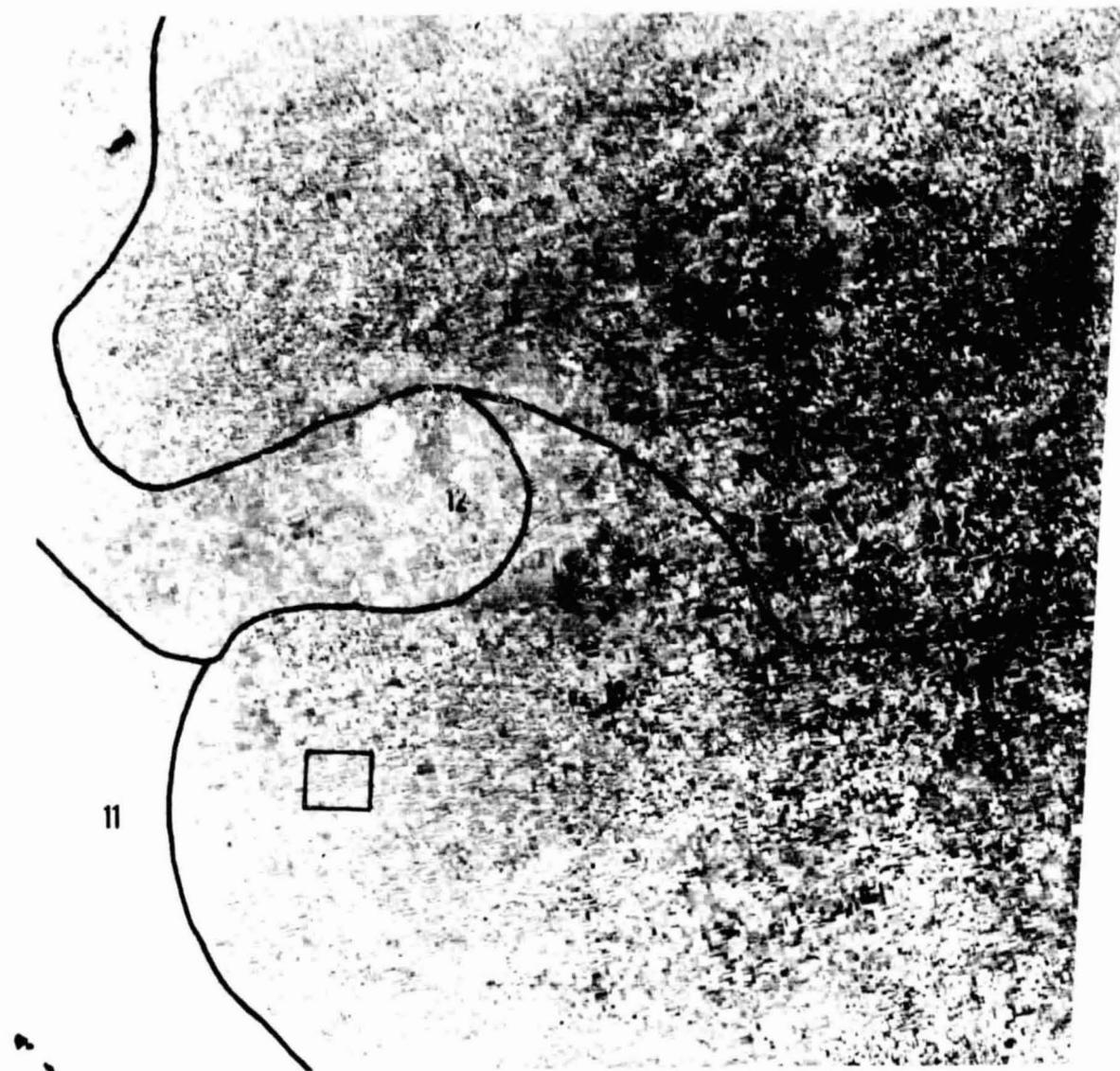
- About three-fifths of this partition is cropland. Most of the remaining area, consisting of hilly and steep slopes bordering drainage ways, is in native grasses and shrubs.
- This partition is a major dryfarming region.
- Winter wheat is the main cash crop, but other small grains, grain sorghum, some corn, alfalfa, and other hay crops occupy large acreages.
- Many kinds of crops are grown in the narrow band of irrigated land along the Arkansas River.

SAMPLE SEGMENTS

- Partition 10 includes sample segments 1025, 1032, 1035, 1041, 1850, 1852, 1854, 1857, 1859, 1861, 1863, 1864, 1866, 1867, and 1868.

(a)

FIGURE 11.—Example partition. (a) Map and general information. (b) Landsat full frame and nominal crop calendar. (c) October 18, 1974, to June 19, 1975. (d) November 10, 1975, to April 20, 1976. (e) June 13 to September 29, 1976.



E5015-16371

May 4, 1975

(b)

FIGURE 11.—Continued.

3. Locate the appropriate reference segment in the key—Using table I in Volume II (a portion of which is reprinted herein as table II), the analyst should find segment 1854 in the left-hand column and read the corresponding partition (10) in the right-hand column.

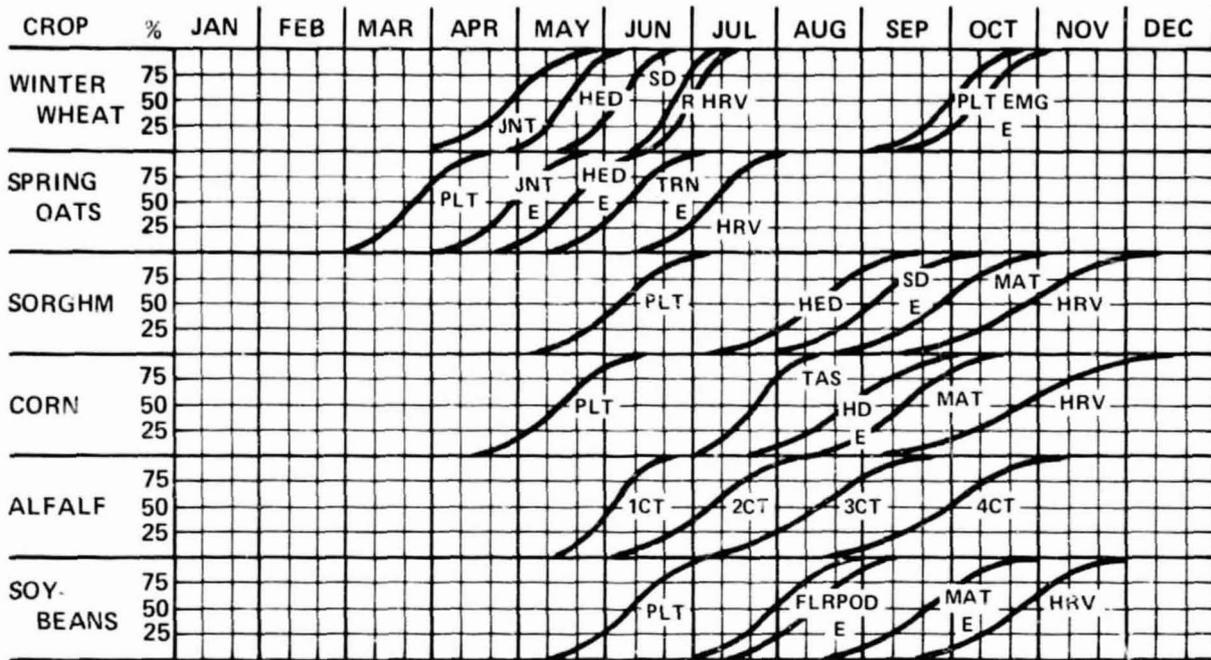
4. Locate fields to be identified—Suppose the analyst must identify dot number 57 (the upper left pixel of the intersection of grid lines 190 and 30) and

dot number 69 (the upper left pixel of the intersection of grid lines 120 and 40). He must first determine that dot number 69 is a part of field A and dot number 57 is a part of field B.

5. Compare these fields to signatures annotated on the reference segment—Since the acquisition has been identified as biostage 5, fields A and B should be compared to fields annotated 5 on the reference segment. If the signatures are not similar, the analyst

CROP CALENDARS PLOTTED 01/06/76
 PERCENT OF AREA IN DEVELOPMENT STAGE BY SPECIFIED DATE FOR

KANSAS 1963 TO 1973 AVERAGE CROP CALENDARS CRD 02



AN E UNDER STAGE NAME INDICATES ROUGH ESTIMATE OF DATE

(b)

FIGURE 11.—Continued.

should compare the fields with other possible biostages (4, 6, and 7). In this example, fields A and B are similar to fields labeled 5 on the reference segment.

6. Follow the logic diagram—Follow the logic diagram (fig. 1) to determine whether these fields are small-grains or non-small-grains fields.

7. Repeat steps 4 through 6 for all available acquisitions and for each field to be identified—In figure 12, acquisition 75/168, both fields should be regarded as similar to fields annotated as biostage 5 on the reference segment. In figure 13, acquisition 75/115, field A appears similar to fields annotated 2 on the reference segment, whereas field B does not appear to be similar to fields labeled 2 or 3, which are the only possible biostages. Following the decision logic diagram, field A should be identified as a small-grains field and field B should be identified as a non-small-grains field.

AI KEYS TEST AND EVALUATION

The introduction of new or modified procedures into the LACIE environment is generally preceded by testing in a quasi-operational mode using a cross section of LACIE analysts. Although the keys concept evolved from the interpretation methods used in LACIE, there are procedural elements different enough to warrant thorough testing and evaluation before implementation into ongoing LACIE analyses.

The test was designed and subsequent ground-truth comparison evaluation conducted by the Research, Test, and Evaluation Branch (ref. 7). The objective of the test was to determine the type and pattern of influence on wheat/small-grains identification accuracy resulting from the introduction of the interpretation keys and associated decision logic into operational use in LACIE.

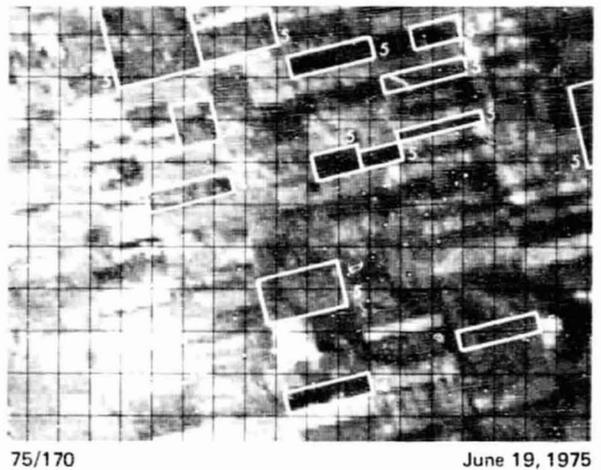
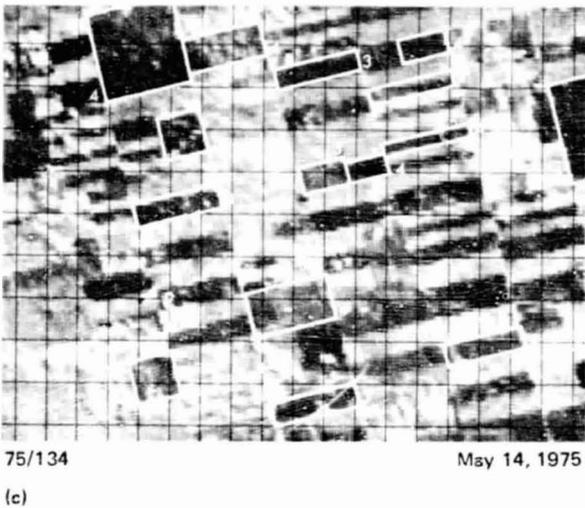
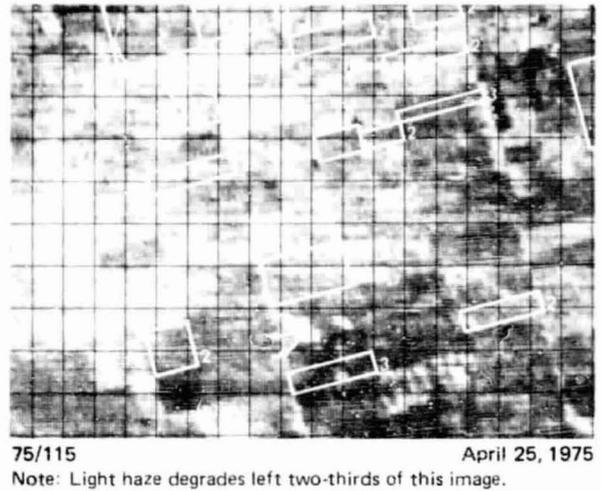
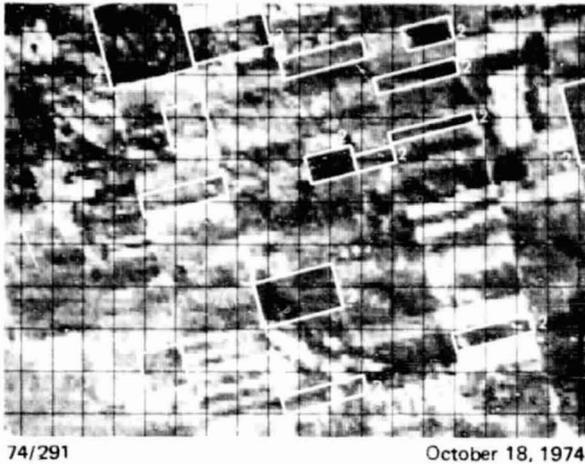


FIGURE 11.—Continued.

Three segments were chosen randomly from each of the 4 U.S. spring/mixed wheat partitions, providing a total of 12 segments for testing. Twenty to thirty fields were selected in each segment so that the full range of wheat and nonwheat signatures was represented. Each segment had from two to five acquisitions. Each segment was interpreted eight times with the keys by a group of analysts and eight times without the keys by a different set of analysts.

The test approach was to use 16 analysts grouped according to 4 levels of LACIE experience. The four teams consisted of the following.

1. Analysts with little LACIE experience who had no familiarity with the U.S. spring wheat regions

2. Analysts with little LACIE experience who had some familiarity with U.S. spring wheat regions

3. Analysts with LACIE experience in areas other than the U.S. spring wheat regions (e.g., U.S.S.R. or PRC)

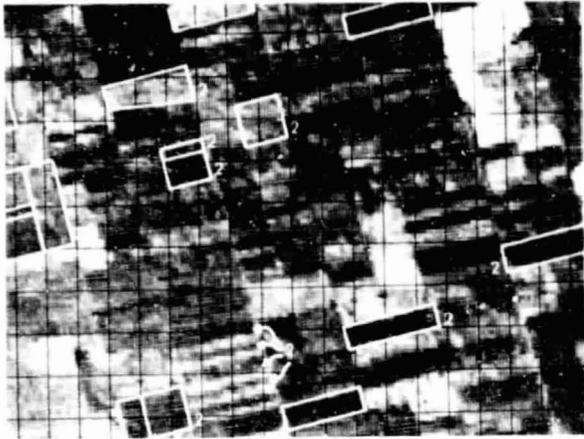
4. Analysts with LACIE experience in the U.S. spring wheat regions

Total errors in field labeling (small-grains/non-small-grains) were tested using analysis of variance (ANOVA) methods (ref. 8). The findings and conclusions from the test are summarized as follows.

1. The interpretation accuracy for all four groups of analysts improved significantly with the use of the AI Keys. The total error was reduced in each group,

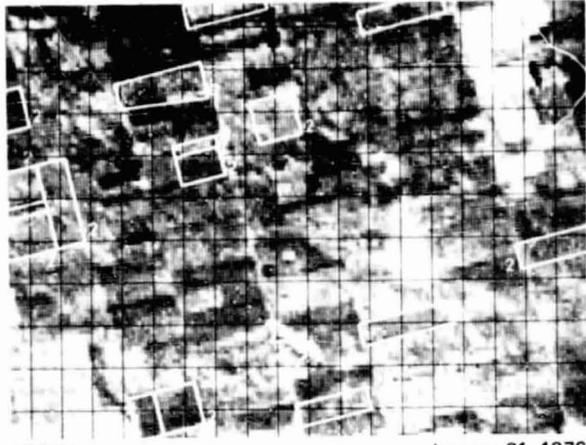
REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

1976 CROP YEAR



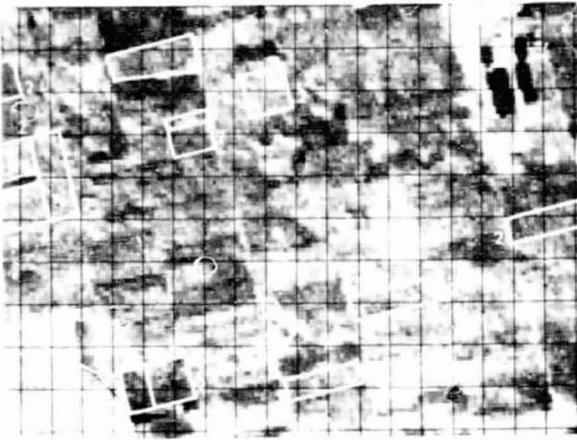
75/314

November 10, 1975



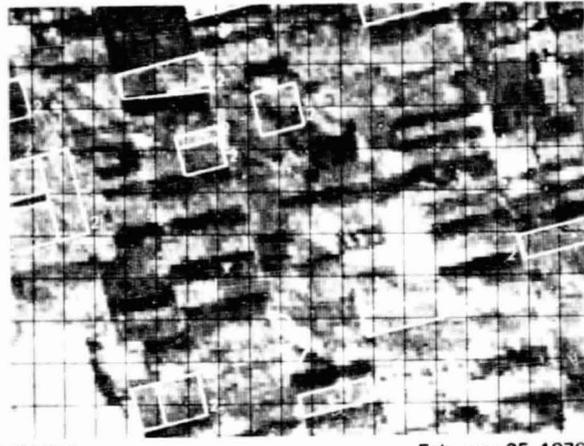
76/021

January 21, 1976



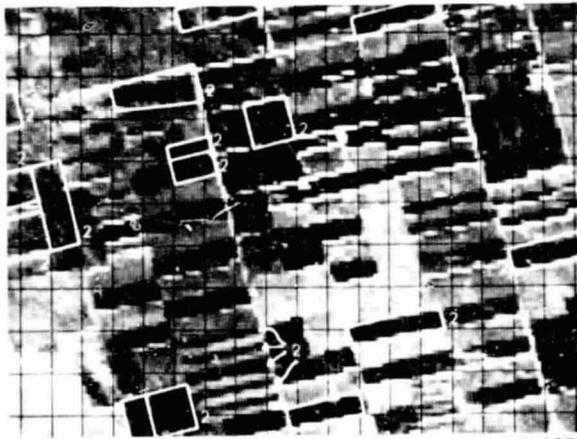
76/039

February 8, 1976



76/056

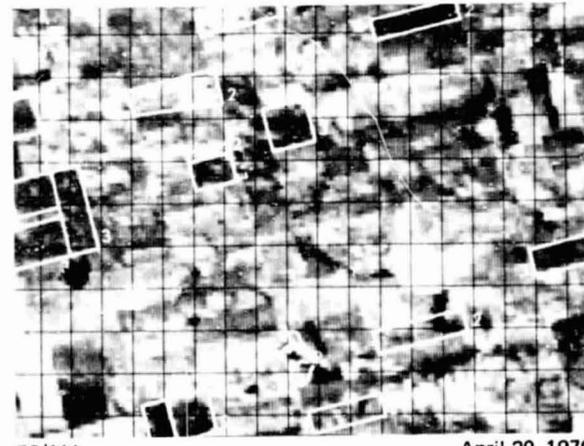
February 25, 1976



76/092

April 1, 1976

Note: A cold front caused freezing in northwestern Kansas. White signatures are probably frost; no snow was reported this date.



76/111

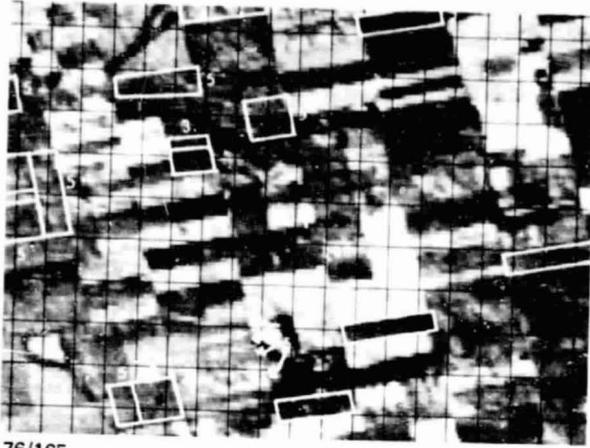
April 20, 1976

Note: Drought resulted in thin, spotty fields. Note difference in 1975 and 1976 crop years; some winter kill may be evident in upper portion of the image.

(d)

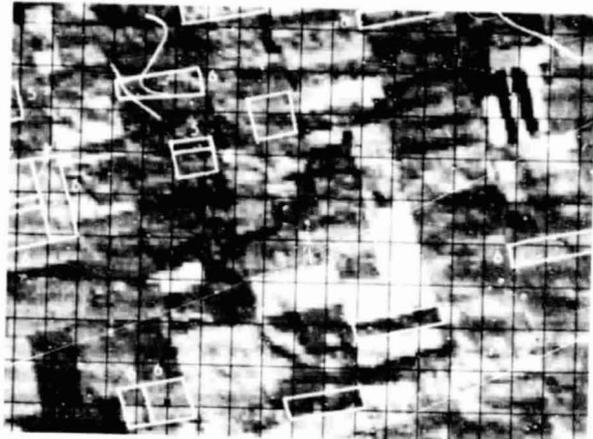
FIGURE 11.—Continued.

1976 CROP YEAR



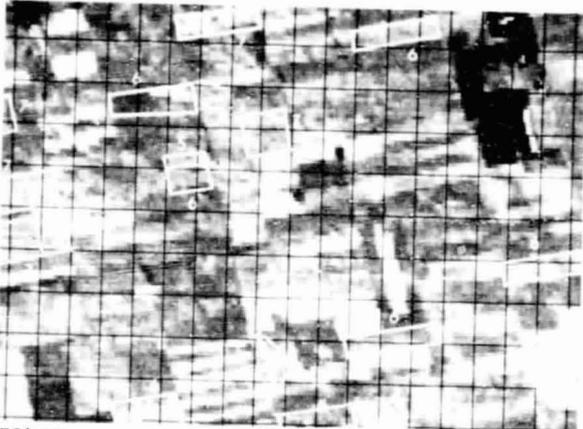
76/165

June 13, 1976



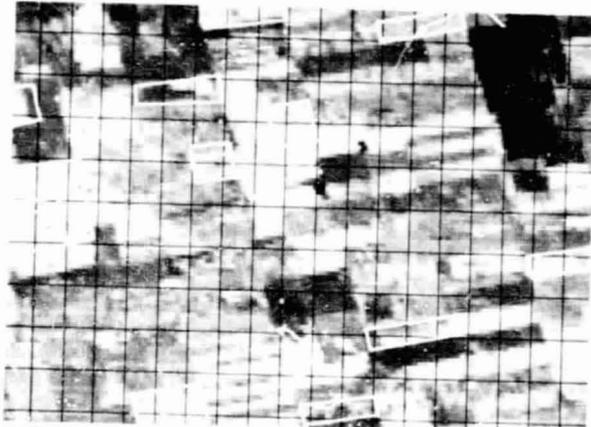
76/173

June 21, 1976



76/182

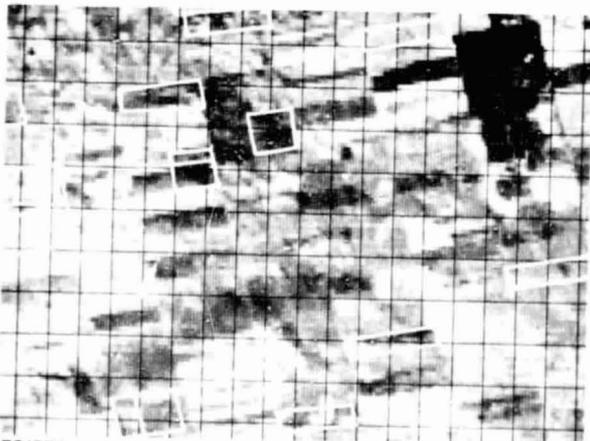
June 30, 1976



76/200

July 18, 1976

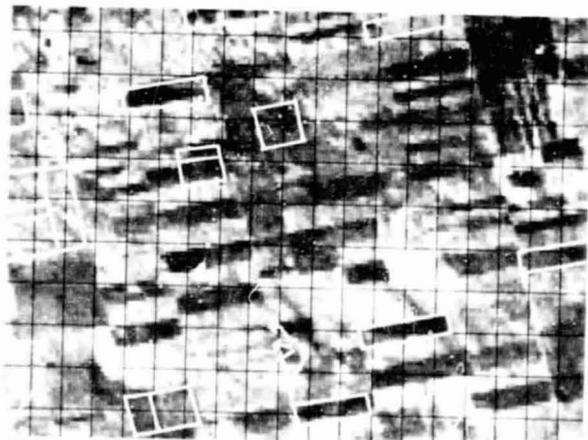
Note: County wheat crop failed in April; the unharvested wheat was plowed under and sorghum was planted, most of which failed late in the season. Biostage numbers were omitted to show fields greening up again.



76/254

September 10, 1976

(e)



76/273

September 29, 1976

FIGURE 11.—Concluded.

TABLE 1.—Regional Partitions and Reference Segments

<i>Partition number</i>	<i>Partition description</i>	<i>Reference segment number, Phase II</i>	<i>Reference segment location</i>
1	Northern Texas Blackland Prairie	1274	Fannin Co., Tex.
2	Central Rolling Red Prairie	1241	Woods Co., Okla.
3	Central Rolling Red Plains (East)	1259	Baylor Co., Tex.
4	Central Rolling Red Plains (West)	1230	Greer Co., Okla.
5	Southern High Plains	1084	Swisher Co., Tex.
6	Central Rolling Red Plains (North)	1232	Kiowa Co., Okla.
7	Cherokee Plains	1178	Bourbon Co., Kans.
8	Great Bend Sand Plains	1889	Edwards Co., Kans.
9	Northern third of the Southern High Plains	1865	Stevens Co., Kans.
10	Southern Central High Tableland	1025	Greeley Co., Kans.
11	Upper Arkansas River Valley Rolling Plains and South Central High Plains	1005	Cheyenne Co., Colo.
12	North Central High Tableland	1093	Yuma Co., Colo.
13	Central High Tableland	1851	Graham Co., Kans.
14	Rolling plains and breaks	1875	Osborne Co., Kans.
15	Central Loess Plains, Bluestream Hills, and Central Kansas Sandstone Hills	1181	Cowley Co., Kans.
16	Nebraska and Kansas Loess Drift Hills	1574	Colfax Co., Nebr.
17	Central Nebraska Loess Hills	1588	Adams Co., Nebr.
18	Mixed sandy and silty tableland/Middle Central High Plains	1562	Cheyenne Co., Nebr.
19	Wyoming-South Dakota-Upper Platte River Valley	1682	Haakon Co., S. Dak.
20	Rolling Pierre Shale Plains/South Dakota-Nebraska Eroded Tableland	1694	Lyman Co., S. Dak.
21	Eastern Black Glaciated Plains	1674	Faulx Co., S. Dak.
22	Loess, till, and sandy prairies	—	—
23	Western Minnesota forest-prairie transition	1521	Grant Co., Minn.
24	Red River Valley of the North	1681	Roberts Co., S. Dak.
25	Central Black Glaciated Plains	1622	Ramsey Co., N. Dak.
26	Rolling Soft-Shale Plains and Southern Dark-Brown Glaciated Plains	1629	McLean Co., N. Dak.
27	Northern Rolling High Plains and Rolling Soft-Shale Plains	1555	Fallon Co., Mont.
28	Northern Rolling High Plains	1556	Powder River Co., Mont.
29	Northwestern Black Glaciated Plains	1606	Ward Co., N. Dak.
30	Dark-Brown Glaciated Plains	1538	McCone Co., Mont.
31	Northern Rolling High Plains and Northern Smooth High Plains	—	—
32	Northern Rocky Mountain Foothills	1732	Glacier Co., Mont.
33	Brown Glaciated Plain	1739	Teton Co., Mont.
34	Southwestern Saskatchewan and northern Montana	3081	Saskatchewan
35	Southeastern Saskatchewan and southwestern Manitoba	3129	Saskatchewan
36	Eastern Saskatchewan and western Manitoba	3158	Saskatchewan
37	South-central Saskatchewan	3122	Saskatchewan
38	Southwestern Saskatchewan	3133	Saskatchewan
39	Alberta	3256	Alberta
40	Central Alberta and central Saskatchewan	—	—
41	Northwestern Alberta	—	—

TABLE II.—Segment/Partition Cross Reference

Segment	Partition	Segment	Partition	Segment	Partition
1745	33	1856	13	1879	15
1746	28	1857	10	1880	14
1747	32	1858	13	1881	14
1748	32	1859	10	1882	14
1749	32	1860	13	1883	15
1750	32	1861	10	1884	15
1751	32	1862	9	1885	15
1752	32	1863	10	1886	14
1753	28	1864	10	1887	14
1850	10	1865	9	1888	15
1851	13	1866	10	1889	8
1852	10	1875	14	1890	8
1853	14	1876	15	1891	8
1854	10	1877	14	1892	15
1855	14	1878	14	1893	8

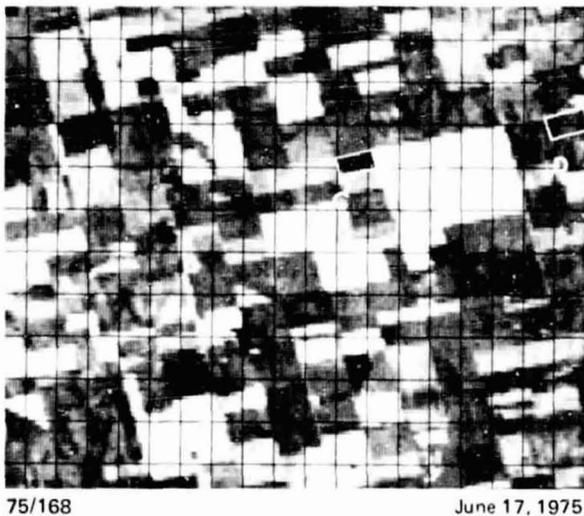
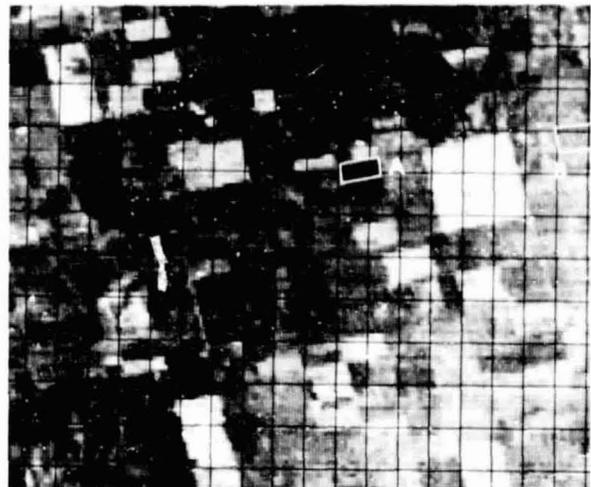


FIGURE 12.—Segment 1854 (Scott, Kansas) in biostage 5; other possible biostages are 4, 6, and 7.

as is shown in figure 14(a). Group 4 had the most improvement (3 percent), and groups 2 and 3 showed the least improvement.

2. The omission and commission error rates were reduced by using the AI Keys. The greater reduction, approximately 8 percent, was in the omission error rate. The commission error rate improved by approximately 2 percent (fig. 14(b)).



75/115 April 15, 1975

FIGURE 13.—Segment 1854 in biostage 2; the only other possible biostage is 3.

3. The AI Keys improved accuracy in each of the four agrophysical partitions, as is shown in figure 14(c). The general locations of the partitions are given in table I.

4. Use of the AI Keys improved labeling accuracy during all four biowindows. The greatest increase in accuracy occurred in biowindow 1 (8 percent), with the least improvement in biowindow 2 (2 percent). Figure 14(d) shows the accuracy improvement for each biowindow.

CONCLUSION

The LACIE Analyst Interpretation Keys Volume I, "Image Analysis Guide for Wheat/Small Grains Inventories," and Volume II, "United States and Canadian Great Plains Regional Keys," were developed during Phase II of LACIE (1976) and implemented during Phase III (1977). The AI Keys were tested using operational LACIE data, and the results demonstrate that use of the AI Keys provides improved labeling accuracy in all analyst experience groupings, in all geographic areas within the U.S. Great Plains, and during all periods of crop development (biowindows).

To document the complete range of signature variability and temporal sequences, several additional years of data may be necessary. Volume II currently contains the 2 years of segment imagery which were available during the development of the keys in

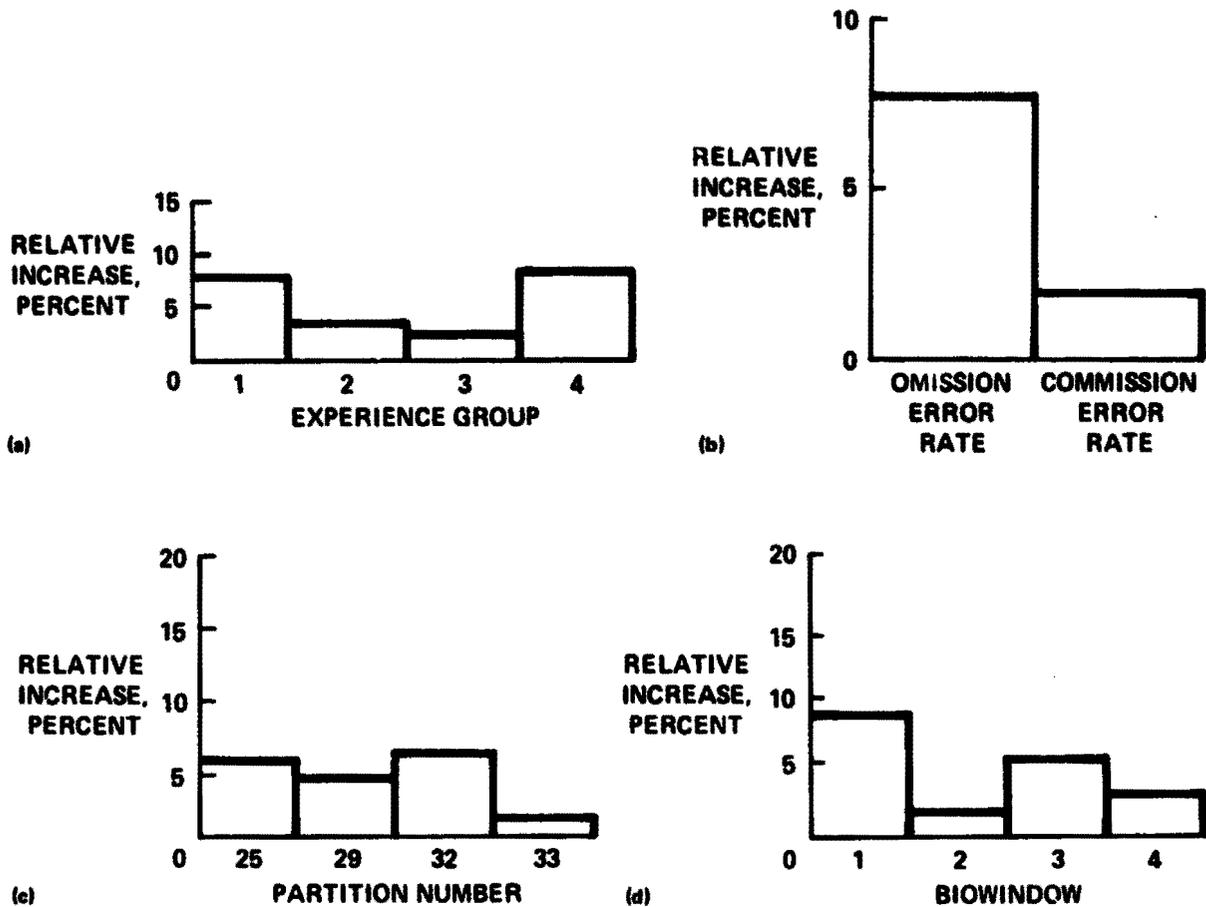


FIGURE 14.—Accuracy improvement using the AI Keys. (a) Accuracy improvement by analyst experience group. (b) Accuracy improvement for omission and commission error rate. (c) Accuracy improvement for each agrophysical partition. (d) Accuracy improvement by biowindow.

1976. As improved analysis techniques, imagery products, and spectral aids are developed and tested, these should be incorporated into the AI Keys volumes to reflect current technology in use for satellite agricultural surveys.

ACKNOWLEDGMENTS

The design and development of the Analyst Interpretation Keys were conducted under NASA Contract NAS 9-15000 by Lockheed Electronics Company (LEC) with developmental support from Lockheed Missiles and Space Company (LMSC). In addition to the authors, LEC technical contributors to the design and development of the keys were E. E. Edwards, D. C. Helmer, M. A. Lundelius, C. V. Nazare, L. D. Pearl, J. C. Prill, and D. E. Urbassik.

The LMSC contributors were T. L. Anderson, R. R. Gilliam, G. S. Mansfield, and A. W. Richmond.

Technical consultants to the design and development of the Analyst Interpretation Keys were R. I. Welch of Texas A & M University, H. L. Hansen of the U.S. Department of Agriculture (USDA), and J. W. Clifton (retired) of the USDA. For technical and administrative guidance, the help of J. L. Dragg, W. E. Hensley, and L. C. Wade of the NASA Johnson Space Center is especially acknowledged.

REFERENCES

1. Wheat Identification Aid for Image Interpreters. LEC-3730, Lockheed Electronics Co., Houston, Tex., June 1974.

2. Departments of the Army, Navy, and Air Force: Image Interpretation Handbook, Vol. I. Dec. 1967.
3. USDA: The Look of Our Land—An Airphoto Atlas of the Rural United States: North Central. Agriculture Handbook No. 384, 1970.
4. USDA: The Look of Our Land—An Airphoto Atlas of the Rural United States: The Mountains and Deserts. Agriculture Handbook No. 409, 1971.
5. USDA: The Look of Our Land—An Airphoto Atlas of the Rural United States: The Plains and Prairies. Agriculture Handbook No. 419, 1971.
6. Large Area Crop Inventory Experiment (LACIE) Phase III CAMS Detailed Analysis Procedures. LACIE-00720 NASA Johnson Space Center (Houston, Tex.), Aug. 1977.
7. Analyst Interpreter (AI) Keys Test Field Labeling Accuracy. LEC-9922, Lockheed Electronics Co., Houston, Tex., Mar. 1977.
8. Cochran, W. G.; and Snedecor, G. W.: Statistical Methods. Iowa State University (Ames), 1974.

Colorimetric Consideration of Transparencies for a Typical LACIE Scene

R. D. Juday^a

INTRODUCTION

In the LACIE operations, analyst interpretation of transparencies made from Landsat data was an important factor. The Earth Observations Division of the NASA Johnson Space Center (JSC) embarked on schemes designed to provide the analyst with optimal film products. Two such products, designated Product 1 and Product 3 (the latter also being known as the Kraus product), were the ones principally used. Certain aspects of those products are reviewed here from the standpoint of color theory. The examination of the two products has led to a clarification of some principles of color image generation for remote sensing and has given a quantitative basis for the development of a new class of imagery based on those principles, reported elsewhere in the LACIE symposium proceedings.

Details of the mathematical calculations are given in the appendix; an attempt is made in the text to give short, simple explanations of some perhaps unfamiliar colorimetric terminology. The reference list is a guide to a more thorough education.

A complete colorimetric evaluation includes at least two subtasks: consideration of the color presentation of information within a single transparency and consideration of the image-to-image stability of the presentation. In this report, the former is quantitatively treated and the latter is qualitatively discussed.

COLORIMETRY

The machine used to produce the LACIE imagery is the FR-80 manufactured by Information International, Incorporated, and installed at the Johnson

Space Center. It is locally known as the production film converter (PFC). It operates by imaging a black-and-white cathode-ray-tube display sequentially through colored filters onto color reversal film, with the result being thought of as writing red, blue, and green images through independent channels. Each of the channels is configured so that color densitometry on the developed film shows a linear relationship between transmission density (being the logarithm of the transmission, measured at a wavelength near the spectral peak for the channel under consideration) and the input for that channel (being a number of digital counts in the range 0 to 255).

Normal color vision is three dimensional (ref. 1). That is, three properly chosen primary lights are sufficient to match any other colored light by an additive process in which the primaries are added in varying proportion (including negative contributions, in which the negative contribution is attained by adding that primary to the light to be matched). Many coordinate systems, all three dimensional, are used to describe color. Examples include the CIE¹ system (luminosity and two chromatic coordinates) and the Munsell system (hue, chroma, and value) (ref. 2). There are mathematical relationships allowing passage between the systems. The three-dimensional system having the counts in the three channels of the PFC as basis vectors can be related to the standard color systems by appropriate measurements; this has been done for the PFC, and the details of the mathematical model are given in the appendix. The considerable body of colorimetric theory is then accessible for a discussion of the PFC and the LACIE film products.

One color system that is particularly apropos to

^aNASA Johnson Space Center, Houston, Texas.

¹The Commission Internationale de l'Éclairage, or the International Commission of Illumination, which sets the international standards for the specification of color.

this discussion is a uniform chromaticity scale (UCS) space. The characteristic of a UCS space is that Euclidean distance in that color space is directly proportional to the perceptibility of the difference between neighboring colors. Theory shows that even though color vision is three dimensional, the distortion of the usual color spaces required to match the stated quality of the UCS space cannot be displayed in a Euclidean space of three dimensions because of negative Gaussian curvature. However, various approximations to a true UCS can be displayed in a three-dimensional Euclidean space, and the ($L^*a^*b^*$) approximation, being evaluated by the CIE (ref. 3), will be used. The parameter L^* is associated with lightness; a^* , with a balance between the complementary colors green and magenta; and b^* , with the complementary colors blue and yellow. The unit of distance is the just noticeable difference (JND) discussed in reference 4. The proportionality between Euclidean distance and perceptibility of color difference holds only locally (i.e., for small color differences such as those between various shades and lightnesses of green) rather than globally (as for large color differences, such as those between bright pink and navy blue). But under the rationale that large color differences are made up of small ones, the proportionality will be applied to large color differences where required. The ($L^*a^*b^*$) system is attractive for its analytic invertibility, a property that will be seen to be highly advantageous.

One last approximation should be noted: the transformation between counts and colors is treated as continuous, even though the transformation exists only for integer count values.

A colorimetric model of the PFC has been generated, as outlined in the appendix.² The model describes the relationship between PFC input count vectors and the ($L^*a^*b^*$) space, with the transparency being viewed on a cool white fluorescent light table (i.e., the viewing illumination forms part of the colorimetric description). Specifying the transformation between PFC input counts and feature space then allows the passage between data space and the UCS space. Some of the results of examining the feature-space/color-space relationship are described in a later section, but some of the concepts that will be used are introduced here.

²Richard D. Juday, "Colorimetric Principles As Applied to Multichannel Imagery," Master's thesis, to be published.

The linear sensitivity of a color product is the number of perceptible color steps (which include brightness differences) moved through for a unit motion in the data space. The dimensions of linear sensitivity are JND's per count. For Products 1 and 3, linear sensitivity is a function of both location in the data space and the direction of motion.

As mentioned previously, the transformation between data space and color space is regarded as continuous, despite the 16-level truncation and the underlying digital nature of the transformation. Under this approximation, which will yield general truths, a planar distribution of points in data space occupies a two-dimensional surface (generally curved) in color space. A differential area in data space dA_d maps into a differential area in color space dA_c . Respective units are square counts and square colors. This step leads to the chromatic expansion ratio $(dA_c)/(dA_d)$, with units of square colors per square count. (There is similarly a chromatic expansion ratio, cubic colors per cubic count, for the full three-dimensional color transformation. For now, however, only planar distributions in data space are considered.) The chromatic expansion ratio is one measure of how the available area (or volume) in color space is budgeted to the corresponding portions of data space.

A finite set of discriminable colors is produced by the PFC. In analogy to the statisticians' term "probability mass," the discriminability mass is defined as the volume of the UCS space accessible to the PFC. On a two-dimensional surface in the UCS space, a smaller number of discriminable colors exists; for that surface, the discriminability mass is approximately one color times the area of the surface. An optimum transformation of data into color space will have a surface area of largest extent, subject to other constraints such as continuity, monotonicity, smoothness in linear sensitivity, and orthogonality (angle-preserving nature) in transforming data from data space into color space.

Visual orthogonality is a concept deserving a little more explanation. It is tied in turn to geodesic paths in color spaces. Imagine all possible curves connecting two color experiences in a particular color space. Now imagine that you are observing a color patch the color of which is changing in the manner described by motion along one of those paths. Keep track of the number of times that the color of the observed patch just barely noticeably changes in color in moving along the path, and assign to the path the number of changes. The geodesic path between the colors is the path that has the minimum number of changes.

The perceptibility of color difference is proportional to the JND's counted along the geodesic path connecting the colors. Around any color, there is, in general, an ellipsoid the points of which are just barely noticeably different from the center. The geodesic path between colors passes through the smallest number of the JND ellipses. For small regions about any color center, the color space can be warped so that the ellipsoids in the region become spheres (this is the local definition of a UCS space). In that local region, with the ellipses distorted into spheres, the Euclidean metric ($s^2 = \Delta x^2 + \Delta y^2 + \Delta z^2$, with s being the distance between points) is proportional to the color difference. In a UCS space, marks made on a color path at JND intervals will be equally spaced along the path. Motions occurring in perpendicular directions in the UCS space will have the greatest difference in the *kinds* of change in color. In psychophysical terms, examples are brightness change compared to hue change or hue change compared to saturation change. Brighter and dimmer are in opposite directions certainly but are the same *kind* of change. The author proposes that a desirable feature of a data-to-color transformation is for different *kinds* of change in data space to correspond to different *kinds* of color change; i.e., the transformation from data space to color space should be orthogonal, or angle preserving. Operations can be performed in data space before the transformation, of course; e.g., to emphasize numerical features of particular interest.

HISTORY OF PRODUCTS 1 AND 3 ALGORITHM DEVELOPMENT

Product 1 predates Product 3. It independently scales three bands of Landsat data (multispectral scanner (MSS) channels 1, 2, and 4) into the three inputs of the PFC so that the six standard deviations centered on the mean of an MSS channel occupy in linear fashion the full range of the PFC's assigned input channel. The goal in its formulation was maximal chromatic expansion with only an acceptable minority of points being saturated (falling beyond the 0- to 255-count input range of the PFC). It is readily apparent that the chromatic representation of a point in feature space will vary with the statistics of the scene; as the means and standard deviations of the data move in feature space, the relationship between a given feature space vector and input counts

for the PFC will be altered. Analyst experiences led to a desire for a "true color" product, in which a given feature space vector would have a more highly consistent appearance than as put onto film under Product 1. In Product 1, each channel has its own gain and bias associated with the means and standard deviations in each MSS channel. For Product 3, the MSS statistics are used to govern essentially only one parameter, an overall gain. There is zero bias, and the gain (change in PFC counts per change in MSS counts) has a constant relationship among the three PFC channels. In Product 3, a feature space vector transformed into the PFC gun cube would be changed in magnitude, but not in direction, by varying scene statistics. Both direction and magnitude could vary under Product 1. The equations for the Product 1 and 3 transformations from feature space to PFC input space are given in the appendix.

COLORIMETRIC ASSUMPTIONS FOR PRODUCT 1 AND 3 ALGORITHMS³

In scaling data linearly into the PFC, one assumes either that the PFC is a UCS machine or that the manufacturer of the machine "knew" somehow the manner in which it was to be used and therefore built into it a feature causing it to concentrate chromatic discriminability into the desired parts of the input cube. Neither is quite the case, actually; the manufacturer (together with the film supplier and the film developer) supplies a machine that has so large a gamut of expressible colors that it is a little difficult to go wrong. That fact and the quality control exercised over the fine tuning of the machine and film processing give a stability of color representation sufficient for much of the analysts' success.

The uniform chromaticity assumption for the PFC arises naturally. The Weber-Fechner law states that to be just noticeably brighter, one patch of light must exceed in luminance that of another patch by a constant fraction. The linearization of the PFC in transmission density (ref. 5) for each input channel gives precisely that result for single activation of

³Some fairly free interpretation on the part of the author is present in this section. Complete accord by all the JSC principals is not to be expected, but the author will happily defend his viewpoints.

each channel. In other words, the PFC is a UCS machine when only one of the channels at a time is activated. Because the result of activating two channels at a time gives a result that looks different than either of the channels that were added together, it is easy to assume that the PFC generalizes from uniform scale along the axes to uniform scale anywhere in the input cube. The author believes that in the development of Products 1 and 3, without its ever having been explicitly stated, the PFC was implicitly regarded as having UCS properties over all its input range. The failure in chromatic consistency in the appearance of a feature type under Product 1, rather than a lack of uniform discriminability, led to the effort of defining Product 3.

The actual nonuniform discriminability of features separated by a constant Euclidean distance in feature space was masked by a truncation of the PFC's input range to 16 levels in each channel for both products. The truncation was done for visually esthetic reasons—many things were tried, and a few were chosen. Field boundaries are reported to be more sharply distinct with the inputs limited to 16 levels per channel. The net result is that a coarser grid of color points is occupied by data, with inter-point discriminability "crowbarred" to be larger. This is a poor man's approach to increased visual separability.

The goal of Product 3 was color fidelity; the underlying assumption is that maintaining relative proportions among the three inputs to the PFC will maintain the chromaticity (hue and saturation) of a picture element (pixel), allowing only lightness to vary. If the PFC inputs were in proportion to the transmission rather than to the logarithm of the transmission, that goal would have been realized. With the existing linearization in density, there is not the chromatic constancy expected. Product 3, however, comes much closer to that concept of chromatic fidelity than does Product 1.

Color vision is three dimensional; that is why there are no more than three inputs to the PFC (or any other color-generating process, though the four-color printing process uses black ink to darken a color). Thus, a dimensionality-reduction scheme is needed in making imagery from Landsat MSS data because the four channels in the MSS are spectrally independent. Of several schemes (such as calculating principal components for four-channel distribution or taking a standard rotation of the four-dimensioned data and displaying the first three in either case), the simplest scheme—dropping one channel—was

chosen. Because of a high degree of correlation between channels 3 and 4 of MSS data, channel 3 was eliminated. The channel assignments between MSS channels and PFC channels were made to achieve a high degree of similarity between the appearance of the final product and the color-infrared aerial photographs with which the analysts were already very familiar. Thus, MSS channel 4 was assigned to red, channel 2 to green, and channel 1 to blue.

QUANTITATIVE RESULTS

A typical LACIE scene was subjected to colorimetric analysis; the details are given in the appendix and this section contains only the quantitative results.

Figures 1, 3, 5, 7, and 9 are, in order, the linear sensitivity for separations in the tasseled cap's greenness direction, the cosine of the angle between the transformed brightness and greenness directions (which are at right angles in data space), the chromatic expansion ratio, and the JND ellipses for Product 1. Figures 2, 4, 6, 8, and 10 are similar quantities for Product 3. The axes are the tasseled cap's brightness and greenness. The curvilinear outline is the envelope of the scatter plot of the data; the full scatter of the data is fairly well confined close to the plane of the figure and within the outline. The rhombohedral outline shows the limits reachable by the PFC without saturation.

Figures 1 to 4 show that Product 1 has a noticeably higher linear sensitivity than does Product 3. (Contour lines are marked with the values of the linear sensitivity ratio.) Significant portions of the data fall outside the saturation limits for Product 1; Product 3 does not saturate high-brightness/low-greenness pixels, but the upper end of the green arm nonetheless saturates. A common feature of figures 1 to 4 is the presence of a marked variation in the linear sensitivity across the scatter plot, with higher values of sensitivity toward larger values of the tasseled cap's brightness. Linear sensitivity, however, does not tell the whole story, as begins to become apparent in the next two figures.

Figures 5 and 6 give the cosine of the angle between transformed unit vectors that in data space parallel the tasseled cap's brightness and greenness directions (and hence are perpendicular in data space). For Product 1, there is a strong antialignment between the transformed unit vectors. As a result, some directions of separation in data space are nearly

indistinguishable, whereas data space separations at right angles to the nearly indistinguishable direction are highly visible. The linear sensitivity is thus a function of the direction of separation in data space, and determination of the sensitivity in only two directions does not describe the full situation. (It

takes three parameters to describe an ellipse—eccentricity, rotation, and major axis, for example.) Incidentally, an orthogonal transformation would have not only $\cos \theta = 0$ but also equal linear sensitivity in all directions.

Figures 7 and 8 give an idea of the budgeting of the

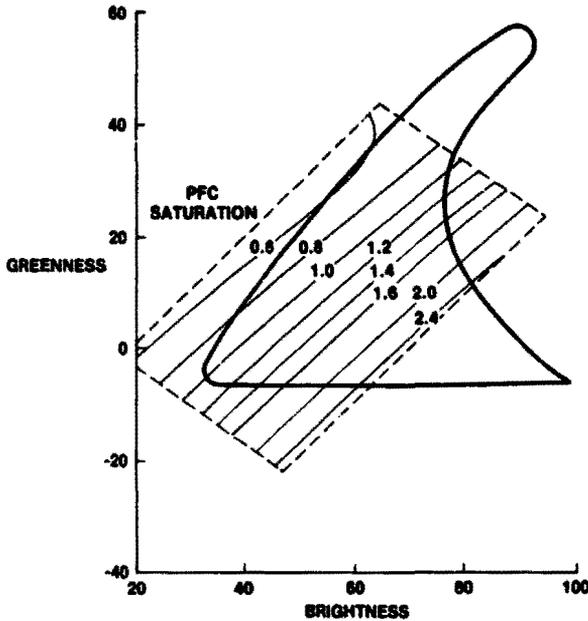


FIGURE 1.—Product 1 brightness sensitivity. Contour values are expressed in colors per count.

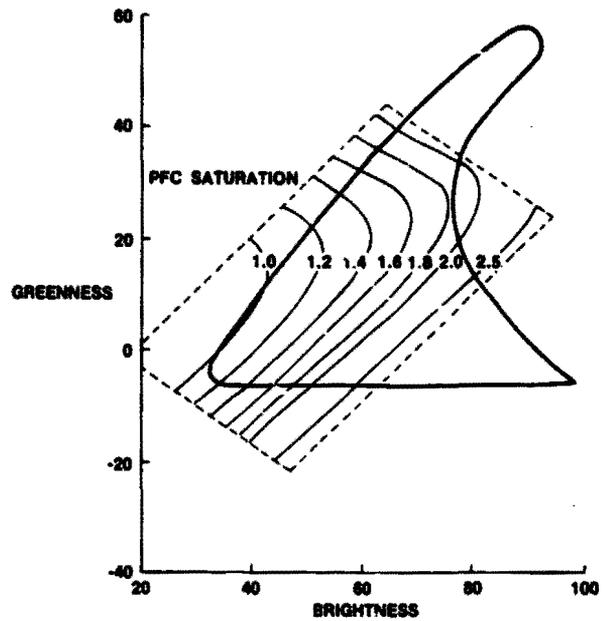


FIGURE 3.—Product 1 greenness sensitivity. Contour values are expressed in colors per count.

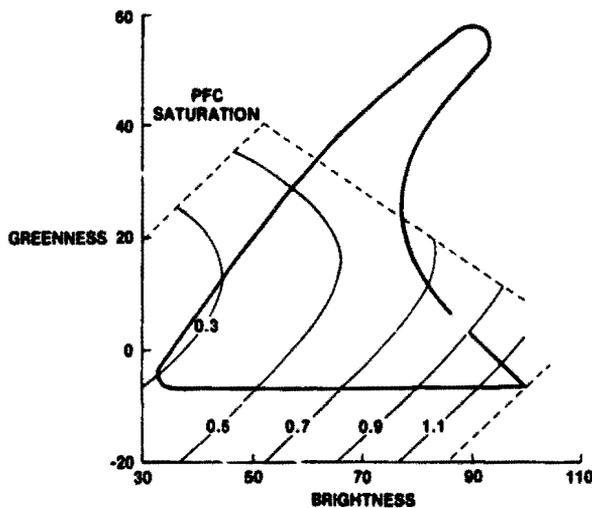


FIGURE 2.—Product 3 brightness sensitivity. Contour values are expressed in colors per count.

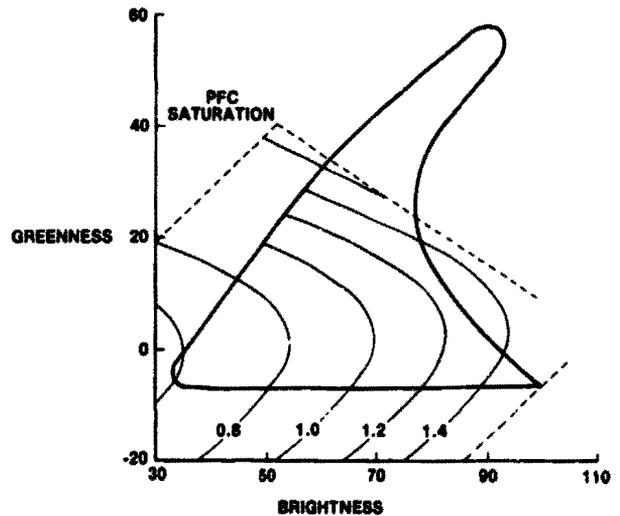


FIGURE 4.—Product 3 greenness sensitivity. Contour values are expressed in colors per count.

available discriminability mass in the data space. The fact that the saturation boundaries more closely conform to the data distribution for Product 1 immediately indicates that less of the discriminability mass is allocated to portions of data space actually unoccupied by pixels; that is confirmed by these two

figures. Both products concentrate the discriminability mass toward higher tasseled cap brightness rather than toward regions rationally chosen as being relevant to an image analysis problem (such as catching the motion of a pixel off the soil line as it begins to green up).

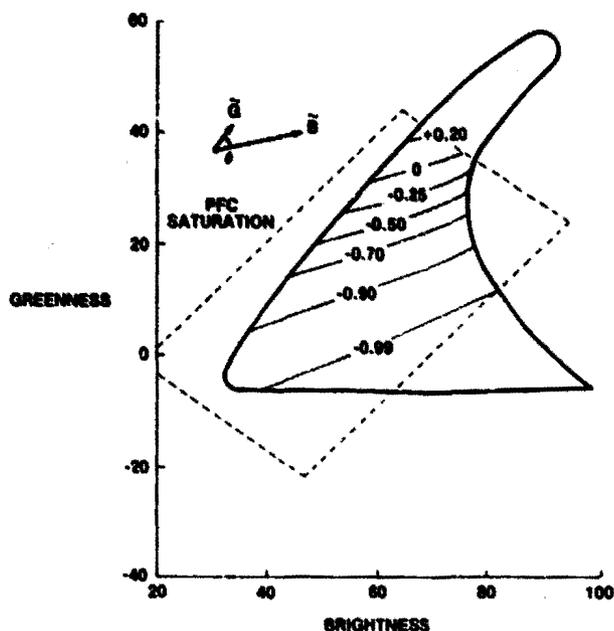


FIGURE 5.—Cosine of the angle between the transformed brightness and greenness directions for Product 1.

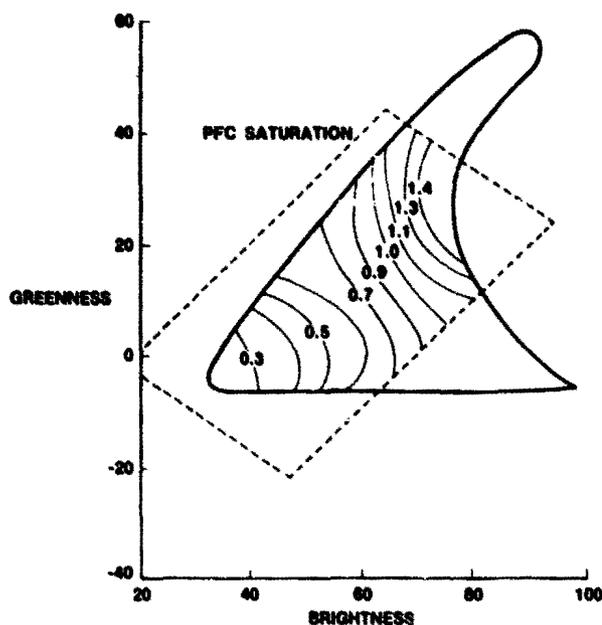


FIGURE 7.—Product 1 discriminability mass distribution. Contour values are expressed in square colors per square count.

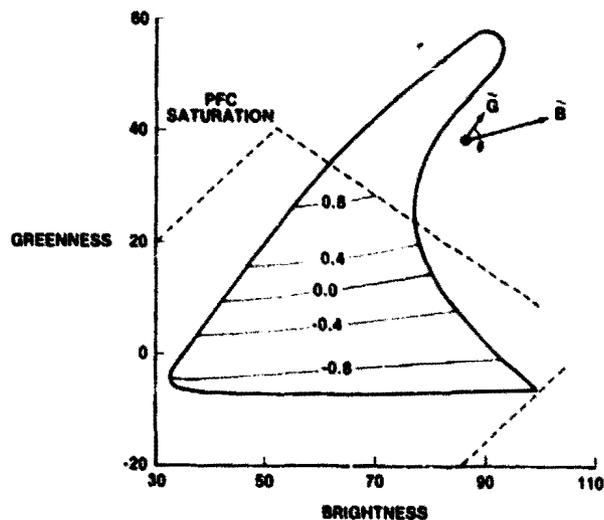


FIGURE 6.—Cosine of the angle between the transformed brightness and greenness directions for Product 3.

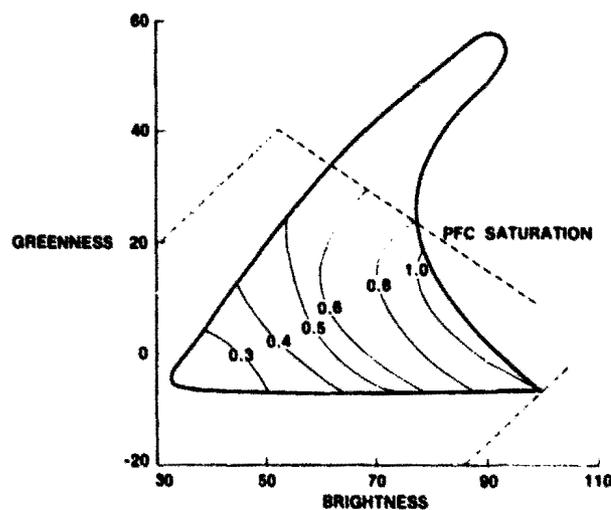


FIGURE 8.—Product 3 discriminability mass distribution. Contour values are expressed in square colors per square count.

The discriminability mass distribution does not tell the whole story, either. The discriminability is not a scalar; a given amount of discriminability mass can be squashed out flat, with the result that a large amount of variation in the directional characteristics of linear sensitivity is introduced without alteration of local values of square colors per square count. The JND ellipses, however, do tell the whole story.

Figures 9 and 10 summarize all the information in the preceding eight figures. The ellipses plotted in these two figures are the intersections of the JND ellipsoids with the brightness/greenness data plane (at the average value of the third and fourth tasseled cap parameters, yellowness and nonesuch). The distance from the center of an ellipse to the curve in a particular direction is how far one moves a pixel in data space before barely being able to perceive the color change—again, under the approximation that there is a continuous transformation between data space and color space. The chromatic expansion ratio is the inverse of the area of an ellipse. The cosine of the angle described earlier is deducible, but the ellipses give all the information contained in $\cos \theta$, and more. The linear sensitivity is the inverse of the distance from center to curve; therefore, all parameters are present.

From previous discussion, desirable features of a counts-to-color transformation would be that the ellipses are as small as possible and that their eccentricity and orientation accommodate realistic discrimination problems. The portion of UCS space accessible to any color-generating machine is limited, and the problem of making optimal color imagery is one of budgeting in the allocation of the color space to the data space. For example, if the task is crop identification, one does not wish to spend color volume discriminating between clear and turbid water. Until another criterion is put onto quantitative basis, the author proposes that small values of eccentricity (near-circular) are desirable because analysts are accustomed to looking at scatter plots in unscaled coordinates; having equal-sized circles for the JND ellipses gives analysts the same visual perspective on the data. In an example of a rationale for other than small, low-eccentricity ellipses, for communication between an analyst and a computer doing classification on the basis of Mahalanobis distance (ref. 6), the JND ellipses giving the computer and the analyst the same perspective would be aligned with the Mahalanobis ellipses and given the same eccentricity.

The major differences in the JND ellipse behavior

of Products 1 and 3 are that the ellipses for Product 3 are typically larger in area and more nearly circular; i.e., there is a trade-off between chromatic expansion ratio and orthogonality in the two products. The JND ellipses for Product 1 are very elongated in parts of data space, with the major axis extending in

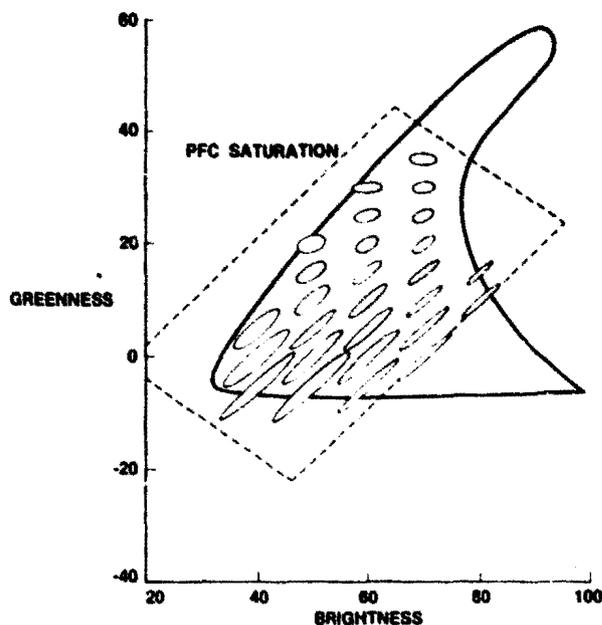


FIGURE 9.—Product 1 JND ellipses (1.5×).

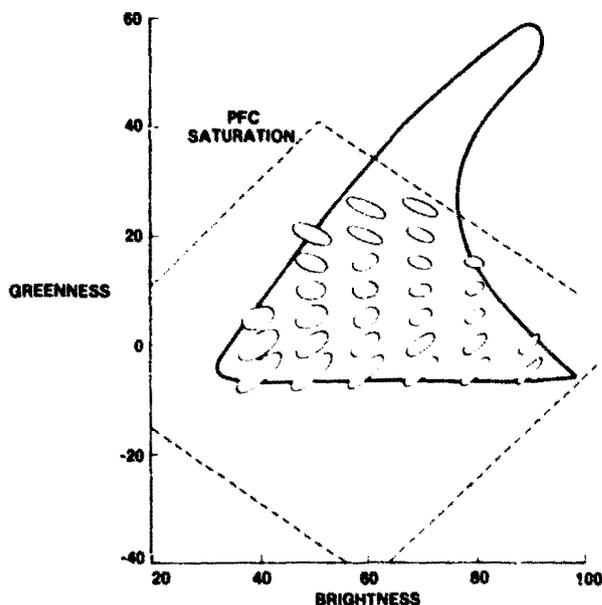


FIGURE 10.—Product 3 JND ellipses (1.5×).

C-10

the direction of very poor linear sensitivity mentioned earlier. The motion of a pixel off the soil line is in a direction to be highly visible, being in a direction of increasing greenness and decreasing brightness. The ellipses are so aligned that the moved pixel cannot be distinguished from a pixel of lower initial brightness that is still lying on the soil line. That comment is true of both products, but more strongly so of Product 1. Both products also show a variation in size and orientation of the JND ellipses over the scatter plot. This variation is a consequence of the non-UCS behavior of the PFC.

The preceding discussion is in terms of a continuous relationship between Landsat MSS count space and the three-dimensional UCS space. In specializing to the many-to-one count vector-to-color relationship introduced by the truncation to 16 levels per channel, the ellipses will be replaced by rectilinear cells. Their general alignment and other features will follow the trends shown by the ellipses.

IMAGE-TO-IMAGE COLOR STABILITY

Both Product 1 and Product 3 perform manipulations on data vectors before sending them off to the PFC for conversion to color. The attempt is to standardize between images against data variations from Sun angle etc. Product 1 subtracts a bias vector and scales each component independently; Product 3 subtracts no bias and applies to the channels scale factors that derive from a single parameter. Thus, from image to image, Product 1 modifies a data vector in both direction and magnitude, whereas Product 3 essentially modifies only its magnitude. (Note that by subtracting a bias vector parallel to the data vector, Product 3 could have had a bias included that would have left the direction of the data vector unchanged. The result would have been increased chromatic expansion.) Product 3's stated goal was chromatic fidelity; by a standardizing transformation that left the data direction unchanged, it is implicit that the PFC was regarded to have some sort of constant chromatic behavior, as an input vector is being changed only in magnitude. Product 3's data standardization comes much closer to chromatic consistency than does Product 1's because of the unchanged direction of the modified vector.

Along a radial path from the origin in the PFC input cube, a constant ratio of counts exists among the activated channels. If the PFC were linearized in transmission rather than in transmission density, the

radial paths would have constant hue and saturation. Because of the logarithmic relationship between input counts and transmission, however, paths of constant hue and saturation are curved. Perhaps the analyst tends to compensate for the curve, but this has not been investigated.

Because Product 3 is governed by a single parameter, one expects that there will be color consistency for a given data vector in images of scenes having similar statistics. Indeed, this trend has been reported for scenes having similar tasseled cap brightness (R. Cicone, private communication).

ACKNOWLEDGMENT

The author expresses his appreciation to F. Johnson of JSC for assistance with computer graphics, to R. Carver of JSC for aid in colorimetric measurements, and to the technical reviewers R. Cicone of the Environmental Research Institute of Michigan and T. Kaneko of International Business Machines for helpful comments.

REFERENCES

1. Wyszecki, G.; and Stiles, W. S.: *Color Science*. John Wiley and Sons, New York, 1967. (Especially Sec. 6.7 ff.)
2. Nimeroff, I.: *Colorimetry*. NBS Monograph 104, Jan. 1968.
3. Proposal for Study of Uniform Color Spaces and Color-Difference Evaluation. CIE's Compte Rendu 18^e Session, Publication CIE N 36, 1976, p. 171.
4. MacAdam, David L.: Visual Sensitivities to Color Differences in Daylight. *J. Opt. Soc. America*, vol. 32, 1941, p. 247.
5. McCamy, C. S.: *Color Densitometry*. Color: Theory and Imaging Systems. Society of Photographic Scientists and Engineers, 1973.
6. Duda, R. O.; and Hart, P. E.: *Pattern Classification and Science Analysis*. John Wiley and Sons, 1973.
7. Kauth, R. J.; and Thomas, G. S.: The Tasseled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops As Seen by Landsat. *Proceedings of the 1976 Machine Processing of Remotely Sensed Data Symposium (Purdue University)*, June 21-23, 1977, p. 77.

Appendix

Equations for Mapping From MSS Data Space to Color Space

The mapping from Landsat data to PFC input is as follows. The subscript *R* refers to the red PFC channel; the subscript *G*, to green; and the subscript *B*, to blue. Numerical subscripts refer to MSS channels 1 to 4. Means are indicated by μ ; standard deviations, by σ .

Product 1

$$u_i = \mu_i + 3\sigma_i \quad i = 1,2,4$$

$$l_i = \mu_i - 3\sigma_i \quad i = 1,2,4$$

Product 3

$$\mu_1 = 2/3 (1.1\mu_i + \mu_2 + 2\mu_4)$$

$$u_2 = u_1$$

$$u_4 = u_1/2$$

$$l_i = 0 \quad i = 1,2,4$$

Scales and biases:

$$S_i = \frac{255}{u_i - l_i}, \quad b_i = -l_i S_i \quad i = 1,2,4$$

PFC channel counts:

$$C_R = S_4 \bar{x}_4 + b_4$$

$$C_G = S_2 \bar{x}_2 + b_2$$

$$C_B = S_1 \bar{x}_1 + b_1$$

where \bar{x} is the MSS count vector. Truncation to 16 levels follows:

$$d_j = \left[\frac{C_j}{16} \right] \times 16 \quad j = R,G,B$$

with $[y]$ being the largest integer $n \leq y$.

The colorimetric model² relates the CIE tristimulus values

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

to the PFC input count vector $\vec{\tau}$

Film transmission density:

$$\vec{D} \equiv \begin{pmatrix} D_B \\ D_G \\ D_R \end{pmatrix} = \begin{pmatrix} 2.513 \\ 2.182 \\ 2.520 \end{pmatrix} - 0.001 \begin{pmatrix} 7.133 & 0 & 0 \\ 0 & 7.357 & 0 \\ 0 & 0 & 6.929 \end{pmatrix} \begin{pmatrix} d_B \\ d_G \\ d_R \end{pmatrix}$$

Film transmission:

$$\tau_j = 10^{-D_j} \quad j = R,B,G$$

PFC channel activations:

$$a_j = \frac{\tau_j}{\tau_{j,max}}$$

where $\tau_{j,max}$ is obtained from $C_j = 255$.

Tristimulus values:

$$\vec{Q} \equiv \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 13.53 & 35.90 & 39.74 \\ 10.86 & 47.09 & 32.04 \\ 46.13 & 6.646 & .09343 \end{pmatrix} \begin{pmatrix} a_B \\ a_G \\ a_R \end{pmatrix}$$

²Richard D. Juday, "Colorimetric Principles As Applied to Multichannel Imagery," Master's thesis, to be published.

The illumination source was found to be

$$Q_o = \begin{pmatrix} X_p \\ Y_p \\ Z_p \end{pmatrix} = \begin{pmatrix} 89.18 \\ 100 \\ 52.89 \end{pmatrix}$$

Reference 3 gives the UCS approximation:

$$L^* = 25 \left(100 \frac{Y}{Y_o} \right)^{\frac{1}{3}} - 16$$

$$a^* = 500 \left[\left(\frac{X}{X_o} \right)^{\frac{1}{3}} - \left(\frac{Y}{Y_o} \right)^{\frac{1}{3}} \right]$$

$$b^* = 200 \left[\left(\frac{Y}{Y_o} \right)^{\frac{1}{3}} - \left(\frac{Z}{Z_o} \right)^{\frac{1}{3}} \right]$$

The Landsat count vector

$$\begin{pmatrix} X_1 \\ X_2 \\ X_4 \end{pmatrix}$$

is obtained from the tasseled cap values

$$\begin{pmatrix} KB \\ KG \\ KY \\ KN \end{pmatrix}$$

by (ref. 7):

$$\vec{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_4 \end{pmatrix} \begin{pmatrix} 0.33231 & -0.28317 & -0.89952 & -0.01594 \\ .60316 & -.66006 & .42830 & .13068 \\ .26278 & .38833 & -.04080 & .88232 \end{pmatrix} \begin{pmatrix} KB \\ KG \\ KY \\ KN \end{pmatrix}$$

The particular scene analyzed had values

$$\begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_4 \end{pmatrix} = \begin{pmatrix} 26.174 \\ 22.146 \\ 19.165 \end{pmatrix}$$

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_4 \end{pmatrix} = \begin{pmatrix} 3.249 \\ 5.233 \\ 4.981 \end{pmatrix}$$

$$\langle KY \rangle = -11.6$$

$$\langle KN \rangle = -0.38$$

which suffice to relate values of

$$\vec{K} = \begin{pmatrix} KB \\ KY \end{pmatrix}$$

to values of $(L^* a^* b^*)$.

$$\vec{e} = \begin{pmatrix} L^* \\ a^* \\ b^* \end{pmatrix} = f(\vec{K})$$

with $f(\cdot)$ being invertible.

The linear sensitivity LS in the direction $\vec{\delta}$ (in the tasseled cap plane) is the limit as $|\vec{\delta}|$ becomes small of

$$LS = \frac{|\Delta \vec{e}|}{|\vec{\delta}|}$$

where

$$\Delta \vec{e} = f(\vec{K} + \vec{\delta}) - f(\vec{K})$$

Note that LS is generally dependent on both \vec{K} and $\vec{\delta}$. LS is plotted in figures 1 to 4.

Obtaining LS for $\vec{\delta}_1$ paralleling KB (giving $\Delta \vec{e}_1$), $\vec{\delta}_2$ paralleling KG (giving $\Delta \vec{e}_2$), and the intermediate direction $\vec{\delta}_3$ (giving $\Delta \vec{e}_3$) yields the cosine of the angle between $\Delta \vec{e}_1$ and $\Delta \vec{e}_2$.

$$\vec{\delta}_3 = \vec{\delta}_1 + \vec{\delta}_2$$

$$\Delta \vec{e}_3 = \Delta \vec{e}_1 + \Delta \vec{e}_2$$

$$|\Delta \vec{e}_3|^2 = |\Delta \vec{e}_1|^2 + |\Delta \vec{e}_2|^2 + 2 |\Delta \vec{e}_1| |\Delta \vec{e}_2| \cos \theta$$

The values of $\cos \theta$ are shown in figures 5 and 6.

The area of the parallelogram having sides $\Delta \vec{e}_1$ and $\Delta \vec{e}_2$ (and diagonal $\Delta \vec{e}_3$) is

$$A_c = 2 \sqrt{g(g - LS_1)(g - LS_2)(g - LS_3)}$$

where

$$g = 1/2 (LS_1 + LS_2 + LS_3)$$

The discriminability mass distribution is then

$$\frac{A_c}{|\vec{\delta}_1| |\vec{\delta}_2|} \left[\frac{\text{Square colors}}{\text{Square count}} \right]$$

which is plotted in figures 7 and 8.

The JND ellipses are plotted in the tasseled cap plane by noting that for the ellipse centered at \vec{K} , the distances to the ellipse in three properly chosen directions define the ellipse. Along the KB axis, the distance is $t_1 = 1/LS_1$; along KG , $t_2 = 1/LS_2$; and in the intermediate direction, $t_3 = 1/LS_3$. Solving for the coefficients of the general centered conic equation

$$Ax^2 + Bxy + Cy^2 + F = 0$$

by means of the determinant equation

$$\begin{vmatrix} X^2 & XY & Y^2 & 1 \\ X_1^2 & X_1 Y_1 & Y_1^2 & 1 \\ X_2^2 & X_2 Y_2 & Y_2^2 & 1 \\ X_3^2 & X_3 Y_3 & Y_3^2 & 1 \end{vmatrix} = 0$$

results in

$$X^2 (LS_1^2) + XY (LS_3^2 - LS_1^2 - LS_2^2) + Y^2 (LS_2^2) - 1 = 0$$

These are the ellipses plotted in figures 9 and 10.

Generation of Uniform Chromaticity Scale Imagery From Landsat Data

R. D. Juday,^a F. Johnson,^a R.A. Abotteen,^b and M. D. Pore^b

INTRODUCTION

This paper documents a method used for generating uniform chromaticity scale (UCS) imagery from Landsat data. A previous study (ref. 1) was made to map multichannel Landsat data into color space using a maximal chromatic expansion method. This study extends the work of reference 1 to the use of a UCS in the form of color film products. (Familiarity of the reader with standard colorimetric nomenclature is assumed; references 2 and 3 are recommended for the novice.)

The motivation behind generating UCS imagery from Landsat data is uniform, controlled perceptibility of color difference caused by differences in Landsat data vectors. One can move a certain distance around any color center before noticing the difference between the color center and the translated point. The distance one can move depends on the direction of the motion and on the location of the color center itself. The generally ellipsoidal surface surrounding the color center is one step in perceptibility from the color center. In the nonlinear transformation to a UCS space, the ellipsoids become spheres with the same radius at all color centers. In this circumstance, the color difference is simply the length of the straight line connecting the points for which color difference is desired.

An orthogonal mapping from three-dimensional feature space into color space will have the property that visual discriminability of picture elements (pixels) in the resulting imagery will be in direct proportion to the Euclidean distance between the pixels in the feature space. Various methods of dimen-

sionality reduction (i.e., color space is three-dimensional, whereas multispectral and possibly multitemporal scanner data have many dimensions) and alignment of feature space to color space are possible.

One of these transformations was used to generate the UCS imagery from the Landsat data reported herein. A description of this transformation is given in the following section. Formulation of the algorithm used for generating UCS imagery is presented in the third section. Conclusions and recommendations of this study can be found in the final section.

UCS TRANSFORMATION USED IN THIS STUDY

The ($L^*a^*b^*$) UCS transformation (ref. 4) was used in this report. In a UCS space, a Euclidean metric is proportional (approximately) to perceptibility of color difference. (See reference 2 for a description of some color-difference formulas and a more thorough exposition of the subject.) The ($L^*a^*b^*$) transformation was adopted by the Commission Internationale de l'Eclairage (CIE) (the International Illumination Committee) for evaluation; a final selection of a universal UCS approximation has not been made.

Let X , Y , and Z be the CIE tristimulus values; the UCS space is generated from the tristimulus values as follows.

$$L^* = 25 \left(100 \frac{Y}{Y_0} \right)^{\frac{1}{3}} - 16; \quad 1 \leq Y \leq 100 \quad (1)$$

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Houston, Texas.

$$a^* = 500 \left[\left(\frac{X}{X_0} \right)^{\frac{1}{3}} - \left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} \right] \quad (2)$$

$$b^* = 200 \left[\left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} - \left(\frac{Z}{Z_0} \right)^{\frac{1}{3}} \right] \quad (3)$$

where X_0 , Y_0 , and Z_0 define the color of the nominally white object-color stimulus; i.e., the illumination spectrum, taken here as that of the light table illuminating the transparencies. The L^* coordinate may be thought of as the lightness or brightness of a color, a^* as the green/magenta balance, and b^* as the yellow/blue balance.

Slices of the ($L^*a^*b^*$) space (at constant values of a^*) that are accessible to the color "guns" of the production film converter (PFC) were generated. Plots of some of those slices are shown in figure 1 for a^* values of -10 , -5 , 0 , 5 , and 10 . In these plots, an edge marked $B = 0$ indicates that the blue gun count of the PFC is 0 at that edge. An edge marked $R = 1$ indicates that the red gun count of the PFC is 255 at that edge. A point in the enclosed area of any of the plots in figure 1 corresponds to red (R), blue (B), and green (G) PFC gun counts between 0 and 255. A PFC product of the ($L^*a^*b^*$) space, sliced at $a^* = 10$, is shown in figure 2. Formulas used for calculating gun counts in figure 2 can be found in the following section.

Other UCS transformations exist in the literature (refs. 5 and 6). These transformations might be considered in later research. The ($L^*a^*b^*$) space was chosen over those discussed in references 5 and 6 for the following reasons.

1. The ($L^*a^*b^*$) space has been adopted by the CIE for evaluation as an approximation to a true UCS space for the object-color solid.

2. Inversion from the ($L^*a^*b^*$) UCS approximation into the CIE (X, Y, Z) system is mathematically tractable. According to D. L. MacAdam (private communication), the (L, j, g) system (ref. 5) is philosophically preferable to the ($L^*a^*b^*$) system for this application. However, there are some considerations, such as the incorporation of Semmelroth's crispening factor (ref. 7), that create difficulties in using the (L, j, g) system.

FORMULATION OF THE UCS ALGORITHM

In generating UCS imagery, the light-table and film parameters are considered. The transformation from PFC counts to transmission values (wavelength dependent) is regarded as stable. The primary colors considered in this study are red (R), green (G), and blue (B). This section is divided into three parts. The first part deals with calculating the primary tristimulus values for light table and film. The second part deals with fitting a prototype Landsat data structure into the ($L^*a^*b^*$) UCS space. The third part presents the algorithm that generates color gun counts from ($L^*a^*b^*$) values for PFC product along with four UCS images of a LACIE segment.

Colorimetric Description of the Transparency Generation Process

The PFC images a cathode-ray tube (CRT) through color filtration onto color reversal film, which is then developed. The CRT display is controlled by numerical input; for each color, the film density is very nearly linear with respect to input counts. The counts-to-density relationship is carefully maintained in exposing and developing the film, and the relationship is herein regarded as stable.

An approximation is made that the manner of addition of sequential images can be expressed in the form

$$X = \sum_i X_i \quad (4)$$

where X is the first tristimulus value and i is an index for the red, green, and blue PFC primaries. The same approximation is made for Y and Z , assuming that the film system produces the same final result as would be obtained by separation images projectively added. The approach avoids certain practical problems such as interimage effects and reciprocity failure (ref. 8). The complete colorimetric model of the PFC is given in an unpublished thesis.¹ For the

¹R. D. Juday, "Colorimetric Principles as Applied to Multichannel Imagery." Master's thesis, to be published.

purposes of the present paper, assume that a known one-to-one relationship exists between the $(L^*a^*b^*)$ space and the set of three input commands for the PFC.

Fitting Landsat Data Space Into the UCS Space

Two multitemporal LACIE segments were used to generate scatter plots. The Landsat data in each unitemporal segment were rotated using the tasseled cap transformation (ref. 9) with a bias. The brightness, greenness, and yellowness biases were taken to be 0, 30, and 53 counts, respectively. The scatter plots of Landsat data in the (K_b, K_g) space for segments 1618 and 1645, respectively, are shown in figures 3 and 4. Figures 5 and 6 are scatter plots of the data in the (K_b, K_r) space. The scatter plots in figures 3 and 4 were used to produce an overall Landsat (K_b, K_g) data space scatter envelope. This data space scatter envelope contains the raw K_b and K_g counts; i.e., the bias was removed and was fitted into the (L^*b^*) UCS space as shown in figure 7. The UCS space, sliced at $a^* = 10$ (fig. 2), was used because it had the best fit to the envelope.

According to color theory, there is no preferential orientation between data space and the UCS space. For subjective reasons, however, the brightness of the data space was aligned with L^* , the lightness direction. Points on the ground that would appear lighter in conventional aerial photographs will tend to appear lighter here also. The portion of color space accessible to the PFC has a greater extent in (L^*b^*) than in (L^*a^*) , and the first two tasseled cap components have the greatest variance. Thus, the first two components were laid into (L^*b^*) .

The component K_r is taken parallel to a^* to complete the orthogonal relationship between the UCS space and the first three tasseled cap components. The two points p_1 and p_2 shown in figure 7 were used to calculate the transformation from Landsat (K_b, K_g, K_r) data space into $(L^*a^*b^*)$:

$$u = Fe + \delta \quad (5)$$

where u = a column vector of L^* , b^* , and a^*
 e = $[K_r, K_b, K_g]^T$, and e' is the transpose of e
 δ = $[A_3, A_6, A_7]^T$

Using the numerical values of p_1 and p_2 of figure 7, one computes F and δ to be

$$F = \begin{pmatrix} -0.3012 & 1.0267 & 0 \\ 1.0267 & .3012 & 0 \\ 0 & 0 & 1.0670 \end{pmatrix} \quad (6)$$

and

$$\delta = \begin{pmatrix} -11.594 \\ -32.288 \\ 18.560 \end{pmatrix} \quad (7)$$

Algorithm That Generates Color Gun Counts From $(L^*a^*b^*)$ Pixel Values for PFC Product

It was shown in the preceding section that given the e vector for a pixel in a Landsat scene, the L^* , b^* , and a^* values for that pixel were computed according to equation (5). The tristimulus values X , Y , and Z for a pixel are computed from the pixel's L^* , b^* , and a^* values by inverting equations (1) to (3) as follows.

$$Y = Y_0 \frac{(L^* + 16)^3}{1562500} \quad (8)$$

$$X = X_0 \left[\frac{a^*}{500} + \left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} \right]^3 \quad (9)$$

$$Z = Z_0 \left[\left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} - \frac{b^*}{200} \right]^3 \quad (10)$$

The colorimetric model of the PFC was then employed to produce the UCS color film products shown in figure 8 for LACIE segment 1618. Figure 8 shows four images of segment 1618 corresponding to four multitemporal acquisitions and covering all the

biological stages of small grains. Figures 3 and 4, with figure 7, provide an interpretive key for the images in figure 8. Small-grains fields in biostage 1 appear to have a blue color (low greenness value K_{g_1}), where they change from light yellow orange ($K_{g_2} > K_{g_1}$) to a darker yellow orange ($K_{g_3} > K_{g_2}$) in going from biostage 2 to 3. In biostage 4 ($K_{g_4} < K_{g_3}$), the small-grains fields appear to have a pinkish color. This is in agreement with how the data were fit into the UCS space as shown in figure 7. The pinkish cast shows an increase in the yellow component; K_y is scaled toward positive a° , which is in the general direction of the PFC red gun.

CONCLUSIONS AND RECOMMENDATIONS

An algorithm for generating uniform chromaticity scale imagery from Landsat data has been presented. A computer program was written to implement the algorithm, and UCS film products were generated. The colors in the film and their temporal change are consistent with those expected for the particular scaling of Kauth components into the ($L^*a^*b^*$) color space. The UCS film product has not been subjected to the practical test of competing with previous transformations. In that competition (to be done outside the purview of this report), the philosophically satisfying notion of transforming Landsat data so that a one-count difference is equally perceptible at all locations in data space will be tested.

The authors recommend that analyst-interpreters test the UCS imagery using a variety of LACIE segments. Preliminary examination indicates that the UCS product offers the following possibilities.

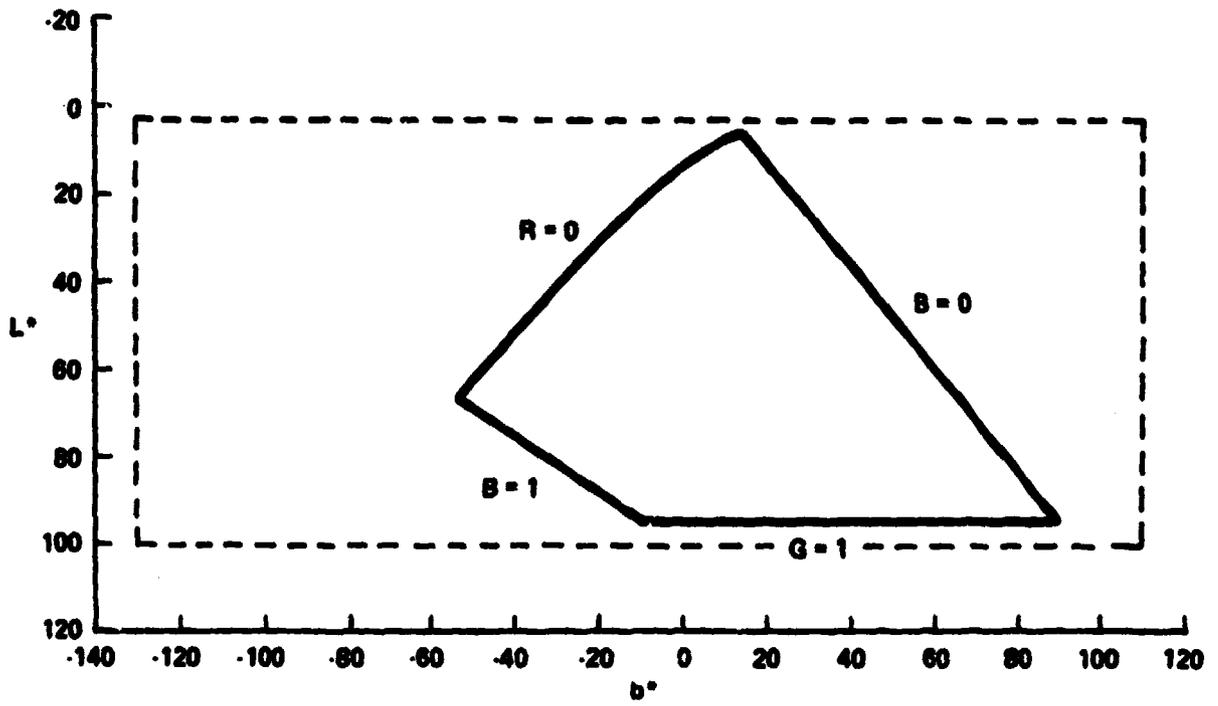
1. A single film product that will supplant two film products in current use

2. Improved visibility of data differences in regions in data space that are critical to crop identification

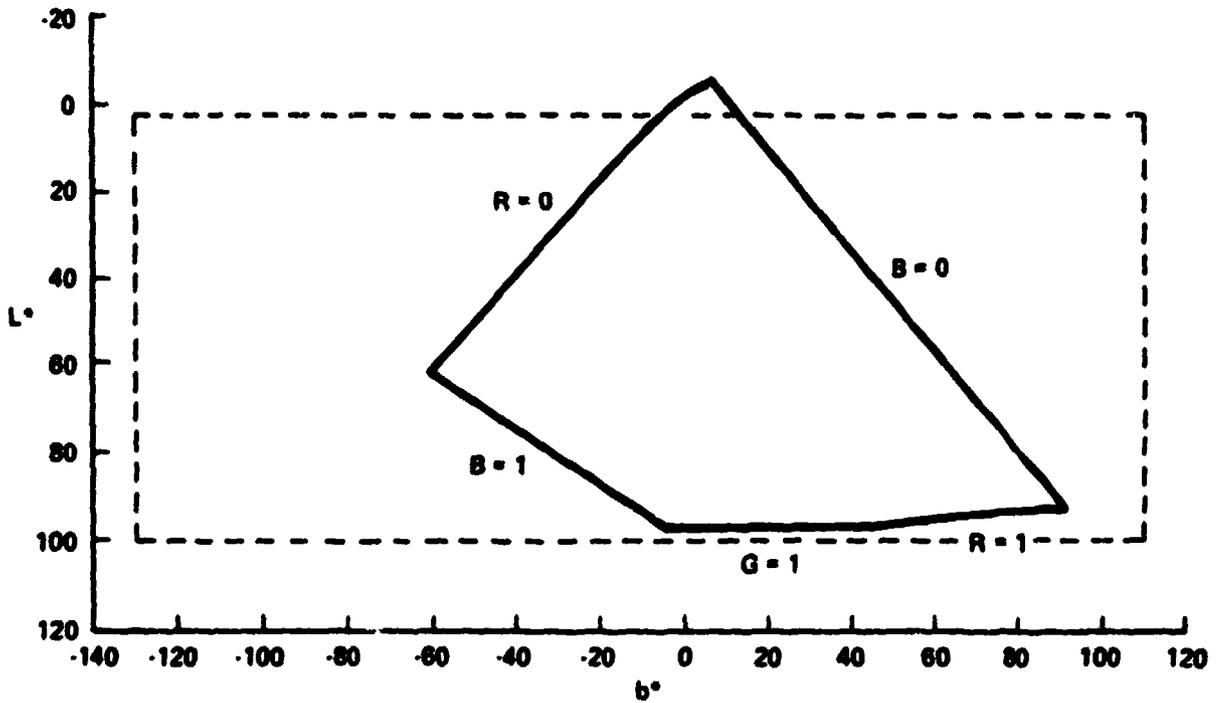
3. An analytic route to the determination of data-space transformations that will be optimal for particular discrimination problems—for example, in another project, the transformation has been used to display water bodies in Landsat data, with encouraging results.

REFERENCES

1. Juday, R. D.; and Abooteen, R. A.: A Maximal Chromatic Expansion Method of Mapping Multichannel Imagery Into Color Space. LEC-10830, Lockheed Electronics Co., Houston, Tex., Jan. 1978.
2. Wyszecki, G.; and Stiles, W. S.: Color Science. John Wiley and Sons, 1967.
3. Nimeroff, I.: Colorimetry. NBS Monograph 104, Jan. 1968.
4. Proposal for Study of Uniform Color Spaces and Color-Difference Evaluations. CIE Compte Rendu 18^e Session. CIE Publication no. 36, 1976, p. 171.
5. MacAdam, D. L.: Uniform Color Scales. JOSA 64, 1691, Dec. 1974.
6. MacAdam, D. L.: Geodesic Chromaticity Diagram. Die Farbe 18, no. 77, 1969.
7. Semmleith, Carl C.: Prediction of Lightness and Brightness on Different Backgrounds. JOSA 60, 1685, 1970.
8. Eynard, R. A., ed.: Color: Theory and Imaging Systems. Society of Photographic Scientists and Engineers (Washington, D.C.), 1973.
9. Kauth, R. J.; and Thomas, G. S.: The Tasseled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue Univ. (W. Lafayette, Ind.), June 29, 1976.

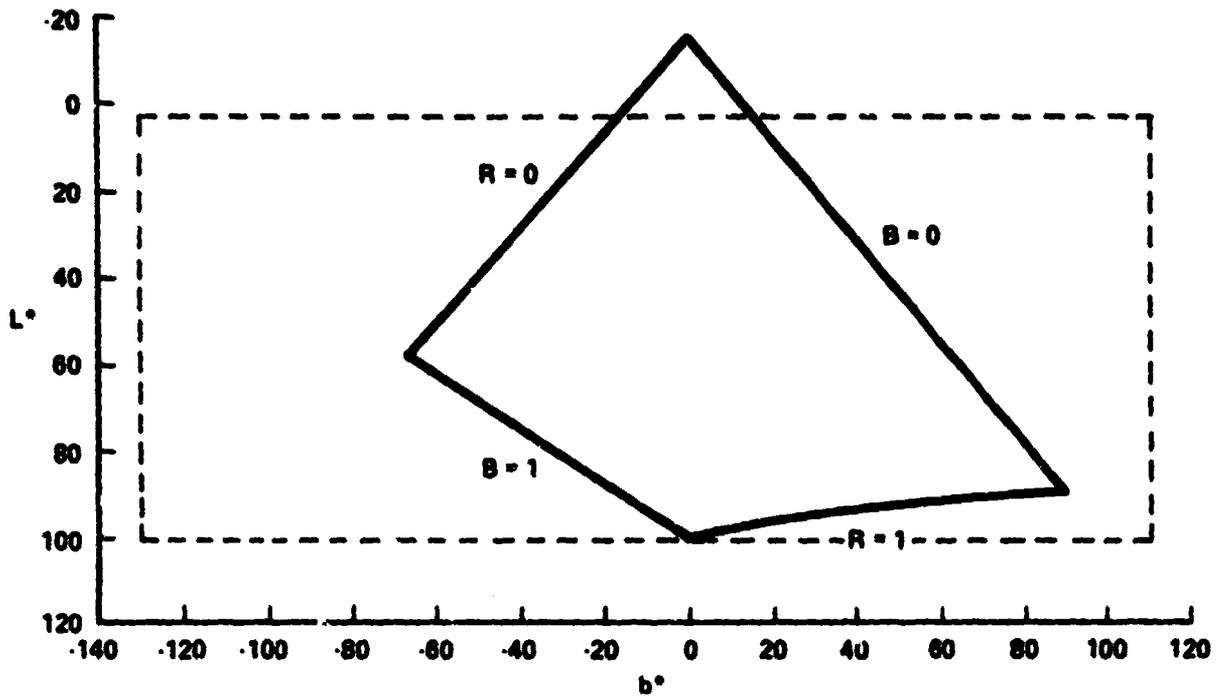


(a)

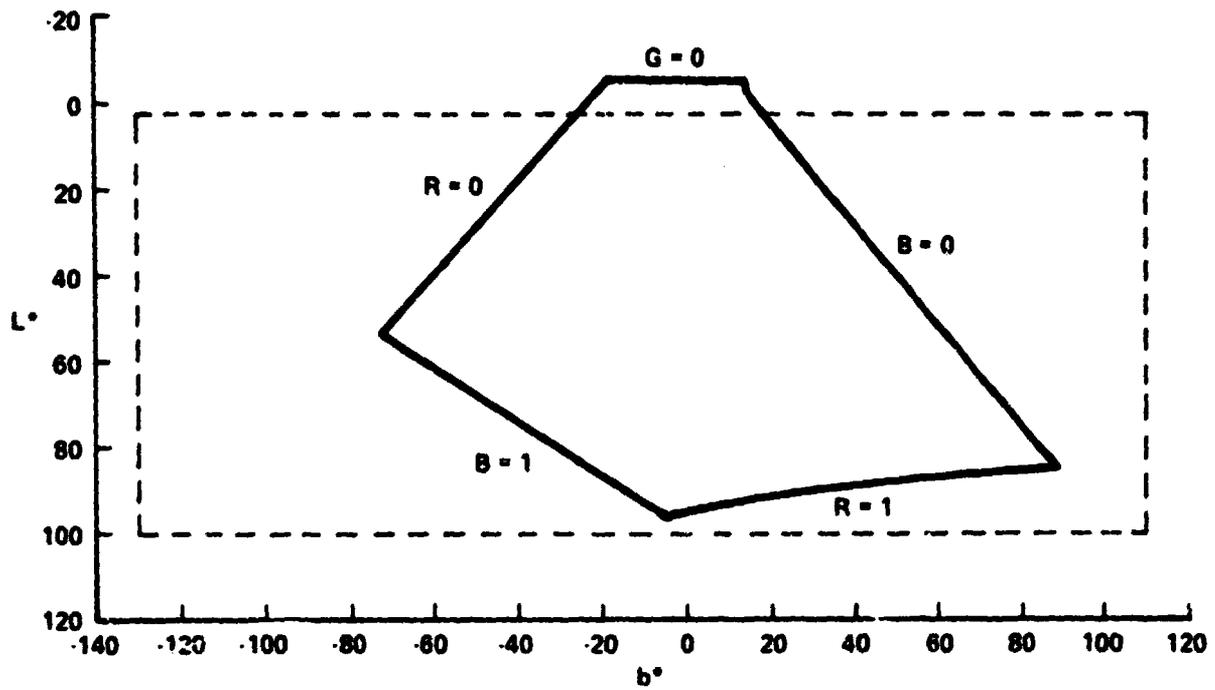


(b)

FIGURE 1.—Slices of the $(L^*a^*b^*)$ space at constant values of a^* that are accessible to the color gun cube of the PFC. (a) $a^* = -10.00$. (b) $a^* = -5.00$. (c) $a^* = 0.00$. (d) $a^* = 5.00$. (e) $a^* = 10.00$.

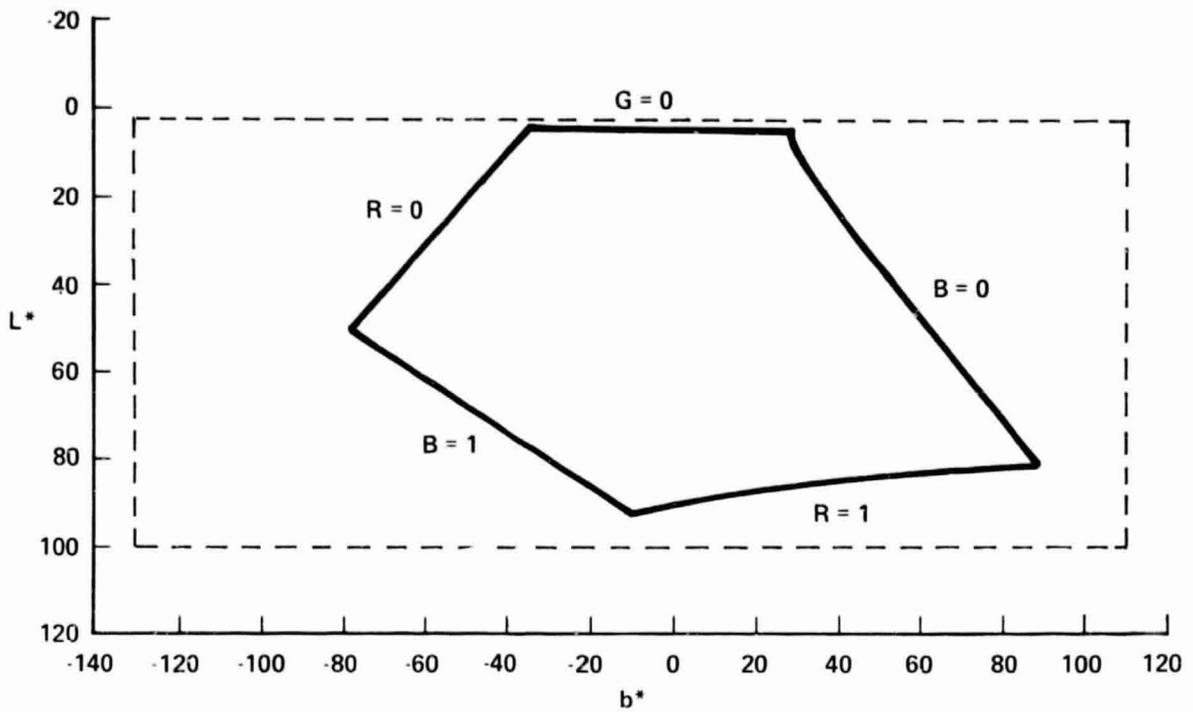


(c)



(d)

FIGURE 1.—Continued.



(e)

FIGURE 1.—Concluded.

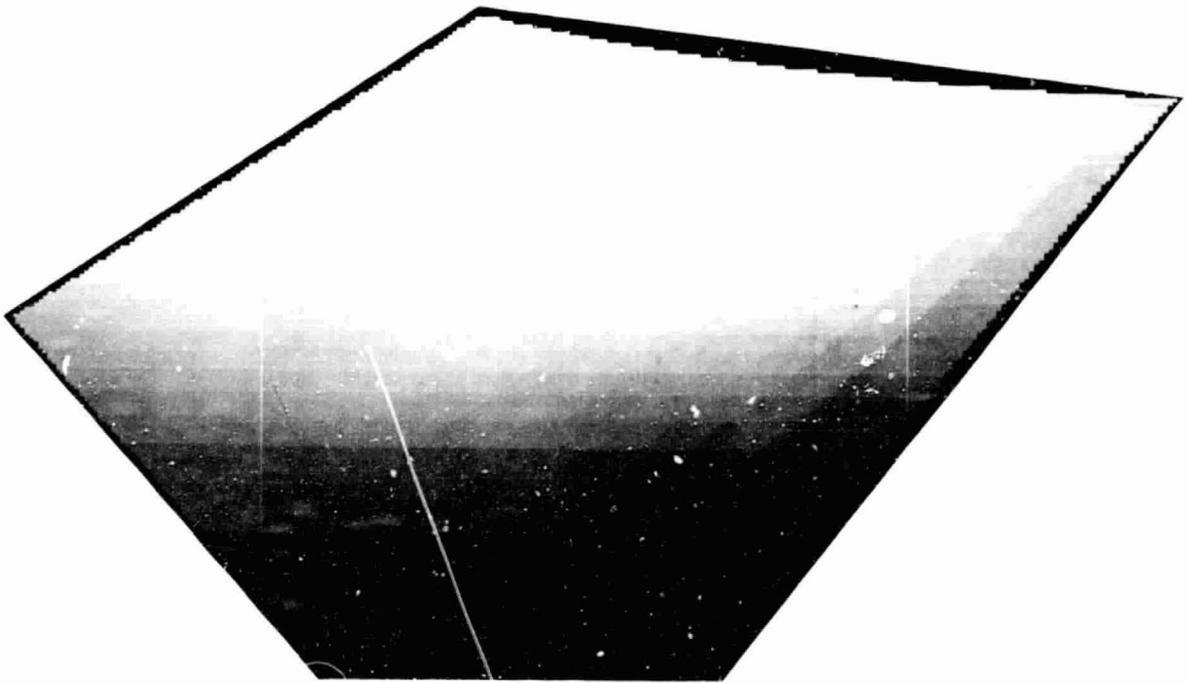
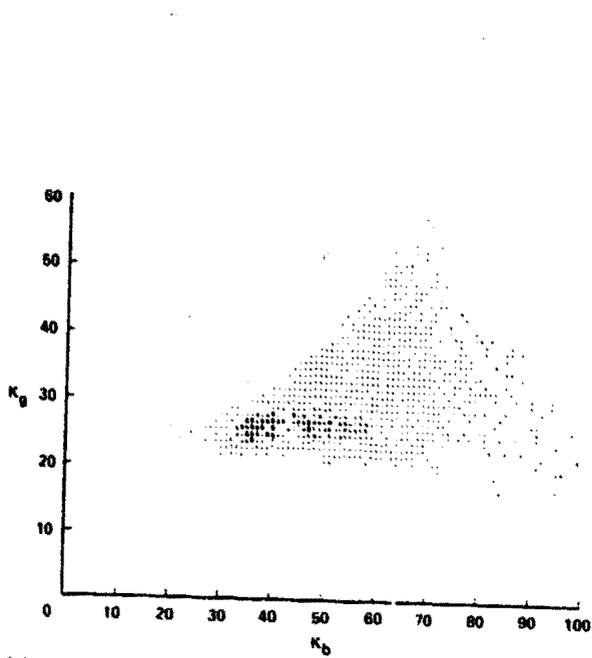
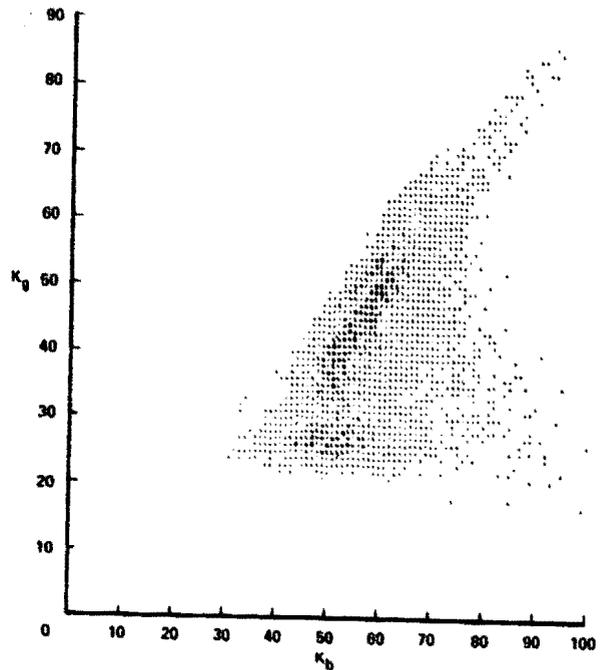


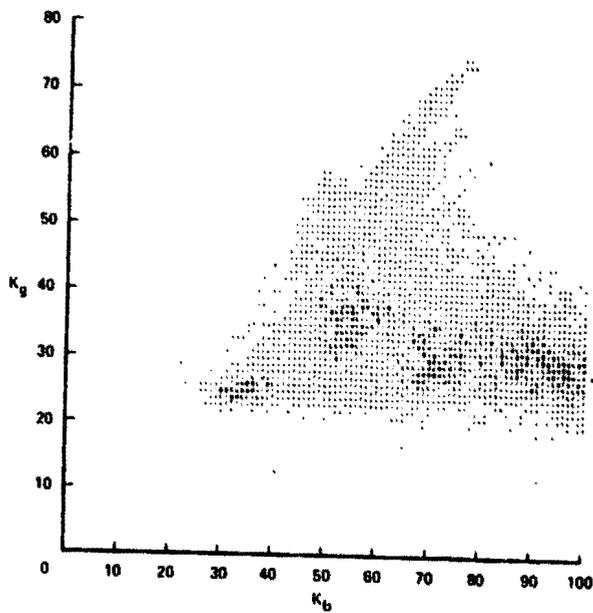
FIGURE 2.—A PFC product of the $(L^*a^*b^*)$ space sliced at $a^* = 10$.



(a)

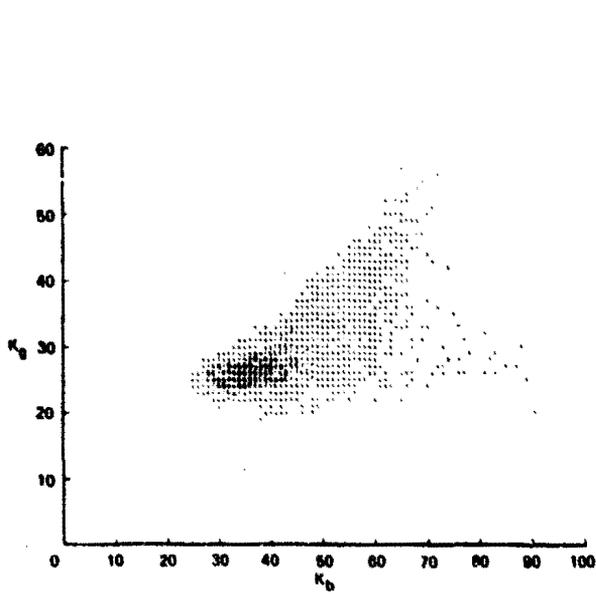


(b)

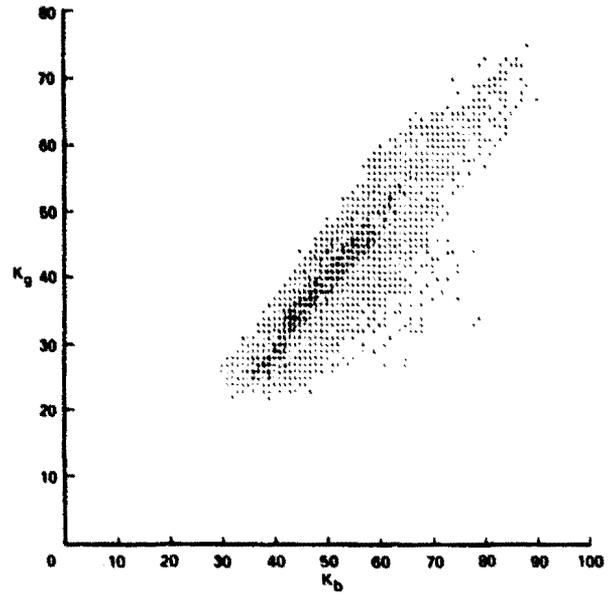


(c)

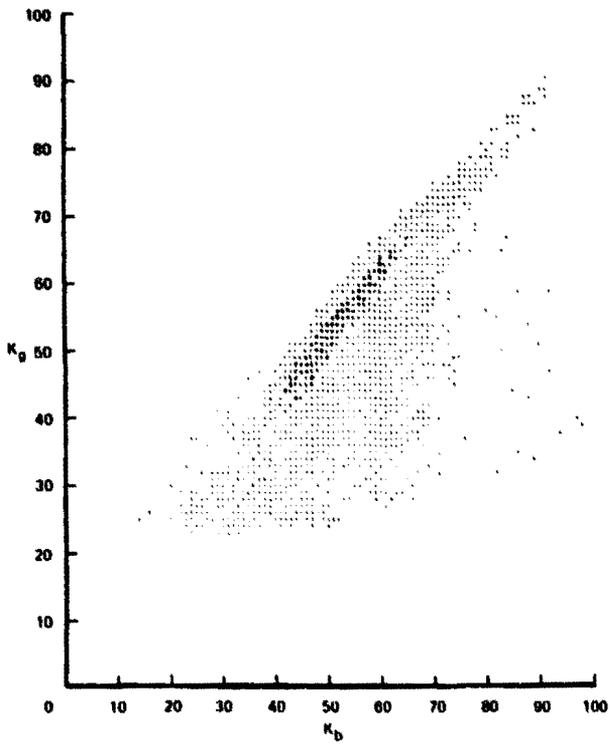
FIGURE 3.—Scatter plots of Landsat data in the (K_b, K_g) space for LACIE segment 1618 at various acquisition dates. (a) May 6, 1976. (b) June 11, 1976. (c) August 22, 1976.



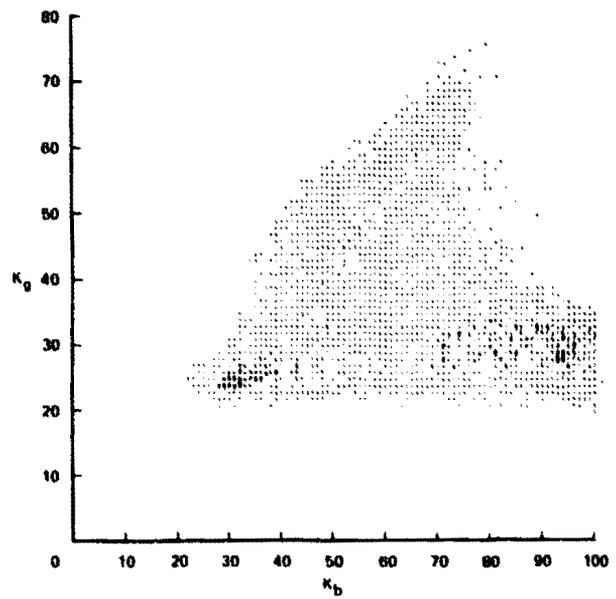
(a)



(b)

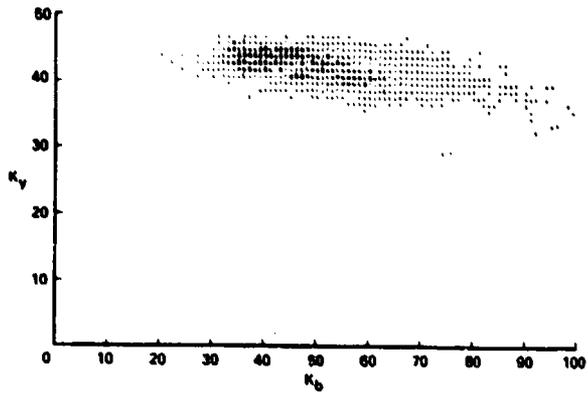


(c)

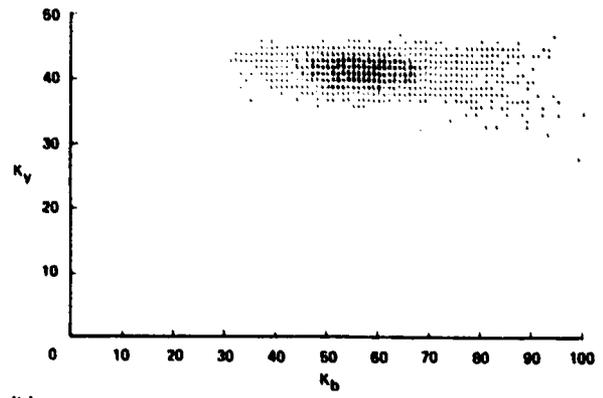


(d)

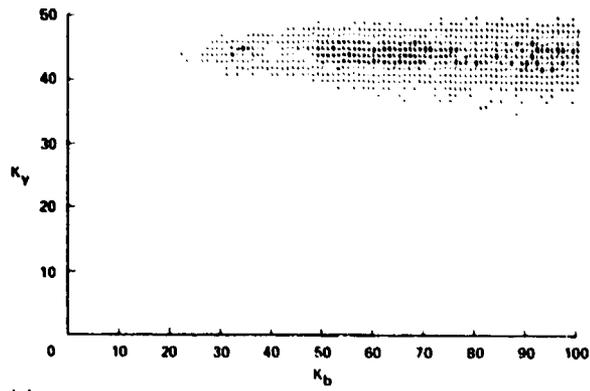
FIGURE 4.—Scatter plots of Landsat data in the (K_b, K_g) space for LACIE segment 1645 at various acquisition dates. (a) May 6, 1973. (b) June 12, 1976. (c) June 29, 1976. (d) August 22, 1976.



(a)

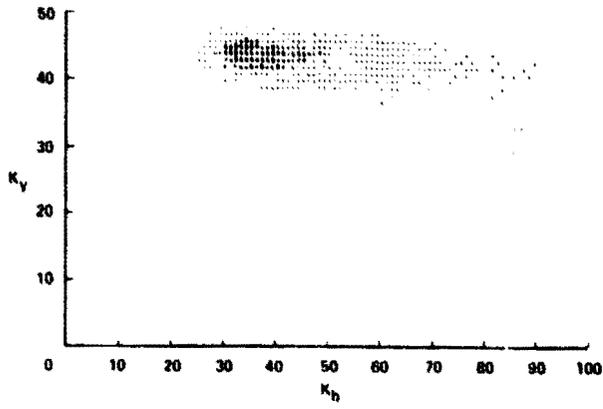


(b)

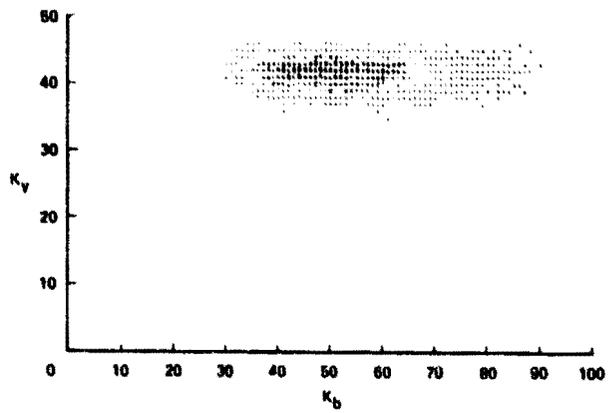


(c)

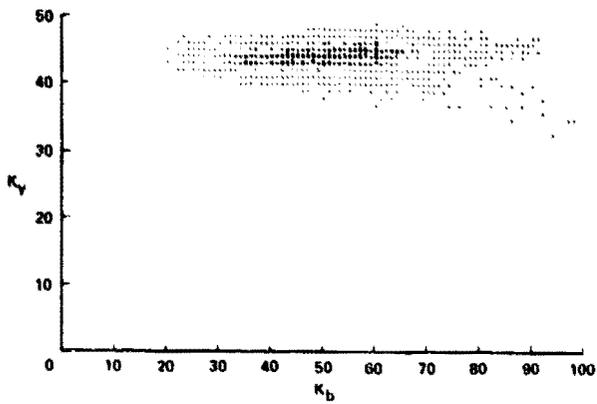
FIGURE 5.—Scatter plots of Landsat data in the (K_b, K_v) space for LACIE segment 1618 at various acquisition dates. (a) May 6, 1976. (b) June 11, 1976. (c) August 22, 1976.



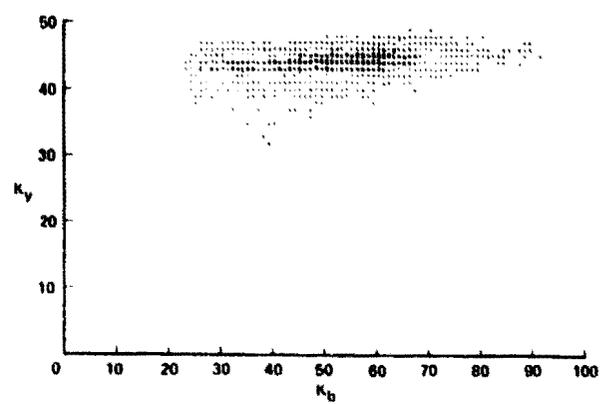
(a)



(b)



(c)



(d)

FIGURE 6.—Scatter plots of Landsat data in the (K_b, K_v) space for LACIE segment 1645 at various acquisition dates. (a) May 6, 1976. (b) June 12, 1976. (c) June 29, 1976. (d) August 22, 1976.

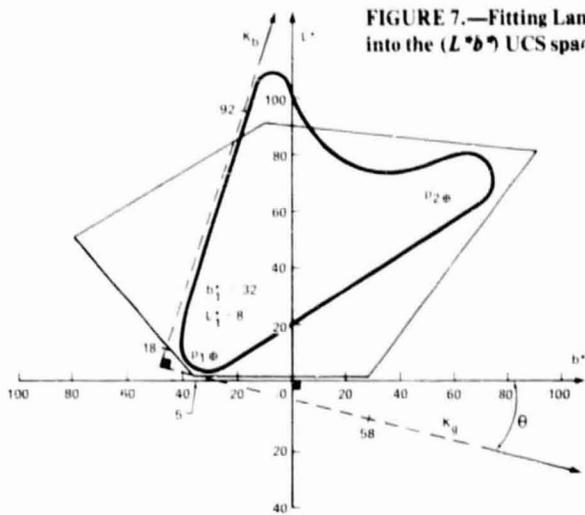


FIGURE 7.—Fitting Landsat (K_p, K_r) data-space scatter envelope into the (L^*b^*) UCS space sliced at $a^* = 10$.

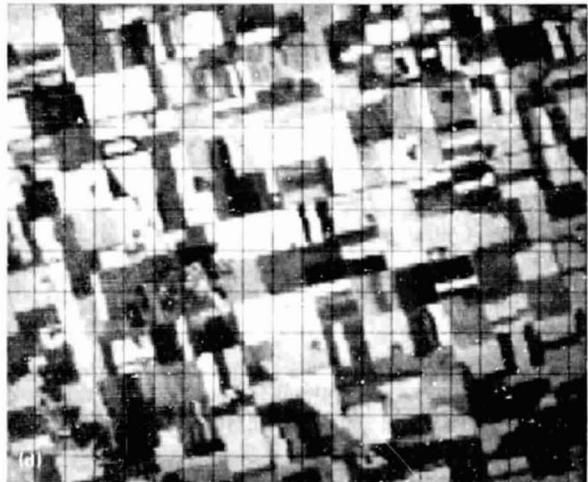
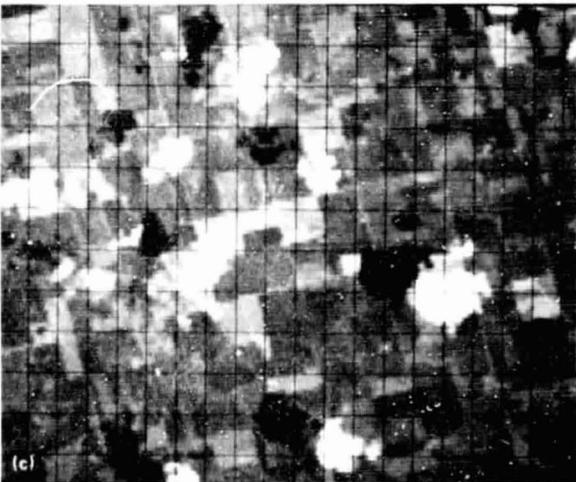
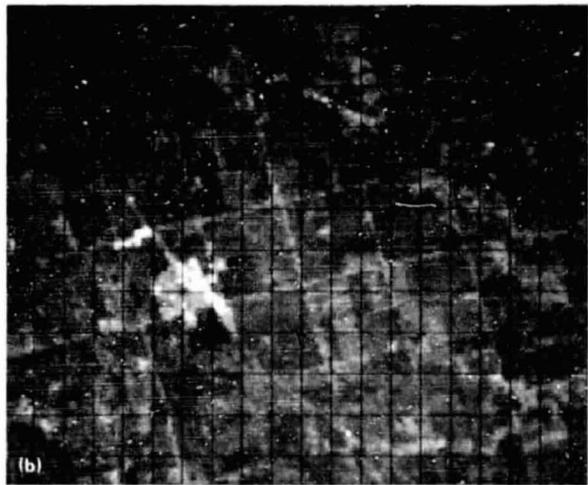
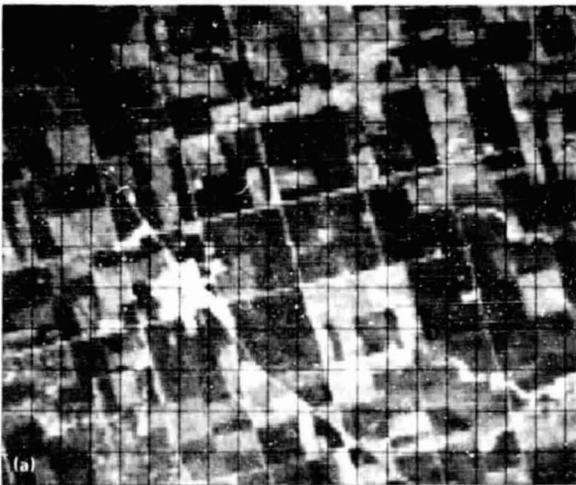


FIGURE 8.—UCS film product of LACIE segment 1618 at various acquisition dates. (a) May 6, 1976. (b) June 17, 1976. (c) July 17, 1976. (d) August 22, 1976.

Image and Numerical Display Aids for Manual Interpretation

R. A. Abotteen^a

SUMMARY

In the Large Area Crop Inventory Experiment, image and numerical display aids for manual interpretation were produced to assist in selecting and/or identifying representative samples of signatures in a given Landsat scene. Four methods for producing numerical display aids were developed and are discussed in this paper. The four methods employed are clustering techniques, data compression, phenological growth pattern extraction, and aggregation of like spectral data on a two-dimensional spectral plot.

INTRODUCTION

Image interpretation is an important method for acquiring training data for classification of Landsat images in the LACIE (ref. 1). Interpreting a scene for classification requires that training samples of all spectral signatures in the given scene be selected and correctly labeled. This procedure becomes especially difficult when multiple passes over a scene are to be interpreted. The variation of the spectral signatures, in a multitemporal sense, makes it difficult to select and identify all the various signatures in a scene. To address these problems, four types of image and numerical display aids were developed.

The first and the second image interpretation aids were obtained by applying nonsupervised pattern recognition techniques (clustering) and data compression. The clustering method identifies the inherent classes in the scene. Color film is generated from the cluster image, with each cluster having a distinct color; the color corresponds to the value of the cluster mean (ref. 2). In interpreting multipass data, a principal component (PCOMP) transforma-

tion is applied to the cluster image. By this means, the multitemporal spectral variation of the scene is compressed, or summarized, into a three-dimensional image which can be displayed as a color image. To track phenological growth patterns of various crops, the principal component greenness (PCG) transformation was introduced as a third aid. The PCG transformation maps each of n multitemporal acquisitions onto the greenness axis and then compresses these n greenness channels into three new channels using a linear combination (i.e., the first three principal components) of the n green channels.

To enable the analyst to view the structure of the Landsat data in spectral space, a two-dimensional spectral plot of the data was developed as the fourth aid. The spectral plot (ref. 2) takes advantage of the inherent two-dimensionality of Landsat data (ref. 3). The plots are constructed to assist in relating picture elements (pixels) in the scene to their locations on the spectral plot.

CLUSTER IMAGE

A cluster image is generated first by clustering the data in the scene and then by replacing each data sample with the mean of the cluster to which it belongs. A simulated color-infrared (CIR) film of the cluster image can be generated by a production film converter (PFC). One of the main features of the cluster image is that two spectrally similar clusters are shown by the film product to have similar colors. This feature could easily be lost if an arbitrary color assignment were used to generate a film product from the one-dimensional cluster map, as shown in figure 1(a). A CIR film product of the cluster image is normally generated using the same gain and bias as that used to produce the original CIR image. This technique results in a CIR film product of the cluster image that resembles the standard CIR

^a Lockheed Electronics Company, Houston, Texas.

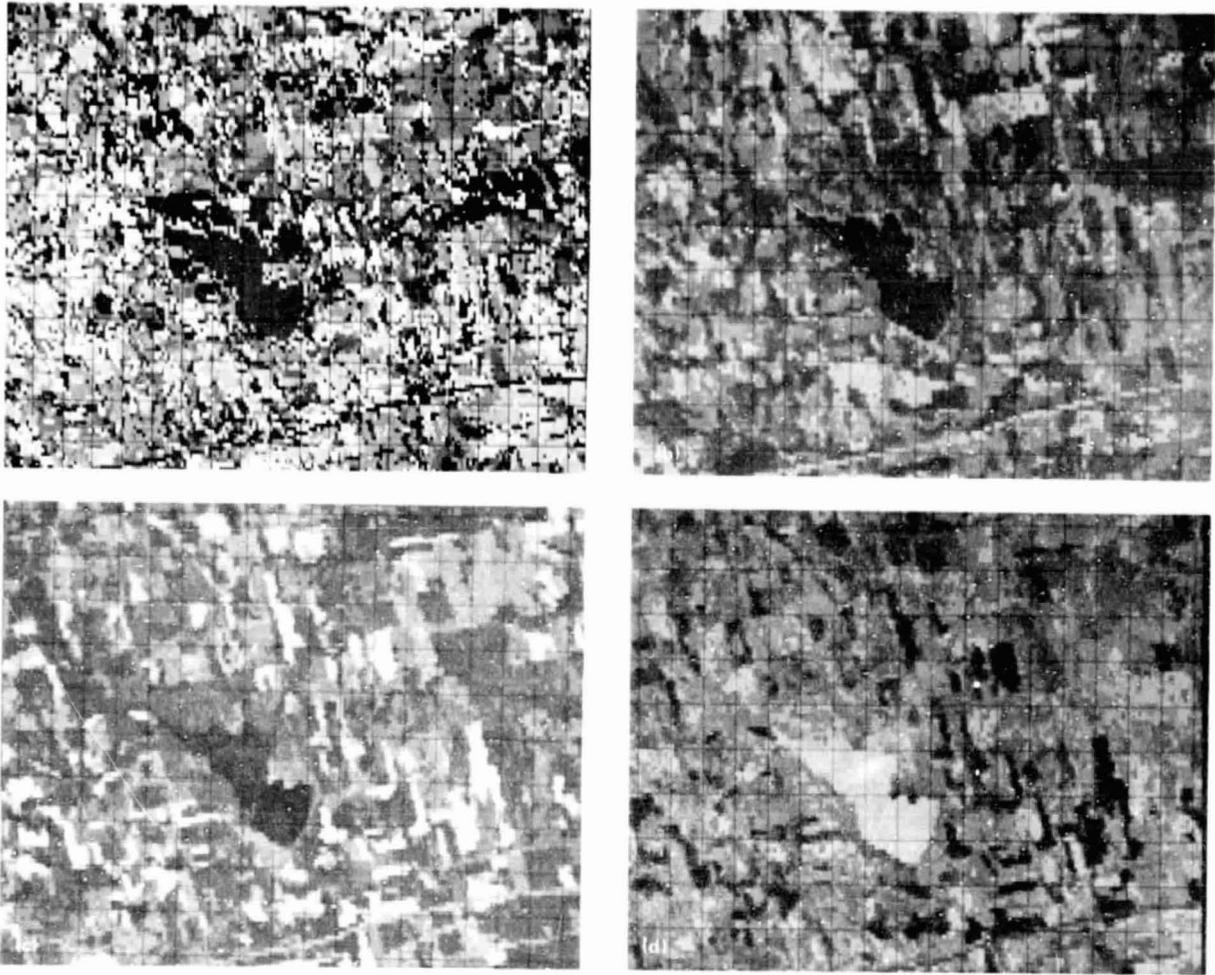


FIGURE 1.—Eight-channel cluster images of segment 1576, Lancaster County, Nebraska. (a) Cluster map; arbitrary color assignment. (b) CIR cluster image; channels 1, 2, and 4. (c) CIR cluster image; channels 5, 6, and 8. (d) PCOMP cluster image; first three principal components.

film product, as shown in figures 1(b) and 1(c).

A color key to the clusters is generated by assigning a square of 100 samples (pixels) to each cluster. Each pixel in the square is then assigned the value of the cluster mean it represents. The color keys are then ordered according to the Kauth greenness number (ref. 3).

It was discovered from observing cluster images on CIR film that they can be used as aids in defining spectral classes and thus in standardizing the image interpretation procedure. In addition, an increase in the contrast of adjacent fields is apparent, which assists in the delineation of training fields.

PCOMP CLUSTER IMAGE

A PCOMP cluster image of a scene is generated by applying principal component transformation to the cluster image described in the preceding section. Figure 1(d) is an example of a PCOMP cluster image. Ready and Wintz (ref. 4) have shown that the PCOMP transformation applied to aircraft- and satellite-gathered multispectral data is very useful for information extraction, since the first few principal components contain essentially all the information present in the original spectral bands. Additional analyses of PCOMP-transformed Landsat data are

available (ref. 5).

The PCOMP transformation is

$$Y = MX \quad (1)$$

where X = a vector of n spectral intensities associated with each pixel

M = an n by n unitary matrix derived from the mixture covariance matrix Σ_X of the spectral bands such that the rows of M are the normalized eigenvectors of Σ_X

Y = a vector of n PCOMP's

The covariance matrix of the PCOMP-transformed data then becomes

$$\Sigma_Y = M\Sigma_X M^T = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \dots & \\ 0 & & & \lambda_n \end{bmatrix} \quad (2)$$

where M^T is the transpose of M and $\lambda_1, \lambda_2, \dots, \lambda_n$ (the variances of the PCOMP's) are the eigenvalues of Σ_X ordered so that $\lambda_1 > \lambda_2 > \dots > \lambda_n$.

Since M is a unitary matrix, the PCOMP transformation preserves the total data variance; i.e.,

$$\sum_{i=1}^n \sigma_{X_i}^2 = \sum_{i=1}^n \lambda_i \quad (3)$$

where the values for $\sigma_{X_i}^2$ are the variances of the original spectral bands. Note that in the PCOMP

transformation, most of the data variance is concentrated in the first few PCOMP's. It was observed that, for Landsat data in four channels, $\lambda_1 + \lambda_2$ contained approximately 96 percent of the total data variance. In eight channels of Landsat data, $\lambda_1 + \lambda_2 + \lambda_3$ contained approximately 91 percent of the total data variance. For 16-channel Landsat data, $\lambda_1 + \lambda_2 + \lambda_3$ contained approximately 82 percent of the total data variance.

The interest in the first three PCOMP's of Landsat data arises from processing color film products. A PFC generating a color product of Landsat data uses three channels, with each channel assigned to the blue, green, or red film converter gun. The first three PCOMP's are associated with these three channels to produce the PCOMP cluster image. Since the first three PCOMP's contain most of the data variance, the PCOMP cluster image seems to be a good means of compressing multitemporal data.

Once the PCOMP transformation has been applied to both the cluster image and the color keys, the transformed data are rescaled to lie between 0 and 255 to allow storage of the image in a standard computer format.

PRINCIPAL COMPONENT GREENNESS IMAGE

In developing the PCG image, it is assumed that a temporal change in the Kauth greenness (ref. 3) of an agricultural crop is an indicator of a phenological change in that crop. The PCG transformation is formulated as follows. Let n be the number of multiple, registered Landsat passes over a segment. Let Z be an N -dimensional feature vector drawn from the segment image, where $N = 4n$ and the segment image is composed of passes 1, 2, ..., n . The feature vector Z is mapped onto the Kauth greenness feature vector G using the following transformation:

$$\begin{bmatrix} K_1 \\ K_2 \\ K_3 \\ \vdots \\ K_n \end{bmatrix} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & v_1 & v_2 & v_3 & v_4 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & v_1 & v_2 & v_3 & v_4 & \dots & 0 & 0 & 0 & 0 & 0 \\ \vdots & & & & & & & & & & & & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots & v_1 & v_2 & v_3 & v_4 \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \\ \vdots \\ Z_N \end{bmatrix}$$

or $G = AZ$ (4)

where $v_1, v_2, v_3,$ and v_4 are entries from the second row of the Kauth transformation matrix. Note that $A = [a_{ij}]$ is an n by N matrix such that for $i = 1$

$$\left. \begin{aligned} a_{ik} &= v_k \text{ when } k = 1, 2, 3, 4 \\ a_{ik} &= 0 \text{ when } k = 5, 6, \dots, N \end{aligned} \right\} \quad (5)$$

For $i = 2, 3, \dots, n$, let $j = 4(i - 1)$; then

$$\left. \begin{aligned} a_{ik} &= 0 \text{ when } k = 1, 2, \dots, N \text{ and } k \neq j + m; m = 1, 2, 3, 4 \\ a_{i,(j+m)} &= v_m; m = 1, 2, 3, 4 \end{aligned} \right\} \quad (6)$$

For Landsat-1

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} -0.290 \\ -0.562 \\ 0.600 \\ 0.491 \end{bmatrix} \quad (7)$$

and for Landsat-2

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} -0.2832 \\ -0.6601 \\ 0.5774 \\ 0.3883 \end{bmatrix} \quad (8)$$

The PCG transformation is

$$P = UG \quad (9)$$

where G = a vector of n greenness values associated with each pixel and corresponding to the n passes as shown in equation (4)

U = an n by n unitary matrix derived from the mixture covariance matrix Σ_G of the greenness bands such that the

rows of M are the normalized eigenvectors of Σ_G

P = a vector of n PCG bands

The covariance matrix of the PCG-transformed data then becomes

$$\Sigma_P = U \Sigma_G U^T = \begin{bmatrix} \lambda_{g_1} & & & 0 \\ & \lambda_{g_2} & & \\ & & \dots & \\ 0 & & & \lambda_{g_n} \end{bmatrix} \quad (10)$$

where U^T is the transpose of U and $\lambda_{g_1}, \lambda_{g_2}, \dots, \lambda_{g_n}$ (the variances of the PCG bands) are the eigenvalues of Σ_G ordered such that $\lambda_{g_1} > \lambda_{g_2} > \dots > \lambda_{g_n}$.

The PCG transformation was applied to 10 four-pass small-grains-growing LACIE segments in the United States. The eigenvalues and eigenvector components of Σ_G for the 10 segments are shown in table I. Table I also contains the percentage of the data greenness variability explained by the first two PCG bands. For most of the segments in table I, note that λ_{g_1} and λ_{g_2} together contain more than 85 percent of the data greenness variability.

A detailed graphical analysis was performed on segment 1988, located in Finney County, Kansas. The four Landsat passes over the segment were acquired on February 7, April 18, May 6, and June 12, 1976, which correspond to winter wheat biowindows 1, 2, 2, and 3, respectively. The four greenness bands were obtained using equations (4) and (8). To maintain the nonnegativity of the greenness bands, a four-dimensional bias vector was added to equation (4). All the components of this bias vector were taken to be 16 counts. A plot of the eigenvector components versus greenness bands 1, 2, 3, and 4 is shown in figure 2. Figure 3 shows the temporal trajectories of training field greenness means of the segment. It is noteworthy that the temporal trajectory of the small-grains field mean in figure 3 resembles the first eigenvector plot in figure 2, whereas the temporal trajectory of the non-small-grains field mean resembles the second eigenvector. This correlation suggests that, for segment number 1988, the components of the first eigenvector define a small-grains trajectory and the components of the second eigenvector

TABLE I.—Eigenvalues and Eigenvectors of Greenness Image Mixture Covariance Matrices for 10 Four-Pass Acquisitions of Small-Grains-Growing LACIE Segments in the United States

Segment	Eigenvalues ($\lambda_{g_1}, \lambda_{g_2}, \lambda_{g_3}, \lambda_{g_4}$) ^T	First eigenvector	Second eigenvector	Third eigenvector	Fourth eigenvector	$\frac{(\lambda_{g_1} + \lambda_{g_2})}{\sum_{i=1}^4 \lambda_{g_i}} \times 100$, percent
1988	170.51	0.063	-0.018	-0.120	0.991	92.27
	22.34	.505	-.214	-.825	-.136	
	14.10	.803	-.230	.549	.012	
	2.06	.309	.949	-.055	-.009	
1046	29.14	-0.010	0.061	-0.307	-0.950	87.08
	23.95	.283	.059	-.911	.295	
	5.84	.771	.579	.259	-.055	
	2.03	-.570	.811	-.096	.089	
1233	85.31	0.145	-0.379	-0.418	-0.813	85.71
	39.28	.713	-.408	.570	.024	
	11.85	.562	.032	-.698	.444	
	8.93	.394	.83	.116	-.377	
1506	159.23	0.048	-0.023	-0.056	-0.997	85.93
	83.26	.965	.054	.255	.031	
	33.89	.239	.181	-.952	.060	
	5.81	-.096	.982	.160	-.036	
1851	115.45	0.071	0.014	0.090	0.993	87.48
	96.86	.793	.210	.561	-.110	
	27.03	.530	.208	-.821	.034	
	3.36	-.291	.955	.054	.002	
1967	300.42	0.012	0.987	0.051	-0.151	88.59
	72.55	.801	.025	-.598	.026	
	35.34	.591	-.076	.780	-.189	
	12.69	.095	.138	.156	.970	
1538	154.40	0.074	-0.006	0.622	0.780	88.62
	39.13	.906	-.407	.030	-.113	
	15.31	.375	.698	-.492	.362	
	9.53	.182	.589	.609	-.499	
1618	240.59	0.019	-0.023	-0.194	0.981	81.47
	100.61	.293	-.847	-.431	-.110	
	70.23	.956	.264	.128	.013	
	7.38	-.008	.462	-.872	-.162	
1655	98.41	0.381	0.087	-0.911	0.135	79.39
	42.02	.877	-.345	.335	.011	
	28.77	.244	.864	.239	.370	
	7.69	.164	.356	-.033	-.919	
1694	78.02	0.325	0.933	0.117	0.096	88.85
	13.26	.799	-.198	-.559	-.105	
	7.63	.458	-.226	.796	-.326	
	3.83	.216	-.197	.203	.935	

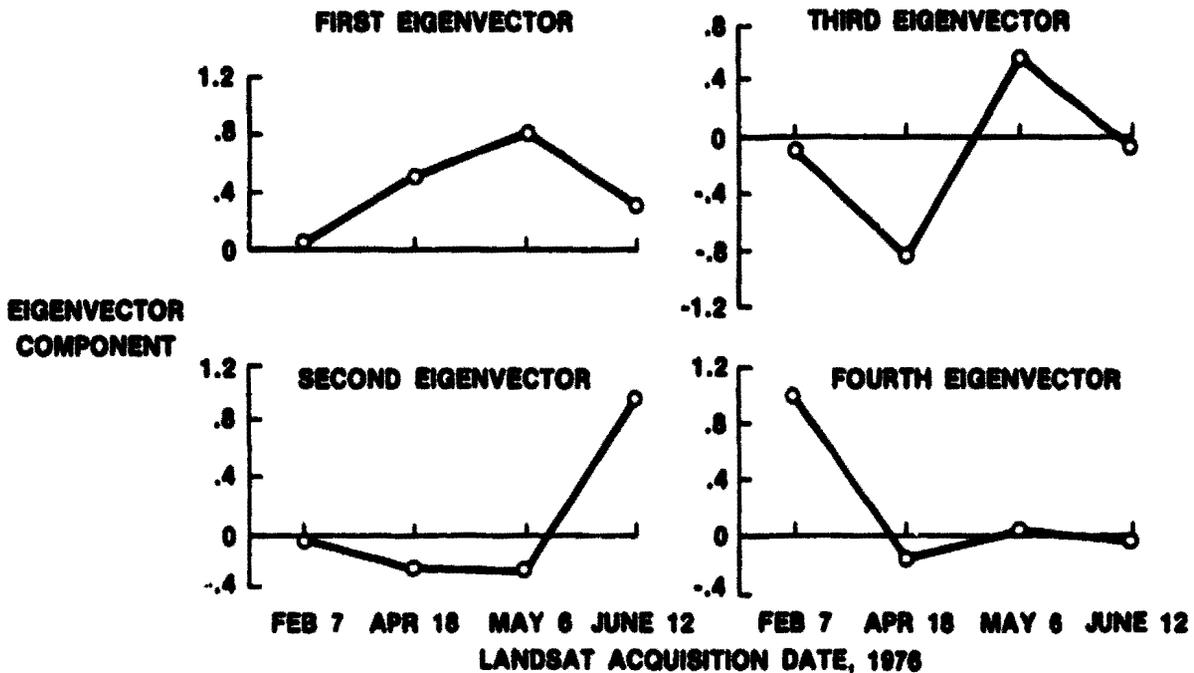


FIGURE 2.—Plots of eigenvector components versus greenness bands 1, 2, 3, and 4 in LACIE segment 1988, Finney County, Kansas.

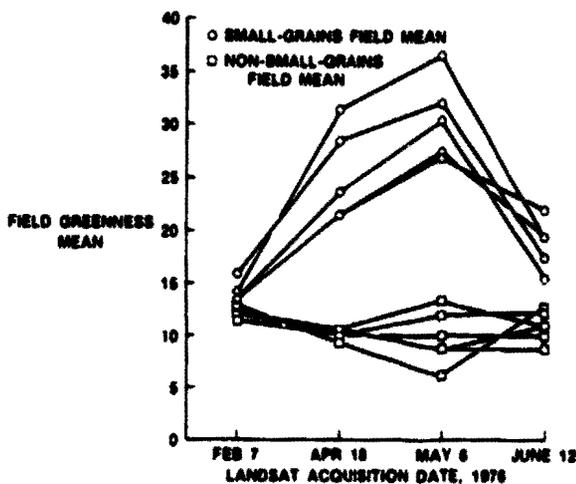


FIGURE 3.—Plots of temporal field greenness means of LACIE segment 1988, Finney County, Kansas.

define a non-small-grains trajectory. The third and fourth eigenvectors correspond to small eigenvalues and contain little information. Hence, the first and second PCG bands would appear to separate the small-grains field means from the non-small-grains field means. This separation is indeed evident in figure 4, which shows PCG band 2 versus PCG band

1 of the training field means. The linear decision boundary drawn in figure 4 is arbitrary. Figure 4 also shows that PCG band 1 alone is sufficient to separate small-grains field means from non-small-grains field means, whereas PCG band 2 alone cannot separate the small-grains field means from the non-small-grains field means.

For most segments in table 1, examining the first and second eigenvector components indicates that the components of the first eigenvector define a small-grains trajectory and the components of the second eigenvector define a non-small-grains trajectory. This is a favorable characteristic of the PCG transformation for applications in multitemporal agricultural Landsat data analysis.

The PCG-transformed data were used to generate color film products for several segments. Because of its phenological interpretability, the product generated from the PCG-transformed Landsat data is referred to as a temporal greenness interpretational film (TGIF) product. The color TGIF product appears to be a very useful way of representing multitemporal Landsat agricultural data. Also, the color TGIF product summarizes n color film products and can be used as an image interpretation aid. Figure 5 illustrates how four acquisitions are summarized by the TGIF product. In generating the

TOIF product, PCG bands 1, 2, and 3 are assigned to the red, green, and blue film converter guns, respectively.

An experiment was conducted to determine the amount of information retained in the PCG image: 5 segments were classified 3 times using all 16 original channels, the 4 greenness bands, and, finally, the first 2 PCG bands. The classification performance is specified here in terms of the probability of correct classification (PCC) and the estimated small-grains proportion in the segment compared with the ground-truth (GT) small-grains proportion. The PCC here is calculated as the average PCC of the training and test samples, which, according to the work of Foley (ref. 6), provides quite a good estimate of the actual error probability even for a small number of samples. For each segment, 50 training and 60 test samples were used to execute a classification run. The classification performance for each of the five LACIE segments is shown in table II. Table III shows the overall classification performance for the five segments, where the mean square error (MSE) in small-grains proportion is computed as follows.

Let $p^{(GT)}$ and p be the GT and a classification-estimated small-grains proportion, respectively; MSE is then

$$MSE = \sqrt{\frac{1}{5} \sum_{i=1}^5 [p_i - p_i^{(GT)}]^2} \quad (11)$$

The overall classification performance using the first 2 PCG bands is close to the overall classification performance using all 16 original channels. This observation shows that the first two PCG bands contain most of the information necessary to separate small grains from non-small grains and agrees with the previously observed correlation in figure 4 for segment 1988.

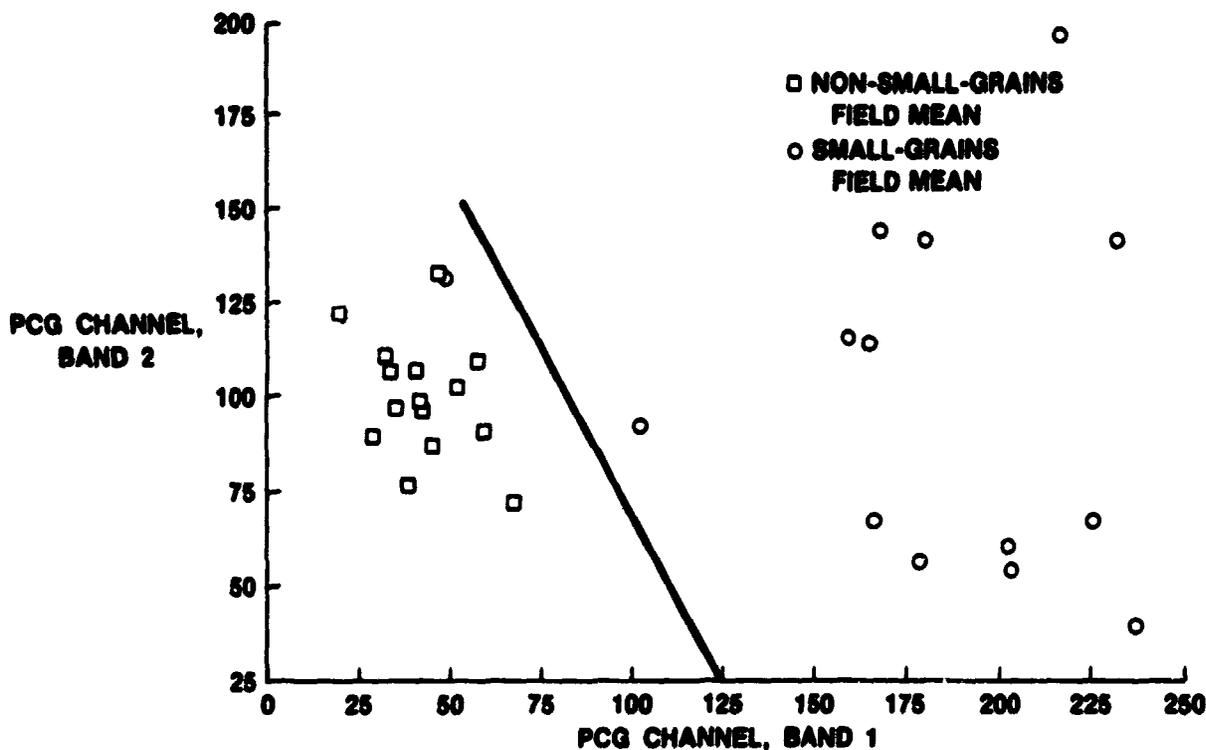


FIGURE 4.—Plots of training field means in the PCG space, using 1976 data from four temporal acquisitions, LACIE segment 1988, Finney County, Kansas.

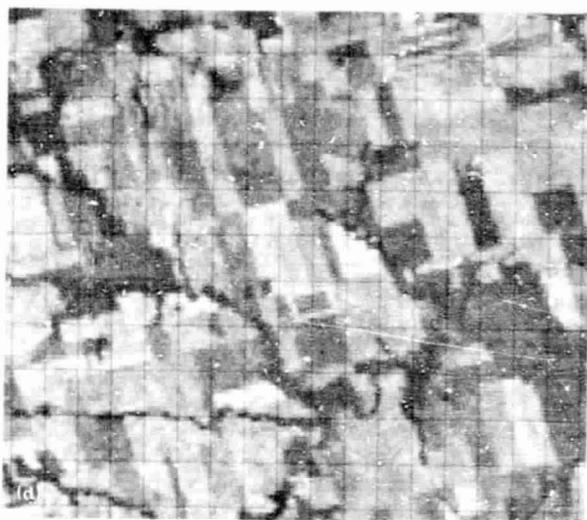
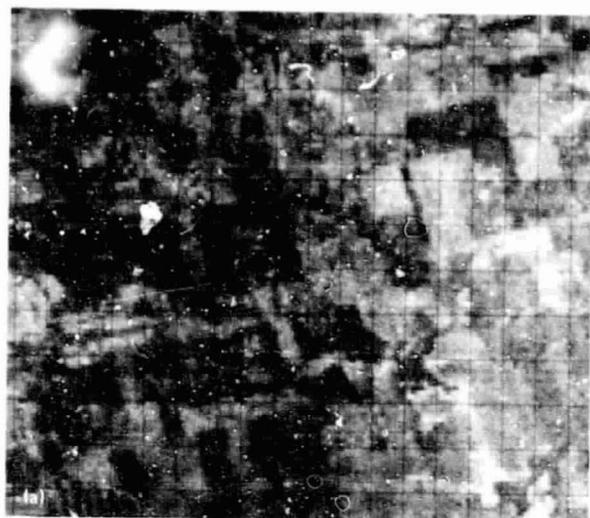


FIGURE 5.—Sixteen-channel-generated TGIF image of segment 1538, McCone County, Montana. (a) April 25, 1976. (b) June 18, 1976. (c) July 23, 1976. (d) August 11, 1976. (e) TGIF image.

TABLE II.—Classification Performance for Five LACIE Segments Using Various Methods

(Percent)

Segment	Ground-truth small-grains proportion	Classification performance using—					
		All 16 original channels (4 passes)		Bands $g_1, g_2, g_3,$ and g_4		PCG bands 1 and 2	
		Small-grains proportion	PCC	Small-grains proportion	PCC	Small-grains proportion	PCC
1988	33.0	34.1	96.1	36.6	95.2	35.7	94.6
1538	38.7	35.0	89.7	31.0	88.0	34.0	86.0
1618	63.5	62.0	96.2	58.0	88.7	58.0	87.7
1655	28.0	28.0	98.2	21.0	88.0	24.0	77.2
1694	27.7	26.0	82.0	26.0	80.7	26.0	74.9

SPECTRAL PLOT

A spectral plot of a Landsat scene is a graph of one channel of the data versus another. The spectral plot relates the image space (i.e., the spatial domain) of the PFC product to the spectral space of the classifier. A schematic diagram showing the relationship between image space and spectral space is given in figure 6. The spectral plot uses the inherent two-dimensionality of Landsat data, where most of the spectral class separability exists (ref. 3). For example, overlaying the spectral plot of training field means over the spectral plot of the scene provides a quick view of missing signatures.

The axes used for generating a spectral plot may be two selected Landsat channels or two linear combinations of channels; i.e.,

$$W = BX + e \quad (12)$$

where B = a two by n transformation matrix of rank 2

e = a two by one bias vector

W = a two by one vector of the channels to be plotted

The transformation matrix B might be formed from the first two rows of the Kauth transformation, from the first two rows of M in equation (1), or from a linear-combination feature-selection algorithm (ref. 7). If B is obtained from the Kauth transforma-

TABLE III.—Overall Classification Performance for Five Segments Using Various Methods

Classification method	MSE in small-grains proportion, percent	Average PCC, percent
All 16 original channels	2.00	92.44
Bands $g_1, g_2, g_3,$ and g_4	5.56	88.12
PCG bands 1 and 2	3.96	84.08

tion, then the dimension of the data n must be four since the Kauth transformation is a four by four matrix.

A color-coded spectral plot contains the locations of the pixels and the channels to be used for coloring them on the spectral plot. The location of each pixel on the spectral plot is computed using the radiance values (or linear combinations of radiance values) of the pixel. Multiple pixel occurrences at the same location on the spectral plot are shown to be the color of the pixel corresponding to the first occurrence. To illustrate what is meant by a color-coded spectral plot, assume that a given pixel on the Landsat image has radiance values of 28, 30, and 50 on channels 1, 2, and 4, respectively, as shown in figure 7. Also, let channel 4 be plotted against channel 2. The color-coded spectral plot is created by assigning the values of 28, 30, and 50, respectively, to the point (30, 50) on channels 1, 2, and 3 of the spectral image. By maintaining the same gains and biases, the color-coded spectral image can be displayed in the same color as

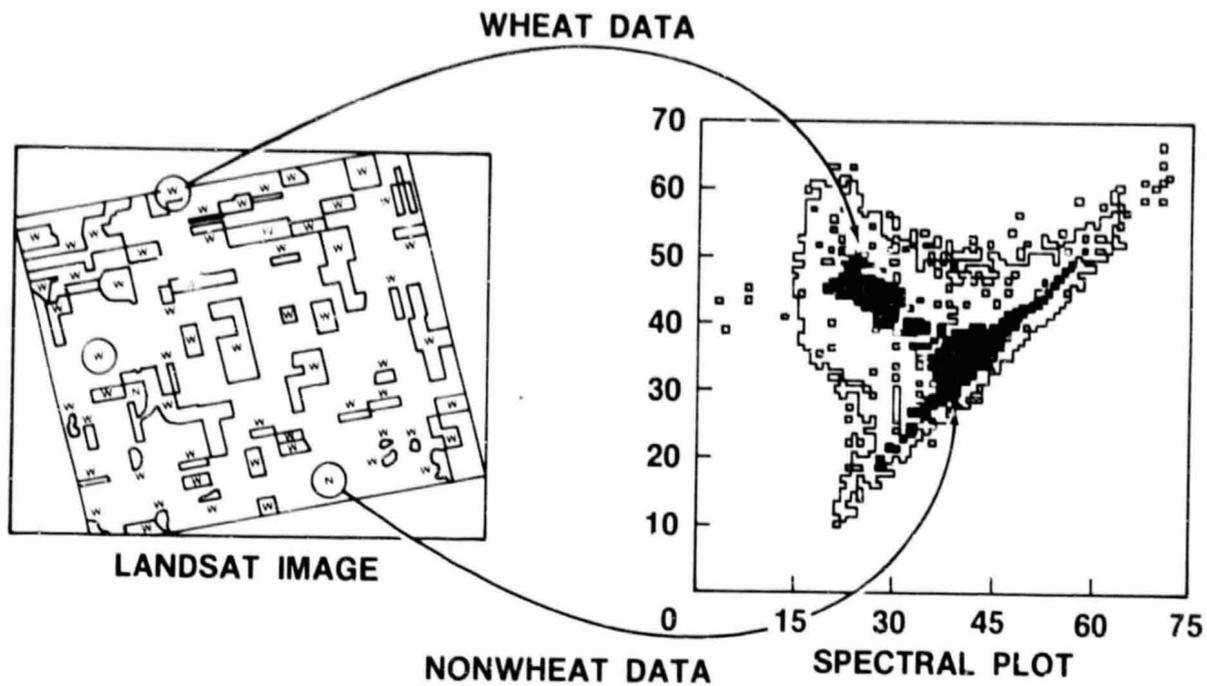


FIGURE 6.—Schematic diagram showing the relationship between image space and spectral space.

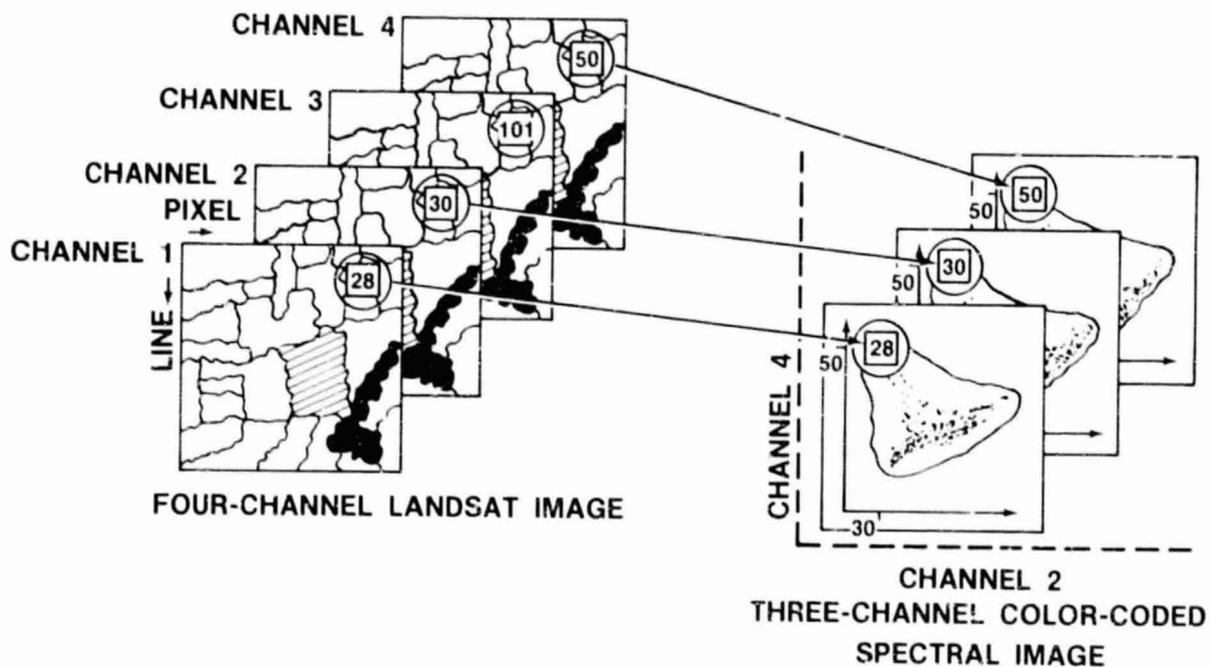


FIGURE 7.—Schematic diagram showing how Landsat radiance values are plotted on the color-coded spectral image.

the original Landsat image. Such a spectral plot easily reveals the spectral distribution associated with original colors in the Landsat scene.

The color of the pixel on the spectral plot is optional. It can be colored according to its original radiance value or the mean of the class, cluster, or field from which it was extracted. Naturally, the pixel location on the spectral plot can also be colored using the PCOMP value for the pixel.

Color-coded spectral plots can be used to observe the partitioning of spectral space imposed by clustering or by maximum likelihood classification. It can be used also to view the spectral locations of training samples. A partition of the two-dimensional spectral space by the maximum likelihood classification rule is depicted on the color-coded spectral plot by coloring the pixel on the plot according to the mean of the subclass to which it was assigned. A change in the color or its intensity on such a plot determines the maximum likelihood decision boundary.

When multiple, registered Landsat images are available for a scene, the location of the pixels to be plotted can be selected from two channels of one pass, whereas another pass is used for color definition. Such a color-coded spectral plot is especially useful for the analysis of multitemporal Landsat data. Through spatial correlation of this color-coded spectral plot and the Landsat image from which the plotting axes are selected, areas of temporal change caused by factors such as growth, disease, severe weather conditions, and harvest can be delineated. A typical application of the color-coded spectral plot, which is currently being considered by LACIE, is to use the plot as an aid in labeling the training samples. This is done by providing a spatial correlation between the spectral plot and the original Landsat image.

CONCLUSION

To aid in interpreting Landsat imagery, four color image display techniques have been developed. These interpretation aids are the cluster image, the PCOMP cluster image, the PCG image, and the color-coded spectral plot. From the results of experimentation, the four display techniques have been shown to be useful in selecting and/or identifying training data from Landsat images. The developed interpretation aids are being considered for implementation in LACIE.

REFERENCES

1. Large Area Crop Inventory Experiment. NASA press release, Nov. 6, 1974.
2. Levy, S. A.: Evaluation of Cluster Images. Lockheed Electronics Co., Inc., Technical Memorandum LEC-9001, NASA Johnson Space Center, Houston, Texas, Oct. 1976.
3. Kauth, R. J.; and Thomas, G. S.: The Tasselled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue Univ. (W. Lafayette, Ind.), June 29-July 2, 1976, IEEE cat. 76, CH1103-1-MPRSD.
4. Ready, P. J.; and Wintz, P. A.: Information Extraction, SNR Improvement, and Data Compression in Multispectral Imagery. IEEE Trans. on Comm., 10, Oct. 1973.
5. Abotteen, R. A.: Analysis of Principal Component Transformed Landsat Data. Technical Memorandum LEC-9003, Lockheed Electronics Co., Inc., Houston, Texas, Aug. 1976.
6. Foley, D. H.: Considerations of Sample and Feature Size. IEEE Trans. Inform. Theory, IT-18, Sept. 1972, pp. 618-626.
7. Decell, H. P.; and Quirein, J. A.: An Iterative Approach to the Feature Selection Problem. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue Univ. (W. Lafayette, Ind.), Oct. 1973, IEEE cat. 73, CM0834-2GE.

A Programed Labeling Approach to Image Interpretation

M. D. Pore^a and R. A. Abotteen^a

INTRODUCTION

The procedure for analyzing Landsat multi-spectral scanner data in LACIE is called Procedure 1 (P-1). One purpose of this analysis of Landsat data is to estimate the acreage of small grains in a 9- by 11-kilometer rectangular area, referred to as a segment. For this purpose, the four-channel vector of spectral values for each picture element (pixel) in the segment is transformed by a production film converter (PFC) to form a simulated color-infrared film product of the segment. There are 22 932 pixels in a segment.

The PFC products are generated for several dates throughout the growing season of small grains; from these, up to four are selected by the analyst-interpreter (AI) for machine processing. Every pixel falling on a 10- by 10-pixel grid is referred to as a grid pixel or grid dot. The first grid dot is on row 10, column 10, and one would assume that it represents the same "piece of real estate" on the ground on all four acquisition dates. It often happens that grid dots appear to switch fields because the registering of the PFC products to each other or, more precisely, to a reference date is not very accurate.

The labeling of pixels for particular acquisitions is a more precisely defined problem than multitemporal labeling since the grid dots may not fall on exactly the same real estate; however, despite this lack of perfect registration, P-1 requires the assignment of multitemporal labels to a subset of the grid dots. The multitemporal advantage of tracking crop growth outweighs the loss due to imprecise registration in P-1. Manual labeling techniques require the AI to use not only PFC products but also agricultural and meteorological (agromet) data and spectral aids in an integrated, judgmental fashion. To control an anticipated high variance in these techniques, a semiauto-

mated labeling technology was developed. The product of this technology—LIST (Label Identification from Statistical Tabulation)—is the subject of this paper.

The LIST operates on a discriminant analysis basis and thus has the potential for several favorable qualities. Among these are the ability to measure the reliability of a label and the ability to introduce an arbitrary bias. The latter property could be useful in offsetting other biases, such as those due to a failure to consistently detect small grains. In summary, automation in labeling can introduce the favorable aspects of using objective information and integrating continuous variable information in a way that an AI cannot.

This paper introduces the LIST properties, presents detail in describing the LIST development, gives numerical results of an application, and discusses the evaluation and conclusions about LIST.

LIST DEVELOPMENT

It has been observed that two AI's can study the same PFC imagery, agree on the various features (agromet information), and yet come to different conclusions about pixel labels. This may be due to different weighting of the information rather than to incorrect labeling or personal biases. However, this labeling phenomenon does create a high variance in labeling that is undesirable.

The LIST asks the AI questions related to simple properties and numerically quantizes the AI information, the agromet data, and the spectral information in variable formats and labels through use of a statistical discrimination process. This process generates a consistent (lower variance) procedure that can be manipulated in biases. The consistency of labeling is twofold: (1) questions relate to simple properties to generate consistent responses among AI's, and (2) a discriminator will give consistent labels from a

^aLockheed Electronics Company, Houston, Texas.

given set of information (responses).

A historical note is in order here. Originally, a 50-item questionnaire was developed by experienced AI's with the aim of including all possible sources of information and current interpretation channels (ref. 1). Because image interpretation via this questionnaire was an extremely lengthy process, very little data were processed by means of the questionnaire. However, with the small amount of data so derived, a more operational questionnaire was developed (ref. 2). This latter questionnaire is the LIST labeling procedure described here.

The questions posed to the AI to be answered from the PFC imagery products (more than one product may be available) are given in table I. The first four questions are answered once for each pixel, whereas question 5 is answered four times for each pixel (once for each acquisition date). Table II is a list of the questions to be answered in an automated fashion. Each question is answered for each acquisition with the exception of questions 3 and 6, which are multitemporal trajectory responses (answered once for all acquisitions).

The AI questions are used to screen the pixels and to allocate labels of "designated other" (DO), "boundary pixel," or "pure pixel." The DO pixels are automatically labeled "other," and the boundary pixels are not dealt with further; they are left with the

boundary pixel label. The pure pixels, however, are labeled by the discriminant algorithm. Only automated responses are used in the discriminator; however, AI question 5 responses are used to answer automated questions 5 and 6. The discriminator will generate labels of either "small grains" or "other."

Any of several discriminant analysis algorithms can be used to generate the labeling classifier. However, the two used in this study will be the only algorithms discussed here. Both are based on minimizing the mean square error (MSE). The first and principal algorithm used is the one in the Statistical Package for the Social Sciences (ref. 3). The discrimination is followed by a classification of mixture densities based on Gaussian distribution assumptions and prior-category probabilities equal to the training proportions. This classification of mixture densities changed very few labels and was not considered a significant attraction (or distraction) from the discriminant algorithm. This algorithm was versatile and contains many useful options; however, it assumes that the within-category covariance matrices are equal (traditional Fisher linear discriminator) and hence can be improved.

The second algorithm is a minimum MSE Bayesian procedure known as the Patterson-Pitt algorithm as implemented by Thadani (ref. 4) and Ahlers (ref. 5). It uses a loss matrix and prior-catego-

TABLE I.—LIST Questions for the Analyst

<i>Question</i>	<i>Response</i>
1. Pixel is obviously in	<input type="checkbox"/> Nonagricultural area. STOP; pixel is "designated other" (DO). <input type="checkbox"/> Agricultural area or indeterminate.
2. Is pixel registered with regard to analyst-chosen registration date (i.e., in the same category)?	<input type="checkbox"/> No. STOP; pixel is not classifiable. <input type="checkbox"/> Yes or indeterminate.
3. Is pixel a mixed pixel (part of more than one field or boundary)?	<input type="checkbox"/> Yes. STOP; pixel is not classifiable. <input type="checkbox"/> No or indeterminate.
4. Is this an anomalous pixel (not representative of most of the other pixels within the field)?	<input type="checkbox"/> Yes. STOP; pixel is not classifiable. <input type="checkbox"/> No.
5. PFC vegetation canopy indication is _____ (Use all available imagery film types.)	<input type="checkbox"/> (0) No vegetation canopy <input type="checkbox"/> (1) Low-density green vegetation canopy <input type="checkbox"/> (2) Medium-density green vegetation canopy <input type="checkbox"/> (3) High-density vegetation canopy <input type="checkbox"/> (4) Senescent (turning) vegetation canopy <input type="checkbox"/> (5) Harvested canopy (stubble)

TABLE II.—Automated LIST Questions for Classification of Small Grains

Question	Response
1. Green number of pixel is (corrected to 60° incidence).	_____
2. Is the green number of the pixel within the range for small grains?	____Yes ____No
3. Do the green numbers follow a small-grains trajectory?	____Yes ____No
4. Brightness number of pixel is	_____
5. Is the vegetation indication of the pixel valid for the Robertson biostage of wheat for the acquisition?	____Yes ____No
6. Does the pixel follow a small-grains vegetation development pattern?	____Yes ____No

ry probabilities equal to the training proportions. The loss matrix used was the traditional zero-one loss matrix.

The final LIST product is a set of pixel labels consisting of "boundary," "small grains," and "other." The following section addresses the problem of what techniques were used to answer the automated questions used to produce these labels.

LIST KEYS

It has been shown by Abotteen (ref. 6) that when the components of the first greenness image eigenvector are plotted as a function of the acquisition date, the shape of the curve is similar to the temporal wheat trajectory. It was also shown in the same work that high values in the first principal component greenness (PCG) band correspond to small-grains pixels. Hence, the PCG statistic is introduced here as a feature for separating small grains from non-small-grains. The PCG statistic of a dot (pixel) is the first PCG band value for that dot. It can be calculated by taking the inner product of the first greenness image eigenvector with the green number vector for the pixel. The PCG statistic of a pixel answers the question: does the pixel's temporal greenness trajectory look like a small-grains dot greenness trajectory?

The calculation of the PCG statistic for a pixel requires knowledge of the first greenness image eigenvector. A model is developed here to estimate this eigenvector given the Robertson biostage of wheat (ref. 7). Two models are developed, one for winter small grains and the other for spring small grains. A plot of the components of the first greenness image eigenvector as a function of the Robertson wheat biostage for seven quadritemporal winter small-grains LACIE segments acquired in the 1976 crop year is shown in figure 1. The Robertson biostage axis in figure 1 is divided into six increments. For every increment, the eigenvector components corresponding to Robertson biostage numbers falling in the increment are recorded and averaged. The Robertson biostage number range for every increment and its corresponding average eigenvector component are shown in table III. A plot of the average first greenness eigenvector components as a function of the Robertson biostage for winter small grains is shown in figure 2. The Sun-angle-corrected (SAC) eigenvector components are also shown in figure 2 as the dotted line. The dotted line in figure 2 is not significantly different from the solid line, which eliminates the need for using SAC first greenness eigenvector components. Instead, the green number vector for a pixel is SAC when used with the estimated first eigenvector for calculating the pixel's PCG

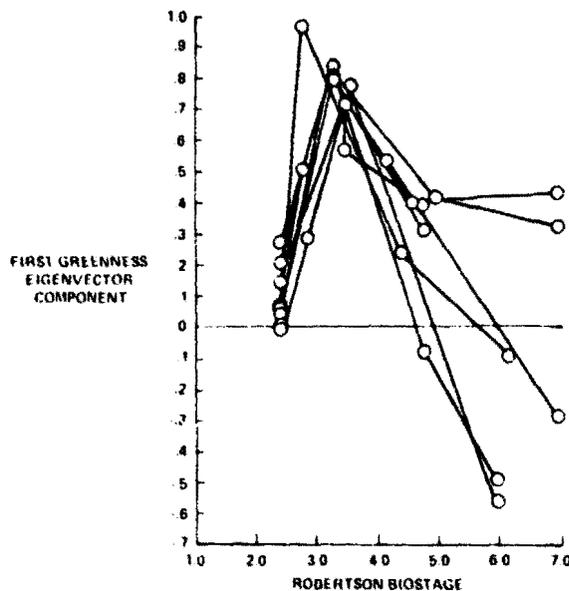


FIGURE 1.—First greenness image eigenvector components as a function of Robertson biostages for seven winter small-grains segments.

TABLE III.— Robertson Biostage Number Range and Corresponding Average First Greenness Eigenvector Components for Winter Small Grains

Robertson biostage range	Average first greenness eigenvector component
2.0 to 2.5	0.09975
2.5 to 3.0	.58433
3.0 to 3.8	.763
3.8 to 4.5	.3845
4.5 to 5.0	.285
5.0 to 7.0	-.1165

statistic. A digitized version of figure 2 is shown in table IV, where Robertson biostage numbers are increased by 0.1.

The spring small-grains model that estimates first greenness image eigenvector components is developed in a fashion similar to that just described. A plot of the first greenness eigenvector components as a function of the Robertson biostage for six LACIE quadritemporal spring small-grains segments acquired in the 1976 crop year is shown in figure 3. The line with dotted ends in figure 3 represents the average first greenness eigenvector components for spring small grains. A digitized version of this line is shown in table V, where Robertson biostage numbers are increased by 0.1.

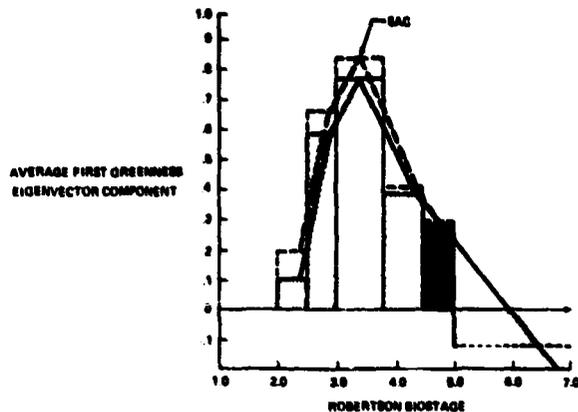


FIGURE 2.—Average first greenness image eigenvector components as a function of Robertson biostages for winter small grains.

The ground-truth-labeled small-grains grid intersection dots on a LACIE segment image were used to calculate a temporal average green number for small grains. A total of 34 segments was used; each segment had several acquisitions in the 1976 crop year. The Robertson biostage number for each acquisition was also obtained. When the temporal average green number was computed, the standard deviation was also computed for small grains. The segments contained either winter, spring, or mixed small grains. The small-grains green number range was developed

TABLE IV.— Principal Component Greenness Statistics Generation Table for Winter Small Grains in LIST

Robertson biostage number	First greenness eigenvector component	Robertson biostage number	First greenness eigenvector component	Robertson biostage number	First greenness eigenvector component
2.4	0.10	4.0	0.53	5.6	0.085
2.5	.24	4.1	.49	5.7	.06
2.6	.33	4.2	.45	5.8	.04
2.7	.42	4.3	.41	5.9	.01
2.8	.51	4.4	.38	6.0	-.01
2.9	.59	4.5	.35	6.1	-.04
3.0	.63	4.6	.33	6.2	-.06
3.1	.66	4.7	.30	6.3	-.08
3.2	.69	4.8	.28	6.4	-.11
3.3	.73	4.9	.25	6.5	-.13
3.4	.76	5.0	.23	6.6	-.16
3.5	.73	5.1	.21	6.7	-.18
3.6	.69	5.2	.18	6.8	-.2
3.7	.65	5.3	.16	6.9	-.23
3.8	.61	5.4	.13	7.0	-.25
3.9	.57	5.5	.11		

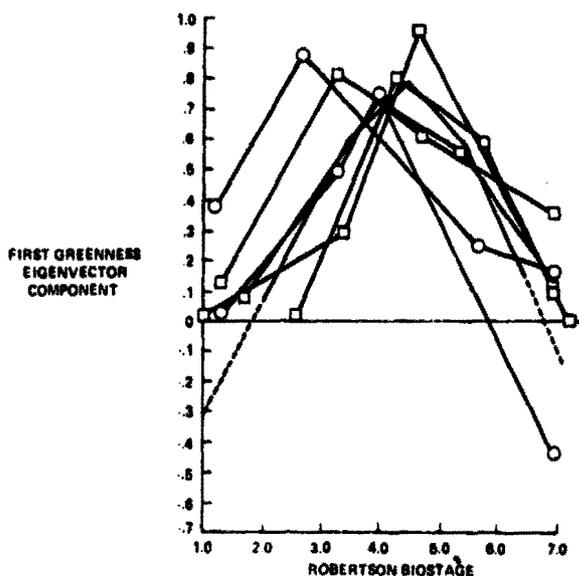


FIGURE 3.—First greenness image eigenvector components and their averages as a function of Robertson biostages for spring small grains.

for winter and spring small grains. The segments designated as mixed had two Robertson biostage numbers and were used in both the winter and spring

small-grains green number range models.

The Robertson biostage numbers for winter and spring small grains were divided into several sections. For each range of Robertson biostage numbers representing a certain section, the average small-grains green number and the standard deviation corresponding to a Robertson biostage number in the section are recorded. The average small-grains green numbers and the standard deviations within a Robertson biostage number range are averaged to obtain a grand small-grains green number mean and a grand standard deviation. For each Robertson biostage number range, the corresponding grand small-grains green number mean and standard deviation are shown in table VI for winter and spring small grains. The winter small-grains green number range as a function of the Robertson biostage number is shown in figure 4 for 1 standard deviation. The lower and upper bounds of the green number in figure 4 are computed by taking the grand green number mean minus and then plus 1 grand standard deviation, respectively. Figure 5 shows the winter small-grains green number range as a function of the Robertson biostage number for 2 standard deviations. The spring small-grains green number range as a function of the Robertson biostage number is shown in figure 6 for 1 and 2 standard deviations.

TABLE V.—Principal Component Greenness Statistics Generation Table for Spring Small Grains in LIST

Robertson biostage number	First greenness eigenvector component	Robertson biostage number	First greenness eigenvector component	Robertson biostage number	First greenness eigenvector component
1.5	-.125	3.4	.057	5.3	.0618
1.6	-.09	3.5	.60	5.4	.595
1.7	-.05	3.6	.62	5.5	.575
1.8	-.018	3.7	.64	5.6	.53
1.9	.02	3.8	.66	5.7	.49
2.0	.059	3.9	.68	5.8	.44
2.1	.09	4.0	.695	5.9	.40
2.2	.13	4.1	.715	6.0	.35
2.3	.17	4.2	.73	6.1	.31
2.4	.205	4.3	.75	6.2	.27
2.5	.24	4.4	.77	6.3	.225
2.6	.28	4.5	.785	6.4	.18
2.7	.31	4.6	.765	6.5	.141
2.8	.35	4.7	.745	6.6	.10
2.9	.39	4.8	.72	6.7	.055
3.0	.42	4.9	.70	6.8	.01
3.1	.46	5.0	.68	6.9	-.03
3.2	.50	5.1	.66	7.0	-.07
3.3	.53	5.2	.64		

TABLE VI.—Grand Small-Grains Green Number Mean and Standard Deviation

Robertson biostage range	Grand green number mean	Grand standard deviation
<i>Winter small grains</i>		
2.0 to 2.6	6.41	2.64
2.6 to 3.0	12.1	5.56
3.0 to 3.5	14.78	5.63
3.5 to 4.0	16.23	6.75
4.0 to 4.5	18.30	7.65
4.5 to 5.5	14.59	5.85
5.5 to 7.0	9.25	4.15
<i>Spring small grains</i>		
1.0 to 2.0	5.12	2.68
2.0 to 2.5	9.28	3.93
2.5 to 3.0	17.03	8.13
3.0 to 3.5	18.85	7.63
3.5 to 4.0	20.65	7.13
4.0 to 5.0	20.57	6.32
5.0 to 6.0	11.52	4.3
6.0 to 7.0	8.25	3.65

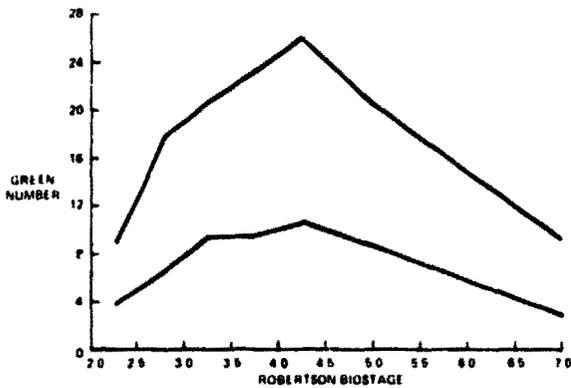


FIGURE 4.—Winter small-grains green number range as a function of Robertson biostages (1 standard deviation).

Figure 7, a chart of the Robertson biostage on the horizontal axis and the vegetation canopy on the vertical axis, describes an automation technique for the vegetation canopy questions. (Table VII provides the same information in tabular form.) For each acquisition, a point is located in figure 7 with the horizontal axis coordinate corresponding to the Robertson

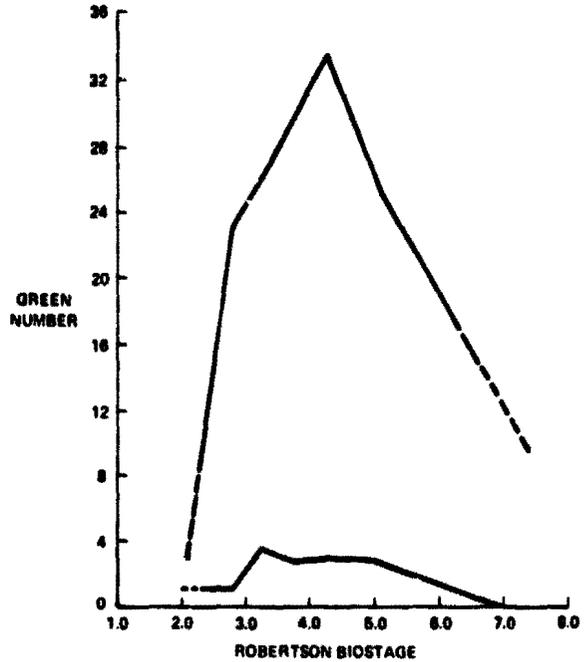


FIGURE 5.—Winter small-grains green number range as a function of Robertson biostages (2 standard deviations).

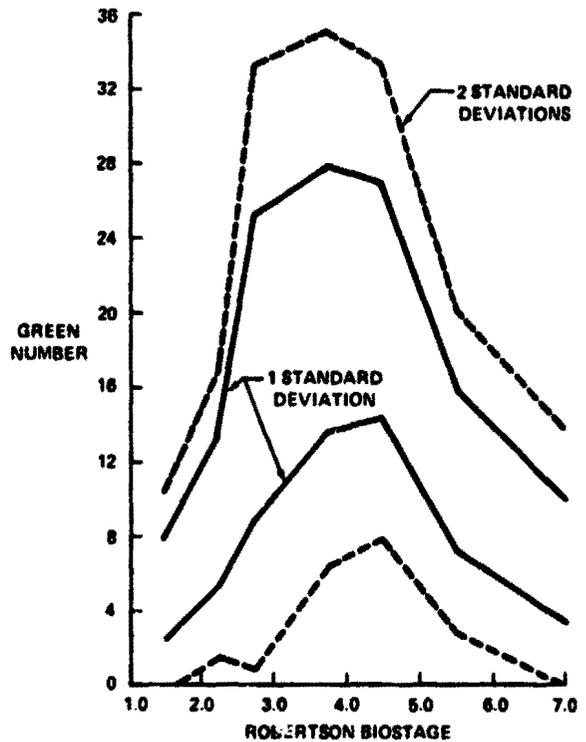


FIGURE 6.—Spring small-grains green number range as a function of Robertson biostages.

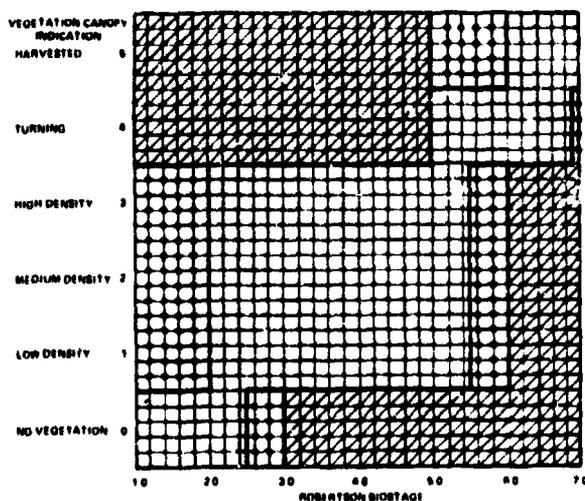


FIGURE 7.—Automation technique.

biostage for wheat for the acquisition and the vertical axis coordinate corresponding to the answer (for a given pixel) to question 5 (table I). If the point is in the blank or the dotted area, question 5 (table II) is automatically answered with a yes. If the point is in the shaded (barred) area, the answer is no (for that pixel and acquisition). Vertical borders belong to the class on the left.

Question 6 (table II) is answered with a yes if the points corresponding to the four acquisitions are all

in the blank or dotted regions and if at least two are in the blank region. This is equivalent to the rule corresponding to table VII: all four acquisitions must be classified, and at least two must be first-class responses for question 6 (table II) to be answered with a yes. Hence, the dotted region in figure 7 (and the second-class response in table VII) is used as a different designation from the blank region (first-class region) for answering question 6 (table II) only.

Figure 7 is like a crop calendar that reflects the growth characteristics of small grains. This chart may need to be revised to reflect the peculiarities of a particular growing season or a particular region of the country. For example, it has been suggested that the region corresponding to Robertson biostages 5.6 to 6.0 and vegetation canopy indication 1, 2, or 3 (parts of the sixth row in table VII) could be shaded (removed from the second-class response) for segments from winter wheat regions. This reflects the necessity for margin in acceptable vegetation canopy indications in mixed and spring wheat areas and for more uniformity in winter wheat areas. The present key (table VII) is intended to be quite general and may be used in this state.

TEST RESULTS

Four AI's were used to test the quality of the questions for discriminating small grains (agricultural

TABLE VII.—Automation Technique in Tabular Form

Robertson biostage range	First-class response	Second-class response
1.0 to 2.0	No vegetation (0)	Green vegetation (1, 2, 3)
2.1 to 2.5	No vegetation or green vegetation (0, 1, 2, 3)	
2.6 to 3.0	Green vegetation (1, 2, 3)	No vegetation (0)
3.1 to 5.0	Green vegetation (1, 2, 3)	
5.1 to 5.5	Green vegetation or turning (1, 2, 3, 4)	Harvested (5)
5.6 to 6.0	Turning (4)	Green vegetation or harvested (1, 2, 3, 5)
6.1 to 6.9	Turning or harvested (4, 5)	
7.0	Harvested (5)	Turning (4)

crops) from non-small-grains. Each AI analyzed 16 segments and took approximately 2.5 hours per segment. Earlier studies, such as the study of Register and Hocutt (ref. 8), have indicated that interpixel correlations decrease with distance and that a distance of 10 pixel widths corresponds to negligible correlation. Hence, dot grids are assumed to be independent samples with respect to crop types.

Separate analyses were performed for the 1976 winter wheat and spring wheat sites (eight of each). All Kansas blind sites in new LACIE stratum 11 with available ground truth were chosen as the winter test sites. The eight spring sites were chosen from the blind sites in new LACIE stratum 21 (fig. 8) for strata locations. Since ground truth was required in stratum 21, segments were chosen to be representative of the three-state coverage of the stratum. The data within each stratum were further partitioned into four training segments and four test segments. Table VIII describes this breakdown.

For each segment, four acquisition dates were chosen arbitrarily (without respect to special area agromet conditions or cloud cover) to cover the 1975-76 growing season of wheat. Table IX gives these dates and the respective Robertson biostages for winter wheat (WW) and spring wheat (SW). Three types of PFC products were generated: Product 1, Product 2, and the Kraus product (reference 9 describes these films). The films were made into research, test, and evaluation packets (separate from LACIE operational packets) to maintain a restricted experimental environment of labeling without full-frame imagery (185-kilometer square area of land) of the broad area of interest and without ancillary agromet information. Hence, accuracies should be below those experienced in an operational labeling system. The discriminants were determined using ground truth for the four training segments and accuracy was determined by using the discriminant function to classify the four test segments. Percentages of pixels correctly labeled were calculated from contingency tables of ground truth by LIST labels.

The particular variables that a stepwise discriminant procedure admits are a function of the number of training samples, the variability of the particular area sampled, and the acquisition dates. Certainly, implementation in LACIE of a training sample of the size used here is not recommended; hence, discriminant vectors and tests for category mean differences are not presented. Instead, tables for test accuracy (on segments not used in training) are presented. Figure 9 is a key for these contingency tables.

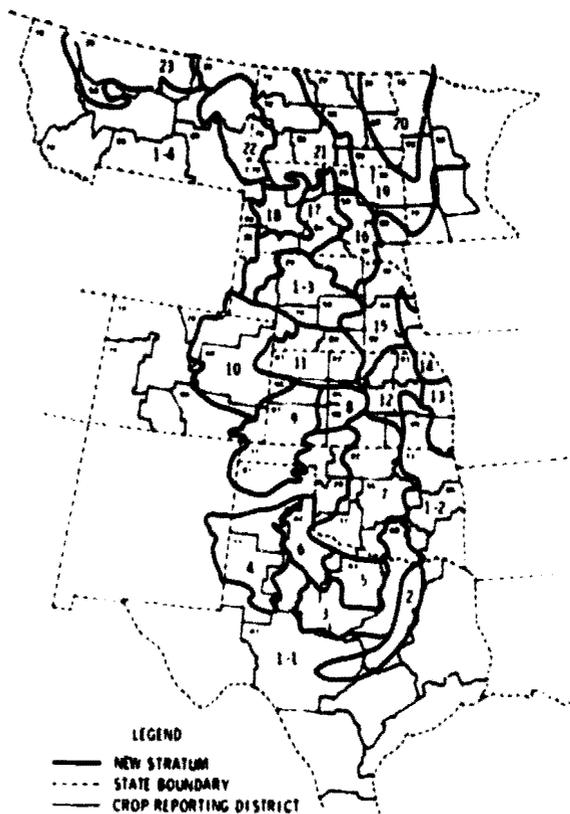


FIGURE 8.—Crop reporting districts and new LACIE strata in the U.S. Great Plains.

TABLE VIII.—LIST Data Set

Stratum	LACIE number	Type	Purpose	County, state
11	1019	Winter	Train	Norton, Kans.
11	1035	Winter	Train	Ford, Kans.
11	1855	Winter	Train	Trego, Kans.
11	1865	Winter	Train	Stevens, Kans.
11	1020	Winter	Test	Rawlins, Kans.
11	1852	Winter	Test	Lane, Kans.
11	1860	Winter	Test	Hodgeman, Kans.
11	1880	Winter	Test	Ellis, Kans.
21	1542	Spring	Train	Roosevelt, Mont.
21	1650	Spring	Train	Hettinger, N. Dak.
21	1651	Spring	Train	Bowman, N. Dak.
21	1667	Spring	Train	Harding, S. Dak.
21	1530	Spring	Test	Phillips, Mont.
21	1656	Spring	Test	Morton, N. Dak.
21	1660	Spring	Test	Logan, N. Dak.
21	1668	Spring	Test	Perkins, S. Dak.

TABLE IX.—LIST Data Acquisition Dates (1976)

Segment	County	Dates (WW blorage, SW blorage, if applicable)	Segment	County	Dates (WW blorage, SW blorage, if applicable)
1019	Norton, Kans.	Jan. 19 (2.4) Feb. 6 (2.5) June 12 (4.6) June 30 (5.4)	1852	Lane, Kans.	Mar. 31 (2.6) May 7 (3.2) June 20 (5.8) July 17 (6.0)
1020	Rawlins, Kans.	Feb. 25 (2.5) Apr. 10 (2.7) June 3 (3.7) July 18 (6.0)	1855	Trego, Kans.	Mar. 13 (2.6) Apr. 18 (3.0) June 20 (5.7) July 17 (6.0)
1035	Ford, Kans.	Mar. 13 (2.6) May 6 (3.4) June 1 (4.1) July 8 (6.0)	1860	Hodgeman, Kans.	Mar. 13 (2.5) May 6 (3.3) June 2 (4.1) July 8 (6.0)
1530	Phillips, Mont.	June 1 (3.5, 3.1) June 18 (3.9, 4.0) July 7 (5.5, 5.0) Aug. 12 (7.0, 6.0)	1865	Stevens, Kans.	Feb. 7 (2.4) May 15 (3.6) June 20 (5.8) July 8 (6.0)
1542	Roosevelt, Mont.	Apr. 25 (2.5, 1.1) June 18 (4.3, 3.4) July 6 (5.7, 5.0) July 24 (6.0, 6.0)	1880	Ellis, Kans.	Mar. 13 (2.6) May 6 (3.2) June 10 (4.9) July 16 (6.0)
1650	Hettinger, N. Dak.	May 9 (3.2, 2.0) May 27 (3.8, 3.0) Aug. 7 (6.0, 6.0) Aug. 25 (6.0, 6.0)			
1651	Bowman, N. Dak.	May 10 (3.3, 2.2) May 29 (4.0, 3.0) July 21 (6.0, 6.0) Aug. 8 (6.0, 6.0)			
1656	Morton, N. Dak.	May 9 (3.0, 2.0) July 2 (6.0, 4.4) July 20 (7.0, 6.0) Aug. 7 (7.0, 7.0)			
1660	Logan, N. Dak.	May 7 (3.1, 2.0) June 12 (4.2, 3.7) Aug. 6 (6.0, 6.0) Aug. 23 (6.0, 6.0)			
1667	Harding, S. Dak.	May 10 (3.4, 2.3) May 29 (4.3, 3.2) July 21 (5.9, 6.0) Aug. 8 (6.0, 6.0)			
1668	Perkins, S. Dak.	Apr. 22 (2.6, 1.7) May 9 (3.3, 2.3) May 28 (4.0, 3.1) Aug. 7 (6.0, 6.0)			

TYPE OF LABELER

	SG	NON	?
GROUND TRUTH	a	b+c	d
	e	f+g	h
	i	j	k

PCL

a,b,c,d = RAW PIXEL COUNTS FOR THE FOUR TEST SEGMENTS
e,f = RAW PIXEL COUNTS FOR THE DO PIXELS
h = a + b + c + d + e + f
g,h,i,j = MARGINAL PROBABILITIES (EXPRESSED AS PERCENT) OF CORRECT LABELING

$$g = \frac{a}{a+b+c} = 100 \cdot (1 - P(\text{COMMISSION})) = 100$$

$$h = \frac{d+f}{e+d+f} = 100 \cdot (1 - P(\text{COMMISSION})) = 100$$

$$i = \frac{a}{e} = 100$$

$$j = \frac{d+f}{e+d+f} = 100$$

$$k = \frac{b+c}{e} = 100 = \text{THE GROUND TRUTH PERCENTAGE OF SMALL GRAINS}$$

$$m = \frac{a+b}{e} = 100 = \text{THE LIST-LABELED PERCENTAGE OF SMALL GRAINS}$$

$$PCL = \frac{a+b+c+d+f}{k} = 100 = \text{THE PROBABILITY (EXPRESSED AS A PERCENTAGE) OF CORRECT LABELING}$$

FIGURE 9.—Contingency table key.

Four analyses were performed on the winter segments: two using the quadratic discriminator, one using the stepwise discriminant, and one using the AI labels. These results and the variables used are given in figure 10 for all four AI's, each responding to the four winter test segments. (The appendix gives variable definitions for all analyses.) As presently programed, the quadratic discriminator was determined to accrue numerical analysis errors of computation at an unacceptable rate and was not used in the spring site analyses.

All spring sites were treated as mixed wheat sites, even when winter wheat analysis was patently unnecessary. The mixed wheat philosophy was to give positive responses automatically where indicated for either spring or winter wheat. For example, if the canopy trajectory for a pixel was similar to a winter wheat trajectory (SUM is high for winter biostage numbers) but dissimilar to a spring wheat trajectory (SUM is low for spring biostage numbers), then KEYS and SUM were based on winter biostages for that pixel. The results for the spring sites are given in figure 11.

The AI percentage of small grains and the LIST percentage of small grains were consistently below the ground-truth percentage of small grains ($m < 1$ in fig. 9) regardless of the type of discriminant used. This is partly attributed to the facts that omission rates are apparently always less than commission rates ($b < c$ in fig. 9) and that there is a fairly consistent tendency for nearly 4 percent of the DO pixels to be small grains ($e/(e + f) \approx 0.038$).

Although midseason estimation cannot be effectively analyzed since acquisition date selection for end-of-season estimation is usually inappropriate for midseason estimation and specialized midseason questions (e.g., automated prototype green number trajectories) have not been developed, such an analysis is presented here, recognizing that lower than realistic accuracy is expected. Such an analysis indicates the efficacy of present keys and may be heuristically valuable in pointing to new developments. A rather high accuracy (PCL in the terminology of fig. 9) and a moderate decrease in the percentage of small grains reported ($m < 1$ in the terminology of fig. 9) is demonstrated in figure 12.

AI LABELS

	SG	NON	21.1%
GROUND TRUTH	SG	44 + 65	82%
	NON	586 + 1553	97%
	19.8%	87%	95%
			2803

93.5%

VARIABLES: ANALYST LABEL

(a)

LINEAR DISCRIMINANT

	SG	NON	21.1%
GROUND TRUTH	SG	35 + 65	83%
	NON	573 + 1553	96%
	22.7%	85%	96%
			2803

93.4%

VARIABLES: G1, CANOPY TRAJECTORY, B4, GREEN3, B2, G4, KEY4, B1, G2, PCGW, G3, KEY3, GREEN2, KEY2, P114, GREEN4, BIO2

(b)

Q WITH B AND G ONLY

	SG	NON	21.1%
GROUND TRUTH	SG	61 + 65	79%
	NON	578 + 1553	96%
	19.5%	85%	94%
			2803

92.6%

VARIABLES: B1, B2, B3, B4, G1, G2, G3, G4, AND ALL POSSIBLE INTERACTIONS

(c)

Q17

	SG	NON	21.1%
GROUND TRUTH	SG	50 + 65	81%
	NON	574 + 1553	96%
	20.0%	85%	95%
			2803

92.9%

VARIABLES: BIO2, BIO4, B1, B2, B4, G1, G2, G3, G4, GREEN2, GREEN3, GREEN4, PCGW, KEY2, KEY3, KEY4, CANOPY TRAJECTORY, AND ALL POSSIBLE INTERACTIONS

(d)

FIGURE 10.—LIST test accuracy on winter sites. (a) AI labels. (b) Linear discriminant. (c) Q with B and G only. (d) Q17.

C-11

AI LABELS

	SG	NON	8.8%
GROUND TRUTH	SG	113	32 + 79
	NON	47	166 + 2106
	6.3%	71%	96%
			2543

93.8%

VARIABLES: AI OPINION

(a)

LINEAR DISCRIMINANT

	SG	NON	8.8%
GROUND TRUTH	SG	105	40 + 79
	NON	67	146 + 2106
	6.8%	61%	96%
			2543

92.7%

VARIABLES: CANOPY TRAJECTORY, G1, G3, B4, B1, GREEN1, G2, G4, GREEN4, PCGW, B3, KEY3, B2

(b)

LINEAR WITH B-G-BIO STEP

	SG	NON	8.8%
GROUND TRUTH	SG	105	40 + 79
	NON	47	166 + 2106
	60%	69%	96%
			2543

93.5%

VARIABLES: CANOPY TRAJECTORY, GS1, GW3, BS4, BW1, GREEN1, GW1, GS4, GS2, GREEN4, BS3, KEY3, PCGW, BW3, BS2

(c)

LINEAR WITH B-G-BIO DIRECT

	SG	NON	8.8%
GROUND TRUTH	SG	100	45 + 79
	NON	41	172 + 2106
	5.5%	71%	96%
			2543

93.5%

VARIABLES: GW1, GW2, GW3, GW4, GS1, GS2, GS3, GS4, BW1, BW2, BW3, BW4, BS1, BS2, BS3, BS4, PCGW, PCGS, CANOPY TRAJECTORY, KEY1, KEY2, KEY3, KEY4, GREEN1, GREEN2, GREEN3, GREEN4

(d)

FIGURE 11.—LIST test accuracy on spring sites. (a) AI labels. (b) Linear discriminant. (c) Linear with B-G-BIO step. (d) Linear with B-G-BIO direct.

WINTER SITES

	SG	NON	21.1%
GROUND TRUTH	SG	409	117 + 65
	NON	113	546 + 1553
	18.6%	78%	92%
			2803

89.5%

VARIABLES: BIO1, BIO2, G1, B1, G2, B2, KEY2, GREEN1, GREEN2, GW1, GW2, BW2

(a)

SPRING SITES

	SG	NON	8.8%
GROUND TRUTH	SG	85	60 + 79
	NON	38	175 + 2106
	4.8%	69%	94%
			2543

93.0%

VARIABLES: SBIO1, SBIO2, WBIO1, WBIO2, GW1, GW2, BW1, BW2, GS1, GS2, BS1, KEY1, KEY2, GREEN1, GREEN2

(b)

FIGURE 12.—Midseason test accuracy. (a) Winter sites. (b) Spring sites.

EVALUATIONS AND CONCLUSIONS

The phenomenon of a nearly 4-percent DO being small grains constitutes a source of bias that is apparently consistent over diverse geographic regions and that is readily measurable. The unexpectedly high PCL (high means close to AI label accuracy) in the "undeveloped discriminator" for midseason

labeling analyses suggests that a directed development of a midseason LIST labeler (as opposed to a casual byproduct of an end-of-season LIST labeler) would yield a highly accurate operational labeling system.

The present Classification and Mensuration Sub-system procedural philosophy is for the AI to select a reference acquisition date (film) and to mentally ad-

just registration discrepancies of other acquisitions to give label accuracy to the "real estate" represented in the reference film. It is becoming increasingly evident that LIST, and in fact any labeling procedure that relies on spectral aids (e.g., trajectories), is inherently based on a different philosophy. Since acquisitions are usually not identically registered, spectral values for a pixel, across several acquisitions, therefore represent the area about the "real estate" and not a precise pixel of one date. Boundary pixels and mixed pixels (across a boundary) have spurious spectral trajectories; i.e., the trajectory is not sampled from a single category of interest but rather is switched from one category to another. Such trajectories tend to confuse the labeling process and reflect a basic modeling error in image interpretation. LIST, on the other hand, labels what is represented by the spectral trajectory (in this case, the grid dot intersection on the PFC (film) product).

To make this more meaningful, LIST first filters out the boundary (and mixed) pixels and treats these pixels as a nonlabelable class to be proportioned. In summary, LIST does not label real estate; it does label film grid intersection pixels. This philosophical change is implied by the increased reliance on spectral trajectories.

The high accuracies demonstrate that the concept of a programmed statistical discrimination approach to pixel labeling is valid and, in particular, that the LIST procedure performed comparably with the AI's in the restrictive environment of these test conditions. This is a highly successful result that confirms the efficacy of the LIST questionnaire. However, it can be easily and obviously improved through the further development and training of the automated keys, particularly green number ranges and trajectories.

Appendix

Variable Definitions for Analyses*

<i>Variables</i>	<i>Definition</i>	<i>Variables</i>	<i>Definition</i>
BIO1, BIO2, BIO3, BIO4 or WBIO1 through WBIO4 SBIO1 through SBIO4	Winter wheat Robertson biostages for the respective acquisitions	GW1 through GW4	The products of $G_i \times WBIO_i$ for $i = 1,2,3,4$
G1, G2, G3, G4	Spring wheat biostages	GS1 through GS4	The products of $G_i \times SBIO_i$ for $i = 1,2,3,4$
B1, B2, B3, B4	Green numbers	BW1 through BW4	The products of $B_i \times WBIO_i$ for $i = 1,2,3,4$
Brightness numbers		BS1 through BS4	The products of $B_i \times SBIO_i$ for $i = 1,2,3,4$
GREEN1 through GREEN4	Yes/no answer: Is green number in the small-grains range?		
KEY1 through KEY4	Yes/no answer: Is canopy in the small-grains range?		
Canopy trajectory	Yes/no answer: Is canopy trajectory acceptable for small grains?		
PCGW, PCGS	PCG statistic for winter and spring wheat, respectively		

*See reference 2 for the numerical derivations.

REFERENCES

1. Pore, M. D.: Label Identification From Statistical Tabulation (LIST). LEC-10825, Lockheed Electronics Co., Houston, Tex., Oct. 1977.
2. Abotteen, R. A.; and Pore, M. D.: Development of Semiautomated, Operational Label Identification From Statistical Tabulation (LIST). LEC-11457, Lockheed Electronics Co., Houston, Tex., Feb. 1978.
3. Nie, N. H.; Hull, C. H.; Jenkins, J. G.; Steinbrenner, K.; and Bent, D. H.: SPSS: Statistical Package for the Social Sciences. Second ed., McGraw-Hill, 1975.
4. Thadani, S. G.: A Nonparametric Loss-Optimal Pattern Classification System. LEC-11451, Lockheed Electronics Co., Houston, Tex., Jan. 1978.
5. Ahlers, C. W.: "As-Built" Design Specification for the Patterson-Pitt-Thadani Minimum Loss Classifier. LEC-12285, Lockheed Electronics Co., Houston, Tex., May 1978.
6. Abotteen, R. A.: Principal Component Greenness Transformation in Multitemporal Landsat Imagery. LEC-10333, Lockheed Electronics Co., Houston, Tex., May 1977.
7. Robertson, G. W.: A Biometeorological Time Scale for a Cereal Crop Involving Day and Night Temperatures and Photoperiods. Intern. J. Biometeorol., vol. 12, 1968, pp. 191-223.
8. Register, D. T.; and Hocutt, W. T.: Proportion Assessment Using Systematic Sampling (PASS) Feasibility Study. LEC-8049, Lockheed Electronics Co., Houston, Tex., Mar. 1976.
9. Austin, W. W.: Detailed Description of the Wheat Acreage Estimation Procedure Used in the Large Area Crop Inventory Experiment. LEC-11497, Lockheed Electronics Co., Houston, Tex., Feb. 1978.

Status of Yield Estimation Technology: A Review of Second-Generation Model Development and Evaluation

R. G. Stuff,^a T. L. Barnett,^a G. O. Boatwright,^a D. E. Phinney,^b and V. S. Whitehead^a

SUMMARY

Multiple regression models were selected as the LACIE yield estimation baseline, primarily on the basis of experience and expediency. Their requirements for long, region-specific historical records of yield and weather data and inherently damped responses were recognized as a priori limitations relative to LACIE objectives. These limitations and the potential improvements claimed for other approaches were the principal motives for initiating a Research, Test, and Evaluation (RT&E) program to evaluate and develop more amenable models for the agricultural-meteorological (agromet) estimation of yield.

Candidate alternatives were labeled as second or third generation, based on their progressively more detailed resolution elements and the effort needed to make them operational. The objectives of research in second-generation models were to obtain (1) yield estimation capability for any arbitrary unit of area and (2) greater responsiveness and accuracy in yield estimates through the use of additional data sources applied at smaller spatial and temporal scales. Also, candidate second-generation models were to be evaluated in a research mode by characterizing their expected performance relative to the multiple regression (first-generation) models.

Limits on independent data available to test candidate models required the use of different evaluation procedures. Test runs of the Baier model (ref. 1) displayed inadequate results when applied outside the Canadian spring wheat area, and major revisions would be required to adapt it to winter wheat.

Historically based time-series estimates for individual areas were used when replications over years or other accuracy statistics were not available. Time-series projections were found to be better predictors of the 1975 distribution of spring wheat yields across counties and districts in four states than predictions by the Earth Satellite Corporation (EarthSat) moisture stress model (ref. 2). The Feyerherm model (see the paper by Feyerherm and Paulsen entitled "A Universal Model for Estimating Wheat Yields"), which was developed as a follow-on to the Baier model test, was compared to the regression models through 10 years of test predictions in the U.S. Great Plains states. Performances were estimated to be equivalent for spring wheat but not as good as the first-generation models for winter wheat. Additional test predictions for states in the U.S. Corn Belt and the U.S. Pacific Northwest, for India, and for an oblast in the U.S.S.R. were used to verify the quasi-universal applicability of the Feyerherm model. Other evidence for universality was found in development of the Haun, Cate-Liebig, and Center for Climatic and Environmental Assessment (CCEA) II models (see reference 3, the paper by Cate et al. entitled "The Law of the Minimum and an Application to Wheat Yield Estimation," and the paper by LeDuc entitled "CCEA Second-Generation Wheat Yield Model for Hard Red Wheat in North Dakota").

Candidate models which use the Landsat-derived leaf area index (LAI) in transpiration and growth-based yield models were developed by Kanemasu (ref. 4), and other basic yield-Landsat relationships were investigated. It was concluded that data base inadequacy was the factor limiting performance in all the second-generation models considered and that each of the models has more yield-predicting capability than was reached during LACIE.

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Inc., Houston, Texas.

INTRODUCTION

The basic objectives of LACIE were to evaluate and demonstrate remote crop production estimation technology in a practical application. Also, a supporting Research, Test, and Evaluation (RT&E) function was established (ref. 5) to develop and/or validate improvements for the Applications Evaluation System.

For remote estimation of wheat yields, multiple regression equations derived from historical crop and weather data were considered the most expedient approach (see the paper by McCrary et al. entitled "Operation of the Yield Estimation Subsystem"). These models had been developed since the early 1900's (ref. 6) to analyze historical yield variability for individual states, but their prediction accuracy for large areas was not known. Several inherent weaknesses in the regression models were identified initially; the key weaknesses of concern were as follows:

1. Restriction of applicability to areas with long historical records

2. Insensitivity because of (a) averaging "out" effects of local and short-duration phenomena by state and monthly variables, (b) limited numbers of parameters which could be estimated independently for any given length of record, (c) lack of attention to crop calendar changes, and (d) use of surrogate variables not directly related to crop functioning (such as precipitation for soil or crop moisture and year for technological trend)

Since it would require at least LACIE Phase I (1974-75) to determine whether the regression models would support the LACIE accuracy goal (estimates of regional production within 10 percent of the true value 90 percent of the time), the feasibility of alternative yield models was an issue for RT&E from the beginning.

The use of analog areas or the acquisition of a universal model were the options considered for overcoming the first (area specificity) limitation. Theoretical claims and empirical evidence to support each option were available, but neither had been tested. In support of yield, RT&E was designed to address both problems. Descriptions and examples of research in the analog region method are given elsewhere (see reference 7 and the paper by Strommen et al. entitled "Development of LACIE CCEA-1 Weather/Wheat Yield Models"); this paper will summarize the research and evaluation of models designed to be applicable to any given region.

A special conference was held at the beginning of LACIE to review the state of the art in wheat yield models (ref. 8). Haun (ref. 9), Robertson (ref. 10), and others presented evidence that models with universally applicable yield-weather coefficients were possible if they accounted for factors such as the following.

1. Defining environmental variables in biological time rather than in calendar time

2. Using soil moisture rather than precipitation as the moisture supply variable

3. Use of varieties, fertilizer application, irrigation, etc., to explicitly explain yield trends

4. Natural differences in soil fertility, water-holding capacity, etc.

5. Variable representation of daily to weekly weather and soil series to family levels of detail

Models with these characteristics were called second generation in contrast to the less-detailed area-specific but senior first-generation models. A discussion of second-generation models and how they compare to other models is given in the appendix.

The more detailed and theoretical system (third-generation) models of crop growth mechanisms were considered, but their expansion to predict yields with conventionally reported agricultural and meteorological data was not practical within the scope of LACIE. Existing third-generation models (see appendix) were designed primarily for research rather than for operational yield prediction. Especially lacking were the submodels necessary for the operation of these models with conventional meteorological data (air temperatures and precipitation) and the extension of grain yields to aggregable units of area. One source estimated the cost of developing a third-generation model to be \$25 million.

Another initial proposal was that crop appearance as observable in satellite data could readily contribute much to yield estimation (ref. 8). The basic rationale for using appearance variables comes from the fact that mathematical models cannot account for the number and complexity of environmental and agricultural factors which are integrated into crop yields. Appearance could provide an estimate of integrated results at any point in time if correctly interfaced with agromet information. Since spectral-yield relationships should change with stage of crop development, growth history, and cultural practices, the agromet components were assessed a leading or host role in a combined model. Also, early in the season and under cloudy conditions, only agromet data may be available. Thus, research to derive yield

information from Landsat data was conducted concurrently with investigations using meteorological and other agricultural data.

The overall objective of supporting research in second-generation yield models was to attain a step improvement over the LACIE baseline yield estimation technology. Specific objectives were as follows.

1. To obtain a yield estimation capability for any arbitrary unit of area such as the 5- by 6-nautical-mile LACIE sample segment, the LACIE agrophysical stratum, or a foreign region without sufficient historical data to develop adequate multiple regression models.

2. To acquire a model that can be readily expanded to use additional data (temporal and spatial resolution) and data sources as the corresponding functional relationships to yield are developed. Examples of data and information sources not used in the baseline models were Landsat, soil survey, soil moisture models, pest models, crop calendar models, and nitrogen use models.

3. To obtain a model that uses variables more directly related to yields than the ones used in the baseline models to provide estimates that are more responsive (and correspondingly more accurate) to actual yield fluctuations.

The objective of comparative testing and evaluation of yield models was to characterize the probable performance of candidate models relative to the LACIE baseline. The minimum requirements to meet this objective were considered to be the following.

1. Comparison of overall accuracies of yield predictions for an independent set of test data

2. Validation of model-generated prediction errors (variance)

3. Identification of model strengths and weaknesses relative to the specific objectives given previously

Also, it was considered desirable to periodically examine model development to assess additional research requirements and likelihood of success. This latter objective was one of the responsibilities of the yield procedures advisory group (see the paper by McCrary et al.).

TECHNICAL APPROACH

The overall approach was to first evaluate the existing yield models that could potentially meet the research objectives. The candidate models initially

found were developed by Baier (ref. 1) and Haun (ref. 3), and contracts were awarded to Kansas State University (KSU, Feyerherm) and Clemson University (Haun) during Phase I for their evaluation and adaptation to LACIE requirements. Additional research for the development of Landsat-spectral-yield relationships was begun at Texas A & M University (TAMU, Harlan, ref. 11), the Environmental Research Institute of Michigan (ERIM, Colwell and Suits, ref. 12), and KSU (Kanemasu, ref. 4). As new problems appeared, other research or evaluation was initiated. The major efforts are described in separate symposium papers and other publications; only a summary of the individual approaches is given here.

The Baier model consists of a product of regression functions for three meteorological variables with each function containing 12 coefficients to accommodate polynomial weights for stage of development. Fitting to experimental yields was accomplished with an iterative procedure, and the coefficients and model software were supplied by Baier. After it was found that the model would have to be rederived (see results section) and difficulties with the fitting algorithms were encountered, the model was abandoned. The yield equation eventually reached a linear regression form and became known as the Feyerherm model (see the paper by Feyerherm and Paulsen). An overview of the fundamental model components, as described in the appendix, is given in figure 1.

The unique feature of the Haun model is the submodel of a growth-development index based on daily observations of relative leaf size and leaf numbers for plants in experimental plots. Maximum and minimum temperatures and a Thornthwaite estimated soil moisture parameter are then integrated through models for predicting the indexes. The indexes, along with preseason precipitation, formed independent variables for regression against county yields.

EarthSat (ref. 2) used submodels to estimate daily precipitation, modified Penman potential evapotranspiration (PET), soil moisture, actual evapotranspiration (ET), and crop calendars at 22.5- by 22.5-kilometer (12.5- by 12.5-nautical-mile) cells from first-order stations and meteorological satellite inputs. Spring wheat yield for each cell was predicted with a linear regression equation containing a linear trend term (year number) and a squared moisture stress term ($1 - ET/PET$) averaged from planting to ripening. Coefficients in the equation were estimated

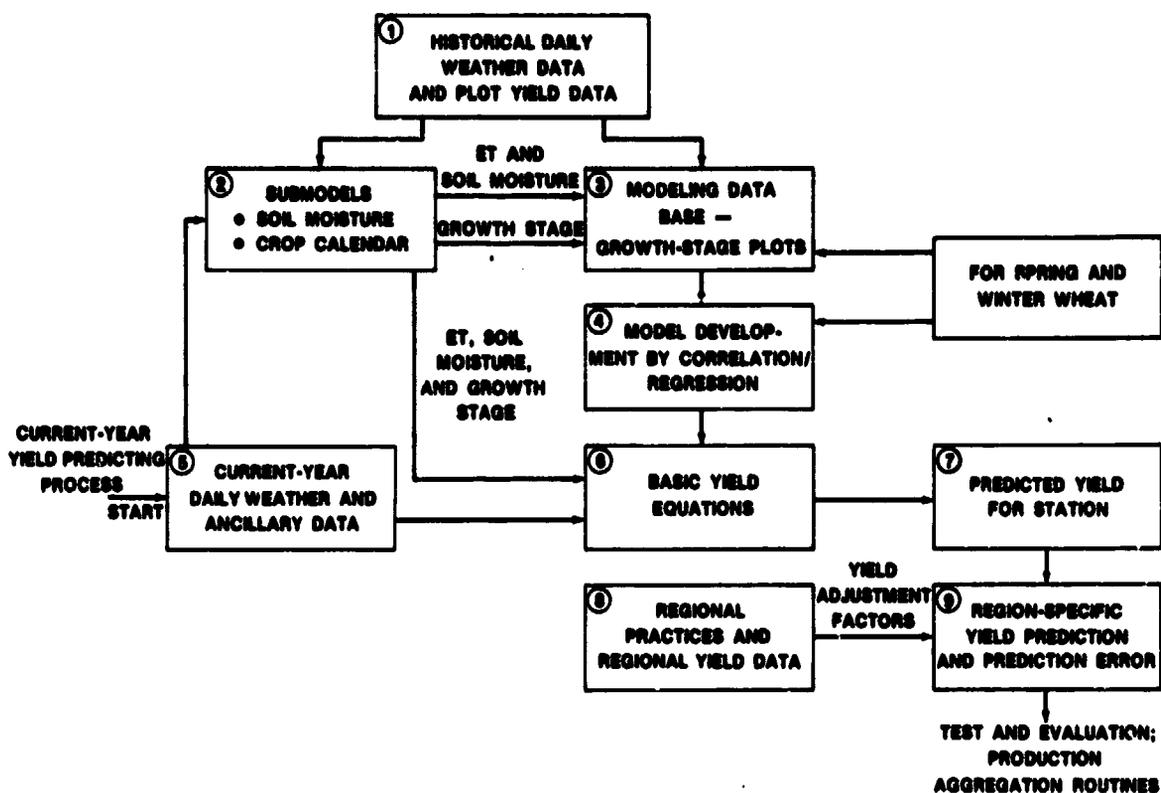


FIGURE 1.—Description of the fundamental components of the Feyerherm wheat yield models.

from 22 years of historical data from Williams, Burleigh, and Cass Counties, North Dakota. Stratum (county, crop district, and state) yield values were obtained by direct averaging of cell yields.

During Phase II, developmental efforts were concentrated on the Feyerherm model. As it became evident that additional revision of the Feyerherm model was necessary, new exploratory efforts were initiated in each of the cooperating agencies during Phase III. Also, an additional effort was made to develop error propagation procedures for the Feyerherm model, since that requirement was not included in the model development contract.

The approach explored by the National Oceanic and Atmospheric Administration (NOAA) was to use crop-district-level historical data and to define weekly weather variables on a crop time scale for use in a regression model (see the paper by LeDuc). Palmer water balance functions were used to generate the moisture supply variable. Cultural practices, such as irrigation, fallowing, varieties, and fertilization, were considered as variables to explain trends. North Dakota was used to evaluate the feasibility of this approach.

A wheat yield modeling team was formed in the Science and Education Administration (SEA) of the U.S. Department of Agriculture (USDA). The proposed approach would ultimately equate yield to its morphological components (heads per acre times kernels per head times weight per kernel). Each component would contain empirical functions of weather, agronomic data, spectral reflections (Landsat), and other factors derived by submodels. The model development was designed to project beyond the time frame of LACIE and include the collection of detailed experimental field data. Initially, a field study on winterkill was conducted to support LACIE and model development.

Cate et al. proposed during Phase III that the Law of the Minimum (Liebig) be used to relate the effects of some variables to crop yields and tested algorithms for obtaining such functional relationships (see the paper by Cate et al.). The basic theory is that yield is determined separately (without substitution) by the value of the individual variable in the involved set which is most limiting. The capability to add another variable, which has a known relationship to yield when it is the limiting

factor, without changing existing functions is considered a major advantage.

Through analysis of many data sets, including those used by CCEA and Feyerherm, the proposed Law of the Minimum model currently exists in terms of submodeled variables corresponding to synthesis and loss of yield matter. To demonstrate the concept, a model with synthesis based on a single variable and with coefficients derived from research results was used to generate a set of spring wheat test predictions. With this approach, a successful model could be considered akin to third-generation models.

Attempts to determine which yield-related crop parameters are "viewed" in Landsat data were made in both theoretical and field indicator studies. At ERIM, crop growth and bidirectional reflectance simulation models were used to infer spectral-yield relationships (ref. 12). TAMU acquired helicopter-based field spectrometer measurements in the LACIE intensive test sites and subsequently analyzed the data for yield information (ref. 11). Other correlation studies between intensive test site yields and Landsat data were performed at the NASA Johnson Space Center (JSC).

Two approaches for using Landsat-based predictions of leaf area were investigated by Kanemasu (ref. 4). One was an attempt to improve on the traditional evapotranspiration yield models by separating transpiration from soil evaporation with the aid of satellite-estimated leaf areas. Secondly, leaf area estimates were used to compute light interception by the crop and corresponding growth (net carbon exchange from photosynthesis and respiration). Adjustment factors applied to grain yield (head weight) could then give yield.

Comparative evaluations were based primarily on the statistical analyses of independent test predictions by the yield models (candidate and baseline) and their departures from USDA values. Predictions for 10 or more independent years were the preferred test data since year-to-year yield variability is the target of interest. Mean differences and mean squared differences were subjected to the paired t-test for bias and to the Wilcoxon paired rank test for relative accuracies. Ratios of modeled to observed prediction errors (variances) were compared with the standard F-test. Details of the Wilcoxon paired rank test are given by Snedecor and Cochran (ref. 13) and by Seeley et al. (ref. 14).

To evaluate models which could not be operated retroactively to obtain test replications over years, time-series-based yield estimates for separate area

units were used as an alternative baseline. For example, in order to compare the prediction to the corresponding values from the LACIE models in a case with only 1 year of test predictions, it would have to be assumed that the model-propagated error (accuracy) estimates are valid and are independent between regions. Conversely, comparing yield predictions from a universal agromet model to those from area-specific time-series estimates is valid only outside areas used to derive coefficients in the agromet model. Time-series estimates are defined as projections on trend lines fitted to yields reported by the USDA for years prior to the test case. Additional tests to evaluate error characteristics and model responsiveness based on trend projections are being documented by Stuff and Houston.

Various other analytical and sensitivity analyses were used separately or in conjunction with the previously described tests to evaluate a model's responsiveness to actual inputs as well as to systematic and random errors in the inputs. A sensitivity analysis was done by Hildreth (ref. 15) using the Feyerherm model.

RESULTS

Test runs of the Baier model using U.S. data produced major divergences between predicted and actual (plot) yields or realistic daily contributions to yields. Especially erratic results were obtained when winter wheat yields were estimated using data from Kansas. The erratic results were first attributed to the fact that input data were outside the range of that used to develop the model; however, censoring inputs to the developmental ranges provided little improvement. Separate evaluation of the soil moisture submodel gave acceptable results; however, the crop calendar submodel was found inadequate for winter wheat. After collecting the data set for rederivation of model coefficients, it was decided that the functional form of the yield model could be improved; thus, the Feyerherm model evolved.

Efforts to adapt and upgrade the Haun model indicated that several improvements were necessary or possible. Variables representing effects of postheading conditions, submodels for predicting planting and emergence dates, and variables representing technological trends were considered the major weaknesses and limitations. Also, the model had not been developed or calibrated for winter wheat. A plan to collect new data from experimental fields

over a wide variety of wheat-growing areas was initiated by Haun.

The EarthSat model for spring wheat was evaluated during Phase II. Because of the model's requirement for meteorological satellite data, only a current year (1975) of test predictions could be generated, and no model-propagated values of prediction accuracy were formed. Thus, predictions for Montana, North Dakota, South Dakota, Minnesota, and their counties and crop districts were compared to time-series models as an alternative predictor of the actual USDA (preliminary) values. Linear trends based on 1948-74 crop district data were added to the base yields per county given by Larson and Thompson (ref. 16) to derive the time-series estimates. Ratios of root mean squared error (RMSE) indicated that the time-series projections were 57 percent more accurate than the model at the county level and 77 percent more accurate at the district level (table I).

A map of relative errors for the EarthSat county predictions ($(Y - Y_{USDA})/Y_{USDA}$, fig. 2) shows that errors became larger and more erratic with distance from the counties used to calibrate the model. Simple least squares comparisons between predicted and reported yields (table I and fig. 3) indicated that the predictions were independent of the USDA values. (The hypothesis that slope is equal to 0 is not rejected at the 0.05 level of probability.)

Clearly, the EarthSat model did not meet the second-generation model accuracy objective. By comparison, the LACIE models predicted about 70 percent more accurately than the time-series models, (i.e., $RMSE_{LACIE} = 2.0$ for these four states in 1975). Some of the critical deficiencies were con-

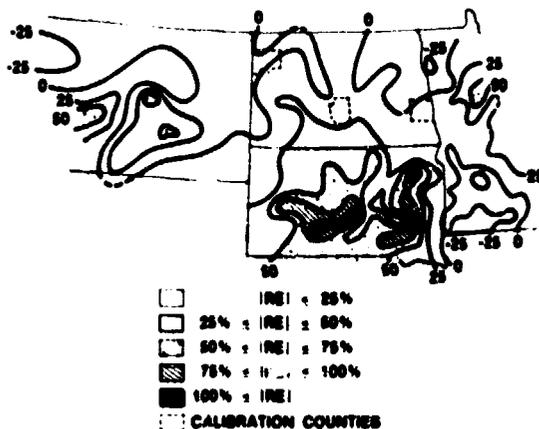


FIGURE 2.—Relative error (RE) for the EarthSat model predictions of 1975 spring wheat yields.

TABLE I.—Summary Statistics for Comparing 1975 Spring Wheat Yields (USDA) to Predictions by EarthSat Model and Normal Yields

Model	Statistic	County	District	State
EarthSat	Correlation coefficient	0.27	0.45	0.45
	Regression coefficient (slope)	.08	.15	.21
	Mean difference	2.0	2.1	.8
	RMSE	7.0	5.8	4.3
Time series	Correlation coefficient	.75	.90	.94
	Regression coefficient (slope)	.55	.74	.85
	Mean difference	3.1	3.4	3.2
	RMSE	5.6	4.3	3.6
Comparison	$MSE_{EarthSat} / MSE_{Time\ series}$	1.57	1.77	1.38

sidered to be allowances for the differential effects of stress during the season, temperature effects, and soil fertility variability. The preparation of data for some of these improvements was begun by EarthSat before completion of the evaluation.

The lack of response in the EarthSat model at the county and district levels provides an example of imbalance between detail in the model and its application scales as discussed in the appendix. Also, a separate analysis of predictions of precipitation showed that the meteorological submodel was not as accurate as the first-order stations for predicting precipitation at cooperative meteorological stations.

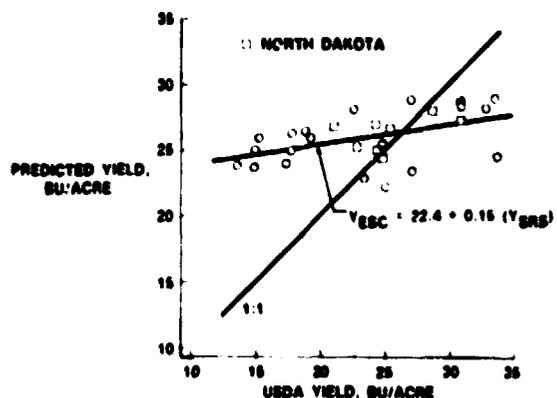


FIGURE 3.—Comparison of USDA reported and EarthSat model predicted 1975 district yields.

The Feyerherm model was compared to the LACIE baseline during Phase III. Ten years of model predictions were made at the crop district level and were aggregated to LACIE pseudozones using USDA acreages (see the paper by McCrary et al.). Statistics for individual pseudozones and aggregation by wheat types are given in table II. The Wilcoxon statistic calculated for crop types was 0.05 and 1.78 for spring and winter wheat, respectively. The nonsignificance at the 0.05 level of probability does not reject the hypothesis that the accuracy of the two spring wheat models was equal.

A restriction on the equality of the test statistics because of different meteorological station densities should also be noted. The use of the denser cooperative station network in the calculation of test yields for the CCEA model can be expected to significantly increase estimation precision (lower RMSE) in some cases. Since the cooperative station data are not available on a real-time basis, the Feyerherm ac-

curacy more correctly simulates operational conditions.

The approximate equivalent accuracy obtained by the Feyerherm model was considered a significant technical demonstration of the concept of universality in second-generation models. To evaluate the extendability of the models to a greater extent, test runs were made for several areas outside the U.S. Great Plains. Comparisons of these predictions with USDA values are given in table III and figure 4. The results are comparable to those obtained for the U.S. Great Plains. However, since 4 to 10 years of historical data are used to adjust the model in each region, the model should be considered only quasi-universal.

The sensitivity analyses by Hildreth (ref. 15) showed that yield predictions by the Feyerherm model were stable with respect to all variables in the yield equations. A large temperature sensitivity in the biocalendar estimates of stage 2.5 for winter wheat suggested that another stage may be more ap-

TABLE II.—Results of the 10-Year (1967-76) Comparative Test of the Feyerherm and LACIE Phase III Yield Models^a

Pseudozone	LACIE Phase III (b)		Feyerherm (c)	
	Bias	RMSE	Bias	RMSE
<i>Spring wheat</i>				
Montana	0.6	2.2	0.1	2.6
North Dakota	1.2	2.9	1	2.5
Red River	1.4	4.0	9	2.7
Minnesota	.6	3.8	2.6	5.6
South Dakota	.8	3.0	.9	5.0
Total	1.0	2.6	4	2.1
<i>Winter wheat</i>				
Montana	0.3	2.7	0.1	2.2
Badlands	.1	4.6	2.0	4.6
Nebraska	2	2.9	2.2	4.5
Colorado	8	3.4	1.6	4.6
Kansas	3	3.4	9	3.5
Oklahoma	1	2.2	1.0	2.2
Oklahoma-Texas Panhandle	.5	2.7	8	3.6
Texas	.5	2.3	1.3	2.2
Total	1	1.9	6	2.4

^aModel predictions relative to USDA reported yields.

^bBased on cooperative meteorological station network.

^cBased on synoptic meteorological station network.

TABLE III.—Preliminary Tests on Extending the Feyerherm Winter Wheat Model

Region	Extent	Bias, bu/acre	RMSE, bu/acre
U.S. Great Plains	7 states, 10 years	0.6	2.4
U.S. Corn Belt	5 states, 12 years	.8	4.4
U.S. Pacific Northwest	3 states, 12 years	1.0	4.5
India	5 states, 4 years	0	2.3
U.S.S.R.	1 oblast, 10 years	.6	3.4

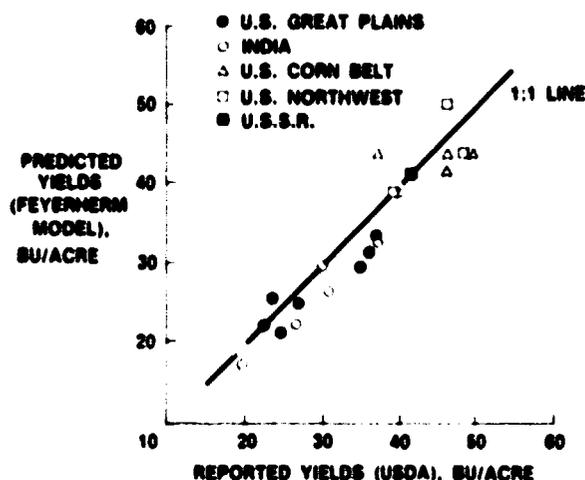


FIGURE 4.—Comparison of mean predicted and reported winter wheat yields for strains used to test the Feyerherm model.

propriate for defining yield variables. The temperature and precipitation responses were reasonable.

Deficiencies in the modeling data base appeared to be a major limiting factor in the performance of the Feyerherm model. In particular, the data contained few observations of actual crop calendars and no measurements of soil moisture or other soil parameters. The use of submodels to generate these data (fig. 1) can introduce large errors in any given case. Also, in some cases, the meteorological stations were considerable distances from the yield plots; therefore, considerable "noise" would be introduced in the precipitation values and soil moistures assigned to the plots.

The Cate-Liebig exploratory spring wheat model was run for 10 years of independent test cases. The mean squared errors (see the paper by Cate et al.) were generally smaller than those obtained with the LACIE or Feyerherm models, but the difference was not statistically significant according to the Wilcoxon test ($z = 1.17$ versus $z(0.05) = 1.645$). The practical significance of the results was the predictive power illustrated by the two variables, the equivalence between a model coefficient and experimental results, and the applicability of the model to an extended region.

The CCEA-II North Dakota prototype model was evaluated initially by examining test predictions for years in which the LACIE model registered major errors. Comparisons indicated that these errors with the CCEA-II model were smaller than with the LACIE models (see the paper by LeDuc).

Examples of simulated leaf areas and bidirectional spectral reflectance in the infrared (IR) and red Landsat channels are given in figure 5. The theoretical grain yields were based on the assumption that yield is proportional to the duration of green leaf area after crop heading (ref. 12). Also, the yield relationships were constrained by other assumptions necessary for the crop-growth model and the lack of data to test underlying assumptions or submodels.

In practice, it was found that Landsat-predicted leaf areas and measured leaf areas correlated with coefficients between 0.5 and 0.8 when ranges over the entire season were involved. Comparisons of various transformations and channel combinations of Landsat data gave approximately the same results. Sample models are given by Kanemasu (ref. 4) and results are reproduced in figure 6.

Tests performed in conjunction with the development of the Kanemasu transpiration model showed

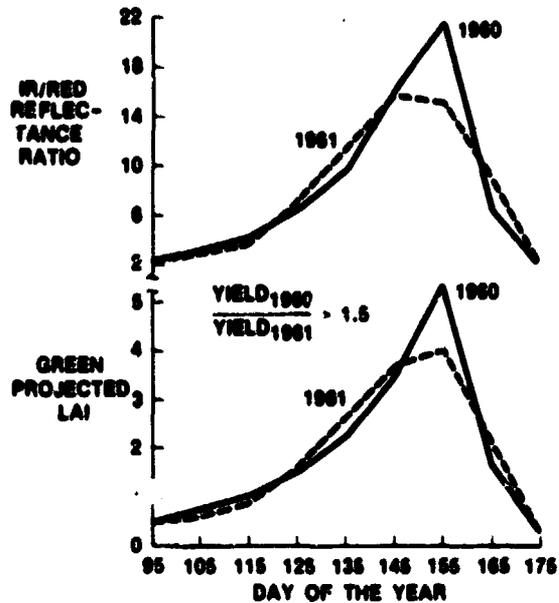


FIGURE 5.—Simulated crop and spectral values using ERIM growth and vegetation reflectance models.

significant correlation between actual and predicted yields on a field basis. No aggregated results were derived to compare the procedure to a time-series of other model. Exploratory versions of the Landsat leaf area growth model indicated that a considerable amount of work remains.

Other indicator studies based on correlations between Landsat and crop data showed that reflectances correlated as well with yield as did other crop parameters. Sample correlations taken from Thompson (ref. 17) are given in table IV. The typically lower correlations between the intermediate crop parameters (ground cover in this case) and yield have been interpreted to indicate that Landsat data measure multiple yield-related factors (ref. 18).

TABLE IV.—Landsat^a Crop Data Correlation Coefficients

Sample	No. of fields	Plant height	Ground cover	Yield detrac-tants	Estimated yield
A	30	0.28	0.26	-0.13	0.45
B	23	.45	.77	.02	.75
C	23	.25	.70	.16	.73
Combined ^b	76	.54	.77	.02	.62

^aSouth green number

^bCorrelation between estimated yield and ground cover is 0.50

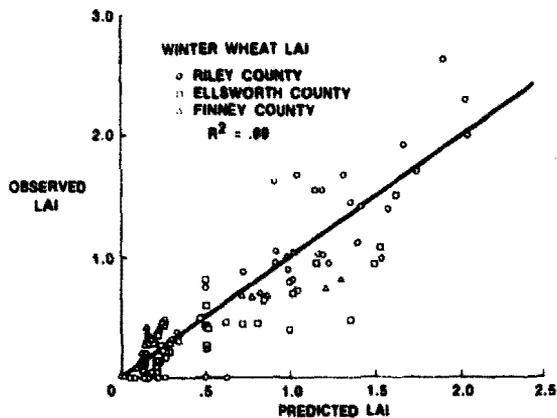


FIGURE 6.—Comparison of observed and predicted LAI using models developed by Kanemasu (ref. 4).

CONCLUSIONS

The most notable overall result of second-generation yield model RT&E was that none of the models evaluated performed significantly better than the LACIE baseline models. Because of the limitations of the regression models described in the introduction, it could be argued that equal accuracy by the Feyerherm or Cate-Liebig models in the U.S. Great Plains would represent more accuracy than the regression models in foreign areas. However, the objective of significantly better performance than the LACIE models in the U.S. Great Plains “yardstick” region is valid because of the uncertainties associated with foreign yield data used in comparative testing.

Model-building data inadequacies could be identified as a limiting factor in each of the models evaluated. The scope and quality of the data bases appeared to be more important in second-generation than in first-generation models. Certainly, any of the models could be improved with a better data base, but the relative accuracy of the resulting models still cannot be projected.

The Feyerherm and Cate-Liebig models provided estimates of the degree of universality that may be possible with second-generation models. Although they provide evidence for universal weather-yield relationships, they indicate that a few (4 to 10) years of historical data may be required to adjust the relationships to local conditions.

If the second-generation yield models were compared to biological entities, estimates of their relative development within the LACIE conditions and time

span would be as summarized in figure 7. Although the Feyerherm model reached the most advanced status, it is likely that it would be rated only at an infantile or juvenile level for an ideal data base. In terms of evolutionary potential, each model is considered embryonic.

In addition to the adequacy of data bases for building second-generation models, other technical issues remain unsolved. For example, the optimum area representation has not been determined for any particular model. The extent to which weather information should be averaged or sample sizes assigned to data from weather stations or satellites should also be addressed. The potential for early-season yield predictions using serial correlations (both yield and related factors in space and time), economic factors, and pre-season conditions was not assessed relative to weather uncertainties in different parts of the crop season. The relationships between expected yields, areas not harvested, and classification errors are factors to be investigated. Certainly some direct accounting should be made for biological plagues (diseases, insects, and weeds) in second-generation models, even though eventually they may be assessed mainly by satellite variables.

It is strongly recommended that a set of criteria or prerequisites be developed for screening models in LACIE follow-on research or test programs. Partially as a result of LACIE, there are now more candidate models and several aspects should be considered in their screening. The scope and quality of the data base should be reviewed, as well as factors in the model. Internal procedures for propagating errors or variances for the predictions should be included in the model development, and basic indications of model competence should be provided.

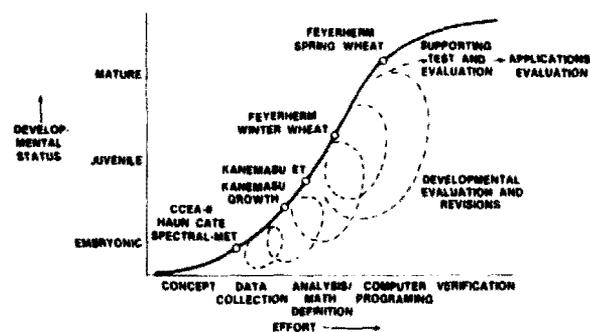


FIGURE 7.—Schematic diagram of yield model development and relative status for LACIE data sets.

Appendix

Description and Relation of Second-Generation to Other Yield Model Terminology and Concepts

When considering the universe of yield-determining factors—from the within-field variability of soil, chemical, and physical properties; the genetic variability in a crop species; and the cultural and management options employable by each different farmer to all the possible combinations of daily weather; the dynamics of insect, weed, and disease populations; and the variable and unforeseen capabilities of particular fields of plants to adapt—the number is, for all practical purposes, infinite. A first-order breakdown of these factors is illustrated in figure 8. To mathematically relate observations of a finite number of variables to yield is by definition a statistical problem or abstraction. Even yield-determining biochemical processes that can be described by deterministic equations require stochastic parameters to appropriately quantify their rates, inputs, or outputs. Consequently, numerous approaches have been proposed for estimating crop yields from information on one or more of the related factors at different levels of biological detail, functioning, and scope.

Crop prediction models have been classified according to various criteria but a standardized taxonomy is not apparent. Stanhill (ref. 19) divided general approaches into statistical or experimental (controlled environments, simulation, etc.), and Baier (ref. 20) added an intermediate category which he called weather analysis models. With more basic criteria such as the nature of functional relations between variables, models are commonly classified as empirical or mechanistic (ref. 21). Other divisions used jointly or sequentially with the above are stochastic or deterministic (influence of probability parameters on inputs or outputs), analytical or numerical (equation solving methods), continuous or discrete (degree of continuity in possible variable quantities), and dynamic or static (dependence of model components on time coordinates).

Models frequently are described as physiological or phenomenological and by other mainly subjective terms used to indicate the physical abstractness and/or biological hierarchy involved—not model validity. A mechanistic simulation model may contain environmental variables, “controlling” rates of

photosynthesis, respiration, etc., but be physiologically less valid than an empirical statistical model. An example is found in moisture functions. Three empirical growth-moisture relationships are given in figure 9. When water concentration of the root environment is unconditionally increased, detrimental levels are reached because of the exclusion of oxygen to the roots (rice is an exception since it can internally transport oxygen to the roots from the atmosphere); thus, curve B should be observed in cases where precipitation exceeds soil holding capacities. Two separate functions representing the positive effect of increased moisture and the negative effect of root asphyxiation would be the most physiologically correct. Yet, curve-A-type functions are frequently used in mechanistic simulation models (ref. 22), and those illustrated in figure 10 commonly occur in empirical statistical models. Clearly, the empirical statistical model using the more abstract precipitation variable (more removed from plant moisture than soil moisture) may be physiologically more valid than the mechanistic simulation model that does not account for asphyxiation or other negative effects of excessive moisture (such as associated disease, hail, and lodging damage). It is safe to assume that a physical/physiological explanation can be found for any empirical yield model which has predicting

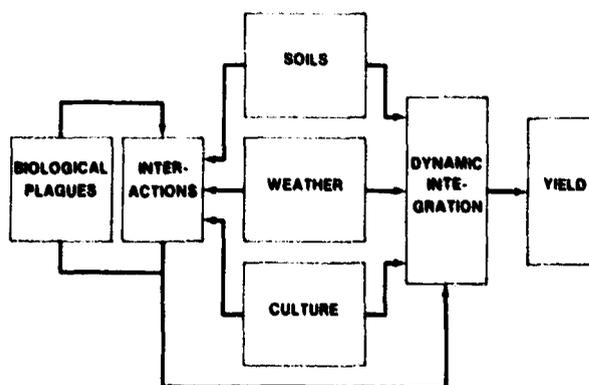


FIGURE 8.—First subdivision of yield factors.

capability; and, conversely, a mechanistic simulation model completely designed around direct physiological processes may not have any yield-predicting capability.

In LACIE, yield models were classified as first, second, or third generation according to their readiness for implementation or resources necessary to achieve readiness. Since readiness depended on status of model development, availability of input re-

quirements, comparative testing, etc., the LACIE classifications closely paralleled the three categories defined by Baier (ref. 20). Some general characteristics associated with the three classes of models are summarized in table V.

A key factor in development and operational costs is the amount or level of detail that a model is designed to capture. Corresponding to the increased level of detail included in second- and third-genera-

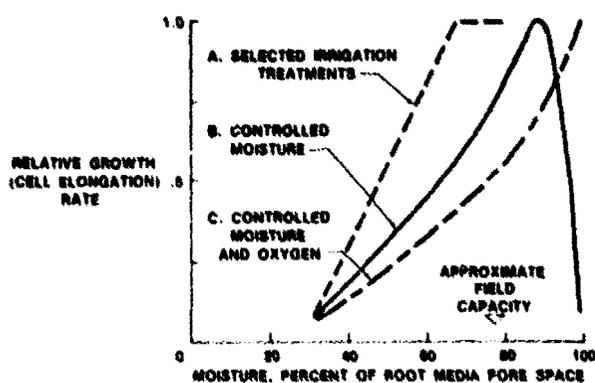


FIGURE 9.—Characteristic plant responses to moisture supply observed under different experimental conditions.

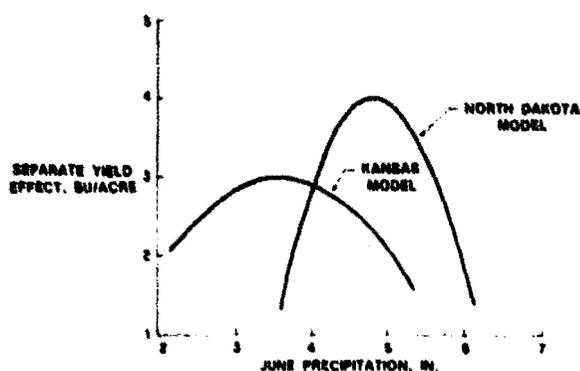


FIGURE 10.—Examples of nonlinear moisture effects found in LACIE Phase II yield models.

TABLE V.—General Characteristics of Model Types

General characteristics	First generation	Second generation	Third generation
Level of detail	Arbitrary subdivisions of ecosystems	Field	Plant to biochemical
Overall structure	Holistic	Modular	Modular
Physiology	Implicit relationships	Recognized relationships	Explicit relationships
Soils	Constant term or differential equations	Water-holding capacities and strata factors	Explicit variables and relationships
Management/technology	Static coefficients	Constants and explicit variables	Explicit variables and relationships
Seasonal time/integration	Calendar static coefficients	Biological phase weighting (Static)	Biological/dynamic growth
Examples	Thompson LACIE/CCFA Fisher polynomial Hanks ET Stress indexes	Feyerherm CCFA-II Cate-Liebig Haun Kanemasu ET USDA SFA	SIMAIZE Kanemasu growth GOSSYM SOYMOD SORGIF

tion models are general differences in the spatial and temporal scales that their inputs and outputs typically represent. The general scale ranges covered by models in each category are sketched in figure 11. The most cost-effective scales for a first-generation model should be the largest level of aggregation at which the temporal and spatial variations of acreages and yield-determining factors do not cause model predictions to exceed the desired tolerance limit. If a sufficient amount of the year-to-year yield variability is caused by the factors operating at smaller scales, success in yield estimation will depend on having such a level of detail built into a model. Given a detailed third-generation model of known precision for fields or groups of fields, the application scale issue is one of determining the tolerable sampling error and feasibility of the required sampling.

A third axis could be added to figure 11 depicting the level of biological detail from a total crop

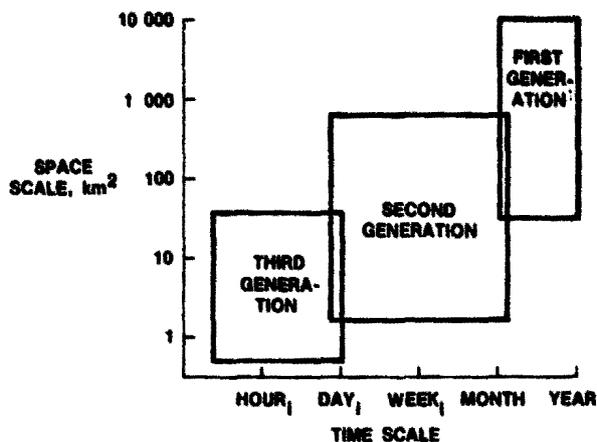


FIGURE 11.—Relationship of yield model types to input/output scales.

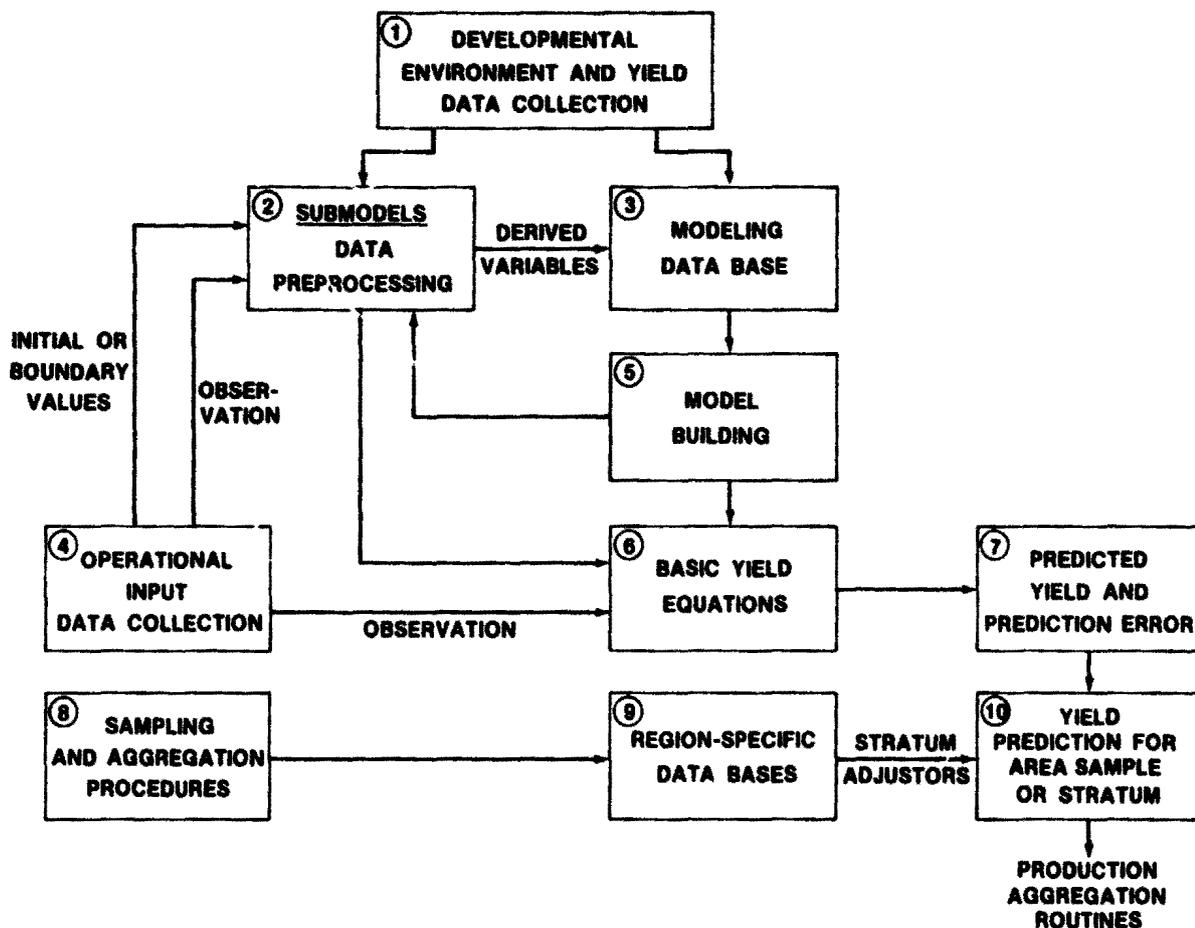


FIGURE 12.—Fundamental components in yield modeling for production estimation.

ecosystem to molecular levels with approximately the same distribution of model types. Stanhill (ref. 19) related meteorological phenomena to time and space scales. Likewise, a hierarchy of soil classification detail could be identified for the spatial scale.

The fundamental components involved in models for yield estimation are relationally illustrated in figure 12. There may be salient or subtle differences in the nature of these components for each model type which could be used as a classification criterion. Corresponding descriptors for the three classes of models are summarized in table VI. For example, second- and third-generation models use progressively more submodels to generate variables that are more directly related to yield-forming processes than first-generation models. A third-genera-

tion model designed to use conventional data should have submodels to estimate canopy structure; to derive subcanopy values of environmental variables; and to estimate nitrogen application, weed, insect, and disease development, photosynthesis, respiration, and translocation for any given field or sample "points" in a crop stratum.

Generally, in second-generation models, a heuristic approach which combines biological theory with empirical results is used to define the cause-effect relationships between yields and available information. The LACIE objective of universality allows the models to be less dependent on historical data; however, some actual data are still assumed to be necessary for the local conversion from "modeled" to normally harvested yields.

TABLE VI.—Descriptions of Fundamental Components for Three Types of Models

Fundamental component (a)	First generation	Second generation	Third generation
Data source	① Survey (historical) for specific geographical areas	Survey or field experiments	Controlled environments, designed experiments
Submodels	② None or simple indexes	Soil moisture, crop development	Canopy, micrometeorology, soil moisture, crop development, physiological processes, morphological components, etc.
Input data	③ Average values for discrete periods (weeks, months, or season)	Daily or weekly environmental measurements	Initial conditions for soils, crops, and management variables—hourly to daily values of environmental variables
Model building	④ Intuitive definition and/or statistical selection of variables Statistical fitting of coefficients	Heuristic definition or variables and cause-effect relationships Empirical results and statistical fitting	Mathematical fit to empirical results Theoretical relationships, differential equations
Yield equation	⑤ Simple algebraic	Simple algebraic	Simple algebraic to extended series of differential equations
Strata adjustments	⑥ Dependent on data base	Adjustments for various soils management or other unique strata influences required	Theoretically not required

^aNumbers are cross-referenced to figure 12.

REFERENCES

1. Baier, W.: Crop-Weather Analysis Model: Review and Model Development. *J. App. Meteor.*, vol. 23, 1973, pp. 937-947.
2. Earth Satellite Corporation. EarthSat Spring Wheat Yield System Test, 1975. Final Contract Report 1052 (Washington, D.C.), 1976.
3. Haun, J. R.: Prediction of Spring Wheat Yields From Temperature and Precipitation Data. *Agron. J.*, vol. 66, 1974, pp. 405-409.
4. Kanemasu, E. T.: Estimated Winter Wheat Yield From Crop Growth Predicted by Landsat. Final report, Contract NAS 9 14899-35, NASA Johnson Space Center, Houston, Tex., 1977.
5. MacDonald, R. B.: LACIE Project Plan. Baseline Document C00605, NASA Johnson Space Center, Houston, Tex., 1975.
6. Sanderson, F. H.: *Methods of Crop Forecasting*. Harvard Univ. Press, 1954.
7. Hildreth, W. W.; and Barger, G. L.: Wheat Phenology: Analogous Region Analysis. Lockheed Electronics Co., Inc., Tech. Memo 644-816, Houston, Tex., 1976.
8. Pitts, D. E.; and Barger, G. L.; eds.: Proceedings of the 1974 Lyndon B. Johnson Space Center Wheat-Yield Conference. NASA TM X-58158, 1975.
9. Haun, J. R.: Wheat Growth/Environment/Yield Relationships. Proceedings of the 1974 Lyndon B. Johnson Space Center Wheat-Yield Conference. NASA TM X-58158, 1975.
10. Robertson, G. W.: Wheat-Yield Estimates Based on Weather: Research and Applications in Canada. Proceedings of the 1974 Lyndon B. Johnson Space Center Wheat-Yield Conference. NASA TM X-58158, 1975.
11. Harlan, J. C.: On Winter Wheat Yield From Landsat and Landsat Follow-On Satellites. LACIE-00468, NASA Johnson Space Center, Houston, Tex., Mar. 1978.
12. Colwell, J. E.; and Suits, G. H.: Yield Prediction by Analysis of Multispectral Scanner Data. Final Report, ERIM 109600-17-F, ERIM (Ann Arbor, Mich.), 1975.
13. Snedecor, G. W.; and Cochran, W. G.: *Statistical Methods*. Iowa State Univ. Press (Ames, Iowa), 1976.
14. Seeley, M. W.; Phinney, D. E.; Cate, R. B.; and Trenchard, M. T.: Comparative Evaluation of Wheat Yield Models in LACIE. Paper presented at American Society of Agronomy Annual Meeting (Chicago), 1978.
15. Hildreth, W. W.: Sensitivity Analysis of Feyerherm's Wheat Yield Models: Temperature and Precipitation Effects on Yield. Lockheed Electronics Co., Inc., Tech. Memo 642-2439, Houston, Tex., 1977.
16. Larson, K. D.; and Thompson, L. S.: Variability of Wheat Fields in the Great Plains. USDA Economic Research Service, Tech. Pub. No. 287 (Washington, D.C.), 1966.
17. Thompson, D. R.: Monitoring Growth and Yield Components by Landsat. LACIE-00468, NASA Johnson Space Center, Houston, Tex., Mar. 1978.
18. Erickson, J. D.: Feasibility of Assessing Crop Condition and Yield from Landsat Data. LACIE-00468, NASA Johnson Space Center, Houston, Tex., Mar. 1978.
19. Stanhill, G.: Quantifying Weather-Crop Relations in Environmental Effects on Crop Physiology. Proceedings of the Fifth Long Ashton Symposium, Bristol, J. J. Landsberg, ed., Academic Press (New York), 1977.
20. Baier, W.: Note on Terminology of Crop-Weather Models. World Meteorological Organization Expert Meeting on Crop-Weather Models (Ottawa), Oct. 11-13, 1977.
21. Thornley, J. H. M.: *Mathematical Models in Plant Physiology. A Quantitative Approach to Problems in Plant and Crop Physiology*. Experimental Botany, Vol. 8, Academic Press (New York), 1976.
22. Arkin, G. F.; Vanderlip, R. L.; and Ritchie, J. T.: A Dynamic Grain Sorghum Growth Model. *Trans. ASEA*, vol. 19, 1976, pp. 622-630.

A Universal Model for Estimating Wheat Yields*

A. M. Feyerherm^a and G. M. Paulsen^a

INTRODUCTION

The development of universal wheat yield models to show separate and joint effects of weather and culture on yields applicable to fall- and spring-planted wheat on a global basis is discussed. The model was built with the restriction of using only weather-related variables (WRV's) based on meteorological variables currently observable globally. Therefore, only daily minimum and maximum temperatures and precipitation were used.

Early in the research, it was decided to use varietal-trial-plot-yield data from state experiment stations in a wide range of climates in the U.S. Great Plains to build basic relationships among yields, weather, and culture. Discussed here are steps taken to develop the model; its application on a macroclimatic scale in the United States, the U.S.S.R., and India; and potential improvements. Added details of the work described here can be found in final reports of contract work (refs. 1 to 3). The work of Robertson and Baier with crop calendars (ref. 4), moisture budgets (ref. 5), and yield/weather modeling (ref. 6) had the greatest influence on the research efforts.

MODEL DEVELOPMENT

Model Form

Consider the problem of estimating wheat yield (production per unit area) in a given region such as a county or crop reporting district (CRD) with weather information from a station (S) and

knowledge of specified cultural practices applied in region (R). The estimate of yield for year y is designated by $MOD_y(R,S)$ and is calculated as follows.

$$MOD_y(R,S) = A(R,S) + B * WAC_y(R,S) \quad (1)$$

where $MOD_y(R,S)$ = model-estimated yield in year y for region (R) with weather at station (S)

$A(R,S)$ = a constant, calculated from historical yield and cultural data for region (R) and weather at station (S)

B = a universal constant (0.75 for winter wheat; 0.50 for spring wheat)

* = multiplication sign

$WAC_y(R,S)$ = contribution to yield of weather and culture (WAC)

More specifically,

$$WAC_y(R,S) = \sum_{j=1}^3 p_{jy}(R) * VYA_y(R) * [W_{jy}(S) + W_{0y}(S) * NI_{jy}(R)] \quad (2)$$

where, for region (R) and station (S) in year y,

$VYA_y(R)$ = a varietal yielding ability (VYA) component, which is an average of VYA values for varieties planted in year y

$p_{jy}(R)$ = proportion of wheat under cropping practice j (j = 1 = continuous; j = 2 = fallow; j = 3 = irrigated)

$NI_{jy}(R)$ = amount of elemental nitrogen applied for cropping practice j (j = 1,2,3)

*Contribution No. 78-268-A, Department of Statistics and Statistical Laboratory and Department of Agronomy, Kansas Agricultural Experiment Station, Manhattan, Kansas.

^aKansas State University, Manhattan, Kansas.

$W_{jy}(S)$ = weather-generated component of yield for wheat under cropping practice j ($j = 1,2,3$)

$W_0(S)$ = weather-generated coefficient of N

The last two quantities are mathematical functions of WRV's calculated from daily readings of precipitation and minimum and maximum temperatures. A major part of model development involved determining the relationship between wheat yields and WRV's to generate values of $W_j(S)$ and $W_0(S)$, which is discussed in the following sections.

Data Base

The weather-related yield components were assumed to be linear functions of WRV's, designated by X . Thus,

$$W_j(S) = B_0X_0 + B_1X_{1j} + \dots + B_rX_{rj}; \quad j = 1,2,3 \quad (3)$$

The coefficients denoted B in equation (3) are universal constants with separate sets for fall-planted (winter) and spring-planted (spring and durum) wheat. Likewise, $W_0(S)$ was assumed to be a linear function of WRV.

Plot data from intrastate and regional nurseries for varietal trials and meteorological data (daily precipitation and minimum and maximum temperatures) from nearby weather stations provided the basic data to estimate the values of B in equation (3) and in $W_0(S)$ and to calculate VYA values for specific varieties. Plot data included yield by variety, cropping practice (continuous or fallowed), and amount of elemental nitrogen applied. In addition, planting and heading dates from the plots were used to calibrate crop calendars.

Plot data were secured from 64 state agricultural experiment stations in the major wheat-producing states of the U.S. Great Plains and the Eastern Great Plains. Estimates of the winter wheat coefficients (B -values) in equation (3) were based on 1034 location-years; those for spring wheat, on 306.

Once the universal constants (coefficients) in equation (3) and $W_0(S)$ were estimated, it was possible to generate historical values of $W_j(S)$ ($j = 1,2,3$) at any weather station. The $W_j(S)$ values were combined with historical data of the U. S. Department of

Agriculture (USDA) Statistical Reporting Service (SRS) and Economic Research Service (ERS) on varieties planted, proportion of wheat under the three cropping practices, and amount of nitrogen applied to generate $WAC(R,S)$ values, by year, for CRD's in the U.S. Great Plains. The historical WAC values were combined with USDA SRS estimates of yield, first to estimate the universal constant B and then to estimate $A(R,S)$ values in equation (1).

Standardization of Yields and Crop Calendars

To relate yield variation using many varieties to weather variation using data over a wide range of climates, Kansas State University (KSU) had to adjust yields to a "standard" variety and to calculate WRV's over common phenological phases (e.g., jointing to heading) rather than coincident calendar days (e.g., April).

To accomplish the first task, KSU developed VYA values for varieties that became popular with producers. The VYA values were computed by first comparing yields of each variety with every other variety over all location-years for which data were available within regions of varietal adaptability. The VYA value for a variety (v) was an expression of the yield capability of variety (v) to that of a standard variety (s) as a ratio. The final value assigned to variety (v) incorporated not only the direct comparison of (v) and (s) but also indirect comparisons through application of a chain rule with other varieties as intermediaries. Some representative VYA values are shown in table I.

The need to identify common phenological phases in different climates was satisfied by Robertson's biometeorological time scale (BMTS) (ref. 4) for spring wheat and an adjusted form of the BMTS (ref. 1) for winter wheat. Correspondence of points on the BMTS to phenological stages is as follows.

BMTS	Stage	BMTS	Stage
0.0	P—Planting	3.0	H—Heading
1.0	E—Emergence	3.5	M—Milk
1.5	T—Tillering	4.0	D—Dough
2.0	J—Jointing	5.0	R—Ripe
2.5	F—Flag leaf		

(Robertson's BMTS included stages with whole numbers; KSU added names to 1.5, 2.5, and 3.5 to facilitate the discussion.)

A Soil Moisture Budget

Soil moisture conditions are better indicators of water supply or stress for plants than is precipitation. Baier and Robertson's versatile soil moisture budget (VSMB) (ref. 3) was used by KSU to simulate soil moisture conditions. The VSMB needs only daily precipitation and temperature extremes to operate and provides a tool to reconstruct historical soil moisture conditions from data available worldwide. In application, KSU used a plant-available water capacity of 10 inches for all seasons and locations.

TABLE I.—VYA Values of Some Representative Varieties

U.S. Great Plains region	Variety	VYA	Release date
<i>Winter wheat</i>			
Northern	Kharkof	1.00	1900
	Cheyenne	1.04	1933
	Winalta	1.13	1961
	Scout	1.25	1963
Central	Turkey	.85	1875
	Kharkof	.88	1900
	Comanche	1.00	1942
	Bison	1.07	1956
	Scout	1.22	1963
	Sage	1.23	1973
Southern	Kharkof	.81	1900
	Comanche	1.00	1942
	Triumph	1.07	1940
	Concho	1.12	1954
	Triumph 64	1.21	1964
Eastern	Trumbull	.93	1916
	Pawnee	1.00	1943
	Butler	1.09	1947
	Ben Hur	1.17	1966
	Arthur	1.35	1968
<i>Spring and durum wheat</i>			
Northern	Marquis	0.85	1907
	Thatcher	1.00	1934
	Canthatch	1.09	1959
	Crim	1.15	1963
	Wells	1.19	1960
	Era	1.42	1970

Estimation of Universal Constants

Values for the constant B in equation (3) were estimated by regressing standardized plot yields (yields divided by VYA) on the WRV's (X 's). Definitions for the WRV's are as follows.

$AE(a,b)$ = VSMB simulated evapotranspiration from stage a to stage b

$PE(a,b)$ = VSMB simulated potential evapotranspiration from stage a to stage b

$RE(a,b)$ = $AE(a,b)/PE(a,b)$

$SM(a,b;\alpha)$ = $[1 - RE(a,b)/\alpha]^+$, a soil moisture stress term

$CNT(a)$ = contents of zones 4 and 5 in the VSMB at stage a

$SSM(a;\alpha)$ = $[1 - CNT(a)/\alpha]^+$, a subsoil moisture stress term

$PR(a,b)$ = precipitation from stage a to stage b

$XPR(a,b;\alpha)$ = $[PR(a,b) - \alpha]^+$, an excess precipitation term

$TP(a,b;\alpha)$ = $PR(a,b)$ if $PR(a,b) \leq \alpha$

= α if $PR(a,b) > \alpha$, a truncated precipitation quantity

TN = daily minimum temperature

TX = daily maximum temperature

$ATX(a,b)$ = average daily maximum temperature from stage a to stage b

$ATX(a,b;\alpha)$ = $[ATX(a,b) - \alpha]^+$

$TN(a,b;\alpha)$ = average daily value of $[TN - \alpha]^+$ from stage a to stage b

$TX(a,b;\alpha)$ = average daily value of $[TX - \alpha]^+$ from stage a to stage b

JT = long-term average daily temperature for January

FL = 0 for continuously cropped wheat; 1 for wheat on fallowed ground

In the preceding definitions, $[X]^+ = X$ if $X \geq 0$, but $[X]^+ = 0$ if $X < 0$. A number of the definitions involved thresholds, designated by α ; and the values of the variable are constant for arguments either above or below the threshold values.

The entries in table II combine to express the year-to-year and location-to-location variations in yield due to meteorological variation. The signs on the B -values and their magnitudes appear agronomically acceptable. The winter wheat model reflects some known facts; namely, the deleterious effects of moisture stress, particularly from jointing to the milk stage; excess precipitation after heading; and warm temperatures throughout the season. For spring-planted wheat, warm temperatures are cer-

tainly deleterious. The moisture stress situation is not so clear-cut. The model indicates that precipitation is most beneficial if "dryout" periods occur between adequate rains.

The subscript *j* on some of the WRV's in table II indicates that the variables assume different values depending on cropping practice. The KSU computer version of the VSMB carries two moisture budgets, one for continuous wheat and one for wheat planted on fallowed soil. The beneficial effects of fallowing are expressed indirectly through the soil moisture terms and directly, in the winter wheat model, through the *FL* term.

TABLE II.—Formulas to Calculate Weather Components of Yield^a

[Eq. (3)]	
<i>B</i>	WRV (<i>X</i>)
<i>Winter wheat</i>	
50.8	1
-10.3	[SSM(D:2)] ²
-16.55[1 - 0.02 * JT]	SM(E:7:0.6)
-.37 * JT	[SM(J:F:0.8)] ²
-.15 * JT	[SM(F:H:0.8)] ²
-6.40	[SM(H:M:0.9)] ²
-1.05	XPR(H:D:4)
-.06	ATX(E:7)
-.17	ATX(T:J)
-1.01[1 + 0.23 * PR(J,F)]	TN(J,F:50)
-.40	TN(F,H:50)
-.60	TX(H,M:86)
-.47	TX(M,D:86)
.37[1 - 0.025 * JT]	FL _{<i>j</i>}
<i>Spring wheat</i>	
154.4	1
3.66	TP(P,J:3)
3.18	TP(P,M:5)
-2.45	TP(P,H:9)
-9.16	RE(T:J)
3.86	AE(H,M)
1.89	CNT(M)
-.47	ATX(J,F)
-.37	ATX(F,H)
-.34[1 + 0.11 * PR(H,M)]	ATX(H,M)
-.59	ATX(M,D)
-.29	ATX(D,R:78)

^aYield is measured in bushels per acre, temperature in degrees Fahrenheit, and precipitation in inches.

For irrigated winter wheat (*U* = 3), all *SM* and *SSM* terms are set equal to zero and *FL* = 0. For irrigated spring wheat, moisture-related variables are fixed as follows: $TP_3(P,J:3) = 3.0$; $TP_3(P,M:5) = 5.0$; $TP_3(P,H:9) = 5.0$; $AE_3(H,M) = 2.5$; and $CNT_3(M) = 4.0$.

The coefficient of *N* in equation (2) is weather related, and $W_0(S) = 0.17 + XPR(P,H:8)[0.016 - 0.007 * JT]$ for winter wheat and simply 0.09 for spring wheat. For winter wheat, a leaching effect due to excess precipitation reduces the contribution of each pound of nitrogen to yield. Here, *N* is measured in pounds per acre. That completes discussion of the universal constants in equation (2).

To estimate the constants in equation (1), *B* and *A*(*R,S*), KSU chose and matched, approximately one to one, CRD's with first-order U.S. Weather Bureau, Federal Aviation Administration, and military weather stations in the U.S. Great Plains. KSU then computed *WAC*(*R,S*) values, given in equation (2), for as many as 22 years (1955-76) for most of the region-station (*R,S*) combinations. Government-reported (USDA SRS) yields *GOV*(*R*) were retrieved for each region (CRD) and *b*(*R*) was computed as follows:

$$b(R) = \frac{\sum_{y=1}^n (Z_y - \bar{Z})(X_y - \bar{X})}{\sum_{y=1}^n (X_y - \bar{X})^2} \quad (4)$$

for each (*R,S*), where $Z_y = GOV_y(R)$ and $X_y = WAC_y(R,S)$ and the sums were over years *y*. Then, *B* was calculated as a weighted average of *b*(*R*) as follows.

$$B = \sum_{R=1}^{R_0} q(R) * b(R) \quad (5)$$

where *q*(*R*) = proportion of U.S. Great Plains harvested acres allocated to region (*R*) and *R*₀ = total regions. The *q*(*R*) value was calculated from average USDA SRS acreage estimates for 1971-75. Results gave *B* = 0.75 for winter wheat and *B* = 0.50 (0.51 before rounding) for spring wheat.

Finally, a value for the regional constant $A(R,S)$ was calculated from historical data by the formula

$$A(R,S) = \overline{GOV}(R) - B \cdot \overline{WAC}(R,S) \quad (6)$$

where the means were calculated over as many years as were covered by data. In application, only the historical means of government-reported yields are used for estimation. All other constants in the model are universal and were derived independently of the region (R) for which an estimate was desired.

APPLICATIONS

Application of the model requires values for the WRV's and the cultural variables. Values for all WRV's can be generated from daily readings of precipitation, minimum and maximum temperatures at a station (S), tabled values of Q (solar radiation at the edge of the atmosphere), and day length. The cultural variables needed for a region (R) are VYA, amount of nitrogen applied, and proportions of wheat.

Historical values for cultural variables may be more difficult to determine than are values of WRV's. However, the model is relatively insensitive to modest errors in observation, and estimates from "experts" can be used to good advantage. The model has some self-correcting capability in that consistent overestimates or underestimates of $WAC(R,S)$ values over seasons can be partly offset by the $A(R,S)$ values, which may be recalculated periodically. Further, cultural variables change slowly from year to year. It is weather that produces abrupt shifts in yield in any given semiarid region (R) from one year to the next.

United States

Results of KSU application of the model in all the major wheat-growing areas of the United States are summarized in tables III and IV. The model was applied to weather station-region (CRD) combinations with a density of less than one station per CRD, and the yields were aggregated upward to multistate results. Average acreages during 1971-75 were used as weights in the aggregation. The USDA SRS yields were aggregated upward with the same weights.

TABLE III.—Comparison of Model (MOD) and SRS Estimates of Yields for U.S. Great Plains

[Differences > 1 bu/acre are italicized.]

Harvest year	Winter wheat, ^a bu/acre		Spring wheat, ^b bu/acre		All U.S. bu/acre	
	MOD	SRS	MOD	SRS	MOD	SRS
1965	22	23	24	25	23	24
1966	24	22	22	22	23	22
1967	21	20	21	24	21	21
1968	26	24	27	27	26	25
1969	27	28	27	28	27	28
1970	27	30	23	24	25	28
1971	26	30	26	31	26	30
1972	27	29	29	28	28	29
1973	29	32	24	28	28	31
1974	28	26	20	21	25	24
1975	27	28	23	25	26	27
1976	27	27	25	25	26	26

^aSeven states: Montana, South Dakota, Nebraska, Colorado, Kansas, Oklahoma, and Texas

^bFour states: Minnesota, North Dakota, South Dakota, and Montana

TABLE IV.—Comparison of Model (MOD) and USDA SRS Winter Wheat Yield Estimates for the Eastern Great Plains and the Northwest

[Differences > 4 bu/acre are italicized.]

Year	Eastern Great Plains, ^a bu/acre		Northwest, ^b bu/acre	
	MOD	SRS	MOD	SRS
1965	32	33	36	39
1966	36	40	36	38
1967	34	36	42	39
1968	36	36	41	38
1969	37	37	42	39
1970	35	37	40	44
1971	42	44	43	44
1972	42	44	43	45
1973	39	32	39	34
1974	42	34	47	43
1975	41	39	50	46
1976	43	37	51	46

^aMissouri, Illinois, Indiana, and Ohio

^bWashington, Oregon, and Idaho

The results for the U.S. Great Plains (table III) show that, despite a sparse weather network, model-generated yields were within ± 1 bushel per acre of SRS estimates in years of lowest yields. On the other hand, the model underestimated SRS yields by 3 to 4 bushels per acre in the years of highest yields.

Table IV, for the Eastern Great Plains, shows an increase in yields from 1965 to 1972 due largely to increased nitrogen applications and the introduction of high-yielding semidwarf varieties like Arthur. The model overestimated yields in 1973 and 1974 because of a septoria epidemic and in 1976 because of late freezes at heading time. The model does not account for losses due to disease, epidemics, and late freezes. For those years, the model estimates what yields could have been without diseases and late freezes.

Application to the Northwest, shown in table IV, provides a challenging test for the model because no data from that region were used to estimate universal constants. In addition, historical data on cropping practices were unavailable for Washington and Oregon, and "rough" estimates were used in the calculations. Except for a few years, the model and the SRS estimates agree well. The overestimate for 1973 is partly accounted for by a large area of winterkill in Washington. In 1976, poor yields in Idaho reduced the average. The underestimate for 1971 probably resulted from general underestimation of high yields by the model.

U.S.S.R.

The KSU model was applied to three oblasts (states) in the U.S.S.R.; comparisons between model estimates and yields reported by the U.S.S.R. are shown in table V. The poor winter wheat yields in 1968 in the Khmel'nitskiy oblast (part of the Ukraine) were detected by the model. Although U.S.S.R. winter wheat yields were low in 1972, such was not the case in Khmel'nitskiy; the model substantiated that fact.

For spring wheat, there is a difference of 7 bushels per acre between MOD and GOV for Kurgan in 1970; a difference that large can appear with data from only one weather station for such a large region. The agreement between MOD and GOV estimates is extremely good for Tselinograd in the Kazakhstan region. Especially notable was the model's detection of a relatively high yield in 1972.

TABLE V.—Comparison of Model (MOD) and Government-Reported (GOV) Yields for Three Oblasts in the U.S.S.R.

Year	Winter wheat, bushels, for Khmel'nitskiy		Spring wheat, bushels, for—			
	MOD ^a	GOV	Kurgan		Tselinograd	
	MOD ^a	GOV	MOD ^b	GOV	MOD ^c	GOV
1965					9	3
1966			23	23	12	14
1967	45	41	22	22	8	8
1968	33	34	25	30	11	10
1969	41	43	19	24	12	14
1970	37	39	31	24	13	11
1971	44	41	23	19	14	14
1972	43	42	30	29	25	24
1973	41	43			17	15
1974	41	39				

^aBased on seven weather stations.

^bBased on one weather station.

^cBased on two weather stations.

India

The model was applied to five wheat-growing states in India for the 4 years from 1972 to 1975; comparisons of model estimates and yields reported by the Government of India are shown in table VI. The model was run with a normal crop calendar because Robertson's BMTS was not applicable at the lower latitudes. Irrigated wheat was assumed to be composed of high-yielding varieties (VYA = 1.30) with 30 pounds per acre of nitrogen applied. The universal constant B in equation (1) was set equal to 0.70 because the analysis was run before the final decision to use 0.75 was made.

With irrigation and high-yield varieties, India achieved rather uniform year-to-year yields in those five states. Table VI shows not only yields but also the proportion of wheat irrigated P_i and not irrigated P_n . The weather components ($W_{ij}(S)$ of eq. (2); $y = 1972, 1973, 1974, 1975$; $j = 1, 3$) of WAC values were averaged over the weather stations within a state and indicate the importance of irrigation in some states. The $A(R)$ values, averaged over $A(R,S)$ values (eqs. (1) and (6)), reflect relatively poor soils in Uttar Pradesh.

In summary, the model shows how to combine weather and culture to explain yield variation from state to state in India.

TABLE VI.—Comparison of Winter Wheat Yields in India

(a) Estimated and reported yields

Harvest year	Yield, bulacre. for—									
	Punjab		Rajasthan		Haryana		Uttar Pradesh		Bihar	
	MOD ¹	GOV	MOD	GOV	MOD	GOV	MOD	GOV	MOD	GOV
1972	32	36	19	19	26	30	17	19	22	26
1973	34	33	18	19	26	26	16	18	—	—
1974	36	33	16	16	27	23	16	14	21	18
1975	36	36	19	19	27	26	19	17	21	20
A(R) ^a	-0.7		-2.6		-3.2		-9.9		-0.3	

(b) Proportion of wheat area irrigated and nonirrigated^b

Harvest year	Area, percent. for—									
	Punjab		Rajasthan		Haryana		Uttar Pradesh		Bihar	
	P _c	P _i	P _c	P _i	P _c	P _i	P _c	P _i	P _c	P _i
1972	13	87	33	67	17	83	33	67	47	53
1973	12	88	27	73	16	84	31	69	—	—
1974	12	88	34	66	14	86	30	70	41	59
1975	12	88	30	70	14	86	30	70	40	60

(c) Yield attributed to indicated weather components^c

Harvest year	Yield, bulacre. for—									
	Punjab		Rajasthan		Haryana		Uttar Pradesh		Bihar	
	W _c	W _i	W _c	W _i	W _c	W _i	W _c	W _i	W _c	W _i
1972	31	35	18	26	20	29	24	30	19	26
1973	28	35	5	25	19	30	20	29	—	—
1974	34	37	3	25	13	31	17	31	12	27
1975	32	37	1	28	13	32	23	32	15	26

^a B = 0.70

^b P_c = continuously cropped (dryland), P_i = irrigated

^c W_c = dryland component, W_i = irrigated component

Stability of A(R) Over Time

With sufficient historical yield and weather data, it is possible to study the stability of the local constant term A(R,S) over time. Equation (6) gives the basic relation between A(R,S), Government-reported yields, and average WAC(R,S) values.

When results over CRD's and states are combined, the (S) for stations is dropped and A(R) is used. Thus,

$$A(R) = \overline{GOV}(R) - B \cdot \overline{WAC}(R) \quad (7)$$

For U.S. data, $A(R)$ is calculated for 1955-64 and 1967-76, centered approximately on 1960 and 1972; and the differences are considered.

$$A_{72}(R) - A_{60}(R) = \left[\overline{GOV}_{72}(R) - \overline{GOV}_{60}(R) \right] - B * \left[\overline{WAC}_{72}(R) - \overline{WAC}_{60}(R) \right] \quad (8)$$

If one can assume that 10-year means average out much of the weather variation, then

$$\overline{GOV}_{72}(R) - \overline{GOV}_{60}(R) = \text{a measure of the effect of cultural changes on yields} \quad (9)$$

except when either 10-year period includes severe disease epidemics. If $B * [\overline{WAC}_{72}(R) - \overline{WAC}_{60}(R)] = [\overline{GOV}_{72}(R) - \overline{GOV}_{60}(R)]$, then $A_{72}(R) - A_{60}(R) = 0$. The latter result not only indicates that $A(R)$ is quite stable but also that the model explains most of the increase in yields from cultural changes. Results of this type of analysis are shown in table VII.

For winter wheat in the U.S. Great Plains, the model explains 5.0 of a 6.9-bushel-per-acre change, leaving a 1.9-bushel-per-acre increase due to non-modeled causes and/or weather variation not completely averaged out. For the Eastern Great Plains, $\overline{GOV}_{72}(R) - \overline{GOV}_{70}(R) = 6.3$ becomes a poor measure of cultural change because of the septoria epidemics in 1973 and 1974 (table IV), which reduced yields that were not detected by the model.

Cultural changes, as measured by the model, explain all the increase in yield in the Northwest.

For spring wheat, the model explains about 50 percent of the 6.3-bushel-per-acre increase. KSU believes that a major portion of the remaining increase resulted from planting later and using more herbicides to control weeds between 1960 and 1972.

Weather Station Density

Applications discussed previously were based on a rather sparse network of weather stations. The effect of more weather stations using Kansas data from 1955 to 1976 was investigated. Increasing the number of weather stations from 7 to 42 reduced the root mean square error (RMSE) = $\{N^{-1} * [MOD(R) - GOV(R)]^2\}^{1/2}$ from 3.1 to 2.6, or by 16 percent.

Disease Losses

In conjunction with the station density study, KSU considered how much improvement could be achieved if losses due to stem and leaf rust were known. With data supplied by the USDA Cereal Rust Laboratory in Saint Paul, Minnesota, KSU reduced model yields by the percentages indicated, recalculated its constants to adjust WAC values to a regional level, and further reduced the RMSE from 2.6 to 2.3. Thus, the combined benefit of more weather stations and knowledge of rust losses reduced the RMSE by 26 percent.

After application of a high-density weather network and rust loss information, the remaining "large" differences between the model and the USDA SRS resulted from underestimates of high yields in 1970-73 and an overestimate in 1966 due to freezes at heading.

TABLE VII.—Change in Yield of Winter Wheat Averaged Over Two 10-Year Periods

[Eq. (8)]

Regions (R)	$A_{72}(R) - A_{60}(R)$, bulacre	$\overline{GOV}_{72}(R) - \overline{GOV}_{60}(R)$, bulacre	$B * [\overline{WAC}_{72}(R) - \overline{WAC}_{60}(R)]$, bulacre
U.S. Great Plains	1.9	6.9	5.0
Eastern Great Plains	-3.0	6.3	9.3
Northwest	-0.5	9.8	10.3
U.S. Great Plains	3.3	6.5	3.2

POTENTIAL IMPROVEMENTS

Applying the model with diverse climatic data showed several shortcomings and pointed to needed improvement. KSU identified some decision points in model development in which alternative procedures would have been more productive. Alternative procedures include the following.

1. For winter wheat, use a "normal" crop calendar calibrated to each weather station location divided into phases by the following nine stages: planting, beginning of dormancy, end of dormancy, jointing, flag leaf, heading, milk, dough, and ripe. The normal crop calendar would be fixed over years.

2. Screen plot yield data to eliminate yields abnormally low because of disease epidemics, heavy insect losses, and/or unknown causes.

3. Include a soil index in the model with one or more soil variables based on soil types on which varietal trials were conducted.

4. Build WRV's that reflect that the contribution of 1 inch of precipitation on yield depends on the status of soil moisture budget when the precipitation occurs.

Some of the shortcomings of the model, which hopefully will be remedied with these procedures, are underestimating high yields, not detecting and measuring yield losses due to winterkill and freezes at heading, and overestimating when disease epidemics occur.

In conclusion, there is no technical barrier to real-time application of the model either for selected regions or on a global basis. The computer software, WHYMOD (wheat yield model), has been in operation at the National Oceanic and Atmospheric Administration (NOAA) Center for Climatic and Environmental Assessment (CCEA) and can be operated on a real-time basis. Preharvest forecasts are programmed into WHYMOD with the strategy of substituting mean values for variables that are generated after forecast time.

As indicated in the section entitled "Applications," the model in its current form can produce useful results and provide insights into causes for increases and decreases in yields despite its deficiencies.

KSU is continuing work by retracing the steps in model development in an effort to produce an improved model.

ACKNOWLEDGMENTS

The authors gratefully acknowledge support from LACIE management and Research, Test, and Evaluation personnel under Contracts NAS 9-14282 and NAS 9-14533; consultations with agronomic specialists at KSU; data collection by weather observers, experiment station personnel, crop reporting services, and NOAA CCEA workers; the services of the KSU weather data library directed by L. Dean Bark; the programming services of Don Wagner, Steve McFarland, Ken Laws, Bob Owens, Mike Franzblau, Mike Frerichs, and John McKean, who worked under Jeanne Sebaugh's leadership; the technical service of Dale Fjell; and clerical and typing services of Betty Skidmore and Kathy Elliott.

REFERENCES

1. Feyerherm, A. M.: Response of Winter and Spring Wheat Grain Yields to Meteorological Variation. Final Report, Contract NAS 9-14282, Kansas State University, Feb. 1977.
2. Feyerherm, A. M.: Planting Date and Wheat Yield Models. Final Report for Period Feb. 1975 to Mar. 1977, Contract NAS 9-14533, Kansas State University, Sept. 1977.
3. Feyerherm, A. M.: Application of Wheat Yield Model to United States and India. Final Report for Period Mar. 1977 to Nov. 1977, Contract NAS 9-14533, Kansas State University, Dec. 1977.
4. Robertson, G. W.: A Biometeorological Time Scale for a Cereal Crop Involving Day and Night Temperatures and Photoperiod. *Intern. J. Biometeorol.*, vol. 12, no. 3, 1968, pp. 191-223.
5. Baier, W.; and Robertson, G. W.: A New Versatile Soil Moisture Budget. *Canadian J. Plant Sci.*, vol. 46, 1966, pp. 299-314.
6. Baier, W.: Crop-Weather Analysis Model: Review and Model Development. *J. Appl. Meteorol.*, vol. 12, 1973, pp. 937-947.

The Law of the Minimum and an Application to Wheat Yield Estimation

R. B. Cate,^a D. E. Phinney,^a and M. H. Trenchard^b

INTRODUCTION

The law-of-the-minimum (LOM) concept dominated agricultural science throughout the 19th century. Its most famous proponent was a German chemist, Justus von Liebig, although several other prominent scientists contributed to the gradual evolution of the principle (ref. 1). In the early 20th century, the LOM concept was expanded by Blackman to include rate or flux variables, specifically in photosynthesis (ref. 2). At about the same time, Shelford developed the idea of ecological maximums and minimums; these constitute the "Limits of tolerance" that control the distribution of organisms (refs. 3 and 4). It is likely that the scientists who pioneered in the application of the LOM to biological systems were aware that they were merely extending the law of multiple proportions, which had already become the basic principle of chemistry, crystallography, and other fields dealing with the structure of matter.

Despite its fundamental theoretical importance, the LOM was not applied mathematically. As quantification became even more important in scientific research, the LOM declined in prestige because it was incompatible with conventional analytical methods involving calculus, analysis of variance, and multiple (additive) regression. However, in 1963, Swanson pointed out the relationship between the LOM and linear programming (ref. 5). Perhaps coincidentally, increasing attention has since been given to quantified application of the LOM. Several algorithms now exist for fitting the model (refs. 6 to 11).¹⁻³ None of these algorithms is wholly satisfactory, but sufficient progress has been made to permit

fairly rigorous application and testing. The purpose of this paper is to illustrate the LOM concept in a variety of contexts and then to report on how it is being adapted to estimate wheat yields.

THEORY

Algebraic Formulation

Mathematically, the LOM can be expressed as

$$Y = \min f_i (X_i)$$

For example,

$$\hat{Y} = \min \left\{ \begin{array}{l} Y_{x1} = a_0 + a_1 X_1 \\ Y_{x2} = b_0 + b_1 X_2 \\ Y_{xn} = n_0 + n_1 X_n \end{array} \right\}$$

¹M. J. Hartley and H. O. Hartley have developed an unpublished algorithm using maximum likelihood estimation for calculation of the LOM parameters. Although untested at this time, the approach is a significant advance.

²R. B. Cate, and T. Y. Hsu, "An Algorithm for Deriving Linear Programming Activities Using the Law of the Minimum." North Carolina Agriculture Experiment Station Technical Bulletin, to be published.

³P. E. Waggoner, "Liebig's Law of the Minimum and the Relation Between Weather, Pathogen and Disease." Connecticut Agriculture Experiment Station, to be published.

^aLockheed Electronics Company, Houston, Texas.

An example might be that predicted yield \hat{Y} equals the minimum value \hat{Y} predicted by the following equations.

$$\hat{Y}_T = 80 - 0.5 (\text{mean maximum temperature})$$

$$\hat{Y}_R = 0 + 4.0 (\text{total rainfall})$$

$$\hat{Y}_M = 40 (\text{maximum historical yield})$$

The prediction process, using hypothetical data, is shown in table I. The parenthetical \hat{Y} values determine \hat{Y} , the final prediction.

TABLE I.—Example of the Application of a Law-of-the-Minimum Model

Case number	Temperature	Rainfall	\hat{Y}_T	\hat{Y}_R	\hat{Y}_M	\hat{Y}
1	90	10	(35)	40	40	35
2	80	5	40	(20)	40	20
3	70	15	45	60	(40)	40

The equivalent representation for an additive multiple regression model might be

$$Y = a_0 + \sum a_i X_i$$

For example,

$$\hat{Y} = a_0 + a_1 X_1 + a_2 X_2 + \dots + a_n X_n$$

whereas a multiplicative, multiple regression model can be expressed as

$$Y_i = (fX_1)(fX_2) \dots (fX_n)$$

For example,

$$\hat{Y} = (a_0 + a_1 X_1) (b_0 + b_1 X_2) \dots (n_0 + n_1 X_n)$$

Interactions

The treatment of interactions differs with the three preceding types of functions. In the additive model, an interaction between X_1 and X_2 is usually represented as a new variable consisting of the product of the two, $X_1 X_2$. This variable has its own coefficient and is included in the overall additive equation. Since the number of possible interactions increases at an exponential rate as more variables are considered, a large model may become cumbersome and unintelligible. The structure of the LOM function avoids this difficulty by making the interactions absolute; i.e., the effect of a limiting amount of one variable is to suppress completely the response to another factor. In other words, the slope of the second factor becomes zero rather than its coefficient. (Statistically, interactions can be defined as the effects of variables on the slopes of other variables.) The multiplicative model involves a more extreme treatment of interactions than does the LOM. For example, the Mitscherlich-Baule-Spillman-Bray model predicts that if three variables are each present in sufficient quantity to produce 50 percent of maximum yield, the yield will be 12.5 percent of maximum since $0.5 \times 0.5 \times 0.5 = 0.125$ (ref. 12).

Substitution

An inherent property of additive variables is the capability of substituting one variable for another, so that a sufficient amount of one variable can completely overcome even the total absence of another. The LOM does not permit substitution. However, when substitution does exist, a new variable may be created which is the sum of two additive variables. An example of this technique is given later in this paper.

Graphic Representation

The relationship between the LOM and the linear programming model (ref. 13) can be seen by comparing figures 1 and 2, which use the formation of water as an example. In figure 1, only points A and E represent optimum combinations of hydrogen and oxygen. At points B, C, and F, there is insufficient hydrogen to balance the extra oxygen; whereas at points D, G, and H, there is an excess of hydrogen relative to the oxygen supplied. The result is a series of right-angled isoquants that define the production diagonal, or expansion path, along which efficient output of water occurs. Figure 2 shows the yield of water plotted against hydrogen and oxygen individually. Note that the hydrogen plot is identical to that in figure 1, whereas the oxygen plot is a rotated mirror image of the oxygen portion of figure 1. However, in figure 2, it becomes more evident that the points of excess oxygen (B, C, and F) can be used as replicates for definition of the hydrogen response line, 0-EF, whereas the points of redundant hydrogen can be used to define the oxygen response line, 0-EH. This property is the basis for some of the algorithms currently used to fit the LOM.

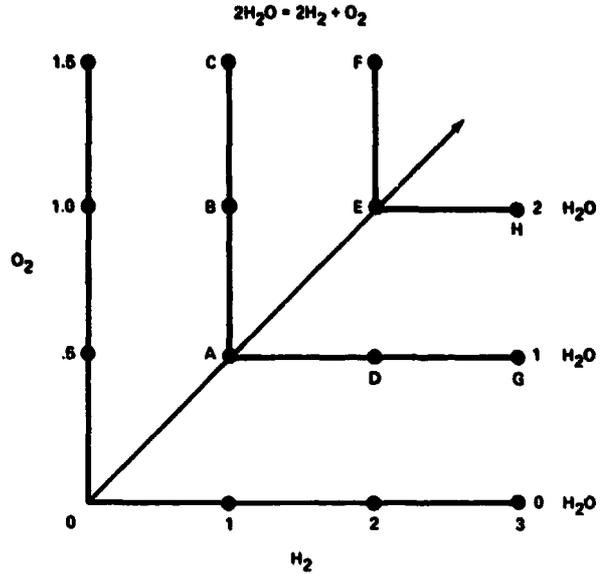


FIGURE 1.—Linear programming model.

The remainder of this paper is devoted to a detailed discussion of a trial application of the LOM to yield modeling.

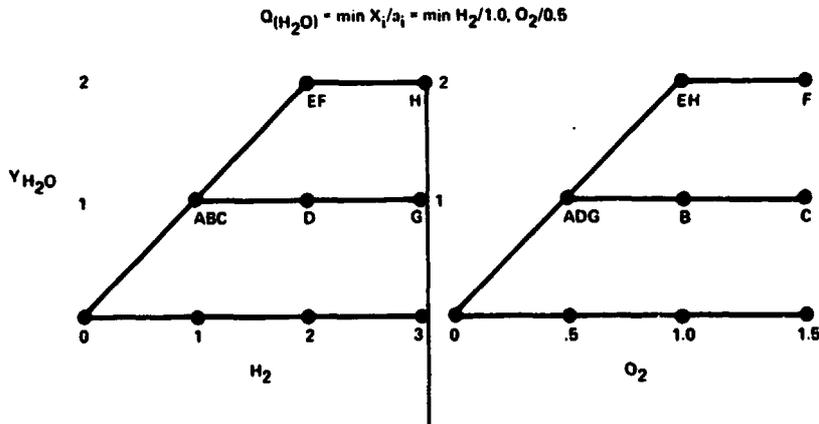


FIGURE 2.—Law-of-the-minimum model.

USE OF THE LOM IN YIELD MODELING

Interpretation of Individual Experiments

The form of the LOM function means that the coefficients of the individual variables are independent. This independence permits the direct use in

yield models of individual experimental results obtained under controlled conditions with replicated factorial treatments. For this example, the results of a typical experiment were used; they are plotted in figure 3 (ref. 14). The left portion indicates that the nitrogen response followed the LOM reasonably well, since treatments 1, 2, 3, and 4 form a relatively horizontal line corresponding to the lowest level of

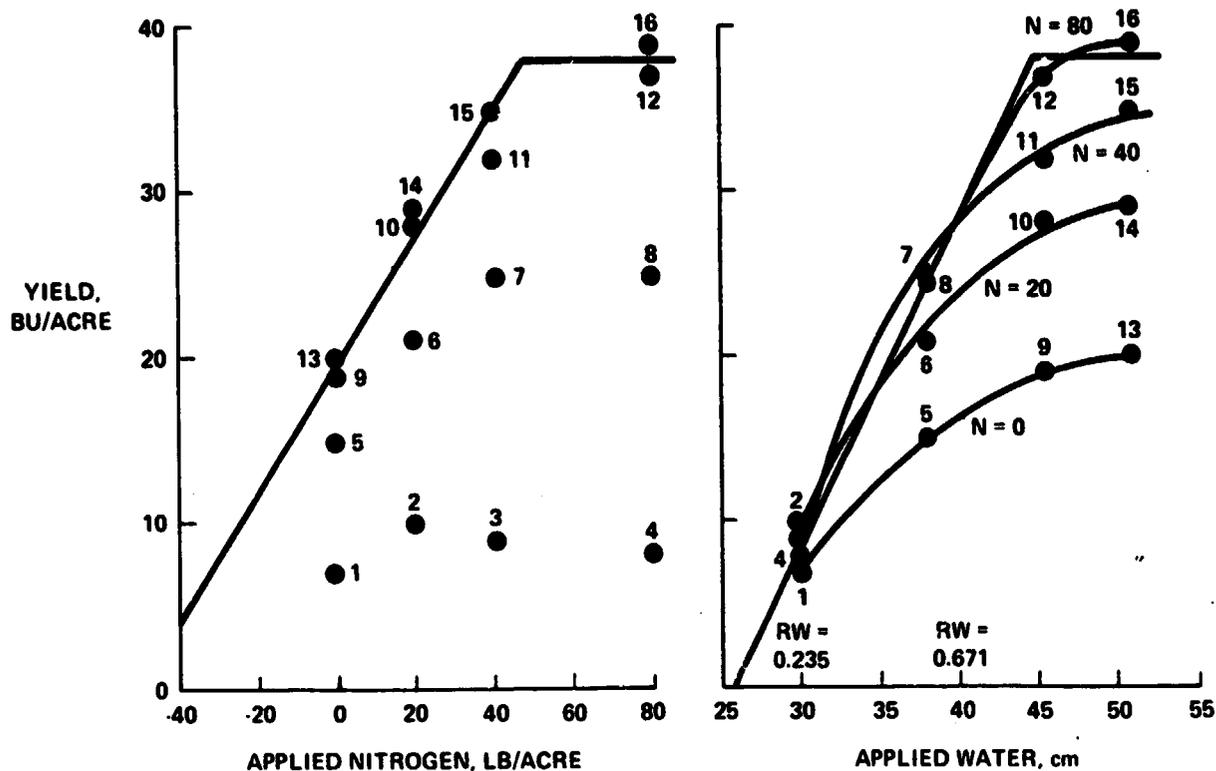


FIGURE 3.—Sample application of the LOM model to the joint effects of nitrogen and moisture on grain yield in which the original data suggest substitution of water for nitrogen. The N values are applied nitrogen in pounds per acre; RW is relative water.

applied water. Similarly, treatments 7 and 8 form a horizontal line corresponding to the next level of applied water. However, the right portion of figure 3 indicates that water has substituted for applied nitrogen since there are curved positive responses to water at each applied nitrogen level. These responses are not as steep (i.e., efficient) as the response to water when nitrogen is not limiting, but the LOM does not appear to be the proper representation. However, this discrepancy can be resolved, using the logical assumption that the intercept value of 20 bushels per acre on the nitrogen graph represents the contribution of soil nitrogen under optimum moisture conditions. If the nitrogen response line is extrapolated to zero yield, the amount of soil nitrogen can be estimated to be approximately 50 pounds per acre. Thus, about 1 pound of soil nitrogen per acre is being made available by each additional centimeter of applied water. To obtain a true LOM representation of the data, it is necessary to create a new variable (which is called total available

nitrogen) that is defined as the sum of applied nitrogen plus soil nitrogen multiplied by relative water availability (RWA). (Relative refers to water level as a percentage of maximum applied.) The results of this transformation are plotted on the left side of figure 4. The LOM model now provides a satisfactory description of the data. All treatment yields are determined by available nitrogen except at the highest applied nitrogen levels (treatments 4, 8, 12, and 16). The nitrogen response equation is

$$\hat{Y} = 3.16 + 0.36N$$

where \hat{Y} is yield in bushels per acre and N is nitrogen in pounds per acre. This equation is very close to other published nitrogen response coefficients (refs. 7 and 15).

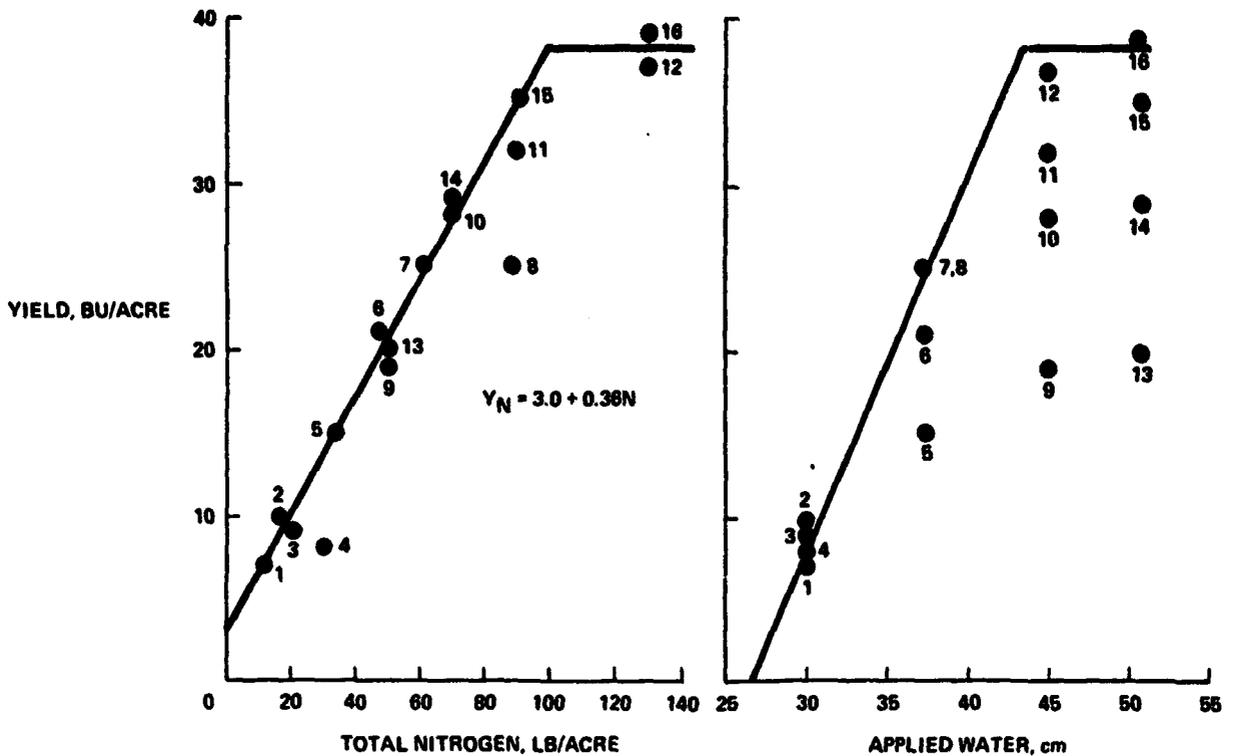


FIGURE 4.—Sample application of the LOM model to the joint effects of nitrogen and moisture on grain yield when apparent substitution phenomena are removed by combining soil and applied nitrogen.

Use of the Experimental Model

The experimental model was used in the following applications.

Data base.—Once a basic model has been built from experimental data, the problem becomes one of adaptation to available data. This example consisted of a data set for spring wheat compiled by Dr. A. M. Feyerherm (personal communication). The basic variables used were (1) total precipitation by crop stage, (2) mean maximum and mean minimum rainfall by crop stage, (3) applied nitrogen by crop reporting district (CRD), (4) percentage of fallow by CRD, (5) the relative yielding potential of the dominant variety planted in each CRD, and (6) yields by CRD. Information on these variables covered the period 1955-76 for most of the spring wheat producing districts in the U.S. Great Plains (USGP).

RWA.—The fundamental assumption of the model is that nitrogen uptake is determined by the amount of nitrogen present and by the availability of

water to transport the nitrogen into the plant. Data limitations precluded the use of a soil moisture budget to estimate water availability. The most practical alternative seemed to be an estimate of atmospheric water balance. This estimate was calculated as the difference between precipitation and estimated pan evaporation. The latter was calculated as a function of the estimated atmospheric vapor pressure deficit; i.e.,

$$E_p = b_0 + b_1 E_s - b_2 E$$

where $b_0 = 0.2163$, $b_1 = 0.3473$, and $b_2 = -0.2644$; E_p is pan evaporation in inches; E_s is the vapor pressure function value in millibars of the daily maximum air temperature; and E is the vapor pressure function value in millibars of the daily minimum air temperature. The functions E_s and E were calculated using a form of the vapor pressure function as derived from the Clausius-Clapeyron equation with

compatible units (ref. 16). The coefficients b_0 , b_1 , and b_2 were derived by fitting observed data for pan evaporation and temperatures at several stations in the USGP. The values of the coefficients were found to be remarkably stable over broad areas. The range of the estimates was -14 to 0 inches. This was converted to a percentage scale to represent RWA and then calculated for the three periods of plant growth defined in the paragraph on nitrogen uptake.

Nitrogen.—Soil nitrogen for each CRD was estimated by dividing the maximum historical yield in a CRD by 0.36 (the slope of the nitrogen response in the basic model) and by the relative yielding ability of the dominant varieties for each year and by subtracting the amount of nitrogen applied in that year. This figure was then divided by a factor consisting of $1.0 + 0.5$ (percentage fallow). The purpose of this division was to consider the accumulation of nitrogen in the soil during fallow. The maximum value obtained in the period 1955-66 was then used as an estimate of soil nitrogen. The resulting figures for soil nitrogen combined with the applied nitrogen and the nitrogen due to fallowing to give a total nitrogen figure for each year in each CRD.

Nitrogen uptake.—Because nitrogen uptake and plant growth are known to follow approximately the logistics curve, nitrogen uptake was broken down into 20 percent during the period planting to jointing, 60 percent during jointing to heading, and 20 percent during heading to ripe. The RWA for each period was used to calculate the total nitrogen available during each period. These figures were in turn multiplied by the uptake coefficient for the period; the sum was considered to be total nitrogen uptake. The basic model equation was then solved for this total, giving a yield prediction based solely on nitrogen uptake. This was done for the period 1955-66, and the results were compared with the actual values. Two major types of systematic errors were identified. The first type was a consistent tendency for actual yields to be anomalously high in years with cool summers, probably due to a decreased respiration rate during grain formation. As an interim substitute for a respiration submodel, a critical level of 71.5° F was simply defined for mean temperature during the milk to ripe period and 4 bushels were added to the yield estimates when the mean temperature fell below this critical level.

The second type of systematic error was a consistent bias over years which varied by CRD, presumably due largely to soil differences. This bias was in-

corporated into the model as a simple additive term consisting of the mean CRD error over the preceding 8-year period.

Figure 5 summarizes the form of the model that was submitted for testing. The respiration and soil components will be improved as data and time permit.

Test Results for Baseline Model

The model was tested by calculating the mean air temperature from milk to ripe and the RWA from planting to jointing, jointing to heading, and heading to ripe at each synoptic weather station in the spring wheat region for the 1955-76 period. An objective analysis using variational analysis with low-pass filtering constraints as described by Wagner (ref. 17) was used to interpolate the four weather-related variables to a 0.5° grid network. All grid points falling within each CRD were averaged to obtain a mean value for the weather variables for each CRD-year combination.

The weather parameters were combined with the appropriate cultural and soil information. A 10-year bootstrap test (1967-76) was performed with a local adjustment factor fitted for each CRD. The resulting CRD yield estimates were aggregated to the Center for Climatic and Environmental Assessment (CCEA) model regions and ultimately to the entire spring wheat area using the U.S. Department of Agriculture Statistical Reporting Service⁴ (SRS) acreages.

Table II presents a comparison of the Cate-Liebig results with the baseline performance of the Feyerherm and CCEA Phase III yield models. Clearly, the Cate-Liebig model has performed well. Figure 6 shows the year-by-year performance of these models and the SRS yields for the spring wheat region.

Relatively poor performance of the Cate-Liebig model in Minnesota resulted in additional error analyses in that region. The marked change in performance for a more humid climate caused speculation that an additional weather and/or soil parameter may be required. However, to date, no satisfactory parameter has been developed. It has been determined that radical shifts in acreage in Minnesota may also be related to this problem.

⁴Now called Economics, Statistics, and Cooperatives Service.

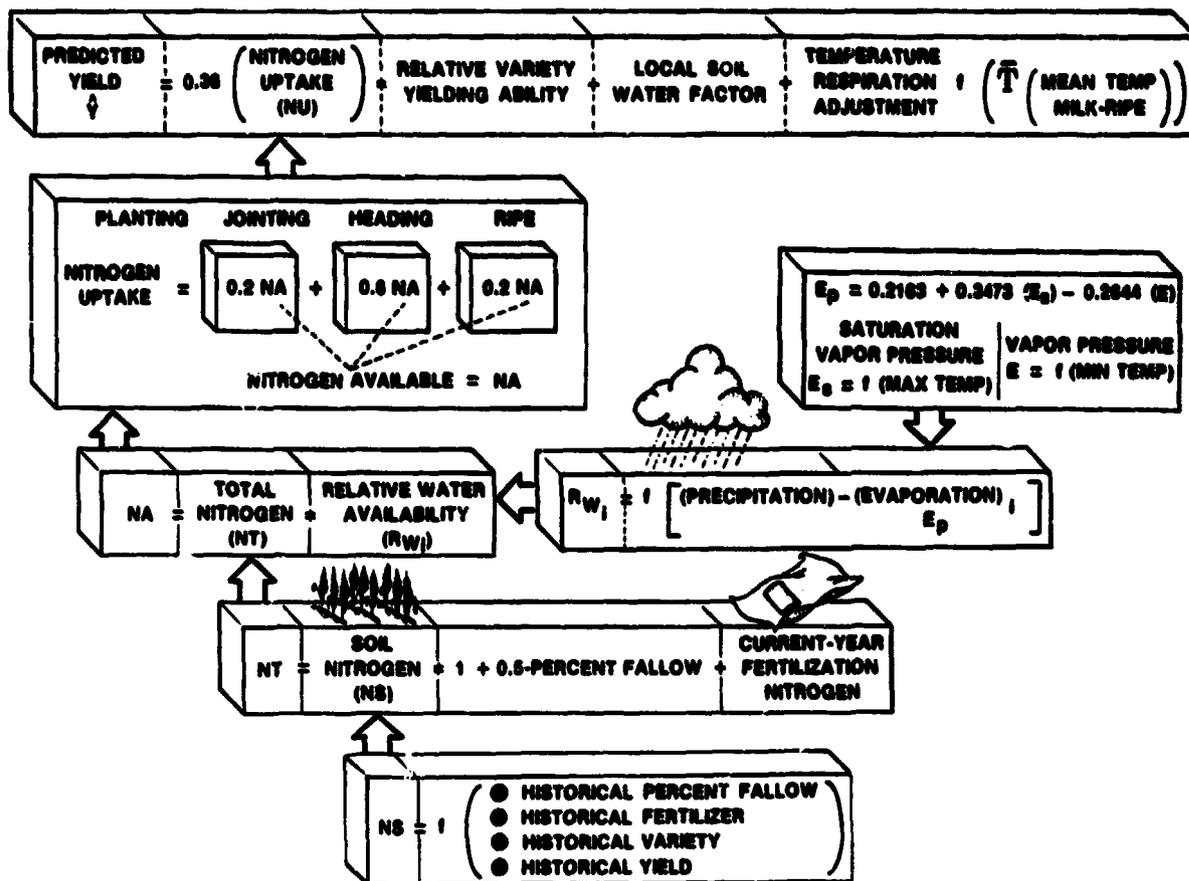


FIGURE 5.—Cate-Liebig baseline wheat yield model.

TABLE II.—Results of 10-Year (1967-76) Bootstrap Test on the Cate-Liebig Yield Model for Spring Wheat With Comparison to Baseline Feyerherm and CCEA Phase III Yield Models

Zone	CCEA Phase III		Feyerherm		Cate-Liebig	
	Bias	RMSE ^a	Bias	RMSE	Bias	RMSE
Montana	-0.6	2.18	-0.1	2.57	0.8	3.44
North Dakota	-1.2	2.94	-1	2.55	.1	1.37
Red River	-1.4	3.95	.9	2.70	-.8	3.19
Minnesota	-.6	3.81	2.5	5.45	-1.3	5.84
South Dakota	-.8	3.00	.9	4.96	.8	4.14
Total spring wheat	-1.0	2.56	.3	2.08	.0	1.29

^aRoot-mean-square error

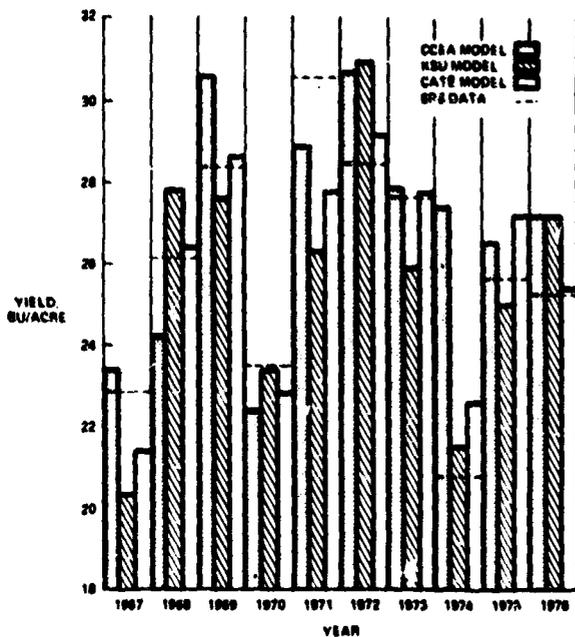


FIGURE 6.—Spring wheat model comparison.

Figure 7 shows the variation in acreage experienced in the south-central CRD of Minnesota. The source of the variation appears to be economic. The problem lies in the quality of land being moved in and out of wheat cultivation. An examination of the Soil Conservation Service (SCS) Conservation Needs Inventory gives a percentage of the area of the state in field crops where the land capability (other than erosion) can limit productivity. Table III gives these figures together with the errors observed for the Cate-Liebig model. The limitations considered by the SCS in compiling the percentages include soil tex-

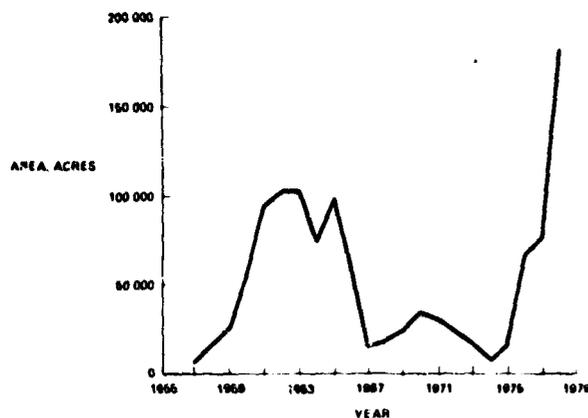


FIGURE 7.—Acreage variation with time in the south-central CRD of Minnesota.

ture, drainage, and climate (dryness and coolness). Evidently, model performance is degraded by the failure to consider these differences. Although the problem is not yet resolved, it is believed that more details on the soils under cultivation are needed, such as nitrogen content, water-holding capacity, and drainage class. In all likelihood, the use of a detailed soil moisture model will be required for a more complete development of this modeling approach.

CONCLUDING REMARKS

The LOM shows promise as a yield-modeling tool that may supplement the common regression-type modeling techniques. Its primary advantages are as follows.

1. Extremes are better predicted because LOM avoids the averaging or dampening of effects that usually results from multiple regression fitting.
2. Coefficients are stable over wide ranges of conditions because variable effects are modeled independently.

3. Coefficients can be derived from experimental work performed under controlled conditions.

4. Additional variables can be added to a model without affecting the coefficients of variables already included.

The relationships of the LOM to other modeling techniques are summarized in table IV, which is based on a table originally presented by Baier (ref. 18). The intent of this paper is to demonstrate that the LOM concept can be a valuable tool for model building when regression tools are inadequate.

Additional data, which may be of interest, on the indication of biological discontinuities as in the LOM, the confirmation of the LOM by a simulation model, and a review of the application of the LOM in tropical agricultural development are available in references 21, 22, and 23, respectively.

TABLE III.—Percentage of Total Area in Field Crops Where Land Capability Limitations Are a Predominant Problem

State	Area, percent	Cate-Liebig RMSE
Minnesota	55	5.84
Montana	40	3.44
North Dakota	33	1.37

TABLE IV.—Comparison of Model Features

Feature	Simulation models	IOM	KSU ^a	CCEA (b)
Type	Deterministic	Deterministic	Semistochastic	Stochastic
Time scale	Dynamic	Semidynamic	Semidynamic	Steady-state
Data source	Factorial experiments	Factorial experiments	Variety trials	Regional means
Approach	Physiological/causal	Empirical/causal	Empirical/correlative	Empirical/correlative
Purpose	Analysis	Prediction/analysis	Prediction	Prediction

^aKansas State University, from reference 19

^bFrom reference 20.

REFERENCES

- Browne, C. A.: Liebig and the Law of the Minimum. In Liebig and After Liebig, AAAS Publication 16, F. R. Moulton, ed., 1942.
- Blackman, F. F.: Optima and Limiting Factors. *Ann. Bot.*, vol. 19, 1905, pp. 281-295.
- Shelford, V. E.: *Animal Communities in Temperate America*. Univ. of Chicago Press, 1913.
- Odum, E. P.: *Fundamentals of Ecology*. Ch. 4, Principles Pertaining to Limiting Factors. Third ed. Saunders (Philadelphia), 1971.
- Swanson, E. R.: The Static Theory of the Firm and Three Laws of Plant Growth. *Soil Science*, vol. 95, 1963, pp. 338-343.
- Cate, R. B.; and Phinney, D. E.: A Test of the Utility of the Law of the Minimum in Yield Modeling. Paper presented at 69th Annual Meeting, American Society of Agronomy (Los Angeles), Nov. 13-18, 1977.
- Boyd, D. A.; Yuen, T. K.; and Needham, P.: Nitrogen Requirement of Cereals I. Response Curves. *J. Agr. Sci. (Cambridge)*, vol. 87, 1976, pp. 149-162.
- Anderson, R. L.; and Nelson, L. A.: A Family of Models Involving Intersecting Straight Lines and Concomitant Experimental Designs Useful in Evaluating Response to Fertilizer Nutrients. *Biometrics*, vol. 31, 1975, pp. 303-318.
- Perrin, R. K.: The Value of Information and the Value of Theoretical Models in Crop Response Research. *American J. Agr. Econ.*, vol. 58, 1976, pp. 54-61.
- Aigner, D.; Lovell, C. A. K.; and Schmidt, P.: Formulation and Estimation of Stochastic Frontier Production Function Models. *J. Econometrics*, vol. 6, 1977, pp. 21-37.
- Waugh, D. L.; Cate, R. B.; and Nelson, L. A.: Discontinuous Models for Rapid Correlation, Interpretation, and Utilization of Soil Analysis and Fertilizer Response Data. *Tech. Bull. 7, Int. Soil Fertility Evaluation and Improvement Program, North Carolina State Univ.*, 1973.
- Tisdale, S. L.; and Nelson, W. L.: *Soil Fertility and Fertilizers*. Second ed. The MacMillan Co. (New York), 1966, p. 26.
- Henderson, J. M.; and Quandt, R. E.: *Microeconomic Theory—A Mathematical Approach*. Second ed. McGraw-Hill, 1971.
- Ramig, R. E.: Relationships of Soil Moisture at Seeding Time and Nitrogen Fertilization to Winter Wheat Production. Ph. D. dissertation, Univ. of Nebraska, 1960. (University microfilms 21/06, p. 1320, no. 60-04508.)
- Bartholomew, W. V.: Soil Nitrogen—Supply Processes and Crop Requirements. *Tech. Bull. 6, Int. Soil Fertility Evaluation and Improvement Program, North Carolina State Univ.*, 1972.
- Hess, S. L.: *Introduction to Theoretical Meteorology*. Ch. 4, Clausius-Clapeyron Equation. Holt, Rinehart, and Winston (New York), 1959, pp. 46-51.
- Wagner, K. K.: Variational Analysis Using Observational and Low-Pass Filtering Constraints. Master's Thesis, Univ. of Oklahoma, 1971.
- Baier, W.: Note on Terminology of Crop-Weather Models. Paper presented at WMO Expert Meeting on Crop-Weather Models (Ottawa, Canada), Sept. 11-15, 1977.
- Feyerherm, A. M.; Kanemasu, E. T.; and Paulsen, G. M.: Response of Winter and Spring Grain Yields to Meteorological Variation. Final Report. Contract NAS 9-14282, Kansas State Univ., 1977.
- Wheat Yield Models for the United States. NOAA CCEA Tech. Note 77-1 (LACIE-00431), Columbia, Mo., Apr. 1977.
- Specht, A. W.: Evidence for, and Some Implications of, a Variation-Control Mechanism in Plant Composition. *Soil Sci.*, vol. 89, 1960, pp. 83-91.
- Smith, O. L.: Nitrogen, Phosphorus, and Potassium Utilization in the Plant-Soil System: An Analytical Model. *J. Soil Sci. Soc. America*, vol. 40, 1976, pp. 704-714.
- Sanchez, P. A.: Properties and Management of Soils in the Tropics. Ch. 9, Soil Fertility Evaluation. John Wiley and Sons (New York), 1976, pp. 295-345.

CCEA Second-Generation Wheat Yield Model for Hard Red Wheat in North Dakota*

S. K. LeDuc^a

INTRODUCTION

A wheat yield model was developed for hard red spring wheat in North Dakota using historical yields for crop reporting districts in conjunction with meteorological predictor variables based on weekly data. The meteorological data were aggregated according to observed phenological stages. The overall goal of this approach was to determine whether yield estimates for years with unusual planting dates and/or unusual phenological development could be estimated more accurately than in the original monthly models developed by the Center for Climatic and Environmental Assessment (CCEA) of the National Oceanic and Atmospheric Administration (NOAA).

CCEA FIRST-GENERATION MODELS

The first yield models developed by CCEA were regression models using monthly average temperature and precipitation for climatological districts as the basic meteorological variables. The averages for the districts were based on the average from a dense network of cooperative stations. Models were for states or for areas the size of states. To obtain variables for model areas, district data were weighted by relative harvested area for a specified year.

The use of monthly data in the first models had several shortcomings. If the crops were planted very early or very late, the stage of development in a particular month would not be what the model was expecting; i.e., the normal stage of development. For example, early planting might allow the wheat to

develop more rapidly to the extent that the wheat might be ripe during the month when heading would normally occur. In this case, the model would assume the crop was in the heading stage and interpret dry, hot weather as being detrimental to yield when in fact it should be advantageous to drying the ripe crop. Also, phenological development varies over a state, and areas smaller than states should be able to more effectively utilize stage-of-development information in a model.

Variables used in the first models were not capable of assessing the delayed and cumulative effects of a moisture deficit. Soil moisture should be a good indicator of this. Variables that are averages also do not reflect the full impact of extreme conditions.

MODEL VARIABLES

A model that would hopefully alleviate some of the problems was developed for hard red spring wheat yield in the nine crop reporting districts (CRD's) of North Dakota. Production and harvested acreage data for hard red spring wheat for the CRD's were available from the Statistical Reporting Service (SRS) of the U.S. Department of Agriculture (USDA). The most recent revision was used.

The quality of the historical yield data for the CRD's is not the same as that for the state based on objective yield surveys. The yield estimates for CRD's and counties are based mainly on responses from SRS mail surveys, which are adjusted using data from the agricultural census. State yields are revised based on check data such as state assessors' reports on acreage.¹

Actual observed phenological stages for each CRD in North Dakota from 1950 to 1975 were

*Paper presented to the Crop Modeling Workshop, October 3-5, 1977, Columbia, Missouri.

^aNOAA Center for Climatic and Environmental Assessment, Columbia, Missouri.

¹Louis T. Steyaert, "Quality of United States Wheat, Corn, and Soybean Crop Statistics," report to the Charles F. Kettering Foundation under Grant Number ST 76-30, Mar. 1977.

smoothed and the median date was determined. The observed stages for wheat are planting, emergence, jointing, heading, milk-to-dough, turning, swathing, and combining. The smoothing was both spatial and temporal when possible. Data were spotty in some of the early years. Attempting to utilize the phenological stages necessitates using periods shorter than a month for these smaller areas. The natural choice was to use periods of a week for the individual CRD's.

The basic aggregation of daily station data into weekly meteorological data for each CRD was accomplished by Dr. Amos Eddy of the Department of Atmospheric Science at the University of Oklahoma.² Data for 59 stations (fig. 1) were combined as follows.

1. Average total precipitation for the week (*PCP*) in hundredths of inches
2. Maximum number of days in which more than 1 inch of precipitation fell (*NPH*)

²Amos Eddy, Final Report on Federal Grant USDC (NOAA) 047-158-44000 concerning development of second-generation yield models, June 1977.

3. Maximum number of days in which more than 0.2 inch of precipitation fell (*NPM*)
4. Maximum number of days in which more than 0.1 inch of precipitation fell (*NPL*)
5. Average weekly maximum temperature (*MX*) in °F
6. Average weekly minimum temperature (*MN*) in °F
7. Maximum number of days in the week when the maximum temperature exceeded 100° F (*MXH*) or 90° F (*MXL*)
8. Maximum number of days in the week when the minimum temperature was less than 32° F (*MHN*)
9. The sum of the average daily growing degree days (*GDL*) in the CRD for the week; e.g., for location *i*, where the daily maximum temperature for day *j* is TX_{ij} , the daily minimum temperature is TN_{ij} , and the number of locations on day *j* is *M*

$$GDL = \sum_{j=1}^7 \frac{1}{M_j} \sum_{i=1}^{M_j} \left\{ \left[\frac{TN_{ij} + \min(TX_{ij}, 86)}{2} \right] - 40 \right\}$$

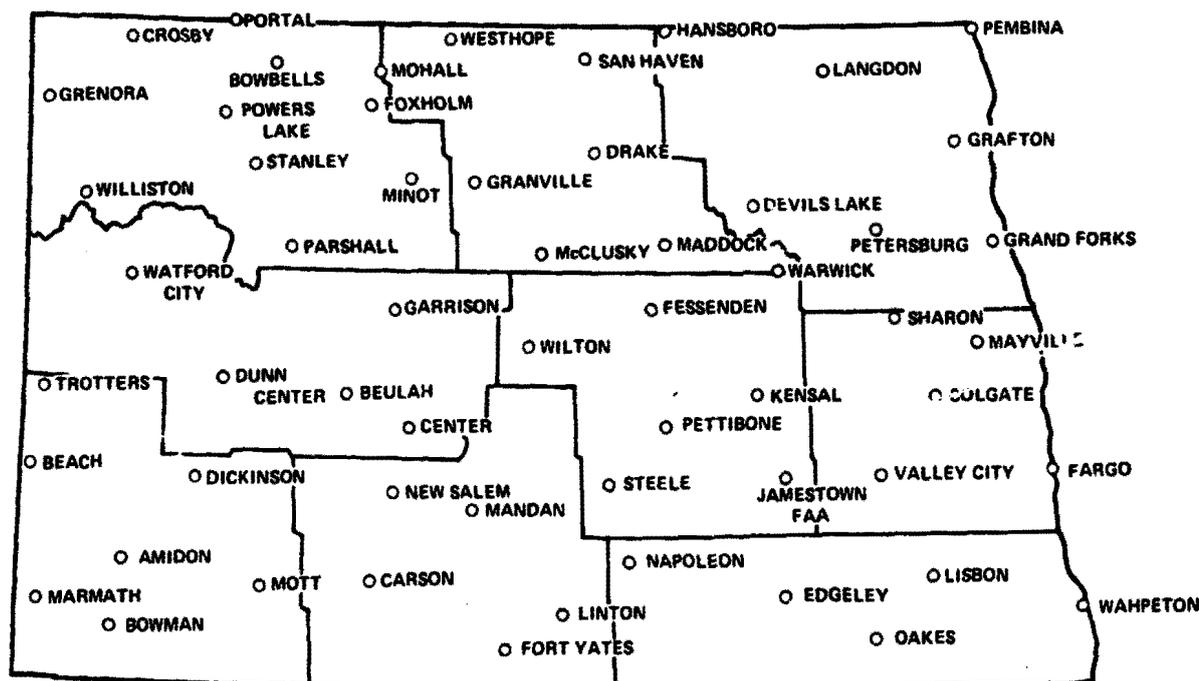


FIGURE 1.—North Dakota meteorological stations used in the aggregation of daily data.

10. Average weekly surface moisture amount where the maximum is 1 inch (SS)

11. Average weekly subsurface moisture amount (SU) where the maximum is the available water capacity minus one (available water capacity was 8 inches for CRD's 4 to 9, 7 inches for CRD's 2 and 3, and 6 inches for CRD 1)

12. Average weekly runoff in inches (RO)

The last three variables were determined from a hydrologic accounting system similar to the one reported by Palmer (ref. 1). Soil moisture is previous storage plus precipitation minus evapotranspiration, up to a set maximum. Excess precipitation is runoff. A surface layer can supply up to 1 inch to evapotranspiration, but only a fraction of demand beyond that can be supplied by the underlying layer. Evapotranspiration is that part of potential evapotranspiration (PET) that is satisfied. Thornthwaite (ref. 2) gives

$$PET_i = \frac{1.6 \left[\frac{5.5556 (T_i - 32)^A}{B} \right] \times \frac{HOURS}{12} \times \frac{7}{30}}{2.54}$$

where T_i = weekly CRD average temperature in °F (PET = 0 where $T < 32$)

T = long-term weekly CRD average temperature in °F

HOURS = number of daylight hours

7/30 = transformation from monthly values used by Thornthwaite to weekly values used in this study

B = heat index computed from long-term record

$$B = (1/4) \sum_{i=1}^{52} \left[\frac{(\bar{T}_i - 32)}{5} \right]^{1.514}$$

where \bar{T}_i is set equal to 32 if it is climatologically < 32

$$A = 0.49239 + 0.01792B - 0.0000771B^2 + 0.000000675B^3$$

MODEL FORM

The model is a multiple regression model. A constant shift term was considered for each CRD; however, none of the shift terms were significant. The model may be expressed as

$$y_j = \alpha + \delta YR + \gamma T_j + \sum_{i=1}^n \beta_i W_{ij} \quad (1)$$

where j varies for each crop district and n is the number of weather terms, selected from those listed in the previous section. The year YR is a variable defined as year minus 1950. The trend variable T_j is of the following form

$$T_j = 28 \cdot A_1 \cdot \exp \left\{ \left[.001 \cdot A_2 \cdot \left[(\text{year} - 1920) (50 \cdot A_3) \right] \right]^2 \right\} \quad (2)$$

This functional form was chosen to allow for the exponential rate of increase in the mid-1950's and to account for the apparent slowdown in the rate of change in the 1970's.

The coefficients A_1 , A_2 , and A_3 were determined from a nonlinear programming algorithm that fitted a linear trend from 1929 to 1949 and T (in the above form) from 1949 to 1976, with a forced juncture at 1949. A trend was fitted to each individual CRD yield time series; i.e., for each series T . The estimates of the parameters are given in table I. The expected state yield without regard to weather variables is shown in figure 2. This yield is derived by omitting the weather variables W_i from the full model and estimating the parameters α , δ , and γ using the years 1953 and 1957 to 1973. These parameter estimates were then used to estimate the expected yield for each CRD for each of the years from 1950 to 1976. Using actual harvested acreages, the expected CRD yields were aggregated to determine an expected state yield. Also shown in figure 2 is the expected state yield using the same process except that the trend function (eq. (2)) was fitted using the 1929-73 period; i.e., no linear trend was used for the early part. Estimates of the parameters for equation (2) are shown in table I.

The meteorological variables W_i are determined as follows. For a particular stage, the week in which

half the crop in a given CRD passed that stage was used. The weather variable for that week and the weather variable for both the week before and the week after are averaged together with weights of 0.50, 0.25, and 0.25, respectively. This averaging was used because, even within a single CRD, not all the crop is in the same stage during a given week.

TABLE 1.—North Dakota Hard Red Spring Wheat Trend Coefficients for Exponential Distribution

(a) Exponential form (eq. (2)) through entire period

CRD	A1	A2	A3
1	0.81	1.01	1.10
2	.80	1.09	1.10
3	.82	1.16	1.10
4	.76	.93	1.10
5	.71	1.15	1.10
6	.78	1.12	1.10
7	.78	.93	1.10
8	.69	.90	1.10
9	.66	1.19	1.10

(b) Exponential form (eq. (2)) 1949-76, linear during prior period,
 $T = B1 + B2^*(year - 1928)$

CRD	A1	A2	A3	B1	B2
1	0.89	1.72	1.07	8.0	0.25
2	.88	1.52	1.07	8.0	.10
3	1.13	1.34	1.06	10.5	.25
4	.87	1.56	1.08	4.5	.25
5	.88	1.52	1.07	5.5	.25
6	1.06	1.51	1.10	9.0	.25
7	.89	1.50	1.03	4.3	.25
8	.70	1.50	1.08	8.0	.01
9	.82	1.58	1.11	7.9	.01

The final variables were selected using a stepwise regression procedure with the restriction that the physical interpretation of the signs of the coefficients was correct with regard to the known response of wheat yields to climatic factors during a particular stage of development. Because of missing phenological data, only the years 1953 and 1957 to 1973 were used in selecting the variables. Truncated models were determined for the different phenological stages using only variables for that stage or previously occurring stages, allowing predictions to be made early in the growing season as the crop reached

each development stage. In each of these preharvest models, the trend was allowed to remain an independent variable and its coefficient was derived for each truncation to maximize the fit to the data.

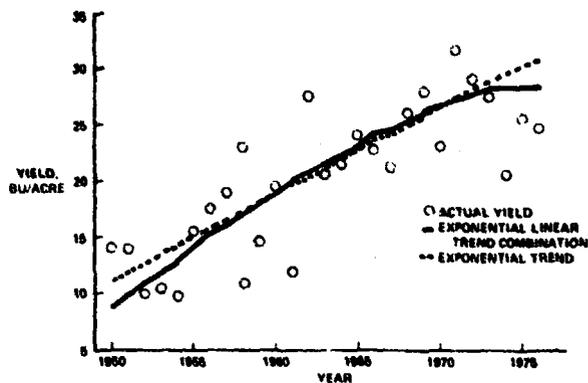


FIGURE 2.—North Dakota hard red spring wheat yield and fitted trend.

RESULTS

Table II contains the variables that were selected for a preliminary model. The first alphabetic characters of the weather variable names refer to the previously described 3-week weighted average of variables where the codes appear. The numeral that follows refers to stage of the wheat in natural order; i.e., planted is 1, emergence is 2, jointing is 3, heading is 4, milk-to-dough is 5, turning is 6, swathing is 7, and combining is 8. The variables, the estimates and the standard errors of their coefficients, and the *F* statistic with its significance level are also included in table II. Maximum temperature *MX*, minimum temperature *MN*, and precipitation *PCP* have the mean subtracted. The means are included in table II.

The trend variable has a coefficient varying from 0.81 to 0.91, indicating that scalar adjustment to the trend estimated in equation (2) is needed. High maximum temperatures are detrimental to yield from jointing through turning. Precipitation is advantageous to yield from emergence to heading except that runoff can cause a decrease in yield during jointing and heading. The milk-to-dough stage has three variables that appear to be important: (1) subsurface soil moisture at this time can be beneficial to yield; (2) maximum temperature has an effect that is difficult to interpret in the model; and (3) the average maximum temperature has a negative coeffi-

TABLE II.—Coefficients (Standard Deviations) and Statistics Pertaining to Preliminary Models for North Dakota Hard Red Spring Wheat

Variable	Truncation stage ^a						Statistics on variables ^b				
	Trend	Planted (1)	Emergence (2)	Jointing (3)	Heading (4)	Milk-to-dough (5)	Turning (6)	Mean	SD ^c	Min	Max
Constant	-0.61 (1.57)	1.86 (1.61)	1.12 (1.63)	7.93 (2.31)	1.78 (2.54)	-5.66 (1.64)	4.11 (2.51)				
Trend	.90 (.10)	.82 (.10)	.81 (.09)	.86 (.07)	.91 (.07)	.91 (.06)	.85 (.06)	20.04	5.73	8.26	31.64
YR	.23 (.08)	.20 (.08)	.22 (.08)	.23 (.06)	.26 (.06)	.20 (.05)	.18 (.05)	13.00	7.80	0	26.00
MX1		.23 (.05)	.20 (.05)	.75 (.15)	.51 (.14)			60.24	6.82	45.13	76.20
GDL1				-.13 (.03)	-.08 (.03)			62.21	33.08	6.45	153.38
NPH2			2.01 (.97)					.21	.30	0	1.75
NPM2							.56 (.29)	1.27	.75	0	3.25
PCP2				.02 (.01)	.02 (.01)			45.21	34.34	.05	192.10
GDL3							-.06 (.01)	161.26	25.19	93.93	229.20
NPL3						.52 (.24)	.54 (.23)	3.37	.96	.50	6.00
MX3				-.49 (.06)	-.35 (.06)	-.25 (.06)		75.72	4.57	62.48	86.63
RO3							-3.62 (2.14)	.02	.09	0	.74
NPM3					.82 (.37)			1.86	.69	.25	3.50
NPM4						.88 (.28)	.88 (.29)	1.81	.75	.50	4.00
MX4					-.33 (.06)			78.72	4.40	66.40	89.44
RO4							-5.31 (2.87)	.01	.07	0	.71
MXH4						-5.71 (1.75)	-6.41 (1.67)	.27	.80	0	5.00
SU4							.36 (.15)	3.99	1.47	.66	7.00
MX5							-.64 (.07)	82.49	3.62	72.23	91.51
MXH5							3.58 (1.05)	3.32 (1.00)	.39	.99	6.00
SU5						.35 (.16)		3.53	1.40	.54	7.00
MX6							-.17 (.07)	83.66	4.00	73.02	92.85
R ²	0.61	0.65	0.65	0.78	0.82	0.87	0.88				
S _e ²	18.2	16.4	16.1	10.2	8.6	6.5	5.8				

^aVariables selected and coefficients estimated using only data for the years 1953 and 1957-73.

^bStatistics on variables use all years 1950-76.

^cStandard deviation.

cient whereas the number of days that the temperature is above 100° F has a positive coefficient. These variables are highly correlated. If temperatures get above 100° F, the yield does not continue to decrease at the same rate but is lessened. This is similar to a quadratic effect in that yield losses are not simple linear functions of temperature over the entire range.

Coefficients for the preliminary models defined in table I were recalculated using additional years. The reestimated models were used to obtain predictions for each CRD for each year from 1950 to 1976. These CRD predictions were then aggregated to the state level using the actual acreages. The mean squared errors or the average squared difference between the aggregated predictions and the actual yield are shown in table III. The mean squared errors appear to be stable and are of course smaller when more years are used in estimating the coefficients.

The quality of the phenological data for the years from 1950 to 1956 is poor. Some of these years had few observations on the phenological stages. The missing phenological data were estimated in an attempt to use the data that were reported in addition to the development stage as indicated in the "Weekly Weather and Crop Bulletin." This estimation is, however, a source of error in evaluating the models, but these questionable years were not considered when selecting variables for the models. The phenological data for 1976 were obtained from the actual planting date and, for later phenological stages, from Robertson's biometeorological time scale (ref. 3).

Yield estimates for the truncated models in table IV, with variables as listed in table II, were generated for independent data for the years not considered in selecting the variables; i.e., 1950 to 1952, 1954 to 1956, and 1974 to 1976. For the early years from 1950

**TABLE III.—Mean Squared Error (1950-76)
Preliminary Truncated Models for North Dakota
Hard Red Spring Wheat Yield**

Years used to estimate coeffi- cients	Trend	Plant- ing	Emer- gence	Truncation		Milk to dough	Turn- ing
				Joint- ing	Head- ing		
1953, 1957-73	13.6	11.4	11.2	6.4	7.0	5.1	6.1
1951-76	11.9	10.5	10.1	5.3	4.6	3.5	2.9
1950, 1952-76	11.9	10.5	10.0	5.1	4.4	3.5	3.0
1950-51, 1953-76	11.8	10.4	9.9	5.0	4.3	3.4	2.8
1950-53, 1955-76	11.9	10.6	10.0	5.2	4.4	3.4	2.8
1950-54, 1956-76	11.8	10.4	9.9	5.0	4.5	3.3	2.8
1950-55, 1957-76	11.8	10.4	9.9	4.9	4.3	3.4	2.8
1950-73, 1975-76	12.0	10.6	10.1	5.4	4.6	3.4	2.9
1950-74, 1976	11.8	10.4	9.9	5.1	4.4	3.4	2.8
1950-75	11.9	10.5	9.9	5.1	4.4	3.3	2.8
1950-73	12.7	11.5	11.0	5.7	4.6	3.5	3.4
1950-74	11.9	10.5	10.0	5.1	4.4	3.3	2.7

to 1955, the yield estimates are not very encouraging. In fact, at the time of turning, the model under-predicted twice by at least 2 bushels per acre and once by 5.5 bushels per acre. For 1952 and 1954, the model overpredicted by 2.6 and 3.6 bushels per acre, respectively. The model, however, did show a decrease in the yield estimate for 1952 that was lower than the actual yield for 1951. The estimates for 1956 were good from the milk-to-dough stage through turning, the last estimate. Omitting individual years,

the estimates for 1974 and 1976 were good. The estimate for 1975, however, is closer in the early phenological stages. If all years after an individual year are left out, the model prediction for 1974 is underestimated by 4.2 bushels per acre.

An alternate model using the number of days above 90° F for the later stages is presented in table V. The results of the independent yield predictions are included in table VI. The fit of the milk-to-dough truncation of this model is shown in figure 3. As with the preliminary model, the independent years 1950 to 1952, 1954 to 1956, and 1974 to 1976 were not used in model development. The dashed line represents yield estimates from the model when coefficients were estimated using the years 1953 and 1957 to 1973. Yield estimates for the independent years were also derived using all other years except that particular year to estimate the coefficients; these estimates are indicated by the label "independent test." In the "bootstrap test," all previous years were used to estimate the coefficients for the model.

The variables in the model with days above 90° F in addition to having a meaningful physical interpretation are statistically significant. The model for the jointing stage shows high maximum temperatures to be helpful at planting but harmful to yield at jointing. Precipitation occurring around jointing increases yield, but yield decreases if there are too many days of precipitation around the planting date. If the minimum temperatures are too high at emergence, the model indicates there will be a

**TABLE IV.—Independent Yield Predictions for North Dakota Hard Red Spring Wheat
Using Preliminary Truncated Models**

Year ^a	Trend	Planting	Emergence	Jointing	Heading	Milk-to- dough	Turning	Actual
1950	10.2	11.3	11.4	9.2	9.5	10.2	11.3	14.0
1951	11.2	12.3	11.7	13.3	13.0	10.8	11.6	14.0
1952	12.8	15.0	14.5	13.5	13.9	12.7	12.6	10.0
1954	14.6	13.5	13.3	12.7	11.5	13.4	13.6	10.0
1955	15.2	15.8	15.7	13.9	11.7	11.0	10.0	15.5
1956	16.1	15.1	14.9	13.4	13.4	17.0	17.1	17.5
1974	26.9	27.4	27.3	24.9	22.8	20.2	21.4	20.5
1975	26.1	27.6	25.6	26.7	25.2	24.9	24.2	25.5
1976	26.2	25.9	25.4	24.4	25.4	25.2	24.3	24.7
^b 1974	27.8	28.5	28.4	25.6	23.4	18.0	16.3	20.5
^b 1975	26.5	26.3	25.8	24.5	25.5	25.4	24.5	25.5

^aYear is left out of calculations for coefficients; also, the years included here were *not* considered when independent variables were selected.

^bYear and following years are left out of calculations for coefficient estimates.

TABLE V.—Coefficients (Standard Deviations) and Statistics Pertaining to the Alternate Models for North Dakota Hard Red Spring Wheat

Variable	Truncation stage ^a			Statistics on variables ^b			
	Jointing (3)	Heading (4)	Milk-to-dough (5)	Mean	SD	Min	Max
Constant	1.37 (1.72)	-1.61 (1.46)	-0.46 (1.23)				
Trend	.85 (.07)	.85 (.06)	.82 (.06)	20.0	5.7	8.3	31.6
YR	.24 (.06)	.29 (.06)	.30 (.05)	13.0	7.8	0	26.0
MX1	.19 (.05)	.12 (.04)	.07 (.03)	60.2	6.8	45.1	76.2
NPL1	-.62 (.31)			2.6	1.0	0	5.0
MN2	-.18 (.07)			41.3	4.2	30.4	55.0
MN3		-.31 (.07)	-.23 (.06)	50.3	3.6	39.0	59.3
MX3	-.23 (.08)			75.7	4.6	62.5	86.6
MXL3	-1.97 (.47)	-1.09 (.33)	-1.14 (.29)	.6	.8	0	4.0
NPM3	.83 (.40)	1.42 (.33)	1.13 (.29)	1.9	.7	.3	3.5
MXL4		-1.94 (.25)	-1.42 (.24)	1.1	1.1	0	5.0
MX5			-.53 (.07)	82.5	3.6	72.2	91.5
MN5			.23 (.09)	55.4	2.9	48.1	63.7
R ²	0.79	0.84	0.88				
S _e ²	9.9	7.7	5.8				

^aVariables selected and coefficients estimated using only data for the years 1953 and 1957-73.

^bStatistics on variables use all years 1950-76

decrease in yield. The model at the heading stage has the same variables as before, except that precipitation around planting and minimum temperatures at emergence are not included and the minimum rather than the maximum temperature at the jointing stage

TABLE VI.—Independent Yield Predictions for North Dakota Hard Red Spring Wheat Using Alternate Models With MXL Variable (Number of Days > 90° F)

Year ^a	Trend	Jointing	Heading	Milk-to-dough	Actual
^b 1976	26.4	25.7	24.6	25.2	24.7
^c 1975	26.8	28.9	28.6	27.4	25.5
1974	28.3	24.0	21.0	20.7	20.5
1950	8.9	8.7	7.6	8.1	14.0
1951	10.0	13.2	10.8	10.3	14.0
1952	11.0	11.6	9.6	10.7	10.0
1954	13.0	11.9	8.1	9.4	10.0
1955	14.1	12.4	8.3	8.7	15.5
1956	15.3	13.2	12.5	15.0	17.5

^aYear is not used in estimation of coefficients, nor were these years considered in the selection of variables

^bData from 1974 and 1975 added to basic data set for calculation of coefficients.

^cData from 1974 added to basic data set for calculation of coefficients.

is used. High maximum temperatures at heading are detrimental to yield. The model for the milk-to-dough stage includes the maximum temperature with a negative effect and the minimum temperature with a positive effect. These two variables are highly correlated, the coefficient on the maximum temperature is of larger magnitude, and both are measured as deviations from normal. If both max-

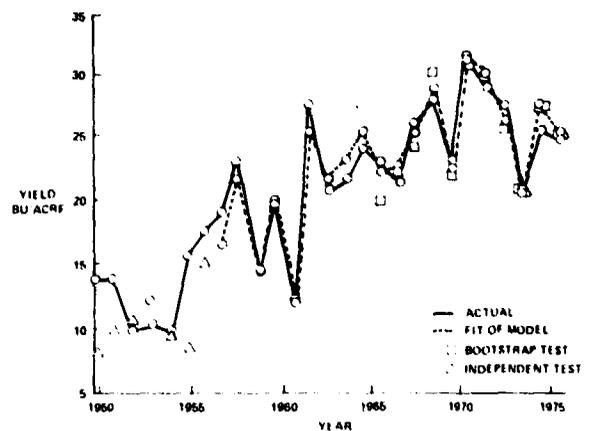


FIGURE 3.—Yields for North Dakota hard red spring wheat, actual and estimated from alternate model.

imum and minimum temperatures are 2° F above normal, the impact on yield would be estimated as $-0.53_2 + 0.23_2 = -0.3$ bushel per acre. If the temperatures deviate in opposite directions from the normal, the contributions would be in the same direction. For example, the maximum being higher than normal and the minimum lower than normal

would indicate not only higher day temperatures but also cooler nights; i.e., more diurnal variability. The model would estimate these conditions to be more detrimental to yield than if both day and night temperatures were higher than normal.

The CCEA first-generation Phase III model, which includes seven of the nine CRD's in North

TABLE VII.—North Dakota Spring Wheat Model

(a) Weighting factors ^a		(b) Definition of constants	
Crop district	Weight	Parameter	Definition
10 Northwest	0.2509	PET	Potential evapotranspiration estimated from Thornthwaite's method
20 North Central	.1558	PET A	1.051
40 West Central	.1178	PET I	34.813
50 Central ^p	.1616	April daylength	1.1297
70 Southwest	.0948	May daylength	1.2573
80 South Central	.0834	Latitude	48° N
90 Southeast	.1357	June deg-days >90° F	1 if deg-days >2; otherwise 0
		July deg-days >90° F	1 if deg-days >15; otherwise 0
		Deg-days stations	Bismarck, Dickinson, Fargo, Grand Forks, Jamestown, Minot, and Williston

^aWeights based on 1973 spring wheat harvested acreage.

(c) Analysis

Variable	Truncation ^a						
	Normal ^b	Trend	March	April	May	June	July
Overall constant	1.00	4.25220	5.07375	5.00981	5.12948	6.66911	7.83411
Linear trend 1932-55	24.00	0.17973	0.12701	0.14318	0.13230	0.10454	0.08950
Linear trend 1955-65	11.00	0.58914	0.68875	0.65567	0.65543	0.68523	0.69733
Linear trend 1965-72	8.00	0.26745	0.10780	0.22519	0.23068	0.27639	0.24716
Aug. to Mar. precipitation, ^c mm	176.67		0.02966	0.02716	0.02660	0.02589	0.02357
Apr. precipitation - PET, ^c mm	10.47			-0.00009	-0.00658	-0.00297	0.00181
Apr. precipitation - PET, ^d mm	10.47			-0.00042	-0.00035	-0.00046	0.00041
May precipitation/PET, ^c mm	0.77				1.24698	1.44860	0.70176
June precipitation, ^c mm	89.26					0.04159	0.03738
June precipitation, ^d mm	89.26					-0.00044	-0.00045
June deg-days >90° F						-1.29241	-0.88576
July deg-days >90° F							-1.55418
<i>R</i> ²		0.69401	0.74860	0.76076	0.76959	0.87180	0.88325
Standard error, ql/ha		2.83090	2.59865	2.60266	2.58938	2.01738	1.95503
Standard variance, ql/ha		8.01400	6.75300	6.77386	6.70489	4.06982	3.82215
Standard deviation of yields = 4.93589 ql/ha							

^aYields based on 1932-75; measured in quintals per hectare.

^bMeteorological normals based on 1931-75.

^cDeparture from normal.

^dSquared departure from normal.

Dakota, is detailed in table VII. It is difficult to compare coefficients of the first- and second-generation models because of inherent differences. Later stages show the negative effects of hot weather and the need for moisture in both models. Results of the test of the first-generation model are shown in figure 4. Basic differences must be considered when comparing figures 3 and 4. The yield shown for the second-generation model is for hard red spring wheat in all

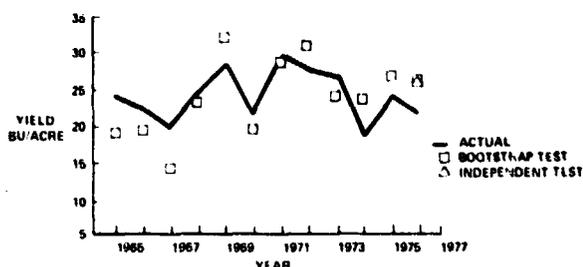


FIGURE 4.—Phase III first-generation model yields for North Dakota spring wheat in CRD's 1, 2, 4, 5, 7, 8, and 9.

of North Dakota, whereas the first-generation model is for the yield of all spring wheat in seven of the CRD's. The higher yielding Red River Valley is not included in the area included in the first-generation model. Both models seem to miss in the same direction. The second-generation model appears to give an improved estimate in 1974 and 1976. The latter year is an independent test year for both models, but 1974 is an independent test only for the second-generation model.

CONCLUSION

The second-generation model appears to provide estimates that are improved in two of the three independent test years. The earlier years analyzed using the bootstrap test do not show considerable improvement. The second-generation models are much more difficult to use operationally because of the variability of truncations caused by the rate of development of the wheat crop. More data must be collected, quality controlled, and used to calculate the derived variables. Thus, more resources are demanded for assessments using the second-generation model. The cost/benefit comparison of the two models is based on meteorological data that are routinely observed. The timeliness of providing the estimates also needs to be considered in the com-

parison. The occurrence of the phenological stages for estimates from the second-generation model may not be convenient for release and dissemination of yield estimates because the dates are variable. If the models are to provide estimates at fixed calendar dates for an operational system, the second-generation model needs to be evaluated for its ability to accomplish this. The comparative value of the two types of models will be determined by operational results of each on independent data.

REFERENCES

1. Palmer, Wayne C.: Meteorological Drought. Research Paper No. 45, U.S. Dept. of Commerce, Weather Bureau, 1965.
2. Thornthwaite, C. W.: An Approach Toward a Rational Classification of Climate. *Geog. Rev.*, vol. 38, 1948, pp. 55-94.
3. Robertson, G. W.: A Biometeorological Time Scale for a Cereal Crop Involving Day and Night Temperatures and Photoperiod. *Intern. J. Biometeorology*, vol. 12, 1968, pp. 191-223.

Prediction of Wheat Phenological Development: A State-of-the-Art Review

M. W. Seeley,^a M. H. Trenchard,^a D. E. Phinney,^a J. R. Baker,^b and R. G. Stoff^c

INTRODUCTION

The objectives of this paper are to describe the supporting research in crop development modeling (crop calendars) and, more specifically, to discuss the relative merits and shortcomings of various models for the development of wheat (*Triticum aestivum*) which emerged during the 3 years of LACIE. The models described herein represent a joint research effort of NASA and contractor scientists. The incorporation of these models into LACIE operations is discussed elsewhere in this volume (McCrary and Rogers, "Operation of The Yield Estimation Subsystem").

Crop phenology, the study of the expression of genotypic and environmental interactions, has been a key concept in the evolution of quantified crop development scales for many crops. Wheat, corn, peas, sorghum, and soybeans are a few of the crops for which development, from emergence through maturation, has been described using a phenologically based numeric scale. The history of agriculture shows that man has always used phenological characteristics to identify stages of development for particular crops. In fact, improvements in crop husbandry are still being made as a better understanding of crop phenology is gained.

Phenotypic characteristics of a crop may be divided into those which manifest growth and those which manifest development. Crop growth and development are frequently confused, although they are distinctly different concepts. Growth traditionally refers to an increase in plant size (roots, shoots, stems, and leaves) and represents one component of plant development. The concept of development includes the sequence of life cycle

events which lead to changes in tissue structure and/or function. Because there is some disagreement over this concept among agricultural scientists, three variant categories of crop development require distinction. The potential rate of crop development is determined genetically and can only be observed in the laboratory under controlled optimum conditions. The actual rate of development is the result of a system of genotypic-climatic-nutritional interactions which occur at the biochemical level in natural environments. Lastly, the observed rate of development depends on the degree to which a crop expresses changes in tissue structure or function and the frequency and accuracy of the observations of such changes.

The quantification of a crop phenological scale is based on the observed rate of development. Too frequently, vigorous growth obnubilates a clear picture of the ontogenetic status of a crop, which is mistakenly identified as an advanced stage of development. The observed rate of development, as the dependent variable in the model, may contain significant error caused by inadequate expression of ontogenetic changes by the crop or observation error by the scientist. When this is the case, it is futile to attempt a crop development model of the form $Y = f(x_1, x_2, \dots, x_n)$, because it is likely that the dependent variable (Y) will have a greater magnitude of error than the predictor variables (x_i), which are typically measured climatic or nutritional characteristics. In order to limit such errors in phenological data, careful dissection of plant parts and frequent field observations have been used in some crop development studies, such as those by Robertson, Hanway, Vanderlip, Williams, Major et al., and Seeley (refs. 1 to 6, respectively).

Traditionally, the quantification of a biological time scale for a crop has been accomplished by numbering, in ascending order, the phenotypic characteristics as they appear. The terms "developmental stage" or "biostage" refer to a particular point on the

^aLockheed Electronics Company, Houston, Texas.

^bFort Lewis College, Durango, Colorado.

^cNASA Johnson Space Center, Houston, Texas.

biological time scale. Terms such as "subperiod" or "biophase" are used to refer to the time from one stage to another. The rate of development is generally expressed in numeric stage units per unit time; occasionally, however, a heat unit, such as growing degree days, is a surrogate for time in the expression of crop development.

A key motivation to derive crop development (crop calendar) models is the idea that, with such a model, environmental variables measured throughout some critical subperiods in the crop life cycle would provide better predictors of grain yield. In addition, crop calendar models were needed to provide a tool for the analyst-interpreter (AI) to use in identifying spectral signatures of wheat fields throughout the growing season.

The initial step in crop calendar modeling must be a review of the physiology of development. This review points out the required assumptions that must be made in order to model crop development over a large geographic region.

PHYSIOLOGY OF DEVELOPMENT

A search of the literature on the physiology of development reveals that there are many theoretical considerations to be taken into account in building a model of crop development. In the general developmental process, photosynthesis, respiration, translocation, and differentiation are key mechanisms which are controlled by a highly complex and interactive system of climatic and nutritional factors. The complexity of this system is magnified by the fact that experimental results show that, for any of these mechanisms and the overall developmental process, the limiting values of temperature, moisture, and soil nutrients frequently change over subperiods in the life cycle. For example, those climatic or nutritional conditions which do not sharply limit early vegetative development may drastically inhibit floral development.

Because of this complexity, specific developmental mechanisms are difficult to model without laboratory controls and frequent measurements of important climatic and soil characteristics. Some simulation models exist for these mechanisms, but their required inputs are inappropriate for use in a large regional modeling effort such as LACIE. Many of these simulation models require frequent and detailed measurements of the crop, soil, and climate.

Crop development models applicable to large regions, such as the U.S. Great Plains, should not be

overparameterized or made to require types of input variables not frequently available from the network of climatic stations. Because the vast majority of climatic stations only record daily maximum and minimum temperatures and precipitation, comprehensive climatic variable input for a crop development model is impractical. Solar radiation, humidity, wind, soil temperature, and soil moisture—all important factors to crop development and the mechanisms which govern it—must be excluded from consideration in building crop models, except to the extent that they can be submodeled. Despite these severe limitations of input data, the problem of quantifying and modeling wheat development over a large geographical region was addressed by LACIE scientists. Some of the pertinent assumptions required for this research were (1) the phenotypic characteristics of wheat express the developmental process well and are observable; (2) wheat is relatively stable ecotypically, and genetic variance does not confound the quantitative development scale; (3) the development of wheat can be modeled with a minimum of climatic data; and (4) the within-year spatial variability in the occurrence of specific stages is relatively uniform.

THE WHEAT PHENOLOGICAL SCALE USED IN LACIE

Stages of development for cereals have been defined by Feekes, Large, Jensen and Lund, Robertson, Williams, Haun, and, most recently, Waldren (refs. 7, 8, 9, 1, 4, 10, and 11, respectively). Some of these scales of development have not been accepted by farmers and agricultural scientists because they are based on small morphological changes which are not readily apparent, especially at the later stages of crop development. Still others require careful field observations to identify stages. The developmental scale used by Robertson (ref. 1) for spring wheat was chosen by LACIE because it uses the predominantly visible phenotypic characteristics to identify a limited number of stages in wheat and because it evolved from years of research in several climatically diverse locations.

The different stages (biological times) and their corresponding numbers on the Robertson quantified development scale appear in table I. Pictures of the successive stages of development are presented in figure 1.

Primary candidate crop calendar models for initial testing in LACIE were the Nuttonson (ref. 12)

TABLE 1.—The Robertson Phenological Scale for Wheat

Phenological characteristic	Stage
Planting	1.0
Emergence	2.0
Jointing	3.0
Heading	4.0
Soft dough	5.0
Ripe	6.0

photothermal linear model and the Robertson (ref. 1) triquadratic spring wheat model, which included nonlinear and interaction expressions of maximum and minimum temperatures and day length. Stuff (ref. 13) analyzed the Nuttonson data and concluded that there were significant nonlinear effects in the temperature and day-length variables. Robertson's treatment of these variables conformed with the findings of Stuff.

Another study which motivated LACIE to begin with the Robertson spring wheat model was Feyerherm's (ref. 14) finding that it accurately predicted heading and maturity of winter wheat in Kansas. Subsequent application to winter wheat development in other regions has shown some large errors. Refinements are being made in the application of the Robertson model to winter wheat. A function to account for vernalization in this crop is needed to improve this model.

Since both planting dates and varieties of wheat vary across the U.S. Great Plains and other large producing regions, there is a problem of uniformity in

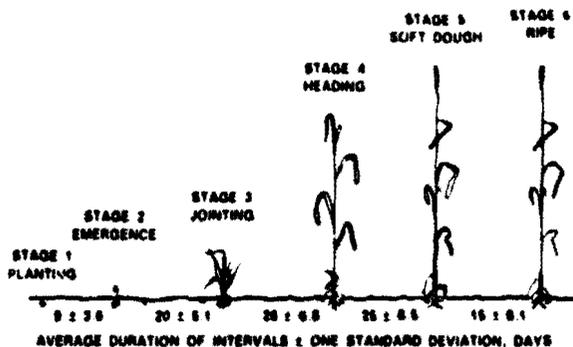


FIGURE 1.—Illustration of Robertson's biological time scale with interval spacing for Marquis spring wheat in Canada.

the occurrence of specific stages. Thus, for the purpose of defining the status of wheat in any particular crop reporting district (CRD), the dates of successive stages of development refer to the times when 50 percent of the crop in that district has reached the specified stage. The types of field observations which the U.S. Department of Agriculture Economics, Statistics, and Cooperatives Service (USDA ESCS) keeps are compatible with this definition. Their estimates of crop stage in each CRD are always given as the percentage of total crop acreage which has reached a particular stage. The USDA ESCS data were valuable to LACIE in testing the accuracy of various crop calendar models built on the Robertson phenological scale.

LACIE APPROACH TO CROP CALENDAR MODELS FOR WHEAT

The types of wheat development models and their extent of sophistication have been limited primarily by the available data. The operational aspect of LACIE requires these models to use a minimum of meteorological variables as daily inputs. The most common and readily available of these are daily maximum and minimum temperatures and precipitation totals. A description of the model types which have been developed or applied during LACIE follows.

Multivariate Least-Squares Techniques¹

Phenological data as reported at the CRD level and environmental data from the National Climatic Center were used in developing and testing an adjustable crop calendar model for winter wheat. Generalized least-squares techniques were applied for parameter estimation of functions to predict the winter wheat phenological stage, with environmental values as independent variables. The independent variables investigated included daily maximum temperature (T_p), daily minimum temperature (T_m), daily day length (D_L), and daily precipitation (P_p).

The outstanding feature of the generalized multivariate least-squares procedure used for parameter estimation is the fact that the sums of squares of residuals for all independent variables are simultaneously minimized. In using this approach, a

¹J. R. Baker.

condition equation set is linearized by Taylor's series. These linearized equations are used, with Lagrangian multipliers, to augment the least-squares condition, resulting in a general constrained minimum to be enforced. Solution of the resulting total normal equation set by matrix partitioning results in parameter estimates and their associated variance-covariance matrix. The method may be used in an iterative manner for nonlinear functional forms.

After parameter estimation, tests were conducted on independent data. From these tests, it may generally be concluded that exponential functions have little advantage over polynomials. Precipitation was not found to significantly affect the fits. The Robertson's triquadratic form, in general use for spring wheat, was found to show promise for winter wheat; however, special techniques and care are required for its use. In most instances, equations with nonlinear effects were found to yield erratic results when used with daily environmental values as independent variables. Thus, as of this writing, the linear function of the form

$$R = H_1 + H_2 T_e + H_3 T_m + H_4 D_L$$

is recommended, where R represents the daily rate of development. Specific coefficients, designated H_n , recommended for inclusion and testing in the LACIE project are given in table II.

Specific recommendations for further work include preparation and inclusion of additional data in the least-squares programs; preparation of more extensive testing programs and data, to include investigation of the effects of using averaged environ-

mental data for predictions; further work on the Robertson's triquadratic model; and variance propagation studies.

Iterative Regression²

The effort to rederive the Robertson spring wheat model by fitting it to phenological data for winter wheat began with assembling the 50-percent-stage dates for 23 CRD's in 7 states. These data varied in definition as well as completeness. For example, some states reported booting instead of jointing and turning instead of soft dough. A summary of the stage data available appears in table III.

A single representative station was selected for each CRD, and a meteorological data base was built which included the growing season for each location year. These data were daily observations of maximum and minimum temperatures. Day length was computed from the Julian date and the latitude of the station.

The iterative regression fitting technique developed by Robertson is described elsewhere in this volume (Whitehead et al., "Growth Stage Estimation"). The original coefficients for the spring wheat model (table IV) were used as seeds in fitting each stage. An improved fit was found in the emergence-to-jointing and the soft-dough-to-ripe stages. The rederived model for emergence to jointing was particularly successful in reducing the bias at jointing (table V). However, the accumulating positive bias was disastrous if the model was run from the observed planting date (table VI). The original Robertson coefficients for planting to emergence, jointing to heading, and heading to soft dough were retained. The complete set of coefficients for the rederived form appears in table VII.

The latest modeling attempt has been an effort to identify and incorporate a moisture variable in a triquadratic model. Instead of precipitation amounts, precipitation occurrence (rain days) was selected as a single-station variable which could represent an entire CRD. This variable was transformed from a [0,1] form to a decimal value by means of a low-pass-filter function designed to simulate a 30-day moving average. This function computed a daily value which explains roughly 95 percent of the variance of occurrence in the preceding 30 days. This filtered rain-day

TABLE II.—Coefficients Recommended for Inclusion and Testing in LACIE

Stage (a)	H_1	H_2	H_3	H_4
P-E	-0.014919	0.0038970	0	0
E-J	-0.0039918	0.0043509	0	0
J-H	-0.216419	0	0	0.19021
H-S	0.314583	0	0	-0.18610
S-R	0.244711	0.0046211	0.0015439	-0.022684

¹P-E = planting to emergence, E-J = emergence to jointing, J-H = jointing to heading, H-S = heading to soft dough, S-R = soft dough to ripe

²M. H. Trenchard.

TABLE III.—Phenological Data Available for Winter Wheat by Stages and State

State	Years	No. of CRD's	Development stage of crop				
			P-E	E-J	J-H	H-S	S-R
Colorado	1972-75	5	20	11	10	0	0
Idaho	1973-75	2	5	0	0	0	0
Kansas	1967-73	5	0	0	30	31	34
Missouri	1970-73	1	0	0	0	4	4
Montana	1971-75	2	0	0	0	10	10
North Dakota	1971-75	5	0	0	10	19	0
Oklahoma	1967-73	2	21	21	21	19	0
Total		23	46	32	71	83	48

TABLE IV.—Characteristic Coefficients Developed by Robertson for the Spring Wheat Crop Calendar

Coefficient	Development stage of crop				
	P-E	E-J	J-H	H-S	S-R
a_0	$V_1 = 1.0$	8.413	10.93	10.94	24.38
a_1	$V_1 = 1.0$	1.005	.9256	1.389	-1.140
a_2	$V_1 = 1.0$	0	-.06025	-.08191	0
b_0	44.37	23.64	42.65	42.18	37.67
b_1	.01086	-.003512	.0002958	.0002458	.00006733
b_2	-.0002230	.00005026	0	0	0
c_1	.009732	.0003666	.0005943	.00003109	.0003442
c_2	-.0002267	-.000004282	0	0	0

variable applied to the same data set as the rederivation except that it was substituted for the day-length variable in the Robertson form. The results showed improvement over the original form in all stages except planting to emergence (table V). When all models were run sequentially from planting, the results were quite reasonable (table VI). Further testing of these models is underway with expansion of an independent test set. Further development is still possible in the planting-to-emergence stage.

However, now it may be concluded that a moisture variable may be successfully substituted for day length in the Robertson triquadratic form. The stages for which these coefficients were calculated appear in table VIII. The three stages in which the moisture quadratic appears have coefficients which describe a concave-down quadratic which expresses a decrease in growth as precipitation occurrence increases. This is logical in experience with maturity rates of moisture-stressed wheat.

TABLE V.—Errors in Days Given the Observed Stage and Starting Date

Error	P-E	E-J	J-H	H-S	S-R
	1-2	2-3	3-4	4-5	5-6
<i>Original</i>					
Bias	6.78	16.97	-8.46	-1.77	6.71
RMSE ^a	9.32	20.45	9.66	6.03	7.72
<i>Rederived</i>					
Bias	12.07	6.25	0.51	0.61	0.19
RMSE	13.38	11.58	4.77	4.31	3.84
<i>New (RDCC^b)</i>					
Bias		-1.47	0.62	-0.08	-0.25
RMSE		10.06	5.04	4.23	3.54
n ^c	46	32	71	83	48

^aRMSE = root-mean-square error.

^bRDCC = rain-day crop calendar.

^cn = number of observations.

TABLE VI.—Error in Days When Run From the Observed Planting Date

Error	P-E	E-J	J-H	H-S	S-R
	1-2	2-3	3-4	4-5	5-6
<i>Original</i>					
Bias	6.78	20.00	0.57	-1.19	7.15
RMSE	9.32	24.73	8.20	7.00	11.40
<i>Rederived</i>					
Bias		28.57	24.50	24.40	22.53
RMSE		35.92	33.60	28.67	27.71
<i>New (RDCC)</i>					
Bias		1.65	2.26	2.18	0.49
RMSE		17.57	12.83	11.18	11.74
n	46	94	102	103	55

TABLE VII.—Characteristic Coefficients for the Winter Wheat Crop Calendar Based on CRD Data

Coefficient	Development stage of crop				
	P-E	E-J	J-H	H-S	S-R
a_0	$V_1 = 1.0$	0.8413×10^1	0.1093×10^2	0.1094×10^2	0.1262×10^2
a_1	$V_1 = 1.0$	$.1005 \times 10^1$.9256	$.1389 \times 10^1$	$.1224 \times 10^1$
a_2	$V_1 = 1.0$	$-.8311 \times 10^{-1}$	$-.6025 \times 10^{-1}$	$-.8191 \times 10^{-1}$	0
b_0	0.4437×10^2	$.1971 \times 10^2$	$.4265 \times 10^2$	$.4218 \times 10^2$	$.4779 \times 10^2$
b_1	$.1086 \times 10^{-1}$	$.2202 \times 10^{-3}$	$.2958 \times 10^{-3}$	$.2458 \times 10^{-3}$	$.6146 \times 10^{-6}$
b_2	$-.2230 \times 10^{-3}$	$-.3376 \times 10^{-5}$	0	0	$.3178 \times 10^{-9}$
c_1	$.9732 \times 10^{-2}$	$.2707 \times 10^{-4}$	$.5943 \times 10^{-3}$	$.3109 \times 10^{-4}$	$.1511 \times 10^{-5}$
c_2	$-.2267 \times 10^{-3}$	$.1215 \times 10^{-6}$	0	0	$-.5998 \times 10^{-7}$

TABLE VIII.—Characteristic Coefficients of New (RDCC) Winter Wheat Crop Calendar

Coefficient	Development stage of crop				
	P-E	E-J	J-H	H-S	S-R
a_0	$V_1 = 1.0$	-1.272×10^0	-8.482×10^{-2}	$V_1 = 1.0$	-5.015×10^{-1}
a_1	$V_1 = 1.0$	1.409×10^0	1.391×10^1	$V_1 = 1.0$	6.373×10^{-4}
a_2	$V_1 = 1.0$	-6.099×10^{-1}	-1.607×10^1	$V_1 = 1.0$	-2.478×10^{-4}
b_0	4.437×10^1	2.207×10^1	6.090×10^1	5.042×10^1	3.632×10^1
b_1	1.086×10^{-2}	4.331×10	2.277×10^{-3}	2.930×10^{-3}	-3.461×10^0
b_2	-2.230×10^{-4}	-6.646×10^{-9}	-3.377×10^{-5}	-4.425×10^{-5}	9.340×10^{-2}
c_1	9.732×10^{-3}	1.035×10^{-4}	1.966×10^{-4}	9.090×10^{-6}	1.183×10^1
c_2	-2.267×10^{-4}	2.953×10^{-9}	2.230×10^{-5}	-6.813×10^{-6}	-2.712×10^1

Working Day Concept³

In order to utilize phenological development models, knowledge about the planting date is required. In general, sufficient information on the actual planting date is not readily available in a timely manner. Hence, the starter model should be considered an integral first stage of a complete phenological development model. Complementary studies by Feyerherm, Stuff and Phinney, and Lytle et al. (refs. 15 to 17, respectively) have used meteorological and simple agronomical information to predict planting dates for spring and winter wheat.

Feyerherm's study considered the effects of temperature and precipitation on accumulated warming/planting (WP) days. The general form of the model was as follows.

1. $WP = 0$ if $TA \leq 32$
2. $WP = \alpha (TA - 32)(PRE)$ if $32 < TA \leq 32 + 1/\alpha$
3. $WP = 1$ if $TA > 32 + 1/\alpha$

³D. E. Phinney.

where TA = the average daily air temperature (°F), α = the threshold value, and PRE = a value between 0 and 1 as a function of the previous 3 days of precipitation. His study found that for spring wheat, $\alpha = 0.1$. No statistically significant precipitation effect was found, and $PRE = 1$ was ultimately used.

The date for 50-percent planting of spring wheat was estimated from a degree-day-type summation, beginning on January 19. When the accumulated warming/planting days reached 35.5, it was assumed that 50 percent of the crop had been planted.

Stuff and Phinney developed an equation for the daily rate of spring wheat planting based on temperature, precipitation, and the normal planting date.

$$R = -0.77 + 0.045(T) - 0.032(P) + 0.053(N)$$

where R = the daily rate of planting, T = the average daily temperature, P = the total daily precipitation, and N = the normal planting date (actual date).

Lytle et al. (ref. 17) derived area-specific equations for each CRD in South Dakota as a function of temperature, precipitation, trend, and the difference

between precipitation and the Thornthwaite (ref. 18) potential evapotranspiration.

To date, a comparative test of spring wheat starter models over the same test set has not been carried out. A preliminary analysis indicates that both the general models have a standard error of estimate near 6.5 days in North Dakota. The corresponding figure for the area-specific Lytle model is approximately 4.5 days. An analysis of the errors imparted by incorrect start dates is required to evaluate whether or not the more universal formulations are satisfactory for general application.

To date, no starter model for winter wheat has been developed which shows improvement over the use of the normal planting date.

SUMMARY AND CONCLUSIONS

The success of the LACIE crop calendar models was remarkable, considering the number of factors which limited their potential accuracy. Over several years, the use of USDA ESCS normal planting dates as initiators of the biometeorological time scale (BMTS) may not impart significant error to a crop calendar model. However, for any specific year, there may be a large error component associated with the use of normal planting dates as starters for the BMTS. For this reason, further work in starter models, such as the Feyerherm model, is recommended.

Because of the geographic scale to which these models were applied, much of their inaccuracy can be attributed to spatial errors in the variables. The degree to which point-source meteorological data from selected stations adequately describe the conditions in a CRD requires further evaluation. An objective analysis procedure is currently being developed at the NASA Johnson Space Center to study the sensitivity of crop calendar models to station density. Additional model error which is spatial in origin may be attributed to the estimates of crop stages at the CRD level made by the USDA ESCS.

Other spatial sources of error are in the differences in soils, management practices, and varieties over large regions. These factors may have unaccounted for effects on crop development. Unfortunately, they are difficult to quantify and incorporate into a crop calendar model. An effort to solve this problem has been undertaken by LACIE.

Lastly, the model form itself may contribute a large proportion of error. The original Robertson ap-

proach was the triquadratic multiplicative model using temperature and day length to predict the stage of crop development. Further development of additive models in which one variable may override or substitute for another is recommended. Nix (ref. 19) has developed a model in which temperature and day length are completely substitutive. On the other hand, the law-of-the-minimum approach (Cate et al., discussed elsewhere in this volume) to crop development may be valuable as an alternative model since it incorporates the use of critical levels for the input variables.

In physiological research, there is no conclusive evidence to show that the integration of environmental effects manifested by crop development follows an interactive, additive, or law-of-the-minimum mode exclusively. For this reason, all model forms should be comparatively evaluated.

REFERENCES

1. Robertson, G. W.: A Biometeorological Time Scale for a Cereal Crop Involving Day and Night Temperatures and Photoperiod. *Int. J. Biometeorol.*, vol. 12, 1968, pp. 191-223.
2. Hanway, J. J.: How a Corn Plant Develops. Iowa State Univ. of Sci. and Technol. Spec. Rep. 48, 1971.
3. Vandertip, R. L.: How a Sorghum Plant Develops. Kansas State Univ. Bull. C-447, 1972.
4. Williams, G. D. V.: Deriving a Biophothermal Time Scale for Barley. *Int. J. Biometeorol.*, vol. 18, 1974, pp. 57-69.
5. Major, D. J.; Johnson, D. R.; Tanner, J. W.; and Anderson, I. C.: Effects of Day Length and Temperature on Soybean Development. *Crop Sci.*, vol. 15, 1975, pp. 174-179.
6. Seeley, M. W.: The Development and Applications of Biological Time Scale Models for Field Corn Hybrids in Nebraska. Ph.D. dissertation, Univ. of Nebraska (Dept. of Horticulture), 1977.
7. Feekes, W.: DeFarwe en haar milleu. Vers. XVII Tech. Farwe Comm. (Groningen), 1941, pp. 560-561.
8. Large, E. C.: Growth Stages in Cereals — Illustrations of the Feekes Scale. *Plant Path.*, vol. 3, 1954, pp. 128-129.
9. Jensen, L. A.; and Lund, H. R.: How Cereal Crops Grow. *North Dakota Agr. Ext. Bull.* 3, 1971.
10. Haun, J. R.: Visual Quantification of Wheat Development. *Agron. J.*, vol. 65, 1973, pp. 116-119.

11. Waldren, R. P.: Growth Stages of Wheat (*Triticum aestivum*). Ph. D. dissertation, Univ. of Nebraska (Dept. of Agronomy), 1977.
12. Nuttonson, M. Y.: Wheat-Climate Relationships and the Use of Phenology in Ascertaining the Thermal and Photo-Thermal Requirements of Wheat. American Inst. Crop Ecol. (Washington, D.C.), 1955.
13. Stuff, R. G.: Preliminary Evaluation of Phenological Models for Spring Wheat. Technical Memorandum LEC-6134, Lockheed Electronics Co., Houston, Texas, 1975.
14. Feyerherm, A. M.: Adjusting Robertson's Biometeorological Time Scale to Winter Wheat Environments and Varietal Maturities. Kansas State Univ. (unpublished report), Feb. 23, 1976.
15. Feyerherm, A. M.: Planting Date and Wheat Yield Models. Final Report on Contract NAS 9-14533, NASA Johnson Space Center, 1976.
16. Stuff, R. G.; and Phinney, D. E.: Climatic and Statistical Analysis of Spring Wheat Planting in the Dakotas. Proceedings of the 68th Annual Meeting of the American Society of Agronomy, 1976.
17. Lytle, W. F.; Strommen, N. D.; and LeDuc, S.: Meteorological Variability of the Planting Date of Wheat. Final Report on Federal Grant USDA (NOAA) 04-6-158-4400, 1977.
18. Thornthwaite, C. W.: An Approach Toward a Rational Classification of Climate. Geog. Rev., vol. 38, 1948, pp. 55-94.
19. Nix, H. A.: Climate and Crop Productivity in Australia. Unpublished paper for the Division of Land Use Research, CSIRO (Canberra, Australia), 1977.

New Developments in Sampling and Aggregation for Remotely Sensed Surveys

A. H. Feiveson^a

BASIC SAMPLING AND AGGREGATION IN LACIE

In a large-scale crop inventory project like LACIE, where procedural and resource constraints severely restrict the total sample size, it is of prime importance to optimize the placement of samples and to achieve the greatest possible accuracy when constructing a large-area crop acreage estimate from individual sample observations. If the crop distribution in a country is fairly stable from year to year and a comprehensive set of historical data exists, then that data can be used effectively for distributing samples and for ratio-estimating acreage for regions where sampling is sparse or nonexistent.

In the United States, the distribution of wheat does not vary greatly from year to year and the 5-year census of county-level crop statistics is readily obtainable and of good quality. For these reasons, county-level historical data have been used throughout LACIE to allocate samples in the U.S. Great Plains.

Counties were divided into three categories of sampling density (Group I, Group II, and Group III) on the basis of data from the last available agricultural census (either 1969 or 1974), so as to minimize the best a priori estimate of the variance of the total U.S. Great Plains wheat production estimate. Details of the guiding philosophy and actual mechanics of the categorization and sampling are given in the paper by Hallum et al. entitled "Sampling, Aggregation, and Variance Estimation for Area, Yield, and Production in LACIE" and in the paper by Feiveson et al. entitled "LACIE Sampling Design."

In addition to placing the samples, the census data were also used to obtain ratioed acreage estimates for counties or groups of counties not sampled (either by

design or because of loss of data). Essentially, this was done by assuming that the ratio of the wheat acreage of a nonsampled county to that of nearby sampled counties was the same for the year of LACIE estimation as it was for the historical census year. For details, the reader is again referred to the papers by Hallum et al. and Feiveson et al.

PROBLEM AREAS IN SAMPLING

In some foreign areas (Brazil, China, Argentina, India, and the U.S.S.R.), historical data are not available at the same level of detail as in the United States; furthermore, the accuracy of the data is unknown. For example, in the U.S.S.R., the smallest political region for which published historical wheat acreage and production data exist is the oblast, which is considerably larger than a U.S. county. By relying solely on the historical data for the U.S.S.R., one could do no better than throw samples randomly or perhaps systematically within an oblast. In Phases II and III of LACIE, it was observed that, because some oblasts were very large, their effectiveness as strata was diminished; i.e., large within-stratum sampling variation was found. Even in the United States, where county-level information was available, large sampling errors were suspected in certain areas. Consequently, an effort was made to develop and test a stratification procedure based on natural topographic or climatic variables rather than on political boundaries.

This experimental effort, called the Natural Sampling Strategy Test (NSST), was run in parallel with the standard LACIE procedure during Phase III to determine whether areas of uniformity with respect to climate, soil type, and distribution of agriculture would make better strata than oblasts or their equivalent in countries without detailed historical data.

To determine the validity of the hypothesis, areas having the above uniformity characteristics were

^aNASA Johnson Space Center, Houston, Texas.

delineated in North Dakota and Kansas and in three oblasts in the U.S.S.R. The areas were typically larger than U.S. counties but smaller than states. These homogeneous areas or "agrophysical units" (APU's) were used as strata for subsequent sampling and aggregation.

In order for the U.S. experiment to simulate the situation in a foreign country without detailed data, only state-level historical data were used to allocate samples and construct ratio estimates where necessary. Because the APU boundaries did not generally coincide with state boundaries, it became necessary to apportion the historical wheat acreage of a state to those parts of the APU's that were included in the state. This was done by assuming that a particular APU's share was proportional to its agricultural area, which was known approximately.

One problem with having strata cross political boundaries is that, in some countries (such as the U.S.S.R.), the amount of wheat planted might be determined by political considerations. If this is the case, the wheat distribution might change drastically when a political boundary within a country is crossed. To protect against this possibility, it was decided in the U.S. simulation to construct substrata by intersecting APU's with states. These substrata, called refined strata, were then treated as strata if they were large enough to be allocated at least one sample unit; otherwise, their wheat acreage was estimated by ratioing. Since only state-level data were used to construct the ratios, this procedure could have caused considerable bias in a particular refined stratum; however, the amount of wheat involved was only a small fraction of the total.

Results of the NSST were inconclusive. In the United States, the NSST estimate of wheat acreage in Kansas was not significantly different from the LACIE estimate in terms of standard errors. (The NSST estimate was actually closer to the U.S. Department of Agriculture (USDA) estimate than to the LACIE estimate.) In the U.S.S.R., no reliable third estimate of wheat acreage was available. In two of the three oblasts estimated, the official LACIE and NSST estimates of acreage were less than 2 standard errors apart; however, in Kurgan, the NSST estimate was about 3 times the LACIE estimate, a difference of about 20 standard errors! It was suspected that large errors were caused by the apportionment procedure for computing ratios. For details, see the paper by Hallum and Basu entitled "Natural Sampling Strategy," which describes the NSST in depth.

In summary, it appeared that the NSST might be

feasible for stratifying countries without detailed historical data but that great problems still remained in determining sample allocations and in ratio estimation for unsampled areas.

Although not reported in these proceedings, R.W. Thomas and C.M. Hay of the University of California at Berkeley tried to improve the approach of the NSST by using a two-phase sampling scheme. This approach obtained crude estimates of wheat acreage for areas covered by a full Landsat frame (about 90 by 90 nautical miles) and then used the crude estimates to allocate samples for intensive study within the area. This two-phase sampling scheme was presented at the 1977 Symposium on Machine Processing of Remotely Sensed Data.

PROBLEM AREAS IN AGGREGATION

Wheat proportion estimates were probably much more reliable in some LACIE segments than in others. For example, not all segments have the same acquisition history; some have all the data necessary to make a good estimate, whereas others may have data from only one Landsat pass. This large variance in reliability has been ignored until now; i.e., in the aggregation process, wheat proportions estimated from a minimal amount of data were treated the same as those estimated from complete sets. As a result, some very poorly estimated segments had too large a part in determining the final large-area wheat production estimate. To alleviate this problem, supporting research has been conducted to develop a weighted aggregation scheme in which each stratum estimate would be a weighted average between "direct" (based only on current sample segments) and "historical" (based on Group III ratio) estimates. This scheme is designed to give more weight to the historical estimate when segment estimates are thought to be unreliable, and vice versa. For details of this procedure, see the paper by Feiveson entitled "Weighted Aggregation."

Another way in which a LACIE-type survey could be improved is to use segment proportion estimates from previous years in obtaining the current-year estimate. At Texas A & M University, supporting research has been in operation to develop a procedure for both sampling and aggregation that uses data from previous years to the best advantage. This procedure is described in the paper by Hartley entitled "Multiyear Estimates for the LACIE Sampling Plans."

CONCLUSIONS

In summary, there have been four sampling and aggregation research tasks carried out in parallel with LACIE operations—the natural sampling strategy, two-phase sampling, weighted aggregation, and

multiyear estimation. Although little or no operational testing of theoretically derived procedures has been done to date, it appears that future LACIE-type surveys should consider a thorough testing of some or all of these methods for possible implementation.

Natural Sampling Strategy

C. R. Hallum^a and J. P. Basu^b

INTRODUCTION

Background

The LACIE sampling strategy is designed to estimate cost-effectively and with predesignated precision the wheat area and production in countries of interest. The level of precision depends on the sample size and is adversely affected by the variability or heterogeneity of wheat density and yield. Stratification is a means of effectively reducing this heterogeneity and improving the efficiency of the LACIE sample design.

During Phase II of LACIE, the methodology was developed for using Landsat imagery and agrophysical data to permit an improved stratification in foreign areas by ignoring political boundaries and restratifying along boundaries that are more homogeneous with respect to the distribution of agricultural density, soil characteristics, and average climatic conditions. These considerations formed the basis for the decision in Phase II to redesign the sampling strategy for the purpose of having an improved foreign sampling strategy based on natural stratification and using Landsat imagery with less dependence on historical data. The primary motivation, then, for a redesign of the initial sampling strategy was to use a stratification in foreign areas that is more homogeneous than political subdivisions. The resulting strategy would provide a common approach for all countries and should permit the same precision (as achieved with the initial design) but with fewer segments. Its use domestically was planned to permit better applicability of the "yardstick" region as a quantifier of foreign results. The former LACIE sampling strategy is referred to as the "initial" or "first-generation" sampling strategy, the latter as the

"natural" or "second-generation" sampling strategy. Because of the manner in which the natural strata were created, they are commonly referred to as "agrophysical units" (APU's).

The remainder of this paper is a description of the natural stratum-based sampling scheme and the aggregation procedures for estimating wheat area, yield, and production and their associated prediction error estimates. A summary of test results will be given including a discussion of the various problems encountered.

Phase III Scope of the Natural Sampling Strategy

The natural sampling strategy was implemented for Phase III in an off-line mode for two states (Kansas and North Dakota) in the U.S. Great Plains and three oblasts (Kurgan, Kustanay, and Tselinograd) in the U.S.S.R. spring wheat indicator region. The initial sampling strategy was retained in an operational mode over these areas for the purpose of comparing the estimates from the two strategies. The natural sampling strategy design for the U.S. yardstick area was developed using procedures and data input requirements so that the performance parameters estimated from the U.S. evaluation would be as applicable as possible to the U.S.S.R. region. Moreover, the Phase III evaluation of the natural sampling strategy was conducted in parallel with and over the same regions as the Feyerherm yield estimation model (see the paper by Feyerherm and Paulsen entitled "A Universal Model for Estimating Wheat Yields"). In summary, the motivation for the Phase III scope of the natural sampling strategy was to describe and test (using Kansas and North Dakota as quantifiers from the U.S. yardstick region) the sampling scheme as well as the procedures for aggregating estimates of wheat area, yield, and production and their associated prediction error estimates in LACIE foreign areas.

^aNASA Johnson Space Center, Houston, Texas.

^bLockheed Electronics Company, Houston, Texas.

Phase III Tests

Several tests were performed during Phase III to evaluate the effectiveness and efficiency of the natural sampling strategy. These included evaluations of the extent of increased homogeneity of yield as well as that of the agricultural and wheat density achieved by the restratification; also, comparisons were made of various aggregations as summarized in the following paragraphs.

Estimates of wheat area and production and their associated variances were obtained at different times during the growing season in the two states in the United States and the three oblasts in the U.S.S.R. The natural sampling strategy area and production estimates/statistics were then compared with those of the initial sampling strategy and of the Statistical Reporting Service (SRS) (in the United States). Estimates were made with the natural sampling strategy using, first of all, a completely new set of segments (i.e., no attempt was made to use any previous LACIE segments). These segments were those resulting from a direct application of the standard stratified random sampling scheme applied directly to the second-generation strata. Estimates were also made using a statistically feasible mixture of first- and second-generation strategy segments¹ (i.e., determinations were made in regard to which of the first-generation sample segments together with a subset of the second-generation sample segments result in a random sample within each stratum of the natural sampling strategy—this procedure is detailed in appendix A of this paper). The primary motivation for attempting to use as many of the first-generation strategy segments as possible was to permit the analysts to use the collected history available on such segments.

The previously described inputs were made in combination with (1) the use of the Feyerherm yield model applied at the natural stratum level (over Kansas) and (2) the use of the Center for Climatic and Environmental Assessment (CCEA) yield model applied at the political subdivision level. (Details of the CCEA model are included in the paper by Strommen et al. entitled "Development of LACIE CCEA-I

¹The sample segments randomly distributed within the second-generation strata are referred to as second-generation segments; those randomly distributed within the political subdivision (county, oblast, etc.) strata of the first-generation strategy are referred to as first-generation segments.

Weather/Wheat Yield Models.") Comparisons of the production estimates from these two inputs provided a further evaluation of the advanced (Feyerherm) yield model in conjunction with the area estimator.

SAMPLING, ESTIMATION, AND AGGREGATION FOR AREA

With the exception of the use of a different stratification and differences in the associated sets of aggregation logic, the initial and natural sampling strategies are similar. Consequently, parts of the following discussion will be redundant relative to some of the material in the Experiment Design section.

Sample Selection Procedure

The second-generation sampling strategy uses

1. A stratified random sampling without replacement scheme
2. "Natural" strata developed according to specifications oriented toward achieving homogeneity in regard to the distribution of agricultural density, soil characteristics, and average yearly climatological conditions
3. The 5- by 6-nautical-mile segment as the sampling unit

The total sample size allocated to the area of interest is such that enough segments will be available for Classification and Mensuration Subsystem (CAMS) processing to achieve a preassigned coefficient of variation for the at-harvest estimate of production allowing for errors due to (1) sampling, (2) classification, (3) yield prediction, and (4) loss of data.² The sampling frame is generated using the same procedure as described in the paper by Hallum et al. entitled "Sampling, Aggregation, and Variance Estimation for Area, Yield, and Production in LACIE."

²The choice of the preassigned value for the production coefficient of variation is dependent on the desired probabilistic accuracy of the production estimate for the area of interest. For example, if the 90/90 criterion is to be satisfied at harvest at the country level, then the goal is to obtain a country production estimate, at harvest, which is within 10 percent of the actual production with a probability of 0.9.

Stratification

The strata are developed based on soil types, climatic conditions, and agricultural density. The suitability of different soils is ranked for growing wheat and is rated on the basis of several soil characteristics such as texture, depth, water-holding capacity, drainage, salinity, and slope. The stratification procedure is oriented toward achieving the same soil suitability rating and similar agricultural density within each stratum. Also, the annual precipitation at any two areas in a given stratum is not to differ by more than 50 millimeters and the average growing season temperature is not to differ by more than 1° C. The resulting strata are referred to as the "natural strata" or the "agrophysical units." Further details of the stratification effort are included in appendix B of this paper.

To obtain estimates at levels such as state or oblast, the intersections of the political subdivision with the natural strata are used in the aggregation. The strata that result from these intersections are referred to as the "refined strata." In a country such as the U.S.S.R., considerable differences in agricultural practices frequently exist, for political or other reasons, between two contiguous oblasts. Consequently, resorting to the refined strata as the base-level strata is the step taken to include "political influence" as a stratification variable.

Sample Allocation

Sample allocation refers to the determination of the total number of segments to be distributed among the strata. These determinations are completed, first of all, for the natural strata. The sample sizes determined for the natural strata are then apportioned to the refined strata based on a proportional allocation using the proportion (relative to the natural strata) of historical wheat, from the epoch year, present in the refined strata.

Determination of total sample size.—The total sample size allocated to the area of interest is such that the LACIE precision goal for the production estimate will, expectedly, be met allowing for errors due to (1) sampling, (2) classification, (3) yield prediction, and (4) loss of data. The best available a priori knowledge of the magnitude of these errors was used together with the following assumptions.

1. Segment-level wheat area estimates are mutually independent and unbiased.

2. Yield estimates are unbiased, are mutually independent (at the yield stratum level), and are independent of the acreage estimates.

On this basis, it is straightforward to approximate the mean-squared prediction error for production as a function of the total sample size n . In accordance with the "Neyman" or "optimal" allocation (ref. 1), minimizing this expression as a function of n subject to the constraint

$$\sum_{i=1}^L n_i = n$$

i.e., such that the summed stratum sample size is the same as the total sample size, results in the following choice for n .

$$n = \frac{\left[\sum_{j=1}^J N_j \sqrt{S_j^2 (Y_j^2 + T_j^2)} \right]^2}{\delta \left[P^2 CV^2(P) - \sum_{j=1}^J A_j^2 T_j^2 \right]} \quad (1)$$

where n = total number of segments allocated to the area of interest

N_j = total number of agricultural (ag) segments in the j th natural stratum

S_j^2 = estimate of segment-to-segment variation of wheat area within the j th natural stratum

Y_j = average yield potential determined from soil characteristics in the j th natural stratum

T_j = standard deviation of yield potential in the j th natural stratum

J = total number of strata in the area of interest

$CV(P)$ = preassigned value of the coefficient of variation of the production estimate

$$P = \sum_{j=1}^J A_j Y_j$$

A_j = estimate of wheat area in the j th natural stratum based on historical information

δ = a conservative lower-bound estimate of the expected sample acquisition rate by the end of a crop season (determined from previous experience with loss of segments due to cloud cover or other reasons)

In case n is not an integer, it is rounded upward to the integer just larger than n .

To compute the wheat area variance S_j^2 , the following procedure is performed. If the j th natural stratum contains at least M_j LACIE-processed segments from the previous year (M_j defaults to the value of 3 in the LACIE software if no overriding value is specified), S_j^2 will be the computed segment-to-segment variation of the previously estimated LACIE wheat area. In case the number of LACIE segments is less than M_j , the estimate of the within-stratum wheat area variance will be obtained from the following regression equation.

$$S_j^2 = (\text{segment area})^2 \left[r_j^2 S_{a_j}^2 + c(a_j^2 + S_{a_j}^2) r_j (1 - r_j) \right] \quad (2)$$

where a_j = ag area in the collection of ag segments of the j th natural stratum computed from a complete enumeration of ag proportion in each 5- by 6-nautical-mile segment in the stratum as determined from a 5-by-6 grid overlay onto Landsat color-infrared imagery

r_j = g_j / a_j , where a_j is as defined previously and g_j is the historical wheat area in the j th natural stratum

$S_{a_j}^2$ = the segment-to-segment within-stratum ag area variance in the collection of ag segments of the j th natural stratum computed from the enumeration of the 5-by-6 ag proportions

and where c is a constant obtained by performing a least-squares fit to other strata (each of which contains at least M_j LACIE segments) in the country containing LACIE-processed segments from the previous year. (If insufficient LACIE segments are available in the country of interest, the least-squares fit is performed on strata from analogous areas from other countries having LACIE-processed segments from the previous year.) The form of equation (2) is identical to that discussed in the paper by Feiveson et al. entitled "LACIE Sampling Design"; details of its derivation are included therein.

Finally, in countries having no historical data available, since S_j^2 cannot be computed from equation (2) in this case, S_j^2 is replaced by $S_{a_j}^2$ in equation (1).

Distribution of sample sizes among strata.—After determining the total number n of segments to be allocated to the area of interest, the sample size n_j to allocate to the j th natural stratum is computed as follows. (Again, the following is a result of the Neyman allocation procedure (ref. 1).) Let t_j be the weight associated with the j th natural stratum, where

$$t_j = \frac{N_j \sqrt{S_j^2 (Y_j^2 + T_j^2)}}{\sum_{j=1}^J N_j \sqrt{S_j^2 (Y_j^2 + T_j^2)}} n \quad (3)$$

and where the quantities to be input into the right side of equation (3) are as defined for equation (1). Let n_j be the provisional allotment to the j th natural stratum, where n_j equals the integer part of t_j . Also let

$$d_j = t_j - n_j$$

and

$$d = n - \sum_{j=1}^J n_j$$

Assign one additional sample segment to each of d of the natural strata with probabilities proportional to d_j ; i.e., let

$$n_j = n_j + \delta_j, \quad j = 1, 2, \dots, J \quad (4)$$

where $\delta_j = 1$ if the j th natural stratum receives an extra sample segment; $\delta_j = 0$ otherwise. The Hartley-Rao procedure (ref. 1) is used to perform this allotment.

Performing the previously described procedure results in the determination of the natural stratum sample sizes. However, wheat area and production estimates are made at the refined stratum level and

require determinations of sample sizes for each refined stratum. To determine the sample sizes for each refined stratum within a given natural stratum, a proportional allocation is performed as follows. The sample size for each refined stratum is determined by multiplying the natural stratum sample size by the proportion (relative to the natural stratum) of historical wheat, from the epoch year, present in each given refined stratum. The fractional part of these weights accounts for one or more segments; consequently, these segments are assigned to the refined strata using the Hartley-Rao proportional procedure with sample sizes proportional to the fractional parts of the weights in the same manner as indicated in the previous paragraph.

Area Estimation

Area estimation is performed as follows.

Designation of Group A and B refined strata.—The j th refined stratum is designated as "Group A" or "Group B" depending on whether it is allocated at least M_j or less than M_j segments, respectively (M_j defaults to the value of 3 if no overriding value is specified). The qualitative definitions are as follows: Group A refined strata represent marginal to high wheat-producing areas historically, whereas Group B refined strata represent areas having very little or no wheat historically. There is one primary exception to this general qualification: since LACIE does not get coverage on every segment for every pass, the j th refined stratum is placed in the Group B category if it has less than M_j segments after such losses. If M_j or more segments subsequently become available for aggregation, the j th refined stratum is reinstated as Group A.

Guidelines for locating segments.—The location of segments within a refined stratum is performed by simple random sampling without replacement within the previously designated ag area of the stratum. All the segments with 5 percent or more agricultural area within the refined stratum are candidates for selection. After the selection of the sample, the selected segments are located on a mosaic and the latitude and longitude of the center of each segment are obtained. Except for application to a different set of strata, the guidelines for locating segments are identical to those applied in the initial sampling strategy; consequently, the reader should refer to the paper by Liszcz entitled "LACIE Area Sampling Frame and Sample Selection" for further details.

Apportionment of political subdivision data to strata.—As noted in the section entitled "Sample Allocation," it may be necessary to know the natural stratum historical wheat area g_j to compute r_j for input into equation (2). This quantity will also be needed for each refined stratum in the aggregation. In countries having historical data available only on one level smaller than the country (e.g., the oblast level in the U.S.S.R.), an apportionment of the political subdivision wheat proportional to the natural/refined stratum is performed. It is particularly important to note that historical data do not exist at the refined stratum level in any country; consequently, until a historical data base can be built (i.e., with the passage of time), the apportioning procedure is the outlet taken to estimate the historical wheat area to associate with each refined stratum. The underlying assumption made in this approach is that the wheat in a political subdivision is uniformly distributed over the agricultural area in that political subdivision. Of course, this is not always the case and it is fully recognized that apportioning has its difficulty and represents an initial attempt at resolving the missing-data problem. The apportioning is accomplished as follows.

$$g_j = \frac{a_{ij}}{a_i} W_i \quad (5)$$

where g_j = the apportioned estimate of historical wheat area in the j th natural/refined stratum

a_{ij} = the ag area in the collection of ag segments common to the j th natural/refined stratum and the i th political subdivision

W_i = the wheat area in the i th political subdivision based on historical data

a_i = the total ag area in the collection of ag segments of the i th political subdivision

Apportionment of area estimates to yield strata.—The natural sampling strategy requires that yield and area (APU) strata be coincident. The situation may arise, however, where the strata do not coincide. In such cases, an apportionment of the acreage estimates to the yield strata is needed to permit the estimate of production at political subdivision (or other) levels. The following discussion specifies the approach taken in this case.

Another set of area strata is generated consisting of the collection of all the area "substrata" that result from the intersection of the refined strata with the yield strata. This procedure results in each yield stratum containing one or more area substrata. The LACIE wheat area estimate for each area substratum is obtained by apportioning the LACIE wheat area estimate of the refined stratum according to the proportion of ag area in each substratum relative to the refined stratum from which it came.

With reference to figure 1, if the indicated yield strata covered the area of interest with $\{A_1, A_2, A_3, A_4, A_5\}$ being the collection of refined strata over this same area, then, in view of the lack of coincidence of area and yield strata, a new collection of area substrata is generated consisting of $\{A_{11}, A_{12}, A_{21}, A_{22}, A_{31}, A_{32}, A_{41}, A_{42}, A_{51}, A_{52}, A_{53}, A_{54}\}$.

If \hat{W}_{ij} is used to denote the LACIE-estimated wheat area apportioned to A_{ij} , then

$$\hat{W}_{ij} = \begin{cases} \frac{a_{ji}}{a_j} A_{jA} & \text{if the } i\text{th substratum is contained} \\ & \text{within the } j\text{th Group A refined} \\ & \text{stratum} \\ \frac{a_{ji}}{a_j} A_{jB} & \text{if the } i\text{th substratum is contained} \\ & \text{within the } j\text{th Group B refined} \\ & \text{stratum} \end{cases} \quad (6)$$

where a_{ji} = the total ag area of A_{ji} determined from the product of the ag proportion of A_{ji} (determined from the complete enumeration of ag in the gridded segments) with the planimeted area of A_{ji} , a_j = the total ag area of A_j (determined from the complete enumeration of ag in the gridded segments), and where A_{jA} and A_{jB} are defined in equations (7) and (8), respectively.

The previously described procedure identifies the area "substrata" and the associated wheat area esti-

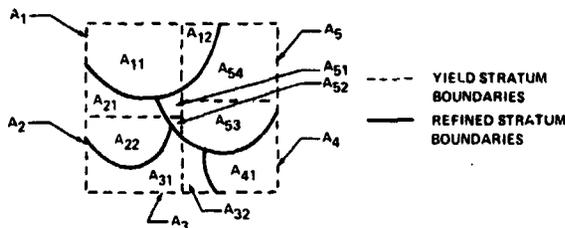


FIGURE 1.—A refinement of acreage strata to the yield strata.

mates to be aggregated within each yield stratum to compute production. Again, it should be emphasized that this procedure is used in estimating production and production prediction error only when the yield strata are not coincident with the natural strata; however, the wheat area estimate for a given area of interest will be the same regardless of whether the area aggregation is performed from the refined strata or from the substrata.

Group A refined stratum area estimate.—The wheat area estimate for the j th Group A refined stratum is calculated as follows.

$$A_{jA} = \frac{N_{jA}}{M_{jA}} \sum_{k=1}^{M_{jA}} A_{jkA} \quad (7)$$

where A_{jA} = the LACIE wheat area estimate for the j th Group A refined stratum

N_{jA} = the total number of ag segments in the j th Group A refined stratum

M_{jA} = the total number of sample segments for which wheat acreage estimates are made in the j th Group A refined stratum

A_{jkA} = the LACIE wheat area estimate for the k th sample segment in the j th Group A refined stratum

Group B refined stratum area estimate.—The wheat area estimate of the j th Group B acreage stratum is calculated as follows

$$A_{jB} = \sum_{k=1}^m C_{kj} A_{kA} + (n_{jB}/M_j) A'_{jB} \quad (8)$$

where

$$C_{kj} = \left(1 - \frac{n_{jB}}{M_j}\right) \left(\frac{d_{kj} W_{jB}}{\sum_{k'=1}^m d_{k'j} W'_{k'A}}\right)$$

W_{jB} = the most recent epoch year (or average over previous 2 or 3 years) harvested wheat area in the j th Group B stratum

$$d_{kj} = \begin{cases} 1, & \text{if the } k\text{th Group A stratum is used in the estimate of the } j\text{th Group B stratum} \\ 0, & \text{otherwise} \end{cases}$$

$$\sum_{k=1}^m d_{kj} A_{kA} = \begin{cases} \text{the sum of the LACIE estimates of wheat area in all Group A strata that are used to estimate the given } j\text{th Group B stratum} \\ \text{the sum of the most recent epoch year (or average over previous 2 to 3 years) harvested wheat area in the Group A strata that are used to estimate the } j\text{th Group B stratum} \end{cases}$$

$$\sum_{k=1}^m d'_{kj} W'_{kA} = \begin{cases} \text{the sum of the most recent epoch year (or average over previous 2 to 3 years) harvested wheat area in the Group A strata that are used to estimate the } j\text{th Group B stratum} \end{cases}$$

m = the total number of Group A strata in the area of interest plus any additional Group A strata that are not contained within the area of interest but that are used in the ratio estimation of the Group B strata contained in the area of interest

n_{jB} = the total number of sample segments for which acreage estimates are made in the j th Group B stratum

and where $A'_{jB} = 0$ if $n_{jB} = 0$; otherwise

$$A'_{jB} = \frac{N_{jB}}{n_{jB}} \sum_{k=1}^{n_{jB}} A_{jkB} \quad (9)$$

where N_{jB} = the total number of ag segments in the j th Group B refined stratum

A_{jkB} = the LACIE wheat area estimate for the k th sample segment in the j th Group B stratum

In the Crop Assessment Subsystem (CAS) software, the capability exists for M_j to default to the value of 3 if no overriding value is specified.

It should be reemphasized at this point that the most refined level for which historical wheat acreage data are available in many LACIE foreign areas is one level below the country level (crop region or economic region). Consequently, until a historical data base is generated after passage of time, the historical wheat value for each stratum used in the Group B estimator is apportioned to the stratum as indicated in the section entitled "Apportionment of Political Subdivision Data to Strata."

In the preceding discussion, the Group A strata to be used as a base for the ratio-estimated part of the wheat area estimate of a given Group B stratum are selected according to the following guidelines.

1. First of all, the capability exists in the CAS software to permit interactive selection of the appropriate Group A strata as a base in the ratio-estimated part of a given Group B stratum. This capability is particularly advantageous in allowing the crop analysts to incorporate real-time information and expertise that may be available at the time of aggregation to assist in selecting Group A strata to use as a base in the ratio estimation. For example, knowledge of agricultural practices and/or information concerning the status (such as the presence of an episodic event) of crops in different localities can be very beneficial in deciding which Group A strata to use in ratio-estimation of a given Group B stratum.

2. When option 1 is not used, the CAS software defaults to the use of all Group A strata in the zone containing the given Group B stratum.

Area aggregation to the zone, region, and country levels.—The wheat area estimate \hat{A} of the area of interest (whether it be zone, region, or country) is given by

$$\hat{A} = \sum_{j=1}^m n_j A_{jA} + \sum_{l=1}^b A_{lB} \quad (10)$$

where $n_j = \begin{cases} 1, & \text{if the } j\text{th Group A refined stratum is contained within the area of interest} \\ 0, & \text{otherwise} \end{cases}$

and m is the total number of Group A strata contained in the area of interest plus any additional

Group A strata not contained within the area of interest but used in the ratio estimation of the Group B strata contained in the area of interest. The quantity b denotes the total number of Group B strata in the area of interest.

AREA VARIANCE ESTIMATION

Group A Refined Stratum Variance

From equation (7), it is straightforward to see that the variance of the estimate of wheat area for the j th Group A refined stratum is given by

$$V_{jA} = \frac{N_{jA}^2}{M_{jA}} S_{jA}^2 \quad (11)$$

where V_{jA} = the variance of the estimate of wheat area in the j th Group A refined stratum

$$S_{jA}^2 = \left[1/(M_{jA} - 1) \right] \sum_{k=1}^{M_{jA}} (A_{jkA} - \bar{A}_{jA})^2 \quad (12)$$

and

$$\bar{A}_{jA} = (1/M_{jA}) \sum_{k=1}^{M_{jA}} A_{jkA} \quad (13)$$

The finite population correction factor, $1 - (M_{jA}/N_{jA})$, is omitted from equation (11) since it is almost always insignificantly different from 1. (When this is not so, equation (11) is a conservative estimate—i.e., an estimate on the upper side).

Group B Refined Stratum Variance

The variance of the estimate of wheat area for the j th Group B stratum is directly obtainable from equa-

tion (8) and is given by

$$V_{jB} = \sum_{k=1}^m C_{kj}^2 V_{kA} + (n_{jB}/M)^2 V'_{jB} \quad (14)$$

where V'_{jB} is the estimate of segment-to-segment variance of wheat area computed for the refined stratum in the manner indicated in the section entitled "Determination of Total Sample Size" (i.e., by making use of the allocation data at the refined stratum level) if $n_{jB} \leq 1$; otherwise (i.e., if $2 \leq n_{jB} < M_j$), V'_{jB} is estimated directly (i.e., in the same manner as the estimate of V_{jA} (eq. (11)).

Variance Aggregation to the Zone, Region, and Country Levels

After substituting the expression in equation (8) into equation (10) and simplifying, it is straightforward to see that \hat{A} in equation (10) is expressible as

$$\hat{A} = \sum_{j=1}^m \alpha_j A_{jA} + \sum_{i=1}^b (n_{iB} A'_{iB} / M_i) \quad (15)$$

where

$$\alpha_j = n_j + \sum_{i=1}^b C_{ji}$$

Consequently, the variance V_A for the estimate of wheat area for the area of interest is given by

$$V_A = \sum_{j=1}^m \alpha_j^2 V_{jA} + \sum_{i=1}^b (n_{iB}/M_i)^2 V'_{iB} \quad (16)$$

SPRING AND WINTER WHEAT

Area and Variance Estimation

In a mixed wheat area, separate area estimates are made for the winter wheat and the spring wheat

using the aggregation procedures described in the sections "Group A Refined Stratum Area Estimate" through "Variance Aggregation to the Zone, Region, and Country Levels" with inputs (LACIE estimates as well as historical) of winter wheat for a winter wheat aggregation and those of spring wheat for a spring wheat aggregation. This method provides the spring wheat and winter wheat area estimates and their respective variance estimates at the stratum and zone levels. The aggregation procedures are used to obtain separate winter wheat and spring wheat area estimates and the corresponding variance estimates at the zone, regional, and country levels.

Total Wheat: Area and Variance Estimation

The total wheat area estimate in a mixed wheat area is computed by aggregating the winter wheat and spring wheat area estimates for the area; that is, if \hat{A}_w and \hat{A}_s denote the winter and spring wheat area estimates, respectively, the total wheat area estimate \hat{A}_t is given by

$$\hat{A}_t = \hat{A}_w + \hat{A}_s \quad (17)$$

This computation is done at the zone and higher levels.

The variance estimates for the total wheat at the zone and higher levels are those obtained from the total wheat aggregation made with inputs of total wheat by CAS for the segments and historical data. The procedure is the same as described in the section entitled "Variance Aggregation to the Zone, Region, and Country Levels."

PRODUCTION ESTIMATION

Production and Variance Estimation at the Stratum Level

The estimate of wheat production for the j th area refined stratum is given by

$$P_j = \begin{cases} A_{jA} Y_{jA}, & \text{if the } j\text{th refined stratum is a} \\ & \text{Group A stratum} \\ A_{jB} Y_{jB}, & \text{if the } j\text{th refined stratum is a} \\ & \text{Group B stratum} \end{cases} \quad (18)$$

In equation (18), Y_{jA} (Y_{jB}) is the predicted yield for the j th Group A (Group B) refined stratum as given by the yield estimation model. Two basic assumptions are made to obtain the LACIE production variance estimator.

1. Segment-level wheat area estimates are mutually independent and unbiased.

2. Yield estimates are unbiased and are mutually independent of the acreage estimates.

Under these assumptions, the variance of the production estimator of the j th Group A refined stratum is given by $\sigma_{P_{jA}}^2$, where

$$\sigma_{P_{jA}}^2 = \sigma_{A_{jA}}^2 \mu_{Y_{jA}}^2 + \sigma_{Y_{jA}}^2 \mu_{A_{jA}}^2 + \sigma_{A_{jA}}^2 \sigma_{Y_{jA}}^2 \quad (19)$$

and where $\sigma_{A_{jA}}^2$, $\mu_{A_{jA}}$ and $\sigma_{Y_{jA}}^2$, $\mu_{Y_{jA}}$ are the respective variance and mean for the acreage and yield estimators for the j th Group A refined stratum. The production variance estimator is the one resulting from the replacement of the parameters in the right side of equation (19) by their estimates; the sign of the last term in equation (19) is changed from positive to negative to obtain an unbiased estimator. Consequently, the resulting estimator is given by Z_{jA}^2 , where

$$Z_{jA}^2 = V_{jA} Y_{jA}^2 + T_{jA}^2 A_{jA}^2 - V_{jA} T_{jA}^2 \quad (20)$$

where A_{jA} and Y_{jA} are the estimates of wheat area and yield, respectively, of the j th Group A refined stratum, V_{jA} is an estimate of the wheat area variance for the j th Group A refined stratum, and T_{jA}^2 is the estimated squared prediction error of yield in the j th Group A refined stratum. The variance of the j th Group B refined stratum production estimate is similarly obtained.

In case the yield strata are not coincident with the natural strata, the estimate of production and its associated prediction error estimate at the substratum

level are obtained as follows. In particular, if P_{ij} denotes the production estimate of the i th substratum (apportioned from the j th yield stratum which has yield Y_j), then

$$P_{ij} = \hat{W}_{ij} Y_j \quad (21)$$

where \hat{W}_{ij} is defined in equation (6). Moreover, denoting the production variance estimate of P_{ij} by Z_{ij}^2 ,

$$Z_{ij}^2 = V_{ij} Y_j^2 + T_r \hat{W}_{ij}^2 - V_{ij} T_r^2 \quad (22)$$

where \hat{W}_{ij} and Y_j are the estimates of wheat area and yield, respectively, of the i th substratum (contained in the j th refined stratum), and

$$V_{ij} = \begin{cases} \left(\frac{a_{ji}}{a_j}\right)^2 V_{jA} & \text{if the } i\text{th substratum is contained} \\ & \text{within the } j\text{th Group A refined} \\ & \text{stratum} \\ \left(\frac{a_{ji}}{a_j}\right)^2 V_{jB} & \text{if the } i\text{th substratum is contained} \\ & \text{within the } j\text{th Group B refined} \\ & \text{stratum} \end{cases} \quad (23)$$

The quantities a_{ji} and a_j are as defined in the section entitled "Apportionment of Area Estimates to Yield Strata."

Production Estimate at a Zone, Region, or Country Level

The wheat production estimate P_A for the area of interest (whether it be a zone, a region, or a country) is given by equation (24) or equation (25), respectively, depending on whether or not the yield strata coincide with the natural strata.

$$P_A = \sum_{r=1}^H \left[\sum_{j=1}^m \gamma_{rj} \eta_{jA} A_{jA} + \sum_{i=1}^b \beta_{ri} A_{iB} \right] Y_r \quad (24)$$

or

$$P_A = \sum_{r=1}^H \left[\sum_{j=1}^m \gamma_{rj} \eta_{jA} A_{jA} + \sum_{i=1}^b \beta_{ri} A_{iB} \right] Y_r \quad (25)$$

That is,

$$P_A = \sum_{r=1}^H A_r Y_r$$

where

$$A_r = \sum_{j=1}^m \left[\gamma_{rj} \eta_j + \sum_{i=1}^b \beta_{ri} C_{ji} \right] A_{jA} + \sum_{i=1}^b \left(\beta_{ri} \eta_{jB} A_{iB} M_i \right) \quad (26)$$

or A_r is given by this same expression after replacing $\gamma_{ij} \eta_j$ by $\gamma_{ij} \eta_{jA}$ and β_{ri} by $\beta_{ri} A_{iB}$. In the situation where the natural strata and yield strata coincide, equation (24) applies, where Y_r is the LACIE-predicted yield in the r th yield stratum, H is the total number of yield strata in the area of interest

$$\gamma_{ij} = \begin{cases} 1, & \text{if the } j\text{th Group A stratum lies within (or} \\ & \text{coincides with) the } i\text{th yield stratum} \\ 0, & \text{otherwise} \end{cases}$$

$$\beta_{ri} = \begin{cases} 1, & \text{if the } i\text{th Group B stratum lies within (or} \\ & \text{coincides with) the } r\text{th yield stratum} \\ 0, & \text{otherwise} \end{cases}$$

and M_i is a variable that defaults to the value of 3 if no overriding value is specified.

In the situation where the natural strata and yield strata do not coincide, equation (25) applies with Y_r and H as defined previously; however,

$$\gamma_{ij} = \begin{cases} 1, & \text{if the } j\text{th Group A stratum intersects the} \\ & \text{ } i\text{th yield stratum} \\ 0, & \text{otherwise} \end{cases}$$

$$\beta_{ri} = \begin{cases} 1, & \text{if the } r\text{th Group B stratum intersects the } r\text{th} \\ & \text{yield stratum} \\ 0, & \text{otherwise} \end{cases}$$

Also

$$q_{jA} = \sum_{q=1}^{m_j} (a_{jq}/a_j) \quad (27)$$

and

$$q_{iB} = \sum_{q=1}^{b_i} (a_{iq}/a_i) \quad (28)$$

In equations (27) and (28), the small a 's are as given in the section entitled "Apportionment of Area Estimates to Yield Strata"; m_j is the total number of substrata that are apportioned out of the j th refined Group A stratum and that lie within the r th yield stratum. Similarly, b_i is the total number of substrata that are apportioned out of the i th refined Group B stratum and that lie within the r th yield substratum.

Production Prediction Error for a Zone, Region, or Country Level

The estimate of the squared prediction error S_p^2 of the production estimate for the area of interest (whether it be zone, region, or country level) is given by

$$S_p^2 = \sum_{r=1}^N \left[Y_r^2 V_{rA} + A_r^2 T_r^2 - V_{rA} T_r^2 \right] + 2 \sum_{r=2}^N \sum_{r'=1}^{r-1} Y_r Y_{r'} \text{COV}(A_r, A_{r'}) \quad (29)$$

and

$$\text{COV}(A_r, A_{r'}) = \sum_j \alpha_{jr} \alpha_{j r'} V_{jA}$$

where

$$\alpha_{jr} = \gamma_{rj} \eta_j + \sum_{i=1}^b \beta_{ri} C_{ji}$$

T_r^2 = the estimated squared prediction error of yield in the r th yield stratum

and

$$V_{rA} = \sum_{j=1}^m \left[\gamma_{rj} \eta_j + \sum_{i=1}^b \beta_{ri} C_{ji} \right]^2 V_{jA} + \sum_{i=1}^b (\beta_{ri} \eta_{jB} / M_i)^2 V_{iB} \quad (30)$$

The preceding equations apply, of course, to the situation in which the natural strata and yield strata are coincident. If this should not be the case, the same equations would still apply provided $\gamma_{rj} \eta_j$ and β_{ri} are replaced with $\gamma_{rj} \eta_{jA}$ and $\beta_{ri} q_{iB}$, respectively, throughout.

SPRING AND WINTER WHEAT PRODUCTION

Production Estimation in Mixed Wheat Areas

In a mixed wheat area, separate production and predicted production error estimates are made for the winter wheat and spring wheat using the procedures described in the sections entitled "Production and Variance Estimation at Stratum Level," "Production Estimate at a Zone, Region, or Country Level," and "Production Prediction Error for a Zone, Region, or Country Level." The total wheat production estimate in a mixed wheat area is computed by aggregating the winter production and the spring production; that is, if \hat{P}_w and \hat{P}_s denote the winter and spring wheat production estimates, respectively, the total wheat production estimate, \hat{P}_t , is given by

$$\hat{P}_t = \hat{P}_w + \hat{P}_s \quad (31)$$

Production Error Estimates in Mixed Wheat Areas

The estimate of the production prediction error at the level of interest (zone, region, or country) in a mixed wheat area is given by equations (29) and (30) with the following modifications; if the i th Group A refined stratum contains mixed wheat and is supplied with both a spring Y_r^s and winter Y_r^w yield estimate, then Y_r and T_r^2 in equations (29) and (30) are replaced by

$$Y_r = \frac{A_{wr}Y_r^w + A_{sr}Y_r^s}{A_{wr} + A_{sr}} \quad (32)$$

and

$$T_r^2 = \left(\frac{A_{wr}T_r^w + A_{sr}T_r^s}{A_{wr} + A_{sr}} \right)^2 \quad (33)$$

respectively, where

A_{wr} = the epoch year harvested winter wheat area in the i th yield stratum

A_{sr} = the epoch year harvested spring wheat area in the i th yield stratum

T_r^w = the root mean square of the prediction error of the winter wheat yield estimate for the i th yield stratum

T_r^s = the root mean square of the prediction error of the spring wheat yield estimate for the i th yield stratum

AREA OF INTEREST YIELD AND PREDICTION ERROR ESTIMATION

When there is a single yield model covering the area of interest, the squared prediction error of the yield model provides the variance needed for evaluation; i.e., for computation of standard statistics. In case of more than one yield model in the area of interest (whether it be zone, region, or country), the average yield estimate and the associated prediction error estimate is computed using the following formulas.

The yield estimate for the area C of interest is obtained by the average yield $\bar{Y}_C = P_C/A_C$, where P_C is the production estimate and A_C is the area estimate for the area of interest. An estimate of the squared prediction error of this average yield estimate is obtained by resorting to the assumptions stated in the section entitled "Production and Variance Estimation at Stratum Level" and applying the standard formula for the variance of the ratio of two random variables. This estimate is given by

$$U_C = \bar{Y}_C^2 \left[\frac{S_C^2}{P_C^2} + \frac{V_C}{A_C^2} - \frac{2 \sum Y_i V_i}{P_C A_C} \right] \quad (34)$$

where S_C^2 = estimated squared prediction error of P_C , the production estimate

V_C = estimated prediction error of A_C , the area estimate

Y_i = yield estimate for the i th yield stratum

V_i = estimated prediction error of the area estimate in the i th yield stratum (computed from use of eq. (16), where the area of interest referred to by that estimate is the i th yield stratum in this case)

SUMMARY AND CONCLUSIONS

The results from the testing conducted in Phase III would have to be labeled as "encouraging," particularly in regard to their supporting the following objectives for the use of a natural sampling strategy.

1. Increase sampling efficiency: improve stratification by making use of information contained in Landsat and agromet data.
2. Reduce bias caused by high incidence of cloud cover over large regions.
3. Permit better estimates of precision (more sample segments per stratum).
4. Provide a common approach for all countries.
5. Permit better applicability of the yardstick region as a quantifier of the foreign sampling strategy.

In particular, some of the more important results that should be highlighted include the following.

1. Aggregation results over the test areas indicated that similar precision results are obtainable (relative to the initial strategy) with 20 to 30 percent fewer segments.

2. In comparison with crop-reporting-district competing strata, the gain in efficiency from the use of the refined strata over six states in the U.S. Great Plains was

- a. Uniformly better in regard to wheat density
- b. Better for five of six states in regard to agricultural density
- c. Better for only three of six states in regard to yield

See reference 2 for further details.

The predominant key issues resulting from the use of the natural sampling strategy were as follows.

1. Some strata are not yet sufficiently homogeneous to be considered beneficial—the stratification procedures need further fine tuning.

2. Probably the biggest difficulty was that of attempting to estimate those areas where little to no satellite coverage (i.e., the nonresponse areas) was available; adequate historical data to support making

such estimates was simply unavailable. (The apportioning procedure described earlier in this paper was the approach taken to make historical data available at the appropriate levels needed; however, it was realized from the outset that considerable improvement would be needed.)

ACKNOWLEDGMENTS

The authors wish to express their appreciation to A. H. Feiveson and J. Downs for preparation of appendixes A and B, respectively.

REFERENCES

1. Cochran, W. G.: *Sampling Techniques*. Third Ed. John Wiley & Sons, Inc. (New York), 1977.
2. Basu, J. P.; and Dragich, S. M.: *Second-Generation Sampling Strategy Evaluation Report*. LACIE CD-00465, NASA Johnson Space Center, Houston, Texas, July 1978.

Appendix A

Determination of First- and Second-Generation Segment Mixture

MOTIVATION

Because of cost and time constraints, it may not be possible to order or, even if ordered, to process all second-generation sample segments. To obtain an estimate of wheat production with a precision comparable to the precision specified in the sampling plan, it is necessary to process a certain minimum number of sample segments. One way to fulfill the sample size requirement is to supplement the list of available second-generation segments with the available first-generation segments. Unless the randomness of the distribution of sample segments in each refined stratum (second-generation strategy strata) is preserved, any statistical statement concerning the sample estimates made according to the second-generation strategy will be invalid. A scheme has been devised for the selection of supplementary first-generation segments preserving the randomness of distribution in each refined stratum.

METHOD FOR SELECTING SUPPLEMENTARY FIRST-GENERATION SEGMENTS

For a detailed mathematical discussion, see the section in this appendix entitled "Method for Using First-Generation Sample Segments in the Second-Generation Sampling Scheme."

Segments chosen under the first-generation strategy and segments chosen under the second-generation strategy (second-generation segments) are available, labeled with both county name and stratum number. The following procedure is used for selecting supplementary first-generation segments.

1. Count the number of second-generation segments in each unit (county intersection refined strata).
2. Count the number of first-generation segments in each unit.

3. For each unit, perform the following operation.
 - a. If the number of first-generation segments is greater than or equal to the number of second-generation segments, randomly replace each second-generation segment with a first-generation segment.
 - b. Otherwise, randomly replace second-generation segments by first-generation segments until all first-generation segments in the unit have been used. Some second-generation segments will remain.

4. Note that step 3b differs from the theoretical proposal. Theoretically, all first-generation segments should be used, and the remaining number necessary for the unit should be selected at random from all segments possible within the NASA Goddard Space Flight Center (GSFC) constraints. However, the second-generation segments have already been chosen and must be used.

5. Check the spacing between segments. When the first-generation and second-generation segments were chosen, the segment density in each case was constrained by GSFC. The same constraint should be presented in the composite allocation.

RESULTS

The method was followed for the segments currently available (first-generation segments and second-generation segments). Segments chosen are listed in the section in this appendix entitled "Sample Segments." When these segments were checked for closeness, it was found that the plan cannot be executed if the GSFC constraints are to be preserved.

METHOD FOR USING FIRST-GENERATION SAMPLE SEGMENTS IN THE SECOND-GENERATION SAMPLING SCHEME

The method for using first-generation segments in the second-generation scheme was developed by A. H. Feiveson of the NASA Johnson Space Center. The following definitions apply to this procedure.

- S = new stratum
- $\{\theta_k\}_{k=1}^L$ = collection of first-generation strata which intersect S
- N_k = total number of segments in θ_k
- n_k = number of selected segments in θ_k under the first-generation strategy

- M_k = total number of segments in $\theta_k \cap S$
- m = number of segments to be selected in S under the second-generation strategy
- M = total number of segments in S
- m_k = number of segments to be selected in θ_k

The steps in the procedure are as follows.

1. Generate $\{m_k\}_{k=1}^L$.
 - a. Let $T_0 = 0, T_1 = M_1, T_2 = M_1 + M_2, \dots, T_L = M_1 + M_2 + \dots + M_L$.
 - b. Define $J_k = \{T_{k-1} + 1, T_{k-1} + 2, \dots, T_{k-1} + T_k\}$.
 - c. Choose a random subset of m from the integers $1, 2, \dots, M$. Let I be that random subset.
 - d. Let m_k = cardinality of $I \cap J_k$. Note that m_k has a hypergeometric distribution. Hence,

$$\alpha_{kj} = (P \ m_k = j)$$

$$= \frac{\binom{M_k}{j} \binom{M - M_k}{m - j}}{\binom{M}{m}} \quad (A1)$$

and

$$E(m_k) = \frac{mM_k}{M} \quad (A2)$$

2. Let I_k = number of first-generation selected segments in $\theta_k \cap S$.
 - a. If $I_k \geq m_k$, choose m_k segments at random among the I_k originally selected ones.
 - b. If $I_k < m_k$, choose all of the I_k originally selected segments plus $m_k - I_k$ additional ones randomly selected from the remaining $M_k - I_k$ in $\theta_k \cap S$.
3. Prove that this procedure selects m segments out of M with equal probability.
 - a. Let $\theta_{k,i} = Pr\{I_k = i\}, i = 0, 1, \dots, M_k$.
 - S = second-generation sample.
 - S_0 = first-generation sample.
 - s_k = any potential segment in $\theta_k \cap S$.

b. Then

$$P(s_K \in S) = \sum_{i=0}^{M_K} \theta_{Ki} P\{s_K \in S/I_K = i, m_K = i\} \quad (\text{A3})$$

$$P(s_K \in S/I_K = i) = \sum_{j=0}^{M_K} \alpha_{Kj} P\{s_K \in S/I_K = i, m_K = j\} \quad (\text{A4})$$

$$P(s_K \in S/I_K = i, m_K = j) =$$

$$P(s_K \in S/I_K = i, m_K = j, s_K \in S_0)$$

$$\bullet P(s_K \in S_0/I_K = i, m_K = j)$$

$$+ P(s_K \in S/I_K = i, m_K = j, s_K \notin S_0)$$

$$\bullet P(s_K \notin S_0/I_K = i, m_K = j) \quad (\text{A5})$$

$$P(s_K \in S/I_K = i, m_K = j, s_K \in S_0) = \begin{cases} 1, & 1 \leq i < j \\ j/i, & i > j \\ 0, & i = 0 \end{cases} \quad (\text{A6})$$

$$P(s_K \in S_0/m_K = j, I_K = i) = i/M_K \quad (\text{A7})$$

$$P(s_K \in S/I_K = i, m_K = j, s_K \notin S_0) =$$

$$\begin{cases} 0, & j \leq i \\ (j - i)/(M_K - i), & j > i \end{cases} \quad (\text{A8})$$

$$P(s_K \in S_0/m_K = j, I_K = i) = (M_K - i)/M_K \quad (\text{A9})$$

$$P(s_K \in S/I_K = i, m_K = j) = \begin{cases} 0 + j/M_K, & i = 0 \\ i/M_K + (j - i)/(M_K - i), & i < j \\ j/M_K + 0, & i > j \end{cases} \\ = j/M_K, \quad i = 0, 1, \dots, M_K \quad (\text{A10})$$

c. Thus, $P(s_K \in S/I_K = i, m_K = j)$ only depends on j , not i . Hence,

$$\begin{aligned} P(s_K \in S/I_K = i) &= \sum_{j=0}^{M_K} \alpha_{Kj} \frac{j}{M_K} \\ &= \frac{E(m_K)}{M_K} \\ &= \frac{m}{M} \\ &= P(s_K \in S) \\ &= \frac{m}{M} \sum_{i=0}^{M_K} \theta_{Ki} \\ &= \frac{m}{M} \quad (\text{A11}) \end{aligned}$$

SAMPLE SEGMENTS

The sample segments in Kansas selected prior to the GSFC constraint are the following.

Segment number	County	Stratum
803	Cowley	3C
1170	Harper	3C
1173	Kiowa	3C
800	Kingman	3C
1174	Pratt	3C
1892	Reno	3C
1175	Sedgwick	3C
1893	Stafford	3C
1176	Sumner	3C
1033	Clark	4B
1168	Barber	4B
812	Kiowa	4B
1035	Ford	5A
1292	Hodgeman	5A
822	Pawnee	5A
825	Pawnee	5A
823	Trego	5A
821	Finney	5B
1857	Grant	5B
1025	Greeley	5B
818	Greeley	5B
819	Greeley	5B
1859	Hamilton	5B
1861	Kearny	5B

Segment number	County	Stratum	Segment number	County	Stratum
1284	Lane	5B	1347	Morris	7B
1864	Stanton	5B	1876	Ottawa	7B
833	Bourbon	6	852	Postawatomie	7B
829	Bourbon	6	843	Republic	7B
1180	Cherokee	6	1888	Salina	7B
1345	Franklin	6	1348	Wabaunsee	7B
839	Franklin	6	1158	Washington	7B
837	Linn	6	1016	Cheyenne	8
830	Miami	6	1017	Decatur	8
834	Miami	6	1880	Ellis	8
1353	Montgomery	6	1024	Gove	8
828	Montgomery	6	1153	Jewell	8
836	Osage	6	1027	Logan	8
840	Osage	6	1019	Norton	8
1184	Wilson	6	1155	Phillips	8
832	Wilson	6	1020	Rawlins	8
842	Lincoln	7A	1281	Rawlins	8
1154	Mitchell	7A	1877	Rooks	8
1349	Butler	7B	864	Rooks	8
847	Chase	7B	1887	Russell	8
857	Chase	7B	1022	Sheridan	8
846	Clay	7B	863	Sheridan	8
1151	Clay	7B	1021	Sherman	8
1879	Dickinson	7B	1282	Sherman	8
1297	Dickinson	7B	1157	Smith	8
1881	Ellsworth	7B	1296	Smith	8
853	Marshall	7B	1023	Thomas	8
1884	McPherson	7B	1031	Wallace	9

Appendix B Stratification

REQUIREMENTS

The LACIE Partitioning Group developed universal strata for use by the Sampling Strategy Team for the Phase III implementation of the natural sampling strategy. These strata were created with the following guidelines.

1. The basic partitions are to be delineated on the basis of agricultural density as determined from full-frame color-infrared (CIR) imagery.
2. These partitions are to be refined with data on soils rated on yield potential and with climatic data.
3. Climatic data considered are to include precipitation and temperature.
4. If the variation in climate over an agricultural density stratum exceeds the allowable threshold, this stratum is to be divided in order to lower the variation over each portion.

5. Whereas the climatic data are to be used to determine whether an agricultural density stratum must be divided, the soils data are to be used to determine where the division occurs.

6. The resulting partitions are to be rechecked against full-frame imagery to adjust and smooth stratum boundaries.

7. Political boundaries that are artificial in an agricultural sense are to be ignored in delineating the strata within a large region. The final partitioning product for such a region is to consist of a group of contiguous strata that includes the region.

STEPS FOR IMPLEMENTATION

The Stratification Team performed the partitioning for sampling strategy based on guidelines 1 to 7. Figure B-1 is a flow chart showing the process.

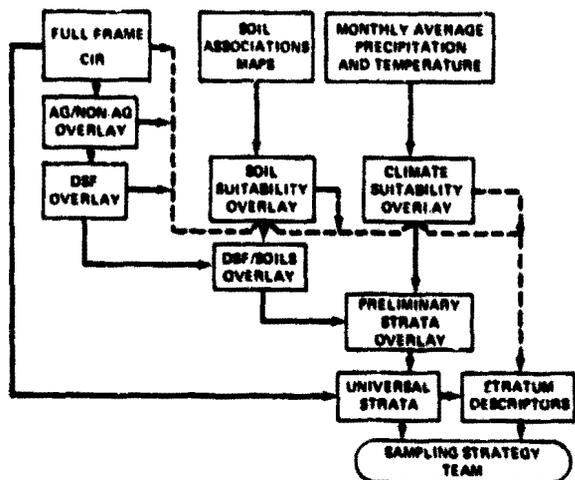


FIGURE B-1.—Stratification for sampling strategy.

Agricultural/Nonagricultural Delineation—Full-Frame CIR Transparencies

The 9- by 9-inch full-frame CIR transparencies used to produce the agricultural/nonagricultural (ag/non-ag) overlay are to be acquired during the optimum growth phase of the majority of crops in the area. Ordinarily, this period will fall during biological window 2 or biological window 3. A record is to be kept of those transparencies containing significant defects such as cloud cover and a list of those that should be reacquired because of unacceptable quality.

Ag/Non-Ag Overlay

The following are guidelines for the delineation of agricultural and nonagricultural areas.

1. All ag/non-ag delineations are to be based solely on the use of 9- by 9-inch full-frame CIR transparencies.
2. No ancillary data, such as agricultural statistics, climatic information, or soil maps, are to be used.
3. An area will be designated as agricultural if it contains recognizable field patterns.
4. Conversely, areas without recognizable field patterns will be designated as nonagricultural.
5. A record will be kept of the acquisition dates for all 9- by 9-inch CIR transparencies that are used.
6. Each state or oblast is to be worked by only one person.
7. A log of the criteria to complete the overlay for each state will be kept.

8. The person working on the overlay for a particular state will produce a short narrative summary including the dates of the 9- by 9-inch CIR transparencies that were used and the percentage of the state incomplete because of clouds.

9. There should be an adequate number of reference points (latitude and longitude) on each overlay so that it can be easily registered to any base map. The steps to be taken in delineating ag/non-ag areas are as follows.

- a. An overlay of 1:1 000 000 scale, covering the state to be worked, will be registered to the corresponding Operational Navigation Chart (ONC).
- b. All 9- by 9-inch full-frame CIR transparencies covering any part of the state will be assembled and examined for evidence of field patterns.
- c. All contiguous nonagricultural areas greater than or equal to 4 square miles in size or with a very sparse sprinkling of agriculture will be delineated on the overlay.
- d. Records for transparencies used should be kept for use in stratum description and evaluation. These records would indicate specific problems in delineation of the transparency such as apparent differences in intensity of agriculture patterns and confusion sources.

Designated Sampling Frame Overlay

The designated sampling frame (DSF) is the region within a country used for the process of partitioning for sampling strategy. The following steps describe this process.

1. The contiguous nonagricultural areas delineated on the ag/non-ag overlay are measured for area.
2. All nonagricultural areas having size greater than or equal to 30 square miles will be eliminated from further consideration. The area on the overlay less the area estimated as nonagricultural comprises the DSF. The DSF, therefore, contains non-agricultural areas ranging in size between 4 and 30 square miles in a segment.
3. On the overlay showing the DSF, areas having approximate agricultural densities falling in the following ranges will be delineated and labeled as to density category.

Density range	Label
0 to 5 percent ag	Low
5 to 40 percent ag	Moderate
40 to 80 percent ag	High
80 to 100 percent ag	Very high

4. Each density category is further refined by outlining high-variance areas (large clumps of non-ag in a predominantly ag area or large clumps of ag in a predominantly non-ag area) and low-variance areas (small clumps of non-ag in a predominantly ag area or small clumps of ag in a predominantly non-ag area). The DSF overlay is thus divided into low- and high-variance areas of each density category. These agriculturally homogeneous areas form the basis for further partitioning.

5. A record of the density category and variance type should be kept for each of these areas for later use in providing stratum descriptors.

Soils Delineation

The A. R. Aandahl map "Soils of the Great Plains" will be used to provide soils data for partitioning the U.S. Great Plains. For the area of the spring wheat indicator region in the U.S.S.R. to be partitioned, the U.S. Department of Agriculture soil survey map "World Soils Map" will be used.

Soil Suitability Overlay

The purpose of the soil suitability overlay is to stratify on the basis of soil yield potential for wheat as follows.

1. The soil mapping units will be rated on their suitability to grow wheat. This will be accomplished by determining a rating in the following six categories of soil characteristics.

Category 1—Texture

Category 2—Depth

Category 3—Drainage

Category 4—Salinity

Category 5—Slope

Category 6—Moisture and Temperature

The numerical rating in each category will range from a value of "1" for best to a value of "4" for worst. The range of values for each rating are provided in table B-1. The overall rating for a particular soil mapping unit will then be determined in the following way. For each i , $1 \leq i \leq 6$, let C_i denote the rating value for the i th category. The overall rating value V is then computed by the formula $V = \max C_i; i = 1, \dots, 6$. In this way, the overall suitability

rating for a soil is determined by its worst rating in an individual category.

2. Adjacent soil mapping units that have the same overall rating will be combined.

3. The resulting 1:1 000 000-scale overlay will contain lines delineating various soil groups, where each of these groups consists of contiguous soil mapping units that have the same rating.

4. The soil characteristics of each soil mapping unit should be recorded for subsequent use in producing stratum descriptors.

TABLE B-1.—Categories of Important Soil Characteristics

Category	Properties—best to poorest
1—Texture	1. Silt, loam, light silty clay loam 2. Sandy loam, heavy silty clay loam 3. Sand, loamy sand, clay, other
2—Depth	1. Deep, 90 cm 2. Moderately deep, 50 to 90 cm 3. Shallow, 25 to 50 cm 4. Very shallow, 25 cm
3—Drainage	1. Well and moderately drained 2. Somewhat poor or somewhat excessive 3. Poor or between somewhat excessive and excessive 4. Very poor or excessive
4—Salinity	1. None, 2 mmhos/cm 2. Slight, 2 to 4 mmhos/cm 3. Moderate, 4 to 8 mmhos/cm 4. Severe, 8 mmhos/cm
5—Slope	1. Level, gently sloping, 0 to 5 percent (symbol L,U,B,F,T,DT) 2. Gently rolling, 5 to 10 percent: (R,DR—lower ranges) 3. Rolling, 10 to 15 percent (R,DR—upper ranges) 4. Hilly, steep, 15 percent (H,C,RD,HD,SU,S,RM)
6—Moisture and temperature	1. Udic mollisols (moist subhumid) 2. Typic mollisols (dry subhumid) 3. Aridic mollisols (semiarid) 4. Aridisols (arid)

Climate

The climatic factors considered in partitioning for sampling strategy are temperature and precipitation. The climatic data used will be based on monthly averages of temperature and precipitation obtained from the World Meteorological Organization (WMO) stations in the areas to be partitioned.

Climate Suitability Overlay

The steps to be taken in producing the climate overlay are as follows.

1. A 1:1 000 000-scale overlay of the area to be partitioned will be registered to the corresponding ONC.
2. For each WMO meteorological station in the area, the annual temperature and precipitation averages will be computed.
3. Using the annual temperature averages from the meteorological stations, temperature isopleths will be interpolated for the area to reflect changes in mean temperature of 3° C.

4. On the same overlay, the annual precipitation averages will be used to produce isopleths of precipitation for the area that reflects changes in mean precipitation of 10 centimeters.

5. The major intent of using the climate overlay is to subdivide large areas of homogeneous soils and uniform agricultural densities.

6. Data used to prepare the overlay (step 1) should be kept for use in stratum description and evaluation.

Ag/Non-ag-Soils Overlay

The ag/non-ag overlay will be checked against the soil suitability overlay for discrepancies and necessary changes made. This will be accomplished by the following steps.

1. A correlation between poor soils rating and nonagricultural areas should be evident. If a non-agricultural area has been rated "1" or "2" on soil suitability, these areas should be rechecked against the appropriate CIR full-frame transparencies to ensure that correct ag/non-ag classifications were made.

Multiyear Estimates for the LACIE Sampling Plans

H. O. Hartley^a

INTRODUCTION

This document presents an approach that may be useful in improving the estimates of the wheat acreages for the LACIE countries for each year by using the short-time series of estimates made in the sequence of consecutive years. Although it may be premature to develop this concept since the series of estimates is just being started, it is of some merit to review this possibility as it will affect future planning.

It is obvious that there will be two types of characteristics of such a survey design; namely,

1. Characteristics that apply to *each* year's survey (such as the size of the sample segments, their stratification, and the sampling procedures with which they are drawn)
2. Characteristics that affect the design and analysis of the survey data arising in a *series* of years

With regard to the characteristics under item 1, in order to fix the ideas, it is assumed that the design is essentially as it is implemented at present. This assumption does not mean that this design is considered the optimum choice for the multiyear estimates.

THE USE OF THE BUREAU OF CENSUS CURRENT POPULATION SURVEY ROTATING DESIGN

The Current Population Survey (CPS) of the Bureau of Census is concerned with sample segments of households and, in the CPS design, these are arranged in "rotation groups." The segments in the same rotation group are surveyed in 4 consecutive months of the first year, then omitted from the survey in the next 8 months, and then again surveyed in 4 consecutive months of the next year. The

estimator of a characteristic y , the so-called composite estimator, is a weighted average of the following two estimator components.

1. The first component simply consists of the best estimator for the current month employing all the data collected in the sample segments for that month.
2. The second component consists of an estimate of δy_t (that is, of the change in y from month $t - 1$ to month t) based only on the matched segments (i.e., the segments that are in the sample in both month $t - 1$ and month t). This change is then added to the composite estimator for month $t - 1$.

Finally, the two components under items 1 and 2 are combined as a weighted average, with weights summing to 1. (Currently, these weights are taken as 0.5 each). It will be seen that the preceding definition of the composite estimator simply defines the composite estimator for month t in terms of the composite estimator for month $t - 1$; that is, in terms of a difference equation. Although this difference equation is used for recurrent computation of the composite estimator of y , the difference equation can be solved to display the composite estimator as an infinite series of monthly estimators with weights exponentially decreasing into the past.

It is the essence of the effectiveness of both the rotating design and the composite estimator that there is a strong positive correlation between the y values in 2 consecutive months. Such a correlation would make the variances of the δy_t small and would thereby increase the effective sample size of segments by those measured in earlier months.

REASONS FOR THE DEPARTURE FROM THE CPS COMPOSITE DESIGNS FOR LACIE

The essential condition for the effectiveness of the composite estimators in rotation designs appears to be well satisfied in LACIE. There is usually a

^aTexas A & M University, College Station, Texas.

strong positive correlation between the wheat acreages of segments observed in consecutive years. In this context, one should remember that since the segment is fairly large (5 by 6 nautical miles), any year-to-year "rotation" of wheat with other crops in accordance with agricultural practices will probably occur *within* segments (apart from boundary effects). Such rotations will therefore generate negative year-to-year correlations of wheat acreages for smaller areas *within* a segment and yet will not destroy the positive year-to-year correlation for segments. However, the following are substantial differences between LACIE and the CPS sampling problems.

1. The LACIE time series is yearly and extremely short (at present, only 2 years) as opposed to the long monthly series in CPS.

2. Whatever the rotation design that is adopted for LACIE, it must be anticipated that a considerable number of matched segments (i.e., segments sampled in 2 consecutive years) will be lost through cloud cover (and possibly other reasons). It will therefore be necessary to replace the composite estimator by a more flexible estimator capable of dealing with unbalanced segment patterns over a moderate number of years.

On this basis, a decision was made to use estimators arising from mixed analysis of variance (ANOVA) models as described in the section of this paper entitled "The Mixed ANOVA Models and the Associated Estimators." These estimators are designed to deal with the completely unbalanced matching patterns that are likely to arise through cloud cover losses of segments, patterns which will differ considerably from any balanced rotation design. Nevertheless, in the next section, suitable rotation designs are developed since they are expected to result at least in *segment* patterns for which the conditions of estimability for the estimators are satisfied.

THE ROTATION DESIGNS

Rotation Designs for Group I Strata

First, the "basic rotation patterns" for segments within a stratum are described; then, the stratum collapsing strategies to deal with strata having only one segment are described.

Basic rotation patterns.—Basic rotation patterns are established as (a.2), two segments per stratum; (a.3), three segments per stratum; (a.4), four segments per stratum; and (a.5), five segments per stratum.

Pattern (a.2), two segments per stratum: Pattern (a.2) is represented by the following table.

Segment no.	Year no.			
	1	2	3	4
1	x			
2	x	x		
3		x	x	
4			x	x
5				x

Pattern (a.3), three segments per stratum: Pattern (a.3) is represented by the following table.

Segment no.	Year no.			
	1	2	3	4
1	x			
2	x	x		
3	x	x		
4		x	x	
5			x	x
6			x	x
7				x

Pattern (a.4), four segments per stratum: Pattern (a.4) is represented by the following table.

Segment no.	Year no.		
	1	2	3
1	x		
2	x		
3	x	x	
4	x	x	
5		x	x
6		x	x
7			x
8			x

Pattern (a.5), five segments per stratum: Pattern (a.5) is represented by the following table.

Segment no.	Year no.				
	1	2	3	4	5
1	x				
2	x				
3	x	x			
4	x	x			
5	x	x			
6		x	x		
7		x	x		
8			x	x	
9			x	x	
10			x	x	
11				x	x
12				x	x

Summary: The general rotation pattern is now clear. If v is the number of segments per stratum, $v/2$ or $(v - 1)/2$ or $(v + 1)/2$ segments are discarded every year in such a way that every segment is in the sample for exactly 2 consecutive years.

Collapsing strategies.—It is first assumed that the total number of strata in Group I is at least two and that each stratum can offer one or more segments to the sample. The collapsing strategies are established as follows.

Strategy (b.2), two strata in Group I: One stratum with one segment per stratum is collapsed with one of the following.

1. The other stratum with one segment to form pattern (a.2)
2. The other stratum with two segments per stratum to form pattern (a.3)
3. The other stratum with three segments per stratum to form pattern (a.4)
4. The other stratum with four segments per stratum to form pattern (a.5)

If both strata have at least two segments, each will be kept separately.

Strategy (b.3), three strata in Group I: The following strategies are employed.

1. Strategy (b.3.0)—If there is no stratum with one segment, all strata will be kept separately.
2. Strategy (b.3.1)—If there is one stratum with one segment, this will be collapsed with another stratum having the smallest number of segments.
3. Strategy (b.3.2)—If there are at least two strata with one segment, all strata with one segment will be

collapsed to form either pattern (a.2) or pattern (a.3); all other strata (if any) will be kept separately.

Strategy (b.4), four strata in Group I: The following strategies are employed.

1. Strategy (b.4.1)—The only stratum with one segment will be collapsed with another stratum having the smallest number of segments.

2. Strategy (b.4.2)—Two strata having one segment each will be collapsed together.

3. Strategy (b.4.3)—If there are three strata with one segment, two will be collapsed together and the third will be collapsed with another stratum having the minimum number of segments.

4. Strategy (b.4.4)—If all four strata have one segment, they will be collapsed in pairs.

Strategy (b.5+), five or more strata in Group I: The collapsing strategy used for five or more strata in Group I will be implemented following the procedure described for strategy (b.3) when the number of strata is odd or strategy (b.4) when the number of strata is even.

Summary: The principle of collapsing is now clear. All strata with one segment are collapsed in pairs, preferably with another stratum having one segment. If the total number of strata having one segment is odd, one triplet of strata is formed to obtain pattern (a.3). If there is only one stratum in Group I, rotation will yield one of the basic rotation patterns ((a.2), (a.3), . . .). If that single stratum has only one segment, no rotation is possible in Group I.

Rotation Design for a Group II Stratum

Each sampled primary unit (county) has only one secondary (segment) precisely. The single Group II stratum in the crop reporting district (CRD) is treated as in the preceding section except that primaries (counties) are rotated in accordance with the basic patterns (a.2), (a.3), If only one primary (county) is in Group II, no rotation is possible. Whenever the same primary is retained for 2 consecutive years, the (single) segment is also retained.

It is clear that any rotation design of primaries would be difficult to implement with a probability proportional to size (PPS) sampling procedure. It is therefore suggested that primaries be sampled with equal probability and without replacement and that the size variable be used as a concomitant variable in conjunction with regression estimation.

THE MIXED ANOVA MODELS AND THE ASSOCIATED ESTIMATORS

Estimation Theory for Group I Strata

The following mixed ANOVA model for the segment wheat acreages in a particular "large area" (say CRD) is adopted.

$$y_{hts} = \alpha_t + \delta\beta_h + \omega x_h + c_{hs} + e_{hts} \quad \text{for } t = 1, \dots, T$$

$$h = 1, \dots, H$$

$$s = 1, \dots, S_h$$

(1)

where y_{hts} = LACIE's "observed" wheat acreage for segment s of stratum h in year t

α_t = average true wheat acreage per segment in year t

$\delta\beta_h$ = differential true effect of stratum h on wheat acreage applicable to all years

x_h = last agricultural-census wheat acreage for county h

ω = regression coefficient for x_h

c_{hs} = true segment variable applicable to all years

e_{hts} = composite segment error variable of segment s of stratum h in year t . This error variable contains two components; namely, the deviation of the true wheat acreage of segment h,s in year t from the additive formula $\alpha_t + \delta\beta_h + c_{hs}$ plus the classification error in y_{hts} .

The mixed ANOVA estimation procedure described in appendix A gives a simple technique of estimating the variance components σ_c^2 and σ_e^2 by "synthesis based" estimators σ_c^2 and σ_e^2 . Moreover, the "fixed" coefficients α_t , $\delta\beta_h$, and ω can be adjoined into a composite regression vector. Without loss of generality, one may assume that the first three terms in equation (1) are of the form $X\gamma$, where X is an orthonormalized form of the fixed-design matrix of equation (1) and γ is an associated reparameterization of the composite vector $\langle \alpha_t, \delta\beta_h, \omega \rangle$. (This follows the procedure described

in section 2 of appendix A.) The maximum likelihood (ML) estimator of γ is then of the form

$$\hat{\gamma} = (X'H^{-1}X)^{-1}(X'H^{-1}y) \quad (2)$$

where the variance-covariance matrix $\sigma_e^2 H$ of the $c_{hs} + e_{hts}$ is given by

$$\sigma_e^2 H = \sigma_e^2 \left\{ I + \frac{\sigma_c^2}{\sigma_e^2} U'U \right\} \quad (3)$$

and U is the design matrix in equation (1) representing the effect of the variables c_{hs} . (Compare also with sections 2 and 3 of appendix A.)

Suppose now that one wishes to estimate the Group I wheat acreage for the CRD in the last year indexed ($t = T$). This estimator is given by

$$\hat{Y}_1(\text{CRD}, T) = N\hat{\alpha}_T + \sum_{h=1}^H N_h \delta\hat{\beta}_h + \omega \sum_{h=1}^H N_h x_h \quad (4)$$

where N_h is the number of segments in stratum h and

$$N = \sum_{h=1}^H N_h$$

is the number of segments in the CRD.

The estimator (eq. (4)) can be written in the form

$$\hat{Y}_1(\text{CRD}, T) = \hat{\gamma}'c \quad (5)$$

and its variance in the form

$$\text{Var } \hat{Y}_1(\text{CRD}, T) = \sigma_e^2 c'(X'H^{-1}X)^{-1}c \quad (6)$$

This variance formula is standard in weighted (Aitken) linear regression theory. The fact that it represents a first-order approximation of $\text{Var } \hat{Y}_1(\text{CRD}, T)$ is proved in appendix B. For an estimation of equations (3) and (6), replace σ_e^2 and σ_c^2 by $\bar{\sigma}_e^2$ and $\bar{\sigma}_c^2$, respectively.

Three questions may arise with this approach. First, it was assumed that the composite errors e_{hts} in equation (1) all have the same variance σ_e^2 . This does not agree with standard sample survey practice, which would have to postulate different within-stratum variances for the true segment wheat acreages in different strata (counties). Similar considerations may also apply to the within-stratum classification variances. The preceding objection could be taken into consideration and different σ_{eh}^2 estimated for each stratum. However, the conditions of estimability described in appendix A are more likely to be violated, in which case a return to a common σ_e^2 would appear to be reasonable. Moreover, the variance-covariance formula (eq. (3)) would now be of the form

$$H = \sum_{h=1}^H U_h U_h' \sigma_{eh}^2 + U_o U_o' \sigma_c^2 \quad (7)$$

where U_o is the design matrix for the c_{hs} and U_h is the design matrix for the e_{hts} .

Strictly speaking, the variables c_{hs} represent samples from finite populations of size N_h . The finite population corrections (fpc's) have here been ignored. Work is in progress to examine the inclusion of the fpc's. The e_{hts} errors are however composed of finite population errors (the within-stratum variation of the true acreages) and an infinite population error variable (the classification error). Again, the fpc's are ignored.

Estimation Theory for a Group II Stratum

The following mixed ANOVA model for the segment wheat acreages in a particular "large area" (say a CRD) is adopted.

$$y_{ts} = \alpha_t + \theta x_s + c_s + e_{ts} \quad \text{for } t = 1, \dots, T$$

$$s = 1, \dots, S \quad (8)$$

where y_{ts} = LACIE's "observed" wheat acreage for segment s in year t . The primary (county) is given the same index as the (single) segment within it.

x_s = known size variable for the primary which contains segment s . Normally, this size variable would be constant over the years.

α_t = average true wheat acreage per segment in year t

θ = regression coefficient of y on x

c_s = true segment variable applicable to all years

e_{ts} = composite segment error variable of segment s in year t . This error variable has three components which will *not* be separately estimated. The three components are (1) the within-primary variable, (2) the primary variable within the Group II population, and (3) the classification error.

The mixed ANOVA procedure described in appendix A can again be used to provide estimates $\bar{\sigma}_c^2$ and $\bar{\sigma}_e^2$ of σ_c^2 and σ_e^2 . Again, one would use the simulated ML estimators of the α_t and θ

$$\hat{\gamma}' = (X'H^{-1}X)^{-1} (X'H^{-1}y) \quad (2)$$

More specifically, the originally fixed design matrix would consist of columns \mathbf{a}_t of 0,1 variables for each of the α_t and a single column of size variables \mathbf{x}_s . This column would have to be orthogonalized on the 0,1 columns and the orthogonal matrix normalized to obtain the new fixed-design matrix X so that the ML estimators $\hat{\gamma}$ for the two matrices are analogous. For the new X matrix, the old \mathbf{a}_t column is replaced by a column \mathbf{f}_t with elements $0,1/\sqrt{n_t}$ where n_t is the number of segments in year t , and the old \mathbf{x} column with elements \mathbf{x}_s is replaced by

$$\frac{\mathbf{x} - \sum_t \bar{x}_t \mathbf{a}_t}{\sqrt{\Sigma}}$$

where \bar{x}_t is the mean of the x_t values in year t and

$$\Sigma = \left(x - \sum_i \bar{x}_{t_i} \right)' \left(x - \sum_i \bar{x}_{t_i} \right)$$

Suppose now one wishes to estimate the total wheat acreage (in the Group II primaries) for the CRD in the last year indexed ($t = T$). This estimator is given by

$$\hat{Y}_{II}(\text{CRD}, T) = N \left\{ \hat{a}_T + \hat{\theta} \bar{X} \right\} \quad (9)$$

where now N is the total number of segments in the population of Group II primaries (counties) and \bar{X} is the population mean of the size variables x_t . The estimator (eq. (9)) can be written in the form

$$\hat{Y}_{II}(\text{CRD}, T) = \hat{y}' d \quad (10)$$

and its variance in the form

$$\text{Var } \hat{Y}_{II}(\text{CRD}, T) = \sigma_e^2 d' (X'H^{-1}X)^{-1} d \quad (11)$$

where $\sigma_e^2 H$ is the variance-covariance matrix of the observation vector y with elements y_{it} . Clearly,

$$\sigma_e^2 H = \sigma_e^2 \left\{ I + \frac{\sigma_c^2}{\sigma_e^2} U'U \right\} \quad (3)$$

where U is the design matrix for the segment variables c_t . For an estimation of equations (11) and (3), replace σ_c^2 and σ_e^2 by $\hat{\sigma}_c^2$ and $\hat{\sigma}_e^2$, respectively.

Of the three questions which were raised previously concerning the appropriateness of the mixed ANOVA model (eq. (8)), the first one, concerning separate strata variances σ_{eh}^2 , does not now arise. However, the other two questions may again be raised about the model and answered in a manner similar to the discussion in the preceding section.

Estimation Theory for a Group III Stratum

The ratio estimator employed in currently used LACIE technology could again be used and would reduce the Group III estimate of the wheat acreage to equations (4) and (9) for the Group I and Group II wheat acreages.

FUTURE WORK

Future work must be concerned with a validation of the assumptions and the theory used in the fifth section of this paper and later with improvements in the rotation designs of the fourth section. In this context, it should be noted that, since the 2 operational years of the past have at least matched some of the segments, it would be possible to estimate the components σ_c^2 and σ_e^2 separately. Moreover, a certain model-monitoring analysis could be performed on the data of the last 2 years and would concern the validity of the model assumptions and specifically the appropriateness of the regression model.

It is the author's intention in the near future to evaluate this method using Phase I to III operational data for the United States.

Appendix A

A Simple "Synthesis"-Based Method of Variance Component Estimation*

H. O. Hartley,^a J. N. K. Rao,^b and L. R. LaMotte^c

1. INTRODUCTION

In this paper, we do not attempt an evaluation of the ever-growing methodology in the estimation of variance components. (For an excellent summary of the literature up to 1971, see reference 1.) Optimality properties are sometimes achieved at considerable computational effort. A case in point is the maximum likelihood (ML) estimation (ref. 2), which is still fairly laborious for large data banks despite the improvements through the W-transformation (ref. 3). Similar observations apply to the *general* case of MINQUE (ref. 4) recently simplified by Liu and Senturia (ref. 5). Other methods, such as the Henderson 3 method (ref. 6) or the abbreviated Doolittle and square root methods (e.g., ref. 7), depend on a subjective ordering of the components (such as with the forward Doolittle procedure), and, if the ordering is unfortunate, the method may fail to yield estimates for certain components, whereas with a different ordering (not attempted), all components may well be estimable. The work involved in attempting all possible orderings of the variance components is usually prohibitive. The present method achieves optimality properties and is nevertheless computationally simple. In fact, not only does it possess MINQUE optimality for a particular choice of norm, but it also simplifies various other optimality properties and necessary and sufficient conditions for estimability associated with MINQUE (see sec. 6). Moreover, we are able to derive sufficient conditions for consistency which also provide estimability conditions of a simpler structure. The

consistency of our estimators makes them convenient as starting points for a single ML cycle to obtain asymptotically fully efficient estimates.

2. THE MIXED ANOVA MODEL

Employing the currently used notation, we write the mixed ANOVA model in the form

$$y = X\gamma + \sum_{i=1}^{c+1} U_i b_i \quad (A1)$$

where y is an $n \times 1$ vector of observations
 X is an $n \times k$ matrix of known coefficients
 γ is a $k \times 1$ vector of unknown constants
 U_i is an $n \times m_i$ matrix of 0,1 coefficients
 b_i is an $m_i \times 1$ vector of normal variables from $N(0, \sigma^2)$.

Specifically, $U_{c+1} = I_n$ and b_{c+1} is an n -vector of "error variables." Moreover, the design matrices U_i have precisely one value of 1 in each of their rows and all other coefficients 0. The total number of random levels is denoted by

$$m = \sum_{i=1}^c m_i$$

We may assume without loss of generality that

$$X'X = I \quad (A2)$$

for, if equation (A2) is not satisfied, we may orthogonalize X by a Gram Schmidt orthogonalization process with a consequential reparameterization

*The material in appendix A was presented at the regional meeting of the Biometric Society, Chapel Hill, North Carolina, April 1977, and a version published in *Biometrics*, vol. 34, no. 2, June 1978, pp. 233-242.

^aTexas A & M University, College Station, Texas.

^bCarleton University, Ottawa, Ontario.

^cUniversity of Houston, Houston, Texas.

of \mathbf{y} , omitting any linearly dependent columns in the Gram Schmidt process. Usually, the first column of X is the column vector with all elements equal to $1/\sqrt{n}$. It is the objective of the method to compute estimates of the variance components σ_j^2 and the vector $\boldsymbol{\gamma}$.

3. THE PRESENT METHOD

The essence of the present method is to

(a) Select $c+1$ quadratic forms $\mathbf{Q}_j(\mathbf{y})$ in the elements of \mathbf{y}

(b) Use the method of synthesis (refs. 8 and 9) to obtain the coefficients k_{ji} in the formulas for $E(\mathbf{Q}_j)$ in the form

$$E(\mathbf{Q}_j) = \sum_{i=1}^{c+1} k_{ji} \sigma_i^2 \quad (\text{A3})$$

(c) Estimate σ_j^2 by equating the computed \mathbf{Q}_j to their expectations; i.e., by inverting the system (eq. (A3)) to compute the vector $\hat{\boldsymbol{\sigma}}^2$ with elements $\hat{\sigma}_j^2$

$$\hat{\boldsymbol{\sigma}}^2 = K^{-1} \mathbf{Q}(\mathbf{y}) \quad (\text{A4})$$

from the vector $\mathbf{Q}(\mathbf{y})$ with elements $\mathbf{Q}_j(\mathbf{y})$, where $K = (k_{ji})$ with rank to be discussed in section 6

(d) Replace any negative elements of $\hat{\boldsymbol{\sigma}}^2$ by 0

We now give more details for (a), (b), and (c).

(a) The $\mathbf{Q}_j(\mathbf{y})$ will be based on contrasts which do not depend on any elements of $\boldsymbol{\gamma}$. Accordingly, we orthogonalize all U_i matrices on X and construct matrices V_i orthogonal on X as follows. Denote by $\mathbf{u}(t,i)$ the t th column vector of U_i and by $\mathbf{x}(r)$ the r th column vector of X ; then, the columns $\mathbf{v}(t,i)$ of V_i are given by

$$\mathbf{v}(t,i) = \mathbf{u}(t,i) - \sum_{r=1}^k \mathbf{x}(r) \{ \mathbf{x}'(r) \mathbf{u}(t,i) \} \quad (\text{A5a})$$

or

$$V_i = U_i - X X' U_i \quad (\text{A5b})$$

We now choose the $c+1$ quadratic forms $\mathbf{Q}_j(\mathbf{y})$ as

$$\begin{aligned} \mathbf{Q}_j(\mathbf{y}) &= \mathbf{y}' V_j V_j' \mathbf{y} \\ &= (\mathbf{V}_j' \mathbf{y})' \mathbf{V}_j' \mathbf{y} \quad j = 1, \dots, c+1 \quad (\text{A6}) \end{aligned}$$

(b) It follows from the method of synthesis (see refs. 8 and 9) that

$$E \mathbf{Q}_j(\mathbf{y}) = \sum_{i=1}^{c+1} k_{ji} \sigma_i^2 \quad (\text{A7a})$$

with

$$k_{ji} = \sum_t (\mathbf{V}_j' \mathbf{u}(t,i))' (\mathbf{V}_i' \mathbf{u}(t,i)) \quad (\text{A7b})$$

Now, since $\mathbf{v}(\tau,j)$ is orthogonal on any $\mathbf{x}(\rho)$ (i.e., since $\mathbf{v}'(\tau,j) \mathbf{x}(\rho) = 0$), we can write the k_{ji} in the alternative form

$$\begin{aligned} k_{ji} &= \sum_t (\mathbf{V}_j' \mathbf{v}(t,i))' (\mathbf{V}_i' \mathbf{v}(t,i)) \\ &= \sum_t \left| \sum_{\tau} \{ \mathbf{v}'(\tau,j) \mathbf{v}(t,i) \} \right|^2 \quad (\text{A8}) \end{aligned}$$

thereby showing that $k_{ij} = k_{ji}$. An alternative form of k_{ji} is

$$k_{ji} = \text{tr} \left\{ (\mathbf{V}_j \mathbf{V}_j') (\mathbf{V}_i \mathbf{V}_i') \right\} \quad (\text{A9})$$

We shall show in section 6 that the symmetrical matrix $K = (k_{ji})$ will have full rank $c+1$ if the $n \times n$ matrices $\mathbf{V}_i \mathbf{V}_i'$ are not linearly dependent.

(c) We shall also show in section 6 that the system of equations

$$\mathbf{Q} = K \hat{\boldsymbol{\sigma}}^2 \quad (\text{A10})$$

is consistent even if the rank of K is degenerate. Solving equation (A10) in the form

$$\delta^2 = K^{-1}Q \quad (\text{A11})$$

we shall, of course, be particularly interested in the full rank case when $K^{-1} = K^{-1}$.

4. THE COMPUTATIONAL LOAD

It may be helpful to give an idea of the computational efficiency of the present method by tabulating the number of products involved in the main operations of the algorithm. To this end, we first note simplified versions for the $k_{c+1,i}$. Observing that $U_{c+1} = I$, we have from equation (A5) that $V_{c+1} = I - XX'$, and, since $X'X = I$, we find that $V_{c+1}V'_{c+1} = I - XX'$ and finally from equation (A9) that

$$\begin{aligned} k_{c+1,c+1} &= \text{tr}(I - XX')(I - XX') \\ &= \text{tr}(I - XX') \\ &= n - k \end{aligned} \quad (\text{A12})$$

Similarly, we find that

$$\begin{aligned} k_{c+1,i} &= \text{tr}\{(I - XX')(V_i V'_i)\} \\ &= \text{tr}\{V_i V'_i - XX' V_i V'_i\} \\ &= \text{tr} V_i V'_i \end{aligned} \quad (\text{A13})$$

Further, we note the form of $V'_{c+1}y$: i.e.,

$$V'_{c+1}y = y - XX'y \quad (\text{A14})$$

Defining now the adjointed matrices

$$\left. \begin{aligned} U &= (U_1 | \dots | U_c) \\ V &= (V_1 | \dots | V_c) \end{aligned} \right\} \quad (\text{A15})$$

the bulk of the work consists of the formation of the elements of the symmetrical matrix $V'V = V'U = U'V$. The elements of this matrix are assembled in submatrices in accordance with the partition (eq. (A15)) as shown in schedule 1, where it must be remembered that the range of the column index i depends on i and is $i = 1, \dots, m_i$, and the range of τ is $\tau = 1, \dots, m_i$, so that the submatrix $V'_j U_i$ has dimensions $m_j \times m_i$. The k_{ij} for $i \geq j = 1, \dots, c$, are then obtained by forming the sums of squares of the elements in each submatrix in accordance with equation (A7).

Finally, we recite the formulas for the remaining coefficients in the equations (A10). The $k_{c+1,c+1}$ and $k_{c+1,i}$ are computed from equations (A12) and (A13), respectively, and the right-hand sides of $Q_j(y)$ from the second form in equation (A6) for $j = 1, \dots, c$, whereas $Q_{c+1}(y)$ is given in accordance with equation (A14) by

$$Q_{c+1}(y) = y'y - (X'y)'(X'y) \quad (\text{A16})$$

Schedule 1: Submatrices of $V'U$

	U_1	U_2	...	U_c
V_1	$v(\tau,1)'u(t,1)$	$v(\tau,1)'u(t,2)$...	$v_1(\tau,1)'u(t,c)$
V_2		$v(\tau,2)'u(t,2)$...	$v(\tau,2)'u(t,c)$
...			
V_c				$v(\tau,c)'u(t,c)$

We can now summarize the approximate number of products involved in the various operations of the algorithms.

Operation	Approximate no. of products involved
Orthogonalization of $X'X$	$(1/2)k^+(k^+ - 1)n$ where k^+ = no. of columns in original X
$XU_i, i = 1, \dots, c$	kmn
$X(X'U_i), i = 1, \dots, c$ (eq. (A5))	nmk
$U'V = V'V$ (schedule 1)	0 Subtotals of elements of $v(i,i)$
$k_{ij}, i, j = 1, \dots, c$ (eq. (A7))	$(1/2)m(m+1)$
$k_{c+1,i}, i = 1, \dots, c$ (eq. (A13))	mn
$k_{c+1,c+1}$ (eq. (A12))	0
$Q_j(y), j = 1, \dots, c+1$ (eqs. (A6), second form, and (A16))	$(m+k+1)(n+1)$

The important point is that the number of products is only a linear function of the number of data lines n . An approximate formula for the total number of products is

$$n \left\{ \frac{1}{2} k^+ (k^+ - 1) + (2m + 1)(k + 1) \right\}$$

5. A NUMERICAL EXAMPLE

A small numerical example with $n = 4$, $k^+ = 3$, $k = 2$, $c = 1$, $m_1 = 2$, $m = 2$, $m_2 = n = 4$ is shown in schedule 2.

Schedule 2: A Numerical Example of a Mixed Model

y	$X(\text{orig.})$	U_1	U_2	$X(\text{new})$	V_1
4	1 1 0 1 0 1 0 0 0	0 0 0	0 0 0	(1/2) (1/2) (1/2)	-(1/2)
2	1 1 0 0 1 0 1 0 0	(1/2) (1/2)	-(1/2) (1/2)		
1	1 0 1 0 1 0 0 1 0	(1/2)	-(1/2)	0	0
2	1 0 1 0 1 0 0 0 1	(1/2)	-(1/2)	0	0

The orthogonalization of X (original) to X (new) follows the standard Gram Schmidt procedure and reduces the $k^+ = 3$ dependent columns to $k = 2$ columns which are orthogonal and standardized. Note that

$$x(2)_{\text{new}} = x(2)_{\text{old}} - (1/2)x(1)_{\text{old}}$$

and

$$x(3)_{\text{old}} = x(1)_{\text{new}} - x(2)_{\text{new}}$$

must be eliminated. Using now $x(r) = x(r)_{\text{new}}$, we orthogonalize U_1 on X and compute (see eq. (A5))

$$x'(1)u(1,1) = (1/2)$$

$$x'(2)u(1,1) = (1/2)$$

and hence

$$v(1,1) = u(1,1) - (1/2)x(1) - (1/2)x(2)$$

Likewise,

$$x'(1)u(2,1) = (3/2)$$

$$x'(2)u(2,1) = -(1/2)$$

and hence

$$v(1,2) = u(2,1) - (3/2)u(1) + (1/2)u(2)$$

This yields the matrix V_1 in schedule 2 which has only one independent column. The elements of $V_1'U_1$ require the computation of

$$v(1,1)'u(1,1) = (1/2)$$

$$\begin{aligned} v(1,1)'u(2,1) &= v(2,1)'u(1,1) \\ &= -(1/2) \end{aligned}$$

and $v(2,1)'u(2,1) = 1/2$ with sum of squares of $k_{11} = 4(1/2)^2 = 1$. Further (eq. (A12)), $k_{22} = 4 - 2 = 2$ and (eq. (A13)) $k_{12} = k_{21} = 4(1/2)^2 + 4(0)^2 = 1$ so that the K matrix is given by

$$K = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

Finally (eq. (A16)),

$$\begin{aligned} Q_2(y) &= 4^2 + 2^2 + 1^2 + 2^2 - \left(\frac{1}{2}9\right)^2 - \left(\frac{1}{2}3\right)^2 \\ &= 25 - \frac{90}{4} \\ &= 25 - 22.5 \\ &= 2.5 \end{aligned}$$

and (eq. (A6))

$$\begin{aligned} Q_1(y) &= \left(\frac{1}{2}2\right)^2 + \left(\frac{1}{2}(-2)\right)^2 \\ &= 2 \end{aligned}$$

The solution of $Q = K\hat{\theta}^2$ therefore yields $\hat{\theta}_2^2 = 1/2$, $\hat{\theta}_1^2 = 1.5$.

6. OPTIMALITY PROPERTIES AND THE CONSISTENCY OF THE EQUATIONS

The estimators described in section 3 may be seen to be "best at $\sigma_i^2 = 0, i = 1, \dots, c, \sigma_{c+1}^2 = 1$ " as defined by L. R. LaMotte (ref. 10). Therefore, the consistency of equation (A10), regardless of the rank of K , is established as Lemma 4 by LaMotte. That the estimators defined by equation (A11) are "best" among invariant quadratic unbiased estimators guarantees that they are admissible in that class; that is, no other invariant quadratic unbiased estimators have uniformly less variance for all $\hat{\theta}$. Further, as noted by LaMotte, the estimators (eq. (A11)) have the property that in any model for which a uniformly best estimator exists, equation (A11) will be uniformly best. Finally, it may be seen that the "synthesis" estimators (eq. (A11)) are also MINQUE as in Rao (ref. 4, section 6) with $V = I$. No claim is made that this choice of the norm has any particular merits among the rather general family of the norms covered by MINQUE formulas. However, it appears reasonable to us that, in the absence of any theoretical criteria for selection of MINQUE norms, a norm leading to simple estimators may be regarded as meritorious.

Following section A5 in LaMotte (ref. 10), it may be seen that the rank of K is equal to the number of linearly independent matrices among $V_i V_i', i = 1, \dots, c + 1$. Thus, a singular K may occur if the $U_i U_i'$ matrices are not all linearly independent or if there exists (see eq. (A5)) a linear combination of the $U_i U_i'$ matrices whose columns are contained in the linear subspace spanned by the columns of X . In the first case, the singularity is caused by the design's leading to the U_i matrices; in the second, the singularity is caused by confounding fixed and random effects. In either case, equation (A10) is consistent but some linear combinations of the variance components cannot then be unbiasedly estimated. We should stress, however, that other special cases of MINQUE (not necessarily invariant to γ) may also deserve particular attention.

The consistent estimator $\hat{\theta}^2$ may serve as a starting value for the iterative maximum likelihood

estimation procedure described by Hemmerle and Hartley (ref. 3). Under certain regularity conditions (not discussed here), one single cycle of the iteration will result in asymptotically efficient estimators of σ^2 and γ . If the iteration is carried to convergence, solutions of the ML equations are reached. If no ML cycles are performed, a consistent estimator $\tilde{\gamma}$ of γ can be computed from the generalized least squares (ML) equations.

$$\tilde{\gamma} = (X'H^{-1}X)^{-1}(X'H^{-1}y) \quad (A17)$$

where

$$H = I_n + \sum_{i=1}^c \frac{\tilde{\sigma}_i^2}{\tilde{\sigma}_{c+1}^2} U_i U_i'$$

It has been shown by Hemmerle and Hartley (ref. 3) that equation (A17) can be computed directly from the $U_i U_i'$ and $X' U_i$ matrices without the inversion of the $n \times n$ matrix H using their so-called W-transformation. In fact, the W_0 matrix (their eq. (19)) is essentially given by the $V_i' V_i$ matrices (see schedule 1) and by the contrasts $V_i' y$ required in the computation of $Q_i(y)$. The variance-covariance matrix of $\tilde{\gamma}$ can likewise be computed through the W-transformation.

REFERENCES

1. Searle, S. R.: *Linear Models*. John Wiley & Sons (New York), 1971.
2. Hartley, H. O.; and Rao, J. N. K.: Maximum Likelihood Estimation for the Mixed Analysis of Variance Model. *Biometrika*, vol. 54, 1967, pp. 93-108.
3. Hemmerle, W. J.; and Hartley, H. O.: Computing Maximum Likelihood Estimates for the Mixed AOV Model Using the W-Transformation. *Technometrics*, vol. 15, 1973, pp. 819-831.
4. Rao, C. R.: Estimation of Variance and Covariance Components—MINQUE Theory. *Journal of Multivariate Analysis*, vol. 1, 1971, pp. 257-275.
5. Liu, L.; and Senturia, J.: Computation of MINQUE Variance Components Estimates. Technical Report 408, Department of Statistics, University of Wisconsin, 1976.
6. Henderson, C. R.: Estimation of Variance and Covariance Components. *Biometrics*, vol. 9, 1953, pp. 226-252.
7. Gaylor, D. W.; Lucas, H. L.; and Anderson, R. L.: Calculation of the Expected Mean Squares by the Abbreviated Doolittle and Square Root Methods. *Biometrics*, vol. 26, 1970, pp. 641-656.
8. Hartley, H. O.: Expectations Variances and Covariances of ANOVA Mean Squares by "Synthesis." *Biometrics*, vol. 23, 1967, pp. 105-114.
9. Rao, J. N. K.: On Expectations Variances and Covariances of ANOVA Mean Squares by "Synthesis." *Biometrics*, vol. 24, 1968, pp. 963-978.
10. LaMotte, L. R.: Quadratic Estimation of Variance Components. *Biometrics*, vol. 29, 1973, pp. 311-330.

Appendix B

The Approximate Variance of a Weighted Least Squares Estimator With "Empirical Weights"

In this appendix, we prove a general theorem concerning weighted least squares sometimes referred to as the Aitken method. We prove the theorem in a somewhat more general form than that required for this paper. Accordingly, the notation for the variance-covariance matrix H is altered as shown.

Consider the linear model

$$y = X\gamma + e \quad (B1)$$

where y is an n -vector of observations, X is an $n \times p$ matrix of constants, γ is a p -vector of unknown parameters, and the residual vector e has an $n \times n$ variance-covariance matrix of the form

$$Eee' = H = \sum_{i=1}^k \theta_i C_i \quad (B2)$$

where the C_i are known matrices and the θ_i are unknown parameters which have to be estimated from the data vector y .

Consider now the weighted least squares estimator

$$\hat{\gamma} = (X' \hat{A}^{-1} X)^{-1} (X' \hat{A}^{-1} y) \quad (B3)$$

where

$$\hat{A} = \sum \hat{\theta}_i C_i \quad (B4)$$

and the $\hat{\theta}_i$ are consistent estimators (computed from y) and are assumed to be approximately unbiased.

Using the so-called Δ -method to obtain the variance of $\hat{\gamma}$, we write

$$E(\hat{\gamma} - \gamma)(\hat{\gamma} - \gamma)' \approx \left[\frac{\partial \gamma}{\partial y} \right]' H \left[\frac{\partial \gamma}{\partial y} \right] + 2 \left[\frac{\partial \gamma}{\partial \theta} \right]' C \left[\frac{\partial \gamma}{\partial \theta} \right] + \left[\frac{\partial \gamma}{\partial \theta} \right]' A \left[\frac{\partial \gamma}{\partial \theta} \right] \quad (B5)$$

where $[\partial \gamma / \partial y]'$ and $[\partial \gamma / \partial \theta]'$ are, respectively, the $p \times n$ and $p \times k$ matrices of partial derivatives of $\hat{\gamma}$ with regard to the elements of y and θ , evaluated at the expected values $Ey = X\gamma$ and $E\theta = \theta$; H and A are, respectively, the variance-covariance matrices of y and of θ ; and C is the covariance matrix of y and θ . We now show that $[\partial \gamma / \partial \theta]' = 0$ so that only the first term in equation (B5) needs to be retained for the Δ -method. In order to obtain the derivatives of $\hat{\gamma}$ given by equation (B3), we note that, to first order,

$$A + \Delta^{-1} \approx A^{-1} - A^{-1} \Delta A^{-1}$$

Therefore, writing $\delta \theta_i = \hat{\theta}_i - \theta_i$ and $H_0 = H(\theta_0)$, we have to first order

$$H^{-1} = H_0^{-1} - \sum \delta \theta_i H_0^{-1} C_i H_0^{-1} \quad (B6)$$

and hence

$$\hat{\gamma} \approx \left[(X' H_0 X)^{-1} + (X' H_0 X)^{-1} X' H_0^{-1} \cdot \left(\sum_i \delta \theta_i C_i \right) H_0^{-1} X (X' H_0 X)^{-1} \right] \cdot \left[X' H_0^{-1} y - X' H_0^{-1} \left(\sum_i \delta \theta_i C_i \right) H_0^{-1} y \right] \quad (B7)$$

If we therefore find the coefficient of $\delta\theta$, at $y = X\gamma$. Further we obtain

$$(X'H_0X)^{-1}(X'H_0^{-1}C'H_0^{-1}X)(X'H_0^{-1}X)^{-1}$$

$$\left[\frac{\partial \gamma}{\partial y}\right]' = (X'H_0^{-1}X)^{-1}X'H_0^{-1} \quad (B9)$$

$$X'H_0^{-1}X\gamma - (X'H_0X)^{-1}(X'H_0^{-1}C'H_0^{-1}X\gamma) = 0 \quad \text{so that} \quad (B8)$$

$$E(\hat{\gamma} - \gamma)(\hat{\gamma} - \gamma)' = (X'H_0^{-1}X)^{-1} \quad (B10)$$

Weighted Aggregation

A. H. Felveson^a

INTRODUCTION

In a sample survey such as LACIE, where one is plagued with measurement errors and loss of data, it might be possible to improve the precision of the overall estimate if observations thought to be of questionable accuracy are downweighted in the aggregation process.

In the LACIE aggregation logic, a "Group III" wheat area ratio estimate is currently made for a stratum, a substratum, or a collection of substrata which, either by design or by loss of data, does not contain a sample segment. This ratio estimate is made by taking the estimated wheat area from surrounding or nearby strata or substrata with "good" data and multiplying it by the historical ratio of the Group III area's wheat to the neighboring area's wheat. If the ratios of crop averages between neighboring political subdivisions do not change radically from year to year, the Group III estimate should be a reasonably accurate means of accounting for missing data.

A flaw in the LACIE procedure is that, if data from even one segment are available, that segment is used to estimate the wheat acreage in its stratum or substratum, no matter how many segments were originally allocated in the sampling design. Furthermore, the wheat proportion estimates from some segments might be of questionable accuracy because of not having acquired the data at the right times in the growing season to discriminate wheat from other crops. As a result, wheat area estimates for strata containing insufficient data or poorly estimated segments can be seriously distorted, thus affecting the large-area country production estimate.

If one knows that a stratum wheat area estimate is likely to be poor because of the preceding factors,

then it is advantageous to replace the suspected estimate by a weighted average of itself with the aforementioned Group III estimate of its wheat area. The size of the weights depends on the degree of confidence one has in the direct segment-based estimates; for example, the estimate for a stratum/substratum containing segments with data acquired on only one Landsat pass would be more heavily weighted with the Group III component than with the direct component. Conversely, a stratum which was thought to have reasonably good acquisition patterns for its segments would be estimated by giving most of the weight to the direct component.

The manner in which a weighted aggregation technique can be implemented given a set of weights is described in the following section. The problem of variance estimation is discussed in the section entitled "Properties of the Estimate \hat{A} ," and the question of how one might obtain the weights in an operational environment is addressed in the section entitled "Determination of Weights."

THE WEIGHTED AGGREGATION PROCESS

Definitions and Notation

In the current LACIE Crop Assessment System (CAS), one has on hand sample-segment-based estimates of wheat acreage for strata containing at least one sample segment with usable data and historical wheat acreages for all strata in a country. In countries with detailed historical data, there are also estimates available for Group I substrata or Group II collections of substrata containing at least one usable sample segment as well as historical wheat acreages for all substrata. To avoid further confusion, the term "domain" will be used to mean a stratum in countries without detailed historical data and to mean either a Group I or III substratum or a Group II collection of

^aNASA Johnson Space Center, Houston, Texas.

substrata in countries with detailed historical data.¹ The following definitions can now be made.

Let d_i^* be the "direct" or segment-based estimate of wheat acreage for the i th domain, ($i = 1, \dots, n$), provided that domain has at least one processed sample segment. For all domains, define

$$d_i = \begin{cases} d_i^*, & \text{data exists for } i\text{th domain} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

In addition, let h_i be the historical wheat acreage and w_i a given weight ($0 \leq w_i \leq 1$) associated with the i th domain. It is assumed in this section that the w_i are known or computable; the problem of obtaining them will be discussed in the section entitled "Determination of Weights."

Finally, for each domain, there is a prescribed set of other domains on which to ratio for the purpose of computing a Group III estimate of wheat acreage. Specifically, for the i th domain D_i , there exists a set of other domains S_i such that z_i , the Group III wheat acreage estimate for D_i , is given by

$$z_i = \left(\frac{\sum_{D_j \in S_i} d_j}{\sum_{D_j \in S_i} h_j} \right) h_i \quad (2)$$

Let the $n \times 1$ vector $b_i = (b_{i1}, \dots, b_{in})^T$ be defined by

$$b_{ij} = \begin{cases} 1, & D_j \in S_i \\ 0, & \text{otherwise} \end{cases}$$

¹For definitions of Groups I, II, and III and designations of LACIE countries, see the paper by Feiveson et al. entitled "LACIE Sampling Design." In the United States, a domain is either a county or a collection of counties within a Crop Reporting District (CRD).

Then, equation (2) can be written

$$z_i = \left(\frac{b_i^T d}{b_i^T h} \right) h_i \quad (3)$$

where $d = (d_1, \dots, d_n)^T$ and $h = (h_1, \dots, h_n)^T$.

Estimation

We now construct a set of wheat acreage estimates $\{\hat{d}_i\}_{v=1}^{\infty}$, where \hat{d}_i will be shown to be the limit of a convergent sequence of iterated estimates $\{a_i^{(v)}\}_{v=0}^{\infty}$. At any stage of the iteration, say the $v+1$ st, $a_i^{(v+1)}$ will be a weighted average between the direct estimate d_i and a "historical" estimate $z_i^{(v)}$, which is defined in equation (5).

To start the iteration process, let $a_i^{(0)} = d_i$, $i = 1, 2, \dots, n$. Then, the $v+1$ st iterated estimate is defined by

$$a_i^{(v+1)} = w_i d_i + (1 - w_i) z_i^{(v)} \quad (4)$$

where

$$z_i^{(v)} = h_i \sum_j b_{ij} a_j^{(v)} / b_i^T h \quad (5)$$

Note that $z_i^{(0)} = z_i$ and hence $a_i^{(1)}$ is a simple weighted average between d_i and z_i . At later iterations, $z_i^{(v)}$ is similar to z_i but uses $a_j^{(v)}$ in place of d_j in equation (3).

By letting the $n \times 1$ vector $a^{(v)} = (a_1^{(v)}, \dots, a_n^{(v)})^T$, and the $n \times n$ matrix

$$\begin{aligned} C &= (c_{ij}) \\ &= w_{ij} (1 - w_i) h_j / b_i^T h \end{aligned}$$

then, equation (4) can be written

$$\mathbf{a}^{(\nu+1)} = \mathbf{u} + \mathbf{C}\mathbf{a}^{(\nu)} \quad (6)$$

where $\mathbf{u} = (u_1, \dots, u_n)^T$ and $u_i = w_i d_i$.

From equation (6), it follows that

$$\mathbf{a}^{(\nu+1)} = (\mathbf{I} + \mathbf{C} + \mathbf{C}^2 + \dots + \mathbf{C}^\nu)\mathbf{u} + \mathbf{C}^{\nu+1}\mathbf{a}^{(0)}$$

hence, $\mathbf{a}^{(\nu)}$ converges to a vector $\hat{\mathbf{a}}$ if and only if the series $\mathbf{I} + \mathbf{C} + \mathbf{C}^2 + \dots$ converges. We now quote two theorems which will show that $\mathbf{a}^{(\nu)}$ converges to $\hat{\mathbf{a}}$ for "reasonable" values of the matrix $\mathbf{B} = (b_{ij})$, where

$$\hat{\mathbf{a}} = (\mathbf{I} - \mathbf{C})^{-1}\mathbf{u} \quad (7)$$

In the following theorems, the function $\rho(\mathbf{M})$, for any matrix \mathbf{M} , is defined by

$$\rho(\mathbf{M}) = \max_i |\lambda_i(\mathbf{M})|$$

where $\lambda_i(\mathbf{M})$ is the i th eigenvalue of \mathbf{M} .

Theorem 1

If \mathbf{M} is an arbitrary complex $n \times n$ matrix with $\rho(\mathbf{M}) < 1$, then $\mathbf{I} - \mathbf{M}$ is nonsingular, and

$$(\mathbf{I} - \mathbf{M})^{-1} = \mathbf{I} + \mathbf{M} + \mathbf{M}^2 + \dots$$

the series on the right converging. Conversely, if the series on the right converges, then $\rho(\mathbf{M}) < 1$.

Theorem 2

If $\mathbf{A} = (a_{ij}) \leq 0$ is an $n \times n$ matrix, and \mathbf{x} is any vector with positive components x_1, x_2, \dots, x_n ,

$$\min_{1 < i < n} \left[\frac{\sum_{j=1}^n a_{ij} x_j}{x_i} \right] < \rho(\mathbf{A}) < \max_{1 < i < n} \left[\frac{\sum_{j=1}^n a_{ij} x_j}{x_i} \right]$$

These theorems may be found on pages 82 and 47, respectively, of reference 1.

In order to show equation (7), we will first show that $\rho(\mathbf{C}) < 1$ and then use Theorem 1. We can attempt to show that $\rho(\mathbf{C}) < 1$ by using Theorem 2 with $x_j = h_j$. Then, for fixed i ,

$$\begin{aligned} \sum_j c_{ij} x_j / x_i &= (1 - w_i) h_i \sum_j b_{ij} h_j / h_i \sum_k b_{ik} h_k \\ &= 1 - w_i \end{aligned}$$

Thus, if $w_i > 0$ for all i , by Theorem 2, we have $\rho(\mathbf{C}) < 1$. In practice, however, some w_i will be zero; hence, some restrictions must be put on the \mathbf{B} matrix to ensure $\rho(\mathbf{C}) < 1$. A natural restriction to require is that $b_{ij} = 0$ if $w_j = 0$; i.e., if the direct estimate in the j th domain is nonexistent or thought to be completely worthless ($w_j = 0$), then that domain should not be used in a ratio estimate for some other domain. With this restriction, the \mathbf{C} matrix now takes on the form

$$\mathbf{C} = \begin{bmatrix} 0 & \mathbf{C}_1 \\ r \times r & r \times (n-r) \\ 0 & \mathbf{C}_2 \\ (n-r) \times r, (n-r) \times (n-r) \end{bmatrix} \quad (8)$$

where the order of the domains has been rearranged so that $w_1 = w_2 = \dots = w_r = 0$ and $w_j > 0$ for $j = r+1, \dots, n$.

Since the nonzero eigenvalues of C_2 are also the nonzero eigenvalues of C , it is clear that $\rho(C) = \rho(C_2)$. By applying Theorem 2 to C_2 as in the previous paragraph, we now have $\rho(C) = \rho(C_2) < 1$.

By Theorem 1, the series $I + C + C^2 + \dots$ converges to $(I - C)^{-1}$; hence, $C^v \rightarrow 0$ and

$$\lim_{v \rightarrow \infty} a^{(v)} = (I - C)^{-1}u = \hat{a}$$

PROPERTIES OF THE ESTIMATE \hat{a}

Bias

Suppose α_i is the true wheat acreage for the i th domain. If $E(d_i) = \alpha_i$ for $w_i \neq 0$ and the ratio of the wheat acreage in domain i to that of its ratio estimation set S_i is the same for the current year as it was in the year which produced h , then \hat{a} is unbiased.

To show unbiasedness under the preceding conditions, an induction argument will be used. First, from equation (4),

$$E(a_i^{(1)}) = w_i \alpha_i + (1 - w_i) E(z_i)$$

From the second assumption,

$$\alpha_i / \sum_j b_{ij} \alpha_j = h_i / \sum_j b_{ij} h_j$$

As a consequence, remembering that $w_i = 0$ implies that $b_{ij} = 0$, we have

$$\begin{aligned} E(z_i) &= h_i \sum_j E(b_{ij} d_j) / \sum_k b_{ik} h_k \\ &= h_i \sum_j b_{ij} \alpha_j / \sum_k b_{ik} h_k \\ &= \alpha_i \end{aligned}$$

It thus follows that

$$\begin{aligned} E(a_i^{(1)}) &= w_i \alpha_i + (1 - w_i) \alpha_i \\ &= \alpha_i \end{aligned} \tag{9}$$

Suppose now that $E(a_i^{(v)}) = \alpha_i$. Then similarly,

$$\begin{aligned} E(a_i^{(v+1)}) &= E(w_i d_i) + (1 - w_i) E(z_i^{(v)}) \\ &= w_i \alpha_i + (1 - w_i) h_i \sum_j E(b_{ij} a_j^{(v)}) / \sum_k b_{ik} h_k \\ &= w_i \alpha_i + (1 - w_i) h_i \sum_j b_{ij} \alpha_j / \sum_k b_{ik} h_k \\ &= w_i \alpha_i + (1 - w_i) \alpha_i \\ &= \alpha_i \end{aligned}$$

Hence, $a_i^{(v)}$ is unbiased for all v , which implies $E(\hat{a}_i) = \alpha_i$.

Variance

Strictly speaking, the w 's are random variables because they are functions of local weather and crop distribution. Since $\hat{a} = (I - C)^{-1}u$, the conditional covariance matrix of \hat{a} given $w = (w_1, \dots, w_n)^T$ is given by

$$\Sigma = (I - C)^{-1} \Psi_u (I - C^T)^{-1} \tag{10}$$

where Ψ_u is the conditional covariance matrix of u given w . The unconditional variance of \hat{a} depends on how w is computed and will not be discussed here.

Variance of the Large-Area Estimate

The large-area wheat acreage estimate is simply the sum of the estimates over the domains; i.e., $\hat{A} =$

\hat{A} . As a result, the conditional variance of \hat{A} is given by

$$V(\hat{A}) = e^T \Sigma e \quad (11)$$

where $e = (1, 1, \dots, 1)^T$. In the following section, we shall examine how this compares with the variance of the LACIE estimate under certain conditions.

Estimation of $V(\hat{A})$

If an estimate of Ψ_u is available, it can be used in equation (10) to obtain an estimate of Σ and hence $V(\hat{A})$. In LACIE-type operations, the d_i can be assumed independent since they are based on independent data sets.

The unconditional variance, say τ_i^2 , of each d_i can be estimated as is currently done in LACIE (see the paper by Chhikara and Feiveson entitled "LACIE Large-Area Acreage Estimation" for details). If the d_i and w_i are "approximately" independent, the estimates $\hat{\tau}_i^2$ can be used to approximate Ψ_u by

$$\Psi_u = \text{diag}(w_1^2 \tau_1^2, \dots, w_n^2 \tau_n^2)$$

Thus,

$$\hat{V}(\hat{A}) = e^T [(I - C)^{-1} \hat{\Psi}_u (I - C^T)^{-1}] e \quad (12)$$

is the estimated variance of A .

DETERMINATION OF WEIGHTS

The General Case

Other than using the obvious choice of $w_i = 0$ for domains with no data, there is no straightforward procedure for determining w . Ideally, one would like

to choose w such that the mean-squared error of \hat{A} is minimized, but this is impossible without knowledge of the bias. If an estimate of Ψ_u is available, it might be possible to minimize the variance of \hat{A} (eq. (12)) with respect to w ; however, this is a formidable task for arbitrary B . For the case $B = J$, ($= ee^T$) some progress can be made as will be shown.

The Case $B = J$

Suppose one is estimating wheat acreage for a relatively small region in a country (such as a CRD in the United States), where the ratio estimation set is the totality of all domains in the region. Then, the matrix B is all ones; i.e.,

$$B = \begin{pmatrix} 1, 1, \dots, 1 \\ 1, 1, \dots, 1 \\ 1, 1, \dots, 1 \end{pmatrix} = ee^T \quad (13)$$

In LACIE, within a CRD in the United States, this is actually the case, as long as there are at least three usable sample segments in the CRD (see the paper by Chhikara and Feiveson for details).

In this situation, $I - C = (I - ve^T)$, where $v = (v_1, \dots, v_n)^T$ and $v_i = (1 - w_i)h_i/(h)$. As a consequence,

$$(I - C)^{-1} = I + \left(\frac{1}{1 - v^T e} \right) ve^T \quad (14)$$

which gives

$$\hat{a} = u + \left(\frac{1}{1 - v^T e} \right) ve^T u \quad (15)$$

It follows that

$$\begin{aligned}
 \hat{A} &= e^T \hat{u} \\
 &= e^T u \left[1 + \frac{e^T v}{1 - e^T v} \right] \\
 &= \frac{e^T u}{1 - e^T v} \\
 &= \frac{\sum_i w_i d_i}{1 - \frac{1}{h} \sum_i (1 - w_i) h_i} \\
 &= \frac{h \sum_i w_i d_i}{\sum_i w_i h_i} \quad (16)
 \end{aligned}$$

where

$$h = \sum_i^n h_i$$

To distinguish equation (16) from the general case of arbitrary \mathbb{B} , we will use the notation \hat{A}_J for \hat{A} when $\mathbb{B} = \mathbb{J}$.

If the w_i are all zero or one, \hat{A}_J is the standard LACIE estimator. To see this, suppose the domains are arranged in order such that $w_1 = w_2 = \dots = w_k = 1$ and $w_{k+1} = \dots = w_n = 0$. Then

$$\hat{A}_J = \frac{h \sum_{i=1}^k d_i}{\sum_{i=1}^k h_i}$$

$$\begin{aligned}
 &= \left(\sum_{i=1}^k d_i \right) \left(1 + \frac{\sum_{i=k+1}^n h_i}{\sum_{i=1}^k h_i} \right) \\
 &= (1 + R) \sum_{i=1}^k d_i \\
 &= \hat{A}_{LACIE} \quad (17)
 \end{aligned}$$

where R is the Group III ratio as shown in equation (2).

Bias and variance of \hat{A}_J .—In general, \hat{A}_J (including \hat{A}_{LACIE}) is a biased estimator of the total wheat acreage A , even if the d_i are unbiased when $w_i \neq 0$. Let α_i be the true current-year wheat acreage for \mathcal{D}_i . If we define $\beta = (\sum \alpha_i) / (\sum h_i)$ and $e_i = \alpha_i - \beta h_i$, then

$$\begin{aligned}
 \sum_i e_i &= \sum_i \alpha_i - (\sum \alpha_i) (\sum h_i / \sum h_i) \\
 &= 0
 \end{aligned}$$

It then follows that

$$\begin{aligned}
 \hat{A}_J &= \frac{h \sum_i w_i d_i}{\sum_i w_i h_i} \\
 &= \frac{(\sum w_i d_i)^{1/\beta} \sum_i (\alpha_i - e_i)}{\frac{1}{\beta} \sum_i w_i (\alpha_i - e_i)} \\
 &= \frac{(\sum w_i d_i) A}{\sum w_i \alpha_i - \sum w_i e_i} \quad (18)
 \end{aligned}$$

since $A = \sum \alpha_i$. If $E(d_i) = \alpha_i$ and $V(d_i) = \sigma_i^2$, we have

$$E(\hat{A}_j) = A \left(\frac{\sum w_i \alpha_i}{\sum w_i \alpha_i - \sum w_i e_i} \right) \quad (19)$$

and, also,

$$V(\hat{A}_j) = \frac{h^2 \sum w_i^2 \sigma_i^2}{\left[\sum w_i h_i \right]^2} \quad (20)$$

If the e_i are such that $|\sum w_i e_i| \ll \sum w_i \alpha_i$, then $E(\hat{A}_j) \approx A$. Consequently, equation (20) can be minimized with respect to w to obtain an approximately minimum mean-squared error for \hat{A}_j . Using the Schwarz inequality, it can be seen that equation (19) is minimum when $w_i \propto (h_i/\sigma_i^2)$. The corresponding optimal variance is given by

$$V_{opt} = \frac{h^2}{\sum_{i=1}^n \left(\frac{h_i}{\sigma_i} \right)^2} \quad (21)$$

Comparison with LACIE estimator.—Suppose we are estimating wheat in a CRD with n counties and $B = J$. Suppose also that the first k counties have usable data but the remaining $n - k$ counties have lost all their data because of cloud cover. The LACIE estimate of A , the total wheat acreage, is given by equation (17), whereas the "optimal" estimate is given by letting

$$w_i = \begin{cases} \frac{h_i}{\sigma_i^2}, & i \leq k \\ 0, & i > k \end{cases} \quad (22)$$

where $\sigma_i^2 = \text{Var}(d_i)$. Substituting in equation (16), we have

$$\hat{A}_{opt} = \left(\sum_{i=1}^n h_i \right) \frac{\sum_{i=1}^k \frac{h_i d_i}{\sigma_i^2}}{\sum_{i=1}^k \frac{h_i^2}{\sigma_i^2}} \quad (23)$$

Comparing variances, we see that

$$V(\hat{A}_{LACIE}) = \frac{\left(\sum_{i=1}^n h_i \right)^2 \left(\sum_{i=1}^k \sigma_i^2 \right)}{\left(\sum_{i=1}^k h_i \right)^2} \quad (24)$$

The variance of \hat{A}_{opt} is similar to equation (21) except the upper limit in the denominator sum is k instead of n ; i.e.,

$$V(\hat{A}_{opt}) = \frac{\left(\sum_{i=1}^n h_i \right)^2}{\sum_{i=1}^k \frac{h_i^2}{\sigma_i^2}} \quad (25)$$

Taking the ratio of variances, we obtain

$$\rho = \frac{V(\hat{A}_{opt})}{V(\hat{A}_{LACIE})} = \frac{\left(\sum_{i=1}^k h_i \right)^2}{\left(\sum_{i=1}^k \sigma_i^2 \right) \left(\sum_{i=1}^k \frac{h_i^2}{\sigma_i^2} \right)} \quad (26)$$

which is always less than or equal to 1, with equality when $h_i = \sigma_i^2$.

Using official LACIE estimates of within-county variances to obtain the σ_i^2 , ρ was evaluated for two CRD's, the North Central in Montana and the North Central in Kansas. In the Montana CRD, ρ was 0.57, which suggests that weighted aggregation would give a considerably more accurate estimate than the current procedure. In the Kansas CRD, ρ was almost 1; here, the LACIE estimate was quite efficient. More work needs to be done with existing LACIE data to evaluate the weighted technique taking into account changes in the σ_i^2 due to acquisition patterns and more general B matrices. Neither factor was considered in the preliminary calculations described herein.

REFERENCE

1. Varga, Richard S.: Matrix Iterative Analysis. Prentice-Hall, 1962.

ACKNOWLEDGMENTS

The author wishes to acknowledge the valuable contributions of H. O. Hartley, Texas A & M University, who originally suggested the use of an iterative weighted aggregation procedure. David Lamb of Texas A & M University also contributed work on this subject.

Design, Implementation, and Results of LACIE Field Research

M. E. Bauer,^a M. C. McEwen,^b W. A. Mallia,^c and J. C. Harlan^d

INTRODUCTION

Major advancements have been made in recent years in the capability to acquire, process, and interpret remotely sensed multispectral measurements of the energy reflected and emitted from crops, soils, and other Earth surface features. As a result of experiments such as LACIE, the technology is moving rapidly toward operational applications. There is, however, a continuing need for quantitative studies of the multispectral characteristics of crops and soils if further advancements in the technology are to be made. In the past, many such studies were made in the laboratory because of a lack of instrumentation suitable for field studies, but the applicability of laboratory studies is generally limited. The development of sensor systems capable of collecting high-quality spectral measurements under field conditions has made it possible to pursue investigations which would not have been possible a few years ago.

A major effort was initiated in the fall of 1974 by the NASA Johnson Space Center (JSC) with the cooperation of the U.S. Department of Agriculture (USDA) to acquire fully annotated and calibrated multitemporal sets of spectral measurements and supporting agronomic and meteorological data (ref. 1). The Purdue University Laboratory for Applications of Remote Sensing (LARS) was responsible for the technical design and coordination of the experiment, as well as for major portions of the data acquisition, processing, and analysis. Other organizations, particularly the Environmental Research Institute of Michigan (ERIM), Texas A & M Remote

Sensing Center, and Colorado State University, contributed to the experiment planning and data analysis.

Spectral, agronomic, and meteorological measurements were made at LACIE test sites in Kansas and North Dakota for 3 years and in South Dakota for 2 years. The remote-sensing measurements include data acquired by truck-mounted spectrometers, a helicopter-borne spectrometer, an aircraft multispectral scanner (MSS), and the Landsat multispectral scanners. These data are supplemented by an extensive set of agronomic and meteorological data acquired during each mission.

The LACIE field measurements data form one of the most complete and best documented data sets required for agricultural remote-sensing research. Thus, they are well suited to serve as a data base for research to (1) determine quantitatively the relationship of spectral to agronomic characteristics of crops, (2) define future sensor systems, and (3) develop advanced data analysis techniques. The data base is unique in the comprehensiveness of sensors and missions over the same sites throughout the growing season and in the calibration of all multispectral data to a common standard.

Continuing analysis of the field data is providing insight into the spectral properties, spectral identification, and assessment of crops. The analyses include development of predictive relationships between spectral variables and leaf area index, biomass, plant water content, and percent soil cover; determination of the effects of cultural and environmental factors on the spectral reflectance of wheat; investigations of the spectral separability of barley and spring wheat; determination of the early-season Landsat threshold for detection of wheat; and comparisons of Landsat MSS and thematic mapper spectral bands for crop identification and assessment.

The remainder of this paper describes the project

^aPurdue University, West Lafayette, Indiana.

^bNASA Johnson Space Center, Houston, Texas.

^cEnvironmental Research Institute of Michigan, Ann Arbor, Michigan.

^dTexas A & M University, College Station, Texas.

objectives, presents an overview of the experimental approach, describes the data acquisition program, and discusses selected results based on field data. The paper ends with a summary of the key accomplishments and results of the experiment and recommendations for future field research.

OBJECTIVES

The overall objective of the LACIE Field Measurements Project was to acquire, process, and distribute to researchers fully annotated and calibrated multitemporal sets of spectral measurements over the wavelength range of 0.4 to 15 micrometers, along with supporting agronomic and meteorological data. These data would serve as a data base for (1) determining quantitatively the temporal-spectral characteristics of spring and winter wheat, the soil background, and surrounding confusion crops; (2) defining future multispectral sensor systems; and (3) developing advanced data processing and analysis techniques.

Specific objectives are listed below for each of these categories. The objectives emphasize analyses to increase understanding of agricultural scenes; however, the data may also be used to pursue sensor design and data processing objectives.

1. Scene-related objectives

a. Determination of the relation of crop canopy characteristics such as percent soil cover, leaf area index, biomass, and plant water content to multitemporal spectral response

b. Determination of the effects of cultural and environmental variables on the spectral properties and spectral identification of wheat

c. Determination of the spectral discriminability of wheat, small grains, and other crops as a function of growth stage

d. Determination of the presence, severity, and extent of crop stresses such as drought, disease, and winterkill from spectral measurements

e. Determination of the year-to-year variation in the condition and spectral response of wheat and other crops

f. Determination of the relation of grain yield to the multitemporal spectral response of wheat

g. Determination of the effects on spectral response of geometric factors such as Sun angle, view angle, and canopy structure of wheat and other selected crops

h. Determination of the effect of the atmosphere on the measured spectral responses of wheat and other crops

i. Determination of the characteristics and use of thermal measurements for discrimination of wheat from other crops

j. Validation of canopy reflectance models

2. Sensor-system-related objectives

a. Determination of optimum or required multispectral sensor system parameters including spectral bands, signal-to-noise ratio (S/N), noise-equivalent difference in reflectance (NE Δ r), noise-equivalent difference in temperature (NE Δ T),¹ and time and frequency of sensor overpasses

b. Comparison and evaluation of Landsat MSS and thematic mapper wavelength bands for crop identification and assessment

3. Data-processing-system-related objective

a. Development of advanced data processing and analysis techniques that use multitemporal, spatial, spectral, transformed spectral, and ancillary data characteristics

OVERVIEW OF EXPERIMENTAL APPROACH

An overview of the experimental approach is shown in figure 1. At the beginning of the project, the technical issues and specific objectives to be addressed with the field measurements data were defined. This led to the experimental design for data acquisition and processing and to the definition of initial data analysis plans and products.

A multistage approach to data acquisition was taken, including areal, vertical, and temporal staging. Areal sampling was accomplished with test sites in different parts of the U.S. Great Plains. Vertical staging, or collection of data by different sensor systems at different altitudes, ranged from mobile towers to Landsat. Temporally, data were collected at 7- to 21-day intervals to sample all important growth stages and during 3 years to obtain a measure of the year-to-year variation in growing conditions and their influence on spectral response.

Measurements were made at three LACIE test sites during 3 crop years, 1975 to 1977. The sites are

¹NE Δ r and NE Δ T are measures of minimum detectable difference in scene reflectance and temperature.

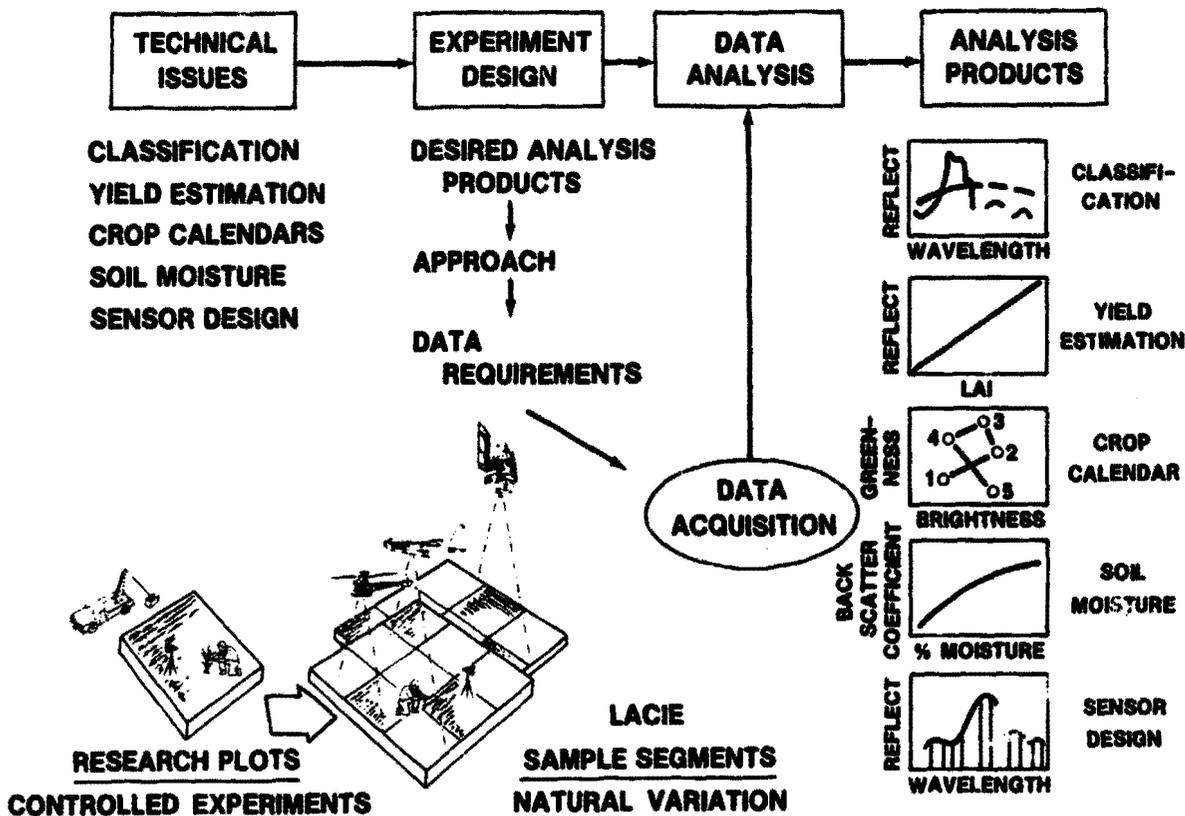


FIGURE 1.—Overview of experimental approach for LACIE field research.

in Finney County, Kansas; Williams County, North Dakota; and Hand County, South Dakota. Finney County and Williams County were chosen to represent winter and spring wheat growing areas, respectively. Hand County is typical of the transitional zone between winter and spring wheat growing areas.

The primary sensors for data collection were truck-mounted spectrometers, a helicopter-borne spectrometer, an aircraft multispectral scanner, and the Landsat-1 and -2 multispectral scanners. Each sensor system has unique capabilities for acquiring spectral data. The spectrometers produce the highest quality reflectance measurements but provide only limited measurements of spatial variability. On the other hand, an aircraft scanner provides spatial sampling of the scene and can obtain data at multiple altitudes, but its spectral coverage, although broader than that of a Landsat MSS, is limited to a fixed set of wavelength bands. The helicopter and aircraft data acquisition systems have the advantage of flexible scheduling and, therefore, provide greater opportunity to obtain cloud-free data at critical crop

growth stages than the Landsat system provides. Landsat provides wide-area coverage but is limited in its spatial resolution and the placement and number of spectral bands.

The staging of data acquisition is summarized in figure 2. Helicopter-spectrometer and aircraft-scanner data were collected in a series of flightlines over commercial fields in the LACIE intensive test site in each of the three counties. Landsat MSS data were acquired and processed for the entire test site, as well as for surrounding areas. These data provide a measure of the natural variation in the temporal-spectral characteristics of wheat and surrounding cover types.

The truck-mounted spectrometers collected spectra of crops in controlled experimental plots at agricultural research stations near the test sites at Garden City, Kansas, and Williston, North Dakota. These data, combined with the more detailed and quantitative measurements of crop and soil conditions which were made at the experiment stations, enable more complete understanding and interpreta-

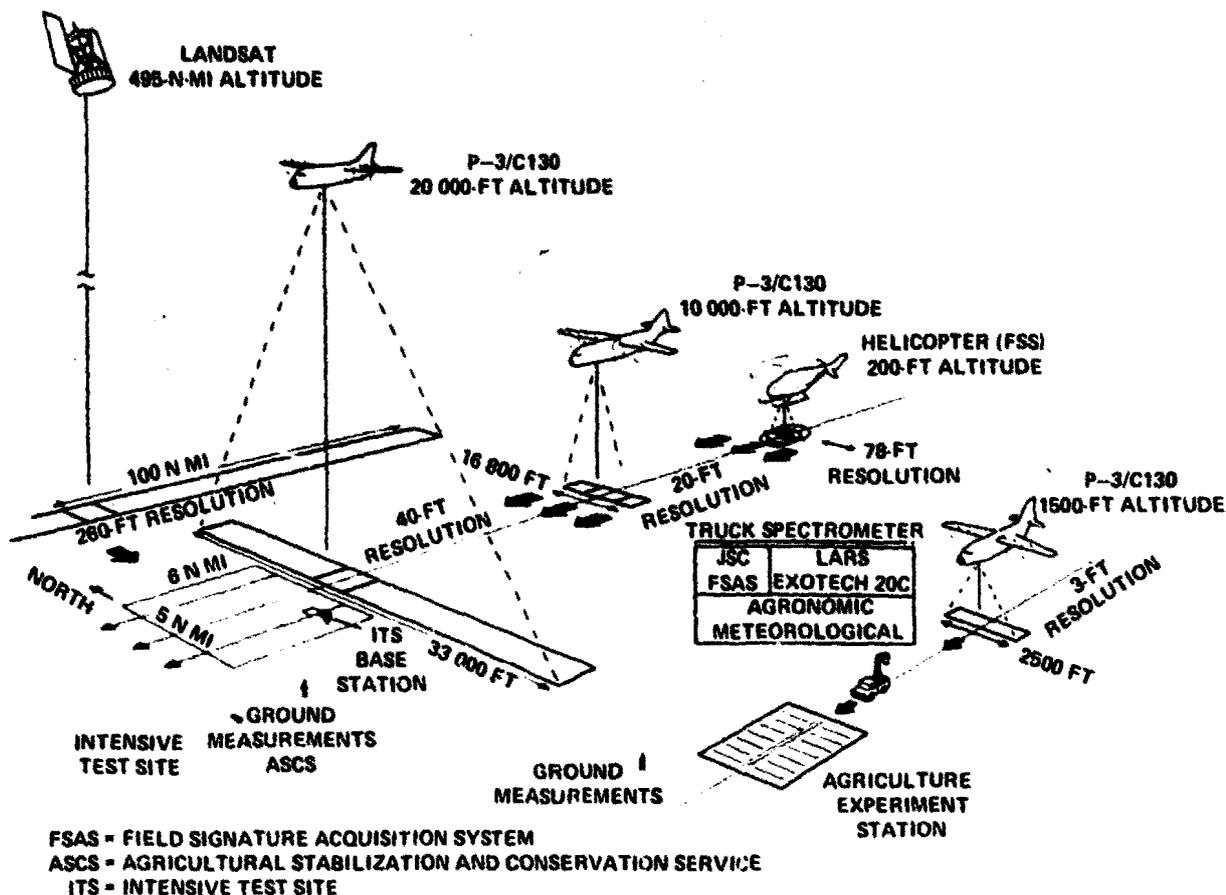


FIGURE 2.—Schematic illustration of LACIE field measurements data acquisition.

tion of the spectra collected from commercial fields. Past experience has shown that there are generally too many interacting variables in commercial fields to determine exact causes of observed differences in spectral response. With data from plots where only two to four factors are varied under controlled conditions, it is possible to determine more exactly and understand more fully the energy-matter interactions occurring in crops.

The spectral measurements were supported by descriptions of the targets and their conditions. The observations, counts, and measurements of the crop canopy include maturity stage, plant height, biomass, leaf area index, percent soil cover, and grain yield. Also included are measurement conditions such as sensor altitude and view angle, as well as measurements of the atmospheric and meteorological conditions. The data are supplemented by aerial photography and ground-level vertical and oblique photo-

graphs of the fields and test plots.

A data library of all spectra, agronomic, meteorological, and photographic data collected is maintained at LARS. The data have been processed in standard data formats and measurement units and made available to JSC-supported investigators and other interested researchers.

DESCRIPTION OF DATA ACQUISITION, PROCESSING, AND DISTRIBUTION

This section describes the acquisition, processing, and distribution of the LACIE field measurements data. It begins by describing the test sites and the experiments at the agriculture experiment stations, followed by descriptions of the sensors and sensor calibration. The procedures for acquiring the

spectral, agronomic, and meteorological data are then described. The section ends with a description of the data processing, library, and analysis systems.

Test Site and Experiment Description

The test sites (fig. 3) were located in Finney County, Kansas; Williams County, North Dakota; and Hand County, South Dakota. Each site consists of a LACIE intensive test site and, in Kansas and North Dakota, an agricultural research station. Measurements were acquired for 3 years at the Kansas and North Dakota sites and for 2 years at the South Dakota site.

The test sites were chosen to include as wide a range of important wheat production areas as possible. In addition, the Finney County and Williams County sites were selected because of their proximity to agricultural research stations. Personnel from the USDA Agricultural Stabilization and Conservation Service (ASCS) were available in each county to collect the required intensive test site ground-truth data.

At the experiment station in Garden City, Kansas, experiments were conducted on dryland and irrigated winter wheat and small grains. At the Williston, North Dakota, experiment station, a small-grains experiment and a cultural practice experiment with spring wheat were conducted.

Intensive test sites.—The intensive test sites are 8.1 by 9.7 kilometers in size. Three flightlines, each 9.7 kilometers long, were located across each site. The

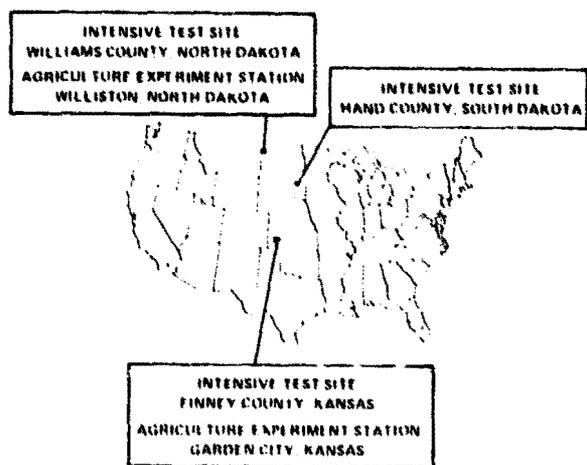


FIGURE 3.—Locations of LACIE field measurements test sites.

number of fields of each major cover type in each site for 1976 is summarized in table 1.

Finney County, Kansas: The test site is located in the High Plains Tableland physiographic area at latitude 38°10' N and longitude 100°43' W. The elevation of the site is 900 meters. The site is overlaid by 3 to 10 meters of loess from the early Wisconsin age.

The soils of the test site are in the Mollisol order, Ustoll suborder, and Argiustolls great group. Mollisols are soils that have nearly black, friable, organic-rich surface horizons high in bases. Ustolls are formed in semiarid regions; they are dry for long periods and have subsurface accumulations of carbonates. The major soil series in the area are Richfield and Ulysses, which are deep, fertile, well-drained, nearly level to gently sloping loamy soils of the upland that are well suited to cultivation.

The area has a distinct continental type of climate characterized by abundant sunshine and constant wind. Most of the precipitation falls during the early part of the year, with a rapid decline in the probability of receiving adequate rainfall during July and August. Thus, the growth cycle of winter wheat is well matched to the available moisture supply. Average annual precipitation for Finney County is 48.5 centimeters: 14.3 centimeters from March through May, 20.1 centimeters from June through August, 9.7 centimeters from September through November, and 4.4 centimeters from December through February.

TABLE 1.—Number of Commercial Fields of Each Crop or Cover Type in the Field Measurements Test Sites, 1976

Cover type	Test site		
	Finney County, Kansas	Hand County, South Dakota	Williams County, North Dakota
Winter wheat	85	43	2
Spring wheat	—	38	222
Barley	—	9	—
Oats	—	15	4
Rye	—	1	1
Fallow	87	32	212
Corn	18	17	—
Grain sorghum	61	5	—
Alfalfa	1	26	1
Other hay crops	—	23	1
Pasture	9	48	61
Other	14	8	29

The major crops in Finney County are wheat and grain sorghum, which account for about 60 and 20 percent, respectively, of the total cropland. The majority of wheat is produced following summer fallow practices, although an increasing amount is being irrigated. Winter wheat is seeded in September or early October, then is dormant from December to February. Green-up occurs in March; the crop is fully headed by mid-May; and harvest is typically completed during the first week of July.

Williams County, North Dakota: This test site is located at latitude 48°19' N and longitude 103°25' W. It is representative of the cool semiarid areas of the northern Great Plains where annual precipitation averages 33 to 38 centimeters. The site is at an elevation of 650 meters and lies in the glaciated area with a drift mantle and an undulating to steep surface.

The soils in the site are of the Mollisol order, Boroll suborder, with Williams and Williams-Zahl being the major associations present. Both occur on undulating to rolling landscapes and are well to excessively drained. Much of the surface drainage is to depressions. The soils were developed from calcareous glacial till and are suitable for cropland and pasture. The soils of the Williams association are very productive.

The climate of the area is typically continental, with long cold winters, short warm summers, wide diurnal ranges in temperature, frequent strong winds, and limited (as well as uncertain and highly variable) precipitation. Average amounts of precipitation are 4.6, 15.5, 12.2, and 4.3 centimeters in the winter, spring, summer, and fall, respectively.

The major crop is wheat, which occupies about 70 percent of the grain crop acreage. Both hard red and durum spring wheats are grown. Most of the wheat is grown on summer fallow land. The major cover types in the site are wheat, summer fallow, and pasture; limited acreages of rye, barley, alfalfa, and flax are also grown. The cropping calendar for the spring wheats begins with seedbed preparation in late April to early May. Planting is generally in mid-May; heading occurs from late June to mid-July; and harvest is from mid- to late August.

Hand County, South Dakota: The test site is in the north-central Great Plains at latitude 44°34' N and longitude 99°00' W. It is a transition area with the Corn Belt to the east, spring wheat producing areas to the north, and winter wheat producing areas to the south. The boundary between the subhumid lowland of eastern South Dakota and the more arid Great

Plains area of central and western South Dakota passes through Hand County. The area is nearly level to gently undulating. The principal soils of the test site are Houdek and Bonila, which are in the Mollisol order, Ustoll suborder. They are dark-colored permeable loams underlaid by slowly permeable glacial till.

Hand County has a continental climate. Winters are long and cold, and summers are warm. The average annual precipitation is 47 centimeters; typically, 33 to 36 centimeters fall between April and September. The county is subject to frequent weather changes, and airmasses that pass through the area bring a wide variety of temperature and moisture conditions.

The principal crops of Hand County are winter and spring wheat, pasture and hay, corn, barley, and oats. Most wheat is grown following summer fallow.

Agriculture experiment stations.—Agronomic experiments with wheat and other crops were available for study at the agriculture experiment stations at Garden City, Kansas, and Williston, North Dakota. The research farms are operated by Kansas State University and North Dakota State University. The advantages of using experimental plots at the stations were that (1) considerable amounts of agronomic data describing the treatments and their effects on the growth and development of the crop could be readily obtained, and (2) sources of difference in spectral response could be more readily determined since only the factors of interest were varied while other factors were held constant. The replicated experiments were designed to provide a range of growing conditions typical of those found in the intensive test sites. The crops were planted, grown, and harvested using conventional practices and equipment.

A small-grains and a wheat experiment were conducted at each location. The treatments and experimental designs for the 1977 crop year for each location are shown in figures 4 and 5.

Garden City, Kansas: The objective of the small-grains experiment was to determine whether various small grains can be discriminated from each other on the basis of their spectral reflectance. The experiment included four winter wheat varieties and one variety each of barley, rye, and triticale (fig. 4).

The principal objective of the dryland and irrigated winter wheat experiments was to characterize crop spectral response as a function of crop maturity and to relate the spectral response to crop variables

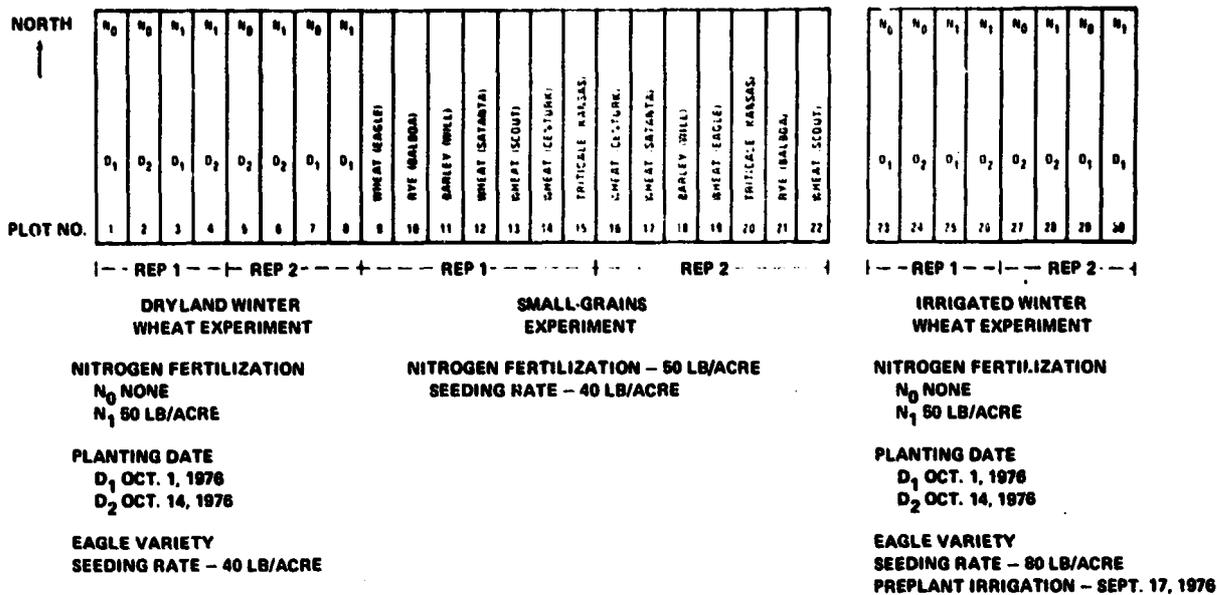


FIGURE 4.—Remote-sensing experiments at the Garden City, Kansas, agriculture experiment station.

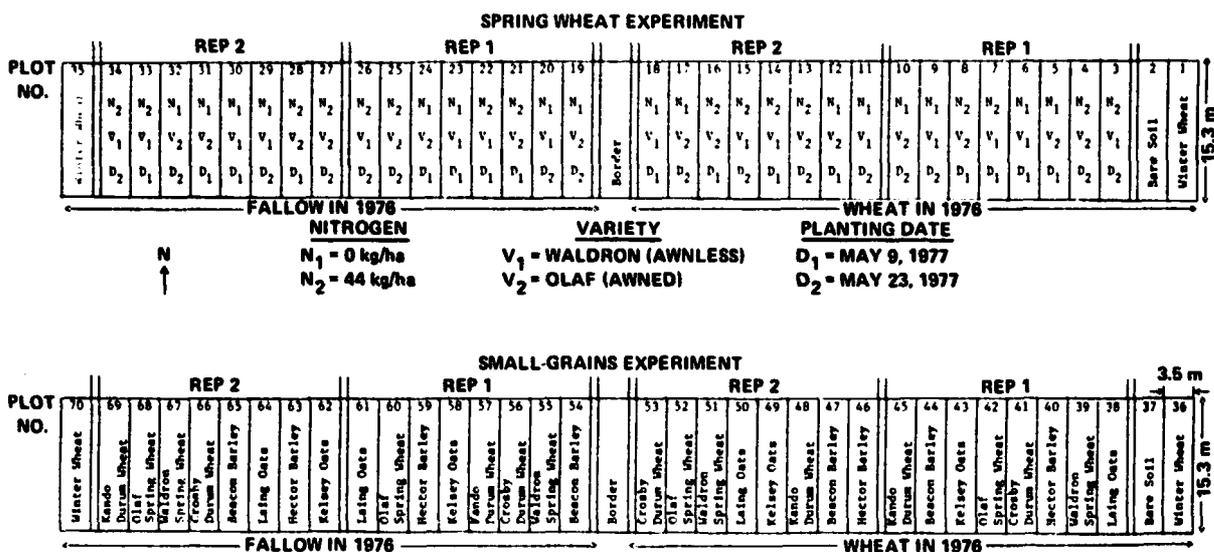


FIGURE 5.—Remote-sensing experiments at the Williston, North Dakota, agriculture experiment station.

such as leaf area index and biomass and to cultural variables such as planting date, irrigation, and nitrogen fertilization. The treatments were selected to give a range of leaf area indexes and biomass at any particular maturity stage or measurement time.

Williston, North Dakota: The objective of the small-grains experiment was the same as for the

small-grains experiment at Garden City. This trial included two varieties each of hard red spring wheat, durum wheat, oats, and barley (fig. 5).

The objective of the spring wheat experiment was to quantify the effects on spectral response of the major variables affecting wheat growth, development, and yield. The factors and levels of each factor

TABLE II.—Characteristics of the Multispectral Scanner Systems

Characteristic	Landsat-1 and -2 MSS	JSC MSS	JSC MMS ^a
Spectral range, μm	0.5 to 1.1	0.4 to 13	0.4 to 1.1 10 to 12
Number of bands	4	24	11
Total field of view, deg	11.56	80	110
Normal operational altitude, km	944	0.5 to 6.1	0.5 to 6.1
Instantaneous field of view, m	79	1 to 12	1 to 15
Precision of data, bits	5 to 6	8	8

^aModular multiband scanner.

were (1) soil moisture—wheat in 1976 and fallow in 1976, (2) cultivar—standard height and semidwarf, (3) planting date—early and late, and (4) nitrogen fertilization—0 and 34 kg/ha.

Sensor Descriptions

The characteristics of the primary sensors used to acquire spectral data over the intensive test sites and agriculture experiment stations are described in this section. The sensors used in the intensive test sites included Landsat MSS, airborne MSS, helicopter-borne spectrometer, and tripod-mounted Landsat-band radiometers. The sensor systems acquiring spectral data at the agriculture experiment stations were the truck-mounted spectroradiometer and interferometer systems operated by LARS and JSC, respectively. General descriptions of the sensor systems are given in the following sections and summarized in tables II and III.

Landsat multispectral scanner.—Landsat 1 and 2 MSS data were acquired at 18-day intervals. The MSS data have four spectral bands from 0.5 to 1.1 micrometers. The sensor scans cross-track swaths of 185 kilometers. Computer-compatible tape (CCT) data and imagery (both color and black and white) were requested for each cloud-free overpass of the intensive test sites.

Airborne multispectral scanners.—During 1975, the 24-channel scanner (ref. 2) operated by JSC was the

TABLE III.—Characteristics of the Spectrometer Systems

Characteristic	JSC FSS ^a	LARS Exotech 20C	JSC FSAS ^b
Spectral range, μm	0.4 to 2.5 8 to 14	0.4 to 2.4 3 to 14	0.4 to 2.5 3 to 14
Spectral resolution at 1.0 μm , μm	0.025	0.025	0.0064
Scan time, scan/sec	1	0.033 to 2.0	10
Field of view, deg	22	0.75, 15	11
Normal operational altitude, m	60	6	6
Data storage format	Digital tape	Analog tape	Digital tape
Camera			
Focal length, mm	76	55	50
Field of view, deg	36	43	46
Film type	70 mm color	35 mm color	35 mm color

^aField spectrometer system.

^bField signature acquisition system.

primary scanner system; during 1976 and 1977, the 11-channel modular multiband scanner (MMS) was used.² Color and color-infrared photography was obtained during the scanner flights to be used as reference data by analysts.

Helicopter-borne field spectrometer system.—The helicopter-borne field spectrometer system (FSS) is a filter wheel spectrometer instrument that is a modification of the S-191 sensor used in the Skylab Earth Resources Experiment Package (EREP) (ref. 3). The FSS has been modified by NASA for mounting on a helicopter (fig. 6). The instrument produces data in 14-track digital format that are converted to CCT's for subsequent reformatting and analysis.

The spectral range of the spectrometer is 0.42 to 2.50 and 8.0 to 14.0 micrometers. The field of view is 22°, which gives a spot size of 24 meters diameter from a 60-meter altitude. The helicopter flies at 100 km/hr. The camera has a 76-millimeter focal length and a 36° field of view, giving 40 meters square ground coverage.

Truck-mounted spectrometer systems.—The Exotech Model 20C field spectrometer operated by LARS acquires spectral data over the visible, reflective infrared, and thermal infrared wavelength regions (ref. 4). The instrument consists of two inde-

²“Modular Multiband Scanner (MMS).” JSC Internal Note No. 74-FM-47, NASA Johnson Space Center, Houston, Tex., 1974.



FIGURE 6.—Helicopter spectrometer system operated by the NASA Johnson Space Center acquiring measurements of the calibration standard. The spectrometer is located just below the window. The calibration standard is a specially coated canvas panel 6.5 by 13 meters in size mounted on a flat level platform.

pendently functioning units. The short-wavelength unit senses radiation from 0.38 to 2.4 micrometers and the long-wavelength unit senses radiation from 2.7 to 5.4 and 7.0 to 13.5 micrometers. The short-wavelength unit is equipped with a translucent diffusing plate, which is used to monitor incident spectral irradiance. Each optical head has a reflective fore-optic system that permits remote selection of the field of view (0.75° or 15°).

The instrument is mounted on a mobile aerial tower that operates with an instrumentation van containing the control electronics and data recorder for the system (fig. 7). The data produced by the instrument are recorded on an analog magnetic tape recorder and later converted into digital information by a laboratory analog-to-digital converter. Calibration sources designed for field use are used to calibrate the spectrometer onsite. Calibrated spectral data and field observations are combined on digital magnetic tapes during the data reformatting process.

The Block wideband interferometer (Field Signature Acquisition System or FSAS) operated by JSC acquires spectral data over the visible and infrared portions of the spectrum (ref. 5). The instrument scans the spectrum rapidly enough to account for environmental variables and is equipped with a self-contained computer system that yields spectral data from the interferograms produced by the instrument. The instrument control electronics and computer are mounted in an instrument van, and the optical head of the instrument is mounted on a mobile aerial tower. The spectral data (expressed as wave

numbers) produced by the instrument were processed by JSC to provide CCT's of spectral reflectance factor calibrated with respect to wavelength.

Landsat-band radiometers.—Four-band radiometers (Exotech Model 100) with the same spectral bands as the Landsat MSS were operated by Purdue University and Texas A & M University to acquire measurements in selected fields at the Finney County and Williams County test sites to support canopy modeling studies. In addition, during 1977, measurements were made throughout the growing season of the plots at the Williston experiment station, using a radiometer mounted on a lightweight van. These measurements, made at hourly intervals, are being used to investigate the diurnal variation in reflectance of wheat.

Meteorological and atmospheric sensors.—Standard meteorological instrumentation was used to obtain measurements of temperature, humidity, windspeed and wind direction, barometric pressure, and total irradiance. Solar radiometers were used to obtain optical depth measurements in six visible and infrared bands during Landsat overpasses and during aircraft and helicopter missions.

Sensor Calibration and Correlation

A key objective of the LACIE Field Measurements Project was the acquisition of calibrated multispectral data. Calibrated data are required to (1) facilitate comparisons of data from different sensors and (2) compare and relate spectral measurements made at one time and location to those made at other times and locations.

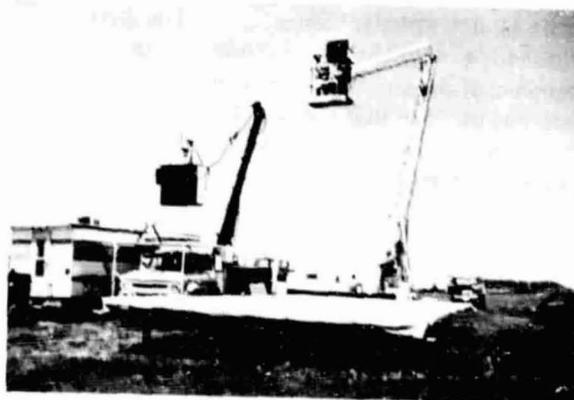


FIGURE 7.—Field spectroradiometer system operated by LARS making spectral measurements of small-grains plots.

To have comparable data, scene reflectance was chosen as the measured property rather than scene radiance. Scene reflectance is a property only of the scene, whereas scene radiance is a property of the illumination also. Calibration largely removes the effects of varying illumination and measurement conditions because of changing Sun angle, atmospheric conditions, and sensor. The bidirectional reflectance distribution function gives the most complete description of the reflectance characteristics of a surface. However, because this property is difficult to measure, more common use is made of the reflectance factor.

Reflectance factor is defined as the ratio of incident radiant flux reflected by a sample surface to that which would be reflected into the same reflected beam geometry by a perfectly diffuse (Lambertian) standard surface identically irradiated and viewed (ref. 6). Because the principal component of the irradiance is direct solar irradiance and the measurement is made in a relatively small cone angle (15° to 20°), the term "bidirectional reflectance factor" is used to describe the measurement. One of the directions is specified by the solar zenith and azimuth angles; the other is specified by the zenith and azimuth viewing angles.

Because no perfectly reflecting diffuser is available, painted barium sulfate (BaSO_4) reference surfaces, which are highly diffuse, were used (ref. 7). The spectral bidirectional reflectance factor of these surfaces was measured in both the laboratory and the field by processes that are traceable to the reflectance of pressed barium sulfate (fig. 8). A correction using the published reflectance of the pressed barium sulfate enables the computation of an approximation of the bidirectional reflectance factor.

Because of the presence of skylight, the measurement is not strictly bidirectional. The process of eliminating skylight by subtracting the spectral response of the shadowed scene and shadowed standard has merit in that it could remove the effects of the skylight. However, the additional measurements and calculations add uncertainty to each computed reflectance. This uncertainty is greater than the effect itself (ref. 8). Furthermore, because the interest of the project was in producing data directly relatable to satellite data, which includes the effects of the skylight, the single comparison method was used. Because the dominant effects are due to the directional nature of the irradiance, the term "bidirectional reflectance factor" is appropriate to describe the measurements.

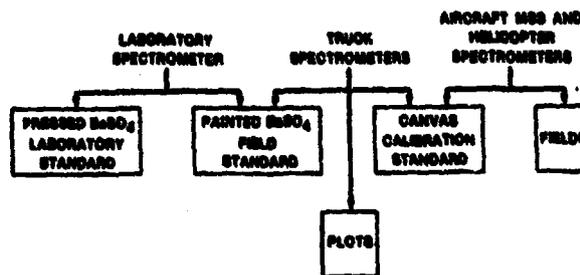


FIGURE 8.—Schematic of the procedures used for reflectance calibration of LACIE field measurements data.

Calibration of truck-mounted spectrometer systems.—Temperature variations, dust, vibration, and other adverse factors associated with field measurements require that calibration be performed at the field site. The procedures chosen reflect the availability of suitable standards and the principle that the calibration measurements be obtained under the same conditions as the target measurements.

The short-wavelength unit was calibrated for spectral reflectance factor. A standard based on the highly reflecting properties of barium sulfate was used as a basis for the reflectance factor calibration. The standards were prepared according to procedures described by Shai and Schutt (ref. 7).

The painted barium sulfate field standard was used to fill the field of view of the instrument under nearly the same conditions as for the measurement of plots. For the simplest calibration, the response to the standard, the response to the scene, the full-dark response (automatically provided during each spectral scan), and the spectral reflectance properties of the standard are used to compute the bidirectional reflectance factor. Since it is inconvenient to make this direct comparison for each measurement, the solar port is frequently used to transfer the reflectance standard for the LARS Exotech 20C system.

The calibration calculation consists of forming the ratio of the instrument response for the target to that for the reflectance standard and correcting for the known reflectance of the standard. This procedure produces a reflectance factor for the given Sun angle and normal viewing of the target.

During the calibration observations, the instrument was aimed straight down at the reflectance standard from a distance of 2.4 meters for the Exotech 20C system and 1 meter for the FSAS system. Care was taken to ensure that the standards were not shadowed and that the illumination conditions were as similar as possible to the conditions of the obser-

vation of the subject. Calibration observations were performed at approximately 15-minute intervals.

Wavelength calibration of the reflective wavelength unit was accomplished by irradiating the solar port with sources having known spectral lines (ref. 9). The primary sources are the General Electric A100.H4T mercury vapor lamp and the helium Pluecker tube. A field wavelength calibrator based on the helium tube was chosen for use because it has at least one strong line in the range of each section of the circular variable filters.

Calibration of the helicopter-borne FSS.—The helicopter-borne spectrometer was calibrated using a 60-percent reflectance canvas panel and the measurements made by the truck-mounted spectrometer of the canvas panel. These in turn were related to the measurements of the barium sulfate painted panels and the pressed barium sulfate standard.

The calibration procedure used deals with limitations imposed by the size and location of the standard by calibrating the instrument at a low altitude (6 meters) and collecting data over the flightlines at 60 meters. This procedure assumes that atmospheric absorption and path radiance are negligible for a 60-meter path.

The absence of an onboard solar sensor integrated into the instrument makes it desirable that calibrations be performed as frequently as possible. Therefore, the reflectance panels were centrally located and procedures were followed which allowed calibration within 15 minutes of any data acquisition (beginning of each flightline of data collection).

The data processing facility converts the FSS data to bidirectional reflectance factor based on the measurements made of the barium sulfate standard and the canvas panel. The calibration calculation consists of forming the ratio of the FSS response for the target to that for the canvas standard and correcting for the measured reflectance of the canvas standard. This procedure produces a reflectance factor for the given solar illumination angle and normal viewing of the subject.

Field calibration of the FSS with respect to emissive radiation was accomplished by recording spectral observations of a blackbody at a temperature below ambient and another blackbody at a temperature above ambient. The subsequent scans of subject scenes were converted to spectral radiance using linear interpolation.

Calibration of airborne MSS data.—The reflective data from the airborne MSS can be calibrated to reflectance using the five gray canvas panels located

at the site and the spectral bidirectional reflectance factor measurements made by the truck-mounted spectrometers over the canvas panels. The nominal reflectances of the panels are 6, 12, 18, 30, and 60 percent.

The gray panel reflectance factor and MSS response data collected at low altitude (500 meters above the panels) can be related through linear regression. The regression equation can then be used to transform the low-altitude airborne MSS data to bidirectional reflectance factor. Fields overflowed at the lower altitude can, in turn, be used as calibration targets to transform higher altitude data to bidirectional reflectance factor.

The emissive MSS data can be calibrated by means of the two blackbodies at known temperatures located in the scanner and viewed with each scan of the scene.

Sensor correlation procedures.—The three major sensor systems—the truck-mounted spectrometers, the helicopter-borne spectrometer, and the aircraft MSS—can be correlated using the spectral data collected by each system over common targets; i.e., five 6- by 12 meter gray canvas panels (fig. 9). The aircraft scanner collected data over the panels during each mission. The helicopter and truck spectrometer systems measured the reflectance of the four darker gray panels during correlation experiments performed during each crop year. The calibration measurements made of the brightest canvas calibration panel by the helicopter and truck spectrometer systems were also used in correlating the sensors.

All the spectrometers were brought together in 1977 for complete calibration and correlation. This included measurement of common targets and reflectance standards (fig. 10), comparison of data collection procedures, and evaluation of instrument performance.

Data Acquisition

The collection of multispectral remote-sensing, agronomic, and meteorological data for the intensive test sites and agriculture experiment stations is described in this section.

Intensive test sites.—This section discusses spectral data collection procedures and the measurements and observations of crop, soil, and meteorological parameters in the intensive test sites.

Spectral data collection: Helicopter-borne FSS, airborne MSS, and tripod-mounted radiometer data



FIGURE 9.—Canvas gray-level panels used for correlation of spectrometer systems and calibration of aircraft scanner data.



FIGURE 10.—Truck-mounted spectrometer systems operated by LARS and JSC preparing to measure the reflectance of the canvas calibration panel in a sensor correlation experiment. The common target was measured simultaneously by these systems, as well as by the helicopter spectrometer, to relate the instrument responsivities.

were collected within a 3- or 4-day mission window. Whenever possible, data were obtained on the same day and time as the Landsat overpasses.

The helicopter spectrometer data were obtained under stable atmospheric conditions with 20 percent or less cloud cover at solar elevation angles greater than 30° . At the test site, six 9.7-kilometer flightlines were flown by the helicopter in three sets of two lines. Flightlines were flown at an altitude of 60 meters, at 100 km/hr groundspeed, and in an east-west direction. Reference panel calibration measurements were made from a 6-meter altitude immediately before flying each set of two flightlines. Correlation of spectra and fields was accomplished

using simultaneously acquired 70-millimeter color photography. A total incidence pyranometer was located at the helicopter calibration site to provide a strip-chart record of the irradiance conditions on the day of data acquisition (usually beginning 1 hour before and ending 1 hour after the data acquisition period). These strip charts provide the data analyst with a visual record of the irradiance conditions at the site during helicopter and MSS data acquisition.

The airborne scanner system acquired data over the intensive test sites and agriculture experiment stations concurrently with data collection by the helicopter spectrometer. The intensive test sites were overflown at 3300- and 7000-meter altitudes and the experiment stations and calibration panels at 500 meters. Collecting data at the two altitudes over the test site flightlines provided different spatial resolutions and different amounts of atmosphere between the scene and sensor. Data collection requirements specified that cloud cover be less than 30 percent and solar elevation greater than 30° . Color and color-infrared photography (23 centimeters) was obtained simultaneously with the scanner data.

A Landsat-band radiometer mounted on a 2-meter tripod was used to collect data from one to three fields in the Finney County and Williams County test sites. The measurements were made at four times during the day to provide four different Sun angles. A painted barium sulfate field standard was measured between the measurements of the canopy. The spectral measurements include wheat canopy reflectance, soil reflectance, the ratio of diffuse to total irradiance, and leaf transmittance. Canopy description data include leaf area index, biomass, number of tillers and leaves, and photographs. The photographs include vertical and 45° views and plant profile scenes. When possible, these data were acquired at five maturity stages (seedling, tillering, jointing, flowering, and ripe) at several locations in typical fields.

Agronomic data collection: Agronomic measurements and observations were acquired describing the condition of each of the fields for which spectral data were collected. These agronomic data describe the condition of each field as fully as possible and are used to account for differences in the spectral measurements. The data were recorded on standard forms, keypunched, and transmitted to LARS for inclusion in the data bank. Data describing all fields in the intensive test sites were collected by USDA ASCS (ref. 10). The following data were collected during the spring and fall inventories: field number,

acreage, crop species and variety, irrigation, fertilization, planting date, and other descriptive information.

Periodic observations coinciding with Landsat overpasses and aircraft/helicopter missions were made to describe the condition of the fields. The variables observed were maturity stage, percent soil cover, plant height, surface moisture condition, stand quality, quality relative to other fields in the site, field operations, density of stand, weed infestation, and growth/yield detractants. Vertical 35-millimeter photographs were taken, and additional descriptive comments were added as appropriate. Grain yields of selected fields were measured at harvest time.

Meteorological data collection: The following atmospheric and meteorological measurements were made in conjunction with FSS and aircraft scanner data collection at the intensive test sites: percentage and type of cloud cover, wet and dry bulb temperature, barometric pressure, total irradiance, windspeed and wind direction, and optical depth at seven visible and near-infrared wavelengths. Daily measurement records of temperature, precipitation, relative humidity, soil temperature, and wind were obtained from the nearest weather station. In addition, rainfall was recorded at six to eight locations in each test site.

Agriculture experiment stations.—The collection of spectral, agronomic, and meteorological data at the agriculture experiment stations is described in this section.

Spectral data collection: The spectral data at the agriculture experiment stations were collected by JSC at Garden City, Kansas, and by LARS at Williston, North Dakota. The primary sensors were the Block wideband field interferometer and the Exotech Model 20C field spectroradiometer. During 1975, an Exotech Model 20D similar to the Model 20C was operated by the NASA Earth Resources Laboratory at Garden City. These were augmented by Barnes PRT-5 precision radiation thermometers boresighted with the spectrometers. To obtain data that could be readily compared, the interferometer and spectroradiometer were operated following similar procedures. The instruments were operated from their aerial towers at 6 meters above the target to minimize the shadowing of skylight and yet ensure that the field of view of the instrument contained only the subject of interest. Care was taken to avoid scene shadowing and to minimize the reflective interaction caused by personnel or vehicles. The routine data-taking mode of the instruments is

straight down. Two measurements of each plot were made by moving the sensor so that a new scene within the plot filled the field of view.

To minimize the effect of solar elevation changes on the spectral response, measurements were made only when the Sun angle was greater than 45° above the horizon in the late spring and summer and greater than 30° in the late fall and early spring.

Data recorded at the time of each measurement included date, time, reference illumination, air temperature, barometric pressure, relative humidity, windspeed and wind direction, percentage and type of cloud cover, field of view, latitude, longitude, and zenith and azimuth view angles. Periodically during the day, spectral measurements of skylight were recorded by spectrometers with a solar port. A 35-millimeter color photograph of each observation was taken from the aerial tower, as were oblique ground-level photographs of each plot.

Agronomic data collection: Crop and soil information for the plots at the research stations were collected at Garden City by JSC with assistance from the agriculture experiment station personnel and at Williston by LARS. At the beginning of the season, information describing the species and cultivar, irrigation practices, fertilization history, soil type, and planting date was obtained for each plot.

Observations made at the time of each mission for each plot included maturity stage; plant height; percent soil cover; surface soil moisture and roughness; stand quality; field operations such as cultivation or harvesting; stress factors (insect damage, disease, nutrient deficiencies, moisture stress, weeds, or lodging); leaf area index; number of stems, leaves, and heads; fresh weight of plants; dry weights of stems, leaves, and heads; and soil moisture profile. Vertical and oblique 35-millimeter color photographs were taken, and grain yields were measured at harvest time.

Meteorological data collection: Percentage and type of cloud cover, wet and dry bulb temperature (or relative humidity), barometric pressure, total irradiance, and windspeed and wind direction were measured in conjunction with the truck-mounted spectrometer data collection. Daily measurement records of air temperature, humidity, radiation, wind, precipitation, and soil temperature were also obtained from the nearest weather station.

Summary of data acquisition missions.—Table IV summarizes the data acquisition for the 1976 crop year at Finney County, Kansas, and Williams County, North Dakota, for the major sensors involved in

the experiment. In each year, as in 1976, an effort was made to obtain data at each of the important growth stages of wheat at each level of the sampling scheme, from controlled experimental plot to Landsat scene. Whenever possible, helicopter spectra and aircraft scanner data were gathered near the time of a Landsat overpass. A complete schedule of acquired data for each location is given in the data library catalogs discussed in the following section.

Data Processing, Library, and Analysis Systems

An important aspect of the project was to prepare the data for later analysis according to uniform formats and to register the agronomic, meteorological, and measurement data with the spectral data. Following processing, data were cataloged in the data library and distributed to interested researchers. Software for interactive plotting and analysis of the data has also been developed.

Data processing.—Before computer processing, spectrometer data were evaluated manually using strip charts of raw data, photographs, records of system parameters, and strip charts of irradiance conditions. Computer processing of spectrometer data included calibration, data-quality evaluation, reformatting, and storage. In order to compare data from the different spectrometers, the spectrometer data were processed in a standardized format with bandwidths of 0.01 micrometer from 0.4 to 2.4 micrometers and of 0.05 micrometer from 2.7 to 14 micrometers. Spectrometer data were merged with ancillary data for storage on nine-track computer-compatible tapes.

Aircraft scanner data were first converted to visicorder imagery for manual evaluation of data quality. This imagery is a 13-centimeter wide, medium-contrast paper strip record of the data for individual channels. Aircraft scanner data were also subjected to computerized examination to validate the performance of the sensors and the data-recording system. In addition, a strip chart of total irradiance was used to verify the irradiance conditions during the overflight. The scanner data were then processed to nine-track computer-compatible tapes in LARSYS format.

Landsat MSS data were previewed from black-and-white transparencies of the image for each band to establish data quality and cloud cover conditions within the intensive test sites and the complete

TABLE IV.—Summary of Data Acquisition by Wheat Growth Stage and Sensor System for Kansas and North Dakota Test Sites, 1975-76

Test site/ mission date	Wheat growth stage	Sensor system			
		Landsat MSS	Aircraft scanner	Helicopter spectrom- eter	Truck spectrom- eter
<i>Kansas</i>					
Sept. 14 to 17, 1975	Preemergence	—	X	X	—
Oct. 2 to 6	Emergence	X	—	X	X
Oct. 20 to 22	Seedling	X	—	X	X
Nov. 11 and 12	Tillering	—	X	X	X
Mar. 13 to 19, 1976	Tillering	X	X	X	—
Mar. 30 to Apr. 2	Tillering	X	—	X	X
Apr. 9 and 10	Jointing	X	—	—	—
Apr. 18 to 21	Jointing	X	X	X	X
Apr. 27 and 28	Jointing	X	—	—	—
May 4 to 7	Preboot	—	X	X	X
May 14 to 16	Boot	X	—	—	X
May 24 to 27	Heading	X	—	—	X
June 2 and 3	Milk	X	—	—	—
June 11 to 13	Dough	X	—	X	X
June 20 and 21	Ripening	X	—	—	X
June 29 to July 2	Mature	X	X	X	X
July 18 and 19	Postharvest	X	—	—	—
<i>North Dakota</i>					
May 10 to 14, 1976	Emergence	X	X	X	—
May 28 to 30	Seedling	X	—	X	X
June 7 to 9	Tillering	—	—	—	—
June 15 to 17	Jointing	—	—	X	X
June 25 to 27	Boot	X	X	X	X
July 4 to 8	Heading	X	X	X	X
July 13 to 17	Dough	—	—	—	X
July 20 to 23	Ripening	X	X	X	—
July 28 to 31	Mature	—	—	X	X
Aug. 6 to 12	Harvest	X	—	X	X
Aug. 17 to 20	Postharvest	—	X	X	—

Landsat frames. Following data-quality evaluation, Landsat MSS data were processed to nine-track computer-compatible tapes.

Data library and distribution.—The multispectral data library maintained at LARS for the LACIE Field Measurements Project contains over 100 000 spectra (corresponding to measurements of over 800 plots and fields) and over 2000 observations made with Landsat-band radiometers (ref. 11). The library

also includes several hundred scenes of aircraft- and satellite-acquired scanner data.

A data library catalog was prepared for each crop year containing summary and detailed schedules of data acquisition by location, sensor system, and mission. Digital data products available for analysis include Landsat and airborne scanner data, helicopter- and truck-spectrometer spectra and ancillary data, and tripod-radiometer spectra and ancillary data (fig. 11). Aerial and ground-level photography acquired concurrently with spectrometer and scanner data is also available.

Data have been routinely distributed to researchers at ERIM, LARS, and Texas A & M in conjunction with the Supporting Research program sponsored by JSC. In addition, data have been provided to the NASA Goddard Space Flight Center, Goddard Institute for Space Studies, and General Electric Corporation. Copies of data sets are provided to qualified, interested investigators for the cost of reproducing the data. Requests for data should be addressed to Chief, Earth Observations Division, Mail Code SF, NASA Johnson Space Center, Houston, TX 77058.

Data analysis systems.—LARSYS (Version 3.1) is a fully documented software system designed to provide the tools for analysis of MSS data (ref. 12). The pattern recognition and interactive data handling techniques in LARSYS have been used worldwide for analysis of aircraft and Landsat scanner data in many applications.

EXOSYS is a specialized software system developed at LARS for analysis of spectrometer data. It provides researchers with the capability to sort spectrometer data by combinations of measurement (e.g., solar elevation) and ancillary variables (e.g., leaf area index). Analysis features of EXOSYS include the ability to compute functions of band-averaged reflectances, perform correlations with crop parameters, and fit polynomial curves to the data using a least squares technique. Initial results are reviewed and then sent to a line printer or a graphics plotter.

RESULTS OF SELECTED ANALYSES OF FIELD MEASUREMENTS DATA

To realize the full potential of remote sensing for crop identification, condition assessment, and yield prediction, it is important to understand and quan-

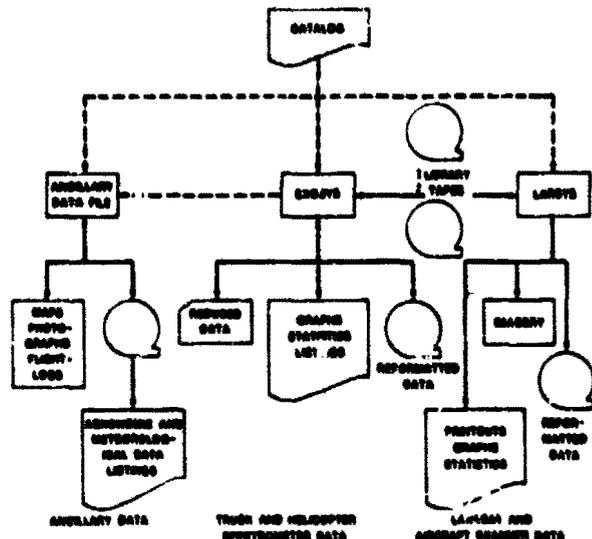


FIGURE 11.—Organization and data products of the LACIE field measurements data library.

tify the relation between agronomic and spectral characteristics of crops. Equally important is the development of improved capabilities for accurately measuring the spectral, spatial, and temporal variations of agricultural scenes and for extracting meaningful information from these data. The LACIE field measurements data make important contributions in both of these areas. The data are particularly useful because many measurements and observations of the crops and their agronomic characteristics were recorded throughout several seasons. Furthermore, complete spectra permit simulation of the response in any specified wavelength band, and the radiometric calibration of the data permits valid comparisons to be made among different sensors and different dates and locations of data collection.

Many of the factors affecting the reflectance properties of plant leaves have been identified and investigated through laboratory measurements. The relationships of physical-biological parameters, such as chlorophyll concentration, water content, and leaf morphology, to the reflectance, transmittance, and absorption of leaves have been well established. Some of the papers and reviews describing these relationships include Gates et al. (ref. 13), Brece and Holmes (ref. 14), Gausman et al. (ref. 15), and Sinclair et al. (ref. 16).

Although knowledge of the reflectance characteristics of single leaves is basic to understanding the reflectance properties of crop canopies in the field,

this information cannot be applied directly to the field situation because there are significant differences between the spectra of single leaves and the spectra of canopies. The reflectance characteristics of canopies are considerably more complex than those of single leaves because in canopies there are many more interacting variables. Some of the more important agronomic parameters influencing the reflectance of field-grown canopies are leaf area index, biomass, leaf angle distribution, leaf color, percent soil cover, and soil color. Differences in these parameters are caused by variations in many cultural and environmental factors, including planting date, cultivar, seeding rate, fertilization, soil moisture, and temperature.

The data analysis phase of the LACIE field research began in 1976 by addressing several of the LACIE critical issues, particularly the discrimination of wheat and small grains. More recently, the analyses have been extended to address objectives related to future crop inventory systems, such as the use of remote sensing to gather information about crop condition and yield. Because it would not be possible to describe adequately the results of all the investigations, several studies that are representative of the types of investigations that have used field measurements data have been selected for this report.

Several of the other studies that have used the LACIE field measurements data are briefly summarized here. Landgrebe et al. (ref. 17) used the spectrometer and MSS data to simulate and evaluate alternative combinations of scanner system parameters for the thematic mapper, such as the instantaneous field of view and signal-to-noise ratio. A comparison of the Landsat MSS and thematic mapper wavelength bands for crop identification has been made by Bauer et al. (ref. 8). The thermal measurements from the helicopter spectrometer have been examined by Harlan et al. (ref. 18), and Bauer et al. (ref. 8) developed a model describing the radiant temperature characteristics of spring wheat canopies in relation to the geometry of the canopy and environmental variables. As part of the LACIE field measurements research, Vanderbilt et al. (ref. 19) developed a method to obtain information on the geometrical properties of crop canopies needed for canopy reflectance and radiant temperature models. Berry and Smith (ref. 20) used LACIE field measurements data to test a canopy reflectance model to predict the spectral response in the Landsat MSS wavelength bands of winter wheat with varying

amounts of leaf area and as a function of Sun angle. A nondestructive method to estimate leaf area index involving analysis of digitized aerial photographs was developed by Harlan et al. (ref. 18).

Objectives

Of the research objectives listed in the second section of this paper, the following are addressed as examples of current research results from LACIE field measurements.

1. To determine the relationship of agronomic variables such as biomass and leaf area index to multispectral reflectance of spring wheat
2. To determine the effects of cultural and environmental factors on the spectral response of wheat
3. To assess the spectral discriminability of spring wheat and other small grains
4. To determine the early-season threshold for detection of wheat
5. To compare and evaluate the Landsat MSS and thematic mapper bands for prediction of crop canopy variables

The results obtained to date for these objectives are presented in subsequent sections, following a summary of the approach used.

Experimental Approach

The data used for the analyses in this report were acquired during 1975 and 1976 in Kansas and North Dakota by the helicopter- and truck-mounted spectrometer systems at approximately 10- to 14-day intervals during the wheat growing seasons. Bidirectional reflectance factor and agronomic data were acquired in approximately 75 fields from each of the intensive test sites and 60 plots from the agriculture experiment stations.

Correlation and regression analyses were used to relate biological and physical variables describing the canopies to spectral response. To relate the reflectance measurements more directly to Landsat, the analyses were performed using reflectance data averaged into bands corresponding to the Landsat MSS and thematic mapper spectral bands. The " tasseled-cap" transformation (ref. 21) was also used to determine the greenness and brightness components of the Landsat MSS band reflectances for some of the analyses.

Relation of Landsat and Field Measurements of Spectral Response

Frequently the question is asked whether results from analyses of field measurements data can be related to and applied to Landsat MSS data. To help answer this question, analyses of the relationship between Landsat MSS data and helicopter spectrometer measurements of the spectral response were performed (ref. 22). Landsat data for 135 fields and 5 dates were correlated with the reflectances measured by the helicopter spectrometer. The Landsat data were first adjusted using the XSTAR algorithm (ref. 23) to minimize differences among the five dates in Sun angle and atmospheric conditions. As shown in figure 12, for MSS band 4 (0.5 to 0.6 micrometer), the two sets of measurements are highly correlated; similar relationships were found for the other spectral bands. Using empirical relationships such as these, or results of radiative transfer modeling (ref. 24), crop discriminability can be predicted by relating measured reflectance differences to corresponding differences in Landsat signals.

Prediction of Crop Canopy Characteristics From Reflectance Measurements

One of the major long-term goals of agricultural remote sensing is to estimate from spectral measurements crop variables that can subsequently be used to assess crop vigor or be entered into a yield prediction model. To achieve this goal, the complex relationship between the spectral reflectance of crop canopies and their biological and physical characteristics must be understood.

One of the LACIE field research objectives (ref. 25) was to determine the relationship of canopy characteristics to reflectance and to assess the potential for estimating these characteristics from remotely sensed measurements of reflectance. The variables selected for analysis are indicators of crop vigor and growth, which could be used to augment agromet models of crop growth and yield.

This section treats the effect of varying amounts of vegetation and of maturity stage on the spectra of spring wheat canopies, the relation of canopy variables to reflectance in different regions of the spectrum, and the potential capability to predict canopy variables from reflectance measurements. As part of the analysis, the wavelength bands of current

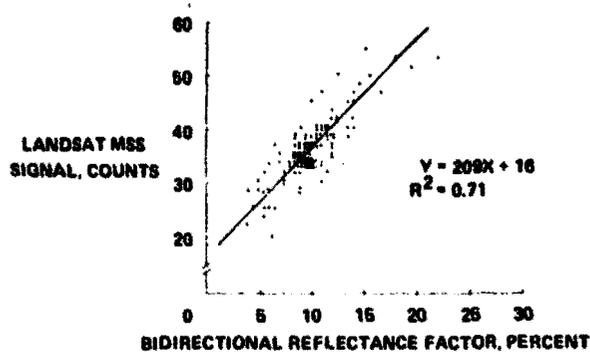


FIGURE 12.—The relation of Landsat MSS and helicopter spectrometer measurements of spectral response for five acquisition dates in Finney County, Kansas (MSS band 4, 0.5 to 0.6 micrometer).

and proposed satellite MSS systems were compared. The measurements were made at the Williston, North Dakota, Agriculture Experiment Station on nine different dates during the summer of 1976.

The amount of vegetation present is one of the principal factors influencing the reflectance of crop canopies. Figure 13 illustrates the effect of the amount of vegetation (as measured by leaf area index, percent soil cover, biomass, and plant height) on the spectral response during the period between tillering and the beginning of heading, when the maximum green-leaf area is reached. As leaf area and biomass increase, there is a progressive and characteristic decrease in reflectance in the chlorophyll absorption region, increase in the near-infrared reflectance, and decrease in the middle-infrared reflectance.

Plant development and maturity (as opposed to growth or increase in size) cause many changes in canopy geometry, moisture content, and pigmentation of leaves. These changes are also manifested in the reflectance characteristics of crop canopies. Figure 14 shows the spectra of spring wheat at several different maturity stages (changes in the amount of vegetation are also occurring).

The linear correlations of five canopy variables with reflectances in the proposed thematic mapper (Landsat D) and Landsat MSS bands are listed in table V. The relationships of percent soil cover, leaf area index, fresh biomass, and plant water content with reflectance in selected wavelength bands are shown in figure 15. The correlations and plots include data from all treatments for the stages of maturity when the canopy is green, seedling through flowering.

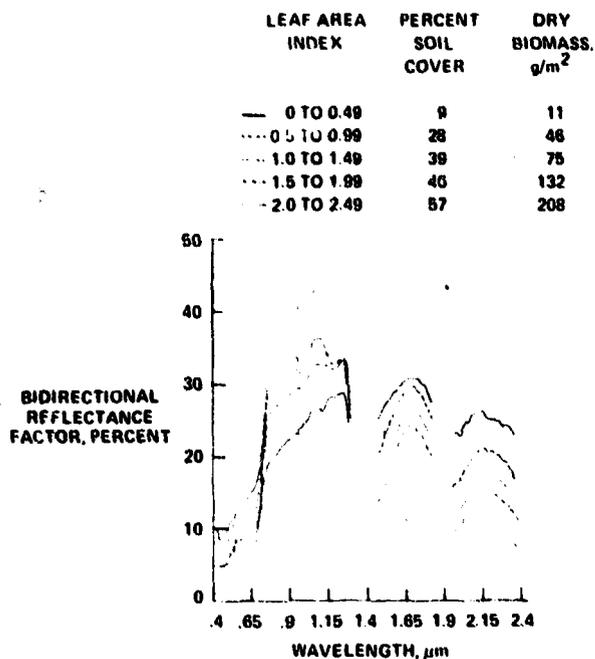


FIGURE 13.—Effect of leaf area index, percent soil cover, dry biomass, and plant height on the spectral reflectance of spring wheat during the period between tillering and the beginning of heading, when the maximum green-leaf area is reached. Data were acquired at Williston, North Dakota, from May 28 to June 18, 1976, and include plots with different soil moisture levels, planting dates, nitrogen fertilization, and cultivars.

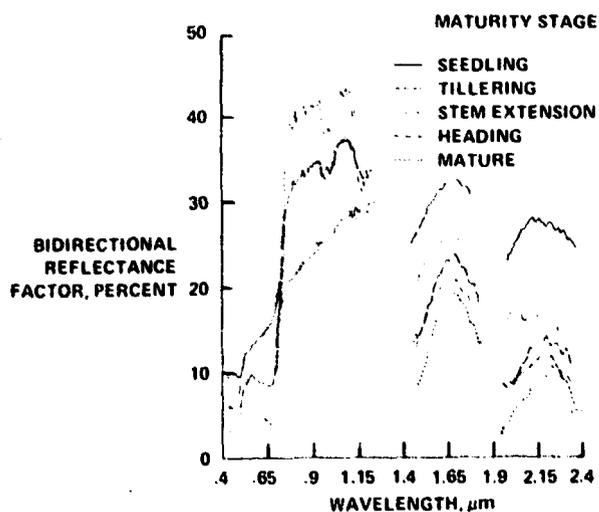


FIGURE 14.—Spectral reflectance of spring wheat canopies at several maturity stages. Measurements were made at Williston, North Dakota, from May to August, 1976, and include plots with different soil moisture levels, planting dates, nitrogen fertilization, and cultivars.

TABLE V.—The Linear Correlations (r) of Reflectances in the Proposed Thematic Mapper and Landsat MSS Wavelength Bands With Percent Soil Cover, Leaf Area Index, Fresh and Dry Biomass, and Plant Water Content

Wavelength band, μm	Percent soil cover	Leaf area index	Fresh biomass	Dry biomass	Plant water content
<i>Thematic mapper</i>					
0.45 to 0.52	-.82	-.79	-.75	-.69	-.76
0.52 to 0.60	-.82	-.78	-.81	-.77	-.82
0.63 to 0.69	-.91	-.86	-.80	-.73	-.81
0.76 to 0.90	.93	.92	.76	.67	.79
1.55 to 1.75	-.85	-.80	-.83	-.79	-.84
2.08 to 2.35	-.91	-.85	-.86	-.81	-.86
<i>Landsat MSS</i>					
0.5 to 0.6	-.82	-.79	-.81	-.76	-.81
0.6 to 0.7	-.90	-.85	-.81	-.74	-.82
0.7 to 0.8	.84	.84	.57	.46	.60
0.8 to 1.1	.91	.90	.77	.68	.79

Fresh biomass, dry biomass, and plant water content correlate most highly (table V) with reflectance in the middle-infrared band, 2.08 to 2.35 micrometers. Percent soil cover and leaf area index correlate most highly with a near-infrared band, 0.76 to 0.90 micrometer. The visible wavelengths were less sensitive to leaf area and biomass; similar results have also been reported by Colwell (ref. 26) and Tucker (ref. 27). Other canopy variables analyzed that were not correlated with reflectance were plant height, percent green leaves, and percent plant moisture.

These and other analyses of the data indicate that the amount of photosynthetically active (green) vegetation has a dominant influence on the reflectance characteristics of crop canopies. This observation is substantiated by the decrease in the correlation of canopy variables and reflectance as the canopy begins to senesce or ripen (refs. 25 and 28).

Understanding the relation of the agronomic properties of crop canopies to reflectance in various regions of the spectrum is the first step in the development of models using spectral measurements. The remainder of this section describes the regression models developed for prediction of crop growth characteristics.

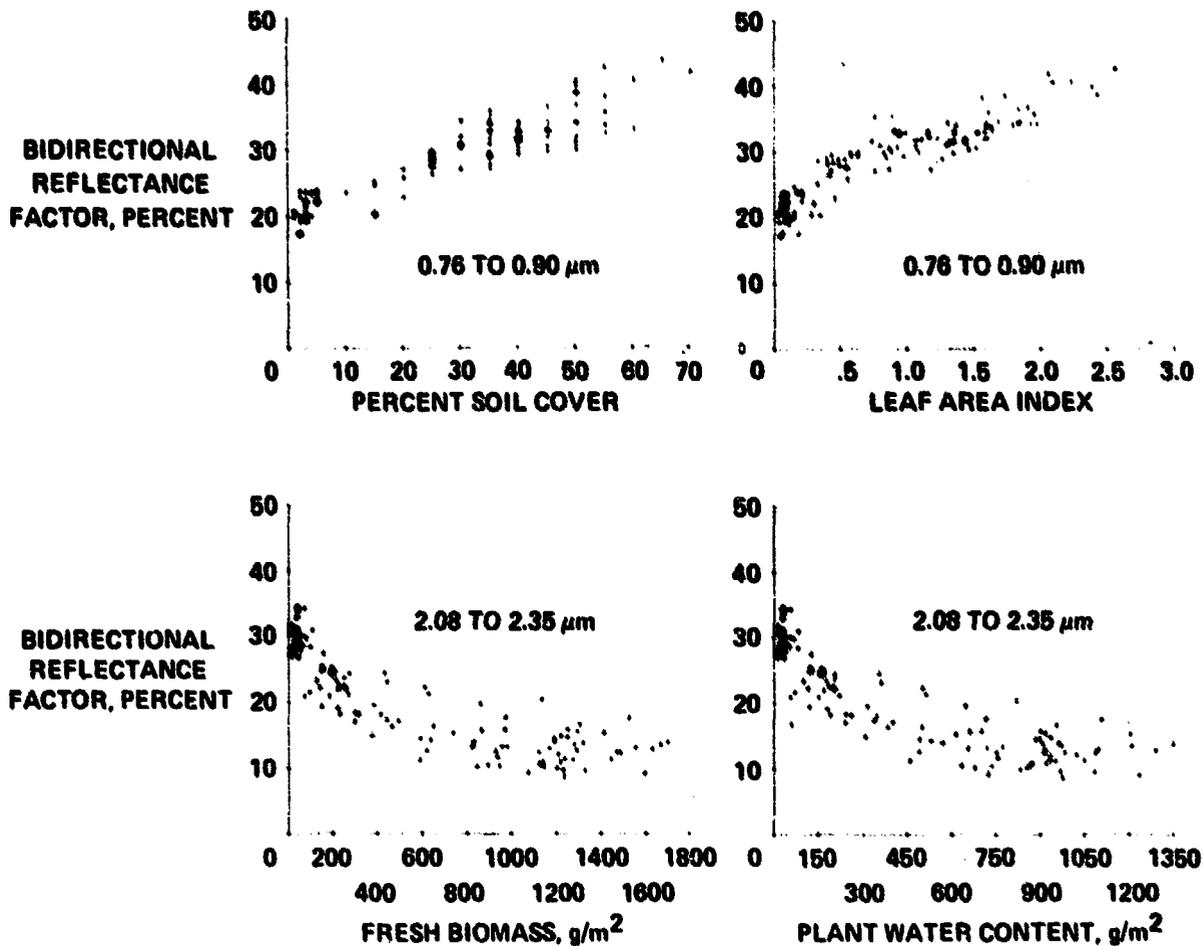


FIGURE 15.—Relationship of percent soil cover, leaf area index, fresh biomass, and plant water content to reflectance in selected wavelength bands.

Table VI shows results for selections of one to six wavelength bands to predict canopy variables. By computing all possible regressions, the best subset of one to six wavelength bands was selected, considering the amount of variability explained and the bias of the resulting regression equation. The near- and middle-infrared bands were found to be most strongly related to the canopy variables. For leaf area index and percent soil cover, the 0.76- to 0.90-micrometer wavelength band accounts for more of the variation than any other single band. The 2.08- to 2.35-micrometer wavelength band is the single most important band for predicting the variation in fresh biomass, dry biomass, and plant water. The 2.08- to 2.35-micrometer wavelength band is one of the two most important bands in explaining the variation in percent soil cover and one of the three most impor-

tant bands in explaining the variation in leaf area index.

The relationships between the measured and predicted leaf area index and percent soil cover are shown in figure 16. Similar results were obtained for the other canopy variables. The results show that reflectance measurements in a small number of wavelength bands in important regions of the spectrum can explain much of the variation in canopy characteristics and can be used to estimate canopy variables such as leaf area index and biomass.

Table VII shows the maximum R^2 value obtained for predictions of each canopy variable using the Landsat MSS bands, the best four thematic mapper bands, and all six reflective thematic mapper bands. In every case, the best four thematic mapper bands explained more of the variation in a canopy variable

TABLE VI.—Selection of Combinations of the Best 1, 2, . . . 6 Wavelength Bands for Estimating Percent Soil Cover, Leaf Area Index, Fresh Biomass, Dry Biomass, and Plant Water Content During the Seedling to Flowering Stages of Crop Development

Canopy variable	No. bands entered	R ²	C _p (a)	Bands entered, μm					
				0.45 to 0.52	0.52 to 0.60	0.63 to 0.69	0.76 to 0.90	1.55 to 1.75	2.08 to 2.35
Percent soil cover	1	.86	132				X		
	2	.92	16				X		X
	3	.92	15	X			X		X
	4	.93	4	X	X		X		X
	5	.93	5	X	X	X	X		X
	6	.93	7	X	X	X	X	X	X
Leaf area index	1	.84	37				X		
	2	.87	7				X	X	
	3	.88	2				X	X	X
	4	.88	4	X			X	X	X
	5	.88	5	X	X		X	X	X
	6	.88	7	X	X	X	X	X	X
Fresh biomass	1	.73	239						X
	2	.76	211		X				X
	3	.83	109		X	X			X
	4	.88	41		X	X	X		X
	5	.90	12	X	X	X	X		X
	6	.93	7	X	X	X	X	X	X
Dry biomass	1	.65	252						X
	2	.67	229	X	X				X
	3	.81	78		X	X			X
	4	.84	44		X	X	X		X
	5	.87	20	X	X	X	X		X
	6	.88	7	X	X	X	X	X	X
Plant water content	1	.75	201						X
	2	.77	175		X				X
	3	.83	98	X	X		X		X
	4	.88	34		X	X	X		X
	5	.90	9	X	X	X	X		X
	6	.90	7	X	X	X	X	X	X

^aThe regression equation is unbiased when the C_p value is less than or equal to the number of terms (wavelength bands) in the equation.

than the four Landsat bands. Addition of the other two thematic mapper bands resulted in small increases in the R² values.

The lower correlations (table V) and predictions (table VII) of the Landsat MSS bands compared to the thematic bands are attributed to the width and location of the bands with respect to the spectral characteristics of vegetation. For example, the data in table V demonstrate a disadvantage of collecting data in the 0.7- to 0.8-micrometer wavelength range. The inclusion in this band of the region (near 0.7 micrometer) of rapid transition from the chlorophyll absorption region of the spectrum to the highly

reflecting near-infrared region (0.70 to 0.74 micrometer) results in a weaker relation between reflectance and crop canopy variables. Similar results were reported by Tucker and Maxwell (ref. 29). This low correlation reduces the usefulness of the data in the 0.7- to 0.8-micrometer wavelength band.

Effect of Agronomic and Environmental Factors on Spectral Reflectance

The crop canopy is a dynamic entity influenced by many agronomic and environmental factors. The

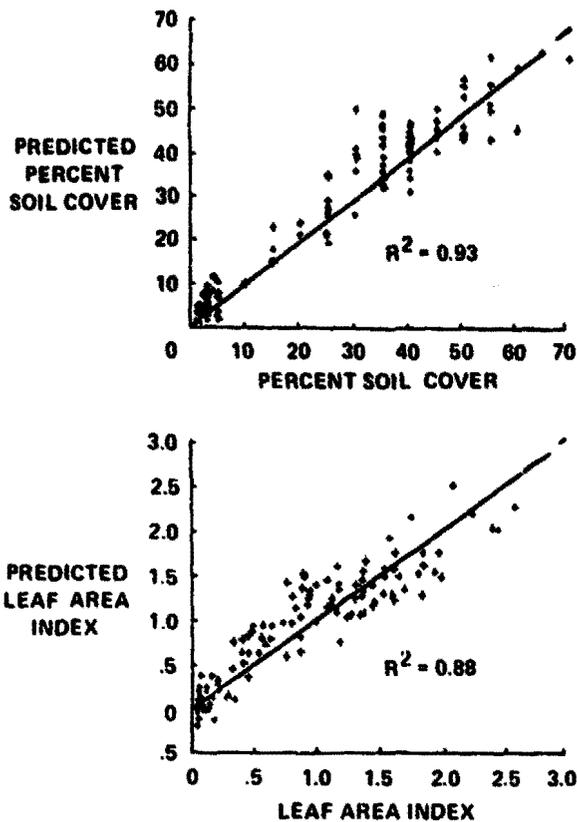


FIGURE 16.—Comparison of measured and predicted percent soil cover and leaf area index of spring wheat.

effects of several cultural and environmental factors on the reflectance of spring wheat were investigated using data acquired at the Williston, North Dakota, Agriculture Experiment Station. The examples of spring wheat spectra from selected measurement dates shown in figure 17 illustrate the effects of available soil moisture, planting date, nitrogen fertilization, and cultivar. Additional examples of crop spectra from the other test sites, crops, and years have been compiled by Hixson et al. (ref. 30).

The Waldron (standard height) cultivar planted early on fallow land with nitrogen fertilization was selected as a standard of comparison. One treatment at a time was varied from this standard, permitting comparisons of reflectance spectra measured on plots with different soil moisture levels, planting dates, fertilization, and cultivar. All treatment comparisons were made using spectra acquired on June 18, 1976, during the stem extension phase of development, except for the comparison of cultivars, which was on July 16, after heading.

In 1976, the effects of available soil moisture on plant growth and spectral response were quite significant. Wheat planted on fallow land had more tillers and, therefore, greater biomass, leaf area, and percent soil cover than the wheat crop grown on land that had been cropped the previous year. These differences account for the decreased visible reflectance, increased near-infrared reflectance, and decreased middle-infrared reflectance in the fallow treatment. The effect of planting date on spectral response is also illustrated in figure 17. The differences are attributed to differences in the amount of vegetation present, as well as differences in maturity stage.

Adding nitrogen fertilizer increased the amount of green vegetation early in the growing season. The fertilized treatment had the spectral characteristics of a greener, denser vegetative canopy—decreased red reflectance, slightly greater near-infrared reflectance, and reduced middle-infrared reflectance.

The two wheat cultivars, Olaf (semidwarf, awned) and Waldron (standard height, awnless), were similar in appearance before heading. After heading, some differences between the two cultivars were apparent but are probably not significant. The greatest spectral differences were in the middle infrared, indicating a difference in the moisture and biomass between the two cultivars at this growth stage.

In one analysis (ref. 22), one-way multivariate analyses of variance were performed on the Landsat MSS band reflectance data from individual plots of

TABLE VII.—The R^2 Values for Predictions of Percent Soil Cover, Leaf Area Index, Fresh and Dry Biomass, and Plant Water Content With Four Landsat MSS Bands, the Best Four Thematic Mapper Bands, and the Six Thematic Mapper Bands

Wavelength bands	Percent soil cover	Leaf area index	Fresh biomass	Dry biomass	Plant water content
Landsat MSS bands	0.91	0.86	0.86	0.84	0.85
Best four thematic mapper bands ^a	.93	.88	.88	.84	.88
Six thematic mapper bands	.93	.88	.91	.88	.90

^aSee table VI.

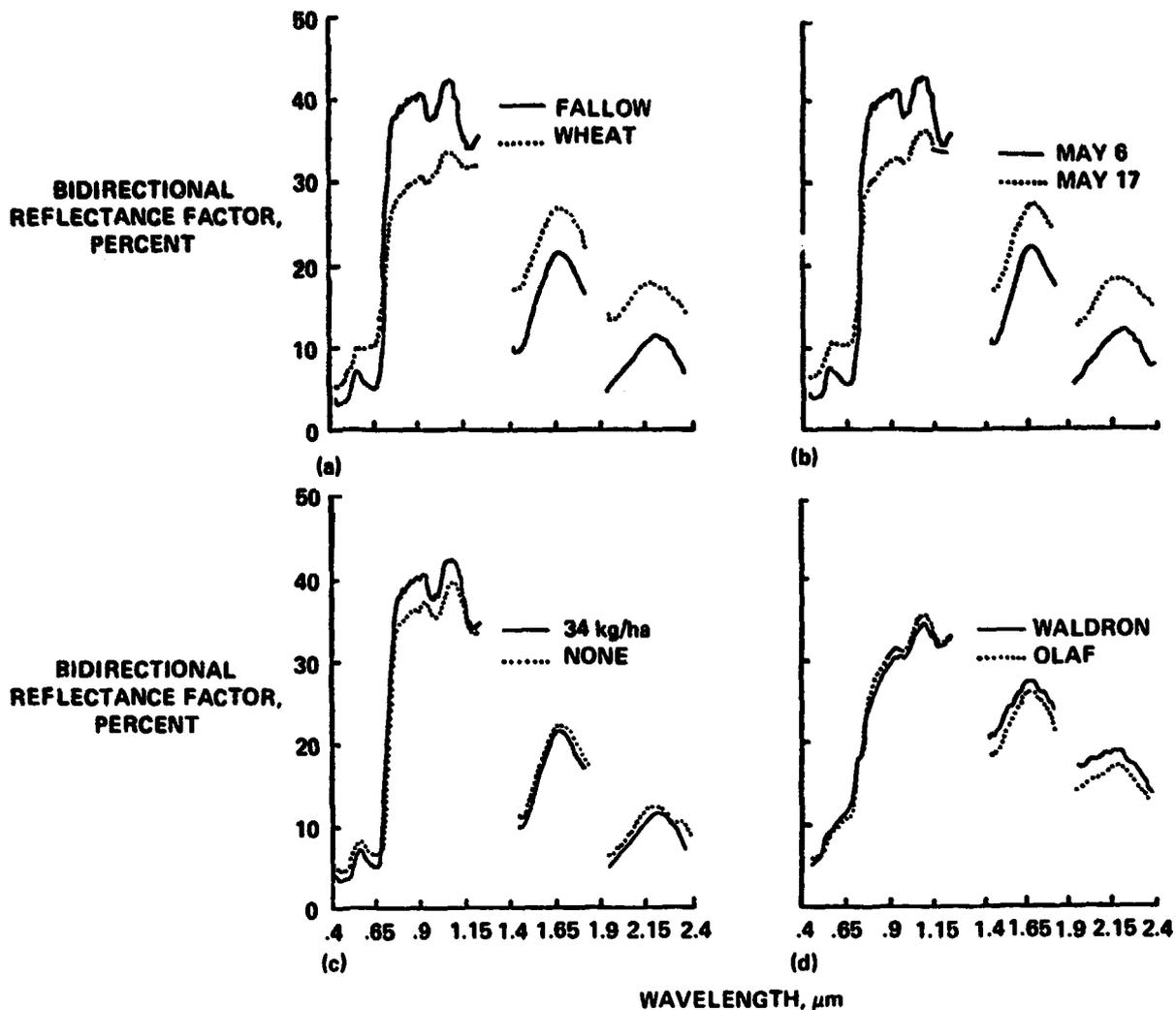


FIGURE 17.—Effects of agronomic treatments on the spectral reflectance of spring wheat. Spectra were measured on June 18 during the stem extension stage of development, except for the spectra of cultivars which were measured on July 16 after heading. (a) Previous crop. (b) Planting date. (c) Nitrogen fertilization. (d) Cultivar.

spring wheat using data from the entire season. Soil moisture availability was found to be the most significant factor. A decreased moisture supply, caused by planting wheat for a second year in succession on the same plot, both decreased the magnitude of green development from that of wheat planted on fallow ground and delayed the date of maximum greenness. A similar delay in maximum greenness was observed when the planting date was delayed by 10 days, but the difference in maximum greenness levels was not as pronounced as in the case where available soil moisture was reduced.

Figure 18 illustrates the effects of the soil moisture and nitrogen fertilization factors on the maximum values attained by the greenness and brightness components of reflectance for the small-grains test plots. Maximum greenness is most affected by soil moisture as plentiful soil moisture produces more vegetation, which covers the soil. Nitrogen fertilization was observed to affect the greenness component in a similar fashion, with the greening value of nitrogen fertilizer being very evident on those plots that were continuously cropped. Soil moisture also affected the brightness component.

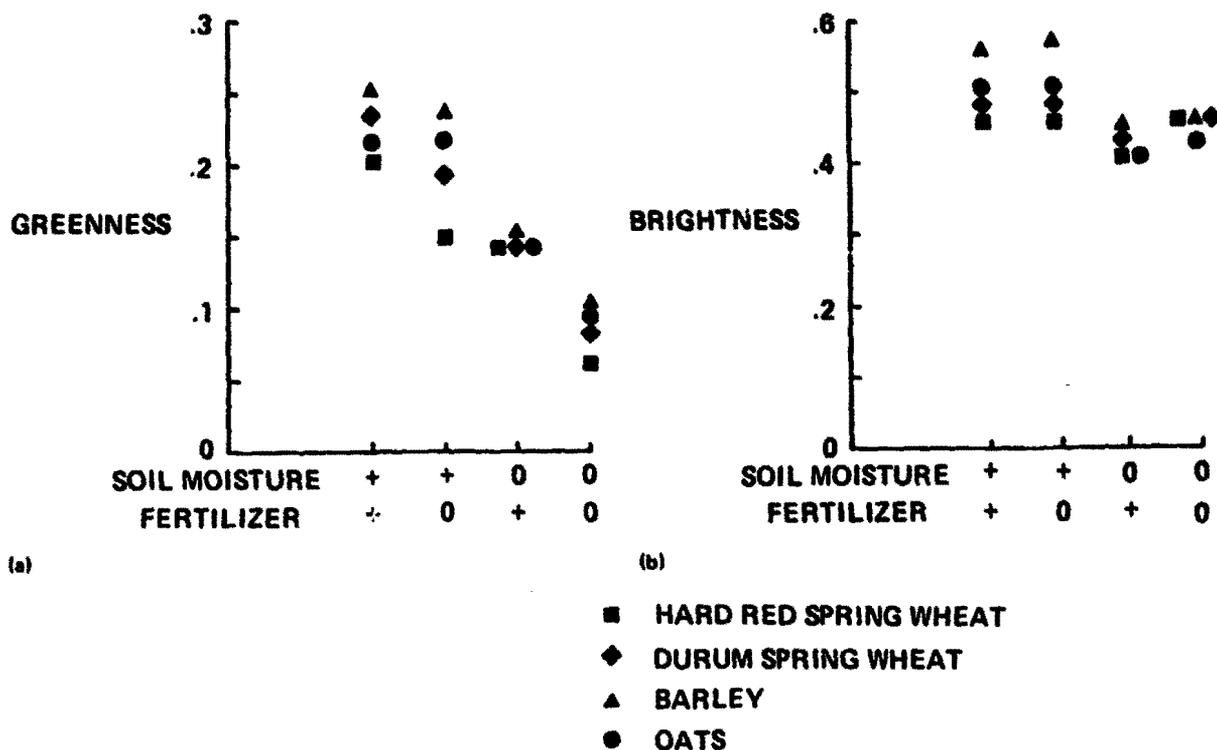


FIGURE 18.—Effects of soil moisture and nitrogen fertilization on the maximum greenness and brightness components of reflectance of small grains. (a) Maximum greenness, June 18. (b) Maximum brightness, July 29 to August 6.

These spectra and analyses illustrate the dynamic character of the canopy and the many factors that influence the spectral reflectance of the canopy. More quantitative analyses of the effect of agronomic treatments and environmental variables on reflectance of wheat are currently being conducted.

Spectral Discrimination of Spring Wheat From Other Small Grains

One of the critical issues that arose during LACIE was spectrally discriminating spring wheat from the other spring small grains. These crops have similar reflectance spectra and crop calendars; consequently, LACIE initially did not attempt to inventory them separately. Instead, a small-grains area estimate was obtained, and historical data on crop production were used to establish spring-wheat-to-small-grains ratios for producing a spring wheat estimate. It was found that these ratios could vary appreciably from year to year, introducing errors in the spring wheat estimates. Consequently, some supporting research

effort was directed toward investigation of spectral techniques for achieving such discrimination (ref. 22). Although major emphasis was placed on analysis of Landsat data from LACIE blind sites, analysis of field measurement data played a strong supportive role, along with analysis of USDA crop statistics. Only data from the first 2 years were included in this analysis. An expanded small-grains experiment was conducted in 1977.

The objectives of the analyses of field measurements data on spectral reflectance of wheat and small grains were (1) to characterize the spectral reflectance of spring wheat and other small grains as a function of time throughout the growing season, (2) to characterize the sources and extent of variability to be expected, and (3) to develop discrimination techniques for distinguishing between spring wheat and other small grains through increased understanding of Landsat signals for these crops.

The change in spectral character of spring wheat reflectance at five maturity stages was illustrated in figure 14, while similarities and some differences of

spring wheat, barley, and oats spectra on three dates are shown in figure 19. The spectral patterns throughout the growing season were determined for spring wheat and other small grains. One technique was to plot the time trajectories of transformed reflectance values for these crops and look for differences that might prove useful for discrimination. Linear combinations of values, analogous to the tasseled-cap transformation of Landsat data (see next section), were used to form greenness and brightness components of reflectance.

Figure 20 presents spectral trajectories for hard red spring wheat, barley, and oats; durum spring wheat is very similar to hard red spring wheat. Each trajectory is for a crop that had been planted on prior-year-fallowed soil (more available soil moisture than continuously cropped soil) and had been fertilized. Thus, they represent spectral patterns for the best growing conditions available at the experiment station. Although the general shapes of the spectral trajectories are similar, several differences can be seen among them; notably, barley attained greater values in both greenness and brightness before heading and its brightness upon ripening was greater than that of wheat. Less distinctiveness was observed in the spectral characteristics of other plots with crops that were grown under less favorable conditions (fig. 18).

An analysis of color photographs (fig. 21) and agronomic measurements (table VIII) made in conjunction with the spectral measurements helps to explain the physical causes of the observed spectral differences and variability. Grown under favorable conditions, the barley had greater biomass, leaf area index, and percent soil cover than spring wheat, resulting in higher maximum greenness values. The barley matured and ripened 1 week to 10 days before the spring wheat. Longer lighter colored awns, drooping heads, and greater soil cover all contributed to a greater maximum brightness for barley at maturity.

In summary, analyses of field measurements data provided insights into the causes of the spectral characteristics of spring wheat and barley that may prove useful for discrimination. For instance, differences in greenness and brightness at heading and brightness at ripening and the timing of these events appear key to their spectral discrimination. One preliminary operational technique for direct spectral classification of spring wheat was tested during LACIE Phase III and improved techniques are currently under development.

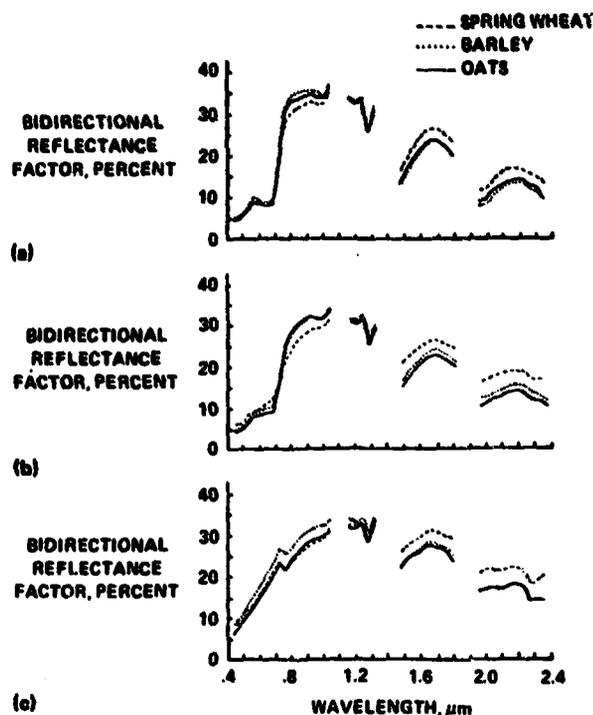


FIGURE 19.—Spectra of spring wheat, barley, and oats at stem extension, heading, and ripe stages of maturity. (a) Stem extension, June 18, 1976. (b) Heading, July 16, 1976. (c) Ripe, July 29, 1976.

Early-Season Detection of Wheat

In LACIE, it was found that early-season estimates of winter wheat area tended to be low and unreliable, because the emergence and development of green vegetative cover on the soil are variable because of differences in planting dates, crop rotation, irrigation and fertilization practices, and local weather.

A study was conducted using LACIE field measurement data to investigate the threshold of wheat detectability in Landsat data (ref. 22). Helicopter-spectrometer and agronomic data acquired for 10 dates during the 1975-76 growing season at the Finney County, Kansas, intensive test site were analyzed.

Figure 12 illustrated the reflectance spectra for fields with different amounts of vegetation. To relate these data to Landsat analysis, reflectance values for the Landsat MSS bands were computed. A useful technique for Landsat MSS data analysis has been to form linear combinations of the bands, defining a new coordinate system for describing the data. One such transformation, the tasseled-cap transforma-

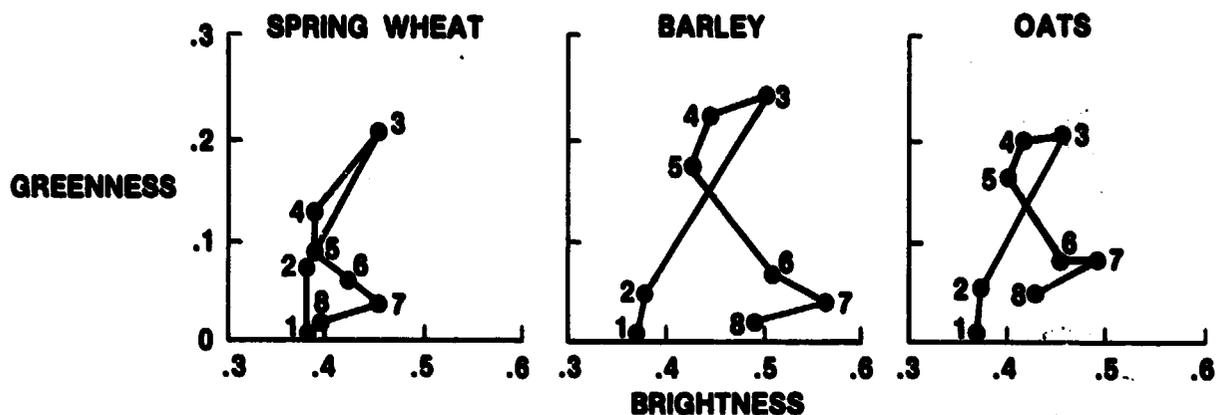


FIGURE 20.—Comparison of selected greenness-brightness trajectories of spring wheat, barley, and oats. The measurement dates and growth stages are as follows: (1) May 28, seedling; (2) June 3, tillering; (3) June 18, jointing; (4) July 8, heading; (5) July 15, milk; (6) July 29, dough; (7) August 1, ripe; (8) August 6, ripe. Note: Barley was in the milk stage on July 8 and was ripe by July 29.

tion, defines a brightness variable that aligns closely with the direction followed by reflectances of varied soils. Orthogonal to brightness is the greenness variable, which is oriented toward the spectral response from healthy green vegetation. These two components describe most of the variability observed in Landsat scanner measurements of agricultural scenes (see the paper by Kauth et al. entitled "Feature Extraction Applied to Agricultural Crops as Seen by Landsat"). A principal components analysis of the reflectances revealed that 98 percent

of the variability was in a plane analogous to the greenness and brightness plane for Landsat MSS data.

Four fields with different management practices were selected to illustrate the relationship of the greenness component of reflectance to measurement data (fig. 22). The absence of fall green development in the late-planted fields and the appreciable fall greening-up of the field that was irrigated and planted at the normal time are apparent.

The proportions of late- and early-planted fields and irrigated and nonirrigated fields will vary from site to site, as will other factors that determine development rates. Yet, it is of interest to determine both how the collection of wheat fields in the Finney County site developed in 1975-76 and how well they would have been detected by a decision rule that called them wheat if their greenness component of reflectance exceeded a given threshold by a given date.

To provide a quantification of the greening-up characteristics of this group of fields, histograms were computed to describe the percentage of fields exceeding a given greenness value as a function of acquisition date. Figure 23 displays these results in two ways: (1) with fixed threshold levels and varied dates and (2) with fixed dates and varied thresholds. With a threshold of 0.06, 95 percent of the wheat fields would have been detected on the eighth mission (May 6, 1976), 63 percent of them on the seventh mission (April 18), and 38 percent on the sixth mission (March 31). For a lower threshold of 0.04, the corresponding percentages would have been 100,

TABLE VIII.—Agronomic Characteristics of Small Grains on Three Measurement Dates at Williston, North Dakota, Agriculture Experiment Station^a

Date	Small grains	Percent soil cover	Leaf area index	Fresh biomass, g/m ²	Maturity stage ^b
June 18	Spring wheat	50	1.5	567	3.3
	Barley	90	2.9	1326	3.4
	Oats	80	2.0	1022	3.4
July 16	Spring wheat	30	.7	1162	5.1
	Barley	70	1.3	1686	5.2
	Oats	50	1.2	1388	5.1
July 29	Spring wheat	30	—	854	5.2
	Barley	60	—	961	5.4
	Oats	50	—	820	5.3

^aThe plots are the same ones shown in figures 19 to 21, were grown on fallow land, and received nitrogen fertilizer

^bMaturity stages: 3.3 to 3.4, stem extension; 5.1, milk; 5.2, soft dough; 5.3, hard dough; 5.4, ripe.

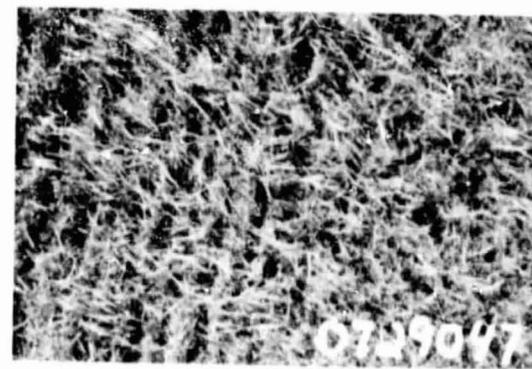
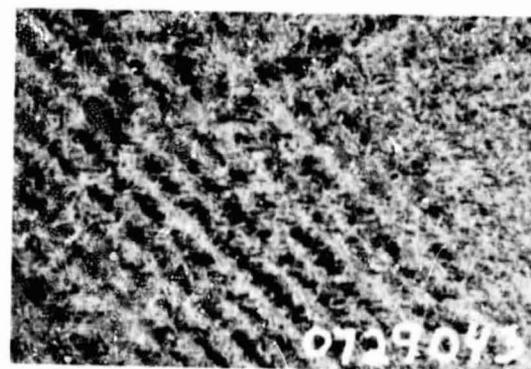
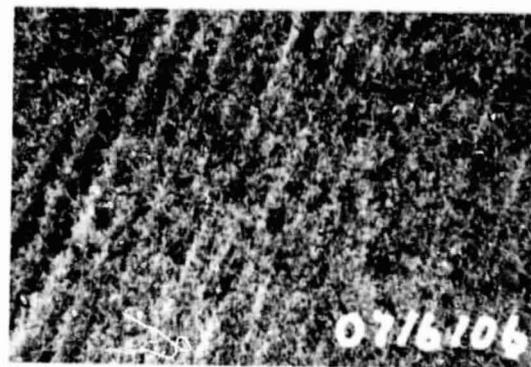
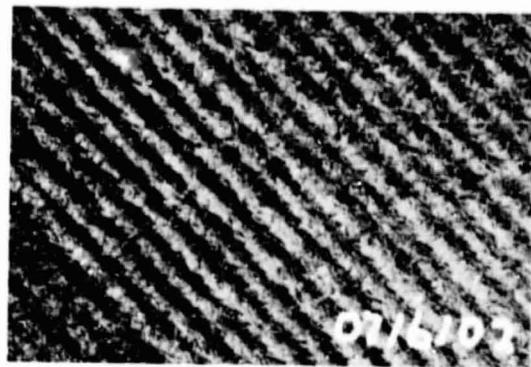


FIGURE 21.—Comparison of oblique and vertical views of spring wheat and barley canopies at two stages of maturity. (a) Spring wheat, July 16, 1976, day 198, milk. (b) Barley, July 16, 1976, day 198, milk. (c) Spring wheat, July 29, 1976, day 211, hard dough. (d) Barley, July 29, 1976, day 211, hard dough.

REPRODUCIBILITY OF THE ORIGINAL PAGE IS POOR

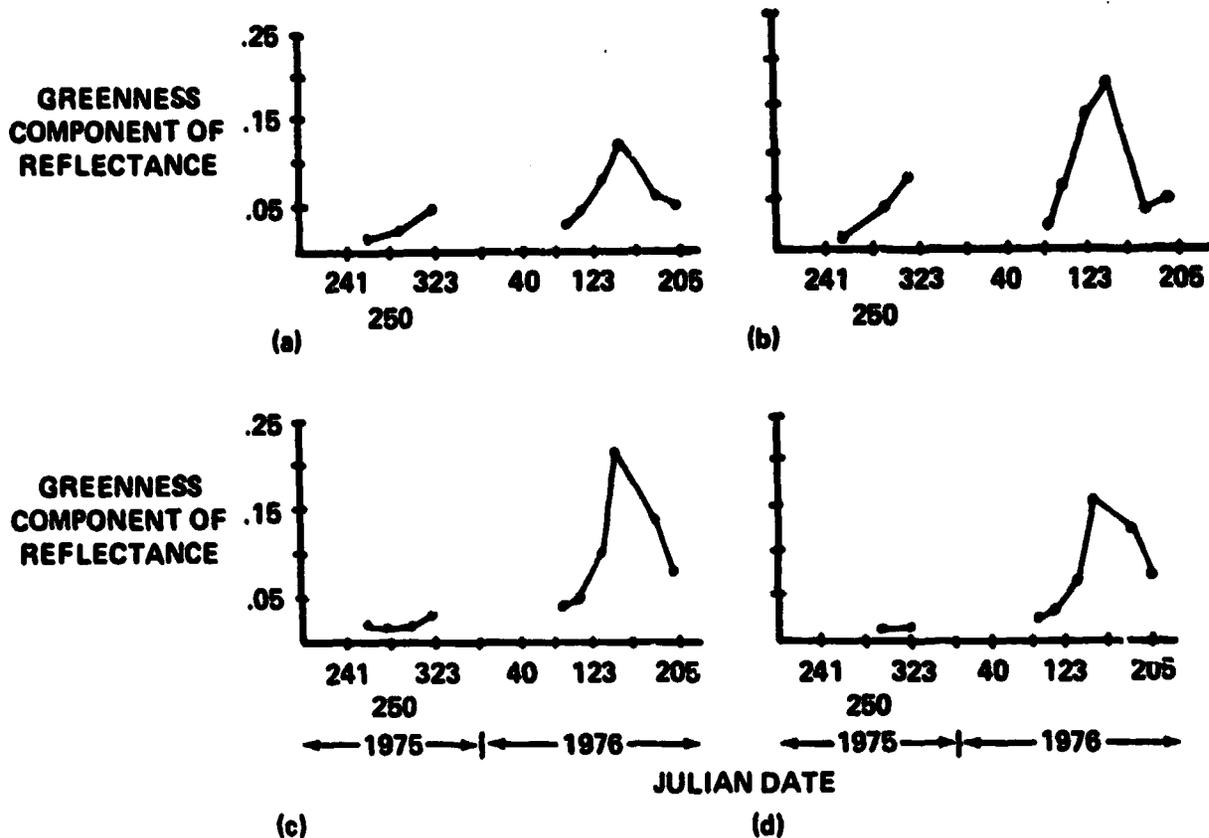


FIGURE 22.—Greenness component of reflectance of winter wheat fields as a function of date for different management practices. (a) Nonirrigated, normal planting date. (b) Irrigated, normal planting date. (c) Nonirrigated, late planting date. (d) Irrigated, late planting date.

86, and 62 for the eighth, seventh, and sixth missions, respectively. On the two next earliest acquisitions, fourth (November 11) and fifth (March 18), 28 and 38 percent of the fields would have been detected, respectively. For a threshold of 0.02, 90 percent of fields would have been detected at acquisitions 4 and 5. The nonwheat fields in this data set were also tested for the greenness threshold crossing, with good exclusion of them by the 0.06 and 0.04 thresholds. For example, for the threshold of 0.06, only one field exceeded the threshold on the eighth acquisition and none on earlier missions. As a point of reference, a root mean square error (RMSE) of 0.018 in greenness would correspond to a two-count uncorrelated RMSE noise level in each Landsat band.

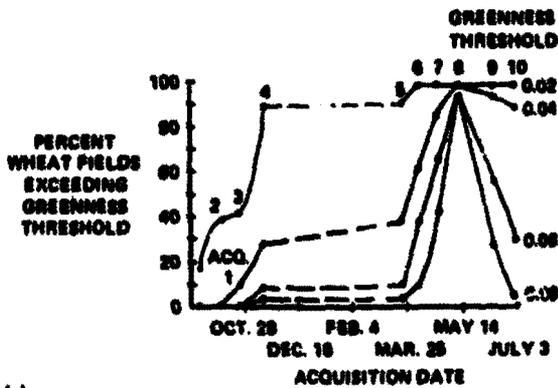
The relationship between the greenness component of measured reflectance and the observed percent soil cover for the wheat fields was also analyzed. A greenness reflectance threshold of 0.02 corre-

sponded to 18 to 25 percent soil cover, one of 0.04 to 30 to 35 percent, and one of 0.06 to 40 to 45 percent soil cover. These values need refinement because only coarse (20 percent) increments of soil cover were recorded for the fields analyzed.

SUMMARY OF KEY ACCOMPLISHMENTS AND RESULTS

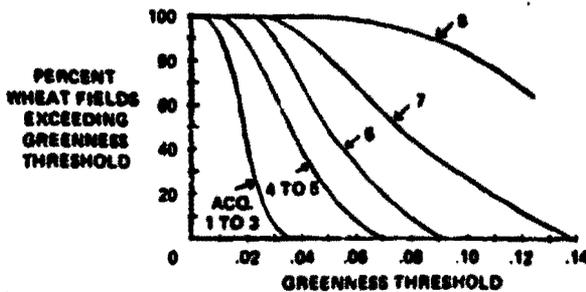
The LACIE Field Measurements Project successfully acquired a large amount of high-quality spectral measurements during 3 years at three test sites in Kansas, South Dakota, and North Dakota. Analyses of these data are providing new knowledge about the spectral properties of crops in relation to their agronomic characteristics.

Spectral measurements were made of controlled experimental plots of wheat and other small grains using truck-mounted spectrometers and of commer-



(a)

ACQ. NO.	DATE
1	SEPT. 18, 1976
2	OCT. 2, 1976
3	OCT. 21, 1976
4	NOV. 11, 1976
5	MAR. 18, 1976
6	MAR. 21, 1976
7	APR. 18, 1976
8	MAY 8, 1976
9	JUNE 12, 1976
10	JUNE 30, 1976



(b)

FIGURE 23.—Determination of the threshold of detection of winter wheat based on the greenness component of reflectance. (a) As a function of acquisition date. (b) As a function of greenness threshold.

cial fields of wheat and other crops by a helicopter-borne spectrometer, an airborne scanner, and the Landsat MSS. The spectral data are supported by extensive agronomic and meteorological measurements and observations. Together, the spectral, agronomic, and meteorological data form the most comprehensive data set now available for agricultural remote-sensing research. The data have been processed and provided to investigators who are now using them in research programs; they are available for use by other investigators.

The capability to acquire and analyze spectral measurements was significantly advanced during the LACIE Field Measurement Project. One of the important attributes of the LACIE field measurements spectral data is that they are radiometrically calibrated. Calibration enables valid comparisons of measurements from different dates, sensors, and/or locations. The procedures for field calibration of data have been defined and tested, and the knowledge gained will continue to be applied in future investigations.

The development of a computerized field research data base and an interactive graphics and statistics software system has significantly increased the capability to analyze and interpret interrelationships of the spectral and agronomic characteristics of crops and soils.

Another result of the LACIE field measurements experience is the definition of specifications of a standardized, flexible, and economical multispectral data acquisition system for field research. The instrument system would consist of a multiband radiometer, including the thematic mapper wavelength bands, and a data recording-handling-playback module. Development and use of these instrument systems will make it possible and economical to acquire and process calibrated spectral measurements from tripods, trucks, or helicopters over a wide variety of crops. This approach to spectral data collection was successfully tested by LARS in 1977.

Analysis of the LACIE field measurements data is providing new knowledge and understanding of the spectral characteristics of wheat and the biological-physical factors affecting spectral response. For example, strong relationships have been found between reflectance and percent soil cover, leaf area index, biomass, and plant water content. These are fundamental measures of crop vigor that can be used in crop growth and yield prediction models. In relating agronomic and spectral characteristics of wheat, it has been found that a middle-infrared wavelength band, 2.08 to 2.35 micrometers, is most important in explaining variation in biomass and plant water content, whereas a near-infrared band, 0.76 to 0.90 micrometer, accounts for the most variation in percent soil cover and leaf area index. In evaluating sensor characteristics, it has been determined that the reflective wavelength bands proposed for the thematic mapper are more strongly related to and better predictors of the canopy variables than the Landsat MSS bands. In other studies, insights for

development of discrimination techniques have been gained through analysis of the spectral differences between spring wheat and barley and the spectral development of wheat fields early in the growing season.

RECOMMENDATIONS FOR FUTURE FIELD RESEARCH

Although the LACIE Field Measurements Project acquired a large quantity of data, the sample of crop, soil, and weather conditions was small, even for wheat. Each of the 3 years in each site was different in terms of the weather and the crop response to it; however, the crop cannot be treated as a constant even if the weather does not vary significantly from year to year. Changing economic conditions and advancements in agricultural technology will bring changes in crop and soil management (e.g., minimum tillage) and even the crop itself (e.g., introduction of semidwarf varieties of wheat). Measurements of wheat and its confusion crops should be continued over additional years if the full potential of the current effort is to be achieved.

As one looks ahead to the development of a global food and fiber information system using remote-sensing techniques, it is critical to begin to make the field measurements required to understand the spectral characteristics of crops other than wheat, such as corn, soybeans, rice, cotton, and rangeland. One of the lessons that should come from the LACIE Field Measurements Project is the importance of conducting field research before the results are needed to design a large-scale effort. Because of the year-to-year variations in weather, several years of data are required.

The primary sensors used for LACIE field measurements were spectrometers capable of producing high-resolution spectra. In the future, a new approach to the collection of field measurements data will be needed because it will not be feasible simply to multiply the current approach by the increased number of crops and regions that should be included in future experiments. Multiband radiometer systems can economically provide the necessary spectral measurements. With these instruments, it will be possible to acquire measurements at more sites than is possible with the currently available high-spectral-resolution spectrometer systems. And, it is more observations of crops and soils under a wide variety of conditions (not detailed spectral measurements of a limited number of locations and

crop conditions) that are needed to increase our understanding of the spectral characteristics of agricultural scenes. There will be a continuing need for the high-resolution spectrometer systems to be used in field research, but less complex systems are also required that can be used to make large numbers of measurements at many sites economically and accurately.

The approach to data acquisition should include cooperative efforts with USDA, land-grant universities, and commercial test stations to make detailed crop, soil, and meteorological measurements in controlled plots, as well as less intensive observations of commercial fields in larger test sites.

In conclusion, field research is an essential component of the development of agricultural remote sensing. A sound field research program can provide the basis on which larger scale satellite experiments and operational systems are constructed. The overall objectives of future field research should be to obtain a quantitative understanding of the radiation characteristics of agricultural crops and their soil backgrounds and to assess the capability of current, planned, and future satellite sensor systems to capture available useful spectral information. Field research is a particularly important component of developing remote-sensing techniques for assessing crop condition and predicting crop yields.

ACKNOWLEDGMENTS

The assistance of the staff at LARS, particularly Larry Biehl, Marilyn Hixson, and Barrett Robinson, in the implementation of the LACIE Field Measurements Project and the preparation of this report is gratefully acknowledged. Dan Womack, Rudy Trabanino, and Michael Heidt of JSC made major contributions to the acquisition and processing of data.

REFERENCES

1. Bauer, M. E.; Robinson, B. F.; Biehl, L. L.; and Simmons, W. R.: LACIE Field Measurements Project Plan, 1976 and 1977. Purdue University Laboratory for Applications of Remote Sensing, West Lafayette, Ind., and NASA Johnson Space Center, Houston, Tex., 1976.
2. Zaitzeff, E. M.; Korb, C. L.; and Wilson, C. L.: MSDS: An Experimental 24-Channel Multispectral Scanner System. IEEE Transactions on Geoscience Electronics, GE-9, 1971, pp. 114-120.

3. Barnett, T.; and Juday, R.: Skylab S-191 Visible-Infrared Spectrometer. *Appl. Opt.*, vol. 16, 1977, pp. 967-972.
4. Leamer, R. W.; Myers, V. I.; and Silva, L. F.: A Spectroradiometer for Field Use. *Rev. Sci. Instrum.*, vol. 44, 1973, pp. 611-614.
5. White, W. P.: Field Signature Acquisition System (FSAS) Operation Procedures. LEC-9619, Lockheed Electronics Co., Houston, Tex., 1976.
6. Nicodemus, F. E.; Richmond, J. C.; Hsia, J. J.; Ginsburg, I. W.; and Limperis, T.: Geometrical Considerations and Nomenclature for Reflectance. Monograph 160, U.S. National Bureau of Standards, Washington, D.C., 1977.
7. Shi, C. N.; and Schutt, J. B.: Formulation Procedure and Spectral Data for a Highly Reflecting Coating. Paper X-762-71-266, NASA Goddard Space Flight Center, Greenbelt, Md., 1971.
8. Bauer, M. E.; Silva, L. F.; Hoffer, R. M.; and Baumgardner, M. F.: Agricultural Scene Understanding. LARS Contract Report 112677, Purdue University Laboratory for Applications of Remote Sensing, West Lafayette, Ind., 1977.
9. Plyer, E. K.; and Peters, C. W.: Wavelengths for Calibration of Prism Spectrometers. *J. Res. NBS*, vol. 45, 1950, p. 452.
10. USDA LACIE Project Staff. Ground Truth Reference Handbook. U.S. Department of Agriculture, Washington, D.C., 1975.
11. Biehl, L. L.: LACIE Field Measurements Data Library Catalogs. Purdue University Laboratory for Applications of Remote Sensing, West Lafayette, Ind., and NASA Johnson Space Center, Houston, Tex., 1977.
12. Phillips, T. L., ed.: LARSYS User's Manual. Purdue University Laboratory for Applications of Remote Sensing, West Lafayette, Ind., 1973.
13. Gates, D. M.; Keegan, H. J.; Schieter, J. C.; and Weidner, V. R.: Spectral Properties of Plants. *Appl. Opt.*, vol. 4, 1965, pp. 11-20.
14. Breece, H. T.; and Holmes, R. A.: Bidirectional Scattering Characteristics of Healthy Green Soybean and Corn Leaves *in vivo*. *Appl. Opt.*, vol. 10, 1971, pp. 119-127.
15. Gausman, H. W.; Allen, W. A.; and Wiegand, C. L.: Plant Factors Affecting Electromagnetic Radiation. Soil and Water Conservation Research Report 342, Agriculture Research Service, U.S. Dept. of Agriculture, Weslaco, Tex., 1972.
16. Sinclair, T. R.; Hoffer, R. M.; and Schreiber, M. M.: Reflectance and Internal Structure of Leaves From Several Crops During a Growing Season. *Agronomy J.*, vol. 63, 1971, pp. 864-868.
17. Landgrebe, D. A.; Biehl, L. L.; and Simmons, W. R.: An Empirical Study of Scanner System Parameters. *IEEE Transactions on Geoscience Electronics*, GE-15, 1977, pp. 120-130.
18. Harlan, J. C.; Rosenthal, W. D.; Clarke, S.; and Welch, J. C.: Analysis of Spectral Data for Wheat. Progress Report RSC 3183-7, Remote Sensing Center, Texas A & M University, College Station, Tex., 1978.
19. Vanderbilt, V. C.; Bauer, M. E.; and Silva, L. F.: Prediction of Solar Irradiance in a Wheat Canopy Using a Laser Technique. *Agricultural Meteorology* (To be published).
20. Berry, J. K.; and Smith, J. A.: Extracting Scene Features Through Modeling. Final Report, NASA contract NAS9-14467, Department of Earth Resources, Colorado State University, Fort Collins, Colo., 1976.
21. Kauth, R. J.; and Thomas, G. S.: The Tasseled Cap—A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat. Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, Purdue University, West Lafayette, Ind., 1976, pp. 4B: 41-51.
22. Mallis, W. A.; and Gleason, J. M.: Investigations of Spectral Separability of Small Grains, Early Season Wheat Detection, and Multicrop Inventory Planning. ERIM Report 122700-34-F, Environmental Research Institute of Michigan, Ann Arbor, Mich., 1977.
23. Lambeck, P. F.: Signature Extension Preprocessing for Landsat MSS Data. ERIM Report 122700-32-F, Environmental Research Institute of Michigan, Ann Arbor, Mich., 1977.
24. Mallis, W. A.; Ciccon, R.; and Gleason, J. M.: Wheat Signature Modeling and Analysis for Improved Training Statistics. ERIM Report 109600-66-F, Environmental Research Institute of Michigan, Ann Arbor, Mich., 1976.
25. Ahiricha, J. S.; Bauer, M. E.; Daughtry, C. S. T.; and Hixson, M. M.: Relation of Agronomic and Spectral Characteristics of Spring Wheat Canopies. Technical Report 080178, Purdue University Laboratory for Applications of Remote Sensing, West Lafayette, Ind., 1978.
26. Colwell, J. E.: Vegetation Canopy Reflectance. *Remote Sensing of Environment*, vol. 3, 1974, pp. 175-183.
27. Tucker, C. J.: Spectral Estimation of Grass Canopy Variables. *Remote Sensing of Environment*, vol. 6, 1977, pp. 11-16.
28. Leamer, R. W.; Noriega, J. R.; and Wiegand, C. L.: Seasonal Changes in Reflectance of Two Wheat Cultivars. *Agronomy J.*, vol. 70, 1978, pp. 113-118.
29. Tucker, C. J.; and Maxwell, E. L.: Sensor Design for Monitoring Vegetation Canopies. *Photogrammetric Engineering and Remote Sensing*, vol. 42, 1976, pp. 1399-1410.
30. Hixson, M. M.; Bauer, M. E.; and Biehl, L. L.: Crop Spectra From LACIE Field Measurements. LACIE Report 00469, Purdue University Laboratory for Applications of Remote Sensing, West Lafayette, Ind., and NASA Johnson Space Center, Houston, Tex., 1978.

omit

The USDA Application Test System

FOREWORD

The U.S. Department of Agriculture (USDA) is aware of the potential for using satellite remote-sensing techniques to support present and future USDA information requirements. The ability of key U.S. decisionmakers, both in government and in private industry, to accurately assess the production potential of major world crops in a timely manner, as well as to assess world market potential, has been economically rewarding to U.S. foreign trade. Commodity experts of the Foreign Agricultural Service (FAS) of the USDA have the expertise to accurately analyze and evaluate foreign crop data but often do not have as timely and/or as complete data for the formulation of their crop production forecasts. Inaccurate and/or untimely crop information can be costly to U.S. foreign trade, to the American farmer, and to the consumer. Crop production potentials can change very quickly because of the vulnerability of crops to the effects of weather and other natural phenomena, such as disease and insect infestation. For example, the impact of an acute event such as an overnight freeze can quickly change the production potential of a given crop and thus alter existing world market conditions. Understandably, more timely and more accurate foreign crop condition information can be of great benefit to U.S. foreign trade.

The significance of the LACIE is that it demonstrates that current Earth resources and meteorological satellite data offer commodity experts and decisionmakers information that can potentially improve the timeliness and accuracies of foreign crop production estimates. The users will determine the cost-effective applications of remote-sensing technology. The Application Test System (ATS) of the USDA was developed as a part of the LACIE and will be one of the vehicles used to transfer remote-sensing technology in the future. Currently, the ATS is testing and evaluating the latest satellite and computer processing and analysis techniques in terms of their future application potential by the USDA. USDA management will review ATS tests and evaluations of candidate techniques prior to making

a decision on their transfer to operational elements. The ATS as part of the USDA will be responsive to changing and expanding user requirements. This year, for example, there were further clarifications of USDA requirements with the issuance of the Secretary of Agriculture's "Initiative for Aerospace Technology," calling for improved information on the "early warning of changes affecting production and quality of renewable resources." As a result of the initiative, the ATS is beginning to test and evaluate present satellite and computer processing and analysis techniques as tools for timely assessment of crop conditions in foreign countries. The ATS is evaluating candidate techniques developed by LACIE as well as techniques developed by the general research community and by private industry.

The purpose of this session is to describe the experience in technology transfer between the LACIE and the ATS; the technical and functional designs of the ATS; the ATS central data base concept and design; and the analysis component of the ATS. The following paragraphs present a brief description of each of the six papers presented in this session.

"The Application Test System: An Approach to Technology Transfer" presents the approach, the achievements, and the shortcomings of the experience in technology transfer between the LACIE and the ATS.

"Functional Definition and Design of a USDA System" describes the design of a USDA prototype system that has many of the same characteristics as the LACIE system. This prototype system has not been implemented, but it is available if USDA management decides to use it.

"The Application Test System: Technical Approach and System Design" describes the requirements for and eventual design of a computer system for large-scale processing of Landsat data. The computer system is composed of modular off-the-shelf components of limited specialization and can readily accommodate state-of-the-art changes in hardware and software technology.

"Data Base Design for a Worldwide Multicrop Information System" addresses the design of the

central data base supporting the ATS. The data base will support multicrop and multicountry information requirements identified by the end users, as well as the everyday functional and analytical needs of the ATS crop and image analysts, management, and system development teams.

"The Application Test System: Experiences to Date and Future Plans" details the data analysis component of the ATS, describing both short- and long-term analysis objectives. The ATS crop analyst uses a state-of-the-art interactive image processing system for the analysis of Landsat multispectral scanner (MSS) data. The analyst has available a central data base that contains valuable data records, such as historical and current Landsat MSS, meteorological, and crop statistics data, based on a

unique 25- by 25-nautical-mile grid system.

"Resource Modeling: A Reality for Program Cost Analysis" describes a tool developed for management of the ATS. Given varying requirements, the cost model can quickly assess, allocate, and manage ATS resources. The model also provides budget projections and comparisons and personnel staffing reports.

The ATS will continue to test and evaluate satellite and computer processing and analysis techniques in terms of their applicability to USDA information needs. The ATS state-of-the-art modular design can readily accommodate changes and therefore can be easily modified or augmented to support future USDA requirements.

The Application Test System: An Approach to Technology Transfer

A. C. Aaronson,^a K. Bitlow,^a F. C. David,^a R. L. Packard,^a and F. W. Raves^b

INTRODUCTION

The Application Test System (ATS) of the U.S. Department of Agriculture (USDA) was implemented to test and evaluate the latest satellite and computer processing and analysis technologies in terms of their application feasibility by the USDA. Technologies to be evaluated include those developed, tested, and evaluated by the LACIE, as well as candidate technologies developed by the research community and private industry. This paper presents some background leading to the implementation of the ATS and discusses the technology transfer experience between the LACIE and the ATS, highlighting the approach, the achievements, and the shortcomings.

CONCEPT AND APPROACH

Technology transfer is a term most often used in the scientific community to define the movement of technical capabilities from a research and development (R&D) environment to a user-oriented group for application in an operational program. Although the basic term is simple in definition, the actual transfer of technology is not a simple, straightforward process. A major problem area is the lack of effective interaction between the R&D community and the end user. Those elements of the technology which must be evaluated by the end user and considered by the R&D community before the transfer is consummated include

1. Technology applicability to user needs
2. Technology cost/benefit trade-offs
3. Personnel training in the use of the technology

4. Impact of changes in technology

The sections that follow discuss some of the LACIE and ATS experiences for each of the preceding user considerations in achieving effective transfer of technology. It is the belief of USDA management that the ATS can be a significant vehicle for transferring satellite remote-sensing technology to the USDA.

Technology Applicability to User Needs

The technologies developed by the LACIE were designed to support an experiment objective of providing end-of-season wheat-production estimates that were within 10 percent of true production 9 of 10 years. When this performance criterion was documented, explicit USDA aerospace and remote-sensing information requirements were unknown. The basic premise throughout LACIE was that a country-level production estimate was an absolutely essential ingredient.

Prior to the LACIE, the USDA had been formulating departmental requirements which could be satisfied by using remotely sensed data. The work of the USDA Remote Sensing User Requirements Task Force (RSURTF) solidified a basic list of departmental remote-sensing requirements in December 1975 after the start of the LACIE. Without a set of specific user requirements, the LACIE established the requirement to inventory wheat production in a number of LACIE-selected countries. The LACIE goals were later modified to accommodate a number of user needs published in the list of RSURTF requirements. The timeliness and accuracy criteria for the LACIE wheat production reports were impacted by this list of user requirements.

Ideally, a well-defined set of user requirements to guide the establishment of project objectives should have existed before LACIE was begun (see the plenary paper by Murphy et al. entitled "Technology

^aU.S. Department of Agriculture, Houston, Texas.

^bNASA Johnson Space Center, Houston, Texas.

Transfer: Concepts, User Requirements, and a Practical Application" for more detail). The LACIE could then have been in a position to test and evaluate system components that were more directly responsive to user requirements. Nevertheless, it was possible for LACIE to essentially respond to a number of user requirements during the 1975-77 time frame.

Recently, the USDA has modified and enlarged their remote-sensing requirements. This is reflected by the Secretary of Agriculture's "Initiative for Aerospace Technology" released in April 1978. The Secretary's Initiative, based on close cooperation with NASA, the Agency for International Development, and the Departments of Interior and Commerce, includes a priority listing of USDA's information requirements that remote-sensing technology could support. The seven information requirements are

1. Early warning of changes affecting production and quality of renewable resources
2. Commodity production forecasts
3. Renewable resources inventory and assessment
4. Land use classification and measurement
5. Land productivity estimates
6. Conservation practices assessment
7. Pollution detection and impact evaluation

These new priorities documented by the user will serve as primary guidelines to future programs using remotely sensed data. In the near future, LACIE technology will be expanded from a single-crop application to a multicrop application. The LACIE techniques and planned follow-on applications essentially address the Secretary of Agriculture's information requirements for commodity production forecasts. In response to the Secretary's Initiative, the ATS is expanding its original design and implementation plans of inventorying wheat production like the LACIE to test a crop condition assessment system that will detect and assess the impact of abnormal events on agricultural production (see the paper by May et al. entitled "The Application Test System: Experiences to Date and Future Plans" for more detail). The ATS will also continue to have the technical and analytical components required to implement a commodity production inventorying system developed and successfully demonstrated in the LACIE (see the paper by Evans et al. entitled "Functional Definition and Design of a USDA System" for more detail). This capability will be tested and evaluated in the ATS during the next two crop seasons.

Technology Cost/Benefit Trade-Offs

The ATS, together with USDA management in Washington, D.C., will assess the immediate and future benefits of the information produced by remote-sensing technology. Of equal importance is an assessment of the cost-effectiveness of the technology (see the paper by Fouts and Hurst entitled "Resource Modeling: A Reality for Program Cost Analysis" for more detail).

Presently, ATS emphasis is on testing and evaluation of information produced by the crop condition assessment system. The ATS will perform tests of the system over selective agricultural areas of the world in order to evaluate ATS output products and system cost. The ultimate decision to transfer selective techniques used by the crop condition assessment system (such as Green Index Number interpretations and yield model estimates) to the end user will be made by USDA management in Washington, D.C. The USDA management will review ATS technical tests and evaluations, as well as cost evaluations, to assist in their decision on the transfer of the selective techniques to operational elements.

The flexibility of the ATS computer hardware design and central data base system makes the ATS cost-effective as well as easily adaptable to applications testing. The modular hardware design is composed of minicomputers, high-density disk drives, graphics terminals, interactive image analysis stations, and other supportive equipment (see the paper by Benson et al. entitled "The Application Test System: Technical Approach and System Design" for more detail). The central data base is geographically oriented and will store historical and current Landsat meteorological and collateral data (see the paper by Driggers et al. entitled "Data Base Design for a Worldwide Multicrop Information System" for more detail). The data base will support a wide spectrum of application needs identified by the image and crop analysts as well as the everyday functional needs of ATS management and system development teams.

Personnel Training in the Use of the Technology

The training of human resources in the use and application of a given technology is as much a part of the technology transfer process as the transfer of concepts, algorithms, and procedures. Comprehen-

sive tests and evaluations of candidate technologies could not be conducted without the technical knowledge needed to understand and implement a given technology. Ultimately, personnel have to put the new procedures and techniques to use. All users of the technology must have the opportunity to become thoroughly acquainted with the new technology before its adoption.

Because of this USDA management philosophy, USDA personnel were assigned to operational project elements during the LACIE to gain a better understanding of remote-sensing technology. Personnel assigned to the Crop Assessment Subsystem (CAS) of the LACIE prepared monthly, unscheduled, and end-of-year crop production reports for each of the active LACIE countries. In addition, USDA personnel gained experience in the acquisition, storage, and retrieval of Landsat multispectral scanner (MSS) data as well as in the interpretation and analysis of the spectral data. A group of USDA analysts participated in the testing and evaluation of an interactive image processing system for sample segment wheat area determination. These analysts became familiar with the latest techniques in MSS data analysis.

The USDA personnel were also actively involved in testing and evaluating the development of crop yield models, crop calendars, and other crop-condition-related programs. Later in the project, USDA crop analysts helped to initiate a LACIE program that used vegetative indexes (transformed four-channel MSS data) to monitor moisture stress and crop condition.

All these experiences in the management, analysis, and reporting components of the LACIE during the period 1974-78 familiarized a core of USDA personnel with the latest satellite and computer processing and analysis technologies.

Impact of Changes in Technology

The ATS will be minimally affected by expected changes in remote-sensing technology due to the implementation of a state-of-the-art computer system that is designed to readily accommodate change. The system relies on off-the-shelf components of limited specialization and is capable of responding to state-of-the-art developments in hardware and software technology. The ATS, when directed by USDA management, can augment the present system configuration with additional minicomputers, or if technological advancements or

additional needs warrant, it can replace present image analysis hardware with new and improved equipment that may be developed in the future.

IMPLEMENTATION APPROACH

User Systems Planning and Applications Test Group

The first organized USDA effort to effect an application test system was in 1976 when the User Systems Planning and Applications Test Group (USPATG) was organized. In this first step, 10 automatic data processing (ADP) experts under the management of a USDA user were dedicated to developing a system to meet USDA remote-sensing goals. They were given the responsibility of designing and implementing a system capable of testing and evaluating LACIE technology with respect to USDA requirements. Initially, the USPATG was primarily composed of ADP personnel. Later, USDA crop analysts formally trained in the use of interactive image processing systems in the LACIE were added to the USPATG. NASA personnel have also been assigned to the USPATG to help facilitate the transfer of LACIE technology.

Initially, the ATS implementation approach called for (1) ATS personnel to establish ATS functional specifications and to be responsible for the practical assessment of technologies and (2) ATS personnel, augmented by contractor support, to be responsible for the detailed design and technical implementation of transferred technologies; ATS personnel will continue to perform these functions, augmented by contracts when required. In all cases, the ATS is responsible for system operation and the preliminary evaluation of output products. Final evaluation of output products will be made by USDA management in Washington, D.C.

Mechanisms for Technology Transfer

The primary mechanisms for the transfer of LACIE technologies were the ATS written requests for proposals (RFP's), the preliminary and critical design reviews (PDR and CDR), the ATS Design Review, the Classification Procedures Advisory Team (CPAT), and the USDA LACIE personnel. During the ATS development period (fig. 1), RFP's

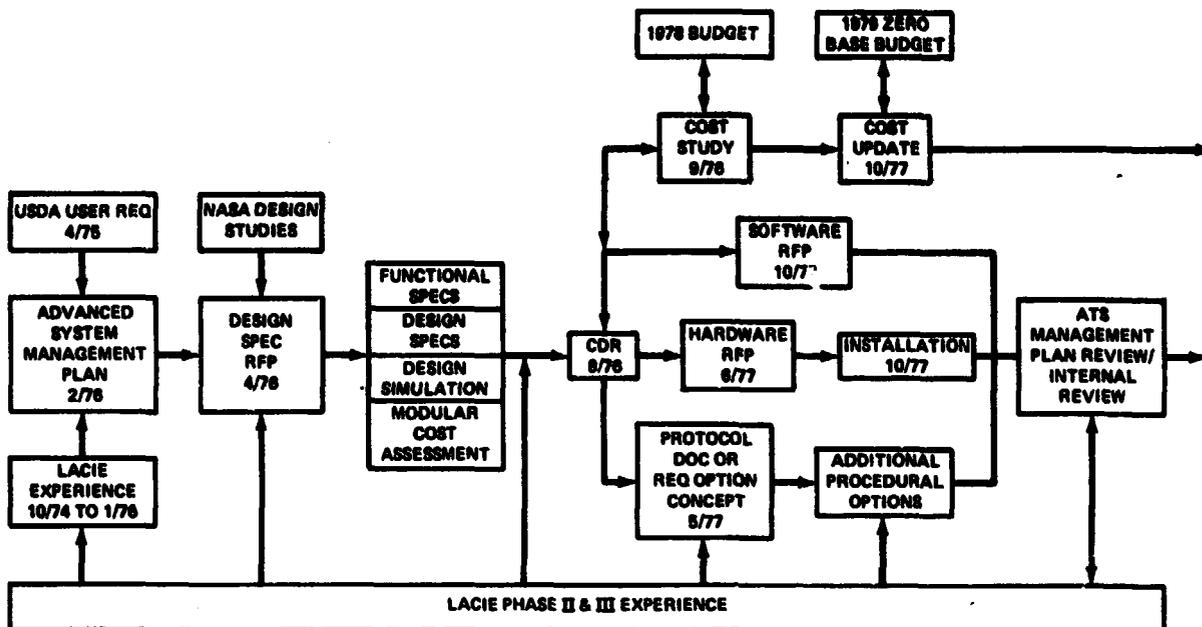


FIGURE 1.— ATS development history.

were written to establish ATS design specifications and the software/hardware composition of the ATS. The specifications included in the RFP's were an outgrowth of the technologies developed by the LACIE, and technical project personnel reviewed the RFP's for consistency with state-of-the-art technology. Open CDR's were held by the ATS contractors to present and review ATS/contractor design specifications. A mechanism was established within the CDR that allowed non-ATS personnel (other LACIE personnel) to submit discrepancy reports on a particular ATS design aspect.

In October 1977, all aspects of the ATS were reviewed by LACIE personnel. The ATS design review was held over a 2-day period. During the first day, ATS personnel presented details of present and planned system capabilities. During the second day, review participants divided into small groups to review particular technical and analytical procedures to be implemented by the ATS, including analyst procedures, yield models, accuracy assessment procedures, data base design, crop condition assessment approaches, early-season estimation approaches, and the LACIE/ATS interface. The resultant review reports from each of the groups were used to redesign many aspects of the ATS and are now part of the current implementation plan.

The CPAT developed early in LACIE was instrumental in facilitating the transfer of LACIE sample segment classification technology to the ATS. The CPAT was composed of personnel from NASA, USDA, and the Lockheed Electronics Company (LEC).

Meetings were held between CPAT and ATS personnel starting in March 1977 to review the LACIE classification procedures and to determine the appropriate design specifications. The knowledge gained by the ATS staff was later used to draft an RFP to acquire these same capabilities. Personnel representing the ATS and the Ford Aerospace & Communications Corporation (FACC) jointly designed and implemented a computer system capable of supporting the testing and evaluation of LACIE classification algorithms, as well as other analytical techniques and procedures. The LEC has been contracted to augment the ATS classification technology initially delivered by FACC to include Procedure 1 (P-1), a key procedure developed in LACIE.

The USDA LACIE personnel were instrumental in transferring the knowledge needed to implement many of the LACIE techniques. As stated previously, USDA personnel were an everyday working part of the LACIE and, through their exposure to

daily operations, became familiar with many LACIE techniques. The successful implementation of many of the LACIE techniques by the ATS can be credited to the ATS technical staff trained in the LACIE.

TECHNOLOGY TRANSFERRED

The ATS has transferred an assortment of technical and analytical capabilities to support the implementation, testing, and evaluation of a crop condition assessment system and a production inventorying system. Limited ATS developmental work has also been done. Those technologies transferred from the LACIE and presently undergoing further large-area testing and evaluation are itemized and summarized in the following paragraphs.

Sample Segment Classification Algorithms

During the LACIE, a number of sample segment classification procedures designed to produce a sample segment wheat area estimate were developed, tested, and evaluated over varying agricultural areas and conditions. The ATS has adopted the following classification procedures from the LACIE: (1) the Analyst-Selected Training Fields procedure, (2) the Designated Crop procedure, and (3) Procedure 1 or the Preselected Training Fields procedure. The ATS philosophy is to utilize all three procedures as analyst options for classifying a sample segment. Each procedure has advantages over the others under certain sample segment conditions. For example, P-1 worked fairly well in areas having small, randomly distributed fields and heterogeneous signatures, but it was less optimal (more time-consuming) in agricultural areas having relatively large fields and homogeneous spectral signatures. In the latter case, the Designated Crop procedure would have been a more optimal analyst procedure to implement.

The ATS acquired the Integrated Multivariate Data Analysis and Classification System (IMDACS) from the FACC in part to implement the Designated Crop and the Analyst-Selected Training Fields procedures. Although IMDACS was not designed for the ATS but rather is off-the-shelf software, ATS analysts have implemented the previously mentioned procedures from the various IMDACS capabilities. The LEC, under a separate contract, has expanded processing options by augmenting IMDACS with the P-1.

Sampling Strategy and Production Aggregation Software

The LACIE implemented a sampling strategy designed to provide end-of-season wheat production estimates that were within 90 percent of true production 9 of 10 years. The LACIE sampling strategy has developed from one which allocated sample segments by political subdivision to the latest strategy that allocates sample segments to relatively homogeneous agricultural areas called agrophysical units or APU's. The latest sampling strategy design was implemented to gain sampling efficiency; i.e., to reduce the number of segments required to achieve an end-of-season 90/90 accuracy goal.

The ATS will perform a large-area test of the LACIE-developed sampling strategies and production aggregation algorithms. The ATS will test and evaluate the APU and political subdivision approaches for sample segment allocation and production aggregation in large areas of the United States and the U.S.S.R. during 1978-79. Although the ATS has not installed the LACIE-developed production aggregation software, the ATS will utilize the algorithms available at the LACIE. Future installation of these algorithms on ATS equipment is presently being considered.

Currently, the ATS is evaluating the sample design and resulting allocation to support a crop condition assessment system. The sample segments will be assessed for crop condition by USDA analysts using various MSS data transforms.

Yield Models

During the LACIE, operational yield estimates were derived from the Center for Climatic and Environmental Assessment (CCEA) regional yield models. The LACIE Research, Test, and Evaluation Group also evaluated the Kansas State University (KSU) yield model for future applications. The ATS will further test and evaluate these models over large areas in the United States and the U.S.S.R.

Crop Calendar Models

The LACIE utilized the CCEA Crop Calendar Model to determine the timing of wheat growth stages. The crop calendar information was used by LACIE crop analysts in the spectral analysis of the

MSS data. Understanding the relationship between the MSS data and the wheat crop calendar made the analyst job of MSS picture element (pixel) labeling easier. The ATS obtained crop calendar results from the LACIE for sample segments in the United States and the U.S.S.R. during the 1978 crop year.

Additionally, the ATS will test and evaluate the KSU crop calendar model, a subroutine of the KSU yield model. The KSU and CCEA crop calendar results will be jointly tested and evaluated against ground-truthed crop calendar information obtained from sample sites in the United States.

Vegetative Index Approach to Crop Condition Assessment

The ATS will further test and evaluate the use of vegetative indexes, such as the Kauth-Thomas green number, for crop condition assessment. A vegetative index number is transformed from the raw MSS digital data and is used for crop vigor assessment. The LACIE used the vegetative index approach to monitor crop condition and soil moisture conditions in a large test region in the United States. These vegetative index numbers and their interpretations will support the operation and evaluation of the ATS crop condition assessment system.

The paper by May et al. entitled "The Application Test System: Experiences to Date and Future Plans" discusses the implementation aspects of each of the LACIE-developed techniques transferred to the ATS.

EVALUATION OF THE TECHNOLOGY TRANSFER EXPERIENCE

A surface-level assessment of the technology transfer process between the LACIE and the ATS indicates a successful transfer. However, a more tenuous assessment of the technology transfer process indicates that the process was extremely difficult, primarily due to the absence of an established mechanism within the LACIE to facilitate the transfer of the technology short of a "turnkey" approach. This approach does not conform to the changing needs of the user. The ATS approach to technology transfer is to test and evaluate technologies the ATS assessed to support USDA information requirements when these requirements became known.

Initially, the ATS design effort was to support the further large-area testing of a wheat production inventorying system similar to the LACIE. With broadening direction from USDA management in support of the Secretary of Agriculture's Initiative for Aerospace Technology, the system has to be capable of responding to a number of information requirements, including the early warning of changes affecting production and quality of renewable resources. Since the ATS approach called for a system design that was flexible, it has been relatively easy to adapt to changing USDA information requirements.

The LACIE/ATS technology transfer experience clearly identified the need to define specific end-user requirements before the design implementation and testing of new techniques and analysis capabilities and/or to provide a mechanism within a project to incorporate changing or modified user requirements (i.e., technology development must be responsive to end-user requirements).

CONCLUSIONS

The LACIE/ATS technology transfer experience is not an example of optimal technology transfer design. Certain aspects of this experience were extremely beneficial. First, the ATS designed, implemented, and tested a computer system capable of supporting the testing and evaluation of LACIE technologies as well as technologies transferred from the research community and private industry. Second, the ATS incorporated many of the LACIE techniques and analytical procedures into its operations, such as classification algorithms, sampling strategy, yield models, and crop calendar models. Third, the ATS is staffed by personnel trained in the use of LACIE techniques and procedures.

The unavailability of specific user requirements before the start of the LACIE complicated the technology transfer experience between the LACIE and the ATS. As a result, the ATS is testing and evaluating those LACIE and non-LACIE techniques and procedures that could support USDA information requirements. In this regard, the ATS implementation of a flexible system design adaptable to changing user requirements is proving to be a cost-effective decision.

Functional Definition and Design of a USDA System

S. M. Evans,^a E. R. Dario,^a and G. L. Dickinson^a

INTRODUCTION

During the initial phases of LACIE development, it was the general intent of the U.S. Department of Agriculture (USDA) to exploit the knowledge derived from the LACIE and to incorporate the verified technology into an operational system. Thus, as the LACIE evolved, the concept of testing the technology in a near-operational environment for transfer to a user system also evolved.

This paper discusses the functional definition and design of a USDA system utilizing the LACIE technology available as of June 1976. The organization and methods described herein are focused on LACIE technology in terms of its transfer for user applications. They are conceptual only and do not necessarily reflect the system that is being implemented on behalf of the USDA.

Constraints

In the design and definition, it is intended that the system be responsive to USDA user requirements and utilize the most cost-effective technology developed and tested during the LACIE. This guideline, as stated in the Management Plan for the User Advanced System Design (ref. 1), necessitated that constraints be placed on the formulation of a design.

The available manpower to operate a USDA system was determined to be approximately 60 persons. To effectuate a system utilizing this number and limited equipment resources, a 5-day, 2-shift operation was provided in the design. The varied agricultural disciplines involved and the probable lack of analysts' familiarity with automatic data processing (ADP) techniques were factors in the decision to use menu-driven software in the user system, where feasible.

Evolving technology in processing remotely sensed data, the need to upgrade equipment, and changes in data sources and user requirements required that the system be flexible. Security precautions for safeguarding crop estimate data were carried through each section of the design. The USDA guidelines for generating, storing, and transporting sensitive data were used.

A 7-day processing time from receipt of multispectral scanner (MSS) data through report generation was levied as a requirement for system throughput. This necessitated that the design provide the capability, on the average, to process a 117-by-196-pixel segment in 1.5 hours.

An additional design constraint provided that Landsat MSS data be utilized with flexibility for future sensor systems. In accordance with the design, nonspectral data would be formatted into a gridded system using a 25-nautical-mile grid with each grid divided into quadrants.

The basic concept for a flexible system was that the total system be operated through the data base. The system also would be kept modular so that changes in algorithms, new data sources, and additional hardware and software might be readily integrated to enhance or replace established components. Also, the use of standard off-the-shelf hardware would provide for relatively easy upgrading.

All system software developed locally would be written in the COBOL and/or FORTRAN languages, in accordance with USDA standards. Exceptions to this rule would be on a case-by-case basis with sufficient justification to support them.

An analyst "team" approach was mandatory in order to ensure availability of regional agricultural, ADP, and meteorological expertise for the processing of Landsat data. The team would be constructed with each of the above-mentioned disciplines represented, according to their approximate respective proportions of usage.

All software for analyst use would be tutorial with emphasis on relieving the analyst of the responsibility of knowing system commands. This required

^aU.S. Department of Agriculture, Houston, Texas.

the use of menus for appropriate process selection and supplemented error condition information, with probable solution steps displayed interactively. These constraints are described in greater detail in terms of functional requirements in a study (ref. 2) done by the MITRE Corporation prior to the USDA design effort.

Objectives and Goals

The USDA goals that were stated in the LACIE Memorandum of Understanding dated September 1, 1974, and the USDA decision to commit resources to this agreement were evolutionary in nature. Simply stated, these goals were

1. To participate in the development of a wheat estimation system through the exploitation of data collected by Earth-orbital satellite systems or by other systems operated by NASA and the National Oceanic and Atmospheric Administration (NOAA) of the U.S. Department of Commerce
2. To validate and assist in optimizing the LACIE technology developed by NASA and NOAA
3. To train multidisciplined USDA analysts in LACIE techniques and related technology
4. To transfer elements of the optimized technology to a USDA operations environment based on proven cost effectiveness
5. To apply experience gained in the LACIE to assess the potential of other feasible projects identified by the USDA Remote Sensing User's Requirement Task Force

While the User Advanced System Design met all the objectives stated above, a set of more detailed goals was established to guide the technical definition and design. These objectives provided direction in terms of system configuration, system reliability, accuracy of results, and methodology employed. They are, in descending order of importance, timeliness, accuracy, objectivity, and continuity.

Microaccuracy was not as important as timeliness of information. This was based upon USDA management's decision that a crop estimation system must be able to deliver a wheat production estimate by late March or early April (ref. 3, sec. 3.0).

Accuracy was not dismissed as unimportant but was, however, treated in a practical manner. At that point in time, the LACIE accuracy criterion was 90/90 (that the LACIE U.S. Great Plains at-harvest estimate be within 10 percent of the true value, with a probability of at least 0.9). It was recognized that

drastic system design modification and reconfiguration might not be predicated on an 89-percent accuracy level for a given crop year.

It was recognized, also, that the design and methodology utilized should not negate the agricultural, economic, geographic, statistical, and other expertise contributed by a USDA analyst within the system. On the other hand, procedures were intended to ensure some continuity in the way data were processed so that subjective input would not tend to distort the end results.

ORGANIZATION

The following sections briefly describe each of the organizational elements in figure 1. Definitions are primarily concerned with basic functions and types of personnel within each element.

Project Management

Project management is a policymaking administrative role with ultimate responsibility for the entire system. The system is complex and sophisticated, requiring specialized and experienced personnel to perform the daily tasks. Top-level management interfaces with the system through departmental executive management.

Technical Staff

The technical staff is administratively controlled by project management and provides management a pool of resources to be utilized as required. The technical staff represents a specific skill mix; i.e., computer specialists, systems engineers, budget analysts, economists, systems analysts, soil scientists, agronomists, meteorologists, and remote-sensing scientists.

Members of this group are responsible for developing and testing applications and special-purpose software and performing systems maintenance. Support to perform these functions is required by the data base management group.

Specialists on the technical staff are responsible for analysis activity. A regionally oriented team concept is planned. Since circumstances and politics could cause sudden and major emphases shifts, the organization is loosely structured to permit

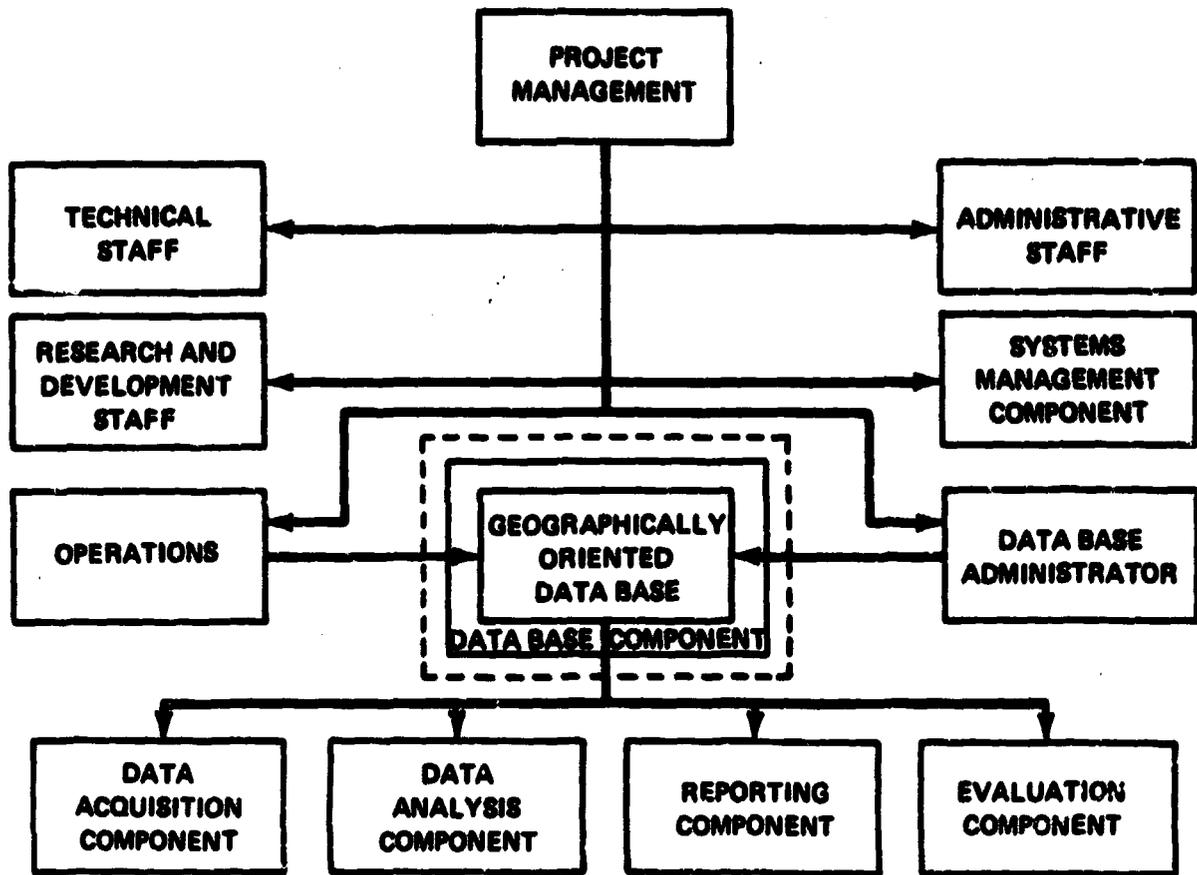


FIGURE 1.—USDA system organizational elements.

specialists in the various disciplines to contribute as priorities demand.

The technical staff supports the research, operations, and systems management components as resources and priorities allow.

Research and Development (R&D) Staff

Senior scientists from all disciplines are required to perform R&D tasks and functions.

Requests for research in a given area may be initiated by any User Advanced System (UAS) staff member, but project management would be the approving authority for the research. Additionally, the R&D staff supports production operations by special studies of episodic events or day-to-day processing, if required. The R&D staff is controlled administratively by project management.

Administrative Staff

The two primary functions of the administrative staff are personnel-related services and management assistance. The personnel functions include such responsibilities as payroll, insurance, and general recordkeeping. The management assistance functions include such responsibilities as budget preparation, contract services, facilities maintenance, and purchasing. The administrative staff reports directly to the Project Manager.

Systems Management Component

The primary function of the systems management component is to serve as a coordinating and integrating unit for advanced system responses to requirements. Data requirements are translated into specific

acquisition, analysis, and processing activities. Responsibilities of the systems management component include the following.

1. Tracking the status of responses to specific requirements through the use of a change control board or panel.

2. Assessing the impact and potential value of implementing new requirements or techniques into the operational system. Through discussions and studies, it determines the additional hardware required; computer software support available or required to be written; data available to meet processing requirements; and impacts on daily operations, on analysis activities, and on the data base. Based on these determinations, a management decision is made to implement or not to implement the proposed requirement or technique. If approved, the systems management component coordinates the implementation and ensures the integration of a thoroughly tested module.

3. Serving as the communications link to the environment with which the operational system must interface. Any changes in legislative policy effecting changes in USDA operations are coordinated by this group.

4. Providing special studies or reports requested by organizations other than those considered part of the production system and the USDA. For example, a request by a member of Congress or a commercial organization could be coordinated with the USDA's Congressional Liaison Office and appropriate marketing and public relations offices as required. This request would then be assessed for impact and tracked through the system until completion of implementation or the decision to not implement.

Operations Section

The operations section is responsible for the following.

1. Scheduling and controlling all day-to-day activities of the production system requiring the use of the analysis stations and their associated general-purpose computers (or special-purpose hardware), or other related services such as keypunching

2. Hardware maintenance, computer system operations and monitoring, and tape and disk pack library establishment and maintenance

The operations staff is administratively controlled by project management. The section has a shift leader and an aid on each shift to review system per-

formance data and ensure that corrective actions are taken.

Data Base Administrator

The responsibilities of the data base administrator include the following.

1. Day-to-day technical control of the production system data base, a large, integrated, complex structure serving users with a wide variety of data and data processing requirements

2. Control of the logical and physical data base structuring, assuring the security and integrity of the data base (including recovery mechanisms), and granting access to the users

3. Control of the purging and subsequent releasing of space for any data item to be removed from the system

4. Assessing the impact of user requests on the entire community and making decisions as to which capabilities are most practical or critical to be implemented, based on management-assigned priorities, implementation costs, resource availability, and other considerations

The data base administrator reports administratively to project management.

COMPONENTS

The USDA requires a closed-loop information system (fig. 2). This loop indicates the use of MSS and meteorological data to perform the identification and mensuration of crop type and condition. Generated reports are then transmitted, along with other system products, to USDA evaluators who refine the information for a final product to be released to the public.

The closing of the loop allows the evaluators to issue requirements, because of product deficiencies or changing missions, to systems management. Management identifies impacts and develops those changes which are justified. Requirements from public policy could also be input to systems management.

The following paragraphs describe each of the components in the diagram with two exceptions.

1. The systems management component has been defined in the preceding section.

2. The data base component, though not indicated

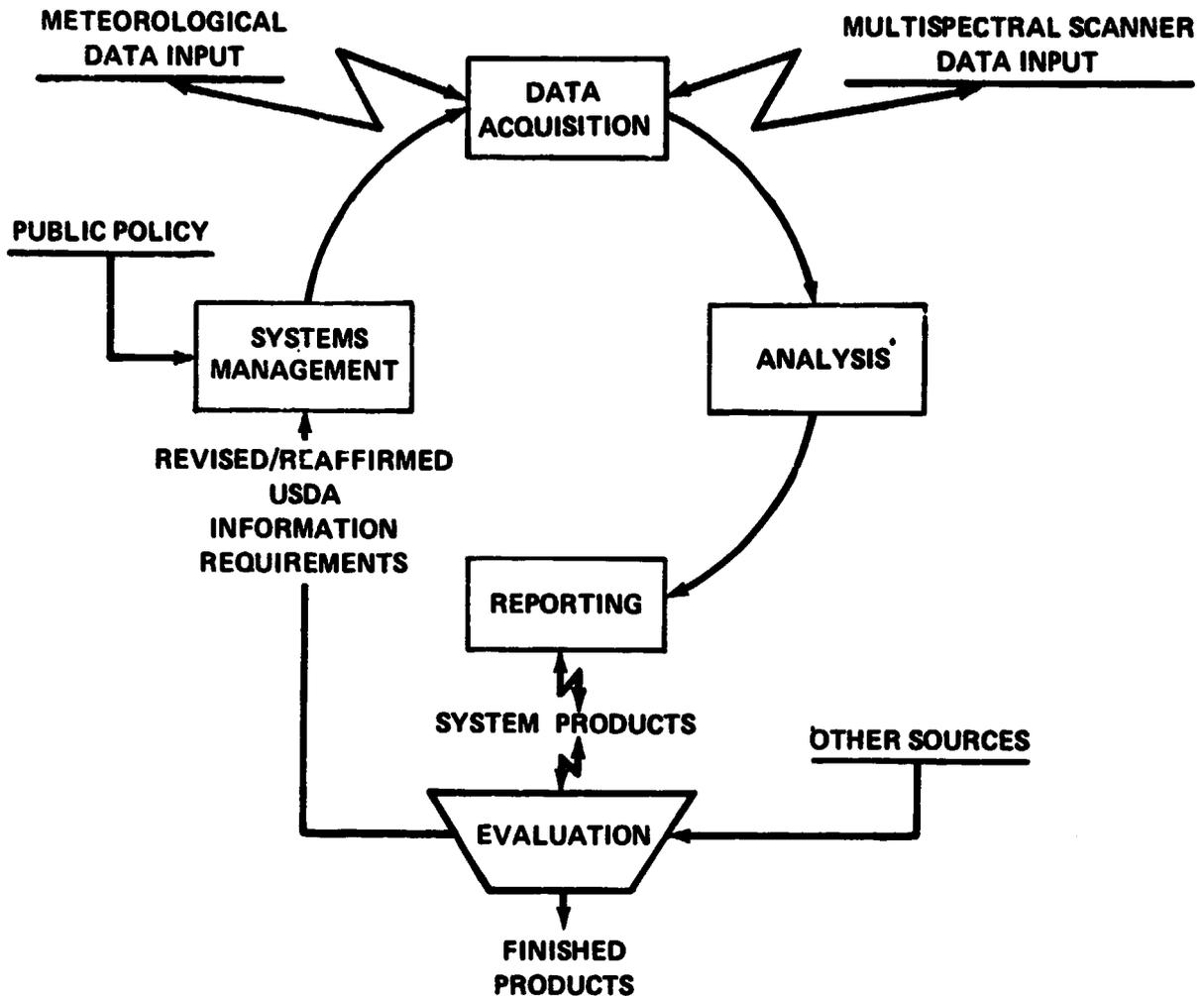


FIGURE 2.—USDA system product and information flow.

on the chart, will be defined since it provides much of the interface between system components.

The components in figure 2 are composed of hardware, software, and procedures, whereas the organizational elements in the preceding section define personnel and policy areas. The interaction between these two structures is evident in figure 1, with the data base component separating the two.

Data Acquisition Component

The functions of the data acquisition component are (1) to serve as the focal point for transmitting data requests to, and receiving data from, sources ex-

ternal to the system; and (2) to perform requirements processing, preprocessing of image, meteorological, and ancillary data, and communications processing. With respect to requests for specific full-frame acquisitions and the transmission of data from the NASA Goddard Space Flight Center (GSFC), the data acquisition component is required to perform the following tasks.

1. Store data on large-capacity random-access devices as they are received.
2. Perform cloud-cover and quality checks on the segments, extract those which passed the editing, and place them on the analysis station data bases.
3. Send reports describing the number and quality of segments to the analysis component.

4. Prepare a data packet containing digital image and ancillary data for the analysis activity.

5. Obtain and store NOAA and/or U.S. Air Force (USAF) agricultural/meteorological (AGROMET) and meteorological data. The transmission of data is via communication link on a periodic basis; the data are temporarily stored on a random-access device. Extraction of the desired data and its required manipulation and placement in the data base is done on a non-real-time basis

6. Handle and control other types of data such as historical data, recent attache reports, and research results available from various agencies.

7. Reformat hardcopy items in digital form, if required, and place them in the data base, or place hardcopy material in the system's reference library.

Evaluation Component

The evaluation component provides the advanced system with the user agency interface, product evaluation support, and the means for initiating new or changed requirements based on product evaluation.

Standard reports produced by the reporting component are available to the user agency facility immediately following validation. The schedule for generating these reports is consistent with the user agency schedule for release of official crop estimates. Generally, the transmittal reports estimate crop area, yield, and production to the country level, but the capability to provide estimates to the lowest level required is available. Historical data reports are also available.

The system provides for generation of nonstandard data requests to meet specific user agency needs. Designated user agency personnel identify specific data needed, such as crop effects from episodic events or correlation of meteorological data with estimates of area, yield, and production. Users may access the data via interactive displays or hardcopy reports.

User agency personnel evaluate results for accuracy, utility, and timeliness with respect to preestablished schedules. System results are compared with data from other USDA sources for accuracy. Utility evaluation was to consider the completeness, accessibility, and usability of system results. In addition, production system support personnel performed an analytical evaluation of results to improve sampling strategy and processing

techniques.

For each crop year, a Product Evaluation Plan is planned covering (1) assessment of system results relative to data from other USDA sources, such as attache reports or foreign publications; (2) assessment of system results using ground-truth (or analogous) data; and (3) simulation studies.

Analysis Component

Responsibilities of the analysis component include generating estimates of crop acreage, yield, and production at all specified geographical hierarchical levels within the seven foreign countries. Standard statistics at these levels are computed for acreage, yield, and production and combined with historical statistics to provide estimates of the analysis component performance accuracy. Computed estimates and statistics are stored in the data base. The analysis component also generates estimates for specified geographical areas associated with episodic events.

Operational requirements include specifications of (1) geographical areas for which periodic and unscheduled reports are requested, (2) sampling strategy, hierarchical definition, and sample unit allocation plan, (3) sizing parameters to control length of tables and memory allocation, and (4) data collection requirements for the data acquisition component, which consist of three functional elements:

1. Classification to estimate wheat acreage for sample segments.

2. Yield to estimate wheat yield for the yield strata.

3. Crop aggregation to combine results from the classification and yield elements and to compute estimates of wheat acreage, yield, production, and standard statistics at specified hierarchical levels. Software is used to make reasonableness checks to assist in producing a valid product.

The classification, yield, and crop aggregation elements are designed to use maximum analyst interaction during initial operations and incorporate techniques requiring less interactive control as such techniques are verified. The goal of the UAS is to make the transition operationally to a system performing the major amount of analysis with minimum analyst interaction. The ratio of minimum to maximum interactive data processing loads is influenced primarily by performance tolerances specified by the analyst or the systems management component.

Data Base Component

The data base component provides the data interface between the system components. Each component receives data from, and places its results in, the data base for access by other components.

High-volume data sets such as sample segment and full-frame imagery for use as an entity are stored in a sequentially organized data base. Storage for other data sets is provided in the data base management/query data base, which permits the storage and retrieval of data in a hierarchical, network, and/or chained manner.

The integrated data base reduces the storage of redundant items and, through its logical structure, provides rapid storage and retrieval of data as required by the various system components. This integrated structure also introduces a common thread to the majority of data in the data base; i.e., a gridded, geographically referenced partition.

The Data Base Management System (DBMS) logically and physically defines the data base and provides storage and retrieval mechanisms. Since the system is oriented toward interactive analysis, rapid data base access is crucial. The hierarchy provided by the partitioned logical structure contributes toward meeting this goal because the user is able to reference various data types with a common attribute and reduce data base accesses and terminal entries. Concurrent access by interactive and batch users provides additional flexibility and increases system throughput. The data base is the responsibility of the data base administrator with software support from the technical staff.

Reporting Component

All scheduled and unscheduled reports are produced by the reporting component. These reports, placed in the data base, are made available to the USDA users via communications link. The automated reporting process has a minimum capability to store report formats, provide reports at varying levels of hierarchy, and provide proper security control for sensitive data.

The reporting component supports predefined formatted and unformatted queries initiated by members of the production staff or by the USDA user. The query results are presented to the USDA user in the same manner as the scheduled reports or to the production system user in hardcopy or display

form.

Both software and procedural checks are applied to the reports prior to release to the evaluation component, with checks on format and completeness being performed manually.

All software which interfaces with the analyst is tutorial in nature, with a menu presentation used as often as feasible. The query language is such that non-computer-oriented professional personnel could use it efficiently. Appropriate error messages and required corrective measures are designed for clarity to the analyst.

The reporting component is under administrative control by the project management and supported by the technical staff, including systems analysts and computer specialists. Analysts also have access to a status and tracking data file and a production system library, which are described in the following sections.

Status and Tracking

The system design provides for a status and tracking data file to be available for the various management and technical groups to use in obtaining information required to efficiently manage and control the production system.

The status and tracking information is provided by the various components of the production system. Required data could be placed in the status and tracking file by an analyst from an interactive terminal or by a software module which is part of a process. For example, the software module which performs clustering writes a record to the status and tracking file after each clustering task. The record contains information identifying the segment, date, and time of processing.

The status and tracking file is designed to provide data concerning all aspects of the system, including system throughput and processing statistics for management. Data base activity is reported to the data base administrator, and reports on nearly all phases of the production are provided to the systems management component.

Production System Library

The production system library is an automated index of all documents, film products, various maps, and other hardcopy ancillary data sources required by the production system. An analyst could enter a

query from a terminal and learn whether or not a given reference is currently available. If available, the item would be logged out to the requesting analyst by modifying the index record on the data base.

DATA PROCESSING SYSTEM

The USDA system design provides for a complex Data Processing System (DPS) for support. The DPS provides the data base, the processing capability, the crop analysis displays and processing, the report generation, and the mode of interface with the user. The DPS would be modular to allow phased implementation and growth to accept expanded support requirements. The modular concept would result in the expected use of multiple small- to medium-class computers, with subsystems to function independently. The data base subsystem would control data flow.

A set of small to medium computers, related peripherals, and operating system software provides support to the DPS. Figure 3 presents a feasible equipment configuration with assigned processing functions.

Computers

All computers are standard products with required interface devices, including necessary timing, logic, and buffering to facilitate the computer-to-computer interface. The computer-to-computer interface provides the capability to pass the up-to-date status and tracking tables between processors. The controlling Central Processing Unit (CPU) flags data requiring analysis or data base support in the status table. All computers monitor the status and tracking tables to determine when processing or a change in resource allocation is required.

The computers have an interrupt structure within a CPU to allow control transfer to a new process. Changes in process control occur as a result of external or internal signals with interrupt logic able to respond to either response requirement. The computer systems have self-diagnostics under operator or technical engineer control and are available to support processing at least 85 percent of the nominal 16-hour day.

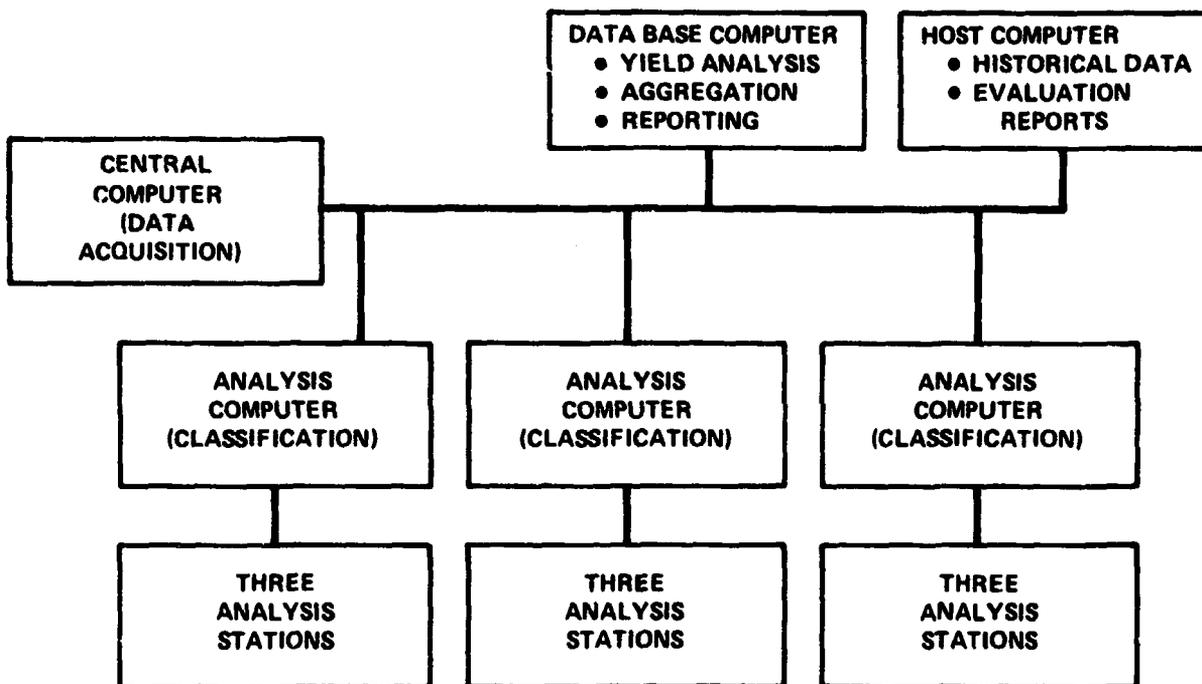


FIGURE 3.—Automatic data processing support equipment.

Operating Systems and Support Software

The design provides for operating systems that are vendor standard products. No uniquely developed code would be implemented into the system, except for the control of nonstandard interface devices. Methods of interfacing special devices or different vendor products would be added using standard "hooks" to the operating system. The design includes provisions for incorporating future revisions of the operating system and user system unique operating system level software using a standard system build.

Each operating system supports up to 12 concurrent jobs in a multiprogramming environment. The jobs are scheduled on a priority basis with the capability to change job priorities dynamically.

The system support software provides for user/computer interaction at the appropriate CPU's. An easy-to-use conversational language is provided to the user at interactive terminals. All applications software modules rely on system software to schedule, control, and translate messages to or from the interactive terminals.

Interactive transaction-oriented processing is provided by the system support software. The transaction processor maintains logs of the transactions conducted at the terminals. These are retained until the data base is updated and reflect activated software processes and the operational revision level of each.

Each operating system monitors and controls the devices assigned to the respective CPU's. Devices within a subsystem may be reassigned by the individual computer operators. All devices can be reassigned from the operations manager console via the subsystem controlling CPU's. Status displays of all DPS resources are maintained for display at the operations manager console on a scheduled or a demand basis.

SIMULATION

Simulation was initiated in June of 1976 to track and verify a design (ref. 4) for the USDA system. As new hardware or design approaches were identified and quantified, they were then simulated to verify the adequacy of their approach. This allowed USDA to assess computer performance prior to making any capital investments. Since time was of essence, a simulation model had to be found that was available locally at little or no cost to the Government. A

thorough search uncovered an IBM proprietary model already installed at JSC that would be made available to USDA. This approach also had the advantage of providing USDA with local IBM personnel who were intimately familiar with the model, thus eliminating the learning-curve time requirements.

Performance prediction and design optimization of the user system required the support of simulation modeling. Simulation was required also, according to the Management Plan (ref. 1), for economic analysis. Initial tasks of simulation were identifying procedural and feedback relationships among functions, identifying major modules and algorithms within each function, and identifying module flows and resource requirements to include frequency of execution.

The parameters required to validate a candidate configuration were hardware configuration, with the relevant performance characteristics; software functions, with their relevant resource demand characteristics; data base designs; and information processing system workloads in terms consistent with the use of the model. The task of simulation modeling then proceeded, and the system design was converted to an input form to begin simulation of both system performance and system throughput.

System Performance

The objectives of system performance modeling were to determine the critical parameters affecting elapsed time and resource utilization for each process. This included determination of input/output activity against data files and "bottlenecks" impeding system performance. Another goal of system performance modeling was to evaluate special-purpose equipment, such as classifiers and mass storage devices.

In order to achieve these goals, it was necessary to define the proposed hardware characteristics, system configuration, software design, operating system characteristics with specific services, and data base management functions. Included in the definitions were the size and rate of data transfers and inter-module communication.

System Throughput

The objectives of throughput modeling were to determine the time required to process a given data cycle and generate a specified report. This would in

turn determine the performance of critical sub-systems necessary to meet throughput criteria. The determination of the time required to process a priority episodic event in a fully loaded system was also an objective. These objectives would allow the evaluation of the processing control algorithm.

The requirements for meeting these goals were to define the processing cycle in terms of each processing step, review and rework cycles, and reporting periods. In support of the processing steps, it was necessary to define data collection cycles, quantities of equipment and personnel, and work schedules.

Results

The simulation of a feasible system design provided timely answers to system design questions, such as the ability of a minicomputer to handle the proposed geometric correction of MSS data. It was determined that this processing function could constrain the types of computers which might be appropriate for the system. However, the use of an external array processor reduced CPU requirements significantly and permitted large arrays of data to be maintained in memory without relying on page or mapping registers.

Because of the time overlap of design and simulation, it was possible to elaborate on simulation details as the design proceeded and to modify the design

based on simulation results. One major verification of the feasible system design was that an average sample segment processing time was approximately 1.8 hours, which supported the required system throughput and associated constraints described in the Introduction (ref. 5).

REFERENCES

1. Management Plan for USDA Advanced System Design, U.S. Department of Agriculture, Houston, Texas, Feb. 1976.
2. A Study of Digital Image Analysis Systems. MTR-7446, The MITRE Corp., Jan. 1977.
3. USDA Advanced System Study Requirements Definition Document. SISO-TR625, Ford Aerospace and Communications Corp. (Houston), Jan. 1977.
4. A System/Subsystem Design for the USDA Which Will Provide a User Advanced System. SISO-TR633, Ford Aerospace and Communications Corp. (Houston), Jan. 10, 1977.
5. USDA Advanced System Model Analysis Results. NAS9-14350, NASA Johnson Space Center (Houston), Aug. 25, 1976.

Data Base Design for a Worldwide Multicrop Information System

W. G. Driggers,^a J. M. Downs,^a J. R. Hickman,^a and R. L. Packard^a

BACKGROUND

Data base design for the U.S. Department of Agriculture (USDA) Application Test System (ATS) was based on a combination of data requirements to meet the needs of end users, remote-sensing analysts working with remote-sensing or crop-reporting procedures, management, and the system development team. These different categories of planned ATS users sometimes view the same data items differently and use them differently. They also have differing needs for access to the data for processing or informational purposes. Furthermore, their needs tend to change at times. One of the primary concepts of the ATS has been to provide a central geographically oriented data base to serve varied application modes, as shown in figure 1 (also see reference 1).

LACIE experience with data needed to support the crop estimation process was of significant value in establishing ATS data base requirements. Processing procedures using the LACIE Earth Resources Interactive Processing System (LACIE/ERIPS) provide ready access to digital imagery, fields definitions, and other data required for statistical separation of spectral classes; these data are managed efficiently by the Information Management System (IMS), a data base management system available on the IBM 360/370 series computers (refs. 2 and 3). Meteorological data used in estimating crop yields and crop calendar adjustments are extracted and processed at sites remote from other LACIE activities. Processing required for estimating production, aggregating results, and reporting results use still another set of computer hardware and software. Interfaces among these LACIE components (and other data sources, both manual and automated)

have at times been awkward, time consuming, and difficult to control.

Analysis of the need for improved data logistics indicated that requirements could be met best through implementation of a central data base, controlled by a generalized data base management system. This approach would make the data accessible both to application software and to direct query by the various users. Use of a data base management system offers the potential for providing greater flexibility to meet changing requirements. Proper design for a central data base also provides an optimum balance among data consistency, redundancy, access, and responsiveness (ref. 4).

The purpose of this paper is to describe the ATS data base design approach and resources. Following a summary of requirements for data and information, the data will be described in more detail by category, with emphasis on those characteristics which influenced the design most. Then the remaining steps of the design process will be discussed briefly.

User Requirements

Current USDA priorities for the use of remote sensing have been defined as follows.

1. Early warning of changes affecting production and quality of commodities and renewable resources
2. Commodity production forecasts
3. Land use classification and measurement
4. Renewable resources inventory and assessment
5. Land productivity estimates
6. Conservation practices assessment
7. Pollution detection and impact evaluation

ATS data base design must provide support for report preparation and information gathering in support of these priorities. For example, early warning analysis of changes affecting production and quality

^aU.S. Department of Agriculture, Houston, Texas.

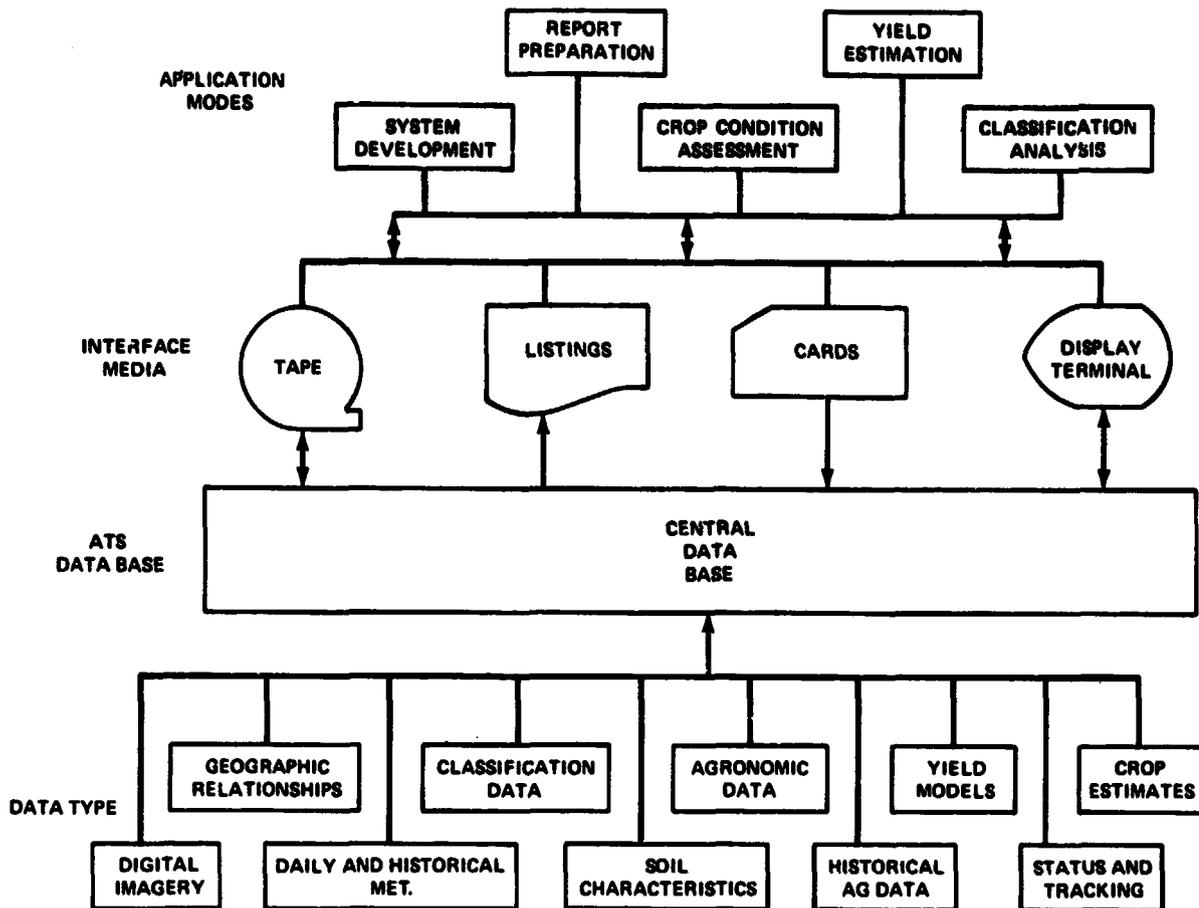


FIGURE 1.—Application/data base interface.

for a specific crop in a specific geographic area might require periodic reports as to climatic alarms in that area; the analyst assigned the early warning analysis task might also need to query selected weather data parameters for that area over some period of time. These needs should be supported by ready access to consistent data.

Depending on the application, the user may need access to ATS data by geographic area or by some combination of geography and date or geography and crop. Weather data, for example, will be accessed only by geographic location and date, whereas a yield estimate is a function of both location and crop. ATS data base design must permit the user ease of access to data according to the most common relationships in which the user views the data; that is, the data relationships, as well as the data itself, must be part of the data base definition.

Analyst Requirements

The term "analyst" is used here to refer to an individual who uses remote-sensing data to produce crop estimates and crop assessments. Several different processes must be supported to assist analysts with one or more of the following tasks.

1. Imagery classification. In addition to storing and accessing the digital imagery and classification data, the ATS data base must support analyst queries of other data types, such as meteorological data or soils data in the vicinity of the segment being analyzed.
2. Area estimation.
3. Yield estimation.
4. Crop calendar adjustment.
5. Climatic alarm detection.

6. Production estimation.
7. Sample allocation.
8. Report generation.
9. Results evaluation.

Similar types of data base support are required for these different tasks. Data relationships are also important to the analyst, just as for the user, although the two are not necessarily interested in the same relationships. Users may need to access data on the basis of administrative boundaries, for example, while analysts may need access to the same data types on the basis of proximity to a specific sample segment.

Management Requirements

Management requirements of the ATS data base cover those data categories needed to assess the current status of remote-sensing processing activities and to plan future activities. Data describing processing backlog, computer system status, and current activity related to specific crops and geographic areas are needed. Status data must cover processing of imagery data and meteorological data and generation of reports.

General Requirements

Large volumes of data are required for the analysis of remote-sensing data. The data category with the greatest volume is the digital Landsat imagery. Each sample segment currently used in crop assessment and estimation comprises 91 728 spectral-intensity values, defining 22 932 pixels; header data bring the total data to over 92 000 bytes (one byte for each intensity value) for each image. During the 1978 and 1979 crop years, an estimated 2500 to 6000 images will be required. It is also important to provide fast display on the color cathode-ray-tube (CRT) screen for images, class maps, and masks.

Significant volumes of meteorological data are also required. In order to assist analysts in yield estimation, crop calendar adjustment, and image analysis, meteorological data parameters (for large numbers of meteorological stations and grid cells) should be retained on-line as long as practical. At least 90 to 120 days of current meteorological data and 10 years of historical meteorological data are required. Historical agricultural data and crop estimate reports also require significant data volumes.

Other general requirements for the data base include minimum redundancy, data consistency, ease of query and maintenance, and flexibility. Data redundancy increases storage requirements and processing time. Redundancy also increases the risk of data inconsistency; that is, when a data element exists in more than one location, the risk of updating one and not the others is higher. Inconsistency, in turn, reduces the usefulness and reliability of the data. Flexibility is needed to accommodate anticipated changes in other data requirements.

DATA CATEGORIES

Geographic Entities

Geographic entities used in the ATS data base structure include the LACIE geographic hierarchy, agrophysical units, meteorological stations, and grid cells. Definition of the relationships among these entities is a key element in the data base structure. Most other ATS data types are defined with relation to one or more of these geographic entities.

The geographic hierarchical levels currently used by LACIE are country, region, zone, stratum, and substratum. The specific number of these levels and their identification with political or administrative boundaries vary from one country to another. For example, in the United States, a zone corresponds to a state, a stratum to a crop reporting district, and a substratum to a county; in the U.S.S.R., a stratum corresponds to an oblast and is the lowest level. LACIE codes are used to identify the hierarchical levels to the ATS data base. The climatic crop region is a specific grouping of some of the hierarchical entities for applying certain LACIE yield models (ref. 5).

Meteorological stations are, for the most part, those World Meteorological Organization (WMO) stations located in the crop areas of interest. Standard WMO codes are used to identify the stations; stations not in the WMO network are identified by call codes.

The grid cell entity used for the ATS data base is taken from an Air Force meteorological data grid, defined as a rectangular mesh on a polar-stereographic plane (ref. 6). When projected onto the Earth's surface, the length of a side of a grid cell is approximately 25 nautical miles at middle latitudes (precise size and shape vary with latitude and

longitude). Each grid cell is identified by an (I,J) pair, representing the matrix location with respect to the two axes on the polar-stereographic plane. The J-axis is parallel to the meridians which define 100° east longitude and 80° west longitude. Algorithms are available for converting from latitude and longitude to an (I,J) pair and vice versa.

As described later in this paper, several data types are recorded for the grid cell mesh. For some of these data types, each cell can be divided further into quadrants (identified as A, B, C, and D); a quadrant is therefore about 12.5 nautical miles on each side at the middle latitudes. For some purposes of the ATS data base, each zone-level area (or stratum level, depending on country) in the geographic hierarchy is defined as the collection of grid cells contained in that area; this definition can be extended easily to higher levels of the hierarchy.

Use of grid cells provides (1) a smaller geographic unit than other geographic areas used in LACIE for recording data and (2) a convenient method for analyzing results over otherwise undefined geographic areas.

The agrophysical unit (APU) provides definition of the two remaining ATS geographic entities. An APU is defined as an area with similar soils, climate, topography, and other agronomic factors such as land use intensity. The intersection of an APU and a zone-level or stratum-level geographic area is identified as a refined stratum. In the ATS data base, an APU is further defined as a collection of grid cells. One might choose to consider the APU, as implemented for ATS, as an irregular polygon which can be converted readily to grid cells.

Crop Samples

The crop sample unit used in the ATS data base is the LACIE-defined sample segment (ref. 7). As in the three phases of LACIE, this sample is an area about 5 by 6 nautical miles (9.26 by 11.11 kilometers) in size. For a specific crop of interest, sample segments are assigned to each geographic area for which the crop is to be analyzed. In LACIE Phases I, II, and III, geographic areas to which sample segments were allocated were defined by administrative boundaries. For the ATS, as in the LACIE Transition Year, sample segments are allocated by APU's and then apportioned to the refined strata comprising each of the APU's.

Digital Imagery (Landsat)

Digital imagery data for each sample segment are extracted from Landsat scenes, each scene being about 185 kilometers square in area on the Earth's surface. The basic unit of imagery data is a pixel, or picture element, referring to one instantaneous field of view (about 1 acre in size) as recorded by the multispectral scanner system (ref. 7). For the three LACIE phases, the digital imagery data for a sample segment contained the four bands of multispectral data for 117 lines, each line containing 196 pixels. All digital images currently in the ATS data base have these dimensions. The data base design, however, provides changeable limits for the number of channels, lines, and pixels per line in an image.

Classification Data

In addition to the digital imagery, several data types which are either used or generated during the classification process are retained in the ATS data base. These include fields data, classification maps, dot definitions, and masks. Fields data are consistent with LACIE definitions; a label is assigned to each field, up to 10 vertices are permitted for each field, and a field class is generated for each field. A class (or classification) map has the same number of lines and pixels as the corresponding image and each pixel is assigned a class; three class maps are permitted for each segment. Dot definitions are consistent with requirements for Procedure-I dots (Procedure I is a specific procedure for assigning pixel classifications). In the expectation that masks may someday be required for excluding pixels assigned to the two classes, designated other (DO) and designated unidentifiable (DU), provisions have been made for storing these masks in the data base.

Meteorological Data

Meteorological data in the ATS data base will include daily and historical parameters both for WMO reporting stations and for the (I,J) grid cell mesh. Daily meteorological parameters now available for WMO stations include maximum temperature, minimum temperature, and 24-hour precipitation. Monthly summaries are prepared for climatic regions, as required for input to first-generation

LACIE yield models, and will be prepared for individual stations as historical data for a 1-year period. Provision is made for future expansion, both in types of meteorological data collected and in the extent of global coverage.

Future plans also call for interpolation of station data to provide the same daily meteorological parameters for (I,J) grid cells in crop areas of interest. Historical summaries will also be prepared and will begin at that point in time when the capability is first available.

Agronomic Data

Agronomic data in the ATS data base define major crops and their densities in the areas of interest. Data on common cropping practices, irrigation and drainage, predominant soil taxonomy, and nominal crop calendars have been estimated and recorded using various information sources.

1. Maps. Operational Navigational Charts (ONC's, scale 1:1 000 000) published by the Defense Mapping Agency are the basic maps for data derivation. Soils maps of the same scale, developed at South Dakota State University under contract to USDA, are also used (refs. 8 and 9).

2. Overlays. Transparent overlays to the ONC maps containing agricultural/nonagricultural delineations, sample segment locations, soils, APU boundaries, and (I,J) grid cell delineations are used in recording the data.

3. Imagery. Digital Landsat imagery is used to assist in the definition and recording of some data parameters and in the refinement of other estimates.

In addition to the grid-oriented agronomic data, information describing soil characteristics has been incorporated into the data base. Developed at Iowa State University under contract to USDA, these data contain many encoded soil properties (such as particle size, mineralogy, available water capacity, permeability, salinity, and land use suitability) for each soil series. From the many encoded properties, those which appear to be of value in crop assessment were extracted for the ATS data base. These soils data are queried by the crop analyst to aid in classifying imagery and in crop assessment.

Crop Assessment Reports

The crop assessment process, including generation of reports, can require data from most of the

other categories maintained in the data base. Depending on the specific analysis being performed, reports of the following types would be required on demand.

1. Crop area, yield, and production estimates for current crop year
2. Climatic alarms
3. Water resources
4. Land resources

Both tabular and graphic forms are required. Retention periods for data in the various reports produced by ATS will vary according to security requirements. Generally, reports generated by the ATS will be retained on-line in the data base for at least 2 crop years.

Historical Data

Historical meteorological data and historical crop estimates (generated by LACIE and by ATS) have been described in previous paragraphs. Historical crop estimates generated by the USDA Statistical Reporting Service (SRS) and Foreign Agricultural Service (FAS) are maintained for specified crops and areas of interest.

Status Data

Processing of data, both digital imagery and meteorological, is tracked from the time the data enter the ATS. Processing status summaries provide information regarding what data are available, completeness of results, and work backlog.

DATA CHARACTERISTICS

Sources

ATS data sources are in general the same as LACIE data sources (table I). For example, Landsat digital imagery and meteorological data by station are extracted either from LACIE sources or from LACIE data files, and map overlays to locate sample segments and to delineate agricultural land use are the same as those used for LACIE. In many cases, however, ATS has established its own data sources; this is particularly true in the case of data recorded for the (I,J) grid cells—the ATS “gridded data” such

TABLE I.—ATS Data Sources

<i>Record type</i>	<i>Data items</i>	<i>Source</i>
Country, crop	Country, crop, and relationship	Encoded manually
Geographic hierarchy	Hierarchical geographic levels	Encoded manually from ONC maps
Sample segment description	Sample segment identification	Encoded manually from ONC maps, segment overlays
	Reference frames	Extracted from LACIE digital imagery data
	Location, active image acquisitions	
	Soil (predominant)	Encoded manually from ONC, soils overlays
	Classification data (climatic) (alarms, fields, etc.)	Generated by application software and analyst definition
Sample segment acquisition	Imagery header data	Extracted from LACIE imagery
	Crop development stages	Nominal—encoded manually Adjusted—generated by application software and analyst observation
Image line	Pixel intensities	Extracted from LACIE imagery
DU mask	Designated unidentifiable pixels for acquisition	Generated by application software and analyst definition
Dots	Dot information	Generated by application software and analyst definition
Class maps	Pixel classification for image	Generated by application software and analyst definition
DO mask	Designated other pixels for segment	Generated by application software and analyst definition
Fields	Field vertices, classification	Generated by application software and analyst definition
Evaluated segments	Segment results from classification	Generated by application software and analyst definition
Crop estimates	Area, yield, production by geographic location	LACIE—transferred from Crop Assessment Subsystem data base ATS—generated by application software
Meteorological (met) station	Met station location	Encoded manually from ONC maps, overlays
Daily met—station	Daily meteorological parameters	Loaded periodically from met data provided by National Oceanic and Atmospheric Administration (NOAA) for LACIE
Station crop data	Crop calendar adjustments	Generated by application software
Climatic crop region	Yield models	Encoded manually
Yield monthly reports	Yield estimates	Generated by application software
Met summary—monthly	Input to yield models	Generated by application software
APU description	APU location, agricultural area	Encoded manually from ONC maps, overlays for APU and agricultural areas
Refined stratum	Geographic location	Generated by load program for APU description
Historical data		
Agricultural	Crop statistics	Extracted from USDA (SRS, FAS) data files
Meteorological	Meteorological statistics	Extracted from NOAA data files
Crop estimates	Estimates generated from remote-sensing data	Summarized from aggregated crop estimates generated by LACIE, ATS before purging from on-line files
Daily met—grid cell	Climatic alarms	Stored on analyst instruction
Agronomic—grid cell	Daily meteorological parameters	Generated by application software
	Agronomic factors	Encoded manually from ONC maps, color imagery, overlays
Geographic hierarchy		
Soils—grid cell	Soil taxonomy, features	Encoded manually from ONC maps, color imagery, overlays
Soils—general	Soil characteristics	Loaded from soils tape developed under USDA contract

as soils and agronomic data. As shown in table I, many of the data are also generated by application software; this applies particularly to classification data.

Volume

Volumes of data to be stored in the ATS data base initially are shown in table II. In order to estimate data volume, several assumptions were necessary. First, it was assumed that the area to be analyzed comprises six refined strata in two states of the U.S. spring wheat growing area (Montana and North Dakota) and one APU in the U.S.S.R. spring wheat growing area. It was also assumed that the digital imagery format would be the same as for Landsat-2, regardless of which satellite provides the data. Another assumption was that both station and gridded meteorological data would be sorted. These assumptions can be translated roughly into the following upper limits: 266 sample segments, 1800 acquisitions, 350 meteorological stations, and 3450 grid cells. The estimates shown in table II were based on these limits.

TABLE II.—ATS Data Volume Estimates^a

Parameter	Data volume, megabytes, for —		
	5 acq/SS ^b	5.5 acq/SS	7.0 acq/SS
Classification data	—	245	310
Support data	—	35	35
Software, working storage	—	50	50
Total	—	330	395
Assumed disk load factors			
60 percent	550	—	660
70 percent	470	—	565
80 percent	415	—	495

^a Assumptions: 266 sample segments, Landsat-2 format (4 bands, 117 lines by 196 pixels).

^b Acquisitions per sample segment.

ATS DATA BASE DESIGN

Configuration

The ATS hardware configuration is based on a Digital Equipment Corporation (DEC) PDP 11-70 mainframe. Hardware features include 256K words (512K bytes) of main storage, packs, and four 9-track tape drives. The computer operates under the IAS operating system, a standard DEC software product described in the paper entitled "The Application Test System: Technical Approach and System Design" by Benson et al.

Data base software includes file management software for large sequential files and a data base management system (DBMS) for the remaining data. Digital imagery and other high-volume data are handled as sequential files by the standard IAS File Control Services. This approach permits an image to be displayed on the color consoles with minimal access time. It also facilitates efficient handling of the classification process.

The DBMS used for other ATS files is IDMS-11, a proprietary product of Cullinane Corporation. IDMS was initially developed for use on the IBM 360/370 series computers and later converted to run on the PDP 11-70. IDMS-11 supports both hierarchical and network types of data structures, as specified in the CODASYL Data Base Task Group Report (ref. 10).

IDMS-11 provides separate language facilities for data definition (DDL) and for data manipulation (DML), both of which are language extensions of COBOL. However, the system can be used easily with the other languages (FORTRAN and macro assembler) which support CALL statements and is, in fact, so used by the ATS development staff. Several concurrent users can be supported by the system.

The entire collection of record types which comprise the data base is defined to IDMS-11 in a schema using the DDL. The schema defines all data elements, record types, physical data storage mapping, and set relationships in the data base. The user can access the data base only through a subschema, a subset of the schema predefined to include all data and data relationships needed for a specific data base application.

In IDMS-11, the set relationship defines most of the logical relationships among the various record types. Each set is a named collection of two or more record types—one "owner" record type and one or more "member" record types. Any record type can

serve as a member or owner record in any number of set relationships. Set relationships can be used to define complex structures among the record types in the data base (ref. 10).

IDMS-11 also provides journaling of all data base accesses which result in changes to the data base. Together with the dump facility, this feature will be used to provide data base recovery in case of a system "crash." Special recovery procedures are also available for crash occurrences.

Privacy provisions are not as complete. Each application is assigned a subschema, which can only be accessed by using a specified user identification code (UIC). A subschema is constrained to use only the record types needed for the specific application, and this constraint can be extended to the data element level. Most of the security provisions are dependent upon operating system capabilities.

Query Capability

In addition to providing access to the ATS data base by means of application software, the crop analyst has direct access to data through a query capability. For example, the analyst might want to review weather conditions at WMO stations nearest a specific sample segment for 10 days prior to the most recent acquisition date. An example of a terminal display resulting from this type query is shown in figure 2. Other query capabilities exist for viewing soils data and various segment-related data at the analyst terminal.

STATION	DATE	MAX TEMP (F)	MIN TEMP (F)	PRECIP (mm)
XXX-YY	24-MAY-77	58	73	0
	25-MAY-77	58	75	0
	26-MAY-77	59	76	0
	27-MAY-77	64	81	6
	28-MAY-77	59	80	10
	29-MAY-77	61	80	0
	30-MAY-77	65	82	0
	31-MAY-77	65	82	0
	1-JUNE-77	68	83	0
	2-JUNE-77	62	79	20
XXX-ZZ	24-MAY-77	54	70	0
	25-MAY-77	55	71	0
	26-MAY-77	55	72	0
	27-MAY-77	55	73	0
	28-MAY-77	58	78	5
	29-MAY-77	54	78	20
	30-MAY-77	56	72	5

FIGURE 2.—Query example.

Query capabilities developed for ATS use the same data base facilities as application software. The initial query packages have, in fact, been developed and implemented as COBOL and FORTRAN applications. These query packages require that the user respond to prompts by furnishing specific parameters to be used in data retrieval.

ATS Data Base Design Approach

Data base design for ATS was an iterative process. Data elements were identified first on the basis of LACIE documentation and discussions with analysts. Record types for the data were then proposed and the resulting structure reviewed with the analysts. Several cycles of review and revision preceded the current data base definitions.

Record Types

ATS record types were designed to accommodate all data categories required for crop analysis. Data elements were identified and grouped according to usage, dependency, and source. Consideration was given to usage both by application software and by means of direct query by the crop analyst operating at the console. Table III lists the record types (defined without regard to the specific implementation), approximate record lengths, and record occurrences.

Structure

Record types defined for the ATS data base are shown in the data structure diagram in figure 3. Key geographic entity record types are bounded by heavier lines in the diagram to emphasize their importance. It should be noted that the names shown for these record types are not precisely those used in the schema definition for the data base because of the need to abbreviate in the schema. Data relationships (not always the same as set relationships in the schema diagram) are also shown; that is, an arrow in the diagram indicates ease of access from one record type to another, but not necessarily through use of a pointer. A single arrow in one direction defines a one-to-one relationship; a double arrow in one direction defines a one-to-many relationship. Omission of an arrow in one direction indicates that the need to access data in that sequence is not expected.

TABLE III.—ATS Record Types

Record type	Length, bytes	Occurrence
Country	32	One record for each country
Crop	24	One record for each crop
Country-crop	32	One record for each combination of country and crop
Geographic hierarchy	60	One record for each stratum or substratum
Sample segment (SS) description	160	One record for each sample segment
SS acquisition	384	One record for each sample segment acquisition
Image line	800	One record for each line (4 bands) of a sample segment image
DU mask	200	One record for each line of a segment acquisition
Dots	20	One record for each dot defined for a segment under Procedure 1
Class map	200	One record for each line of a segment class map (3 class maps for each segment)
DO mask	200	One record for each line of a segment
Fields	80	One record for each field defined for a segment for each of 2 crop years
Met station	128	One record for each met station
Station crop data	24	One record for each computer-generated crop calendar adjustment at a met station
Daily met—station	50	One record for each day for each met station
Climatic crop region	128	One record for each climatic crop region
Met summary—monthly	50	One record for each climatic crop region for each month
Yield results	450	One record for each climatic crop region for each report generated
APU description	64	One record for each APU
Grid cell quadrant	128	One record for each grid cell quadrant
Full grid cell	64	One record for each full grid cell
Agronomic—grid cell	150	One record for each grid cell quadrant for each crop
Daily met—grid cell	96	One record for each day for each full grid cell
Agronomic—grid cell	150	One record for each full grid cell
Soils—grid cell	40	One record for each grid cell quadrant
Soils—general	960	One record for each soil taxonomy family

TABLE III.—Concluded

Record type	Length, bytes	Occurrence
Historical—agricultural	80	One record for each reporting-level unit for each year
Historical—crop estimates	60	One record for each reporting-level unit for each year
Historical—met	64	One record for each met station for each month
Status and tracking—imagery	—	Undetermined
Status and tracking—met	—	Undetermined
Evaluated segments	60	One record each time a segment is classified and evaluated
Aggregated results	300	One record for each reporting-level unit for each report period

Schema Design

Once the inherent data relationships were identified, emphasis shifted to design approach using techniques recommended for the ATS data base management system, IDMS-11 (ref. 11). Additional record types had to be defined wherever a many-to-many relationship occurred. Location mode, record key, and set relationships were defined for each record type. In some instances, a degree of data redundancy was retained with the object of providing more access paths and, as a result, possibly simpler query structure (ref. 12).

The ATS data base schema diagram was then developed on the basis of the foregoing definitions. Figure 4 shows the schema diagram developed for the grid cell area of the data base. Each block represents a record type, defining the record name, record identification, size (in bytes), location mode, location key, disposition of duplicates, and area name. Connectors between blocks represent set relationships and are annotated with set name, linkage (types of pointers), membership option (for storage and removal), and logical order within each occurrence of the set. Record types outside the grid cell area for which set relationships are defined with record types inside the area are also shown, but without annotation inside the blocks.

The final design step prior to coding consisted of mapping the logical definitions to physical storage

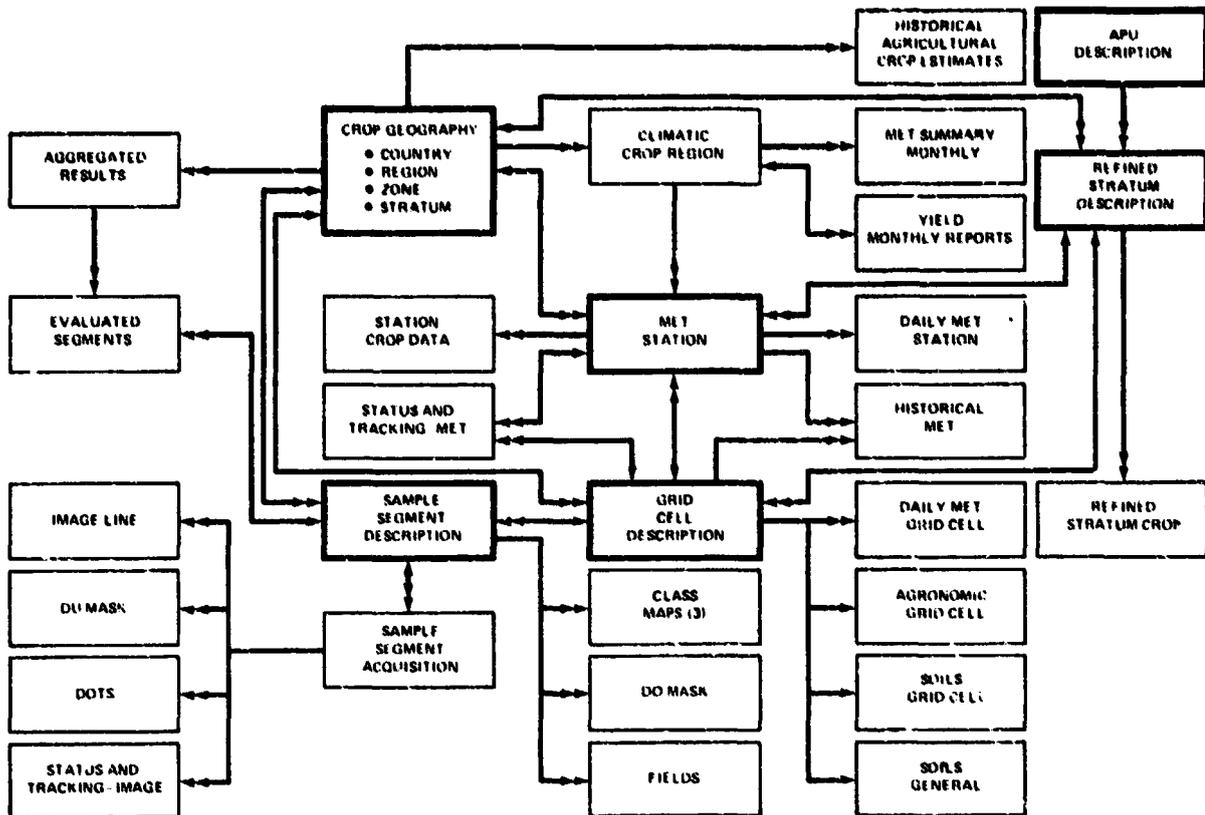


FIGURE 3.—ATS data base structure diagram.

units. Record types were grouped into areas according to the likelihood that they would be used by the same applications. Areas were assigned to physical storage units so as to provide the necessary storage space with the least likelihood of overflow.

EXPERIENCE TO PRESENT

Planned Phases

Implementation of the initial ATS data base was planned to cover the first two Transition Years—FY78 and FY79. Emphasis during FY78 is on demonstration of the usefulness of query capabilities to the image analyst, crop analysts, members of management, and other potential users. Usefulness of crop data representation on a gridded geographical basis for the APU will also be analyzed. Emphasis during FY79 will be extended to analysis of other

aspects of the data base, including different approaches to imagery storage and retrieval, interfacing more application software with the data base, and inclusion of other gridded data.

Initial Phase

Nineteen record types defined in the ATS data base schema were selected for the initial development, primarily on the basis of usefulness and logical load sequence. Subschemas and load programs were developed for these record types. Data were collected or recorded and loaded into the data base. Simple query programs, developed for some of the data expected to be most useful to the analysts, are currently being tested.

Use of IDMS-11 has provided solutions to many of the problems in implementing ATS, as hoped. However, some problems remain. Generalized query capabilities of the system are not yet to the desired

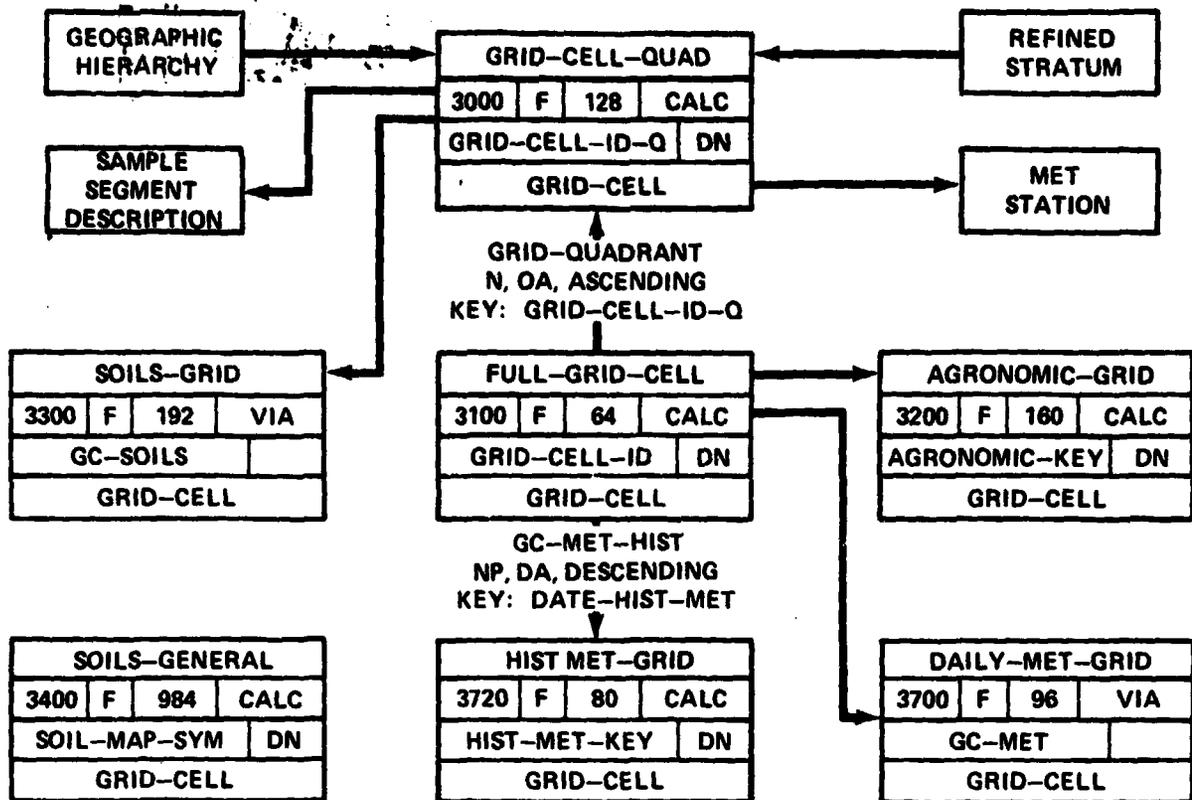


FIGURE 4.—ATS schema diagram.

stage of development, and problems remain in this area. A query language for data retrieval or update in an on-line, ad hoc environment is lacking. This necessitates development of menu-type query programs for each type of query needed. Overhead, primarily with respect to disk storage requirements, is significant. Until implementation was well underway, the ratio of actual data to total storage requirements was not realized. At the present, however, it is not planned to reduce overhead storage requirements by reducing the number of pointers because an increase in processing time would probably result. Some operational problems have occurred, mostly because the system is new to the staff.

The decision to use a generalized data base management system to manage a central data base in support of remote-sensing crop assessment appears to be sound. IDMS-11 appears to be a good minicomputer-base system for this purpose. The USDA ATS staff expects to expand the ATS data base concept in FY79 and succeeding years.

REFERENCES

1. User Requirements. LACIE-T00201, USDA LACIE Project Office, Foreign Agricultural Service, Washington, D.C., Oct. 1975, pp. 61-71.
2. LACIE-2 Design Review Document. IBM Corporation, Houston, Texas, Nov. 1974.
3. Earth Resources Interactive Processing System Requirements. PHO-TR514, Philco-Ford Corporation, Houston, Texas, 1972.
4. Martin, James: Computer Data Base Organization. Second edition, Prentice-Hall, New York, 1977, pp. 34-35.
5. Crop Assessment Subsystem (CAS) Requirements. LACIE-C00200, Vol. IV, NASA Johnson Space Center, Houston, Texas, Oct. 1977.
6. Earthsat Spring Wheat Yield System Test 1975, Final Report. Earth Satellite Corporation, Washington, D.C., Apr. 1976, pp. 3-1 to 3-4.

7. LACIE Project Plan. LACIE-C00605, NASA Johnson Space Center, Houston, Texas, Aug. 1975.
8. Westin, F. C.: Landsat Interpretation—Intensive Use as Cropland. South Dakota State University, Brookings, South Dakota, 1977.
9. Westin, F. C.: Progress Report, Agrophysical Mapping. South Dakota State University, Brookings, South Dakota, Nov. 1976.
10. CODASYL Data Base Task Group Report. Association for Computing Machinery, Apr. 1971.
11. Data Base Administrator's Guide. Digital Equipment Corporation, Maynard, Mass., July 1977.
12. Codd, E. F.: A Relational Model of Data for Large Shared Data Banks. ACM Communications, vol. 13, no. 6, June 1970, pp. 337-387.

The Application Test System: Technical Approach and System Design

J. L. Benson,^a D. R. McClelland,^a J. D. Tarbet,^a and R. F. Purnell^b

INTRODUCTION

The purpose of this paper is to provide insight into the technical approach which was applied to the system design of the U.S. Department of Agriculture (USDA) Applications Test System (ATS). This includes identification of requirements, assessment of remote-sensing contributions, evaluation of existing techniques, and cost-effective development of a system design which utilizes techniques and procedures consistent with requirements.

For many years, scientists and engineers have studied and proposed the potential roles of remote sensing in the management and exploration of Earth resources. It is currently estimated that operational use of Landsat data to derive agricultural crop information has a potential benefit of millions of dollars to the United States alone (ref. 1). The experience gained during the LACIE should result in the development of operational systems for processing Landsat data.

A major function of the LACIE has been the development, testing, and accuracy assessment of techniques derived to extract agricultural information from Landsat imagery. The project has demonstrated that techniques for classifying Landsat data have developed to the point where it is feasible to define systems for testing the LACIE-developed technology in a user application test. LACIE-proven technology has provided the basis for deriving information appropriate to a specific user; for example, the Foreign Agricultural Service (FAS) of the USDA.

The USDA established the User Systems Planning and Applications Test Group (USPATG) with the ground rule of using LACIE technology to define a system within specific USDA requirements. The

USPATG defined the user requirements for a processing system which could evolve from the LACIE and meet the USDA criteria for an operational system in the future.

TECHNICAL APPROACH

Ford Aerospace & Communications Corporation (FACC) participated in a joint study effort with the USPATG to develop a system design for large-scale processing of Landsat data. The study resulted in a series of reports (refs. 2 to 7), the most significant of which was a feasible design for a USDA system.

The USPATG study team followed the classical approach of designing a system. The design approach took the logical steps of

1. Identifying USDA requirements
2. Assessing possible remote-sensing contributions
3. Evaluating existing processing techniques and procedures
4. Developing a system design which utilizes the techniques and procedures to meet user requirements in a cost-effective manner (ref. 5)
5. Providing limited demonstration of an end-to-end system approach
6. Updating system design and developing an operational data system

Along with the user requirements for data content, FAS established general guidelines for the system. The guidelines were prioritized for support of such design approach tradeoffs as timeliness of results, ease in developing the system, cost of operating the system, and accuracy of results.

USDA SYSTEM DESIGN STUDY

The identification of USDA user requirements was the first step toward designing a specific USDA

^aFord Aerospace & Communications Corporation, Houston, Texas.

^bU.S. Department of Agriculture, Houston, Texas.

processing system. The design was to provide the functional capabilities required for the inventory and reporting of agricultural crops and was to include the capability to be transferred to, and to interface with, existing user facilities. The complete USDA system (designated ATS) was to provide functional capabilities for system management, data acquisition, analysis, reporting, and evaluation (refs. 2 and 8). The ATS provides the processing capabilities necessary for the transfer and evaluation of required LACIE technology.

The design of the ATS emphasized system transferability characteristics (i.e., use of high-level computer programming languages) as well as the ability to readily accommodate change. The resultant design was consistent with the major system constraints; i.e., timeliness of results, modularity, total "off-the-shelf" components, cost effectiveness, and accuracy. The design also stressed the development of an operational system responsive to USDA user requirements. Thus, the ATS employed rather than developed state-of-the-art technology. The system relies on off-the-shelf components of limited specialization and is capable of responding to state-of-the-art developments in hardware and software technology through modular changes. This allows for easy expansion of the ATS to provide a worldwide, multicrop information system.

Finally, the ATS was designed for ease of utilization. It is intended for use by skilled resource analysts who normally will not be remote-sensing specialists. It was also important that the system support a non-labor-intensive operation. This implies that, where possible, operations are to be automated with manual intervention kept to a minimum.

The study performed by FACC demonstrated that the strenuous system requirements could be met with the following state-of-the-art systems components.

1. Minicomputers
2. Interactive image processors
3. Low-cost array processors
4. High-fidelity color displays
5. Integrated data base management systems

Based on the study and on other industrial surveys, the USDA issued a Request for Proposal (RFP) for 120-day delivery of the ATS. The system was to conform to the earlier guidelines, but additional performance and capabilities requirements were imposed, including the following.

1. The ATS must be capable of incorporating from one to five analyst stations.

2. A 117- by 196-pixel, four-channel image must be classified into eight classes in less than 10 seconds, using the maximum likelihood classification rule.

3. A 512- by 512-pixel image must be similarly classified in less than 60 seconds.

4. A 117- by 196-pixel, four-channel image must be clustered into 30 classes in less than 30 seconds.

5. The ATS must process full-frame Landsat imagery.

6. The ATS must provide an integrated data base system for the management of massive volumes of data.

7. The ATS must include an integrated interactive query against the data base.

8. The ATS must provide extensive display capabilities, comprehensive analyst support functions, and pattern recognition functions.

THE APPLICATION TEST SYSTEM

The FACC provided the processing system shown in figure 1 to the USDA for its ATS. The ATS represents a cost-effective, expandable system (ref. 8). The host processor of the ATS is a Digital Equipment Corporation (DEC) Programmed Data Processor Model 11-70 (PDP 11-70). The PDP 11-70 is a dual-bus computer capable of data rates from 0.8 megabytes on the massbus, which is required in a data-driven system. The effective utilization of the cache memory buffer provides an effective instruction cycle time of 300 nanoseconds. The main memory in the ATS host computer is 512 kilobytes, expandable to 4 megabytes.

The processing load incurred by the clustering and classification of image data is met by the floating point system's AP120B programmable floating-point array processor. The AP120B, a "pipeline" type processor, is configured to provide the results of an addition and a multiplication every 333 nanoseconds, with an expression capability every 167 nanoseconds. The use of the AP120B and well-balanced system software enabled the ATS to meet the stringent system throughput requirements.

Image display and analyst interaction are provided at each of the three analyst stations. The ATS can support up to five stations if expansion requires. The image display equipment is an International Imagery Systems (IIS) Model 70 with nine 512 by 512 8-bit refresh memories and three graphics planes. It includes two 512 by 512 color monitors with a cursor under trackball control. The ATS was delivered ini-

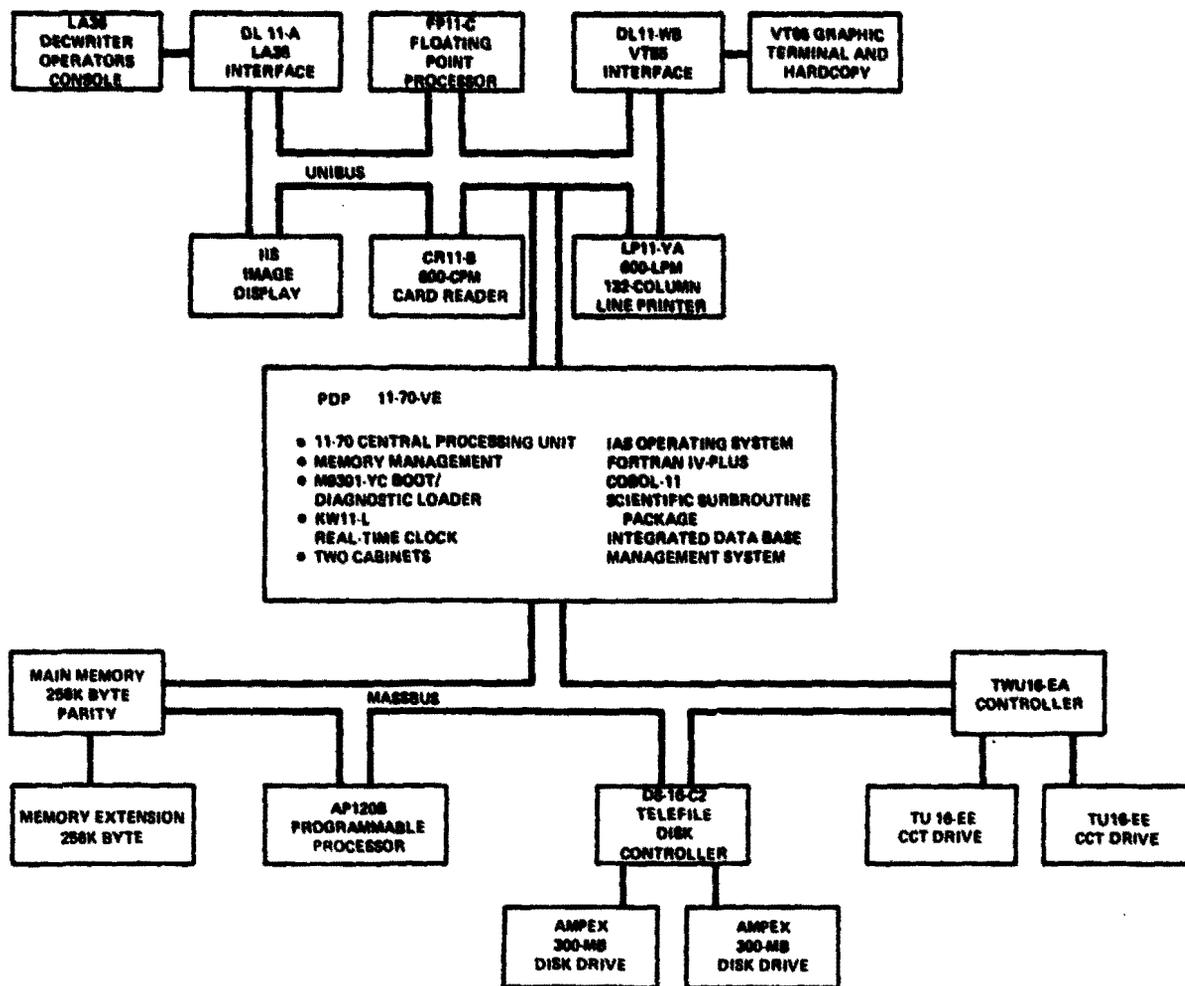


FIGURE 1.—ATS configuration.

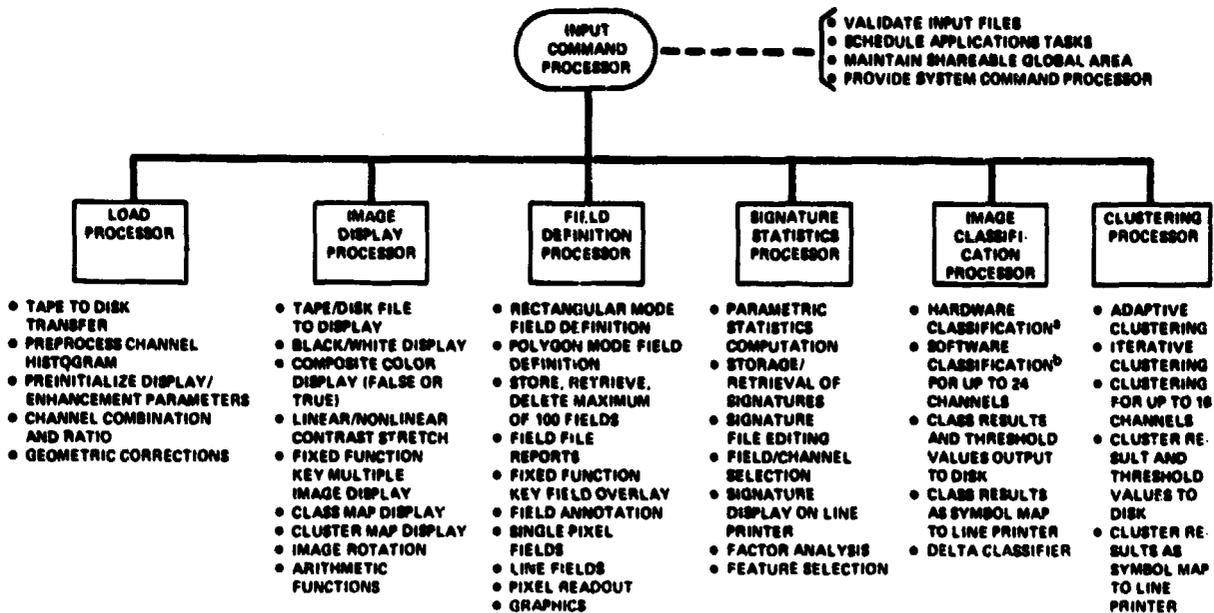
tially with one analysis station and has been expanded to three analyst stations.

The on-line storage of data for the ATS is provided by two 300-megabyte disks. The 3330-type disk drives are capable of read/write operations at a data rate of 1.2 megabytes per second. The system was installed with the disks on the PDP 11-70 unibus and a read/write rate of 600 kilobytes per second. The disk will be relocated to the massbus and upgraded to a full 1.2-megabyte-per-second input/output capability in late 1978.

The key to efficient hardware performance in response to analyst commands is the FACC Integrated Multivariate Data Analysis and Classification System (IMDACS), which has been under continuing refinement for multispectral scanner and seismic

data applications for 5 years. The basic structure of IMDACS is shown in figure 2. The IMDACS operates on the PDP 11-70 under the DEC Interactive Applications System (IAS), providing the user the capability to select and execute the major processing functions interactively via the alphanumeric/graphics terminal. The tutorial menu prompts display to the user the processing options that are available and the definitions of input parameters to be specified for the processing function. All analyst transactions are logged and may be output to the line printer upon conclusion of the processing.

The application software is structured along functional lines to support the processing steps required in performing digital image analysis. The major software functions are summarized as follows.



^aMAXIMUM LIKELIHOOD AND MIXTURE DENSITY CLASSIFICATION IN AP1208

^bMAXIMUM LIKELIHOOD CLASSIFICATION

FIGURE 2.—IMDACS software structure.

1. **Input Command Processor.** The applications software supervisor and common point of interface for all IMDACS processors

2. **LOAD.** Provides the capability to load digitally formatted imagery data from computer-compatible tape (CCT) to disk; also included are the preinitialization of image enhancements and display options, preprocessing of histogram tables, and channel combinations (linear, ratio, or normalization)

3. **IMAGE.** Provides the capability to format, enhance, and display imagery from selected data channels interactively

4. **FIELD.** Provides the capability to define, annotate, and save irregular-shaped fields; also provided are file maintenance utilities and fixed function key capabilities for automatic recall and display of previously defined fields and annotations

5. **STATS.** Provides for the computation, display, and storing of spectral signatures for defined fields; also included are related signature manipulation capabilities and the feature selection function

6. **CLASS.** Performs maximum likelihood and mixture density classification of defined fields and outputs classification results in the form of class map files

7. **CLUSTER.** Performs adaptive and interactive

clustering and outputs cluster results in the form of cluster map files

Communication among the processors is facilitated by a unified file structure. For example, statistics files can be built by either the statistics processor or the clustering processor and can be used to initialize either clustering or supervised classification processes. Thus, in addition to classical pattern recognition processing sequences, IMDACS can control new procedures such as "small fields," a LACIE-developed area classification procedure. In the small-fields procedure, training and test regions or picture elements (pixels) are labeled by the analyst. Statistics of the training regions are computed, and the resulting cluster statistics are then labeled in accordance with the labeled training data which are spectrally nearest the particular cluster mean. Mixture density classification is then initialized with the cluster statistics, and signatures are grouped by class automatically.

The following image processing performance time periods have been measured on the ATS:

1. 117 by 196 pixels, four channels, eight-class runs in 8.8 seconds

2. 512 by 512 pixels, four channels, eight-class runs in 57.8 seconds

3. 117 by 196 pixels, four channels, 30-cluster runs in 16 seconds

The performance times are achieved with a combination of high-speed input/output and high-speed, special-purpose peripherals. This performance, combined with the IAS multitask capability and the IM-DACS throughput efficiency, provides systems capable of testing and evaluating various technologies and operational procedures for the processing of remote-sensing data in a specific user environment.

System capabilities for supporting nonimagery data processing functions and for providing ancillary data support for image processing are provided by the Culliane Integrated Data Management System (IDMS-11) and the FACC Query and Report Writer, both of which have been implemented on the PDP 11-70 under the IAS. The IDMS was developed in strict compliance with the CODASYL standards.

SUMMARY AND CONCLUSIONS

The incorporation of new technology into the user's operations is critical to the development of any application system. The ATS is an example of a system with this capability, where LACIE techniques and procedures were merged with USDA requirements to define the design approach. The goal of the definition and design study was to couple overall feasibility with an extensive and diverse processing capability which minimized manpower requirements. The design was translated into ATS requirements; the ATS was implemented according to these requirements; and evaluation report criteria were defined for technology transfer (ref. 9). The ATS is now successfully supporting USDA activities in Houston, Texas.

The ATS is modular and can be expanded easily and modified piecewise as requirements may change because of changes in quantity or quality of input data or because of the desire and the ability to extract more and/or relevant information.

REFERENCES

1. A Cost Benefit Evaluation of the LANDSAT Follow-On Operational System. GSFC X-903-77-49 (Greenbelt, Md.), Mar. 1977.
2. Benson, J. L., et al.: A System Sub-System Design for the USDA Which Will Provide a User Advanced System. SISO-TR633, Rev. A, Ford Aerospace & Communications Corp. (Houston), Jan. 1977.
3. Benson, J. L.; and Tarbet, J. D.: Technique Validation Approach Document for the USDA Advanced System Study. SISO-TN823, Ford Aerospace & Communications Corp. (Houston), Jan. 1977.
4. Benson, J. L.; and Tarbet, J. D.: Detailed Techniques Flow and Timing Analysis for the USDA Advanced Systems Study. SISO-TN825, Ford Aerospace & Communications Corp. (Houston), Jan. 1977.
5. Benson, J. L.; and Tarbet, J. D.: A Model to Optimize Selection of System Elements. SISO-TN826, Ford Aerospace & Communications Corp. (Houston), Jan. 1977.
6. Benson, J. L.; and Tarbet, J. D.: Report of Simulation Assessment for the USDA Advanced System Study SISO-TN827, Ford Aerospace & Communications Corp. (Houston), Jan. 1977.
7. Benson, J. L.: Minicomputer Data Base Management System Comparison Evaluation. SISO-TR629, Ford Aerospace & Communications Corp. (Houston), Aug. 1977.
8. Benson, J. L., et al.: Functional Requirements Specification for the USDA Advanced System Study. SISTR626, Rev. A, Ford Aerospace & Communications Corp. (Houston), Jan. 1977.
9. Tarbet, J. D.; Bradford, L. H.; and Purnell, R. F.: On the Transfer of Remote Sensing Technology to an Operational Data System. Proceedings, Machine Processing of Remotely Sensed Data, Purdue University (West Lafayette, Ind.), June 1977.

Resource Modeling: A Reality for Program Cost Analysis

L. D. Fouts^a and R. L. Hurst^a

INTRODUCTION

The ever-important question of monetary resources required for the operation of a government program can be presented in several ways. This report conveys the approach, implementation, operation, and utilization of a model to establish capital investment and operational costs based on their interrelationships, dependencies, and alternative actions.

BACKGROUND

From its inception, the LACIE had a stated objective to determine the cost effectiveness of utilizing satellite and surface-derived data to monitor crop production and assess the impacts of agricultural and meteorological conditions affecting potential production. The Office of Management and Budget (OMB) wanted to know the cost of such an operational system. Senior U.S. Department of Agriculture (USDA) management needed to know the costs associated with the implementation and operation of this type of system to make decisions on future commitments to the effort.

The determination of all cost factors, interrelationships, and countless decision alternatives posed a complex problem. The straight analytical approach would accomplish the identification of cost factors and interrelationships, but to calculate the costs based on the interrelationships and countless configurations and decision alternatives still posed a monumental task. Thus, the concept of developing a model to assess the costs provided a logical and viable approach.

The concept of modeling to provide information on which to base decisions is not new, although each model has unique attributes that are dependent on the environment to be modeled. The cost model developed for use in the USDA Applications Test System (ATS) environment was designed using basic cost accounting principles integrated with unique cost attributes. The model is a multiple of major cost elements comprised of interrelated components that contribute directly or indirectly to the total estimated costs. These major cost elements have been categorized into capital investments and operational costs and summarized into the standard government accounting classification object classes.

The model provides a tool for management to analyze potential impacts of alternative scenarios in a timely and efficient manner. The initial use of the model was to provide estimates of the resource requirements, investments, and operational costs associated with a future operational USDA crop assessment program. The model has been continually modified to meet changing requirements and presently provides investment and operating costs by designated scenarios, personnel staffing reports, budget projections by decision package, and required automatic data processing (ADP) information for OMB reports.

METHODOLOGY

The information generated from the model can be presented in various ways, depending on the intended use. The output formats were dictated by an analysis of the various users and consideration of the user's purpose for requiring the data. Two major economic considerations are reflected in the output from the model. The first is the manner in which to present a 10-year cost projection encompassing a system life of 8 years. The second consideration is

^aU.S. Department of Agriculture, Houston, Texas.

that of the "sunk cost" concept. The rationale used in adapting to these considerations is presented in respective order.

One method of presenting the cost projections is the accounting concept of depreciation, which amortizes the cost of capital investment over the life of the system. This is viewing the investment as a pre-paid operating expense; however, a major disadvantage is that this does not reflect the projected actual cash flow in respect to time. Another method is the "present value" rule, which equates the future capital and operating expenditures to the present-day value of dollars. The technique of discounting the future cash flow with respect to the time incurred at an appropriate rate of interest is used to make the adjustments to equate present value.

In accordance with the OMB requirements, the latter method was used to derive the present value of resource costs over the 10-year lifespan of the proposed production system. The required discount interest rate of 10 percent will be used (ref. 1). The present value has been calculated and is reflected in the summary and detailed reports.

The question of "sunk cost" lies in the definition and adaptation as it applies to the environment being modeled. Sunk costs are nonrecoverable resources that have been consumed as the result of a prior decision and have no direct operational benefit (ref. 2). Sunk costs are not altered by a change in the level or nature of an activity and have no bearing on current investment decisions.

The utilization of satellite remote sensing is technology oriented, and the development of this technology is so dynamic that extensive research, followed by application development, is necessary to exploit potential capabilities. Therefore, the costs related to LACIE research and development and Landsat are considered sunk costs and were not included in costing the future USDA system. The costs associated with the application development and test phases were included since the techniques, procedures, capabilities, and equipment would be of direct benefit to the establishment of a future operational system.

ASSUMPTIONS

In order to establish model requirements, assumptions were made to guide the collection and evaluation of pertinent data. These assumptions were used

in the model development to guide the inclusion and manipulation of the various cost factors.

General

1. A timespan of 10 years was used, representing the procurement of hardware with respect to the phase-in of the operational system and the remaining life expectancy of the system.

2. The LACIE costs are classed as sunk costs and therefore are not included in the total cost for the operational system.

3. Costs associated with the procuring and launching of a satellite are not to be included in the total cost. However, the cost of the product (digital image data) as provided by NASA is included in the total cost.

4. Current General Services Administration (GSA) facility rental rates are used for each potential location.

5. Departmental and agency budgeting policies were followed to derive various cost factors used in the resources calculations.

6. Personnel salaries are projected based on actual and projected positions and will be inflated 5 percent each year for cost-of-living increases.

Hardware

The required computer-related hardware will be purchased.

Software

1. Operating system software will be purchased.

2. The application programs will be developed and implemented as a joint effort by contractors and USDA personnel.

3. The conversion programs will be developed and programed by USDA personnel.

Data Base

1. The design and implementation of the data base will be accomplished by USDA personnel.

2. The digital image processing system design provides for one or more resident geographically oriented data bases.

Personnel

1. Total manpower requirements will be dictated by management and USDA ceiling limitations.
2. Operational manpower requirements will be assessed based on hardware configurations.
3. Startup personnel will be fully trained in the experimental environment and transferred to the production system, thus eliminating consideration of major training costs.

Support Services

1. The receiving station and preprocessing of satellite data to USDA requirements will remain at the NASA Goddard Space Flight Center (GSFC).
2. The GSFC will provide imagery data in accordance with USDA requirements.

Facilities

1. The operational system, equipment, and personnel will be located in a USDA facility.
2. Facilities will require a site-preparation charge.
3. Security and utility services will be accounted for in the facility rental rates.
4. Charges for utilities for second- and third-shift operations will be based on trends of actual charges incurred by the existing USDA computer facilities.

ENVIRONMENT

Initial model development and operation was performed using a Digital Equipment Corporation (DEC) computer 11-45. Since the procurement of a DEC 11-70 by USDA, the model has been transferred and is operational on the DEC 11-70. The FORTRAN programming language was used because it lends itself to the concepts of modular programming through the use of subroutines and is more efficient in data manipulation and calculation.

MODEL DESIGN AND DEVELOPMENT

The approach for design and development of the cost model was to identify the cost categories in the form of stated objectives. The objectives are a series of cost elements that, when combined, provide the

total cost. Figure 1 provides a graphic view of this statement.

Each cost element consists of components that contribute directly or indirectly to the costs. These components are identified, the interrelationships are determined, and the components are formulated into a model.

The model has been developed to process cost trade-offs dependent on alternative management decisions and to assess cost variations resulting from incorporation of new technology, optional system configurations, changes in volume of meteorological and satellite imagery data to be processed, and frequency of processing reports. The resulting reports from the model provide the data to derive a range of expected costs.

The objectives which formed the base for the model are the major elements that contribute to the cost of the system. When reported, they are grouped into investment and operating costs. The cost categories are identified as Hardware, Software, Conversion, Data Base, Relocation Expenses, Personnel, Facilities, ADP Services, Support Services, Research and Development, Administrative Support, and Other.

The basic concept of the model is for each major cost element (stated objective) to perform as a separate program in calculating costs. Each program contains data dependency relationships, algorithms for data calculation, and predefined interrelationships between cost element programs. The relationship of one cost element to another within the model as they

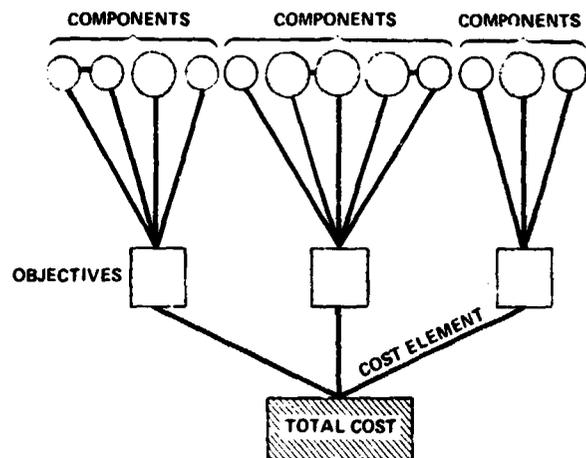


FIGURE 1.— Design and development of the cost model.

provide results to a summary report is depicted graphically in figure 2.

Data are input to the model via computer terminal, although data may or may not be entered for each cost element. The baseline data are maintained in the model's data file.

Each program (cost element) extracts the appropriate data from the file and performs predefined functions. Some data are passed from one program to another and are dependent on a predefined relationship, thus providing the inputs necessary for the receiving program to perform its calculations. Tables are used to test alternative assumptions and to provide cost factors.

The results from each cost-element program are summarized into investment and operating costs by year. In addition, the yearly costs are discounted to present value and are summarized in a report. Detailed procedures for the development of the model are found in a USDA LACIE document, "Approach to Cost Analysis" (ref. 3).

The results obtained from the model (1) are used to assess and influence the design and development aspects of the USDA ATS; (2) provide management with a tool that can increase the competence of management decisions; (3) guide management in decisions on scheduling equipment procurement; and (4) are used to assess and influence future manpower and budget planning.

ALTERNATIVE CONSIDERATIONS

The total costs are based on the combination of

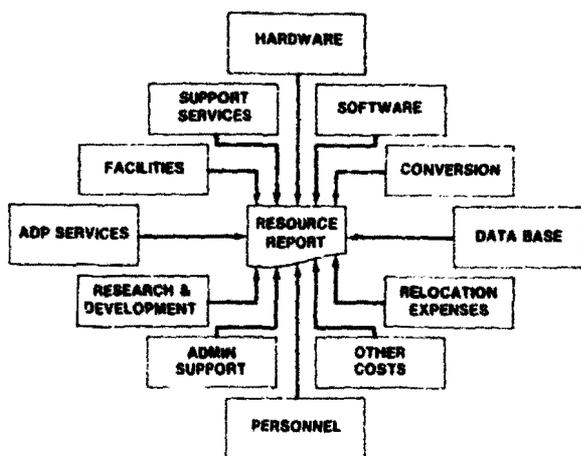


FIGURE 2.— Cost model.

the cost elements bound by the stated assumptions. The sensitivity of these cost elements as they affect the total cost is tested through alternatives. Each alternative represents some degree of impact on the costs. Several of these alternatives are presented here to provide an understanding of model capabilities.

Alternative comparison capability is provided through a Compare Routine. This routine compares the Summary File created by the model for various alternatives against a designated base file and outputs a summary deviation report by cost element. Several tables provide cost factors and algorithms for identified alternatives.

Alternative hardware configurations are tested through the establishment of a file for each configuration. The model is then run for each configuration, and the summary totals are input to a Compare Routine. This routine prepares a report on the deviations from the base configuration as determined by the system design personnel.

Alternative personnel-manpower approaches are tested by varying the numbers and types of positions, creating a file for each alternative. The alternatives for personnel are closely associated with the alternative hardware configuration and management decisions on the extent of goals and countries to be monitored.

Other alternative considerations by cost element are shown in the form of a decision tree.

Software

Costs vary based on the method of procurement, as shown in figures 3 and 4.

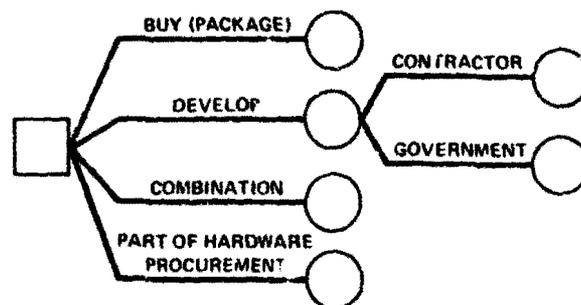


FIGURE 3.— Procurement of systems software.

Support Services

Costs are impacted based on the method of acquiring source data. The Alternatives Table provides cost factors and algorithms regarding method of communication and cost of data. Figures 5, 6, and 7 depict the various types of support services considered.

Facilities

Costs for facilities vary depending on the potential location and extent of modifications (fig. 8). The Alternatives Table provides the GSA lease rates for six different geographical locations. The table can be updated at any time, hence not confining it to any particular location.

MODEL CONSTRUCTION

The model is constructed of 12 separate cost modules linked by a main summary program module. In addition, tables related to alternative approaches are called by the various modules to provide data for the calculation of alternative costs. Subroutine modules provide the data manipulations, calculations, and outputs for the budget projections, personnel staffing reports, and comparison of costs for alternatives. Figure 9 is a simplified flow diagram of the model. The government accounting class codes are incorporated for budget class determination through direct entry or internal programing (ref. 4).

MODEL OPERATION

Each of the cost-element modules has its own data files plus any additional files passed to the module from another module. The various module files are updated based on a scenario to be tested. To provide an understanding of the operation of the cost-element modules, each will be discussed in relation to a simplified data flow diagram and the stated objective of that module. Examples of the detailed outputs from the modules are found in appendix A. All modules output detailed data and summary totals to the Budget and Summary Files, respectively. Appendix B provides examples of the Investment and Operational Cost Summary, Detail Budget Projection

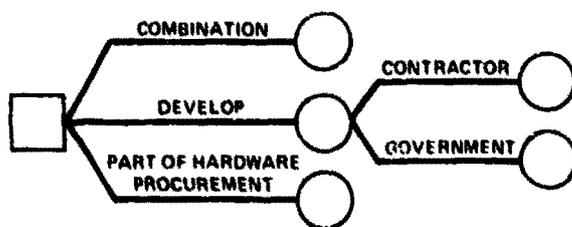


FIGURE 4.— Procurement of application software.

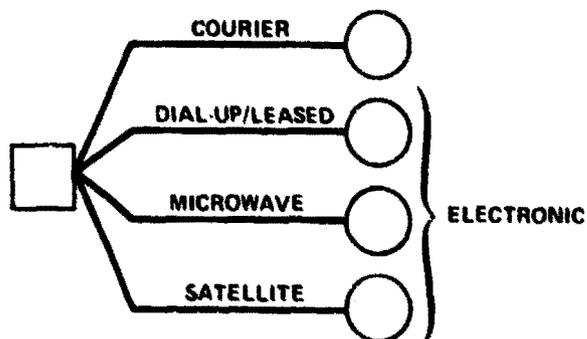


FIGURE 5.— Types of communication media available.

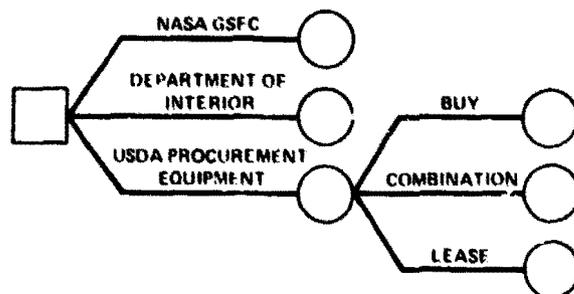


FIGURE 6.— Acquisition of preprocessed satellite data.

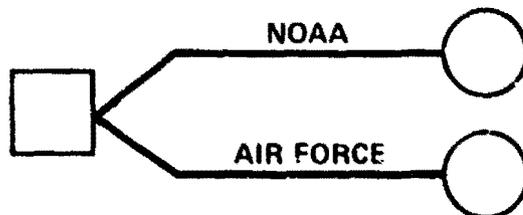


FIGURE 7.— Acquisition of meteorological data.

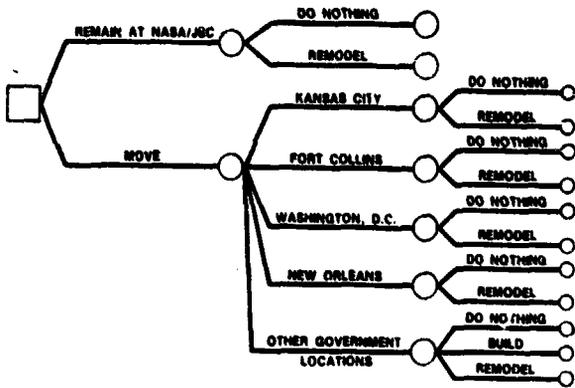


FIGURE 8.— Location/modification alternatives.

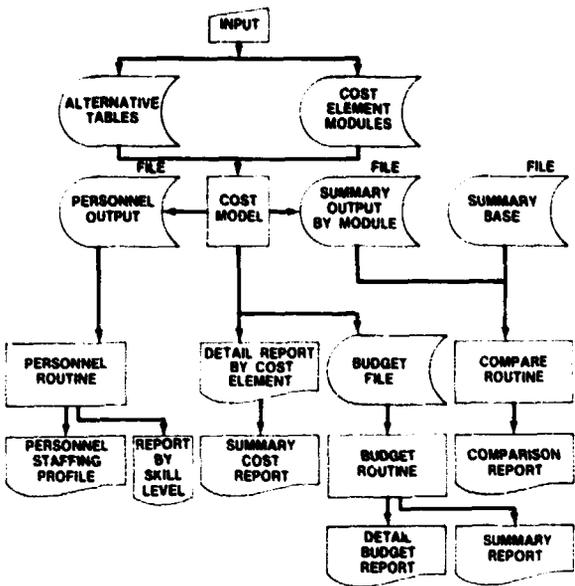


FIGURE 9.— Simplified flow diagram of cost model.

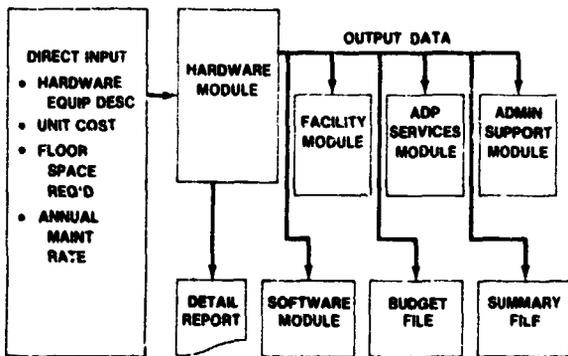


FIGURE 10.— Hardware Module.

Report, Budget Projection Summary Report, Personnel Staffing Profile Report, and Skill Level Summary Report.

Hardware Module

The objective of the Hardware Module (fig. 10) is to identify hardware components to be procured by year and calculate the total cost per year, total overall cost, and other operational impacts. The input of alternative hardware component configurations allows for a cost-effective analysis between configurations and their impacts on operations. Basic data related to the hardware components are input by year of scheduled acquisition and are processed, generating a detailed hardware cost report and output files for use by the Facility, Software, ADP Services, and Administrative Support Modules.

Software Module

The objective of the Software Module is to identify the type of software and calculate the total software costs based on hardware to be procured and on defined software requirements. If application software requirements are not defined, then the module calculates cost based on a percentage of the hardware costs for that year. A report is generated containing detailed costs by year and summary costs.

Conversion Module

The objective of the Conversion Module is to record and provide costs associated with the conversion of data files and application software from LACIE to the USDA environment. This module provides for direct input of defined conversion requirements and estimated costs. The output is a detailed listing of the costs and summary totals.

Data Base Module

The objective of the Data Base Module is to record and provide costs associated with the development, implementation, data collection, and purchase of Data Base Management software programs. This module provides for direct input of defined functions and estimated costs. The outputs include a detailed listing of costs by function and year, plus summary totals.

Personnel Module

The objective of the Personnel Module (fig. 11) is to identify the skill levels required and salaries associated with each position within those skill levels. The manpower level and skills are analyzed based on functions to be performed and the performance goals as defined in the Management Plans (refs. 5 and 6). Each position, grade, step, and salary is input for the year required. The program calculates the succeeding year's salary based on a cost-of-living percentage increase. Promotions are accounted for by entering the new salary in the year of the anticipated promotion. A detailed listing is printed and detailed data are output to a personnel subroutine, which provides a detailed staffing profile report and a summary of positions by skill level. Totals are passed to the files or modules for further calculation of cost impacts.

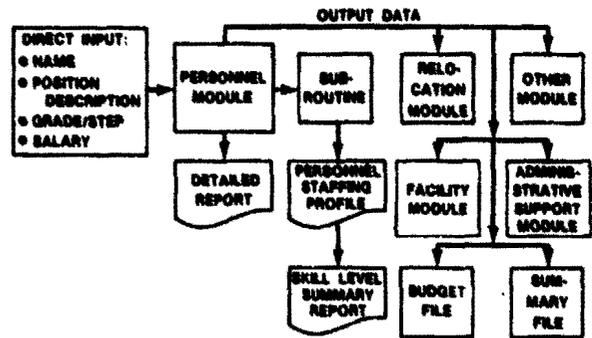


FIGURE 11.— Personnel Module.

Other Investments Module

The objective of the Other Investments Module (fig. 12) is to establish other initial costs incurred in implementing an operational system. Detailed cost components identified in this module include telephone installation, furniture procurement, site preparation, etc. Telephone installation and furniture costs are derived from algorithms using data passed from the personnel module and influenced by the Alternatives Table. Input of known costs can be made through direct entry to the module file.

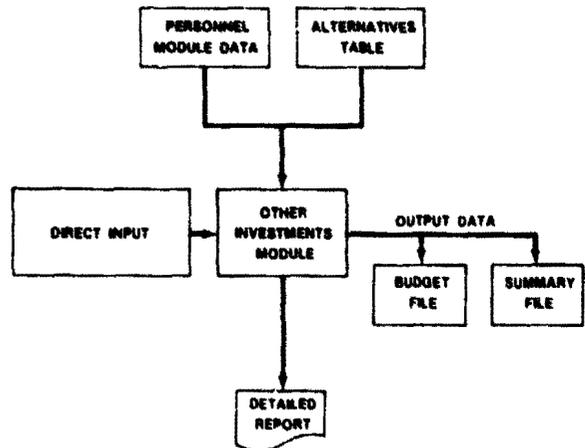


FIGURE 12.— Other Investments Module.

Relocation Module

The objective of the Relocation Module (fig. 13) is to establish the cost to relocate personnel and equipment, depending on the number of personnel and alternative actions. A table is updated to provide the alternatives with regard to time of relocation and number of personnel to be relocated. The cost associated with the equipment relocation is a direct input. The calculations performed involve algorithms utilizing data from the Personnel Module, the Alternatives Table, and cost factors derived from analysis of current moving costs. The outputs include a detailed report of relocation costs and detailed and summary data passed to the Budget and Summary Files, respectively.

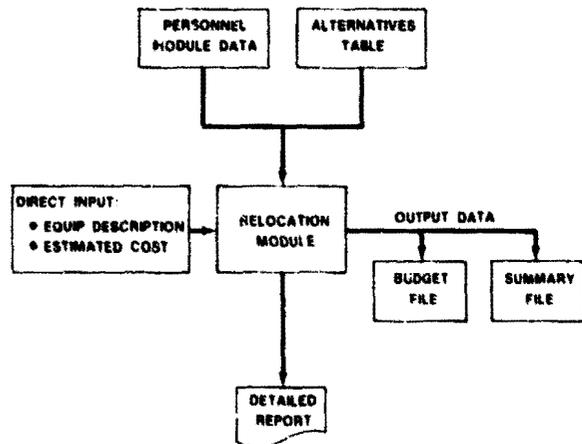


FIGURE 13.— Relocation Module.

ADP Services Module

The objective of the ADP Services Module (fig. 14) is to establish the cost of services directly related to the support of ADP operations. Primarily, these costs are for equipment maintenance and rental. Equipment rental data are entered directly into the file, whereas the calculation of maintenance costs is a function of data passed from the Hardware Module and accumulated for a total yearly cost plus the added cost for each succeeding year of the equipment life. Since facility space is impacted by rental equipment, physical space data are passed to the Facility Module for further processing. A detailed report is printed and the respective data are passed to the Budget and Summary Files.

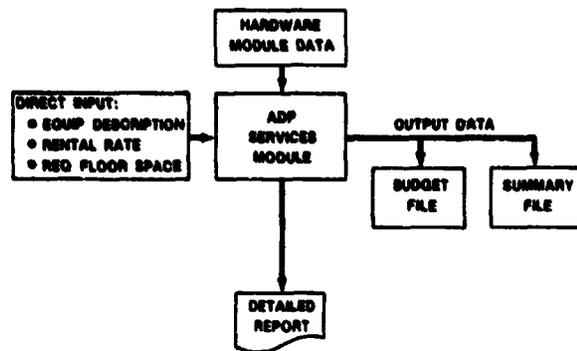


FIGURE 14.— ADP Services Module.

Facility Module

The objective of the Facility Module (fig. 15) is to establish the size and cost of facilities required to house the personnel, equipment, and work areas associated with a USDA environment. Three modules provide input data to this module, which has direct-entry capability. Additionally, the Alternatives Table is accessed to obtain dollar rates for various locations based on type of space. The calculation of the space for personnel is based on GSA allowances. Algorithms are the basis for establishing costs using the table factors and additional-shift utility allowances. A detailed report is printed containing the total square feet of facility required by type of space and the cost for that space.

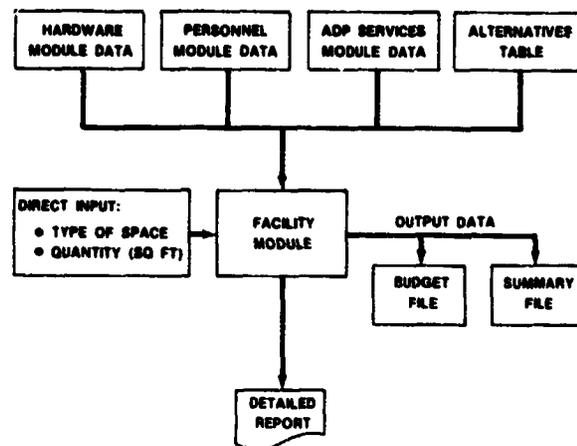


FIGURE 15.— Facility Module.

Support Services Module

The objective of the Support Services Module is to identify, record, and calculate costs incurred for services supplied by other governmental organizations and private enterprise in support of crop assessment operations. Two key alternatives impacting costs are tested in this module. The utilization of satellite communications versus courier service represents significant variances in costs. The calculation of satellite communication costs is dependent on the volume of data and time of transmission. The volume of satellite digital data required also impacts the cost of buying the data and is based on workloads associated with each geographic area to be monitored. Therefore, algorithms using the data volume, which

is a direct input, calculate both the cost of buying satellite data and the cost of transmitting the data. The Alternatives Table supplies the algorithm factors, depending on the alternative to be tested. Other capabilities of this module include direct entry of known required services and associated costs. The output is a detailed report with data passed to the Budget and Summary Files.

Research and Development Module

The objective of the Research and Development Module is to record and provide costs associated with defined research and development functions. Input is direct through creation of several files, and alternative cost approaches are integrated into the total costs.

Administrative Support Module

The objective of the Administrative Support Module (fig. 16) is to identify, calculate, and/or provide the total costs associated with the administrative support functions of a USDA operational environment. Key cost components established in this module are personnel benefits, travel, training, supplies, telephone, and administrative overhead costs. Factors used in formulas to derive the costs are based on historical trend data of the department. Personnel benefits are derived as a percentage of the total personnel salary costs passed from the Personnel Module. Training, supplies, and telephone costs are a function of the number of personnel as passed from the Personnel Module. The administrative overhead costs are calculated on the total operational costs from the Summary File and then added to the total cost. Other known administrative costs may be input directly.

The main summary program is the controlling program of the cost model. Through input data, it determines which table files to access and which files to open for output and calls the subroutine to process the data.

The main summary program accepts the input data, sets up the calling parameters, calls the appropriate subroutine to process the data, and stores any returned data or parameters. The validity checking is performed in the subroutine. After all the input data have been processed, the main summary program produces the summary report. The source code for the main summary program is given in appendix C.

The main summary program does not use overlays, since it and the associated subroutines execute in 26 000 bytes of core. The main summary program uses standard linkages of the CALL and parameter list to interface with the subroutines.

SUMMARY

The utilization of the cost model has provided data to OMB, senior USDA Management, and Project Management and major inputs to the budget process for 1977, 1978, 1979, and 1980. The modular concept of the model simplified its design, implementation, and operation. Approximately 3 man-months were involved in the design, collection of cost factor data, and development of the interrelationships, algorithms, and alternative test capa-

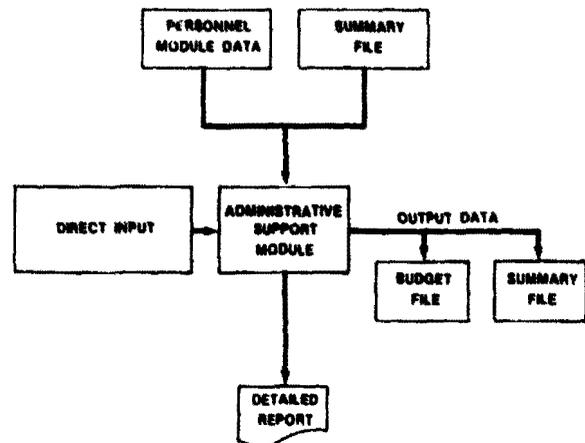


FIGURE 16.— Administrative Support Module.

bilities; programing, implementation, and testing required 3 man-months. The model was operational by July 1976 and provided the detailed ADP cost information for the OMB requested report on projected expenditures for fiscal year (FY) 1977. A special OMB presentation in September 1977 required a detailed analysis of costs to be expected in an operational system. The model was used to generate the data and provided a range of expected costs dependent on alternative management decisions. The model derived the information, together with a detailed Resource Analysis Report which successfully answered OMB's questions concerning costs (ref. 6). The Resource Analysis Report was updated in 1977 and amended in 1978.

During 1977, it became apparent that the extension of the cost model into the budget area would expedite and increase the accuracy of the budget projections. Approximately 2 man-months of design and programing were required to implement the budget routines into the model. The model was used to assess the cost impacts of various hardware design configurations and influenced the selection of a cost-effective design and related specifications used in the procurement of the current system configuration. The implementation of the budget routine categorizes and accumulates the cost components into government accounting classes and provides both detailed and summary budget projections. The budgets submitted for 1978, 1979, and 1980 were directly calculated by the model. With the initial implementation of zero-base budgeting (ZBB) for FY 1979, the model's alternative test capabilities easily

provided the budget levels for the ZBB decision packages.

From early 1977 through the present, the model has been used to assess cost impacts and provide personnel staffing profiles associated with alternative management decisions. It has been instrumental in adding competence to the management decisions in budgeting, project goals, manpower planning, and investments in procuring equipment, software, and support services.

REFERENCES

1. Discount Rates To Be Used in Evaluating Time-Distributed Costs and Benefits. Office of Management and Budget, Circular A-94, Mar. 27, 1972.
2. Glossary for Public Program Analysis. Prepared by the Association for Public Program Analysis under contract to the U.S. Department of Transportation, 1973.
3. Approach to Cost Analysis. Economic Evaluation for the LACIE-Type Crop Estimation System, Vol. II, USDA LACIE Project document, July 19, 1976.
4. Financial Accounting System, Coding Manual FY 1978. FAS Handbook 3, 1978.
5. 1978 Research and Application Test Management Plan. USDA LACIE Project Document 2-MO, Apr. 17, 1978.
6. 1979 Research and Application Test Management Plan. USDA LACIE Project Document 3-MO, Apr. 17, 1978.
7. 1977 Resource Analysis Report. USDA LACIE Project Document, Sept. 1976.

Appendix A

Examples of Detailed Module Output

DETAILED REPORT - HARDWARE					
NAME	QTY	PRICE	CURRENT	FY 2	FY 10
DISK CONTROLLER	15	9.1	9.1	9.1	9.1
DISK UNITS	34	15.0	139.6	139.6	30.0
MAG TAPE CONTROLLER	14	12.5	12.5	12.5	12.5
MAG TAPE DRIVE	25	14.0	28.0	28.0	14.0
GRAPHIC TERM/COPIER	12	7.5	7.5	7.5	0.0
CARD RDR/PUNCH	5	15.0	15.0	15.0	15.0
CARD READER	8	5.6	0.0	0.0	0.0
LINE PRINTER 1200LP	1	31.5	31.5	0.0	0.0
.....SUBTOTAL.....			182		
TOTAL HARDWARE			727.0	727.2	302.1
PRESENT VALUE					

DATA GENERATED FOR EXAMPLES ONLY

DETAILED REPORT - SOFTWARE				
NAME	CURRENT	FY 2	FY 9	FY 10
COMPUTER OS-SW	25.0	25.0	0.0	0.0
ANALYST STATION SW	38.0	70.0	0.0	0.0
COMMUNIC INTERFACE	0.0	20.0	0.0	0.0
HOST COMP INTERFACE	0.0	0.0	0.0	0.0
HOST INTERFACE	0.0	0.0	0.0	0.0
ARRAY PROC-SW	20.0	20.0	0.0	0.0
APPLICATION SOFTWARE	225.9	387.0	0.0	0.0
SOFTWARE DOCUMENTATION	20.7	20.7	0.0	0.0
78 BUDGET ADJ.	0.0	0.0	0.0	0.0
TOTAL SOFTWARE			705.6	263.7
PRESENT VALUE		400.0	329.5	111.8

DATA GENERATED FOR EXAMPLES ONLY

DETAILED REPORT - CONVERSION				
ITEM	CURRENT	FY 2	FY 9	FY 10
MANPOWER RESOURCES	11.0	20.0	0.0	0.0
CPU TEST TIME	0.0	10.0	0.0	0.0
TEST STR DEVICE	0.0	5.0	0.0	0.0
78 BUDGET ADJ.	0.0	0.0	0.0	0.0
TOTAL CONVERSION			0.0	0.0
PRESENT VALUE		35.0	0.0	0.0

DATA GENERATED FOR EXAMPLES ONLY

DETAILED REPORT - DATA BASE				
NAME	CURRENT	FY 2	FY 9	FY 10
DATA BASE MGMT PKG	0.0	25.0	0.0	0.0
DB MOD SW	0.0	0.0	0.0	0.0
78 BUDGET ADJ.	0.0	0.0	0.0	0.0
DATA BASE COST			0.0	0.0
PRESENT VALUE			0.0	0.0

DATA GENERATED FOR EXAMPLES ONLY

DETAILED REPORT - PERSONNEL				
ITEM	CURRENT	FY 2	FY 9	FY 10
JONES, PATRICA	4.2	8.9	11.3	11.5
Z (TEMP)	0.0	32.1	40.9	41.3
LACIE ASSIGNED	0.0	0.0	0.0	0.0
MULTICROP	0.0	0.0	2019.6	2060.2
MULTICROP 2	0.0	0.0	204.0	418.0
LESS DEVEL CTRY	0.0	0.0	208.0	208.0
TOTAL PERSONNEL	64.2	77.0	2244.1	2150.7
LAPSE TIME ADJUSTMENT	56.0	35.0	0.0	239.0
TOTAL PERSONNEL COST			4806.3	5072.3
PRESENT VALUE		1448.3	2244.1	2150.7

DATA GENERATED FOR EXAMPLES ONLY

DETAILED REPORT - RELOCATION				
ITEM	CURRENT	FY 2	FY 9	FY 10
EQUIP PACK & SHIP	0.0	0.0	0.0	0.0
RELOCATION BENEFITS	31.5	58.5	0.0	0.0
RELO HOUSEHOLD GOODS	17.5	0.0	0.0	25.0
TOTAL RELOCATION			238.0	70.0
PRESENT VALUE		82.7	111.1	29.7

DATA GENERATED FOR EXAMPLES ONLY

DETAILED REPORT - OTHER COSTS

ITEM	CURRENT	FY 2	FY 9	FY 10
TELEPHONE INSTALLATION	0.2	0.4	1.0	0.3
FURNITURE COSTS	8.7	12.3	32.3	8.8
DISK PACKS	7.0	7.0	7.0	0.0
LANDSAT DATA S/U CHG	0.0	0.0		0.0
DATA GENERATED FOR EXAMPLES ONLY				
FILM LIGHT TABS	0.0	0.0	0.0	0.0
TOTAL OTD DATA	7.0	103.2	48.3	9.8
PRESENT VALUE	38.4	93.8	18.8	4.2

DETAILED REPORT - ADP SERVICES

EQUIPMENT	CURRENT	FY 2	FY 9	FY 10
MAINTENANCE	58.9	146.0	1148.4	1175.1
RENTALS				
FILM IMAGE PROCESSOR	0.0	0.0	0.0	0.0
KEYPUNCH/VERIFIER	0.0	0.0	8.1	8.1
WORD PROC EQUIP	4.0	9.0	34.4	
78 BUDGET ADJ.	0.0	0.0		0.0
TOTAL SERVICES			1174.5	1204.2
PRESENT VALUE		71.3	548.5	510.8
DATA GENERATED FOR EXAMPLES ONLY				

DETAILED REPORT - FACILITIES

ITEM	CURRENT	FY 2	FY 9	FY 10
COMPUTER SPACE RENTAL	6.2	14.2	160.3	166.7
90 FT	712.6	1618.8	18310.0	18333.8
OFFICE SPACE RENTAL	32.8	33.5	242.3	253.1
90 FT	4850.0	4850.0	33800.0	35100.0
STORAGE SPACE RENTAL	0.0	0.0	6.0	6.0
90 FT	0.0	0.0	123.0	0.0
TOTAL RENTAL	38.7	38.7	249.6	424.8
TOTAL SQUARE FT			53125.0	55308.8
PRESENT VALUE	98.7	43.4	190.8	180.1
DATA GENERATED FOR EXAMPLES ONLY				

DETAILED REPORT - RESEARCH AND DEVELOPMENT

ITEM	CURRENT	FY 2	FY 9	FY 10
CONTRACTS	0.0	0.0	0.0	50.0
REIMBURSEABLES	154.2		150.0	150.0
LACIE REIMB			150.0	0.0
MULTICROP R&D	0.0	0.0	0.0	0.0
LDC R&D	0.0	0.0	475.0	475.0
GROUND TRUTH DATA	0.0	0.0	50.0	50.0
TOTAL R&D	154.2		875.0	725.0
PRESENT VALUE	154.2	110.0	408.8	307.4
DATA GENERATED FOR EXAMPLES ONLY				

DETAILED REPORT - SUPPORT SERVICES

	CURRENT	FY 2	FY 9	FY 10
IMAGERY DATA	49.0	98.0	271.0	271.0
METEOROLOGICAL DATA	0.0	50.0	80.0	80.0
ANCILLARY DATA	0.0	3.2	0.0	0.0
COMMUNICATIONS	0.0	240.2	650.1	650.1
HOST COMPUTER USAGE	0.0	0.0	95.0	95.0
OTHER				
TOTAL	0.0	0.0	0.0	0.0
TOTAL SERVICES			1096.1	1096.1
PRESENT VALUE	49.0	349.9	511.9	464.7
DATA GENERATED FOR EXAMPLES ONLY				

DETAILED REPORT - ADMINISTRATION

ITEM	CURRENT	FY 2	FY 9	FY 10
PERSONNEL BENEFITS	122.6	167.1	504.6	532.6
ADMINISTRATION SUPPORT	150.0	150.0	688.3	682.7
TRAVEL	75.0	75.0	92.0	92.0
SUPPLIES	10.3	12.3	35.0	36.5
TRAINING	75.0	45.0	237.4	248.0
OFFICE EQUIPMENT	0.0	0.0	0.0	0.0
COMMUNICATIONS	10.8	11.2		4.2
OTHER				
TOTAL ADMINISTRATION			1219.8	1262.4
PRESENT VALUE	383.7	98.8	569.8	535.2
DATA GENERATED FOR EXAMPLES ONLY				

Appendix B

Examples of Detailed Data Summary Reports

INVESTMENT AND OPERATIONAL COST SUMMARY						
INVESTMENT COSTS	CURRENT	FY 2	FY 3	FY 10	TOTAL	P.V.
HARDWARE	700.0	800.0	1778.6	712.6	1886.3	12244.3
SOFTWARE	326.6	440.0	840.2	263.7	7432.4	4993.7
OPERATIONAL COSTS						
PERSONNEL	1167.6	1501.1	2100.0	5072.3		19612.4
ADMINISTRATIVE	323.7	108.7	630.8	1382.4		4782.9
TOTAL						
PV		520.4	5784.8	8681.0		40729.0
GRAND TOTAL		6476.2	10340.9			
TOTAL PV		3121.8	7001.4	4696.5		

DETAIL BUDGET PROJECTION REPORT				
	CURRENT	FY 2	FY 3	FY 10
25 0 ADMIN SUPPORT	180.0	180.0	180.0	150.0
25 1 SITE PREPARATION	0.0	26.0	78.0	0.0
25 TOTAL OTHER SERVICES 474.2 655.5 626.8 627.3				
26 TOTAL SUPPLIES 4.3 7.0 2.9				
31 TOTAL EQUIPMENT 111.1 521.7 158.2 0.0				

BUDGET PROJECTION SUMMARY REPORT					
TOTAL	CURRENT	FY 2	FY 3	FY 10	TOTAL
11 - PERSONNEL COMPENSATION	136.4	179.5	263.6	436.5	3512.5
12 - PERSONNEL BENEFITS	32.3	23.3	44.6	45.6	440.7
21 - TRAVEL AND TRANSPORTATION OF PERSONS	25.0	27.5	30.0	15.0	278.0
22 - TRANSPORTATION OF THINGS		3.5	10.0	0.0	40.0
23 - RENTS, COMMUNICATIONS		34.1	53.7	681.0	
25 - OTHER	474.2	655.5	626.8	527.3	6120.7
26 - SUPPLIES	4.3	8.4	7.0	2.8	67.1
31 - EQUIPMENT	111.1	521.7	158.2	0.0	2753.9
GRAND TOTAL	816.7	1446.1	1166.1	1061.2	13774.9

PERSONNEL STAFFING PROFILE REPORT					
NAME	POSITION	GRADE	CURRENT	FY 2	FY 10
PUBLIC, JOHN O	PROJ MGR	301 15-1	38.0	40.7	52.4
SMITH, MARY	EXEC SEC	318 08-1	11.1	11.9	15.3
STOLER, RICHARD AG ECON 110 14-3 32.5 36.1 45.2					
2 (VACANT) AG STAT R/ST 1529 09-1 0.0 0.0 19.4					
2 (VACANT) AG STAT 0.0 32.4 42.0					
2 (VACANT) 1 0.0 0.0 19.4					
2 (VACANT) AGRONOMIST 11 09-1 0.0 0.0 18.1					
2 (VACANT) AGRONOMIST 471 08-1 0.0 0.0 20.8					
2 (VACANT) AG ECON C/S 110 08-1 0.0 0.0 19.5					

SKILL LEVEL SUMMARY REPORT					
POSITION		CURRENT	FY 2	FY 8	FY 10
110	FILLED	1	1	1	1
	VACANT	0	0	7	7
135	FILLED	0	0	0	0
	VACANT	0	0	0	0
150	FILLED	0	0	0	0
	VACANT	0	0	0	0
9000	FILLED	2	2	2	2
	VACANT	0	0	0	0
TOTAL	FILLED	3	3	3	3
	VACANT	1	2	13	13

Appendix C

Source Code for Main Summary Program

```

C  CSTSRM IS THE MAIN CALLING PROGRAM
C  FOR THE COST SYSTEM
C
C  SUB MURST
C
      DIMENSION MNCOST(10),INOTE(78),TBL(9),CVCOST(10),MRCOST(10),
*  OTCOST(10),ADCCOST(10),RDCOST(10),SSCOST(10),AMCOST(10),
*  TUTOP(10),TIPVAL(10),TIPVAL(10),T2PVAL(10),GFCOST(10),
*  ADSOFT(10),TSOFT(10),MNCOST(10),BWCOST(10),DBCOST(10),
*  NPUS(10),TUCOST(10),TICOST(10),FACOST(10),PRCOST(10),IDATES(5),
*  ICOUR(6),TBLA(10),TBLF(10),TBL0(10),IPRTCM(6),IDUM(6)
      INTEGER COMCD
      REAL MNCOST
      DATA ICOUR/'CU','UN','IB','R','MU','DB'/
      DATA IDUM/'DU','RS','AT','M','OD','E'/
      CALL ASSIGN (5,'LP:')
      CALL ASSIGN (6,'T1:')
      CALL ASSIGN (1,'LOC.CD')
      CALL ASSIGN (9,'PERS.UOT')
      CALL ASSIGN (10,'ACCT.CD')
      READ (1,1001) LOCT,COMCD,INOTE
1001  FORMAT (211, /6A1)
      READ (1,1002) TBLA
1002  FORMAT (10F4.0)
      READ (1,1002) TBLF
      READ (1,1002) TBL0
      CALL DATE (IDATES)
      IF (COMCD .NE. 1) GO TO 3
      DO 3 J=1,6
          IPRTCM(J) = ICOUR(J)
3      CONTINUE
      IF (COMCD .NE. 2) GO TO 4
      DO 4 K = 1,6
          IPRTCM(K) = IDUM(K)
4      CONTINUE
      IF (LOCT .NE. 9) IPRCNT = LOCT
      GO TO (10,20,30,40,50,60),LOCT
      IF (LOCT .EQ. 9) GO TO 5
      WRITE (6,2001) LOCT
2001  FORMAT (' WRONG LOCATION CODE ',11)
      STOP
5      IPRCNT = IPRCNT + 1
      GO TO (10,20,30,40,999),IPRCNT
20      CALL ASSIGN (2,'KC.TBL')
      WRITE (5,701) IPRTCM
701  FORMAT ('1','KANSAS C11 ',6A2)
      ILUN = 2
      GO TO 70
10      CALL ASSIGN (7,'FICUL.TBL')
      WRITE (5,702) IPRTCM

```

```

702  FORMAT ('1','FORT COLLINS ',0A2)
      ILUN = 7
      GO TO 70
 40   CALL ASSIGN (4,'WASH.TBL')
      WRITE (5,703) IPRICH
703  FORMAT ('1','WASH D.C. ',0A2)
      ILUN = 4
      GO TO 70
 30   CALL ASSIGN (3,'NEW.D.TBL')
      WRITE (5,704) IPRICH
704  FORMAT ('1','NEW ORLEANS ',0A2)
      ILUN = 3
-----
      GO TO 70
 50   CALL ASSIGN (2,'OTHER1.TBL')
      ILUN = 2
      GO TO 70
 60   CALL ASSIGN (2,'OTHER2.TBL')
      ILUN = 2
 70   READ (ILUN,1003) TBL
1003  FORMAT (F5.2,F4.1,F5.3,F4.1,F7.3,F4.1)
      WRITE (5,2010) IDALS
      WRITE (5,2004) INOTE
2004  FORMAT ('0',78A1)
      CALL HWD (HWCOST,TWHPV,TWNCST,TSOFT,HWCOST)
      CALL SWTR (SWCOST,TSFPV,TSFCST,HWCOST)
      CALL CONVR (CVCOST,TCVPV,TCVCS1)
      CALL DBASE (DBCOST,TDBPV,IDBCST)
      CALL PERSON (PRCOST,IPRPV,IPRCS1,NPUS)
      CALL RECU (RECUST,TRCPV,TRCST,NPUS,IBLU)
      CALL OTHER (OTCOST,TOCPV,TOCST,NPUS,TBLM,IBLU)
      CALL ADPDR (ADCOST,IADPV,IADCST,AUSOFT,HWCOST)
      CALL FAC (FACOST,IFAPV,IFACST,ISOFT,AUSOFT,IBL,NPUS,IBLF)
      CALL RAND (RDCOST,TRDPV,IRDCST)
      CALL SUPSR (SSCOST,TSSPV,TSSCS1,IBL,CUNCU)
      DO 80 J=1,10
          TOTOP(J) = ADCOST(J) + FACOST(J) + RDCOST(J)
          + SSCOST(J) + PRCOST(J)
 80   CONTINUE
      CALL ADMIN (AMCOST,TAMPV,IAMCST,NPUS,ITOTOP,PRCOST,IBLF)
      DO 90 L=1,10
          FICOST(L) = HWCOST(L) + SWCOST(L) + CVCOST(L) +
          * DBCOST(L) + RECUST(L) + OTCOST(L)
          TICOST(L) = PRCOST(L) + ADCOST(L) + FACOST(L) +
          * RDCOST(L) + AMCOST(L) + SSCOST(L)
          GICOST(L) = TICOST(L) + IUCOST(L)
 90   CONTINUE
      CALL PREVAL (IICOST,IIPVAL,GIPVAL,IICST)
      CALL PREVAL (IUCOST,IIPVAL,GIPVAL,IUCST)
      CALL PREVAL (GICOST,IIPVAL,GIPVAL,GICST)
      DO 100 J=5,10
          TSCST = TSCST + IUCOST(J)
100  CONTINUE
      TSCST = TSCST / 6.0
      WRITE (5,2002)

```

```

2002  FORMAT ('1',3X,'PROJECTED RESOURCES FOR THE OPERATIONAL',//
*      '0',T40,'PRODUCTION, AREA, YIELD ESTIMATION SYSTEM',//)
      WRITE (5,2010) IDATES
2010  FORMAT ('0','RUN DATE ',5A2)
      WRITE (5,2003)
2003  FORMAT ('0',4X,'INVESTMENT COSTS',T39,'CURRENT',4X,'FY 2',4X,
*      'FY 3',4X,'FY 4',4X,'FY 5',4X,'FY 6',4X,'FY 7',4X,'FY 8',4X,
*      'FY 9',3X,'FY 10',3X'TOTAL',3X,'P. V. ')
      WRITE (5,2005) HWCOST,THWCST,THWPV,SWCOST,TSFCST,TSPPV,
*      CVCOST,TCVCSST,TCVPPV,DBCOST,TDBCST,TDPPV,RECOST,TRCST,TRPPV,
*      OTCOST,TOTCSST,TOTPPV,TICOST,TICST,TIPVAL,OTPVAL
2005  FORMAT ('0',6X,'HARDWARE',T37,12F8.1,/,
*      '0',6X,'SOFTWARE',T37,12F8.1,/,
*      '0',6X,'CONVERSION',T37,12F8.1,/,
*      '0',6X,'DATA BASE',T37,12F8.1,/,
*      '0',6X,'RELOCATION EXPENSES',T37,12F8.1,/,
*      '0',6X,'OTHER',T37,12F8.1,/,
*      '0',6X,'TOTAL',T37,11F8.1,/,
*      '0',9X,'P V',T37,10F8.1,8X,F8.1,//)
      WRITE (5,2006)
2006  FORMAT ('0',4X,'OPERATIONAL COSTS')
      WRITE (5,2007) PRCOST,TPRPV,ARCOST,TAMPV,
*      ADCOST,TADPV,FACOST,TFAPV,ROCOST,TRDPV,
*      SSCOST,TSPPV,UCOST,TSOST,TZPVAL,GZPVAL
2007  FORMAT ('0',6X,'PERSONNEL',T37,10F8.1,8X,F8.1,/,
*      '0',6X,'ADMINISTRATIVE',T37,10F8.1,8X,F8.1,/,
*      '0',6X,'ADP SERVICES',T37,10F8.1,8X,F8.1,/,
*      '0',6X,'FACILITIES',T37,10F8.1,8X,F8.1,/,
*      '0',6X,'RESEARCH & DEVELOPMENT',T37,10F8.1,8X,F8.1,/,
*      '0',6X,'SUPPORT SERVICES',T37,10F8.1,8X,F8.1,/,
*      '0',6X,'TOTAL',T37,11F8.1,/,
*      '0',9X,'P V',T37,10F8.1,8X,F8.1,//)
      WRITE (5,2008) GTCOST,TTPVAL
2008  FORMAT ('0',4X,'GRAND TOTAL',T37,10F8.1,/,
*      '0',4X,'TOTAL P. V.',T37,10F8.1)
      GO TO (110,120,130,140),IPRCNT
110  WRITE (5,1110),IPRTCH
1110  FORMAT (//,'FORT COLLINS ',6A2)
      GO TO 200
120  WRITE (5,1120),IPRTCH
1120  FORMAT (//,'KANSAS CITY ',6A2)
      GO TO 210
130  WRITE (5,1130),IPRTCH
1130  FORMAT (//,'NEW ORLEANS ',6A2)
      GO TO 220
140  WRITE (5,1140),IPRTCH
1140  FORMAT (//,'WASH D.C. ',6A2)
      GO TO 230
200  CALL ASSIGN (8,'FC.OUT')
      WRITE (8,2020) IPRTCH
2020  FORMAT (' FORT COLLINS ',6A2)
      GO TO 500
210  CALL ASSIGN (8,'KC.OUT')

```

```

WRITE (8,2030) IPRICM
2030  FORMAT (' KANSAS CITY ',6A2)
      GO TO 500
220  CALL ASSIGN (8,'NEW0.OUT')
      WRITE (8,2040) IPRICM
2040  FORMAT (' NEW ORLEANS ',6A2)
      GO TO 500
230  CALL ASSIGN (8,'WDC.OUT')
      WRITE (8,2050)
2050  FORMAT (' WASHINGTON DC',6A2)
500   WRITE (8,2060) MWCOST,THWCST,THWPV,SWCOST,TSFCST,TSFPV,
      * CVCOST,TCVCST,TCVPV,DBCOST,TDBCST,IDBPV,RECCOST,TRECCST,
      * TRPVP,UTCOST,TOTCST,TOTPV,TICOST,TICST,TIPVAL,GIPVAL,
      * PRCCOST,TPRPV,AMCOST,TAMPV,ADCOST,TADPV,FACOST,TFAPV,RCCOST,
      * TRDPV,SSCOST,TSPPV,TOCOST,ISCST,T2PVAL,G2PVAL,GCOST,IPVAL
2060  FORMAT (6(12F8.1/),10(11F8.1/),2(10F8.1/))
      CALL CLOSE (8)
      GO TO 998
998   WRITE (5,2004) INOTE
      IF (LUCT.NE.9) GO TO 999
      REWIND 1
      READ (1,1001) LUCT,COMCD,INOTE
      DO 150 K=1,10
          TSQFI(K) = 0.0
          MWCOST(K) = 0.0
          NPUS(K) = 0
          AUSQFI(K) = 0.0
          PRCCOST(K) = 0.0
          MWCOST(K) = 0.0
          SWCOST(K) = 0.0
          CVCOST(K) = 0.0
          UBCOST(K) = 0.0
          RECCOST(K) = 0.0
          UTCOST(K) = 0.0
          ADCOST(K) = 0.0
          FACOST(K) = 0.0
          RCCOST(K) = 0.0
          TICOST(K) = 0.0
          TOCOST(K) = 0.0
          GICOST(K) = 0.0
          AMCOST(K) = 0.0
          SSCOST(K) = 0.0
          TOTOP(K) = 0.0
          TIPVAL(K) = 0.0
          TTPVAL(K) = 0.0
          T2PVAL(K) = 0.0

```

ORIGINAL PAGE
OF POOR QUALITY

150

CONTINUE

TSCST=0.0

THNCST=0.0

THNPV =0.0

TSFCST =0.0

TSFPV =0.0

TCVCST =0.0

TCVPV =0.0

TUBCST =0.0

TUBPV =0.0

TPRCST =0.0

TPRPV =0.0

TRECST =0.0

TREPV =0.0

TOTCST =0.0

TOTPV =0.0

TADCST =0.0

TADPV =0.0

TFACST =0.0

TFAPV =0.0

TRDCST =0.0

TRDPV =0.0

TAMCST =0.0

TAMPV =0.0

TICST =0.0

G1PVAL =0.0

TUCST =0.0

G2PVAL =0.0

GO TO 5

999

STOP

END

The Application Test System: Experiences to Date and Future Plans

G. A. May,^a P. Ashburn,^a and H. L. Hansen^a

INTRODUCTION

The Application Test System (ATS) was designed to test and evaluate the latest technology in acquisition, storage, retrieval, analysis, and application of remotely sensed data for application feasibility by the U.S. Department of Agriculture (USDA). The purpose of this paper is to describe the ATS analysis component focusing on methods by which the varied data sources are used by the ATS analyst. An integral part of the ATS is the team of USDA multidisciplinary analysts who analyze and interpret varied data sources including remotely sensed data. The ATS analysts have agricultural backgrounds with education and experience in a wide spectrum of disciplines.

Material will be presented in two parts. Analyst training and initial processing of data within the ATS will be discussed first in the section entitled "Experiences to Date." The second section, entitled "Future Plans," will discuss short- and long-term plans for the ATS.

EXPERIENCES TO DATE

LACIE Phase III Activities

During Phase III (1977 crop year), USDA analysts operated and evaluated an interactive computer-linked classification system developed by the LACIE. The system was evaluated in terms of classification accuracy and segment throughput efficiency. USDA analysts gained experience in analyzing Landsat multispectral scanner (MSS) data on an interactive image-processing system. Their image-processing experience played a large role in the

design and implementation of the ATS interactive image-processing system. Many of the analyst-detected inefficiencies in the LACIE system were considered and corrected in the design of the ATS processing system.

The two main components of the LACIE image-processing system were the General Electric IMAGE-100 (I-100) and the Earth Resources Interactive Processing System (ERIPS). The main processing procedure used to analyze data on this system was Procedure 1 (P-1). P-1 was developed to provide estimates of the percentage of a segment devoted to wheat production. Many of the problems encountered in segment classification during LACIE Phase I (1975) and Phase II (1976) were overcome by implementing P-1. Details of this procedure are discussed in reference 1. The analyst used the I-100 for displaying images and classification maps, selecting and labeling training fields, and evaluating and reworking the classification results. All clustering and classification were completed on ERIPS.

Many problems evolved because of the configuration of the I-100/ERIPS system. Interfacing problems created a time delay between initial processing and the receipt of results. It was hoped that the time lag would be a day or two, but experience indicated an average time lapse of a week. Because of this, analysts had to analyze and track up to nine segments at a time, greatly decreasing analyst efficiency. Inherent interfacing problems within the I-100/ERIPS system have been eliminated by the ATS due to its dependence on a single image-processing computer. It must be noted that the I-100/ERIPS system was never meant to be an example of an operationally optimal interactive image-processing system but rather was purposely pieced together to determine whether interactive image processing could improve segment classification results.

The problems and inefficiencies found in the I-100/ERIPS system could be divided into five major categories: (1) the wrong capabilities were stressed,

^aU.S. Department of Agriculture, Houston, Texas.

(2) methods were needed to ease the man-machine interface, (3) unnecessary data were provided to the analysts, (4) additional capabilities were needed, and (5) design performance needed improvement.

The USDA analyst team analyzed on the I-100/ERIPS system selected segments from the United States, Canada, and the U.S.S.R. The varied wheat conditions throughout the three-country study area enabled analysts to become familiar with varying cultural practices, weather conditions, farming methods, and how these variable conditions affect wheat growth and spectral response.

It soon became apparent that a single analyst-processing procedure was not optimal for classifying Landsat data. P-1 worked fairly well when used in areas having small, randomly distributed fields and heterogeneous signatures, but it was inefficient in agriculture areas having relatively large fields and homogeneous spectral signatures. USDA analysts recommended that optional processing procedures should be developed for varying agricultural conditions within a segment in order to make optimum use of analyst time.

An outcome of the classification procedures problem was the application of the direct crop option, which is currently being implemented by the ATS. This procedure gives the analyst the capability of outlining a desired field and obtaining the area within that field directly. Therefore, an areal estimate of a specific crop within a segment can be obtained quickly and does not require intermediate clustering and/or classification algorithms as is the case with the analyst-selected training fields and the P-1 options.

The I-100/ERIPS activity allowed the USDA analysts to conduct several research pilot studies. One of these studies focused on the early-season estimate problem. An early-season spring wheat area estimate was made on a total of 17 segments: 11 from Canada; 3 from the U.S.S.R.; and 3 from the United States. An early-season wheat area estimate is defined as an area estimate of wheat within a scene that is obtained prior to Landsat-detectable emergence of all the wheat grown within that scene. A majority of the early-season spring wheat estimates were made from acquisitions acquired the first week of May 1977. A few estimates were made from acquisitions acquired during the fall of 1976. For the 17 segments analyzed, the mean difference between the early-season estimates and the best at-harvest estimates was 1.8 percent. Additional information on this study is reported in reference 2.

Another study conducted on the I-100/ERIPS was the use of the vegetative index for crop identification. Vegetation indexes are computed from the raw multispectral scanner digital data and are used to determine vegetation density, greenness, and physiological condition within a given area. Considerable research on the vegetation index conducted by LACIE (ref. 3) and other government agencies (ref. 4) revealed that it can be successfully used to detect drought and monitor plant and soil moisture conditions; however, few studies have examined the use of the vegetative index approach for crop identification. One objective of the study on the I-100/ERIPS was to investigate the usefulness of the vegetative index for schemes in crop identification and acreage estimation. The results of this limited study show that the vegetation index can be used cost effectively to identify crops and natural vegetation (ref. 5).

The experience gained from I-100/ERIPS proved to be invaluable to the USDA. The analysts received training in operating an interactive system and were given the opportunity to process and analyze remotely sensed data. The immediate payoff has been in the design and implementation of the USDA ATS.

ATS Processing Activities

The USDA analysts conducted the first operational tests of the ATS interactive image-processing system in December 1977. Originally, 72 U.S.S.R. segments were selected for this test, but only 38 segments were actually processed. (Various circumstances, including cloud cover restrictions, prevented the analysis of the remaining 34 segments.) The data were acquired between seedbed preparation and wheat emergence.

The designated crop option discussed earlier was used for analyzing these segments. The wheat was only partially emerged on the imagery and, therefore, it would have been difficult to obtain meaningful estimates using the conventional clustering and classification procedures. The designated crop procedure enabled the analyst to obtain an area estimate of all the wheat fields within the segment, even though the spectral signatures within the fields were inconsistent due to the partial emergence of wheat. The analyst relied on his interpretation of the spectral signatures within the sample segment to determine the percentage of the segment planted to wheat.

Analyst procedures have been documented and are currently being used in the training of additional image analysts. Each analyst has used these procedures as part of a self-training course on the ATS equipment. Any new analysts to come aboard will be required to participate in a structured training course consisting of all aspects of remote sensing with emphasis on processing and analysis of the data.

FUTURE PLANS

Future plans as well as current programs for the ATS are to develop the ATS Crop Condition Assessment (CCA) Group that will measure the impact of abnormal conditions (e.g., excessive moisture, drought, winterkill) affecting crop production. USDA analysts will assess the impact of events by using remotely sensed data and conventional data sources now used by USDA foreign commodity experts. The CCA Group will focus its efforts on assessing the impact of events in countries where crop shortages and surpluses have a major impact on world commodity markets and prices. Important world crops such as wheat, barley, rye, corn, soybeans, sunflowers, rice, cotton, peanuts, and sorghum will be included in the crop condition assessment program. During crop years 1978 and 1979, ATS personnel will be developing the CCA Group to assess the condition of wheat in important foreign producing countries.

Currently, USDA plans are to have the CCA Group assess the impact of events detected and reported to the CCA Group by the Joint Agricultural Weather Facility (JAWF). The JAWF is composed of personnel from the National Oceanic and Atmospheric Administration (NOAA) and the USDA. The JAWF will monitor and detect abnormal events using meteorological and ancillary data sources. Unusual events detected by JAWF will be reported to the CCA Group in a timely manner to hasten the reporting of impact assessments to key USDA commodity experts and decisionmakers.

The CCA Group is composed of two important components, the data base and analysis/reporting. The remainder of this paper will discuss the format and operations of these components.

Data Base Component

Both historical and current multispectral imagery, meteorological, and agricultural data are required to

support the CCA Group. An efficient and fast system for storage, retrieval, and analysis of the data is crucial for such a large-scale project. The ATS approach to this data handling and analysis problem was to develop an automated, geographically oriented, gridded data base. The data base is expanded as more countries and crops are added to the CCA unit.

The entire agricultural and potential agricultural universe is divided into grid cells. Each grid cell has a unique latitude and longitude address and therefore can be singularly addressed by an "I" and "J" identification. Each cell is 25 by 25 nautical miles and can be further subdivided into quadrants. The following is a brief list of the data stored within each cell.

1. Country, region, zone, and strata locations
2. Five- by six-nautical-mile sample segment locations and associated data
3. Crop types
4. Percent agriculture
5. Current and historical daily meteorological data, including maximum/minimum temperature, precipitation, snow cover, and wind velocity
6. Soil data (quadrant level), including surface texture, depth, slope, drainage, available water-holding capacity, and moisture
7. Yield models
8. Crop calendars
9. Historical agricultural statistics including area, yield, and production
10. Agronomic data, including irrigation type and percentage, fertilization method and percentage, tillage practices, and cultivation practices

The analyst may access these data interactively while working at the analyst station. The data will be presented as maps and/or tables on both the cathode-ray tube (CRT) screens and the line printer. The maps will be displayed at different scales according to the geographical size of the area being displayed. An example of a data information product is a map showing the irrigation distribution and density for a designated area specified by the analyst. A supportive table will appear with the map specifying the types of irrigation within the designated area.

One of the major tasks of the ATS is to construct a data base for each country that has at least one of the major commodity crops listed previously. During crop year 1978, data bases are being constructed for Montana, North Dakota, and a selected area in the U.S.S.R.

Analysis Component

The CCA analysis component will utilize vegetative index numbers to measure health and vigor of the crop or crops of interest. The vegetative index number is a transformation of the MSS data into various descriptive components. One component measures greenness and is commonly known as a "green number." These numbers measure the approximate amount of green biomass in the scene and the relative vigor or health of that green biomass.

Currently, the ATS is testing under varying agricultural conditions six different green numbers. Upon completion of this testing exercise, the ATS will implement a green number(s) that best detects crop vigor or condition. The six green numbers include the Ashburn Vegetative Index (AVI), the Kauth-Thomas Vegetative Index (KVI), the Perpendicular Vegetative Index (PVI), the Transformed Vegetative Index (TVI), the Leaf Area Index (LAI), and the Difference Vegetation Index (DVI). Existing literature on these vegetative index numbers (refs. 3 to 5) has been reviewed and considered by the ATS evaluators.

The green numbers will be used in combination with meteorological data to assess crop condition. Lookup tables showing the relationship between green numbers and (1) soil moisture, (2) crop calendar, and (3) yield will be developed for specific geographic areas. These tables will aid the analyst in his assessment of crop condition. The method by which the green numbers will be used for crop condition assessment follows.

The first step in using green numbers will be to view the current Landsat images of selected sample segments for purposes of creating a map or image of the natural vegetation area (NV) and areas containing the desired crop (DC). The AVI will be used, where possible, to automatically create the NV map. Average green numbers will be calculated and stored, from each of the vegetative index algorithms, for the entire sample segment, for the NV map, and for the DC map. Green-number isoline maps will be plotted and interpreted for crop condition.

Historical Landsat imagery will be acquired for the same segments discussed in the previous paragraph. Green numbers will be calculated and stored for this historical imagery, following the procedure described for current-year imagery. The analyst will compare and evaluate the green numbers derived from the historical and current imagery, the NV map, and the DC map to determine the current-

year crop condition. The crop calendar, soil moisture, current and historical meteorological data, and yields derived for these segments will be included in the analysis. The primary goal of this is to determine and assess the amount of change in the crop. This assessment will address a change in quality, areal extent, yield, and production. A report will then be generated documenting this assessment.

During 1978, ATS analysts will perform the steps just discussed for purposes of assessing the condition of the wheat crop in Montana, North Dakota, and one selected area in the U.S.S.R.

Yield models will be required to support the CCA Group. For 1978 and 1979, two principal wheat yield models are of interest to the ATS. They are the LACIE-tested CCEA model, developed by the Center for Climatic and Environmental Assessment (CCEA) of NOAA (refs. 6 and 7), and the Kansas State University (KSU) model (refs. 8 and 9). These two models will be implemented, tested, and evaluated by the ATS. Results from the CCEA model are produced at 30-day intervals; the KSU model predicts yields at 10-day intervals.

The ATS will implement, evaluate, and apply other crop models as they are developed and documented in the research community.

During 1978, the ATS will implement and operate a wheat crop calendar. The crop calendar subroutine of the KSU spring wheat yield model will be the primary crop development model. Model results at selected weather stations are interpolated to the grid cell units of the data base. The model is run every 10 days with daily meteorological data.

The ATS will implement, evaluate, and apply other crop development models as they are developed and documented in the research community.

During 1978, the ATS will implement and run the Versatile Soil Moisture Budget (VSMB) model. The results are used in the KSU yield model. The VSMB subroutine will be extracted from the KSU yield model and run as a separate program.

SUMMARY

The ATS is chartered to implement, test, and evaluate technologies and capabilities for their application feasibility by the USDA. The analysis and application of remotely sensed and other data is an important component of the ATS. Therefore, the remote-sensing analyst must be highly qualified and trained in order to support this component.

The USDA analysts had an opportunity to gain experience on an interactive image-processing system during LACIE Phase III. Landsat data in 5-b, 6-nautical-mile format from the United States, the U.S.S.R., and Canada were analyzed and wheat area estimates determined. The varied wheat conditions allowed the analyst to study different agronomic and cultural practices. These differences necessitated that more than one processing procedure be developed to handle the varied agricultural conditions. To partially solve this problem, the ATS is currently implementing three processing options, each developed for specific agricultural situations.

While working on the interactive system, the USDA analysts developed a list of recommendations and changes to the system. This list of items was considered during the development of the ATS and has resulted in a system with capabilities and enhancements that are a direct outcome of the USDA analyst experience gained during LACIE Phase III.

Future plans for the ATS call for the development of the ATS CCA Group. The CCA Group will detect in a timely manner changes affecting production and quality of commodities and will assess the impact of the change. The ATS is tasked to develop and integrate the elements of the CCA Group. These elements are the central data base and the analysis component which utilizes Landsat data and yield, crop calendar, and soil moisture models.

The ATS personnel will develop the data base and analysis procedures for this system. The yield, crop calendar, and soil moisture models will be transferred from LACIE and implemented in the ATS.

In 1979 and the early 1980's, the CCA Group will be expanded to include additional crops and crop-producing regions of major importance in world trade. The ATS will coordinate with and rely on the research community to develop the technology

needed to support ATS objectives. Developed technology will be transferred to the ATS for implementation, testing, and evaluation prior to its incorporation into an operational early warning system.

REFERENCES

1. Phase III CAMS Detailed Analysis Procedures. LACIE-00720, NASA Johnson Space Center, Houston, Texas, Aug. 1977.
2. May, G. A.; Frank, A. D.; and Hansen, H. L.: Early Season Wheat Area Estimation—Results and a Recommended Approach. ATS Technical Memorandum No. 3, USDA/FAS Remote Sensing Research Center, Houston, Texas, Mar. 1978.
3. Thompson, D. R.: The Use of Landsat Digital Data To Detect and Monitor Vegetation Water Deficiencies. NASA Johnson Space Center, Houston, Texas, Apr. 1977.
4. Richardson, A. J.; and Wiegand, C. L.: Distinguishing Vegetation From Soil Background Information. Photogrammetric Engineering, vol. 43, no. 12, 1977, pp. 1541-1552.
5. Ashburn, P.: The Vegetative Index Number and Crop Identification. USDA/FAS Remote Sensing Research Center, Houston, Texas, 1978.
6. Wheat Yield Models for the U.S.S.R. LACIE-00430, NASA Johnson Space Center, Houston, Texas, 1977.
7. Wheat Yield Models for the United States. LACIE-00431, NASA Johnson Space Center, Houston, Texas, 1977.
8. Feyerherm, A. M.: Response of Winter and Spring Wheat Grain Yields to Meteorological Variation, Final Report. Contract NAS9-14282, NASA Johnson Space Center, Houston, Texas, Feb. 1977.
9. Feyerherm, A. M.: Planting Date and Wheat Yield Models, Final Report. Contract NAS9-14533, NASA Johnson Space Center, Houston, Texas, Dec. 1977.