E80-10055
NASA CR-160422

AgRISTARS

SR-E9-00404
NAS9-15476

(E80-10055) DEVLEOPMENT OF LANDSAT-BASED TECHNOLOGY FOR CROP INVENTORIES: APPENDICES Final Report, 15 Nov. 1978 - 14 Nov. 1979 (Environmental Research Inst. of Michigan) 303 p HC A14/MF A01 CSCL 02C G3/43
N80-18507
Unclas 00055

...ram for ...and ...nventory Surveys Through Aerospace Remote Sensing

Supporting Research

December 1979

FINAL REPORT

DEVELOPMENT OF LANDSAT-BASED TECHNOLOGY FOR CROP INVENTORIES: APPENDICES

Q.A. Holmes, R. Horvath, R.C. Cicone, R.J. Kauth, W.A. Malila

ENVIRONMENTAL RESEARCH INSTITUTE OF MICHIGAN • P.O. BOX 8618 • ANN ARBOR, MICHIGAN 48107

NOTICES

Sponsorship. The work reported herein was conducted by the Environmental Research Institute of Michigan under Contract NAS9-15476 for the National Aeronautics & Space Administration, Johnson Space Center, Houston, Texas 77058. I. Dale Browne was Technical Monitor for NASA. Contracts and grants to the Institute for the support of sponsored research are administered through the Office of Contracts Administration.

Availability Notice. Request for copies of this memorandum should be referred to:

National Aeronautics & Space Administration
Scientific & Technical Information Facility
P. O. Box 33
College Park, Maryland 20740

SR-E9-00404$_2$
NAS9-15476

FINAL REPORT

DEVELOPMENT OF LANDSAT-BASED TECHNOLOGY FOR CROP INVENTORIES:

APPENDICES

BY

Q.A. Holmes, R. Horvath, R.C. Cicone, R.J. Kauth, W.A. Malila

The research reported here was initiated during the planning of the AgRISTARS Supporting Research Project and was a part of those plans, although this research will stand on its own merit. The benefiting Supporting Research project element is Area Estimation Research.

Environmental Research Institute of Michigan
P.O. Box 8618
Ann Arbor, Michigan 48107

December 1979

PREFACE

This report describes part of a comprehensive and continuing program of research concerned with advancing the state-of-the-art in remote sensing of the environment from aircraft and satellites. The research is being carried out for NASA's Lyndon B. Johnson Space Center (JSC), Houston, Texas, by the Environmental Research Institute of Michigan (ERIM). The basic objective of this multidisciplinary program is to develop remote sensing as a practical tool to provide the planner and decision-maker with extensive information quickly and economically.

Timely information obtained by remote sensing can be important to such people as the farmer, the city planner, the conservationist, and others concerned with problems such as crop yield and disease, urban land studies and development, water pollution, and forest management. The scope of our program includes:

1. Extending the understanding of basic processes.
2. Discovering new applications, developing advanced remote-sensing systems, and improving automatic data processing to extract information in a useful form.
3. Assisting in data collection, processing, analysis, and ground-truth verification.

The research described herein was performed under NASA Contract NAS9-15476 and covers the period from November 15, 1978 through November 14, 1979. I. Dale Browne/SF3 was the NASA Contract Technical Monitor and Thomas Pendleton/SF3 was the primary NASA Technical Coordinator of the activity. The program was directed at ERIM by Richard R. Legault, Vice President and Head of the Infrared and Optics Division, Quentin A. Holmes, Program Manager, and Robert Horvath, Head of the Analysis Department.

The formally identified authors are those individuals who designed and coordinated both the research program and the preparation of this document. Actual authorship should recognize the contributions of the following individuals.

R.J. Balon	W.A. Malila
R.C. Cicone	W.F. Pont
E.P. Crist	D.P. Rice
H.M. Horwitz	W. Richardson
R.J. Kauth	T. Wessling
P.F. Lambeck	

The authors would like to acknowledge the research support provided by E.V. Lamson and D.J. Witte, and indicate their appreciation of the secretarial support provided by P.R. Swonger and D. Dickerson, with assistance from M.N. St.John, G.A. Goike and R.M. Coleman.

ERIM would like to extend its appreciation to others whose efforts facilitated the conduct of the program over the last year: R.L. Nance/SF3 of NASA/JSC for assistance in acquiring Landsat and ancillary data, and D.C. Helmer, C.V. Nazare, R.W. Payne and W.L. West of LEC/Houston for participation in the design and conduct of the analyst interpretation experiment.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Cont.)

LIST OF ILLUSTRATIONS (Concluded)

LIST OF TABLES

LIST OF TABLES (Cont.)

LIST OF TABLES (Cont.)

APPENDIX A
EXTRACTION OF TEMPORAL-SPECTRAL FEATURES

A.1 UTILITY OF TEMPORAL-SPECTRAL FEATURES

A.1.1 BACKGROUND

Assignment of crop type or crop group labels to a set of labeling targets (pixels, fields, etc) often involves recognition of a temporal-spectral pattern that is characteristic of a crop or crop group. Since the labeling targets consist of populations of plants rather than a single plant, there is likely to be some variation in stage of development within the target area at any given point in time. As a result, crop development and the accompanying spectral development at the target level will, in most cases, appear to be a continuous rather than a discrete process. As in most other biological population phenomena, these development patterns could be expected to conform to a Sigmoid curve (see Figure A-1).

Based on these assumptions, one can conclude that the discrete observations obtained from Landsat are samples from a continuous pattern of spectral development. Characterization of that pattern, based on the given set of samples, should allow more complete description of the target and, as a result, more accurate labeling. The term 'profile' is used to describe the mathematical representation of this development pattern.

A.1.2 LEVELS OF USE

Representations of the temporal-spectral development of crops, fields, etc may be used in a variety of ways. Since the degree of precision demanded of the profiles varies with intended use, it is of value to define a set of levels or categories of application. Models may then be developed and evaluated at the level of use for which they

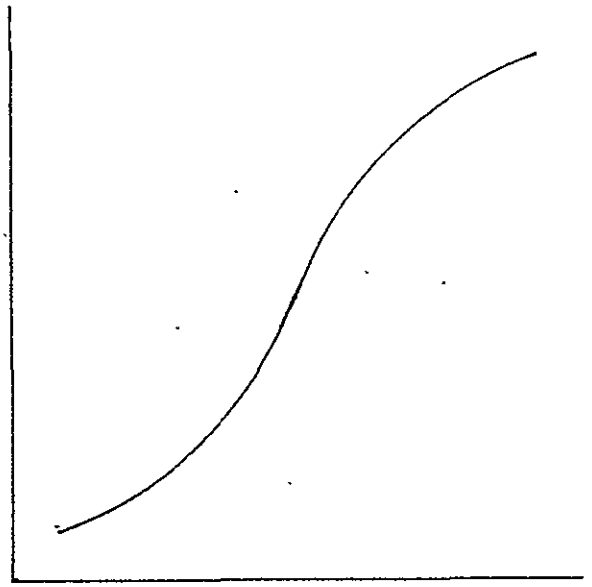


FIGURE A-1. SIGMOID CURVE SHAPE

are intended, and extended to other levels (particularly higher or more demanding levels) only after careful re-testing and, perhaps, modification.

Level 1: At the first level, a profile need only provide a stylized representation of the crop development pattern. Small deviations from the norm that might be observed on a target level needn't be accounted for; a relatively simple model, even one that only connects observations with a series of line segments, may suffice.

Level 2: Estimation of a particular feature of crop spectral development requires a more accurate profile fit, at least in that portion of the profile from which the feature of interest is derived. Thus, a Level 2 model must, on the whole, be more accurate and flexible than one which could be used at Level 1.

Level 3: The next level of application involves characterization of overall crop development and multiple features. Models at this level must be able to accurately portray spectral development throughout the growing season, and as such must be even more flexible and accurate than at the previous levels. This level, which is the focus of the study reported here, will be more fully discussed in Section A.1.4.

Level 4: The final level in the sequence involves a still greater accuracy requirement. If a profile can be fit to a set of observations accurately enough, and the interpolation and/or extrapolation of profile values can be carried out wisely enough, then it may be possible to use the profile values themselves as data, replacing or augmenting the Landsat observations. While precise mathematical fitting is important here, the major need is understanding of the physical and biological processes taking place and their impact on spectral characteristics, and incorporation of that understanding into the fitting process.

A.1.3 PAST APPLICATION OF PROFILE TECHNOLOGY

Multitemporal characterizations of spectral development have been used previously at both Levels 1 and 2. Most of the work has involved spring small grains and Tasseled-Cap Greenness, and utilized a model developed at ERIM. This model is of the form

$$F(t) = at^b e^{ct^2} \qquad \text{(A-1)}$$

where $F(t)$ = Greenness - 25,

t = shifted day of year - 125,

a,b,c = model parameters.

This form is a smooth curve that can fit a series of observations of spring small grains targets. This model was originally applied at ERIM to crop calendar shift estimation [18] in a refinement of a technique developed by G. D. Badhwar [16] in which a simpler profile form was used.

More recently, the same model has been used to describe particular features of spectral development. Appendix G describes a machine labeling technique for distinguishing Spring Wheat and Barley that employs the maximum value of the profile as an indicator of moisture stress conditions on a sample-segment level. A similar application has been carried out by UCB in the context of detection of episodal events [45].

Profile technology, and the described model form, have also been used as a basis for a classifier both of spring small grains and corn [46].

A.1.4 LEVEL 3 PROFILE APPLICATION

Figure A-2 illustrates some of the potentially useful profile-based features which could be extracted at Level 3. The set of features in toto provides a comprehensive description of the target spectral development pattern. The maximum value of the profile can serve as an indicator of crop vigor and percent cover (assuming a vegetation indicator such as Tasseled-Cap Greenness is used). The rates of Greenness increase and decrease, and changes in those rates, can serve as useful features for distingusihing crops whose profiles have similar overall shapes but which develop differently at some points in the season. These same rates can also offer information on crop condition. Similarly, the total development time (the interval between departure from and subsequent return to some nominal base value) or half-amplitude interval could provide information relevant to both crop identification and condition assessment.

The set of features, perhaps with additional profile-derived features, should provide a more cohesive description of the target, allowing broader inferences to be made relative to both labeling and, where the two are separate, assessment of crop vigor.

A.1.5 PURPOSE OF CURRENT INVESTIGATION

The work reported herein was undertaken for two major reasons. First, the promising aspects of Level 3 application (overall development characterization) prompted an interest in determining whether the model currently in use and developed for Level 1 or 2 use, could be extended into the arena of Level 3 application. Second, it was deemed important to study the various steps in the profile-fitting process itself and to define a procedure that could serve to standardize application of the technology in the community.

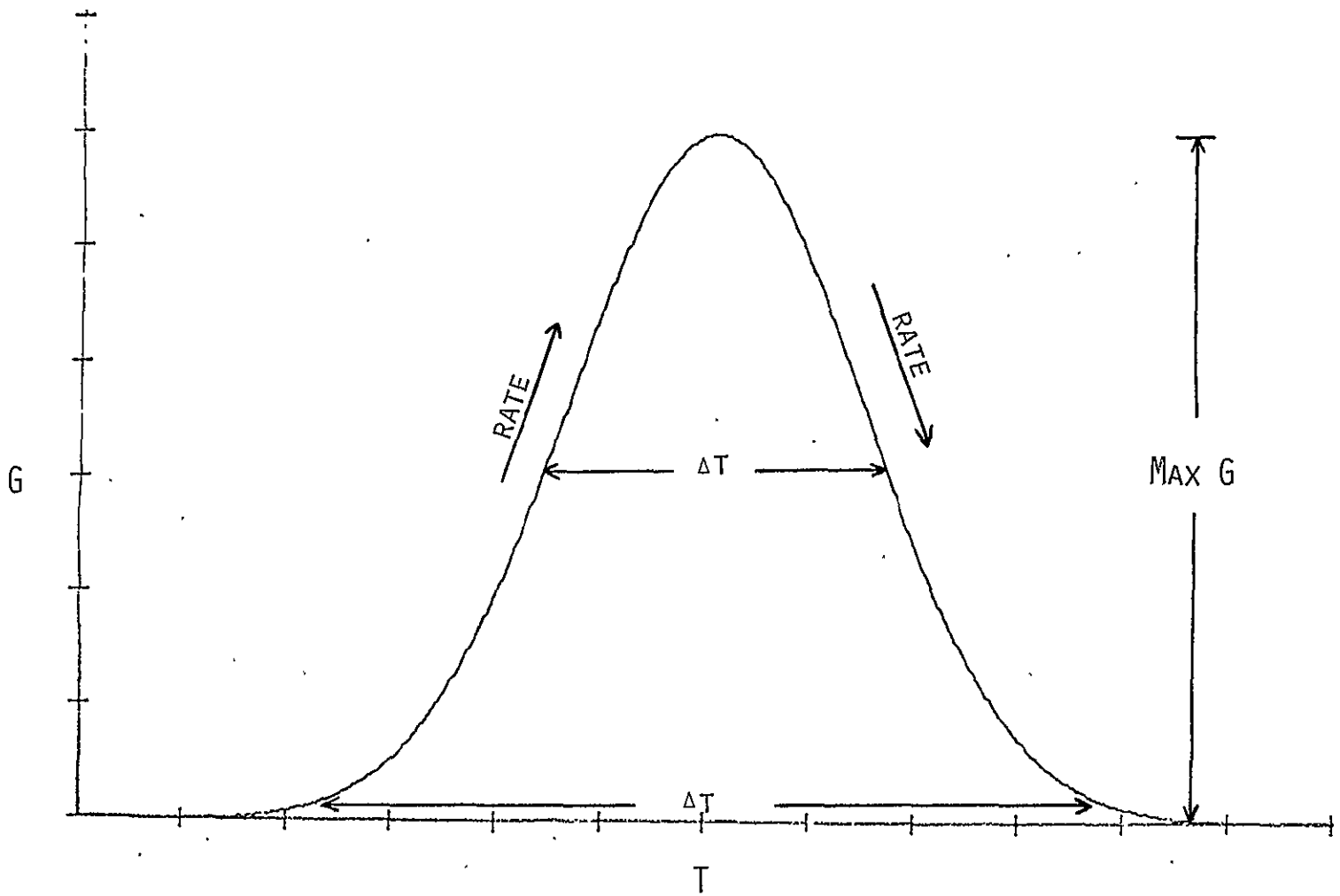


FIGURE A-2. PROFILE-BASED FEATURES

A.2 STUDY OF PROFILE-FITTING FOR OVERALL CROP DEVELOPMENT CHARACTERIZATION

A.2.1 GENERAL STEPS OF A FITTING PROCEDURE

Any procedure by which a profile model is fit to a set of data points would likely include the following steps:

a. Data Selection: The target to which a profile is to be fit must be specified. This target may be individual pixels, all pixels of a field or quasi-field, or the mean of the field or quasi-field.

b. Data Preparation: The Landsat data may require several steps of preprocessing to arrive at a standardized, normalized set of observations that can be converted into a smooth profile.

c. Parameter Estimation: Once the data is prepared, the values of the profile parameters must be estimated.

These steps will be treated separately in the following description of study approach and results, with the data selection step being discussed after the other two steps.

A.2.2 DATA PREPARATION

In order to utilize the information obtained by fitting a profile to a set of observations, to obtain meaningful information about overall crop development, it is necessary that the influence of external phenomena on target spectral appearance be reduced to the greatest possible degree. Variations in profile shape must be tied to crop type or condition and not to sun angle, haze condition, etc.

The series of normalization steps includes:

a. Sensor Calibration (e.g., Landsat 2 to Landsat 3).
b. Sun Angle Correction (cosine).
c. Screening to flag clouds, shadows, etc.
d. Haze Correction (spatially varying XSTAR) [37].

A second stage of data preparation is estimation of crop calendar shift. This procedure standardizes data to an arbitrary but common time scale, thus reducing signal variation on any given day which is the result not of crop appearance differences per se but rather of differences in stage of development at the time of observation. This process not only reduces signal variation on any given day, but provides, in the case of fitting to all pixels of a field, added information that can be utilized in fitting the profile.

The final stage of data preparation involves translation of the data axes such that the origin approximately corresponds to the start of Greenness development of the small-grain target. The need for this translation or offsetting stems from the model form itself. Regardless of the actual starting Greenness value or starting time of Greenness development, the model will consider all times from t=0 through some maximum time, and F(t)=0 through some maximum Greenness. The desired Sigmoid shape of each side of the profile similarly occurs in the range from F(t)=0 through the maximum value. Figure A-3 illustrates the effect of omitting this step. Despite the fact that the first Greenness observations occur well away from the untranslated origin, the model form begins there, and requires parameters that allow for a long, relatively flat tail followed by a relatively rapid increase and decrease in the data range. Figure A-4 illustrates the same data fit with offset applied. Comparison of the two figures clearly points out the model's inability to accommodate both the long tail and the comparatively short span of actual crop development.

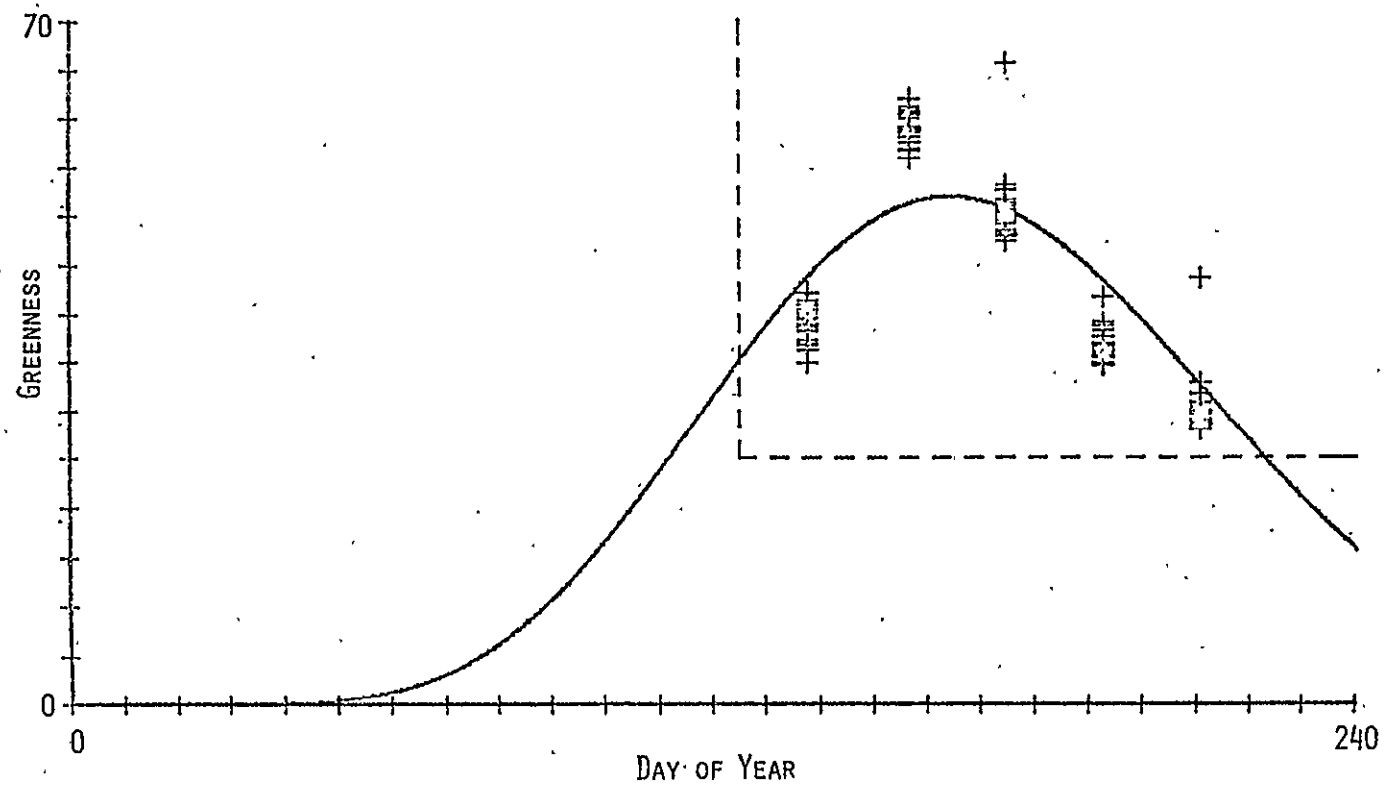


FIGURE A-3. EFFECT OF OFFSET OMISSION
(Segment 1663 Field 5)

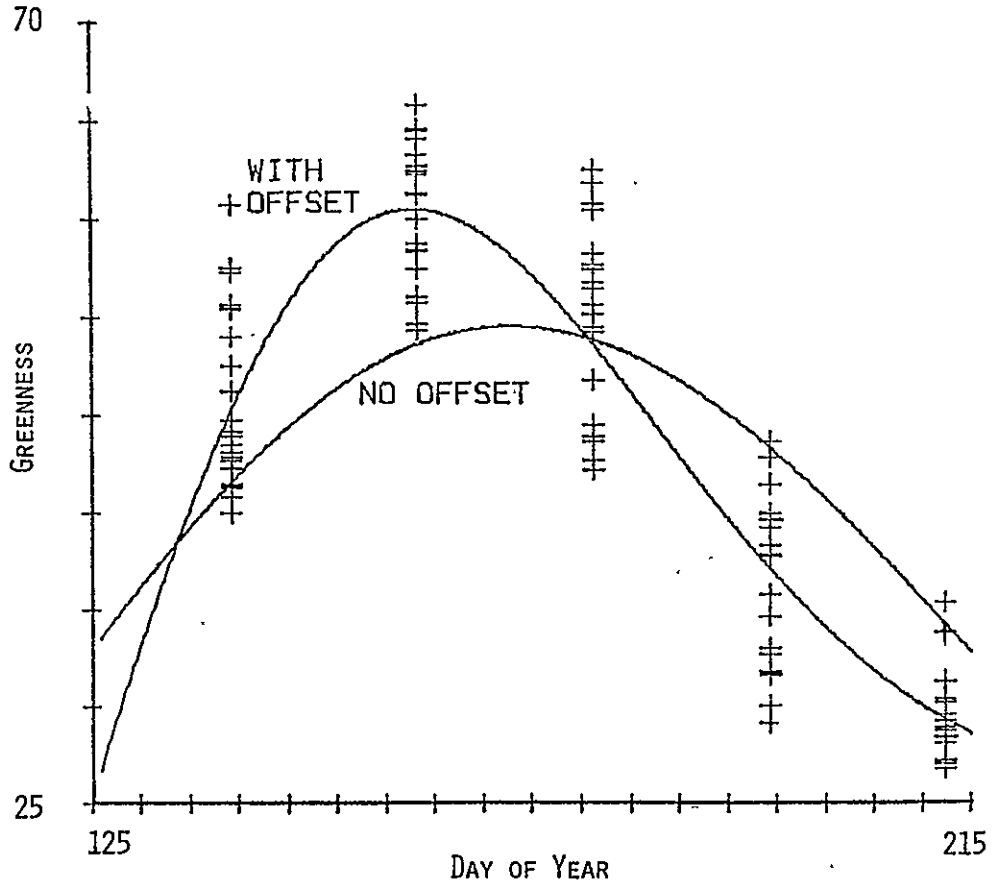


FIGURE A-4. COMPARISON OF PROFILE FIT WITH AND WITHOUT OFFSET
(Segment 1663 Field 7)

The resulting profile is significantly less accurate in overall development characterization than that computed using offset data. Variations in the method of parameter estimation, to be discussed next, can reduce this negative impact, but do not eliminate it.

Base on data from several segments, offsets of 125 days and 25 counts of Greenness (after adding 32 to all Tasseled-Cap channels) have been used.

A.2.3 PARAMETER ESTIMATION

Linear vs. Non-Linear Techniques: The method used to arrive at model parameter estimates has, in the past, exploited the ability to linearize the profile model. A logarithmic transformation produces the following linear model from:

$$\ln F(t) = \ln a + b \ln t + ct^2 \qquad \text{(A-2)}$$

Multiple linear regression can then be applied to produce least squares estimates of the model parameters.

While this method is simple and inexpensive, it has at least two disadvantages. First, the non-linear nature of the logarithmic transformation results in a distortion of the original data, and most importantly a compression of the peak of the set of observations. The least squares fit to these compressed data will, as a result, underestimate the peak of the untransformed data.

The problem is compounded by another aspect of the linear estimation technique. The least squares estimate in log space minimizes the quantity

$$\Sigma \left(\ln F(t) - \ln G_i(t)\right)^2 \qquad \text{(A-3)}$$

where $F(t)$ = profile value,

$G_i(t)$ = data value.

Since the difference of two logarithms is the logarithm of the ratio of the two quantities, the quantity being minimized is, in fact,

$$\Sigma \left(\ln \left(\frac{F(t)}{G_i(t)} \right) \right)^2 \qquad \text{(A-4)}$$

Since a given number of counts difference results in a larger ratio when the profile and data values are small than when they are larger, the linear least squares estimation procedure is giving greater weight to the tails of the profile in determining the best fit. Fit at the peak will be sacrificed, to some degree, in order to improve fit at the lower values. Not only does this accentuate the low peak estimate problem, it gives least importance to those parts of the profile (at and around the peak) that in many applications will be the most important.

An alternative method of parameter estimation involves a non-linear least squares technique. A routine from the IMSL package [47] was chosen which applies a modified Levenberg-Marquardt steepest descent algorithm to seek out the minimum value of a residual sum of squares surface through iterative estimation and evaluation of parameters. This technique avoids the need for a log transformation and thus the technical problem associated with that transformation.

Comparison of the two methods required definition of a common measure of goodness-of-fit since the R^2 from the linear regression is a measure of fit in log space rather than actual data space. As such, the following was used for comparison:

$$\text{Goodness of Fit} = 1 - \frac{\Sigma \left(F(t) - G_i(t) \right)^2}{\Sigma \left(G_i(t) - \bar{G} \right)^2} \qquad \text{(A-5)}$$

where $F(t)$ = profile value,

$G_i(t)$ = i^{th} data value,

$\bar{G}$ = mean of all data values,

and the range of the equation is [0,1].

Figure A-5 and Table A-1 illustrate the results of the comparison. In all cases the non-linear technique provides an improved estimate of profile parameters, as evidenced both visually and empirically.

Offsetting Problems Encountered: In the course of the parameter estimation study, an apparent problem related to data offsetting (see Section A.2.2) was encountered. Figure A-6 illustrates the problem, which appears to be the result of incorrect offset values. In both fields shown, Greenness values at shifted day 125 (offset day 0) are well above 25 (offset value 0). While the result varies with the parameter estimation technique, the impact in both is a clear reduction in profile accuracy. It is apparent, then, that the offset values used as standards are not appropriate in all cases.

Several approaches to field-specific offset determination were considered. First was an approach that assumes a constant Greenness value before crop development begins. A straight line is drawn through a pair of pre-peak observations and the intersection of the line with the constant value is assumed to be the start-up time for crop spectral development. The major drawback of this approach is it's inability to address the non-linear character of Greenness development. Figure A-7 illustrates this problem using simulated data. The lines drawn represent the results of using different 18-day pairs of data points along the profile. While many do seem to point to the same starting date, those that include an observation near the peak significantly mis-estimate the start-up time.

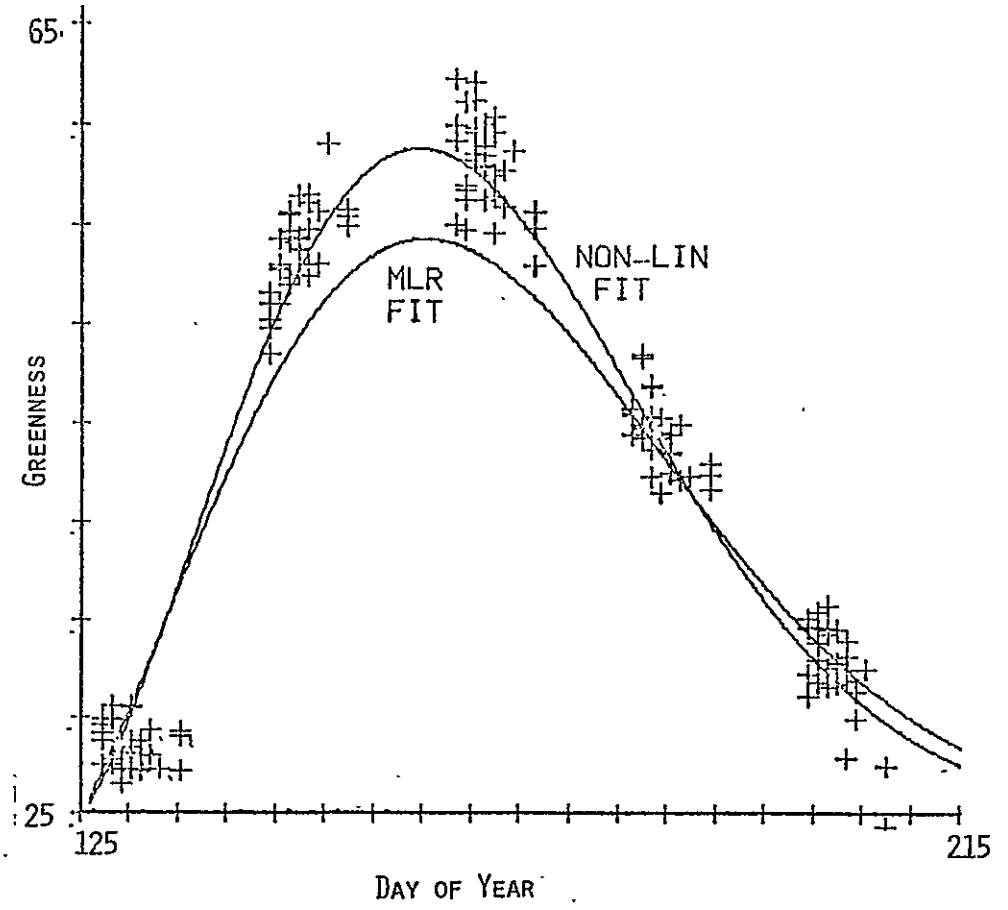


FIGURE A-5. COMPARISON OF PARAMETER ESTIMATION TECHNIQUES - LINEAR VS. NON-LINEAR LEAST SQUARES
(Segment 1663 Field 11)

TABLE A-1. COMPARISON OF GOODNESS-OF-FIT OF LINEAR VS. NON-LINEAR ESTIMATES

Segment	Field	Goodness-of-Fit	
		Linear	Non-Linear
1663	5	.946	.959
	7	.919	.932
	11	.933	.960
	12	.920	.929
	13	.914	.941
	14	.953	.963
1669	28	.819	.834
	30	.774	.860
1929	16	.883	.890
	17	.574	.709
	23	.725	.842
	24	.886	.923

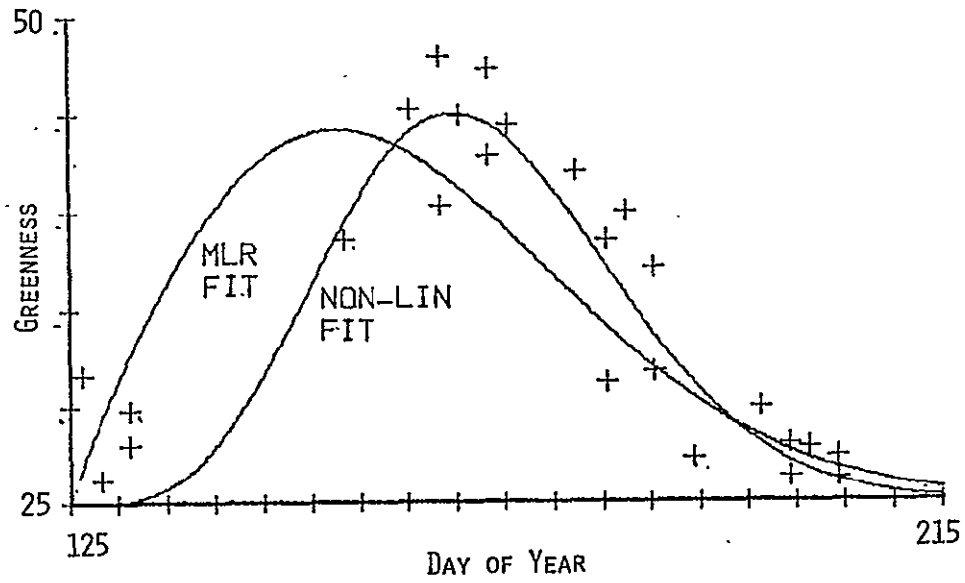


FIGURE A-6. RESULTS OF INCORRECT OFFSETTING
(Segment 1929 Field 23)

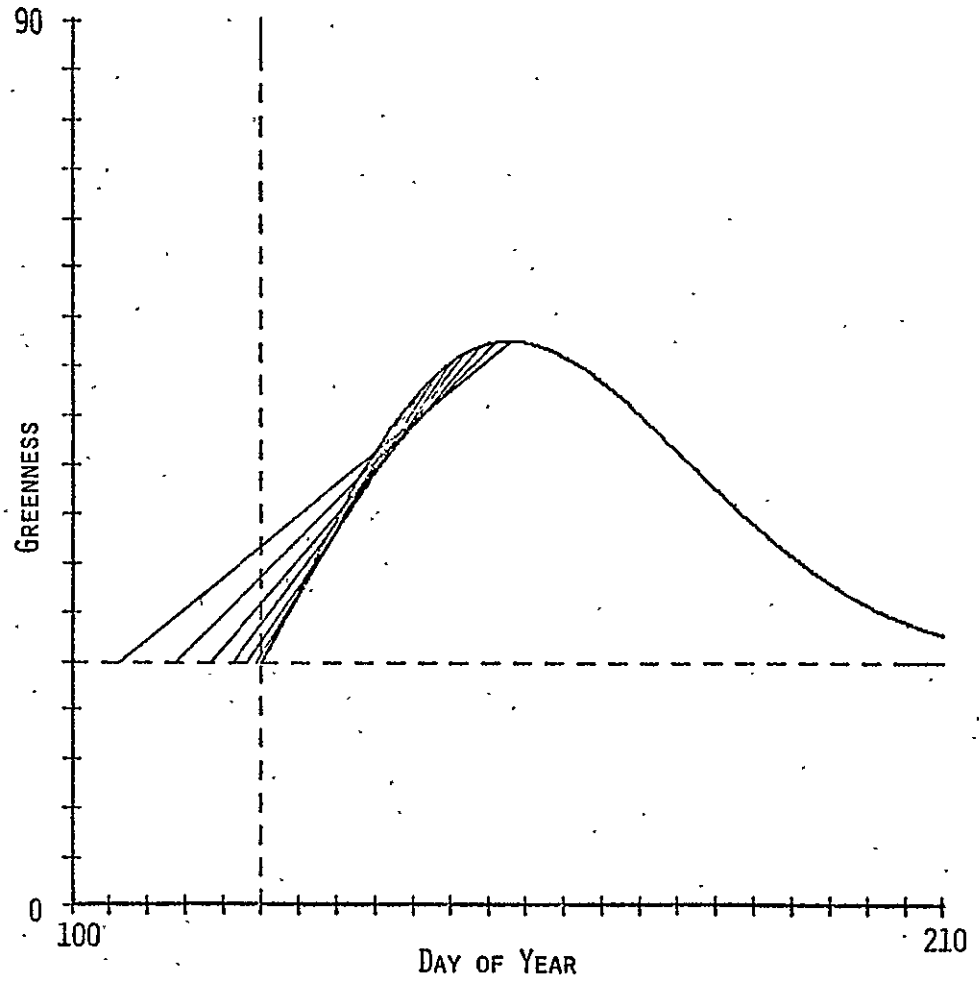


FIGURE A-7. STRAIGHT LINE ESTIMATION OF OFFSET EXAMPLE USING PROFILE AS DATA

This result is further illustrated in Figures A-8 and A-9, using shifted and unshifted actual data.

A second approach considered was inclusion of the offsets in model parameters to be estimated. The least squares fit would then be driven by five parameters instead of three. Attempts to apply the method, however, were unsuccessful, apparently due to parameter interactions.

Finally, an attempt was made to determine field-specific offsets by iteratively calculating profiles while changing offset, and selecting offsets based on the goodness-of-fit measure. However, as illustrated in Figure A-10, successive parameter estimates with a wide range of day offsets produced no clear optimum fit.

An alternative to target-specific offset determination is development of a model that is less constrained by offset requirements.

Development of a New Model: In response to the need for less dependence on offsetting, a new model was developed of the form:

$$F(t) = \begin{cases} ae^{b_1(t-t_p)^2} & ; \; t < t_p \\ ae^{b_2(t-t_p)^2} & ; \; t \geq t_p \end{cases} \qquad \text{(A-6)}$$

where $F(t)$ = Greenness-25

t = shifted day of year,

t_p = reference day of peak Greenness,

a, b_1, b_2 = model parameters,

and furthermore,

$a + 25$ = estimated peak Greenness.

The model is fit to observations before and after the peak independently, while maintaining continuity (to the first derivative) at

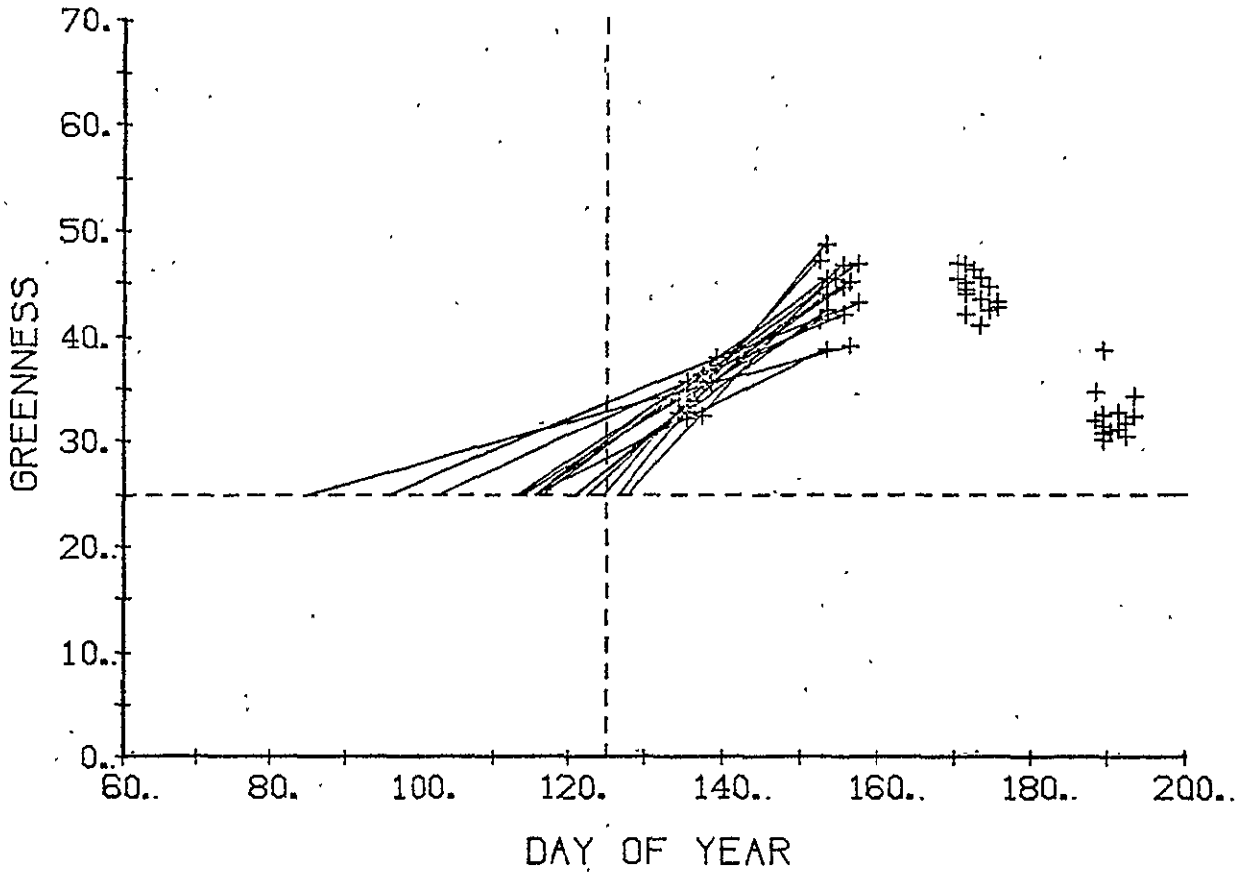


FIGURE A-8. STRAIGHT LINE ESTIMATION OF OFFSET
(Segment 1669 Field 28 - Shifted)

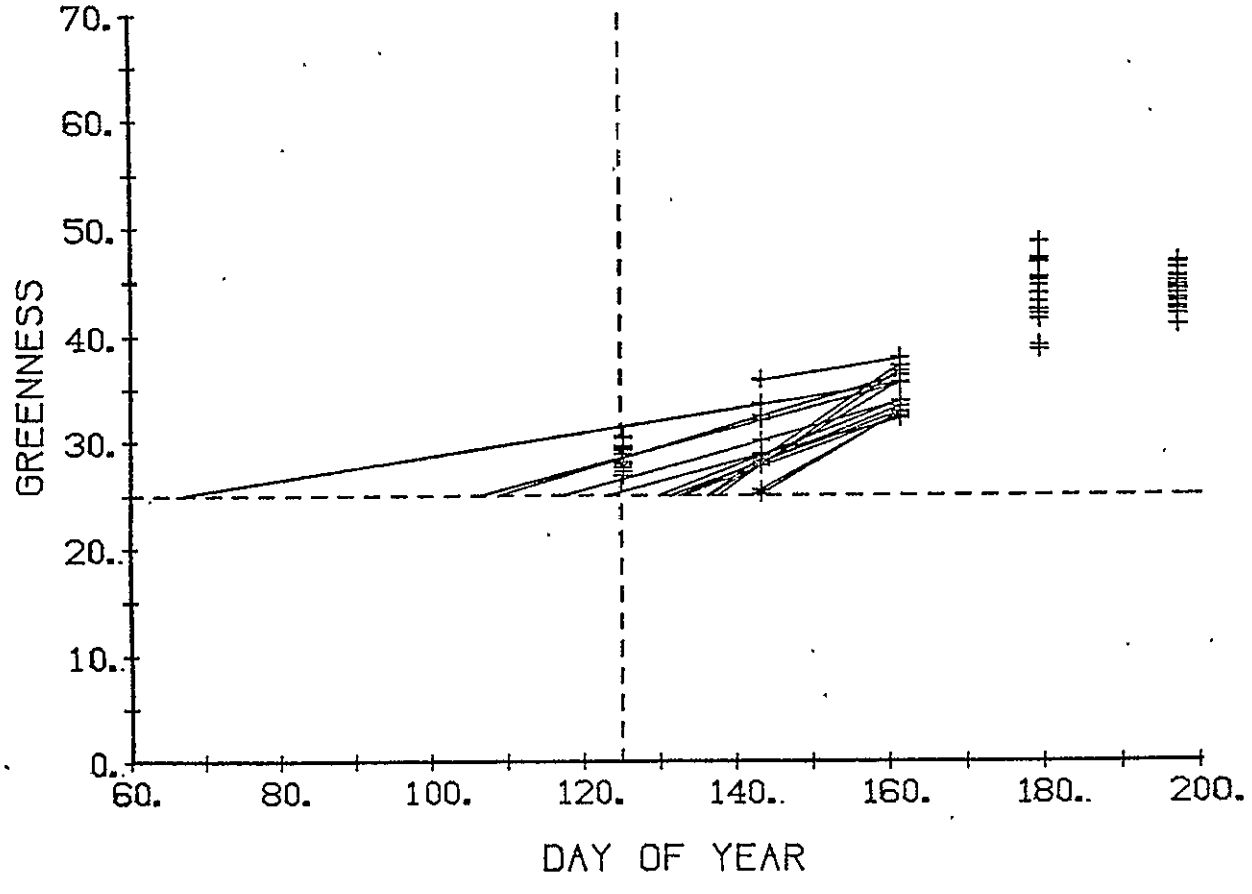
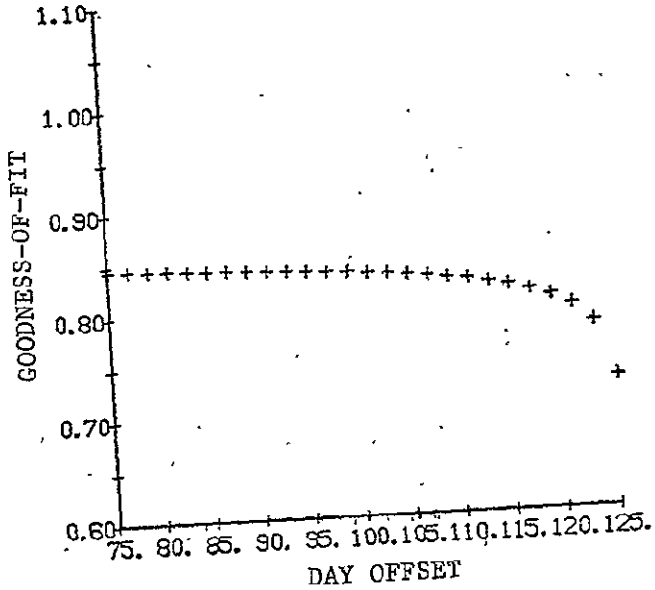
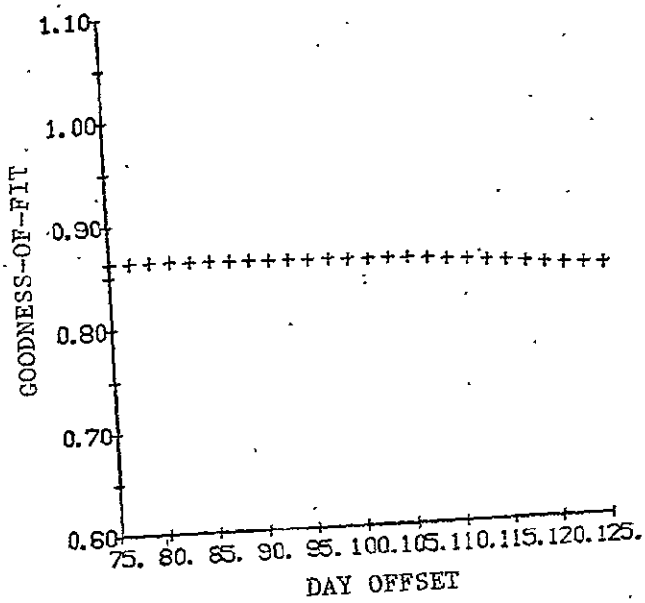


FIGURE A-9. STRAIGHT LINE ESTIMATION OF OFFSET
(Segment 1669 Field 28 - Unshifted)



Linear Regression Estimate



Non-linear Estimate

FIGURE A-10. RESULTS OF FIT USING ITERATIVE CALCULATION TECHNIQUE (Segment 1929 Field 23)

the peak. The "zero-point" for time is now the time of peak Greenness, which is a relatively stable quantity after crop calendar shift estimation.

Since the time value enters into the equation only as $(t-t_p)$, the offset is taken care of automatically, and only crop calendar shift must be estimated beforehand. Further, while an offset in Greenness is still desirable, the elimination of a need for a (0,0) starting point also reduces the requirement for precision of the Greenness offset.

The two models were compared, using the non-linear estimation procedure, using a number of fields from several segments. In most cases, the new model provided a more accurate fit than the original model. In one segment, Segment 1663, many of the fields exhibited a near-linear increase in Greenness early in the season. The new model, which tends to produce a more distinct Sigmoid shape both before and after the peak, proved unable to fit the early acquisitions in many of the fields tested in Segment 1663 (Figure A-11). However, the fit over the remainder of the growing season was accurate.

The improvement with the new model is most distinct on those fields for which offsets were clearly a problem. Figures A-12 and A-13 demonstrate the improvement in fit with the new model.

An additional advantage of the new model form is its flexibility. Since the model is fit in a piece-wise fashion, i.e., before peak and after peak, it is able to adapt to a wider range of profile shapes. Figure A-14 illustrates some of the possible forms achievable with the model.

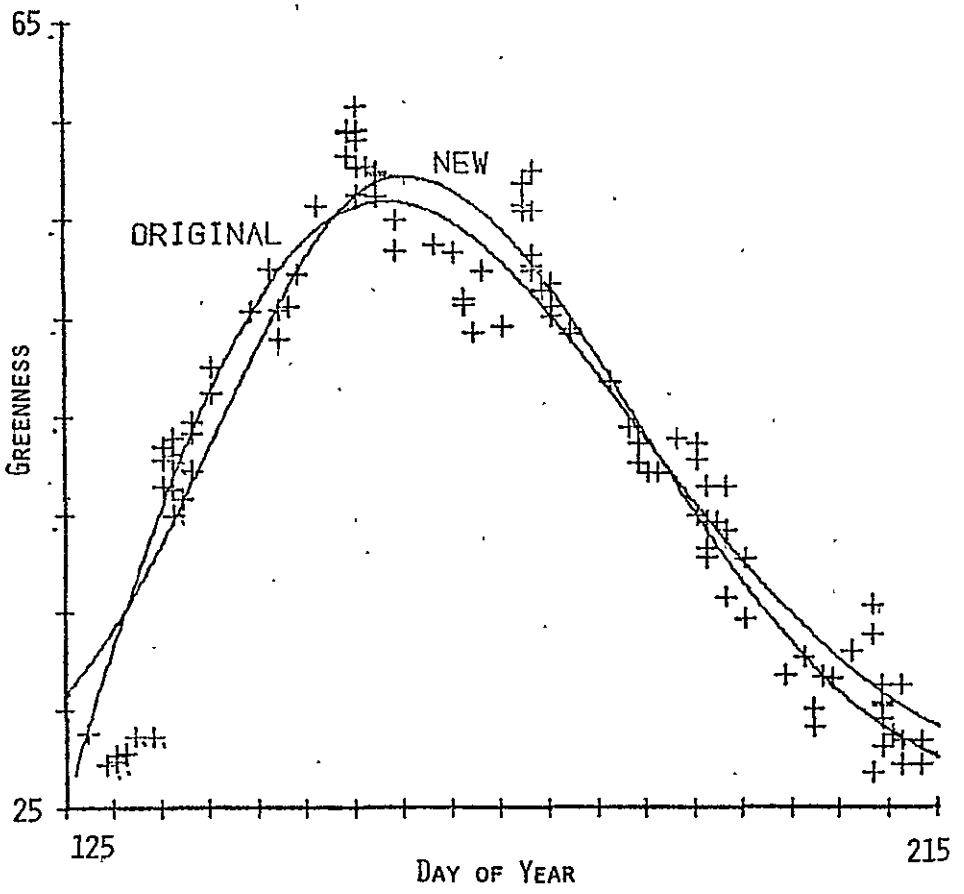


FIGURE A-11. COMPARISON OF MODEL FORMS
(Segment 1663 Field 7)

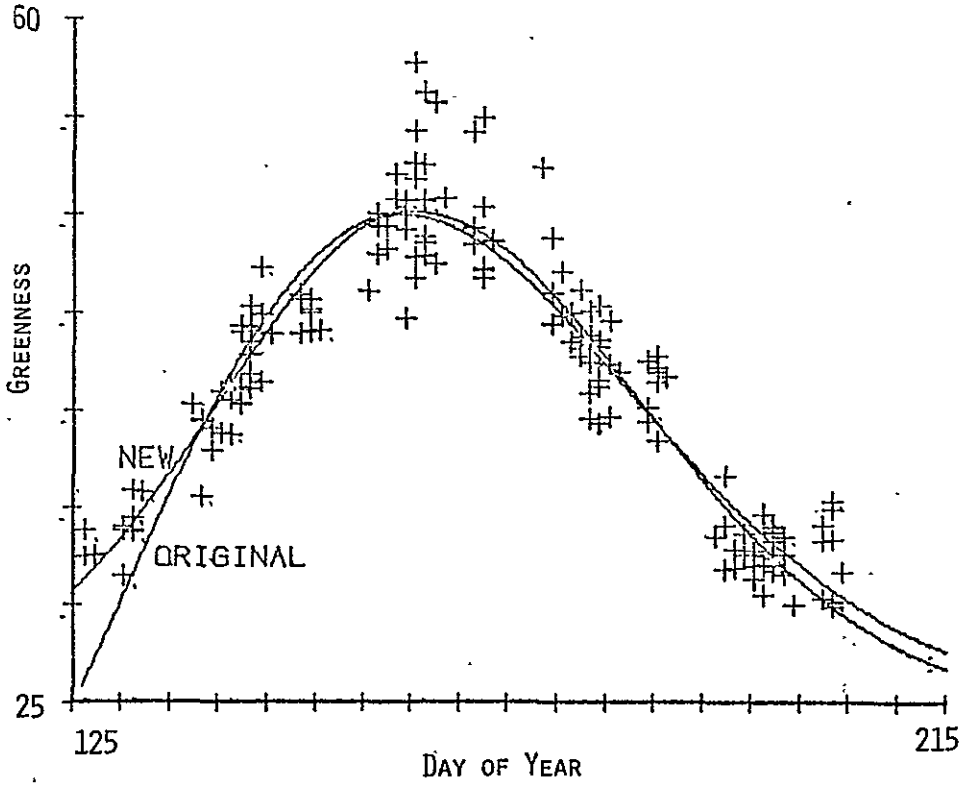


FIGURE A-12. COMPARISON OF MODEL FORMS
(Segment 1669 Field 30)

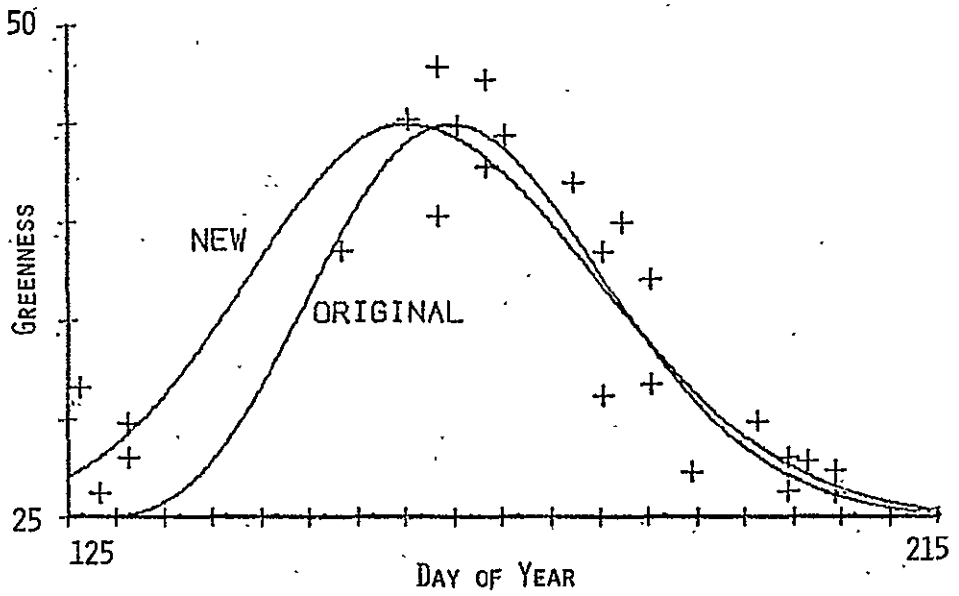


FIGURE A-13. COMPARISON OF MODEL FORMS
(Segment 1929 Field 23)

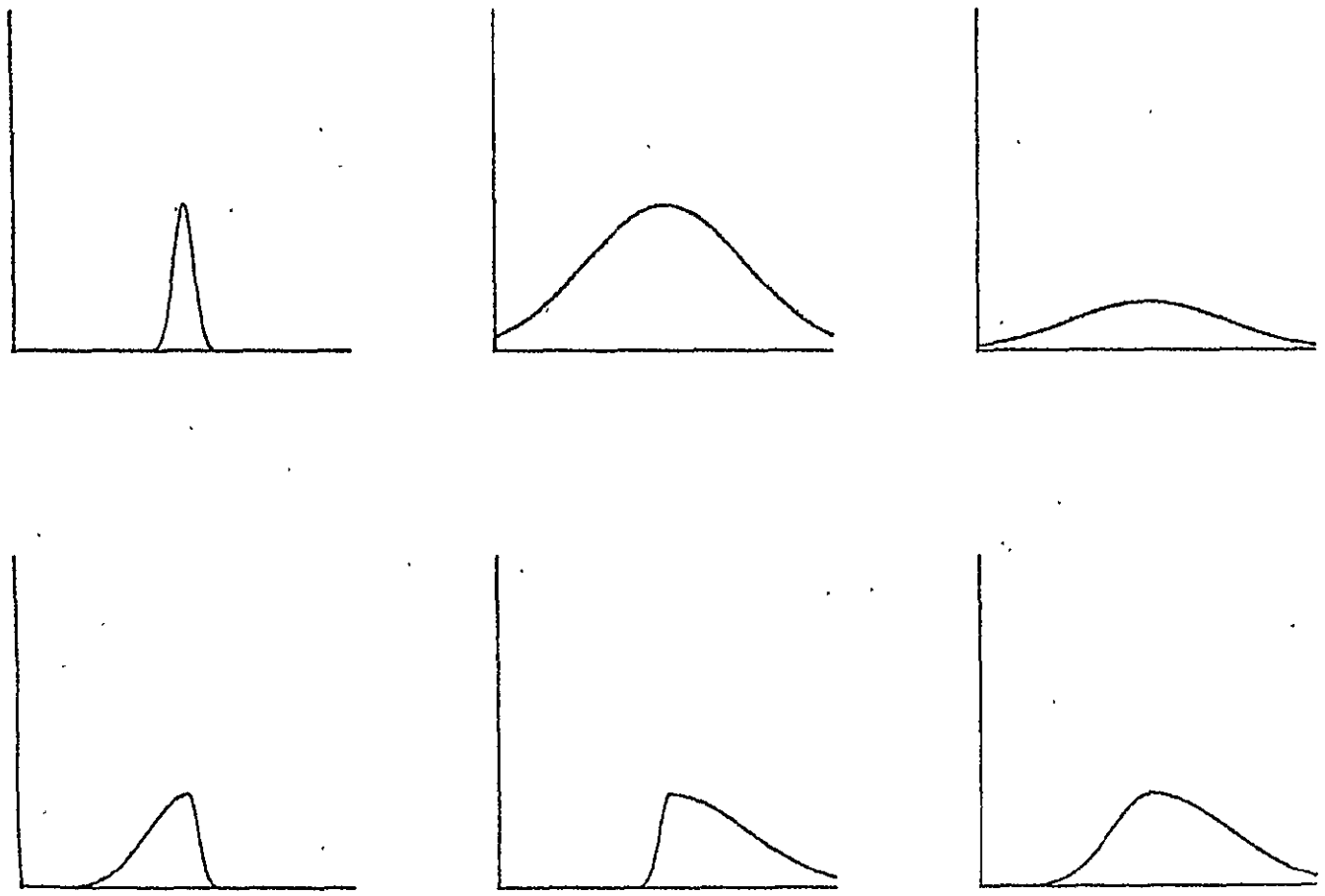


FIGURE A-14. EXAMPLES OF PROFILE SHAPES POSSIBLE WITH NEW MODEL

A.2.4 DATA SELECTION

The final aspect of profile fitting to be addressed, and the first step in the procedure, is data selection. Two alternatives were considered: field interior pixels and field means. Most of the tests, however, are applicable to other possible targets; individual pixels should be much like field means, and quasi-fields should correspond to fields, either means or pixels. It should first be noted that in research efforts where field boundaries are determined by ground truth, mis-registration will likely require that a one-pixel inset be applied at the field edges in order to get a good field signature.

Application of the non-linear fitting technique with either model to field means resulted, in most cases, in a profile that appeared qualitatively the same as one produced using all interior pixels. Two problems, however, are apparent at least on a theoretical level.

First, estimation of crop calendar shift, followed by estimation of three model parameters, computationally requires at least three independent acquisitions in the range of the profile (more than three are desirable). Assuming a growing season of approximately 90 days, and data from only one satellite (i.e., every 18 days), five is the maximum number of observations possible. While two satellites increase the number potentially available to ten, the loss of a few acquisitions due to cloud cover could reduce the total available to a critical level.

In addition, in large fields and to a lesser degree in nearly all fields, there can be expected to be some variation in stage of development within the field. Where this is the case, shifting individual pixels will provide additional information which can be used to better characterize the overall shape of the profile. Thus, one could expect some degree of information loss when using field means. However, since the amount of information lost is unlikely to be critical, by far the

more serious concern is the number of observations available for estimation purposes. If indeed sufficient acquisitions are available, using field interior means in place of field interior pixels should be acceptable.

A.3 PROCEDURE FOR FITTING PROFILES TO TARGET DATA FOR LEVEL 3 APPLICATIONS

A.3.1 DATA SELECTION

Utilize targets defined by field or quasi-field boundaries, with one-pixel insets at borders to account for misregistration. If sufficient acquisitions are available (at least five), means may be used in place of interior pixels.

A.3.2 DATA PREPARATION

The following steps are prescribed for normalizing and standardizing data:

- Satellite calibration correction (see Appendix H).
- Cosine sun-angle correction.
- Cloud, shadow, etc detection [3].
- Spatially-varying XSTAR haze correction [3].
- Estimation of crop calendar shift [3].
- Offsetting of Greenness values by 25 counts.

A.3.3 PARAMETER ESTIMATION

Estimation of parameters for the new model (Equation A-6) should be carried out using a non-linear least squares technique.

APPENDIX B

ANALYSES OF SPECTRAL-TEMPORAL CHARACTERISTICS OF WHEAT

B.1 INTRODUCTION

The overall objectives of the analyses described in this appendix were twofold. The first was to develop discriminative features from spectral data acquired for wheat throughout the growing season. The second was to develop an improved understanding of the spectral characteristics of wheat as functions of the spectral space employed and stage of development.

This appendix has four major topics. Section B.2 addresses methods for analyzing and interpreting patterns in reflectance data in terms of corresponding features in Landsat data. Section B.3 is characteristics of spectral measures of wheat green development, comparing the selected Greenness feature with other measures such as spectral band ratios. The Section B.4 topic is effects of moisture stress and Section B.5 is the estimation of crop development stage from temporal-spectral profiles of Greenness and crop calendar shift calculations.

The data sets analyzed include a sequence of field measurements of wheat reflectance made by USDA personnel in Phoenix, Arizona, [48] and several North Dakota segments from the LACIE Transition Year 1978 Landsat data base. Table B-1 further describes these data sets.

B.2 INTERPRETATION OF REFLECTANCE DATA

It is well established that the majority of variability in Landsat data is confined to two dimensions conveniently described by the Brightness-Greenness plane of the Tasseled-Cap Transformation. It was desired to analyze inband reflectance measurements in a way that corresponds to that transformation. In previous work [49], regression

TABLE B-1. DESCRIPTION OF DATA SETS

A. FIELD-MEASURED REFLECTANCES (Dr. R. Jackson, et al, USDA)

Site: U.S. Water Conservation Laboratory, Phoenix, Arizona

Date: 1977-1978 growing season

Experiment Objective:
Determine water stress effects on wheat yield and spectral characteristics

Experiment Factors:
Moisture treatment - six levels of irrigation

Crop - Two spring wheat varieties and one barley (Planting densities greater than normal for Northern U.S. Great Plains)

Spectral Measurements: Hand-held Landsat-band radiometer

Frequency: Every other day, weather permitting

Agronomic Measurements:
Crop development stage (modified Feekes scale)
Leaf area
Stem length
Wet and dry weights

B. LANDSAT DATA

Sites: North Dakota Segments 1392, 1457, 1461, and 1636

Date: 1978 growing season (See Table B-2 for listing of individual acquisition dates)

Processing Applied:
Calibration adjustments for Landsat 3 to Landsat 2
Sun angle correction
Haze correction using spatially varying XSTAR algorithm
Tasseled-Cap transformation
Selection of field interior pixels
Computation of spectral means for each field
Calculation of day shifts using spectral means

Agronomic Measurements:
Crop development stage (modified Feekes scale)

relationships were established between Landsat values and field-measured (by helicopter-borne spectroradiometer) reflectances for individual fields on several dates. Also, a principal component analysis was conducted of inband reflectances. Ninety-nine or more percent of the variability was found to lie in a plane and a Tasseled-Cap-like transformation was defined by determining two directions, one visually aligned with bare soil data in the principal plane and the other orthogonal to it in the direction of green vegetation reflectance values.

A modified approach was followed this year. The principal component analysis step was repeated, but a different procedure was used to establish the soil line. All bare soil data were isolated and subjected to second principal component analysis; the major component of this analysis was then defined to be the brightness direction, with an orthogonal greenness dimension. In addition, the inband reflectance values were first multiplied by band-to-band calibration ratios based on the calibration values of Landsat. The objective was to provide a weighting of values in the various reflectance bands that was more comparable to the weighting applied by Landsat (for instance the range of MSS7 data is half that of the other bands).

It would be desirable in the future to re-do the regression-type analysis of Reference [49] using data from other sites. One reason is that the prior study did not use the spatially varying XSTAR algorithm. Another is that uncertainties exist about the calibration of those early helicopter-borne reflectance measurements.

As with the earlier principal component analyses, 99% or so of the variability in the data was found to lie in the principal plane. Thus, indications are that senescing vegetation does not lie out of this plane by any significant amount.

B.3 MEASURES OF GREEN VEGETATION DEVELOPMENT

A number of Landsat-derived green measures have been proposed and used by various investigators. Along with Tasseled-Cap Greenness, the list includes MSS7/MSS5, $\sqrt{MSS7/MSS5}$ [50], Green Angle [51], Vegetation Index (VI) [52], and Transformed VI (TVI) [52], among others. We explored the characteristics of these measures using the set of field-measured reflectance data.

Several criteria are appropriate for selecting a green measure for a particular application. These include the shape of its time profile, its stability, its correlation with agronomic variables, and its usefulness for discrimination. The shape is an important factor for use in crop calendar shift operations. For instance, ease of fitting a mathematical form to the profile shape and representation of the overall development process are important. For crop development stage estimation, emphasis of particular development stages can be important. As far as stability is concerned, it is desirable to have both a low variance and a low sensitivity to selected other factors, such as soil color.

Figures B-1 through B-3 present spectral-temporal profiles for the different variables. These profiles were obtained from smoothed reflectance measurements from one plot of data. With cautions that the shapes might be different in Landsat data due to path radiance effects and that only one test plot is represented, a few observations can be made. While there are similarities in the shapes, there also are differences. The peaks of some, especially TVI, are broad and relatively flat. A broad flat peak and steep edges is not the ideal shape for shift calculations; a triangular profile would be better. The peaks of most occur later than that for Greenness.

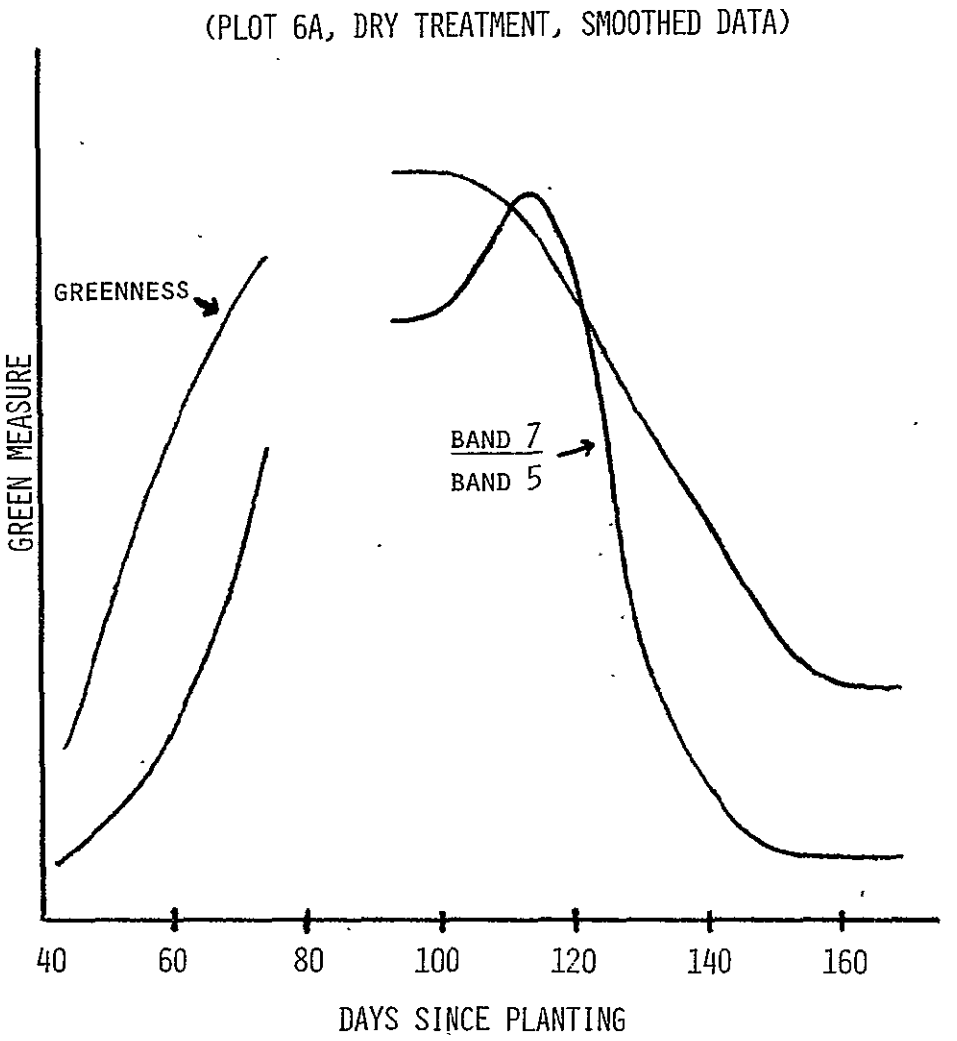


FIGURE B-1. COMPARISON OF REFLECTANCE GREEN MEASURES: GREENNESS AND BAND 7/BAND 5

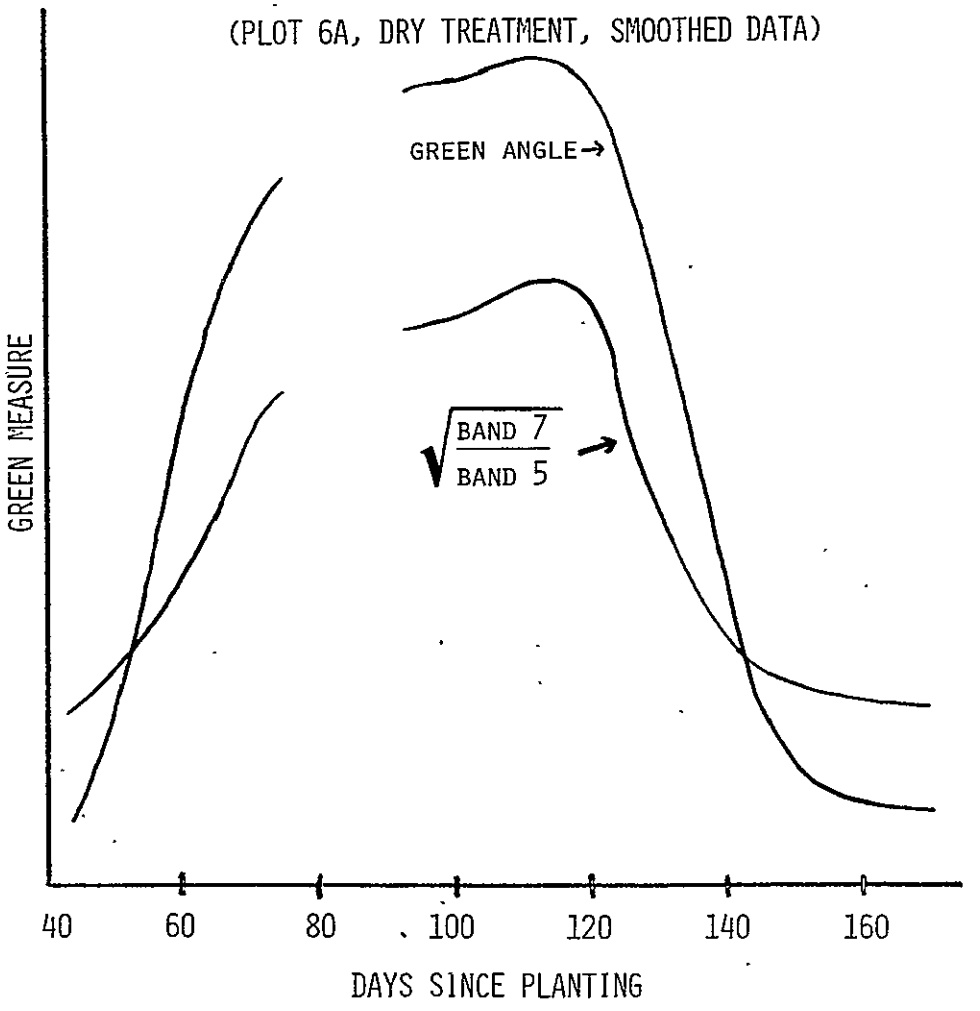


FIGURE B-2. COMPARISON OF REFLECTANCE GREEN MEASURES: $\sqrt{\frac{\text{BAND 7}}{\text{BAND 5}}}$ AND GREEN ANGLE

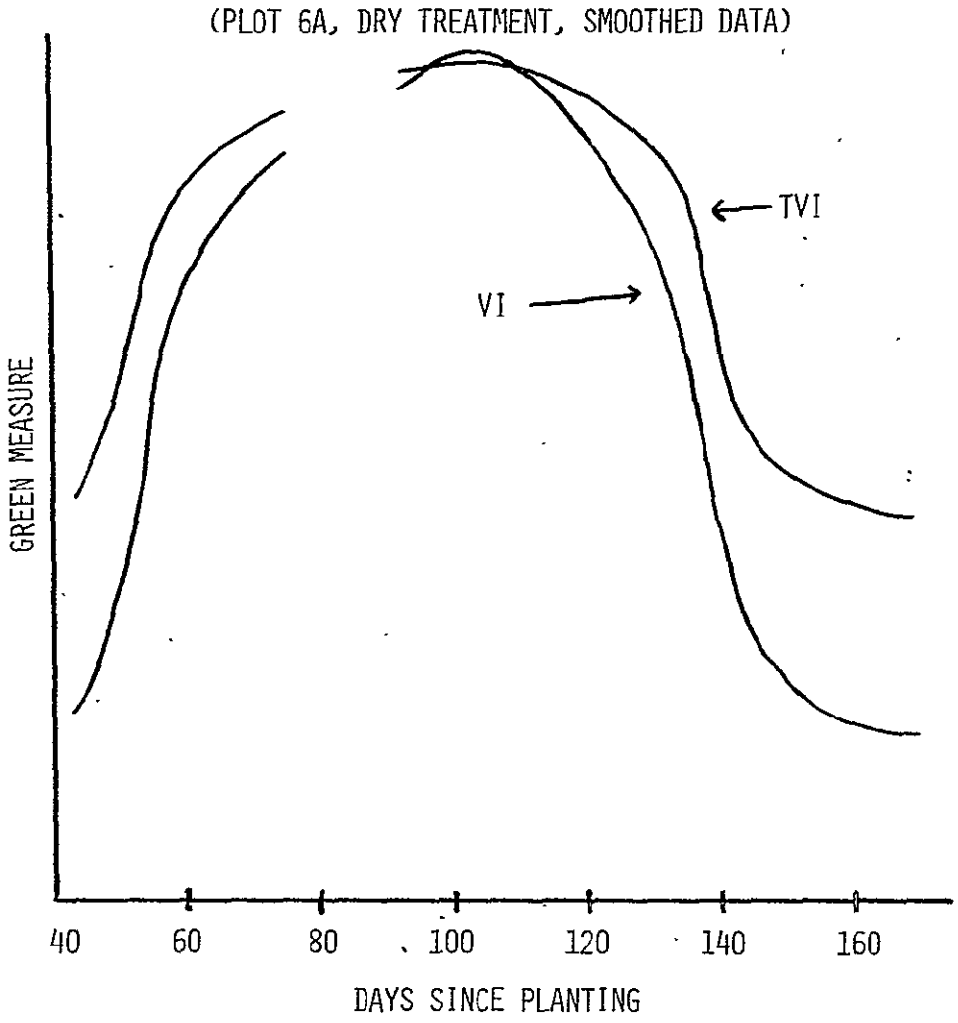


FIGURE B-3. COMPARISON OF GREEN MEASURES: VI AND TVI

The difference in day of peak for Greenness and Band 7/Band 5 is interesting. Insight into the reasons for this can be gained from a plot of the time track of wheat in the plane of Band 7 vs. Band 5 (See Figure B-4). The peak in Greenness corresponds to the peak in Band 7 which occurred at growth stage 9 or 10 which is just prior to heading (See definition of modified Feekes scale in Figure B-5). The Band 7/ Band 5 ratio, on the other hand, reaches its peak when a radial line through the origin is tangent to the left hand side of the loop; this occurred near the end of heading.

It appears that no single green measure will be optimum for all purposes. Landsat data will be needed for final characterization and selection for specific applications. We chose the Greenness variable as the measure of green vegetation for the investigations conducted during the year.

B.4 EFFECTS OF MOISTURE STRESS

The time track in Figure B-4 is for a wheat field that had a dry moisture treatment. Figure B-6 is for a field that had a wet treatment. While both exhibit the looping pattern, the peak value of Band 7 is higher for the wet treatment which had more dense vegetation. The differences were even more pronounced for fields that received more extreme moisture treatments. The wettest field suffered from lodging which caused a marked change in its spectral track.

The Tasseled-Cap-like transformation performs primarily a rotation of the time-track pattern, preserving its distinctive features. Figure B-7 shows the reflectance-space Greenness vs. Brightness plot for the field shown in Figure B-4.

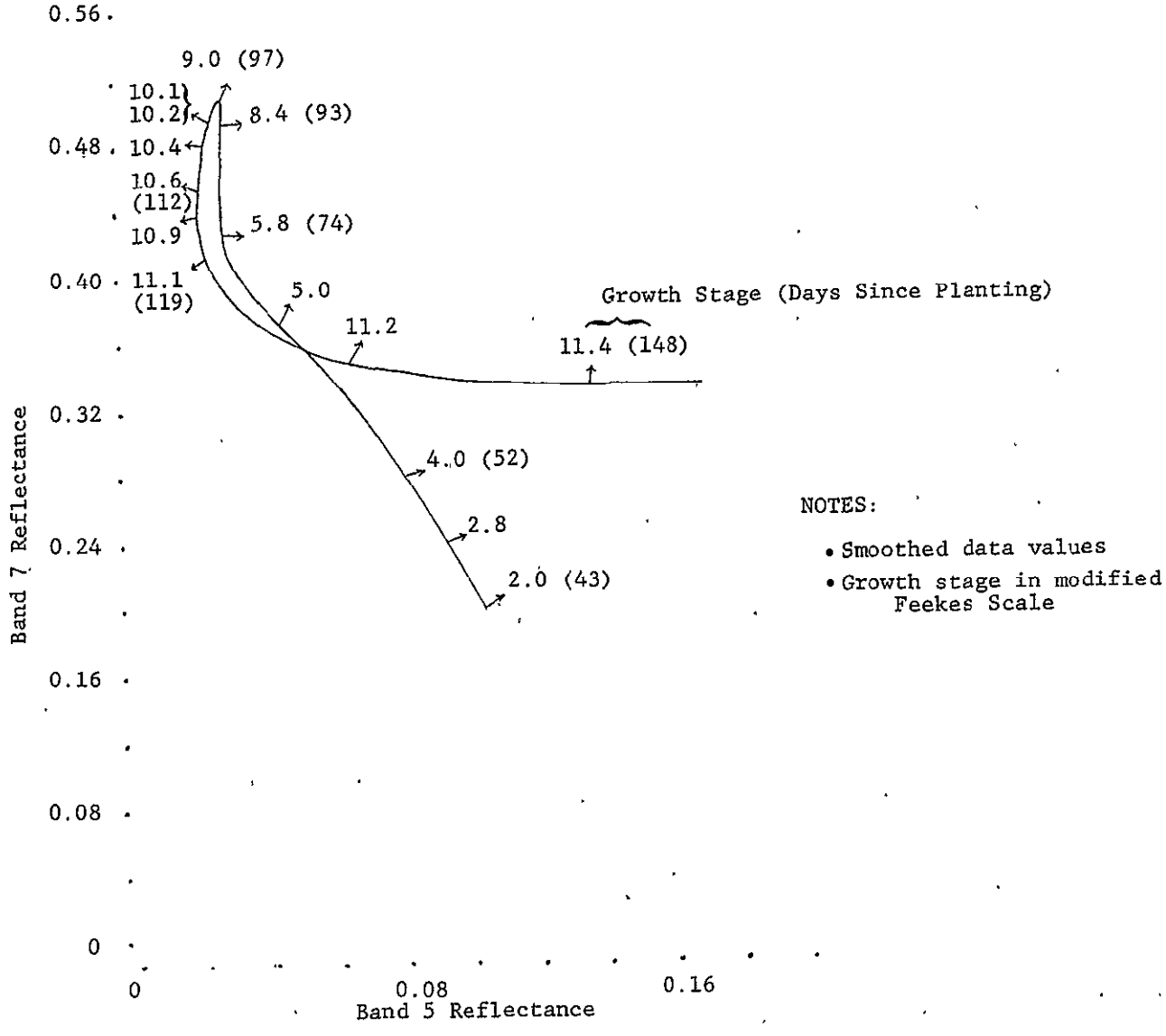
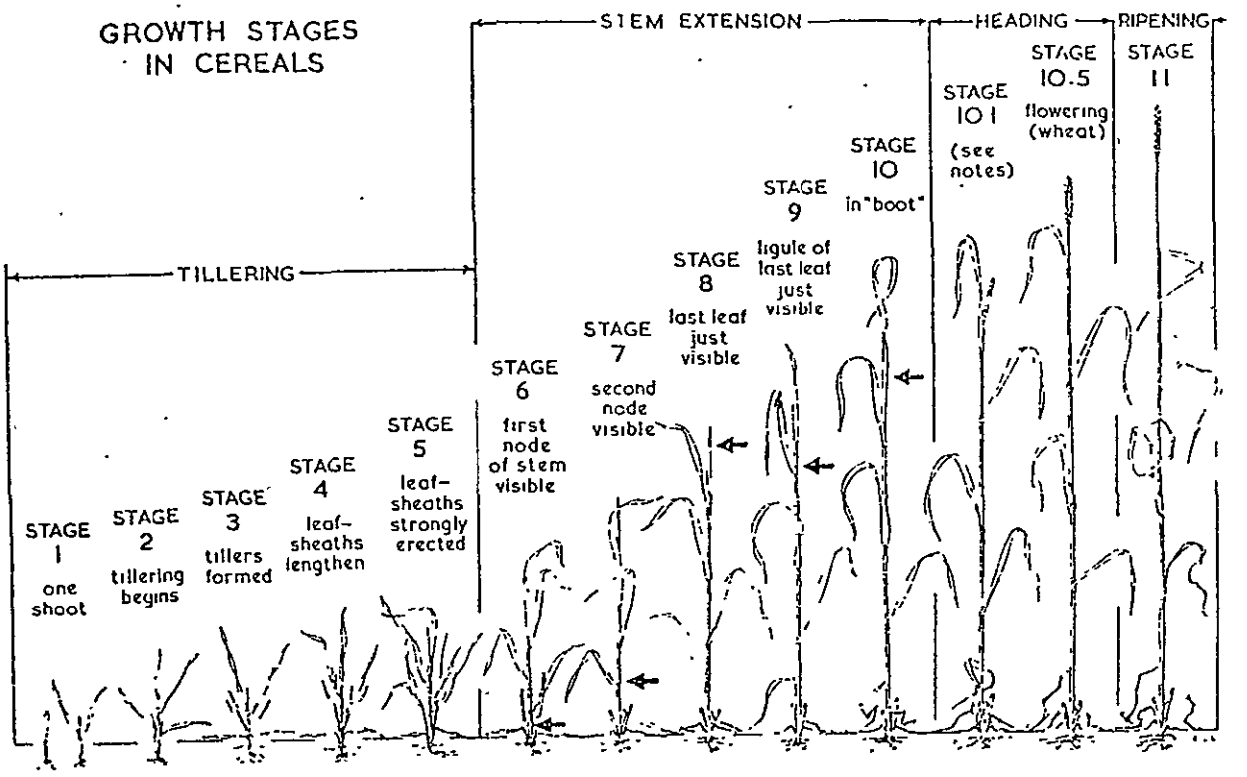


FIGURE B-4. TIME TRACK OF WHEAT REFLECTANCE: PLOT 6A DRY TREATMENT



Stage		
1	One shoot (number of leaves can be added) = 'brairding'	Tillering
2	Beginning of tillering	
3	Tillers formed, leaves often twisted spirally. In some varieties of winter wheats, plants may be 'creeping' or prostrate	
4	Beginning of the erection of the pseudo-stem, leaf-sheaths beginning to lengthen	
5	Pseudo-stem (formed by sheaths of leaves) strongly erected	
6	First node of stem visible at base of shoot	Stem Extension
7	Second node of stem formed, next-to-last leaf just visible	
8	Last leaf visible, but still rolled up; ear beginning to swell	
9	Ligule of last leaf just visible	
10	Sheath of last leaf completely grown out, ear swollen but not yet visible	
10.1	First ears just visible (awns just showing in barley, ear escaping through split of sheath in wheat or oats)	Heading
10.2	Quarter of heading process completed	
10.3	Half of heading process completed	
10.4	Three-quarters of heading process completed	
10.5	All ears out of sheath	
10.5.1	Beginning of flowering (wheat)	Flowering (Wheat)
10.5.2	Flowering complete to top of ear	
10.5.3	Flowering over at base of ear	
10.5.4	Flowering over, kernel watery-ripe	
11.1	Milky-ripe	Ripening
11.2	Mealy-ripe, contents of kernel soft but dry	
11.3	Kernel hard (difficult to divide by thumb-nail)	
11.4	Ripe for cutting. Straw dead	

FIGURE B-5. MODIFIED FEEKES SCALE

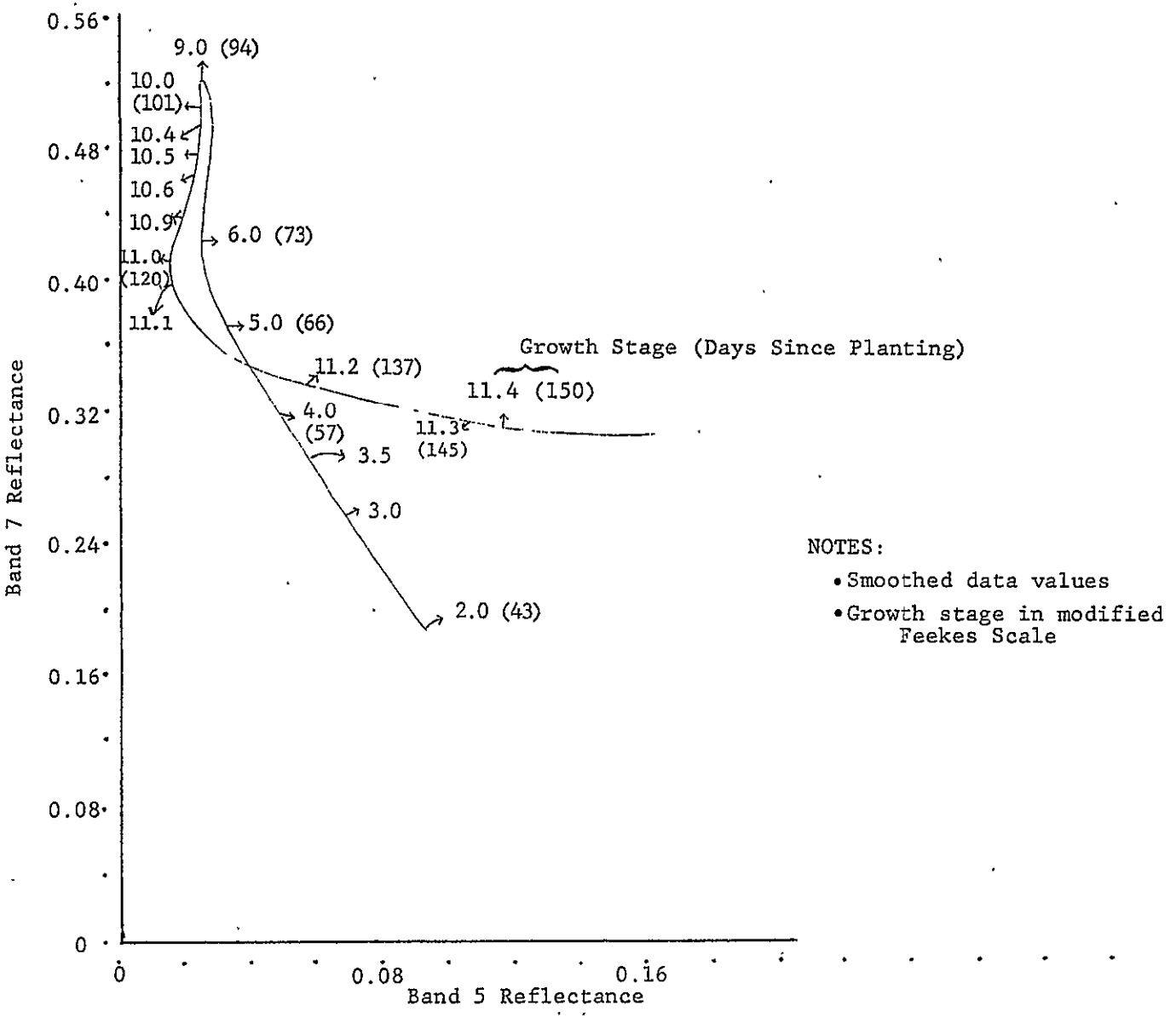


FIGURE B-6. TIME TRACK OF WHEAT REFLECTANCE: PLOT 3A WET TREATMENT

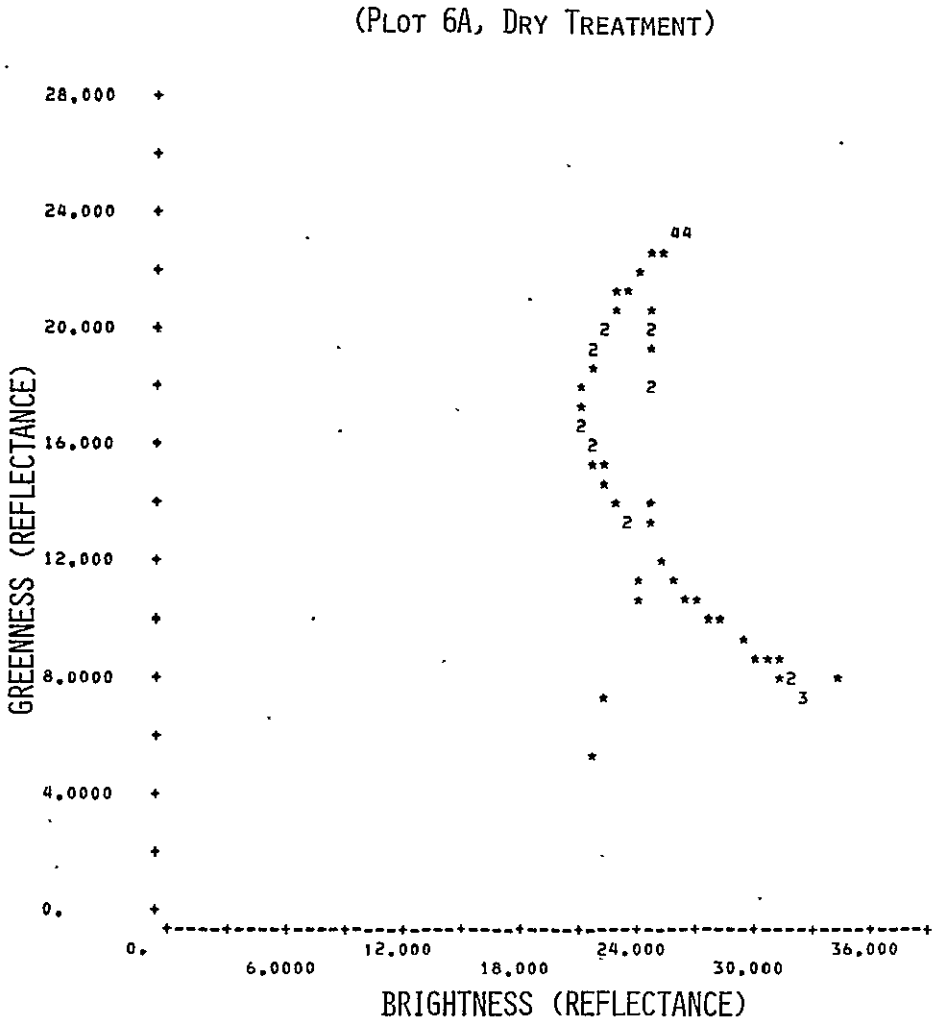


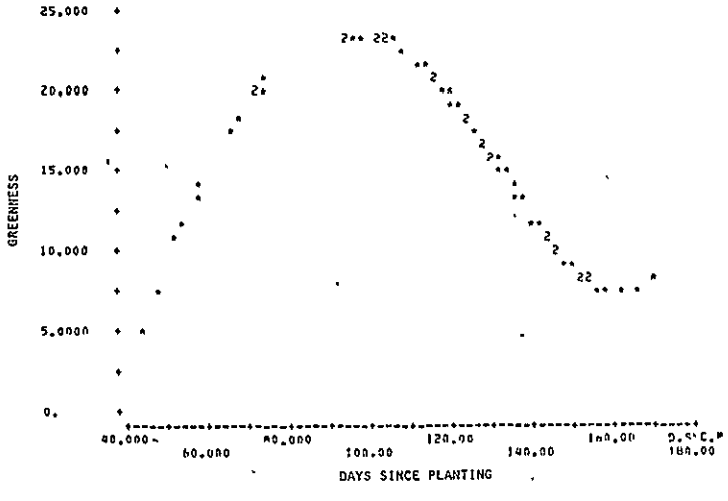
FIGURE B-7. TASSELED-CAP-EQUIVALENT SPECTRAL TRAJECTORY

B.5 ESTIMATION OF CROP DEVELOPMENT STAGE

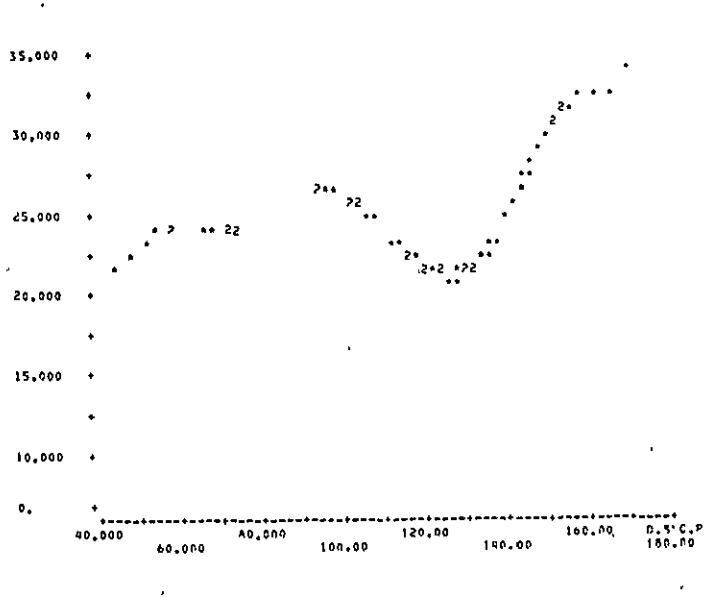
The spectral time track patterns in the preceding figures show a strong correlation with the stage of crop development. We see that peak Greenness occurred just prior to heading, during development of the flag leaf. Reflectances generally decreased proportionally toward the origin during heading to the milky-ripe stage (11.1) of ripening, perhaps due to the opaqueness of heads resulting in increased shadowing. As ripening progressed, the Brightness began to increase again and Greenness continued to decline (See Figure B-8).

Rates of wheat development can vary substantially from location to location. Figure B-9 presents a comparison of selected ground-measured development profiles for spring wheat grown in Arizona and North Dakota. It is such profiles that one wishes to estimate using spectral data. Some of the problems that must be overcome are evident on this graph. First, the Modified Feekes Scale is not linearly related to calendar date; it is compressed at the high end which corresponds to heading and ripening stages. Geographic factors can have an effect, although the planting of spring wheat in mid-December in Arizona is an extreme departure from the more common spring planting in the Northern U.S. Great Plains. The characterization of North Dakota profiles by the 18-day observations is less complete than desirable, but that is the usual interval that Landsat will provide under cloud-free conditions. Underlying both types of curves are uncertainties and variability in ground observations for the fields of interest; some Feekes-scale designations are not readily discernible without detailed examination and within-field variations do exist.

Plots of development stage vs. day of year are presented for 13 fields of North Dakota Segment 1461 in Figure B-10 for which periodic ground observations of development stage were made. The scatter of



(a) Reflectance Greenness



(b) Reflectance Brightness

FIGURE B-8. SMOOTHED SPECTRAL-TEMPORAL PROFILES OF TASSELED-CAP REFLECTANCE VARIABLES (PLOT 6A, DRY TREATMENT)

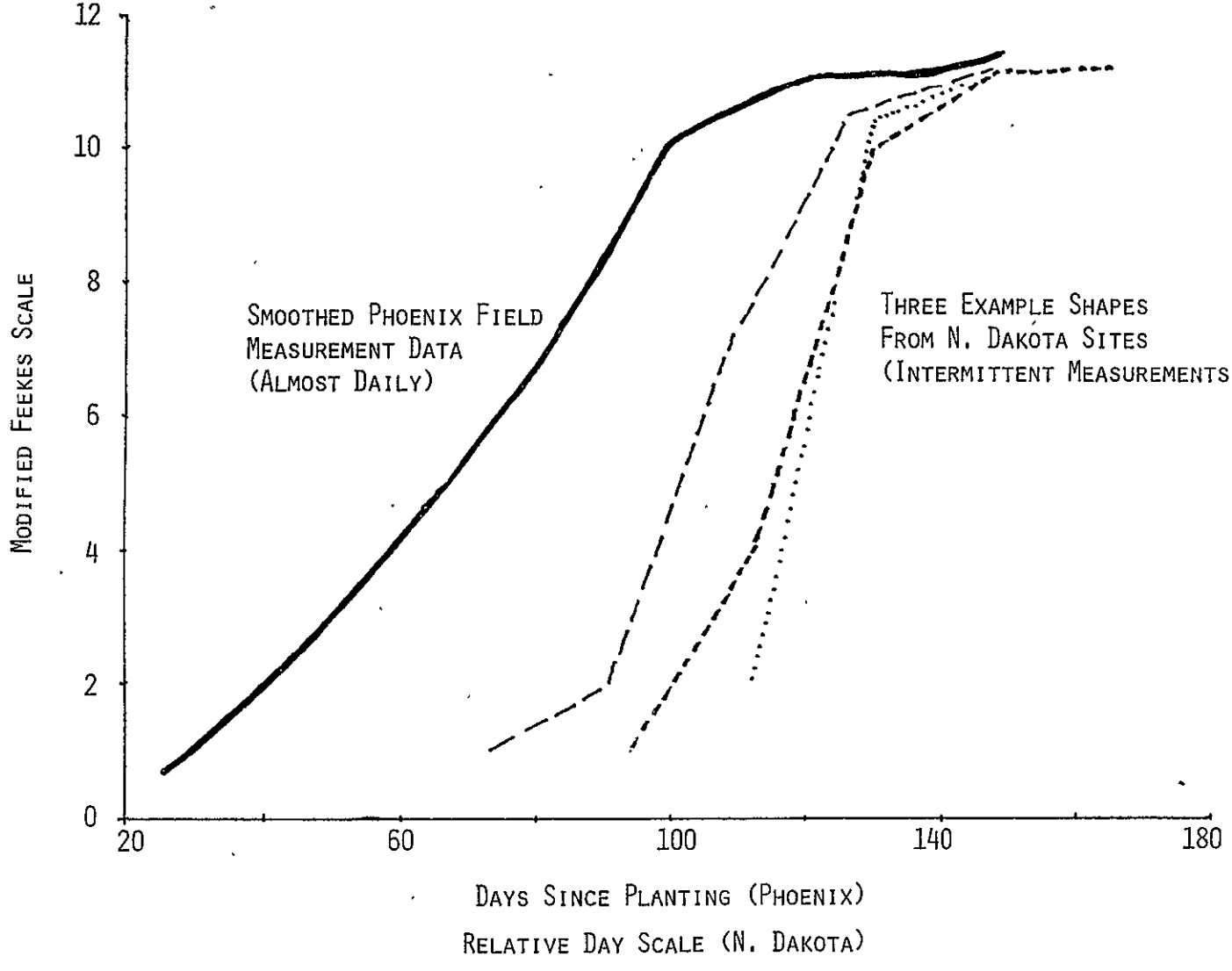


FIGURE B-9. COMPARISON OF TIME PROFILES OF WHEAT DEVELOPMENT

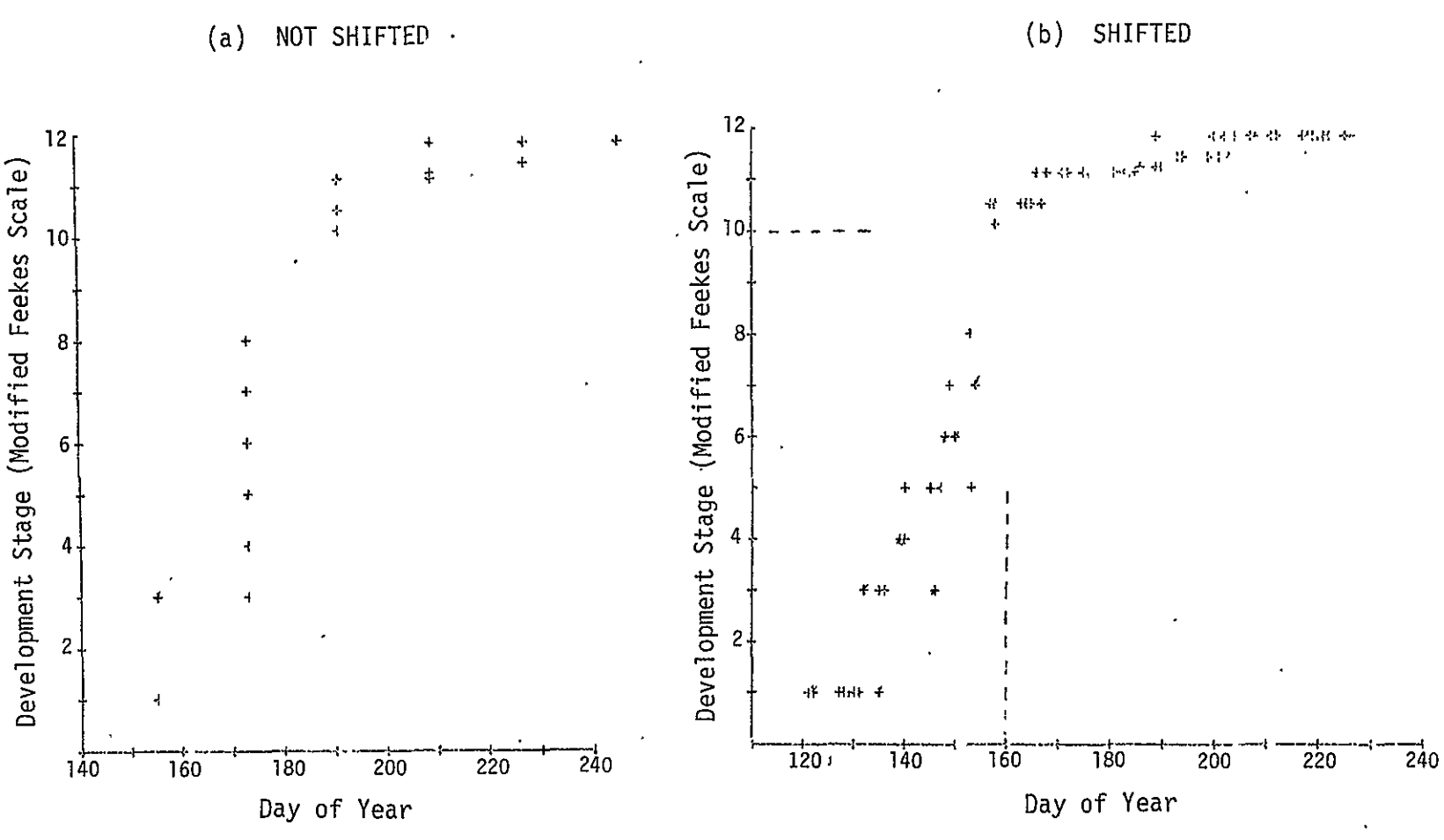


FIGURE B-10. EFFECT OF DAY SHIFT ON WHEAT DEVELOPMENT PROFILES
(Type 1) Segment 1461

points before shifting (Part a) appears greater than afterwards (Part b), especially in the portion corresponding to unshifted day 173. Calculated crop calendar shifts for these fields were used to place the estimated peak Greenness value at shifted day 160 in Part b. In field measurements data, this peak usually occurs just prior to heading, e.g., at or just before Stage 10 (boot). Days shifts calculated for other segments are summarized in Table B-2.

An appreciation for the effectiveness of the crop calendar shift calculation can be gained from Figure B-11. Here, the reduced scatter of the spectral data after shift is striking. Less pronounced improvements were achieved for two other segments.

This initial work in crop development stage estimation did not reach a conclusive stage. However, some issues were identified that could affect future work. It is recommended that ground observations of crop development stage in selected fields in segments be made more frequently than at 18-day intervals, even though spectral data may not be acquired more frequently. The sparseness of Landsat data, especially with missing acquisitions will limit accuracy. The use of green profile shift technology appears helpful but improvements involving more complete characterization of profiles (e.g., level 2 or level 3 applications) and/or use of other spectral features, e.g., Brightness, may be required.

TABLE B-2. ANALYZED SUBSET OF TY 1978 SPRING WHEAT BLIND SITES: LANDSAT DATA (BOTH LANDSAT 2 AND 3)

Segment	Acquisition Days	No. Fields	Range of Shift (Days)
1392	136, 154, 190, 208, 217	12	-27 to -18
1457	156, 174, 228, 246, 264, 273	13	-73 to -35
1461	136, 155, 190, 199, 209, 217, 236	13	-34 to -19
1636	135, 154, 190, 208, 216, 226, 243	12	-47 to -10

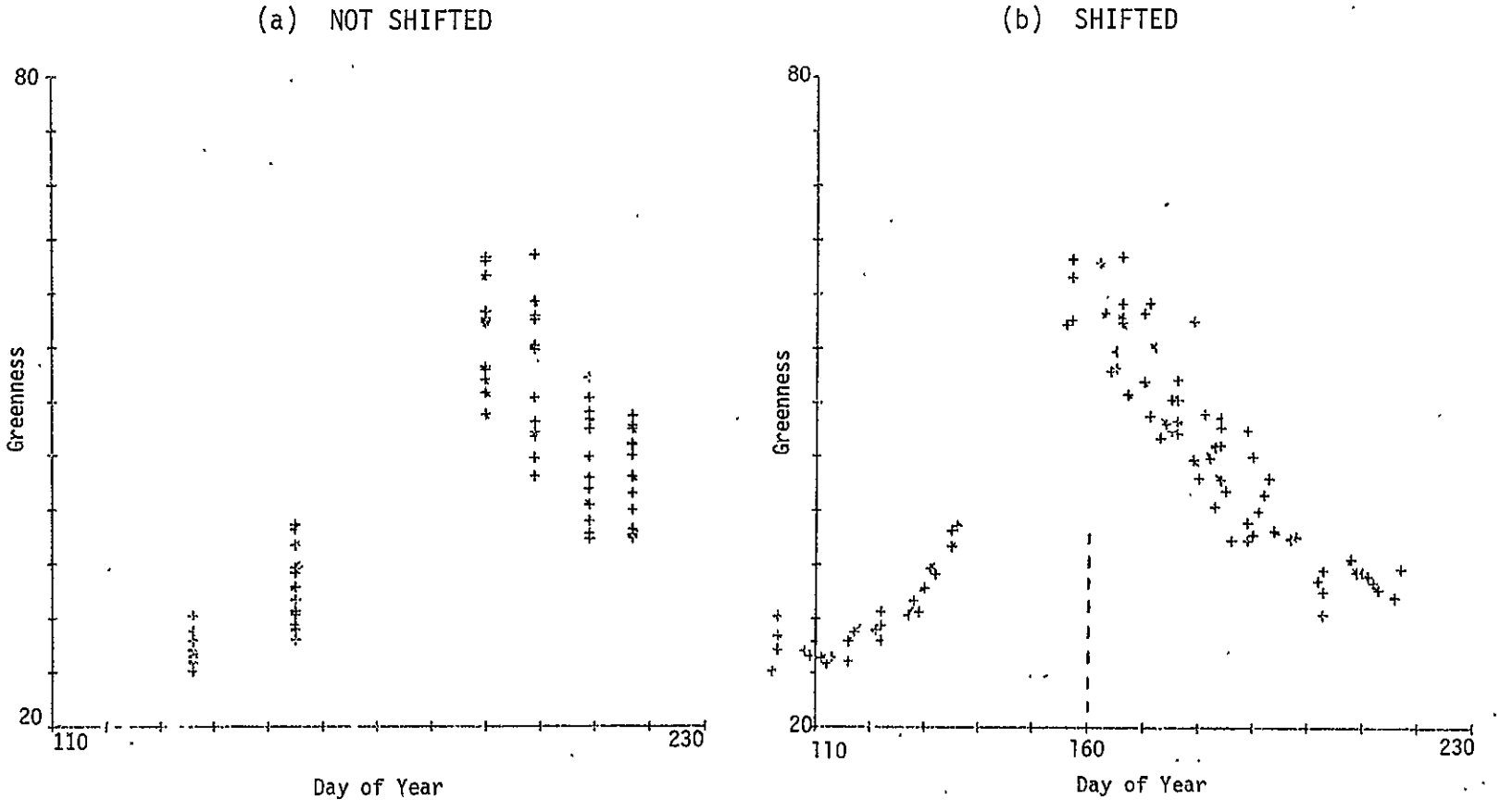


FIGURE B-11. EFFECT OF DAY SHIFT ON LANDSAT GREENNESS PROFILES
Segment 1461

APPENDIX C
ANALYST LABELING/PROCEDURE M EXPERIMENT

The analyst labeling/Procedure M experiment represents a natural continuation of the Procedure M experiment reported in Reference 3. The techniques of testing and evaluating both Procedure M and its components have been refined. The use of analyst labels in this experiment represents an increase in scope over last year's Procedure M experiment. Section 4.9 of Volume I summarizes the results of the analyst labeling study and Section 6.2 describes tests of components of Procedure M. This Appendix provides details of the experiment for completeness.

C.1 PURPOSE AND SCOPE OF EXPERIMENT

1. Gain understanding of analyst labeling of field-like targets.
   a. Examine the use of field-like targets for labeling.
   b. Evaluate individual analyst performance in labeling targets as functions of:
      - target level variables,
      - spectral trajectory strata, and
      - segment-level variables.
   c. Evaluate joint analyst performance in labeling targets.
      - Examination of analyst consistency.
      - Comparison of individual analyst performance to average, vote, and consensus labeling performance.
      - Evaluation of low confidence and mixed target labels.
   d. Evaluate analyst performance and consistency in estimating crop proportions.
   e. Establish a data base of AI labels for evaluation of Procedure M.

2. Test and evaluate Procedure M.
   a. Estimation of performance parameters.
   b. Evaluation of the components of Procedure M.
   c. Study of error propagation by the components of Procedure M.

C.2 DESCRIPTION OF THE EXPERIMENT

C.2.1 PRELABELING PROCESSING OF 18 TY 1978 NORTHERN GREAT PLAINS LACIE SEGMENTS

1. Sun angle correction.
2. Machine screening.
3. Satellite calibration.
4. Haze correction by spatially varying XSTAR.
5. Tasseled-Cap data transformation.
6. Clustering pixels which are close to one another spectrally (Brightness-Greenness) and spatially (line-point) into field-like forms called blobs.
7. Construction of 117 x 196 line printer maps such that:
   - Pixels from little blobs (those with no interior pixels) are printed as ".". This choice of symbol was a mistake; the analysts had much to say about it.
   - Pixels on the boundary of big blobs (those blobs with at least one interior pixels) were printed with a unique character.
   - Pixels in the interior of big blobs were printed with a blank.
   - Production by JSC of PFC image overlays in which all pixels not in the interior of some big blob were blacked out.

C.2.2 LABELING OF TARGETS

1. Sampling of big blobs: Blobs which have at least one interior pixel are called big blobs. These big blobs are currently candidate labeling targets subject to being sampled. In an operational mode, Procedure M would sample on the order of 100 big blobs from each segment.

In order to investigate various sampling strategies and to investigate various analyst labeling attributes, all big blobs were labeled by three LACIE-experienced analysts.

2. Prelabeling training/labeling procedure adjustment: In order to study various parameters associated with error models and analyst consistency, the analysts were asked to work independently. No spectral aids were available in the time frame in which we were conducting this experiment. In order to adjust for the difference in labeling targets, lacks of spectral aids, and the independent labeling, the analysts initially were given about one day to discuss labeling problems and procedures to be used in blob labeling. Segment 1392 was studied during this session, but no labeling was performed at that time. Segment 1392 was the last segment labeled by the analysts. The analysts later were asked to list any segments in which they had to depart from this procedure which they developed. All analysis techniques (i.e., image interpretation), aside from those differences mentioned above were to be the same as those used during LACIE TY operations.

3. Labeling procedure for Procedure M test: The procedural steps for labeling blobs for the Procedure M experiment follow.

a. Each analyst will label the blobs in the 18 TY segments in the order given in Table C-1.
b. Labeling will be performed independently by each analyst.
c. Each analyst will relabel the first three segments after all 18 segments have been labeled (see Table C-1).
d. After labeling each segment, the analyst will fill out the "segment comment forms" shown in Figure C-1.
e. After labeling all 18 segments and prior to relabeling the first three segments, the analyst will complete the "final comment form" shown in Figure C-2.

TABLE C-1. THREE RANDOM PERMUTATIONS OF 18 SEGMENTS

Analyst A	Analyst B	Analyst C
1656	1636	1653
1825	1619	1612
1457	1518	1650
1461	1602	1602
1619	1650	1457
1602	1653	1656
1650	1825	1380
1566	1612	1920
1473	1461	1835
1653	1380	1636
1835	1566	1518
1636	1457	1473
1612	1656	1467
1380	1467	1619
1467	1920	1825
1920	1473	1566
1518	1835	1461
1392	1392	1392

1. Segment Number: _____

2. Analyst's Code: _____

3. Date
   Start: _____
   Completed: _____

4. Number of Hours Required for Labeling: _____

5. Acquisitions:

   A.

Acquisition Date	Primary or Secondary	Comments: (Suitability, Usefulness, or Reasons for Not Using)

   (Use other side if needed)

   B. What are the deficiencies, if any, in the available acquisition history?

6. Comments About the Blob Patterns:

   A. Do the blob interiors seem to be pure?

   B. Do the blob patterns match the field patterns?

   C. Was the choice of blob acquisition optimal? If not, why? Which acquisitions should be used?

   D. Other comments about blob patterns:

7. General Description of Segment. For example, moisture, field size, topography, percent agriculture, etc.

8. Did you have to change your procedure in order to handle this segment? If so, describe how.

9. Other Comments:

FIGURE C-1. SEGMENT COMMENT FORM FOR PROCEDURE M TEST

1. Analyst's Code:

2. Analyst's Background and Years of Experience in LACIE:

3. Analyst's Overall Impression of Blob Interiors as Labeling Targets (Ease of Labeling, Purity, Ease of Finding, Comparison with Dot Labeling):

4. List the Major Problems and Strong Points of Blob Labeling:

5. Comments on Suitability of Image and Map Products:

6. What Procedures Did You Develop to Organize and Use the Product?

7. Recommendations for Improvements in Products, Procedures, etc:

FIGURE C-2. FINAL COMMENT FORM FOR PROCEDURE M TEST

f. Detailed blob labeling instructions are given below.

- For each segment, blob overlays keyed to the LACIE sample segment products (Products 1, 2, and 3) have been generated as well as line printer maps of the blobs for each segment. The line printer maps will be used by the analyst to record the label for each blob.

- A blob is to be labeled spring small grains if it is at least 50% spring small grains, and labeled non-spring small grains if it is less than 50% spring small grains. When there is considerable question about which label to assign a blob, the blob will be labeled based upon the analyst's best guess. If a blob is mixed (i.e., composed of approximately 50% spring small grains and non-spring small grains) it will be flagged and labeled on the line printer map.

- The label for each category will be coded on the line printer map as follows:
  - Spring Small Grains - Red.
  - Non-Spring Small Grains - Green.
  - Questionable Spring Small Grains - Question Mark (?) on Red blob.
  - Questionable Non-Spring Small Grains - Question Mark (?) on Green blob.
  - Mixed blob - The blob will be labeled using the 50% criterion and then outlined in blue.

- The analyst should examine the acquisition listing in Table C-2 in order to become familiar with the acquisitions used for generating the blobs.

- Grid overlays (10 pixels by 10 scan lines) will be keyed to the line printer map in order to facilitate analysis.

TABLE C-2. TRANSITION YEAR DATA[1] USED FOR SPRING WHEAT SEPARABILITY EXPERIMENT

SEGMENT NUMBER	A/D #1	A/D #2	A/D #3	A/D #4	A/D #5	A/D #6	A/D #7
1380	78/169	78/196	78/205*	78/222*	78/232	78/241*	78/249
1392	78/136	78/154	78/190*	78/208	78/217*		
1457	78/156	78/174	78/228*	78/246*	78/264*	78/273	
1461	78/136	78/155*	78/190*	78/199	78/209*	78/217	78/236
1467	78/136	78/154	78/190	78/199*	78/208	78/218*	
1473	78/116	78/197*	78/207*	78/224*	78/269		
1518	78/116	78/135	78/153*	78/188	78/206	78/224*	78/243*
1566	78/115	78/133	78/169*	78/196	78/232*		
1602	78/174*	78/211*	78/228*	78/264			
1612	78/118	78/137	78/155*	78/199	78/218	78/236*	
1619	78/135	78/198*	78/207*	78/216*	78/243	78/252	78/270
1636	78/135	78/154	78/190	78/208*	78/216*	78/226	78/243*
1650	78/156*	78/191	78/209*	78/218	78/228	78/236*	78/246
1653	78/119	78/136	78/155*	78/191*	78/199	78/208*	78/217
1656	78/137	78/155*	78/191	78/209*	78/218	78/263	
1825	78/133	78/169*	78/196	78/206*	78/224*	78/232	78/250
1835	78/134	78/170*	78/187	78/196*	78/224*	78/232	78/241
1920	78/136*	78/199	78/209*	78/218*	78/236	78/271	

[1]Note acquisition dates shown are the subset of available acquisitions which were selected for digital processing.

The acquisition dates marked with an * were used to produce the accompanying blob maps.

- Each of the 18 segments will have a packet consisting of the same material (ancillary data, maps, normal crop calendars, crop calendar adjustment, etc) used during LACIE TY operations. The exception to this is that the spectral aids will not be used for crop identification (because they do not match the blobs).
- All analysis techniques (i.e., image interpretation techniques) will be the same as those used during the LACIE TY operations.

  Note: It is anticipated that the time required to label and complete the evaluation forms will be approximately 10 hours/segment.

C.2.3 DESCRIPTION OF THE DATA BASE

ERIM has constructed data files containing spectral variables, ground truth, and analyst labels on the pixel, blob, and big blob levels. The big blob variables are also contained in an SPSS-formatted system file.

1. Spectral variables: The four Landsat bands were transformed into the Tasseled-Cap variables Greenness and Brightness. All 18 segments in this study have Greenness and Brightness variables for each acquisition on the pixel, blob, and big blob levels.

2. Ground truth: Ground truth for 17 segments (all except segment 1835) is also contained in the same files as the spectral variables on the pixel level, blob level, and big blob level. The ground truth is given in subpixels and aggregated up to the pixel and blob levels.

3. Analyst labels: The analyst labels of big blobs are contained only in the SPSS system file. For each analyst, we have:

   a. grain/other label,
   b. low-confidence label, and
   c. mixed target label.

4. Special variables in the SPSS file: Several variables were created in the SPSS file in order to address the objectives of this experiment.

a. Purity of target.
b. Size of target.
c. Crop category (e.g., summer crop, spring crop, etc).
d. Acquisition history.
e. Crop calendar (profile shift diagnostics).
f. Crop condition (peak green).
g. Spectral trajectory stratum of target.

C.2.4 DESCRIPTION OF THE ANALYSIS OF THIS EXPERIMENT

Analyst labeling: Analyst labeling was investigated both as a component in a segment crop proportion estimation procedure and as a process to be studied for understanding. The analysis of analyst labeling of blobs can be found in Section C.3 of this Appendix. The effect of analyst labeling on grain proportion estimates can also be found in Section C.3. Section C.4 describes how the evaluation of analyst labels as a component within Procedure M will be carried out.

The analysis of analyst labeling of field-like targets was performed for the most part using descriptive statistics. These descriptive statistics included:

a. contingency tables,
b. frequency tables,
c. two-dimensional scattergrams, and
d. plots.

Also, some inferential statistical procedures were used to gain insight into the analyst labeling process. These included:

a. ANOVA
b. Regression
c. Discriminant Analysis

Performance of Procedure M and its components: The reduction of variance of the BLOB and BCLUSTER components are given for each segment in Section C.4 of this Appendix. Histgorams are also given to examine the grain/other separation obtained by the BLOB and BCLUSTER components. The effect of the labeling component and sampling strategies on bias and variance are also studied in Sections C.3 and C.4.

The evaluation of Procedure M's performance is not completed at this time, however, the measurements and methods of evaluation to be used are given in Section C.4.

C.3 INVESTIGATION OF ANALYST LABELING OF FIELD-LIKE TARGETS

C.3.1 EVALUATION OF THE EFFECTS OF TARGET LEVEL VARIABLES

Table C-3 gives the labeling accuracy for various sizes of the blob interior. The labeling accuracy of both grain and non-grain increases as the blob size increases in the range of 1-14 pixels after which the increase in accuracy tends to slow down or to level off.

Table C-4 gives the labeling performance by purity of target. One sees immediately the non-symmetry of the labeling errors. Those blobs with 0-9% grain are (incorrectly) classified grain 2.2% of the time, while those blobs with 90-100% grain are (incorrectly) classified non-grain 100-70.4 = 29.6% of the time. We also see that 46.9 + 19.5 = 66.4% of the big blobs are in these two purity classes. The labeling accuracy of grain within the class 90-100% grain is less than that within the class 80-89% grain. This phenomenon is due to the much lower accuracy of labeling pure oat blobs; 420 of the 1182 grain blobs within the 90-100% grain class are oats. If these 420 blobs are excluded, then the accuracy is about 80%.

TABLE C-3. ANALYST LABELING ACCURACY (PCC) AS A FUNCTION OF QUASI-FIELD SIZE

Size Strata*	Analyst							
	Vote		Green		Red		Blue	
	NG	G	NG	G	NG	G	NG	G
1	91.5	47.5	88.6	48.6	86.8	57.3	93.0	34.6
2	94.5	49.2	90.4	51.6	89.8	57.1	93.9	44.8
3	94.5	66.9	92.4	69.0	89.5	71.5	93.2	46.5
4-6	93.6	63.9	93.8	66.4	87.0	68.2	94.2	47.5
7-9	93.4	67.1	92.6	69.6	87.3	72.5	96.2	47.3
10-14	95.2	72.7	93.5	72.8	88.4	75.5	96.0	51.2
15-19	94.0	71.2	95.7	71.7	88.2	76.3	94.3	57.3
20-28	96.1	76.6	96.7	72.9	91.2	79.1	97.0	61.1
29-46	97.6	76.8	97.9	76.8	93.6	79.3	97.3	58.8
47-120	98.8	70.4	98.8	78.1	95.9	72.8	98.8	55.6

*Number of interior pixels in blob.

TABLE C-4. ANALYST LABELING PERFORMANCE AS A FUNCTION OF QUASI-FIELD PURITY

	Blob Composition (% small grain)									
	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	80-89	90-100
Percent Labeled Grain	2.2	5.8	9.6	18.5	27.6	44.1	63.9	62.9	72.6	70.4
Strata Size	2691	569	250	211	174	195	204	310	453	1182
Weighted Percent Labeled Grain	1.6	4.1	7.0	18.8	28.3	47.5	58.4	66.3	75.5	72.2
Strata Size (Percent of Total Pixels in Big Blob)	46.9	7.1	3.3	3.0	2.6	2.8	3.0	4.4	7.4	19.5

Table C-5 gives the analyst performance on major crop types. We note that, as in the case of Table C-4, non-grains are labeled more accurately than grains. It should be noted that the classification of the strip and unknown category is not known.

Table C-6 gives the percentage of grain labels for each of the grain crops. Spring wheat was labeled as a grain much more often than the others. Barley was more confusing to the analyst than wheat, and oats were more confusing than barley. There were only 14 fields of rye and the majority labeled 12 of them as a non-grain.

Table C-7 gives a breakdown of Table C-6 by segment. Oats in the Minnesota segments (1518 and 1825) were labeled much better than oats in the North Dakota segments. We also note that the very poor labeling of grain in segments 1656 and 1920 is due to the mislabeling of oats. There are wide variations from segment to segment which suggest that the problems encountered by the analysts also vary widely from segment to segment.

Table C-8 gives the percent of non-grain blobs labeled as grain for several crops. Flax turned out to be the only confusing non-grain to Analysts Green and Red, while Analyst Blue labeled all flax blobs as non-grain.

The effect of profile characteristics on analyst labeling accuracy was examined by comparing the Greenness path of each grain blob to an expected Greenness curve. A relationship would imply that analyst labeling accuracy is impacted by crop condition, as reflected through its Greenness trajectory. The standard Greenness curve is given by the expression:

$$F(t) = .65163t^{1.2957}\exp(-1.00052415t^{2})$$

where F = Greenness - 25

t = day of year - 125.

TABLE C-5. PERCENTAGE OF MAJOR CROP QUASI-FIELDS LABELED AS GRAIN*

	Analysts				Percent Purity	Percent of Pixels Considered	No. Segments
	Vote	Green	Red	Blue			
Spring Crops	71.6	73.1	75.3	55.5	93.3	48.2	17
Summer Crops	3.2	3.3	9.0	1.8	94.6	41.6	16
Pasture and Grass	0.8	1.0	3.2	1.5	93.9	74.3	15
Fallow	3.2	4.1	9.3	1.7	91.8	39.8	16
Miscellaneous	1.2	4.7	1.6	1.6	94.3	55.7	15
Unknown and Strip	23.1	11.5	46.2	46.2	90.9	55.9	9

TABLE C-6. PERCENTAGE OF GRAIN CROP QUASI-FIELDS LABELED AS GRAIN*

	Vote	Green	Red	Blue	Percent Purity	No. Quasi-Fields Labeled	No. Segments
Spring Wheat	82.9	84.6	84.9	70.2	91.1	800	17
Barley	64.3	68.2	75.3	38.3	92.6	154	12
Oats	38.5	39.4	40.0	28.0	96.4	267	17
Rye	15.4	21.4	23.1	7.1	93.0	14	5

*Quasi-fields at least 80% pure.

TABLE C-7. AI LABELING ACCURACY (PCC) OF MAJOR GRAINS BY VOTE[1,2]

Segment	Grain	Wheat	Oats	Barley
1380	96.0	94.4	--	--
1392	90.5	90.7	--	--
1457	64.9	81.8	31.7	--
1461	81.9	86.2	--	--
1467	55.3	50.0	30.8	91.7
1473	89.9	97.7	--	78.0
1518	76.9	72.7	92.9	80.0
1566	90.4	100.0	--	--
1602	88.3	91.8	--	--
1612	47.4	84.6	16.1	--
1619	71.2	92.6	--	18.3
1636	59.4	61.4	--	--
1650	76.9	66.7	--	--
1653	71.1	68.0	--	--
1656	9.8	--	8.3	--
1825	82.1	90.9	86.4	64.7
1920	36.7	71.4	20.0	--

[1]Quasi-fields at least 80% pure.

[2]At least 10 quasi-field required.

TABLE C-8. PERCENT OF NON-GRAIN FIELDS LABELED AS GRAIN[1]

	Vote	Green	Red	Blue	Percent Purity	Percent Quasi-Fields	No. Segment
Alfalfa	2.4	2.4	4.9	4.8	91.8	42	11
Corn	2.9	2.2	4.4	1.5	90.4	137	8
Sunflower	0.0	0.0	16.7	0.0	93.5	6	4
Sunflower	2.7	3.2	14.7	0.5	93.7	187	10
Soybeans	1.7	1.7	1.7	2.3	94.8	178	4
Sugar Beets	0.0	0.0	0.0	0.0	93.4	16	3
Flax	21.4	21.4	44.8	0.0	91.3	29	9
Potatoes	0.0	0.0	0.0	0.0	92.4	21	1
Pasture	0.8	0.8	1.0	1.0	93.4	393	14

[1]Quasi-fields are at least 80% pure within the assigned class.

The model for the Greenness of blob i on day t_j+125 is

$$G_i(t_j) = \lambda_i F(t_j - \tau_i) + 25$$

The parameters λ and τ were estimated for each blob.

The estimate of peak greeness for blob i is

$$\lambda_i ||F||_\infty + 25 \approx \lambda_i F(35) + 25 = \lambda_i (34.32) + 25$$

Figure C-3 gives a histogram of labeling accuracy vs. blob shift. The intervals were chosen in such a way that the numbers of blobs per interval were approximately equal. The grain blobs whose shift was between -35 and -12 have an 80% chance of being correctly classified as grain. If the pure grain blobs with low peak Greenness values are excluded, then the labeling accuracy within the interval [-12,100] is about 80%. Those grain blobs whose profile differs by more than 35 days from normal have a much lower probability of being classified as grain. Much of the variation in the values of the blob shifts was explained by variations in the segment mean of these shifts. Figure C-4 gives a histogram of labeling accuracy vs. deviation of blob shift from the segment mean. The intervals were again chosen so that each had approximately an equal number of blobs. Those grain blobs which differed from the segment mean shift by less than 10 days were classified as grain more than 80% of the time. These two figures indicate a tendancy to label according to an expected spectral pattern, with accuracy diminishing as the pattern shifts, due to variations in planting, either early or late.

Figure C-5 gives the histogram of grain labeling accuracy vs. various levels of peak Greenness. Labeling accuracy is under 50% for those pure grain blobs with peak Greenness values under 46. This implies that grain fields that are either stressed or of low canopy cover, tend to be identified as non-grain. It was noted that a high peak Greenness offset the tendancy to mislabel in the presence of late shifts.

Blob Purity ≥80%
Oats Excluded
North Dakota, Segment 1619 Excluded

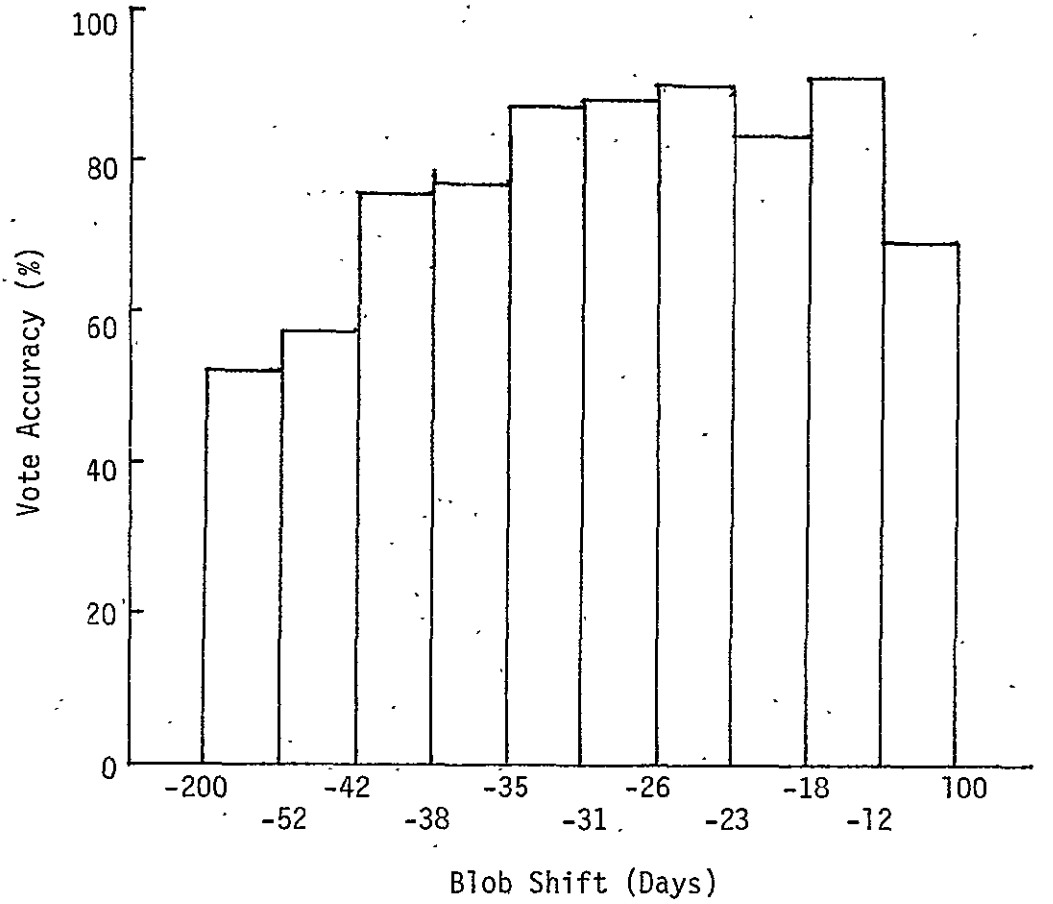


FIGURE C-3. SPRING SMALL GRAINS LABELING ACCURACY VS. SHIFT

Relative Blob Shift = Blob Shift - Average Shift in Each Segment

Blob Purity ≥80%
Oats Excluded
North Dakota, Segment 1619 Excluded

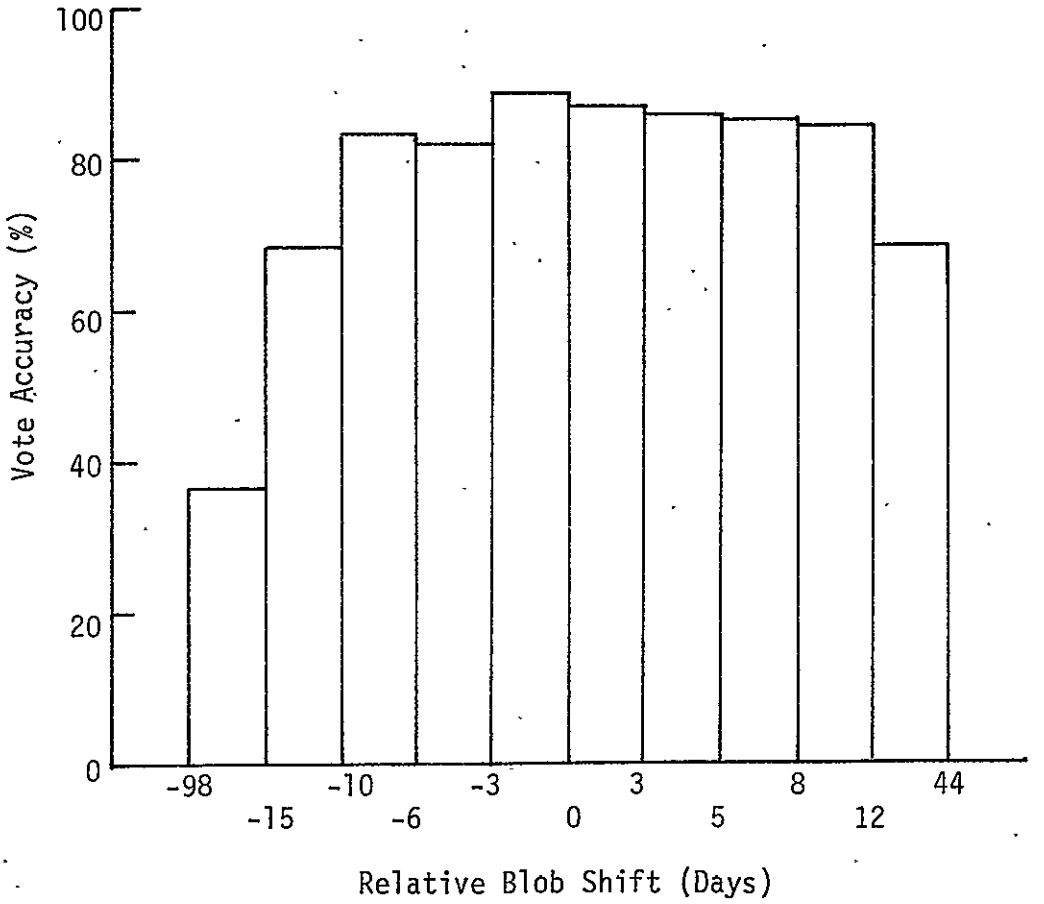


FIGURE C-4. SPRING SMALL GRAINS LABELING ACCURACY VS. RELATIVE SHIFT

Blob Purity ≥80%
Oats Excluded
North Dakota, Segment 1619 Excluded

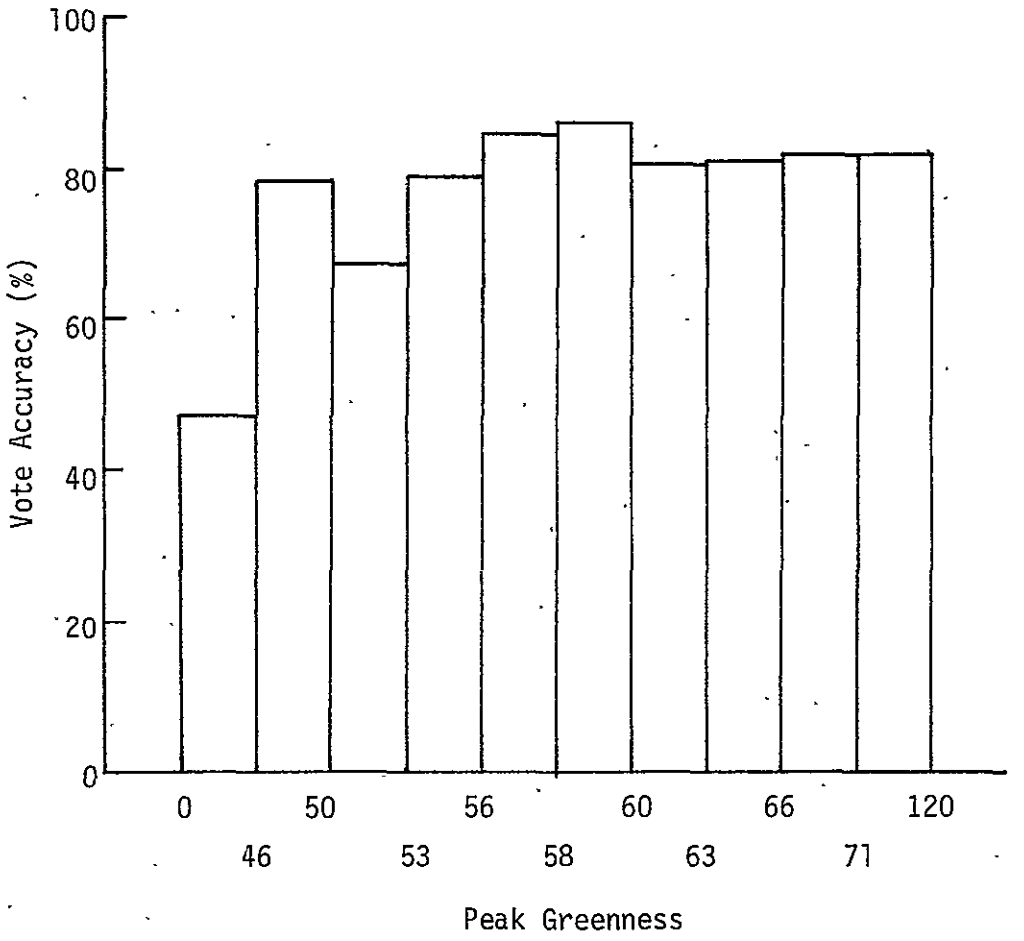


FIGURE C-5. SPRING SMALL GRAINS LABELING ACCURACY VS. PEAK GREENNESS

C.3.2 EVALUATION OF JOINT ANALYST PERFORMANCE

Analyst consistency: Table C-9 gives the percent of decisions on which the pairs of analysts and all analysts agreed for each segment The highest consistency shown by all three analysts occurred in segment 1380 where the analyst agreed on 92.7% of their labels, while the lowest consistency occurred in segment 1467 with only 55.7% agreement.

Table C-10 gives the percentage that pairs of analysts correctly agreed with each other for several crop types. For example, Analysts Green and Red both correctly labeled grain targets as grain 61.2% of the time and Analysts Red and Blue correctly labeled summer crops as non-grain 92.1% of the time. As noted earlier, the analysts do much better with non-grain than grains.

Table C-11 gives, for several crops, the percentage in which none were correct (0 of 3), one was correct (1 of 3), two were correct (2 of 3), and all were correct (3 of 3). For example wheat was mislabeled as non-grain by all analysts (0 of 3 were correct) 7.7% of the time, wheat was labeled as grain by only one analyst (1 of 3 correct) 8.9% of the time, wheat was labeled as grain by two of the analysts (2 of 3 correct) 18.3% of the time, and all three analysts (3 of 3 correct) 65.1% of the time. It is interesting that oats are missed by all three analysts more than half of the time.

The columns of Table C-12 were constructed from the columns of Table C-11 using the following relations:

$$P(\text{correct label 2 analyst agreed}) = \frac{P(2 \text{ of } 3 \text{ correct})}{P(1 \text{ of } 3 \text{ correct}) + P(2 \text{ of } 3 \text{ correct})}$$

and

$$P(\text{correct label 3 analyst agreed}) = \frac{P(3 \text{ of } 3 \text{ correct})}{P(0 \text{ of } 3 \text{ correct}) + P(3 \text{ of } 3 \text{ correct})}$$

TABLE C-9. ANALYST CONSISTENCY - PERCENT OF DECISIONS IN AGREEMENT

Segment	Green Red	Green Blue	Red Blue	Red Green Blue
1380	94.9	96.4	94.2	92.7
1392	82.6	73.9	66.7	61.6
1457	90.6	83.4	83.5	79.1
1461	88.0	83.6	80.6	76.3
1467	79.4	60.5	71.5	55.7
1473	91.2	84.3	83.3	79.5
1518	91.7	82.2	80.5	77.0
1566	91.4	84.8	86.2	81.0
1602	93.0	91.8	91.1	88.0
1612	92.0	90.2	90.5	86.4
1619	88.5	85.9	87.6	80.0
1636	86.7	85.8	84.4	78.5
1650	80.8	79.9	81.3	71.2
1653	73.9	85.2	76.5	68.0
1656	91.0	92.8	91.9	87.9
1825	90.3	89.0	84.4	81.9
1835	92.2	92.9	92.0	88.8
1920	91.9	89.0	89.8	85.2

TABLE C-10. ANALYST CONSISTENCY - PERCENT OF DECISIONS IN CORRECT AGREEMENT

Analyst Pair	Non-Grain	> 50% Grain	> 80% Grain	Wheat	Oats	Barley	Summer Crop	Pasture & Grass
Green Red	87.0	61.2	77.4	79.6	33.6	60.0	99.7	99.1
Green Blue	91.1	44.2	49.8	66.1	23.0	33.1	97.8	99.1
Red Blue	87.9	46.8	48.2	67.4	24.3	37.3	92.1	97.6
Vote	94.7	66.3	71.0	82.9	38.5	64.3	96.8	99.2

TABLE C-11. ANALYST CONSISTENCY - PERCENT OF DECISIONS WITH INDICATED AGREEMENT

No. Analysts Correct	Non-Grain	> 50% Grain	> 80% Grain	Wheat	Oats	Barley	Summer Crop	Pasture & Grass
0 of 3	1.7	19.8	17.1	7.7	52.7	14.9	1.2	0.1
1 of 3	3.5	13.6	11.6	8.9	9.2	19.6	1.9	0.6
2 of 3	9.1	23.5	22.5	18.3	16.9	32.4	6.8	3.9
3 of 3	85.7	43.1	48.8	65.1	21.2	33.1	90.2	95.3

TABLE C-12. PROBABILITY OF CORRECT LABELING AS A FUNCTION OF AGREEMENT

No. Analysts In Agreement	Non-Grain	> 50% Grain	> 80% Grain	Wheat	Oats	Barley	Summer Crop	Pasture & Grass
2	72.2	63.3	66.0	67.3	64.8	62.3	78.2	86.7
3	98.1	68.5	74.1	89.4	28.7	69.0	98.7	99.6

In every case except oats, the probability that a correct label is obtained when all three analyst agreed is higher than when only two agreed. The accuracy of the labeling of oats was very poor in the North Dakota segments and very good in the Minnesota.

Each analyst labeled his first three segments twice. The second labeling occurred after all segments were labeled. Table C-13 gives the analyst consistency for the segments each analyst repeated. Analyse Green made by far the fewest changes while Red made the most changes, and also had the only significant increase in accuracy.

Comparison of average and majority vote labeling to individual analysts: The three grain/non-grain labels were combined to form a majority vote label. The majority vote label was defined to be the label given by the majority of the analysts.

Table C-14 gives labeling accuracies for each analyst and vote for each segment for grain and non-grain. Analyst Red had the highest labeling accuracy for grains (71.1%). Analyst Green's grain accuracy (68.6%) was significantly higher than Analyst Blue's grain accuracy (51.0%). Analyst Blue's non-grain accuracy (95.3%) was the highest followed closely by Analyst Green's (93.9%) and Analyst Red's (89.8%).

Table C-15 gives the ground truth statistics for each segment, normalized for bad ground truth.

Table C-16 gives the segment grain proportion estimates for each analyst, vote, average, and binary ground truth (blobs are given grain/other label depending on ground truth).

Table C-17 gives the bias of the segment proportion estimate, standard deviation and root mean square (RMS) of the errors for each analyst, vote and average. We note the following orderings:

$|\text{bias (red)}| < |\text{bias (green)}| < |\text{bias (vote)}| < |\text{bias (ave.)}| < |\text{bias (blue)}|$,

SD (blue)<SD (ave.)<SD (vote)<SD (red)<SD (green), and

RMSE (ave)<RMS (red)<RMS (vote)<RMS (green)<RMS (blue).

TABLE C-13. ANALYST CONSISTENCY IN LABELING THE SAME SEGMENTS AT THE BEGINNING AND END OF THE EXPERIMENTS

Analyst/ No. Segments	Initial Accuracy		Rework Accuracy		Total Changes	
	G	NG	G	NG	Total	Correct
Green (3)	62.9	96.4	60.0	95.7	74	29
Red (3)	66.7	92.1	78.9	92.0	175	116
Blue (3)	51.2	91.7	52.8	91.6	141	67

TABLE C-14. LABELING ACCURACIES FOR GRAIN AND NON-GRAIN

	Vote		Green		Red		Blue	
Segment	NG	G	NG	G	NG	G	NG	G
All	94.9	66.8	93.9	68.6	89.8	71.1	95.3	51.0
1392	90.1	82.7	87.2	85.8	75.6	91.3	95.4	42.4
1457	97.3	62.6	94.5	67.6	96.7	67.4	97.3	43.8
1461	90.9	79.2	90.9	79.2	88.0	86.0	97.9	57.5
1467	80.4	54.5	75.8	67.6	74.5	57.7	90.8	25.9
1473	98.7	88.0	93.5	86.9	98.7	96.7	98.0	69.4
1602	95.1	74.6	96.6	73.1	93.4	79.7	93.1	83.6
1612	94.5	50.6	95.8	48.1	92.8	51.9	91.6	38.3
1619	98.2	65.7	97.0	70.3	94.4	68.6	97.5	54.9
1636	92.7	57.1	91.6	62.0	86.8	58.5	94.8	58.0
1650	92.4	59.2	93.4	55.4	87.2	61.1	91.3	68.1
1653	95.2	63.6	98.4	33.8	78.9	80.6	91.5	61.9
1656	97.5	10.1	98.3	17.4	94.7	2.9	95.5	17.4
1920	98.4	41.4	97.8	37.8	95.1	45.5	96.7	30.6
1380*	94.1	100.0	92.3	88.5	95.0	92.3	93.8	96.2
1518*	99.1	68.2	97.3	69.7	95.9	74.4	99.6	34.4
1566*	97.4	88.3	93.9	90.1	93.3	91.0	99.0	69.8
1825*	97.5	79.6	95.9	81.0	91.0	83.8	98.3	64.0

*Minnesota Segments

TABLE C-15. SEGMENT GROUND TRUTH STATISTICS

(Grain Proportions, Big Blobs Only)

Segment	Percent Grain	Percent Wheat	Percent Barley	Percent Oats	Percent Rye	Percent Unknown
1380*	6.19	4.91	.005	1.28	0.0	2.9
1392	33.0	26.4	5.4	1.1	0.0	1.5
1457	50.5	37.0	1.2	12.3	0.0	1.0
1461	40.2	31.0	4.6	3.4	1.3	6.0
1467	57.1	35.8	10.7	10.6	0.0	3.0
1473	49.7	31.8	17.0	0.6	0.3	5.3
1518*	34.86	24.44	2.80	7.47	0.16	1.74
1566*	35.28	22.78	6.35	5.84	0.32	4.36
1602	30.4	26.6	1.1	1.9	0.9	0.6
1612	26.6	11.1	0.3	15.0	0.2	0.5
1619	47.7	35.8	11.5	0.4	0.0	1.3
1636	42.5	35.7	2.1	3.7	0.9	5.8
1650	29.5	23.3	1.3	4.6	0.2	5.7
1653	19.0	14.8	0.4	3.7	0.1	3.3
1656	15.8	2.7	0.4	12.7	0.1	0.9
1825*	33.12	17.48	5.44	9.73	0.5	6.5
1920	29.8	14.9	0.5	14.3	0.1	0.4
Average	35.6	24.6	4.1	6.5	.25	2.5

*Minnesota Segments

TABLE C-16. SEGMENT PROPORTION ESTIMATES*

Segment	Binary Ground Truth (% Grain)	Proportion Estimates of Grain (Aggregated Blob Labels)				
		Green	Red	Blue	Vote	Average
1392	.37	.395	.483	.209	.374	.362
1457	.53	.375	.364	.242	.343	.327
1461	.45	.475	.528	.338	.459	.447
1467	.25	.576	.474	.211	.443	.421
1473	.56	.533	.564	.444	.530	.514
1602	.29	.254	.290	.267	.263	.270
1612	.21	.110	.132	.117	.112	.120
1619	.57	.461	.462	.498	.436	.474
1636	.48	.347	.358	.335	.324	.347
1650	.22	.182	.214	.211	.196	.202
1653	.16	.073	.265	.146	.141	.161
1656	.15	.032	.036	.044	.024	.037
1920	.29	.144	.195	.115	.155	.151
1380**	.05	.098	.108	.083	.102	.096
1518**	.36	.274	.312	.114	.267	.233
1566**	.36	.363	.369	.267	.334	.333
1825**	.40	.357	.383	.288	.345	.343

*Blobs which have at most 10% bad ground truth.

**Minnesota Segments

TABLE C-17. BIAS, STANDARD DEVIATION AND ROOT MEAN SQUARE OF ERRORS

Source of Labels	Bias $(\bar{P}-P)$	Standard Deviation of $\bar{P}-P$	RMSE $\sqrt{(\bar{P}-P)^2}$	Performance Ordering
Analyst Green	-.038	.110	.116	4
Analyst Red	-.010	.099	.100	2
Analyst Blue	-.104	.081	.132	5
Vote	-.050	.089	.102	3
Average	-.051	.084	.098	1

The lower variance term of analyst average offsets the lower bias term of Red to give better overall results (RMSE). In terms of performance, analyst average was best, closely followed by Analyst Red, vote and Green.

Simple linear regression was used to predict ground truth grain given the labels from each analyst, vote and average. The results are plotted in Figures C-6 through C-10. The regression lines are all plotted on Figure C-11. A multiple regression was used to predict ground truth grain given the proportion estimates from each analyst. The results are given in Table C-18. Analyst Blue's proportion estimate was best in the sense that the correlation coefficient was higher than those of Analysts Red and Green, and the correlation coefficient of Analyst Blue was almost as high as the multiple regression coefficient using all three analysts. The regression coefficient for Analyst Blue was the only one to be significantly different from zero (5% level). The simple correlation coefficient for Analyst Blue is higher than those for average and vote. This is very interesting since the RMSE performance of Blue is the lowest. RMSE($\hat{P}$) is a measure of the distance between the points ($\hat{P}_i$, P_i) and the 45° line through the origin. The proportion estimates of the other sources of labels fit this line better than the proportion estimate of Blue. Simple linear regressions will fit the best line through the data. Thus, we conclude that Blue's proportion estimate has the best linear trend, but the y-intercept and the difference of the slope from 1 causes Blue's RMSE to be higher than the RMSE of the other sources of labels.

C.3.3 INVESTIGATION OF RELATIONSHIP BETWEEN ANALYST PERFORMANCE AND LABELS OF "LOW CONFIDENCE" AND "MIXED TARGET"

The snalyst, in addition to labeling a blob grain/non-grain, could flag a blob as "low confidence in label" and/or "mixed targets." The analysts used these labels very rarely. There were very few blobs

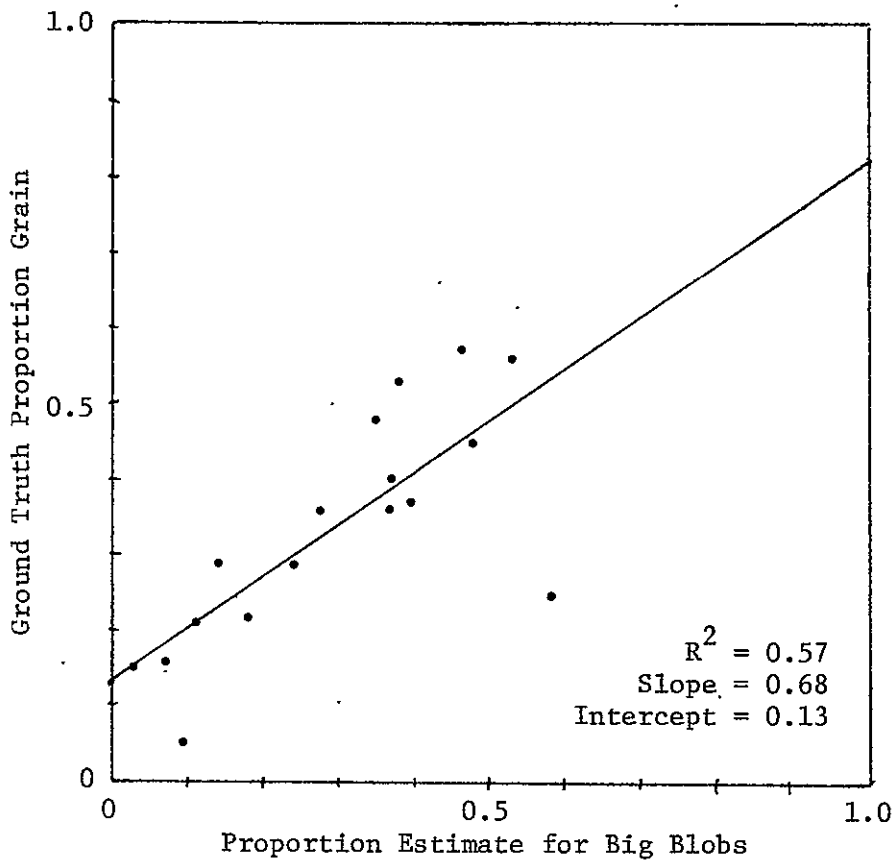


FIGURE C-6. ANALYST GREEN'S PROPORTION ESTIMATE

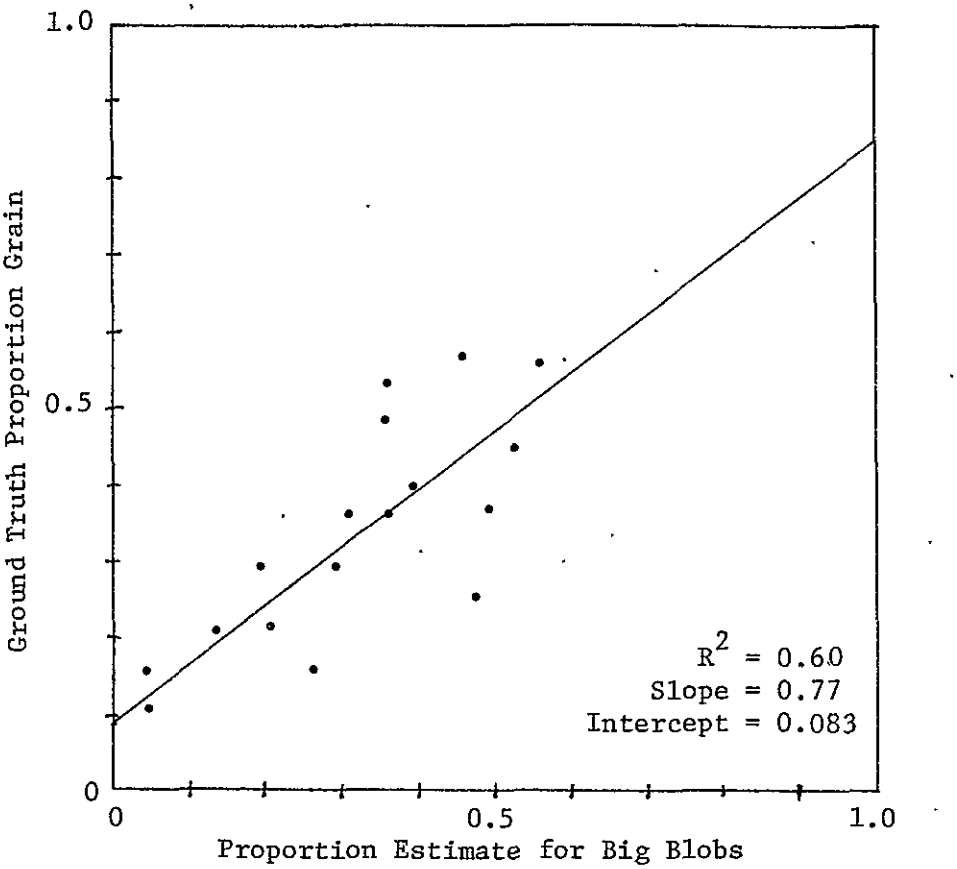


FIGURE C-7. ANALYST RED'S PROPORTION ESTIMATE

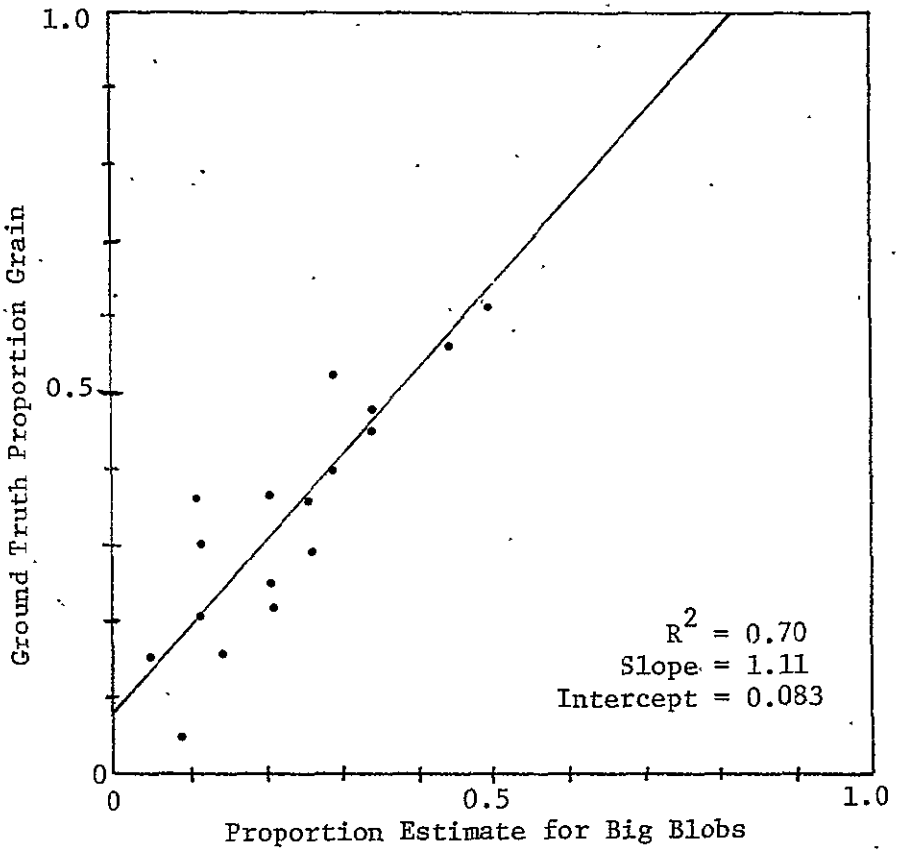


FIGURE C-8. ANALYST BLUE'S PROPORTION ESTIMATE

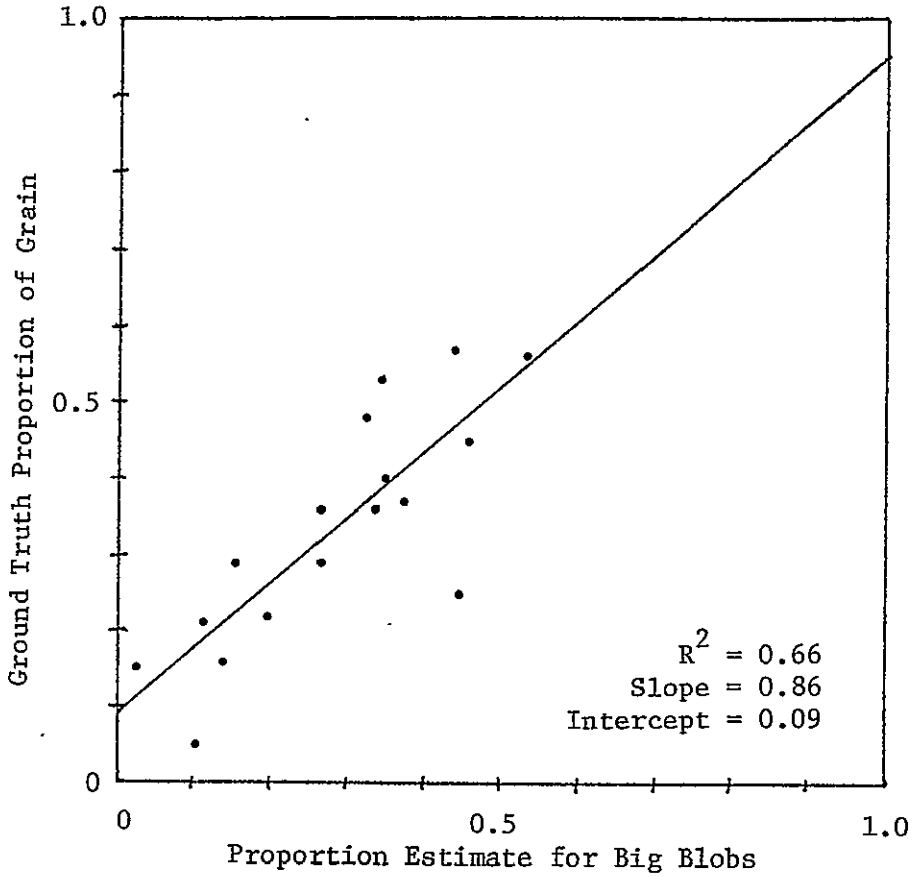


FIGURE C-9. PROPORTION ESTIMATE USING VOTE

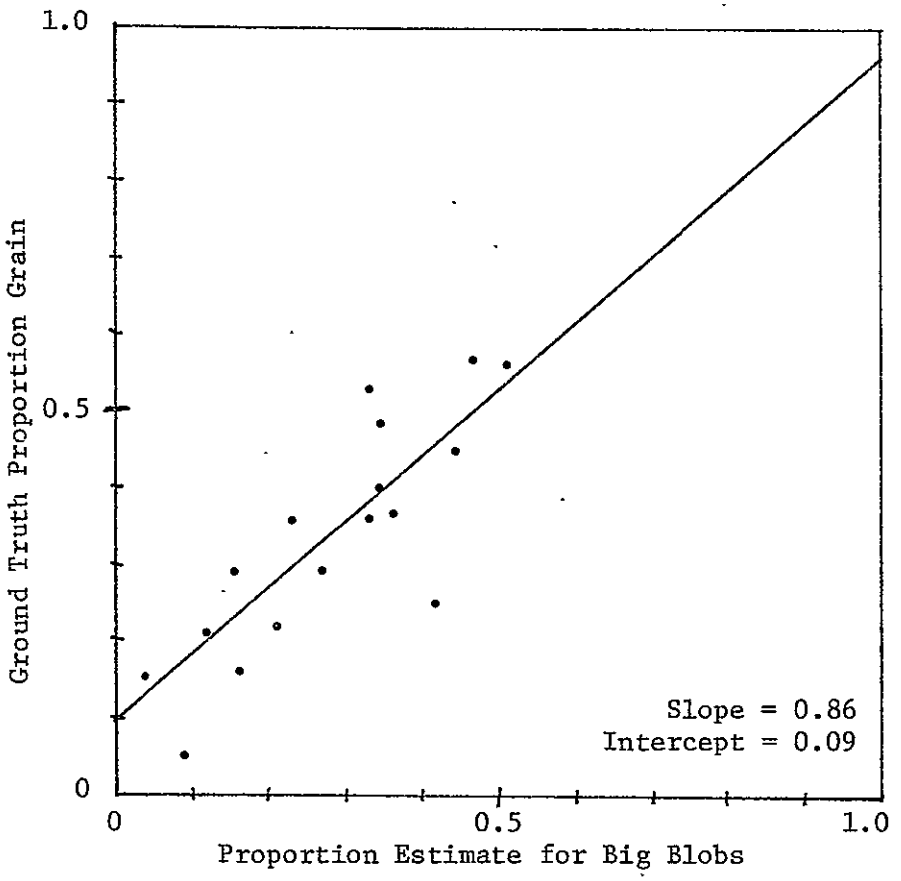


FIGURE C-10. AVERAGE ANALYST PROPORTION

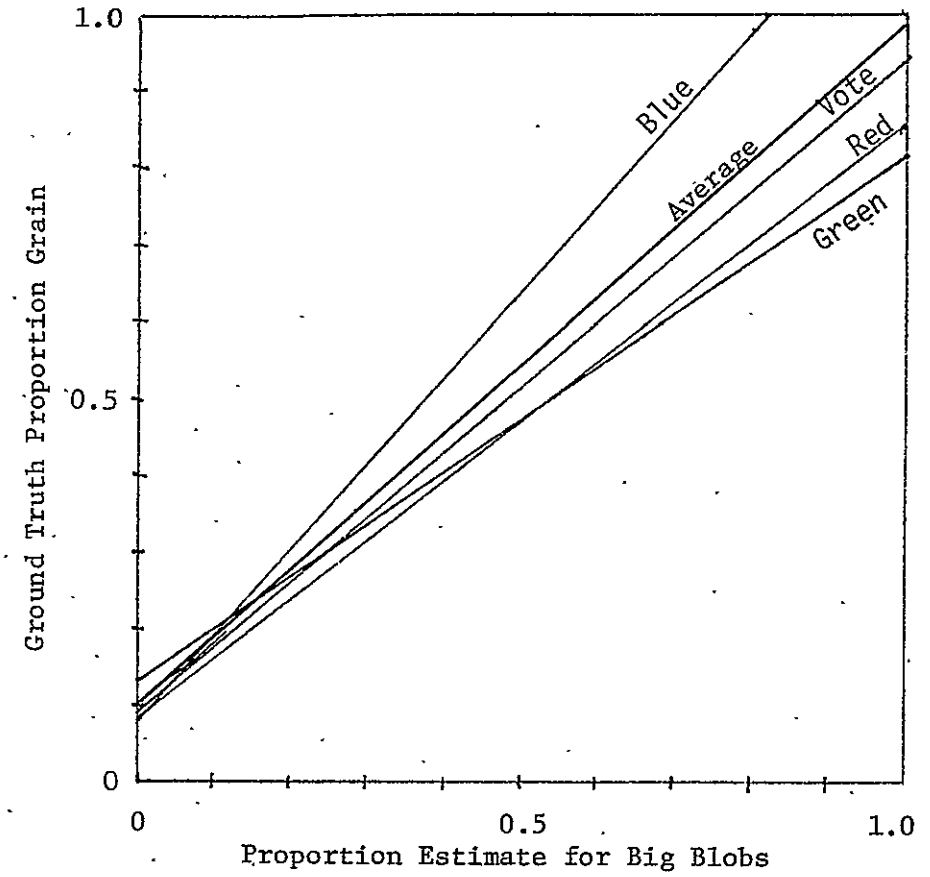


FIGURE C-11. PROPORTION ESTIMATE BY BLOB

C-2

TABLE C-18. REGRESSION COEFFICIENTS FOR ANALYST-BASED PROPORTION ESTIMATES FOR BIG BLOBS

Estimate	Y-Intercept	Slope	Correlation Coefficient
Green	0.132	0.68	0.76
Red	0.084	0.77	0.78
Blue	0.083	1.12	0.84
Average	0.099	0.86	0.79
Vote	0.091	0.86	0.81

Multiple Regression of Ground Truth Grain on Analyst Labels

ANOVA Table

Source	Sum of Squares	df	Mean Square	F	Significance
Regression	.26873	3	.089576	11.384	.0006
Error	.10229	13	.0078688		
Total	.37102	16			

Variable	Coefficients	STD Error	Significance
Constant	.067471	.05115	.2426
Green	.12652	.38938	.7504
Red	.14943	.46177	.7514
Blue	.80616	.34471	.036

that received these labels by more than one analyst. However, the analysts were very consistent in their proportion of flags within the grain purity classes: 0-20%, 20-40%, 40-60%, 60-80%, and 80-100%. When one or more analysts gave a "low confidence" label to a blob, a new label "doubtful" was given to that blob. Table C-19 gives the doubtful labels broken down by grain purity levels and label by vote. The proportion of doubtful labels increased as the grain increased. This result is consistent with the results given earlier that non-grains were labeled more accurately than grain. The ratio of doubtful labels to not doubtful labels of the grain labels decreased as percent grain increased. Thus, the analysts were much less confident of their grain labels given the non-grain targets.

Blobs which received one or more mixed labels were tabulated and are given in Table C-20. The percent of mixed labels increased as percent grain increased from 0-50% and then generally tended to decrease from 60-100%. This is what should be expected; that is, as the labeling targets become less pure, the percentage of mixed labels should increase.

The trend is much less clear in those blobs with grain purity 50-80%. Also, the percentage of mixed labels were much higher for pure grain blobs than for pure non-grain blobs. This seems to indicate that pure grain fields behave in a more confusing fashion than pure non-grain fields. Distributions of grain purity are given in Section C.4 of this Appendix.

TABLE C-19. ANALYST DOUBTFUL LABEL PERFORMANCE

Purity as a Grain	Percent of Doubtful Labels	Analyst Decision: Doubtful		Analyst Decision: Not Doubtful		Fraction Labeled Grain		
		Grain	Non-Grain	Grain	Non-Grain	Doubtful	Not Doubtful	Ratio
0-20%	1.60%	13	36	24	3037	.27	.02	13.5
21-40%	3.60%	6	10	57	388	.37	.13	2.8
41-60%	9.01%	17	14	117	227	.54	.34	1.6
61-80%	10.00%	34	13	274	196	.72	.58	1.2
81-100%	16.54%	185	47	976	427	.79	.70	1.1

TABLE C-20. RELATIONSHIP BETWEEN PERCENT GRAIN AND "MIXED" LABELS

Percent Grain	Percent With One or More "Mixed Target" Flag
0-10%	1.1
10-20%	3.8
20-30%	6.8
30-40%	11.3
40-50%	20.6
50-60%	16.2
60-70%	17.6
70-80%	13.0
80-90%	13.2
90-100%	7.4

C.4 TEST AND EVALUATION OF PROCEDURE M

C.4.1 ANALYST COMMENTS ON BLOB LABELING

Each analyst was asked to complete a segment comment form after labeling each segment and a final comment form after labeling all segments. The comment forms are given in Section C.2 of this Appendix. In addition, the analysts generated a report [32] on analyst labeling of blobs.

General comments on blob labeling: From [32] we obtain a list of strong points, problem areas, and recommendations.

Strong Points

- Blobs are easier to label than the dots used in Procedure 1 (P-1). A blob represents a field center and does not contain border or edge dots as does a P-1 dot.
- BLOBS represent field centers rather well.
- If ERIM reduces the number of blobs to approximately 100, as currently planned, labeling of blobs should be as efficient or perhaps more efficient than the labeling of dots in Procedure 1.

Problem Areas

- The blobbing technique, as currently implemented, produces too many blobs for labeling (400-600).
- Small or stripped fields do not blob or cluster very well.
- Acquisitions selected for blobbing by ERIM for the labeling test were not always optimum.

- Blobs were frequently disjointed which resulted in labeling difficulties.
- Small blobs containing only one to three pixels are difficult to label.

Recommendations

- Acquisition selection for blobbing should be based on multi-temporal/spectral information as well as spatial information (spatial data was used as the primary selection criteria for the test).
- Research should be conducted into the small fields problem.
- Reduce the use of single-pixel blobs whenever possible.
- Modify the line printer blob map and Production Film Converter (PFC) blob overlay. The current format of these two products is not conductive to efficient analyst labeling.

The 1st comment under Strong Points and the 1st under Problem Areas are comments about the exhaustive labeling of all big blobs used in this experiment compared to the labeling of a sample (of size 100) of big blobs in an operational mode of Procedure M. The exhaustive labeling allows us to compute the variances of many different sampling schemes.

The 2nd comment under Problem Areas should be viewed as a comment about big blobs not formed from small or stripped fields. We have found many blobs representing small and stripped fields which would have been good labeling targets but were not included because they did not contain any interior pixels. We are currently conducting research on defining the attributes of those little blobs which would make good labeling targets. The last comment under Problem Areas is almost the opposite of the 2nd, i.e., blobs with only one to three interior pixels are difficult to label.

The 4th comment under Problem Areas has motivated the development of a new blobbing procedure called SUPERB which should eliminate most of the disconnected blobs.

The 3rd comment under Problem Areas would be at least partially addressed by having an analyst involved in the selection of acquisitions for blobbing.

The 1st comment under Recommendations would improve the relationship between fields and blobs. Some of the acquisitions used in blobbing were misregistered by up to three pixels which possibly caused some blobs to be disconnected and caused some to be smaller than what otherwise would be the case.

The 2nd and 3rd comments under Recommendations are very closely related. The more small fields that are sample targets the more blobs with only one interior pixel will be sampled and labeled. On the other hand, bias is introduced by any restriction on sampling. Much more research must be done in this area.

Summarized analyst comments for each segment: Table C-21 summarizes the analysts responses to parts 5b and 6c of the segment comment form Table C-22 summarizes the analysts' responses to parts 6a and 6b. Analysts A and C both flagged Segments 1650 and 1920 as having blobs with impure interiors and as having blobs whose patterns do not match the field patterns.* Table C-23 summarizes the analysts' responses to part 6d. Segment 1650 had many small strip fields while Segment 1920 had poor acquisition history. Table C-24 summarizes the analysts' responses to part 7. Table C-25 summarizes the analysts' responses to part 8: "Did you have to change your procedure?" We note that few changes were made and these were minor. Table C-25 also gives analyst remarks on moisture and elevation.

*The arbitrary association of letter symbols with color symbols for analysts in these tables was established by coin flips.

TABLE C-21. SUMMARY OF ANALYST COMMENTS ON PARTS 5b AND 6c OF THE SEGMENT COMMENT FORM

Segment	(Part 5b) Were Landsat acquisitions deficient? Analyst A		(Part 6c) Was choice of blob acquisitions optimal? Analyst A	
1380	N		N	8205 has poor field definition
1392	Y	need late August, late June	N(?)	inclusion of 8136 might help remove mixed blobs
1457	Y	missing acq. during July	Y	
1461	Y	need late June, early Sept.	Fair	
1467	Y	need late Aug. late June	N	8218 misregistered May, June dates would have helped stream areas
1473	Y	lack acqs. in May/June	Y	
1518	N→	some clouds	Y	
1566	Fair	more July, Aug, Sept.	Y	
1602	Y	need pre-plant, early Sept. acq.	Y	
1612	Y	need late June	Y	few other choices
1619	Y	lacks June, early July	Y	
1636	Y	missing late June	Y	
1650	Y	lacks late June acq.	OK	
1653	N→	late Aug. would have helped	Y	
1656	Y	poor data in late June, early July; poor acq. history	OK	might have used 8218
1825	N		Y	
1835	N		Y	
1920	Y	no June, early July acq.	Y	few other choices

TABLE C-21. SUMMARY OF ANALYST COMMENTS ON PARTS 5b AND 6c OF THE SEGMENT COMMENT FORM (Cont.)

	(Part 5b) Were Landsat acquisitions deficient?		(Part 6c) Was choice of blob acquisitions optimal?	
Segment	Analyst B		Analyst B	
1380	N		N	should use 8169 (vice) 8222 (harv/ripe)
1392	Y	lack of harv, post-harv acq. caused SG/ summer crop confusion	N	should've also used 8136, 8154
1457	Y	need July acq.	Ade-quate	
1461	N		N	should've used 8137, 8190, 8217, 8236
1467	N	need late Aug., early Sept.	N	misregistration; should have used 8155,8199,8217
1473	Y	needed May, late Aug.	Y	
1518	N		N	needed May 15 acq.
1566	B	needed late Aug. early Sept.	N	all 4 acq. should've been used
1602	Y	needed May acq.	Ade-quate	
1612	Y	need Sept. acq.	N	2 acq. not enough; should've used 8137, 8155, 8199, 8236
1619	Y	need June acq.	N	need all four acq.
1636	Y	1977 Fall date would've helped	N	acqs. all too far into growing season
1650	N		Y	
1653	Y	lack ripe, harvest acq.	Y	adequate
1656	N		Y	
1825	Y	need 1 acq. after Sept. 7	N	8133, 8169, 8223, 8250 would've been better
1835	N		Y	
1920	Y	early Sept. acq. needed	Y	adequate

TABLE C-21. SUMMARY OF ANALYST COMMENTS ON PARTS 5b AND 6c OF THE SEGMENT COMMENT FORM (Cont.)

	(Part 5b) Were Landsat acquisitions deficient?		(Part 6c) Was choice of blob acquisitions optimal?	
Segment	Analyst C		Analyst C	
1380	-		Y	
1392	N		N	needed 8154
1457	-		N	only 3 dates; 8174 may have helped
1461	-		N	8236 might have better defined blobs
1467	-		N	needed early date
1473	Y	haze, clouds, cloud shadow	-	
1518	Y	haze, clouds, cloud shadow	N	needed 8206
1566	-		N	needed 4 or more
1602	N		N	should've used 4
1612	N	haze, clouds, need later dates	Basically Yes	needed 8199
1619	Y	clouds, haze on 8198	N	could've used better days
1636	N		N	needed early date
1650	Y	need good pre-planting acq.	N	should've used later harvest acq.
1653	Y	no satisfactory pre-plant or harvest	Y	
1656	-		N	use 1,2 more acq.
1825	-		Y	
1835	N		N	need more spread
1920	-		-	

TABLE C-22. SUMMARY OF ANALYST COMMENTS ON PARTS 6a AND 6b OF THE SEGMENT COMMENT FORM

	(Part 6a) Do blob interiors seem pure?		(Part 6b) Do blob patterns match field pattern?	
Segment	Analyst A		Analyst A	
1380	Y	1 exception	Y	but much smaller than fields
1392	Fair	6 mixed	Fair	patterns on imagery poorly defined
1457	Y	mostly	Y	mostly
1461	Y	3 exceptions	Y	mostly
1467	N	many areas contain more than one field	Fair/Poor	
1473	Y		Y	
1518	Y	but one	Y	
1566	Y		Y	fairly well
1602	Y	except 3	Y	
1612	Y	except 3 in strip area	Y	fairly well except strip area
1619	Y	6 exceptions	Y	
1636	Mostly		Y	
1650	N	many mixed	N	
1653	Mostly	3,4 exceptions	Fair	
1656	Y		Some	patterns poorly defined on imagery
1825	Y	3 exceptions	Y	mostly
1835	Y	mostly	-	difficult to tell
1920	N	many in strip areas	N	field definition poor on imagery

TABLE C-22. SUMMARY OF ANALYST COMMENTS ON PARTS 6a AND 6b OF THE SEGMENT COMMENT FORM (Cont.)

	(Part 6a) Do blob interiors seem pure?		(Part 6b) Do blob patterns match field pattern?	
Segment	Analyst B		Analyst B	
1380	Y	generally	Y	generally
1392	Y	generally	Y	generally
1457	Y	generally	Y	generally
1461	Y	generally	Y	
1467	Y	generally; many exceptions	Y	generally
1473	Y		Y	
1518	Y	generally	Y	generally
1566	Y	generally	Y	generally
1602	Y	some exceptions	Y	generally
1612	Y	generally	Y	generally
1619	Y	generally	Y	generally
1636	Y	in most cases	Y	mostly
1650	Y	generally	Y	generally
1653	Y	some exceptions	Y	generally
1656	Y	generally	Y	generally
1825	Y	generally	Y	generally
1835	Y		Y	
1920	Y	generally	Y	generally

TABLE C-22. SUMMARY OF ANALYST COMMENTS ON PARTS 6a AND 6b OF THE SEGMENT COMMENT FORM (Cont.)

	(Part 6a) Do blob interiors seem pure?		(Part 6b) Do blob patterns match field pattern?	
Segment	Analyst C		Analyst C	
1380	Y	overall	Y	
1392	Y	overall	Y	overall
1457	Y	overall	Y	overall
1461	Y		Y	overall
1467	Y	overall	Y	overall
1473	Y		Y	except a few fields
1518	Y		Y	
1566	Y	overall	Y	as far as could be told
1602	Y		Y	generally
1612	Y	some problem in strip area	Y/N	strip area caused problems
1619	Y		Y	overall
1636	Y		Y	mostly
1650	N		N	
1653	Y		Y	
1656	N →	for spring grains	Y/N	more for non-SG than SG
1825	Y	generally	Y	where patterns were distinguishable
1835				
1920	N		N	somewhat, overall poor

TABLE C-23. SUMMARY OF ANALYST COMMENTS ON PART 6d OF THE SEGMENT COMMENT FORM

Segment	Analyst	Other Comments Re Blob Patterns	Other Comments
1380	A	-	-
	B	None	SG is planted relatively early in harvested by third week in Aug
	C	Whether a blob is developed appears to depend on field width.	Dots still misleading.
1392	A	-	Because of wide differences in p pretation of late SG from summ Appear to have some SG startin Inadequate acquisition to conf
	B	None	None
	C	-	Repeat of symbols; analyst can't greater than 3 pixels as some
1457	A	-	Multiple planting date for SG in of interpretation.
	B	None	-
	C		Dots still confuse; color separa feel comfortable with this one.
1461	A	-	-
	B	Certain areas screened out; wet spots 8155; (see analyst form).	The SG labels include some flax crop in area.
	C	-	Problem with disjoint blobs and
1467	A	-	Little confidence is place on so tations; confusion with possib of lates.
	B	Several areas were screened out (many SG fields).	None.
	C	More mixed blobs than previously encountered; many small.	Border around blobs not constant misleading.
1473	A	Nice blob sizes; easy to work with.	-
	B	Worked like a charm! Blobs pure and umambiguous, helped in fine discrimination.	None.
	C	Blob match to field depends on size; smaller more mixed.	Lack of acquisition increases er tween 26 April and 16 July).

TABLE C-23. SUMMARY OF ANALYST COMMENTS ON PART 6d OF THE SEGMENT COMMENT FORM (Cont.)

Segment	Analyst	Other Comments Re Blob Patterns	Other Comments
1518	A	-	Clouds and haze make interpretation of some signatures shaky.
	B	Couple examples of disjoint blob pixels.	Some SG did not show usual signature trends (see analyst form for discussion).
	C	Some mixed blobs, but less than some other segments.	Some impure and confusing signatures; signature progression hard to follow; some blob ID as SG may be flax.
1566	A	-	-
	B	-	Possibly 2 planting dates some SG might have been called summer crops.
	C	-	Field patterns hard to discern; need more blob acquisitions.
1602	A	-	Size of most strips in segment is as small as will produce a blob center.
	B	-	SG in segment shows classical temporal signatures which facilitated interpretation.
	C	Bigger strip field width fewer mixed blobs.	Winter crops not separable; easy segment to interpret; good signatures; piece of cake.
1612	A	-	Since wheat was in strip area, fewer blobs labeled SG than expected.
	B	-	-
	C	-	Some difficulty determining strip field limits; blob colors a problem.
1619	A	Mostly larger blobs; easier to work with.	-
	B	-	Symbols look too much alike; hard to follow blob boundary.
	C	-	Good segment; high confidence level.
1636	A	-	Two different SG planting dates makes interpretation questionable.
	B	Should eliminate single pixel + small blobs.	Need more PFC scale blob maps.
	C	More mixed blobs than previously.	Small blobs cause some confusion.

TABLE C-23. SUMMARY OF ANALYST COMMENTS ON PART 6d OF THE SEGMENT COMMENT FORM (Cont.)

Segment	Analyst	Other Comments Re Blob Patterns	Other Comments
1650	A	Only areas that did well were large pastures.	Small strip fields (satellite resolution) makes interpretation difficult.
	B	Blobs are small, some not pure.	Natrow strip fields messed things up (see analyst's sheet).
	C	-	Color of blob map and use of "." on computer map confusing.
1653	A	Blobbing was less in wheat areas due to strips.	Blobbing didn't work as well because of narrow strip fields.
	B	Large areas of stripped fields rejected by algorithm.	Segment generally displays inhomogenous signatures; muttled appearance.
	C	-	Use "/" instead of "."; blob overaly colors bad; eliminate small blobs; use grid to check for overlooked blob.
1656	A	-	Small fields without blobs probably results in low SG estimate.
	B	-	-
	C	-	Segment difficult due to small size, lack of distinct field patterns of strips.
1825	A	Many areas of little blobs.	Some "SG" signatures follow a faster progression.
	B	-	Two apparent SG planting dates caused confusion.
	C	-	Confidence Level: Ave. Small fields, no clear field pattern.
1835	A	-	Difficult to interpret because of irregular shapes/ poor field definition.
	B	-	-
	C	-	Grain and non-grain signatures similar during summer.
1920	A	-	Small strip fields; poor blobs; segment is terrible to interpret.
	B	-	-
	C	Considering the segment, it is doubtful any change would have helped.	Poor segment; very low confidence rating; key acquisitions missing; high confusion factor.

TABLE C-24. SUMMARY OF ANALYST COMMENTS ON PART 7 OF THE SEGMENT COMMENT FORM

Segment	Analyst	Percent Agriculture	Percent Small Grains	Major Crops	Field Size	Topography
1380	A	80-90	Small	Corn and Soybeans	120 acres	Flat to gently rolling hills, some stream drainage.
	B	98	Relatively little.	Corn and Soybeans		Central Nebrasks Loess hills, CNL plains, & L, Till, & Sand, Prairie.
	C	100			10-350 pixels	Relatively flat with kettles, except steep banks along river.
1392	A	50-60			80-160 (varied)	3 small lakes, large strip of rangeland.
	B	80		Other 20% rangeland.	Small to medium	Dark wet soils.
	C	60-70			40-140 Ave. 60	Ag areas divided by NW-SE morraine; no definite drainage pattern; slope is south.
1457	A	60-70			100 acres	Flat, many pot-hole lakes, several streams and drainage system.
	B	85				Gently rolling, glaciated plain; some irregular top; numerous wet depressions and ponds.
	C	80-85	20-25%	2000 ac. winter, rest spring/ summer crops	5-10 to 250-300 acres	Slopes down toward NE; morraine belt with numerou kettles.
1461	A	75-80		SG		Flat; several fair-sized lakes; intermediate streams.
	B	90	High	SG		Drift prairie; mostly glacial deposits.
	C	90-95			100-150 acres range; 60-80 acres average	Flat, but glaciated; unestablished drainage.
1467	A	90-95			80-160 acres	Flat, gently rolling hills, several streams.
	B	90			Large	Wet with relatively mild slopes.
	C	100			10-15 to 80-100 ave. 60 acres	Flat with glaciation indicators; drainage SE.
1473	A	90-95		Wide Variety	80-160 acres	Flat.
	B	98	High %		Large to medium	Smooth, part of old lake bed.
	C	100			20 to 200-250 100 average	Flat, part of old lake bed.

TABLE C-24. SUMMARY OF ANALYST COMMENTS ON PART 7 OF THE SEGMENT COMMENT FORM (Cont.)

Segment	Analyst	Percent Agriculture	Percent Small Grains	Major Crops	Field Size	Topography
1518	A	50-60		Large Forest Areas	80-160 acres	
	B	-				Smooth lake basin, poor drainage, small water basin.
	C	60		30% pasture or scrub	5-250; 60-80 ave.	Level, no distinct drainage.
1566	A	70-80				Flat to gently rolling hills; some lakes, marshy areas; minor drainage.
	B	90	High			Drift prairie; almost entirely glacial deposits.
	C	75-85			Range: 10-140 ac. ave. = 60 ac.	Glaciated; gently rolling; numerous lakes.
1602	A	40-50			1/8-1/16 section	Flat terrain; several large lakes.
	B	75-80				Irregular: gently rolling glaciated plains, _____, morraine.
	C	75-80			5-250 ac. strip 1-8 pixels.	Glaciated; moderately rolling, moraines, numerous kettles.
1612	A	20-30			40 acres	Fairly flat; several large lakes.
	B	40				Drift prairie: glacial deposits, knolls, depressions, some streams.
	C	45-55			10 70-80 acres.	Flat; low hills; numerous lakes.
1619	A	90			1/4 section	Flat.
	B					Mostly smooth, nearly level lake basin.
	C	100			Range: 10-200 ac. ave. = 100-160	Gentle to flat.
1636	A	75-80			1/8-1/4 section	Fairly flat; few low areas that hold water.
	B	80			Medium	Drift prairie: glacial deposits, undulating plain.
	C	100			20 300 ac. ave. = 100-150 ac.	Flat with scatter glacial deposits.
1650	A	30-40				Gently rolling hills, few small lakes.
	B	70				Glacial deposits: rolling plain, some badlands top.
	C	75-80			5-125 acres	Rolling terrain; some glaciation.

TABLE C-24. SUMMARY OF ANALYST COMMENTS ON PART 7 OF THE SEGMENT COMMENT FORM (Cont.)

Segment	Analyst	Percent Agriculture	Percent Small Grains	Major Crops	Field Size	Topography
1653	A	20-30			20-30 acres	Fairly flat.
	B	40				Gently rolling; glaciated landscape.
	C	15-20			5-200 acres	Undulating terrain; intermittent creek, lake.
1656	A	20-30		Small Amount SG	30-40 acres	Flat area with agriculture; rest is fairly hilly.
	B	20			Small	Extremely rugged; mostly natural vegetation and rangeland.
	C	85			3-60 acres	Medium to heavy direction near Heart River; gentler in east.
1825	A	60-70			100 acres	Flat with trees, natural vegetation along a river.
	B	80				Generally smooth, nearly level lake basin.
	C	55			10-120 acres	Level with gentle slope west; glaciated.
1835	A	50			Small	Swampy area; hills; several lakes.
	B	80			Small	Wet area; several bodies of water.
	C				10-70 acres	Heavily glaciated; numerous marshes and lakes.
1920	A	20-30			30-40 acres	Agricultural area fairly flat; rest is rather hilly.
	B	50			Medium	Gently rolling glaciated plain with some kames and morraine.
	C	30-40	Low Spring Grain Segment		10-20 acres	Dissected in west; gently rolling to flat in east.

TABLE C-25. SUMMARY OF ANALYST COMMENTS ON PART 8 OF THE SEGMENT COMMENT FORM

Segment	Analyst	Did You Have to Change Your Procedure?	Moisture?	Elevation
1380	A	Yes. Identified small grains blobs first, then non-small grains	Adequate	
	B	Yes. Because of significant distortions, Product 3 used extensively		
	C	-		
1392	A	No		
	B	No		
	C	No		
1457	A	Yes. Slightly because of many discontinuous blobs	Adequate	
	B	No		
	C	No		2000-2300'
1461	A	No	More than adequate	
	B	No	Moist to wet throughout	
	C	No		
1467	A	No		
	B	No		
	C	No		Avg. ∿1600'

TABLE C-25. SUMMARY OF ANALYST COMMENTS ON PART 8 OF THE SEGMENT COMMENT FC

Segment	Analyst	Did You Have to Change Your Procedure?	Moisture?	
1473	A	No	Adequate	
	B	No		
	C	-		Δ=5'
1518	A	-		
	B	No		
	C	No. Just a little more cautious		
1566	A	No		
	B	No	Area is generally wet	
	C	-	Drainage is deranged	00'
1602	A	No	Adequate	
	B	No		
	C	-	Climate semi-ar	
1612	A	No		
	B	No		
	C	Yes. Labeled non-spring grains, then spring grains	Deranged drainage	

of 4)

TABLE C-25. SUMMARY OF ANALYST COMMENTS ON PART 8 OF THE SEGMENT COMMENT FORM (Cont.)

Segment	Analyst	Did You Have to Change Your Procedure?	Moisture?	Elevation
1619	A	Yes. Colored obvious non-grain, then obvious grain	No apparent water problem	
	B	No	Poor drainage	
	C	No	Drainage dendritic, but intermittent	900'
1636	A	No		
	B	No		
	C	No	Deranged drainage	1465'
1650	A	No	Adequate	
	B	No	Relatively dry	
	C	No	Dendritic drainage	2835'
1653	A	No	Adequate	
	B	No		
	C	Yes. Product 3's removed (confusing); blob map used only w/ Product 1	No rain noticed	1950'
1656	A	-		
	B	No		
	C	No		1750-2350'

TABLE C-25. SUMMARY OF ANALYST COMMENTS ON PART 8 OF THE SEGMENT COMMENT FORM (Cont.)

Segment	Analyst	Did You Have to Change Your Procedure?	Moisture?	Elevation
1825	A	No	No apparent problem	
	B	No	Wet in Spring	800-1000'
	C	-	Adequate	1000-1310'
1835	A	No	Adequate	
	B	No	Wet area	
	C	-		1325'
1920	A	No		
	B	No		
	C	-		2100-2450'

C.4.2 PARAMETERIZED PERFORMANCE OF PROCEDURE M

The variance and bias of Procedure M are viewed as functions of the parameters which control sample size, type of stratification, number of strata, and source of labels. Methods of estimating the bias and variance of Procedure M's segment spring small grain (wheat/barley) proportion estimate follow.

The bias from sampling only big blobs is the difference in the ground truth proportions of grain within the big blobs and of grain in the complete segment.

Given a fixed segment, source of labels and stratification denote the label given the j^{th} blob of stratum s as C_{js}. The bias due to the source of labels is

$$b(\hat{P}) = \sum_{s=1}^{BCLUST} \frac{N_{\cdot s}}{N_{\cdot\cdot}} \left(\sum_{j=1}^{M_s} \frac{N_{js}}{N_{\cdot s}} C_{js} - P_s \right)$$

where

C_{js} = 1 if label is grain, C_{js} = 0 otherwise,

N_{js} = number of pixels in blob j of Stratum s,

$N_{\cdot s}$ = number of big blob pixels in Stratum s,

$N_{\cdot\cdot}$ = number of big blob pixels in segment,

P_s = ground truth proportion of grain in big blobs,

and given a sample $S = S_1 \cup S_2 \cup S_3 \cup \cdots \cup S_{BCLUST}$ chosen using the Midzuno sampling technique, $\hat{P}$ is the sample proportion, namely

$$\hat{P} = \sum_{s=1}^{BCLUST} \frac{N_{.s}}{N_{..}} \left(\frac{\sum_{j \epsilon S_s} N_{js} C_{js}}{\sum_{j \epsilon S_s} N_{js}} \right) .$$

If n_s is the s stratum sample size, and M_s is the number of blobs in the s stratum, then there are $\binom{M_s}{n_s}$ possible samples from Stratum s of size n_s which could be obtained. We denote these by

$$S_{si},\ i=1,2,3,\ldots,\binom{M_s}{n_s} .$$

The mean square error of

$$\hat{P} = \sum_{s=1}^{BCLUST} \frac{N_{.s}}{N_{..}} \hat{P}_s$$

is

$$(*) \quad MSE(\hat{P}) = \sum_{s=1}^{BCLUST} \left(\frac{N_{.s}}{N_{..}} \right)^2 \sum_{i=1}^{\binom{M_s}{n_s}} \left(\frac{\sum_{j \epsilon S_{is}} N_{js} C_{js}}{\sum_{j \epsilon S_{is}} N_{js}} - P_s \right)^2 \binom{M_s - 1}{n_s - 1}^{-1} \left(\sum_{j \epsilon S_{is}} \frac{N_{js}}{N_{.s}} \right)$$

where

$$\hat{P}_s = \frac{\sum_{j \epsilon S_{is}} N_{js} C_{js}}{\sum_{j \epsilon S_{is}} N_{is}} .$$

The variance of $\hat{P}$ is

$$V(\hat{P}) = MSE(\hat{P}) - b^2(\hat{P}).$$

A proof of (*) can be found in [ ].

Examination of the bias and variance of Procedure M as a function of type of strata, number of strata, sample size, and source of labels follows.

Denote the bias, mean square error and variance of Procedure M proportion estimate for Segment j as

$$b_j(\alpha,N,n,\beta),\ MSE_j(\alpha,N,n,\beta),\ \text{and}\ V_j(\alpha,N,n,\beta)$$

where

α = static trajectory strata, BCLUSTER, or tolerance blocks,

N = 40, 50, or 60,

n = 80, 100, or 120,

β = analyst green, red, blue, vote, average, or ground truth.

The sampling reduction of variance factor is

$$\frac{V_j(\alpha,N,n,\beta)}{V_j(\cdot,1,n,\beta)}$$

The effect of sampling strategy will be studied by examining

$$\overline{V}_{\cdot}(\alpha,N,n,\text{ground truth}) = \frac{1}{17}\sum_{j=1}^{17} V_j(\alpha,N,n,\text{ground truth})$$

for different combinations of α, N, and n given above.

The effect of the source of labels will be studied by examining

$$\overline{V}_{\cdot}(\text{BCLUSTERs},40,100,\beta) - \overline{V}_{\cdot}(\text{BCLUSTERs},40,100,\text{ground truth})$$

for

β = analyst green,
= analyst red,
= analyst blue,
= vote, and
= average label.

C.4.3 EVALUATION OF THE BCLUSTER, BLOB, AND MACHINE LABELER COMPONENTS

Evaluation of the BCLUSTER: Because of a delay in obtaining ground truth for the Minnesota segments, the evaluation of the BCLUSTER component is performed only on the 13 North Dakota segments. A measure of the value of a stratification is the ratio of the stratified variance to the unstratified variance. This measure is called the reduction of variance (RV) factor. Table C-26 gives the RVs for sampling with replacement. The average RVs for 20, 40, and 60 BCLUSTERs are .637, .544, and .483, which are close to what was obtained in last year's experiment. In the case where the sampling is with replacement we have

$$0 \leq RV \leq 1.$$

This is not always the case if the sampling is without replacement. The stratum finite correction factors have less effect on the stratum variance than on the unstratified variance. The methods in which true RVs will be computed are outlined in the last section for the case where sampling is without replacement (used by Procedure M).

Table C-27 gives the number of BCLUSTERs broken down by their grain purity for 20, 40, and 60 BCLUSTERs. The number of blobs, number of pixels, percentage of BCLUSTERs, percentage of blobs, and percentage of pixels are also given in Table C-27.

TABLE C-26. REDUCTION OF VARIANCE (NORTH DAKOTA SEGMENTS)

	Number of BCLUSTERs		
Segment	20	40	60
1392	0.709	0.479	0.422
1457	0.562	0.516	0.441
1461	0.715	0.504	0.482
1467	0.723	0.683	0.621
1473	0.441	0.335	0.315
1602	0.463	0.434	0.371
1612	0.701	0.654	0.627
1619	0.490	0.342	0.292
1636	0.705	0.616	0.469
1650	0.728	0.634	0.616
1653	0.657	0.613	0.500
1656	0.731	0.652	0.541
1920	0.654	0.608	0.577
$\overline{R}$	0.637	0.544	0.483

TABLE C-27. NUMBER OF BCLUSTERs BY GRAIN PURITY

20 BCLUSTERs

Percent Grain	Number of BCLUSTERs	Number of Blobs	Number of Pixels	Percent of BCLUSTERs	Percent of Blobs	Percent of Pixels
0-10%	127	1758	93203	41.64	36.18	38.23
10-20%	16	544	30484	5.25	11.20	12.50
20-30%	16	301	11241	5.25	6.19	4.61
30-40%	9	289	15771	2.95	5.95	6.47
40-50%	7	68	2328	2.30	1.40	0.95
50-60%	6	39	1653	1.97	0.80	0.68
60-70%	12	639	32262	3.93	13.15	13.23
70-80%	14	412	18157	4.59	8.48	7.45
80-90%	12	431	22640	3.93	8.87	9.29
90-100%	86	378	16043	28.20	7.78	6.58

40 BCLUSTERs

0-10%	261	1817	91250	49.81	37.04	37.43
10-20%	23	518	31202	4.39	10.56	12.80
20-30%	24	503	25105	4.58	10.25	10.30
30-40%	14	208	9092	2.67	4.24	3.73
40-50%	16	122	4627	3.05	2.49	1.90
50-60%	13	117	7266	2.48	2.38	2.98
60-70%	13	173	7353	2.48	3.53	3.02
70-80%	19	263	12403	3.63	5.36	5.09
80-90%	17	288	13792	3.24	5.87	5.66
90-100%	124	897	41692	23.66	18.28	17.10

60 BCLUSTERs

0-10%	386	1987	101122	49.61	40.35	41.49
10-20%	33	545	29748	4.24	11.07	12.20
20-30%	26	324	15611	3.34	6.58	6.40
30-40%	11	84	2968	1.41	1.71	1.22
40-50%	23	186	9496	2.96	3.78	3.90
50-60%	16	188	7491	2.06	3.82	3.07
60-70%	18	111	4992	2.31	2.25	2.05
70-80%	21	202	10242	2.70	4.10	4.20
80-90%	25	381	18054	3.21	7.74	7.41
90-100%	219	917	44031	28.15	18.62	18.06

Figure C-12 gives the graph of the percentages of BCLUSTERs within each grain percentage class for 20 and 40 BCLUSTERs. The graph of 60 BCLUSTERs is very close to that of 40 BCLUSTERs. The differences between the number of BCLUSTERs in grain percentage classes is not significant.

However, when the number of big blobs and number of pixels within those BCLUSTERs are graphed (Figures C-13 and C-14) we see that the stratification obtained with 40 and 60 BCLUSTERs is much better than that obtained with 20 BCLUSTERs. The difference between the 40 BCLUSTER and 60 BCLUSTER stratification is small and perhaps the stratum finite correction factor would reverse this ordering.

Evaluation of the blob component: The purpose of the blob component is to cluster pixels into field-like forms which are relatively pure grain or non-grain. These blobs with interior pixels are sampled and labeled by analyst. The labeling results are given in Section ___ of this appendix. The analyst subjective evaluation of the blob component can be found in this section of the appendix. It was shown that those blobs with grain percentage between 25% and 75% were labeled wrong more often than those outside of this interval. Figure C-15 gives the percentage of blobs at each of the grain percentage levels 0-10%, 10-20%, 20-30%, 30-40%, 40-50%, 50-60%, 60-70%, 70-80%, 80-90%, and 90-100%. We note the vast majority of the blobs are less than 25% grain or greater than 75% grain. If the blob grain percentage is computed using only the interior of the blobs the results are even more striking. Figure C-16 gives the percentage of pixels within the blobs in the same grain levels given above. The two graphs are extremely similar because the grain percentage and the blob sizes are nearly independent in these 17 segments.

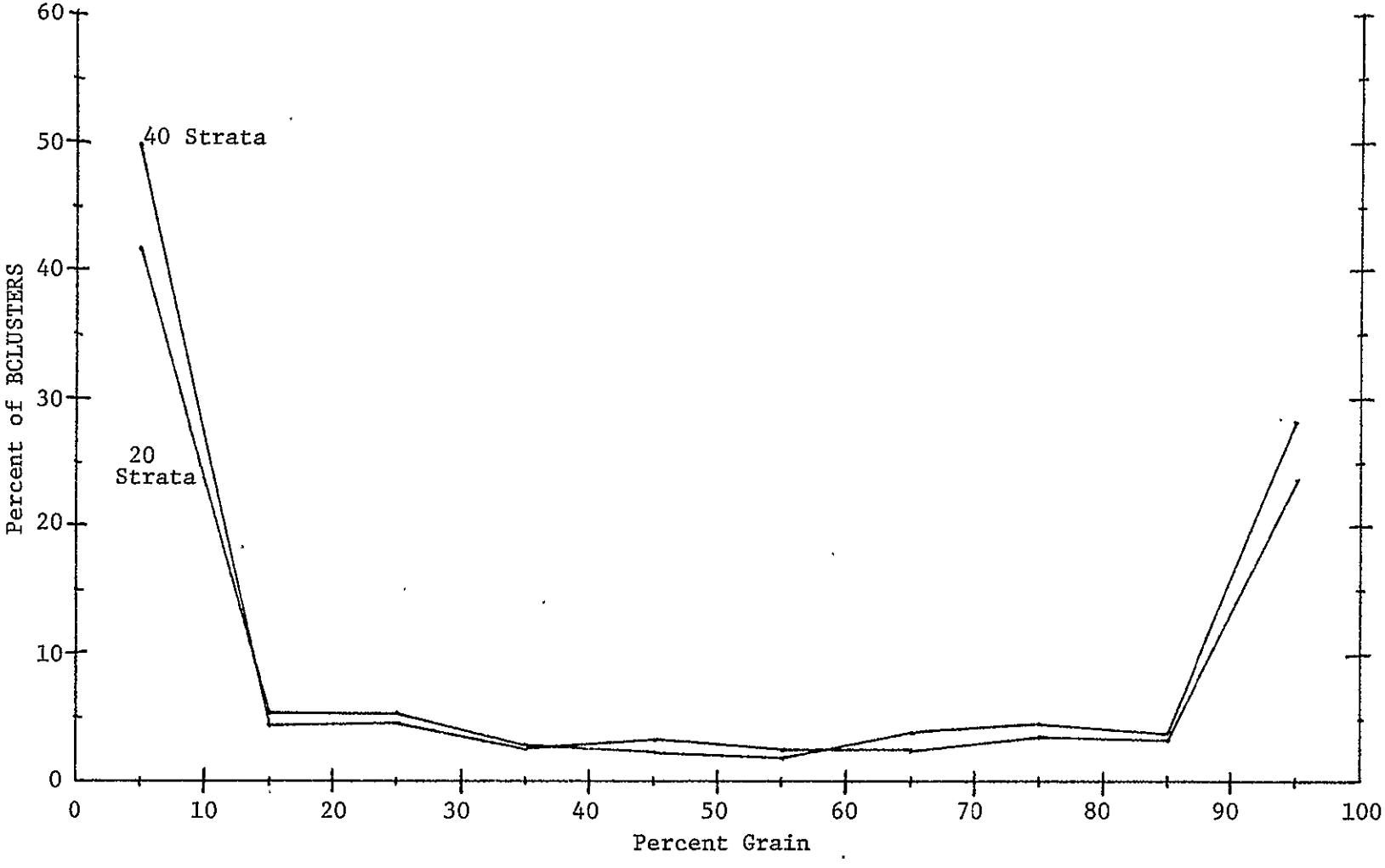


FIGURE C-12. PERCENTAGE OF BCLUSTERS VS. PERCENT GRAIN

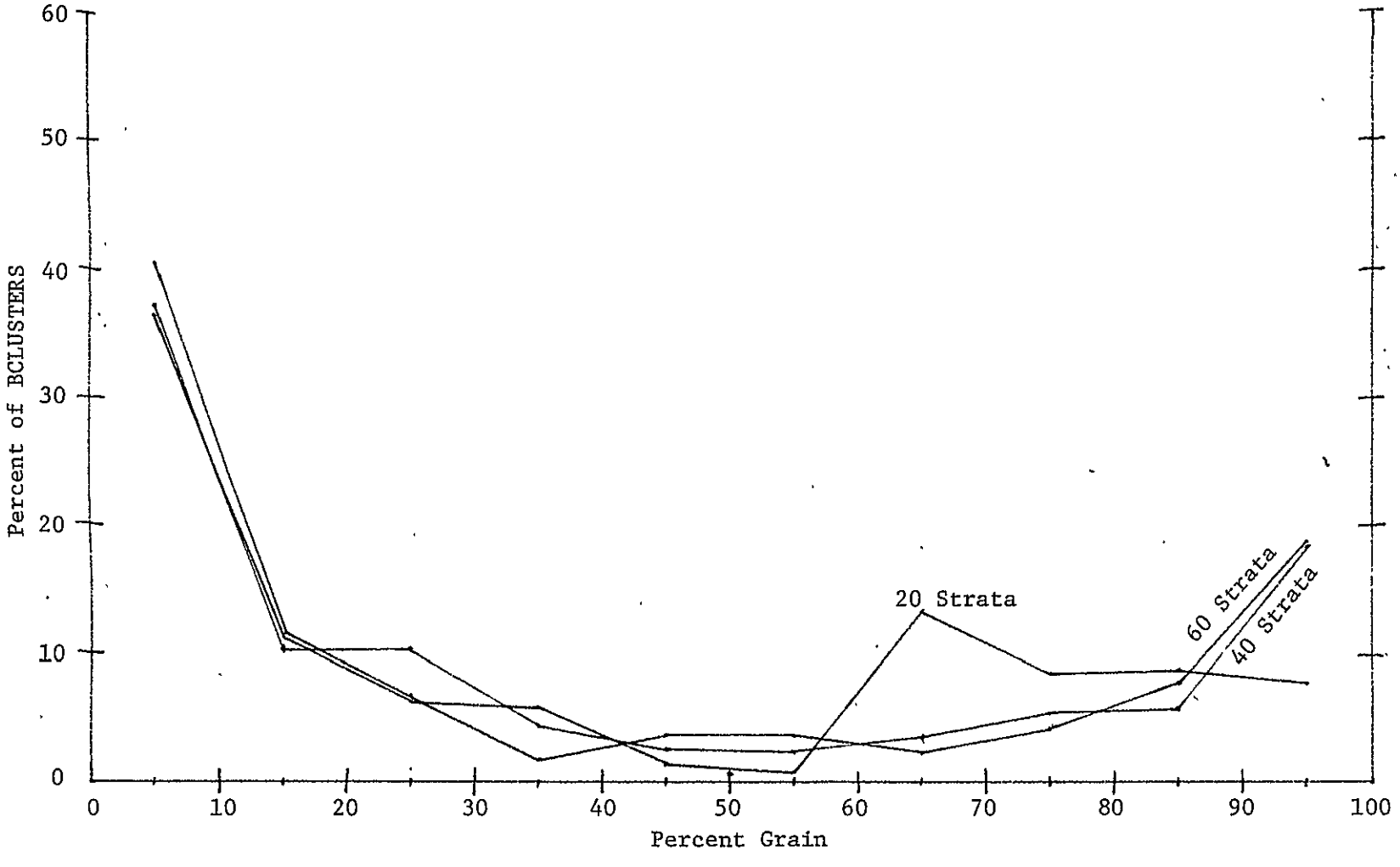


FIGURE C-13. DISTRIBUTION OF THE NUMBER OF BLOBS WITHIN BCLUSTERS VS. GRAIN PERCENTAGE

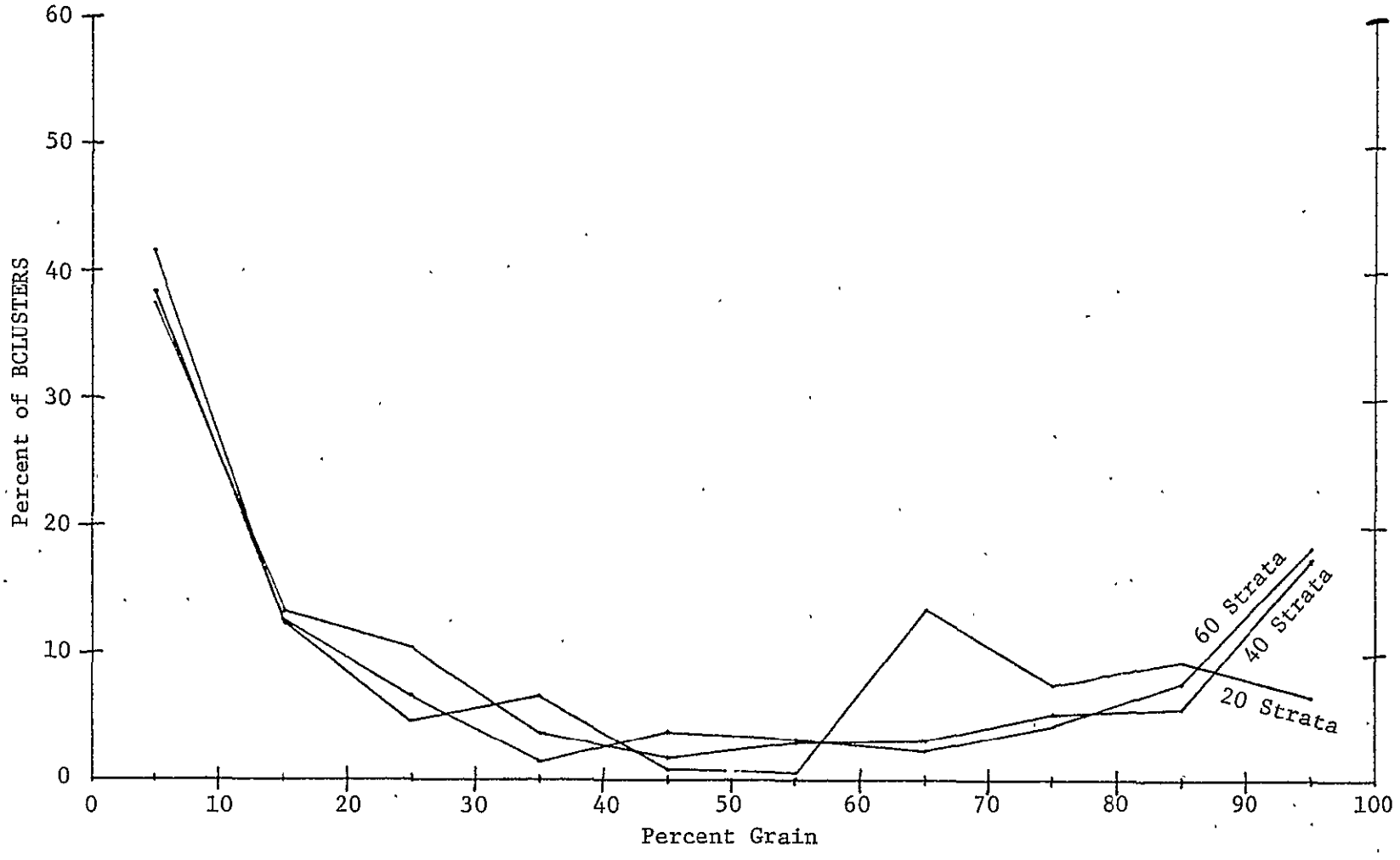


FIGURE C-14. DISTRIBUTION OF THE NUMBER OF PIXELS WITHIN BCLUSTER GRAIN PERCENTAGE LEVELS

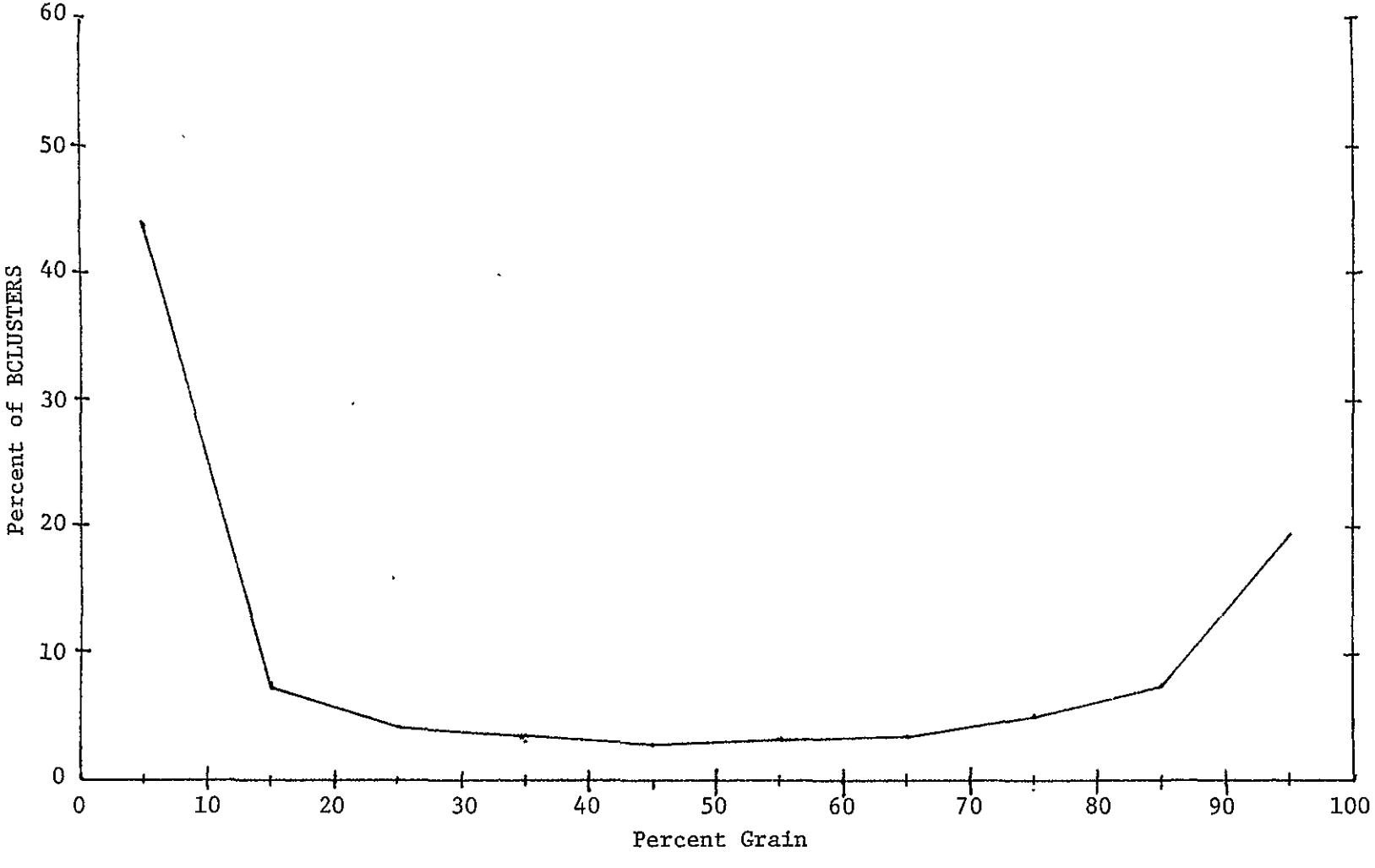


FIGURE C-15. PERCENT OF BIG BLOBS VS. PERCENT GRAIN

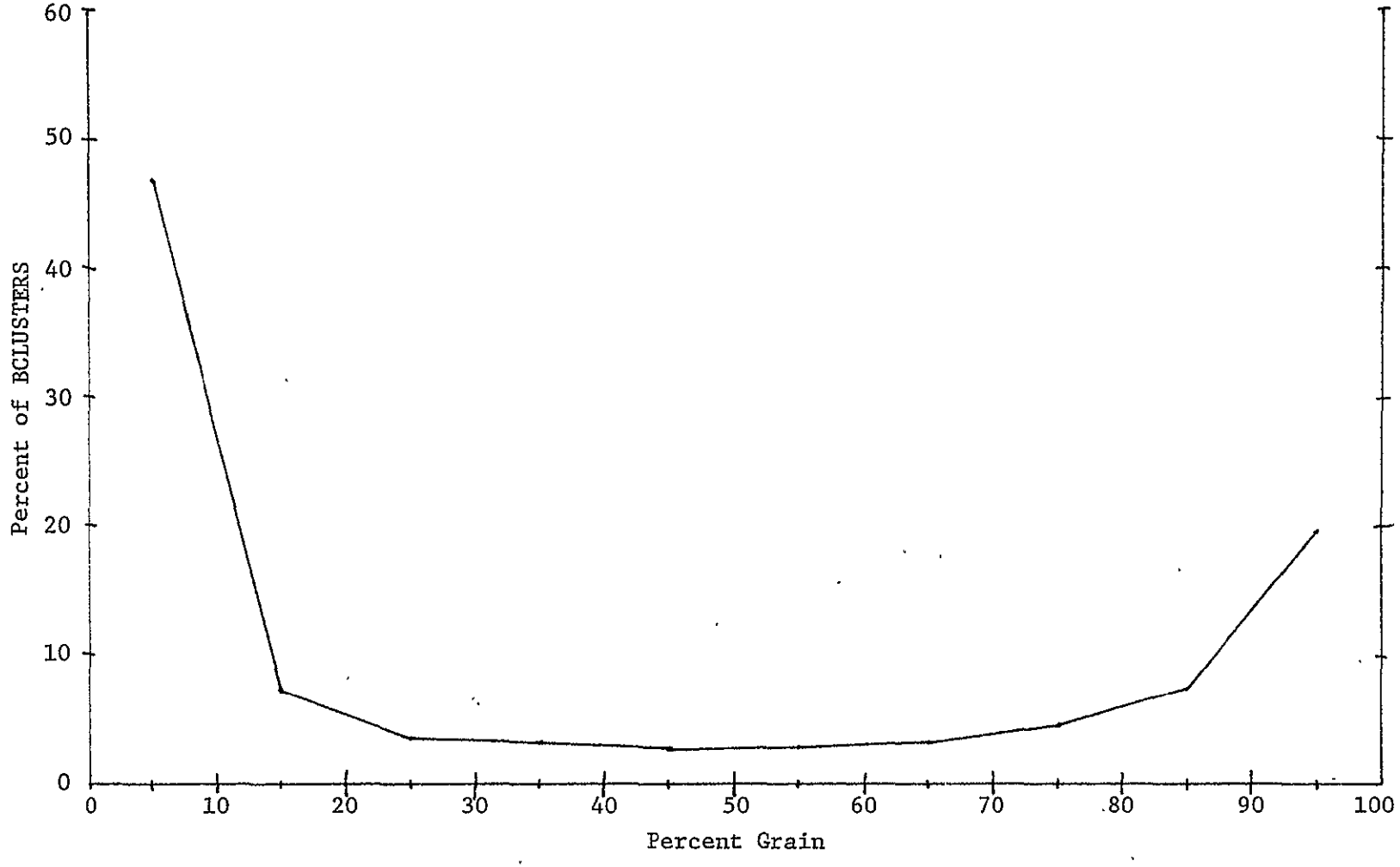


FIGURE C-16. DISTRIBUTION OF PIXELS WITHIN BLOBS VS. PERCENT GRAIN

Table C-28 gives for each segment, the number of blobs, the number of big blobs, the percent of segment covered by big blobs, the blob purity, the reduction of variance factors, and the information theoretic factor discussed in Appendix K and Section 6.

Evaluation of the machine labeler: The machine labeling component uses Brightness/Greenness values to give wheat or barley labels to all of the targets which are labeled grain and have at least three acquisitions falling in the time interval between plant emergence and harvesting.

The performance of the labeler will be evaluated both as a classifier and as a barley proportion estimate. The percentage of ground truth grain targets correctly classified will be computed for each segment. The bias, variance, and root mean squared error (RMSE) of the barley proportion estimate will be computed for each segment.

The relationship between analyst errors and the machine labeler errors will be studied by evaluating the machine labeler's performance on grain targets missed by the analysts and non-grain targets labeled grain by the analyst.

TABLE C-28. SEGMENT QUASI-FIELD STATISTICS

Segment	Blobs	No. Big Blobs	Percent of Segment Covered	Blob Purity	Blob RV	Blob ITF
1380	1633	403	65.64	93.07	.40	.85
1392	688	380	87.6	92.5	.20	.76
1457	900	401	84.1	85.5	.15	.81
1461	1506	465	71.2	96.9	.08	.92
1467	554	400	91.4	89.2	.32	.67
1473	1353	387	72.6	97.6	.06	.95
1518	819	420	87.38	88.22	.27	.93
1566	1264	461	68.39	86.97	.11	.83
1602	1298	399	78.5	97.0	.69	.90
1612	577	329	90.8	94.0	.22	.78
1619	916	379	79.5	96.6	.09	.89
1636	1660	427	80.1	93.1	.20	.75
1650	862	397	81.3	85.7	.44	.58
1653	613	339	89.0	96.6	.19	.85
1656	631	322	87.5	95.6	.19	.83
1825	1114	455	80.79	89.17	.34	.82
1920	436	299	99.2	93.8	.20	.80
		6663				
Ave.		392	79.47	93.03	.244	.8182

APPENDIX D

SUPERB -- AN IMPROVED SPECTRAL-SPATIAL CLUSTERING ALGORITHM

D.1 INTRODUCTION

SUPERB represents a next generation of spectral-spatial clustering beyond ERIM's algorithm BLOB. The new algorithm takes its name from "supervised BLOB". As with BLOB, the purpose is to divide a scene into patches (which we call quasi-fields or blobs) of adjacent or nearly adjacent pixels that are relatively homogeneous spectrally, and to a reasonable extent, represent the fields present within the scene.

Spatial-spectral clustering has been successfully used to specify labeling targets for analysts, to provide a head start for spectral clustering, to flag mixture or edge pixels, and to form a basis to "compress" a scene (forming quasi-field spectral mean values and substantially reducing the data volume). The extent to which the purpose of spatial-spectral clustering has been achieved was discussed in previous reports [3], although more work needs to be done to examine SUPERB.

What follows is a description of the basic algorithm of SUPERB, and a discussion of its special features. As a summary, the principle new features of SUPERB are given in Table D-1, along with the expected desirable effects of using those features. These features represent the main differences from algorithm BLOB.

D.2 ALGORITHM DESCRIPTION

D.2.1 GENERAL

To help understand the system environment, we mention that SUPERB is written in an extended FORTRAN ("Overdrive") as a module within ERIM's

TABLE D-1. DEVELOPMENT HIGHLIGHTS OF SPATIAL-SPECTRAL CLUSTERING PERFORMED BY ALGORITHM SUPERB

Development	Significance
Ground truth supervision is available as an option.	Formation of quasi-fields can be forced consistent with ground truth boundaries. The resulting blobs can be used as ground truth fields, in some applications.
A new distance measure allows separate consideration of spatial and spectral effects.	It is easier to judge the physical significance of threshold parameters values.
A pixel may be assigned only to a quasi-field associated with a neighboring pixel (or form a new quasi-field).	Quasi-fields now can be made to consist of a single group of adjacent pixels. Fewer distance calculations need to be performed.
When deciding whether to assign a pixel to a blob versus form a new quasi-field, pixels having the easiest decisions are processed first (within each scan line).	More candidate neighbor quasi-fields often become available to help improve later decisions. Biasing effects of sequential processing order are minimized.
Newly formed blobs are compared to their neighbor blobs, and combined if sufficiently similar.	If two blobs happen to start in the same homogeneous field, they may be combined.

QLINE system. QLINE is a modular multispectral software architecture in which a module is called (usually) once for each scan line in a scene (the module is called for various kinds of initialization and post-processing as well). A scan line of data is made available to the module in the form of a two-dimensional array (channels by pixels in the scan line). The module may use or modify the data, for example, by adding one or more channels representing processed data.

Having discussed the environment in which SUPERB exists, we now describe the processing carried out for each scan line. First, the pixels are processed sequentially right to left. A distance measure, one of those selected from Table D-2, is computed between the pixel and each candidate blob. The candidate blobs are all blobs that were assigned to pixels neighboring the pixel being processed. (This neighborhood can be modified by a user to any subset of pixels up to two units away, but normally consists of five neighbors--upper left, upper, upper right, right, and left. On this first pass, the left neighbor, and sometimes the right, would not yet have an assigned blob.) For this pixel, the spatial-spectral distance D to the nearest candidate blob is compared to a pair of thresholds d_ℓ and d_h. If $D \leq d_\ell$, the pixel is assigned to the blob and is used to update the stored blob mean vector. If $D \geq d_h$, the pixel forms a new blob. However, if D is between d_ℓ and d_h, the decision is delayed, and the pixel is placed on a deferred list along with its value of D.

Once all pixels are processed in this manner, the list of deferred pixels is sorted in order by descending values of $|D-d_f|$, where d_f is the final decision threshold ($d_\ell \leq d_f \leq d_h$). Each of these pixels are processed in sorted order by computing distances from the pixel to a new set of candidate blobs. (There may be more candidate blobs during the second try for a given pixel, since its neighbors may have been assigned to blobs.) The new smallest distance D is compared to threshold d_f. If $D > d_f$, the pixel forms a new blob. Otherwise, it is assigned to the nearby blob with distance D.

Of the three distance measures described in Table D-2, the first, a composite of spectral and spatial measures, is one of the options in BLOB. The second is a pure spectral measure, while the third is the new distance vector consisting of two elements (D_1 and D_2) representing spectral and spatial components, respectively. In the last case, the three decisions given in the above paragraphs must be specified in terms of D_1 and D_2, as shown in Table D-2. In this case, deferred pixels are sorted in descending order of $D_1 + D_2$.

The foregoing discussion covers the basic mechanism, when no special conditions or options are present. What follows is the discussion of several features and conditions that can modify the basic picture somewhat.

D.2.2 GROUND TRUTH SUPERVISION

If available, ground truth can be used to constrain the algorithm so that each blob must consist of pixels of a single ground class.

In one of two supervision options, the ground truth is supplied in the form of an integer code for each channel. When distance measures are computed between a pixel and its neighbor blobs, only blobs with matching ground truth code are considered. Furthermore, codes of zero are used to mark pixels as mixture or unknown, and such pixels are assigned to a single garbage blob.

In the second supervision option, the ground truth is supplied in the form of several subpixel codes for each pixel, combined with one code that summarizes the several. The summary code is zero if the subpixel codes do not all agree, and is the same as the subpixel codes if they agree. In this option, the candidate blobs for a given pixel are neighbor blobs whose ground truth code matches at least one of the subpixel codes for the pixel. When a pixel forms a new

TABLE D-2. DISTANCE MEASURES THAT MAY BE SELECTED IN SUPERB

	Distance Measure	Decision Rules: Initial Pass — Assign Pixel to Nearest Blob, if*	Decision Rules: Initial Pass — Form New Blob, if*	Decision Rules: Deferred Pass — Assign Pixel to Nearest Blob, if**
1.	$D = \left[\sum_i \frac{(x_i - \overline{x}_i)^2}{v_i}\right] + \left[\frac{(\ell - \overline{\ell})^2}{v_\ell} + \frac{(p - \overline{p})^2}{v_p}\right]$	$D \leq d_\ell$	$D \geq d_h$	$D \leq d_f$
2.	$D = \left[\sum_i \frac{(x_i - \overline{x}_i)^2}{v_i}\right]$	$D \leq d_\ell$	$D \geq d_h$	$D \leq d_f$
3.	$D_1 = \left[\sum_i \frac{(x_i - \overline{x}_i)^2}{v_i}\right]$ $D_2 = \left[\frac{(\ell - \overline{\ell})^2}{v_\ell} + \frac{(p - \overline{p})^2}{v_p}\right]$	$D_1 \leq d_\ell$ and $D_2 \leq s_\ell$	$D_1 \geq d_h$ or $D_2 \geq s_h$	$D_1 \leq d_f$ and $D_2 \leq s_f$
		*Otherwise defer the decision.		**Otherwise form new blob.

Table Notes:

- D = overall pixel-to-nearest-blob distance
- D_1 = spectral pixel-to-nearest-blob distance
- D_2 = spatial pixel-to-nearest-blob distance
- d_ℓ = low threshold
- s_ℓ = low threshold for spatial
- d_h = high threshold
- s_h = high threshold for spatial
- d_f = final threshold
- s_f = final threshold for spatial
- v_i = allowed variance in channel i
- v_ℓ = allowed variance in line coordinate
- v_p = allowed variance in point coordinate
- x_i = pixel value in channel i
- $\overline{x}_i$ = blob mean value in channel i
- ℓ = line coordinate of pixel
- $\overline{\ell}$ = mean line coordinate in blob
- p = point coordinate of pixel
- $\overline{p}$ = mean point coordinate in blob

blob, the ground truth code for the blob is the most frequent subpixel code in the pixel. If there is a tie for most frequent subpixel code, the one-pixel blob is considered "unlabelable", and is combined with the garbage blob. If the most frequent subpixel code is zero, the one-pixel blob is combined with the garbage blob.

Whichever supervision option is selected, the function is to limit the set of neighbor blobs that are considered in computing distance measures to (and therefore possibly being combined with) each pixel. An additional effect of supervision is that ground truth mixture pixels will not form new blobs except during deferred pixel processing. This may reduce the chance of a mixture pixel that results in an unlabelable one-pixel blob, since a neighboring pure or less ambiguous pixel may initiate a blob that would absorb the mixture pixel. A second additional effect is that, for each pixel, a result ground truth code is produced that reflects the ground truth code of the blob containing that pixel. This code may be a stronger indication of the pixel's primary class than the ground truth codes themselves, since slight misregistrations of the ground truth can be overridden for mixture pixels by spectral similarity to neighbor classes.

D.2.3 USE OF SCREEN INFORMATION

Algorithm SCREEN [36] is designed to examine Landsat data and identify those pixels which probably aren't good agricultural data, such as clouds, water, etc, and this can be carried out for each acquisition in multitemporal data. When SUPERB is run, it can examine information from SCREEN and tell, on a pixel-by-pixel basis, which channels contain valid data. Usually, all channels are valid, in which case SUPERB works normally.

The behavior is modified, when some channels are not usable, as follows. The spectral part of the distance measure is computed as:

$$\left[\sum_{i=\{\text{good channels}\}} \frac{\left(X_i-\bar{X}_i\right)^2}{V_i}\right] \times \frac{\text{(number of channels specified)}}{\text{(number of good channels)}}$$

where the good channels are those specified that are not flagged by SCREEN for the pixel, and also are not flagged in SCREEN information maintained for the blob.

When a pixel is used to form or update a blob, care is taken so that channels flagged (as clouds, etc) for the pixel do not change the blob mean in those channels, and channels flagged for the blob cannot be updated but only replaced by unflagged pixel channels. There must be at least a specified minimum number of non-flagged channels in a pixel, or the pixel will be assigned to the garbage blob. It is occasionally possible for a finished blob to occur without any valid data in some channels, in which case those channels would contain bad information (namely the values for the first pixel assigned to the blob). These are indicated with a warning message, but currently are not discarded or otherwise marked.

D.2.4 BLOB NEIGHBOR TABLE

Algorithm SUPERB as an option can prepare a table that lists for each blob, a list of blobs that neighbor it. The definition of neighborhood used to form this table is the same as the neighborhood described previously for establishing candidate blobs for a pixel when determining the pixel's assignment. A blob is adjacent to another if any pixel of one blob is in the neighborhood of any pixel of the other blob. The resulting table can be written and stored for use by blob mapping algorithms and other analysis routines that work with blobs.

D.2.5 BLOB COMBINATION OPERATIONS

Occasionally, it can happen that a group of adjacent spectrally similar pixels which normally would comprise a single blob, may be started from two sides, resulting in two blobs rather than one. One way this can happen is that the upper corner of a field on one scan line contains a pixel that starts a blob, then on the next scan line, a pixel on the opposite side of the same field starts a blob also.

A first attempt to reduce this possibility was based on the fact that rectangular fields (in the U.S.) tend to be rotated a few degrees counter-clockwise to the scan line direction, so that the upper corner of a field normally belongs to the right-hand side of a field. When the order of processing was changed from left-to-right to right-to-left, the problem of multiple starts was reduced, but still occurred for some field geometries.

In order to more fully solve the problem of multiple starts, an option was implemented to consider combining adjacent blobs. After all pixels on a scan line are assigned, any newly formed blob is compared to its neighboring blobs, taken from the blob neighbor table. Using the mean of a newly formed blob as if the blob were a pixel, the distance function is computed to each neighbor blob (assuming that if supervision is in effect that the blob ground truth codes match). If the nearest blob is sufficiently close (using thresholds d_f' and s_f', which may but need not be the same as d_f and s_f as discussed above), the blobs are combined, taking the same care with SCREEN information that is used when assigning a pixel to a blob.

APPENDIX E
LOGIC FOR PLACEMENT OF LANDSAT DATA INTO COLOR IMAGERY

E.1 INTRODUCTION

A first principle in design of imagery is that independent, important features of the data should translate to independent dimensions of color variation as perceived by the human eye. In this way, one allows the visual "information processor" to correctly come into play in analyzing the image. Imagery acts as the interface between the data and the human being, tying us into a psychological processor which is quite impressive for its pattern recognition capabilities. To best use this processor it is important to understand and respect its inherent manner of classifying a scene, literally its way of looking at things. This appendix describes where these considerations have led us in designing imagery for Landsat data, and what logical options appear to be available to us, given our current understanding of the data structure.

E.2 MAPPING DATA FEATURES TO COLOR VARIATION

It is possible to place more than one component of information about a scene into a color image, and still have those components independently apparent to and assessible by an observer. This is possible because perception of color is multidimensional. Colors are perceived to vary in three independent characteristics or along three independent dimensions. If we control each independent way of changing a color with a different information component then the information components remain independently apparent and assessible.

The importance of making our information components control independent dimensions of color is not hard to see. Under this arrangement we are properly linking up with the visual information processor. The

observer (image analyst) perceives the information components in the scene as though the human eye had evolved so as to pick them out. The effort involved in interpretation of color is minimized because much of the work of classification occurs immediately in the visual system. The information is presented in the form required for pattern recognition by the observer.

The predominant dimensions of color variation are three in number. They bear the names hue, saturation and lightness and are defined as follows [53]:

Hue	is the attribute of a color perception which gives rise to the names red, purple, yellow, etc.
Saturation	is the attribute of a color perception determining the degree of its difference from the achromatic (grey) color perception closest to it, i.e., of the same lightness.
Lightness	is the attribute of a color perception permitting it to be classed as equivalent to some member of the series of achromatic color perceptions ranging from black to white (grey levels).

The attributes of hue, saturation and lightness are subjectively defined and are not predefined in any analytic or quantitative sense.

A second important concept in mapping data variation into color is the perceptual scaling of color differences. It is possible to find a sequence of colors between two given colors such that the perceived size of the color change from step to step is uniform, i.e., the same at each step. Naturally, we wish the distance relationship between data

points to be preserved for the observer of the image. The way to do this is by mapping the data interval onto a uniform color scale of the sort described. In this manner, a given distance in the data interval is consistently transformed to color differences of the same size. This is important since large deviations from uniform transformation will certainly give us a distorted view of data variation in our image.

In order to portray more than one variable in color while maintaining uniform distance translation we need an extension of the concept of color uniformity to more than one dimension. The single dimensional uniform color scale gives rise to the concept of placing colors in two- or three-dimensional configurations for which distance between any colors still correlates to the psychologically perceived difference between those colors. The Munsell color specification system is an attempt to set up such a configuration for all object colors. The configuration was arrived at empirically and is defined by cards of color chips. Many attempts have been made to describe the desired spacing of colors mathematically. The usual starting point is the X,Y,Z color space. Transformations are made on these variables to arrive at a color space which is more uniform to the eye's judgement of color differences. Currently the L^*,a^*,b^* color space is the standard Uniform Color Space designated by CIE. Its relatively simple equations obtain a color spacing in good agreement with the Munsell system.

E.3 DIFFERENT COLOR SPACES, THEIR UNIFORMITY AND THE LOCI OF HUE, SATURATION AND LIGHTNESS

E.3.1 X,Y,Z COLOR SPACE [53]

We do not perceive a unique color for every different light stimulus presented to the eye. For example, Figure E-1 shows 12 reflectance curves which are markedly different but which nonetheless give rise to

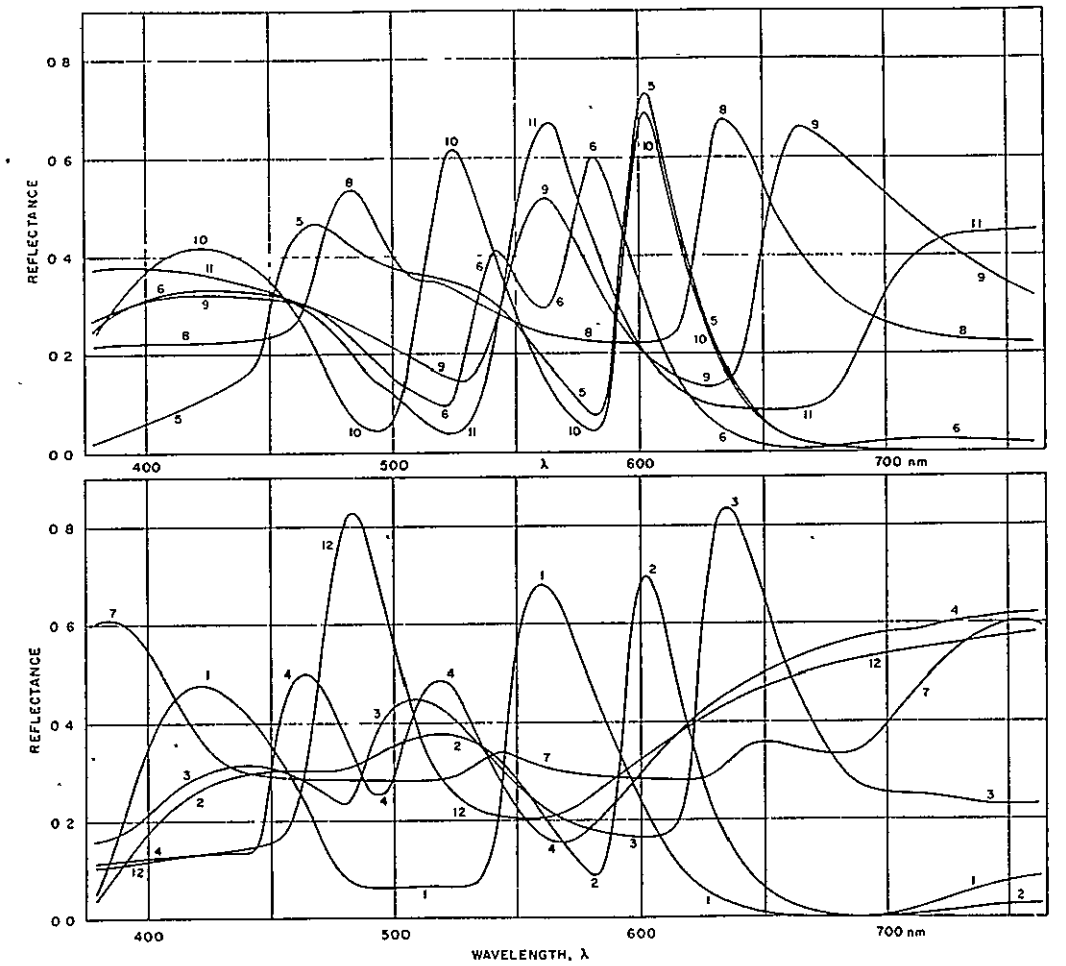


FIGURE E-1 [53]. TWELVE SPECTRAL RESPONSE FUNCTIONS OF OBJECT COLORS METAMERIC WITH RESPECT TO CIE SOURCE C AND THE 1931 CIE STANDARD OBSERVER

the same color perception. The reason for this is that the eye has a finite number of "sensor channels", in fact it has only three. Three pigments, located in the cones of the retina, are involved in color vision. Each pigment has a different spectral sensitivity to light so each acts like a sensor channel with a different bandpass. Many different spectra can produce the same color sensation as long as the trio of channel values, referred to as the "tristimulus values", end up the same.

A set of spectral sensitivity curves for the eye's three receptors have been defined for a standard or ideal observer, based on results of color matching experiments. These three weighting curves appear in Figure E-2. They are labeled $\bar{b}_\lambda$, $\bar{g}_\lambda$ and $\bar{r}_\lambda$ and are referred to as color matching functions. In 1931, the Commission Internationale d'Eclairago or CIE set down a linear combination of the color matching functions with certain specially desirable properties, to be the basis of a standard color space. The color matching functions used by the CIE are shown in Figure E-3 and are labeled $\bar{x}_\lambda$, $\bar{y}_\lambda$, $\bar{z}_\lambda$.

When these weighting curves are applied to spectra, the three resulting numbers are called tristimulus values and labeled X,Y,Z. In constructing the weighting functions $\bar{x}_\lambda$, $\bar{y}_\lambda$, $\bar{z}_\lambda$ the constraint was made that $\bar{y}_\lambda$ would have the same shape as the photopic (cone vision) luminous efficiency function. Thus the Y tristimulus value does double duty -- it also is a predictor of color lightness. Empirical tests have shown that Y is a good, approximate predictor of perceived lightness (in the sense of lightness matching between colors, not in the sense of uniform spacing of color lightnesses), but has a small systematic problem, namely underestimating the lightnesses of saturated colors. The additional constraint was made that the $\bar{x}_\lambda$, $\bar{y}_\lambda$, $\bar{z}_\lambda$ weighting or color-matching functions would be everywhere non-negative. Since spectra are non-negative, this implies the X,Y,Z tristimulus coordinates will always have non-negative values.

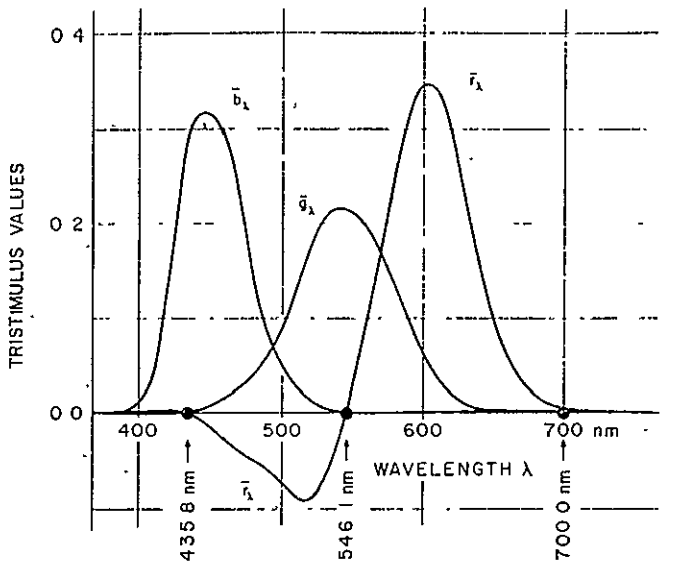


FIGURE E-2 [53]. COLOR-MATCHING FUNCTIONS, $\overline{r}_\lambda, \overline{g}_\lambda, \overline{b}_\lambda$

IN THE PRIMARY SYSTEM R = 700.00 nm,
G = 546.1 nm, B = 435.8 nm

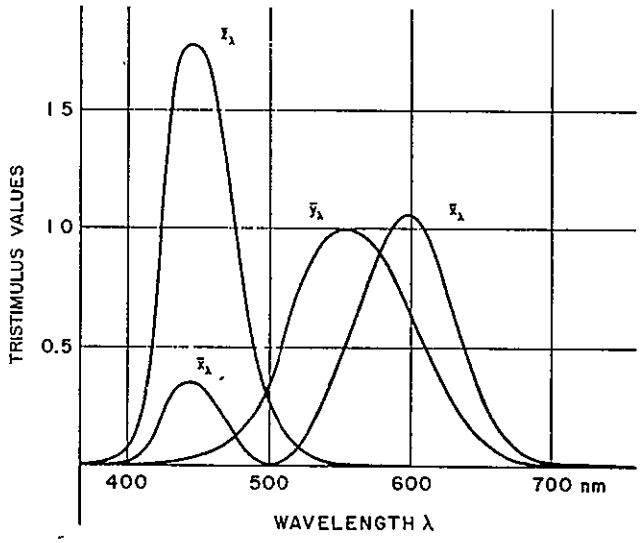


FIGURE E-3 [53]. COLOR-MATCHING FUNCTIONS OF THE 1931 CIE STANDARD OBSERVER

In X,Y,Z space, colors of fixed chromaticity (hue and saturation) fall on lines which emanate radially from the origin. Hue and saturation are determined by the ratio of the X,Y,Z tristimulus coordinates and a constant ratio among the coordinates is found on any line out from the origin. To see why this is so, imagine the following experiment. Light is projected onto a screen from a white light source after passing through a given, fixed color filter. The white light source has a dimmer switch so we can vary its output. As we vary the source from dim to bright the perceived lightness of the projected, filtered light will change correspondingly; however, its perceived chromaticity, i.e., hue and saturation, will remain the same. Now if we calculate the X,Y,Z tristimulus values for the color at several different lightnesses we will find the magnitude of the vector varies, but not its direction. This is a consequence of the fact that the X,Y,Z measurements respond linearly to changes in energy. If you double the energy of a spectra, or if you half it, each of the X, Y, and Z values changes correspondingly and their ratio remains unchanged. Thus the locus of points in X,Y,Z space which could be produced in this experiment fall on a line out from the origin with direction determined by the filter used.

Finding the locus of constant hue with variable saturation is not so simple. Neither is finding the locus of constant saturation with variable hue. Isolating hue or saturation is a psychological judgment without a simple mathematical correlate in X,Y,Z space. Lines of constant hue and lines of constant saturation have been empirically mapped. A typical result of this mapping is shown in Figure E-4, plotted in the x,y chromaticity diagram for a fixed lightness. Notice that hue tends to be radially distributed while saturation tends to be cylindrically distributed.

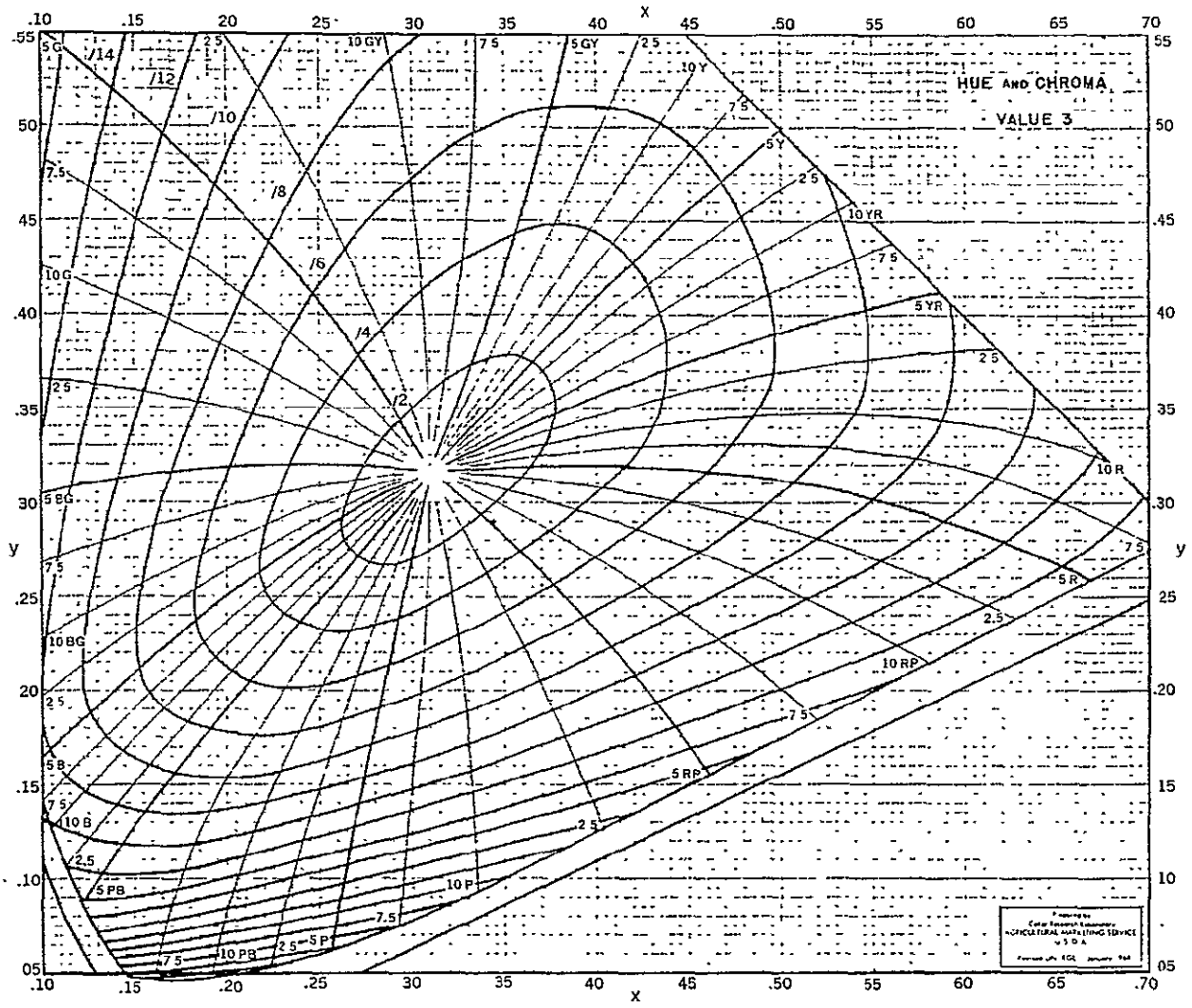


FIGURE E-4 [53]. 1931 CIE CHROMATICITY DIAGRAM SHOWING LOCI OF CONSTANT HUE AND CONSTANT SATURATION AT LIGHTNESS VALUE 3/ OF MUNSELL RENOTATION SYSTEM

E.3.2 L^*,a^*,b^* COLOR SPACE

If the eye gauged differences between colors linearly with respect to differences in energy of the light stimuli, then the X,Y,Z color space would be an exactly uniform color space. However, the eye does not judge color differences linearly with respect to energy differences. Empirical findings show that the psychological magnitude of color differences are a function of energy differences under transformation by a logarithm or fractional power less than 1/2. One formula does not work under all conditions because the background of the colors being compared, i.e., the adaptive state of the observer's eye, is an important factor. A logarithmic transformation works best when the background is a color in between the colors being judged. A cube root transformation works best given a constant middle grey background (reflectance of 20%). A square root transformation works best given a white background for all observations. As a standard transformation, the cube root formula has found favor. It has found a place in the Munsell color specification system and the equations of L^*,a^*,b^* Uniform Color Space.

The L^*,a^*,b^* standard Uniform Color Space is built from the X,Y,X color space in a few straightforward steps. First, coordinates are normalized to the nominal white of your particular color production hardware. Denote the tristimulus values of "white" by X_o,Y_o,Z_o and normalize all other coordinates by forming the ratios X/X_o, Y/Y_o and Z/Z_o. Second, take the cube root of each coordinate axis. This step incorporates the understanding of psychological equispacing discussed above. Third, a linear transformation is applied to the axes to line them up with what are believed to be the directions of chromatic balance which the eye uses in differentiating color. These poles of chromatic balance are blue vs. yellow and red vs. green. The a^* coordinate labels the red/green balance while b^* labels the blue/yellow balance. The transformation, thus far described, has the following form:

$$\begin{pmatrix} L^* \\ a^* \\ b^* \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} (X/X_o)^{1/3} \\ (Y/Y_o)^{1/3} \\ (X/X_o)^{1/3} \end{pmatrix}$$

The fourth and final steps of the transformation weights the coordinates to obtain a color spacing more correlated with psychological judgment of color differences. The complete transformation in its accepted form appears as follows:

$$L^* = 25\ (100 \cdot Y/Y_o)^{1/3} - 16$$

$$a^* = 500\ [(X/X_o)^{1/3} - (Y/Y_o)^{1/3}]$$

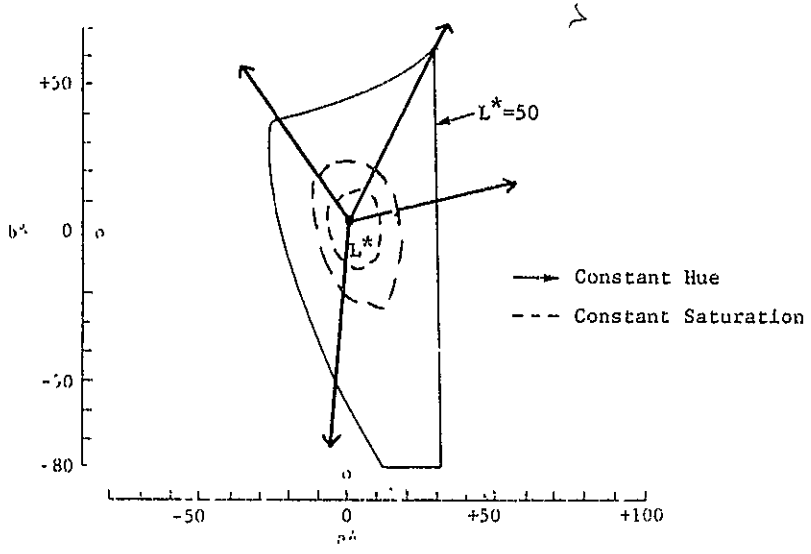
$$b^* = 200\ [(Y/Y_o)^{1/3} - (Z/Z_o)^{1/3}]$$

What should L^*,a^*,b^* space look like in principle? Where are colors located and in what manner are they organized, as far as the psychological dimensions of color perception are concerned? The predominant psychological dimensions of color are lightness, hue and saturation. We will review the structure of L^*,a^*,b^* color space with respect to each of these dimensions.

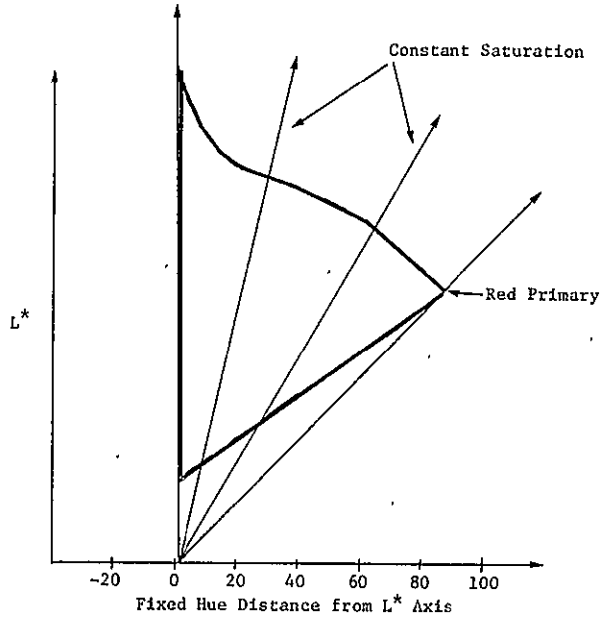
L^* depends solely on the Y tristimulus value. The Y value predicts color lightness to a fair approximation. The L^* function only acts to spread out the lightness measure so that 1) it is more nearly uniform to psychological distance estimation and 2) its scale is approximately commensurate with the scales of the chromaticity axes, a^* and b^*.

Hue in L^*,a^*,b^* space is not well defined and can be held constant only to an approximation. Chromaticity in L^*,a^*,b^* space is specified by the two dimensional, rectangular coordinates a^* and b^*. As mentioned above, this pair of coordinates divide up chromaticities by finding the balance of the color on two scales -- a blue to yellow scale and a red to green scale. In principle, hue should be preserved in moving radially out from the L^* axis as illustrated in Figure E-5(a). As long as the ratio of the chromaticity scales is maintained, hue should approximately be maintained. If L^*,a^*,b^* space was truly a Uniform Color Space this would have to be true -- in order for the straight line between two points of the same hue and on the same radial line to be the geodesic or minimal distance path between them, no hue variation could occur along the path. However, L^*,a^*,b^* space is not, even in principle, a truly Uniform Color Space. L^*,a^*,b^* space is only an approximation to uniform scaling. It was selected as a standard because it worked about as well as other spaces (in matching empirical findings for small color differences) and, at the same time, was mathematically very simple. What this means is that hue need not be precisely radially distributed out from the L^* axis. If the exponent of the L^*,a^*,b^* transformation were changed, say to 1/4 or 1/2 or 0.4, then the colors which are associated by virtue of being in the same radial direction would change. As discussed above, these values of the exponent would not be unreasonable -- there is nothing magic about the exponent 1/3. Therefore, the loci of constant hue are not well defined, i.e., analytically defined in L^*,a^*,b^* space. We can only say hue will be approximately radially distributed.

Saturation in L^*,a^*,b^* space is distributed radially from the origin of the space. Figure E-5(b) shows a plane of L^*,a^*,b^* space (containing the L^* axis) with lines of constant saturation drawn in. In this plane we are approximately fixing hue. The loci of constant saturation for varying hue are not simple. Clearly saturation will tend to vary in cones of increasing opening angle. (Imagine the surfaces



(a) Idealized Hue and Saturation Intervals at $L^*=50$



(b) Idealized Constant Saturation for a Fixed Hue

FIGURE E-5. LIGHTNESS, HUE AND SATURATION IN L^*,a^*,b^* SPACE

generated by rotating the lines of Figure E-5(b) about the L^* axis.) To be more exact than this we would have to translate the empirical findings about judgment of saturation, as displayed in Figure E-4, into L^*,a^*,b^* coordinates. Figure E-5(a) illustrates contours of constant saturation.

In X,Y,Z space the locus points for which chromaticity (hue and saturation simultaneously) is constant is a line radiating from the origin of the space. This was explained in the above section on X,Y,Z space. The transformations involved in passing to L^*,a^*,b^* space preserve this property. We have discussed where the loci of constant hue or constant saturation fall above. The locus of a constant chromaticity, hue and saturation considered simultaneously, is still a straight line radiating from the origin.† The cube root transformation moves these lines around -- basically it draws more saturated colors in towards the diagonal of the space. However, the lines of constant chromaticity remain straight lines radiating from the origin.

E.3.3 THE COLOR GUN CUBE AS A COLOR SPACE

The exposure control system of the Production Film Converter is a means of color specification and, hence, the inputs to the system may be thought of as coordinates in a color space. The coordinates are the three gun numbers we input to the PFC and the resulting color space volume we refer to as the color gun cube. In the past, many assumptions were implicitly made about properties of this color space which were simply false. They were very appealing assumptions to make -- it was assumed that the space was uniform, that lines of constant chromaticity radiated from the origin and that color lightness was determined by the sum of gun counts. A sizeable amount of work, related to designing film products [23,24,25] and to analyzing those products [54,55,56] was

†The origin of L^*,a^*,b^* space is (-16,0,0) to correspond with X = 0, Y = 0, Z = 0.

based on those implicit assumptions. In this section we review what a color science analysis predicts about the PFC color gun color space.

The PFC color gun color space is a logarithm color space. Whereas L^*,a^*,b^* space used a cube root transformation on transmission, the PFC uses a base 10 logarithm. This is the way the PFC was set up to work. A set of three look-up tables were built into the PFC, one for each primary color gun, which establishes a linear relationship between gun counts and the image densities (i.e., the densities of the film to transmission of light at the three wavelengths which represent the primaries of the film, density is the $\log_{10}$ of transmission which is the percentage of light energy the film lets through). Note that the logarithmic transformation is not applied to the axes of X,Y,Z space but to the primaries of the PFC, each of which is expressable as a vector in X,Y,Z space.

In applying a logarithmic transformation to a transmission color space one does not preserve the property that lines of constant chromaticity radiate from the origin. Any exponential transformation preserves this property but the logarithm does not. The logarithm changes fixed ratios to fixed distances. Consider two lines which meet in the origin of the original space. Each of these lines represents a fixed ratio among the coordinates of the space. If we are talking about a color space in which the coordinates measure transmissions, then these are two lines of constant chromaticity. If we now apply a logarithmic transformation to the coordinates of the space the two lines with different fixed ratios become two lines with a fixed distance between them, in other words, parallel lines. Mathematically, it looks like this. Consider two lines in the original space linked with the parameter t:

$$\vec{r}_1 = t\ (x_1,y_1,z_1)$$

$$\vec{r}_2 = t\ (x_2,y_2,z_2)$$

$$0 \leq t \leq \infty$$

Now apply a logarithmic vector transformation, L, to the space in the following manner:

$$u = \log x$$
$$v = \log y$$
$$w = \log z$$

The two lines are transformed as follows:

$$\vec{\ell}_1 = L(\vec{r}_1) = (u_1, v_1, w_1) + (\log t, \log t, \log t)$$
$$\vec{\ell}_2 = L(\vec{r}_2) = (u_2, v_2, w_2) + (\log t, \log t, \log t)$$

where $u_1 = \log x_1$, etc. Now look at the point-for-point difference between the transformed vectors:

$$\vec{\ell}_1 - \vec{\ell}_2 = (u_1 - u_2,\ v_1 - v_2,\ w_1 - w_2)$$

We see that it is a constant vector. The vectors $\overline{\ell}_1$ and $\overline{\ell}_2$ are parallel. By the same token, all lines in the original space which meet in the origin end up parallel to one another.

The currently used alternative image product, the Kraus Product or Product 3, is claimed to maintain consistency in color definition. A basic design principle of the product was that data points with the same ratios among the three channels should be mapped to colors of the same chromaticity (hue and saturation), with different lightnesses. To do this is to preserve "angularity" (a word coined by Richard Juday/JSC). The Kraus Product attempts to do this by preserving data channel ratios in the color gun coordinate ratios. This would work, provided that the lines of constant ratio among the color gun coordinates were indeed lines of constant chromaticity. As we have seen, such is not the case. Lines of constant chromaticity are parallel. Along a vector out from the origin, color saturation does not remain constant. It increases, changing from zero (achromatic) to fully saturated.

Lightness of colors cannot be predicted by summing the color gun counts. For example, using 255 counts on the green gun produces a much lighter color than using 255 counts on the blue gun. This is because the eye is more sensitive to the wavelengths of light permitted by the green primary. A weighted average of the gun counts would do better but would probably not be adequate either. From a color theory standpoint, the iso-lightness contours in the color gun cube will not be planes.

The perceptual uniformity of the color gun cube to visual color difference estimation is problematic. One would expect the logarithmic transformation to improve uniformity over a space which measures transmission. Beyond that, the degree of uniformity is open to question.

Some slices of the PFC color gun cube have been placed on film. The remarks made above about chromaticity and lightness in the color gun cube, and its uniformity, are all consistent with observations which can be made from the film slices of this color space which we have generated.

E.4 MAPPING TASSELED-CAP VARIABLES TO COLOR

The way we would choose to place the Tasseled Cap into color depends upon what we think are really the independent features we want to portray. This comes down to asking whether the features are actually rectangular Brightness and Greenness or polar Green Angle and Brightness Radius.

E.4.1 RECTANGULAR COORDINATES

If we believe that our data space is fundamentally rectangular then the most appropriate color space for us to map it into is similar to the PFC color gun cube. The PFC's logarithmic color space separates changes in saturation and lightness in a rectangular, rather than radial,

manner. If we want lines of constant Greenness to fall onto lines of constant saturation (we assume a fixed hue), then we need a rectangular color space.

Investigations of a modification of the PFC color gun cube to determine if such a space could be a realistic option in our work have been pursued. The modified space, termed L,C,D space, is a linear transformation of the cube defined as follows:

$$\begin{pmatrix} L \\ C \\ D \end{pmatrix} = \begin{pmatrix} .11 & .57 & .32 \\ 0 & 1/2 & -1/2 \\ 1/2 & -1/4 & -1/4 \end{pmatrix} \begin{pmatrix} B \\ G \\ R \end{pmatrix}$$

The C coordinate measures a red/green balance while D measures a blue/yellow balance. The L coordinate is an attempt to come as close as possible to isolating lightness levels within the cube. Classical color theory predicts that iso-lightness levels in the logarithmic space will not be planes. Analysis of imagery filmed on the PFC should be performed to determine how much they deviate from being planar. There is reason to believe the distortion will actually be less than that apparent in the images of L^*,a^*,b^* space we have analyzed. Of course, our version of L^*,a^*,b^* space is based on a model of the PFC with acknowledged inaccuracies, currently being addressed by Juday.

The essential characteristic of the color gun cube, that saturation varies as a rectangular coordinate, is retained under this transformation. It would be interesting to map data into this space as an experiment. The uniformity of the space would have to be assessed empirically, as would the success of isolating lightness levels.

E.4.2 POLAR COORDINATES

It appears that the Landsat data space may be characteristically radial. Intuitively it seems reasonable that spectral data vectors with the same direction but different magnitude represent objects which are somehow similar. Within the resolution of the sensor they represent identical spectral responses with different overall brightnesses. That actually may mean nothing. We have no a priori reason to believe that the two objects in question are really similar in some sense. The modeling work of Reference [57] suggests that, in the case of plant canopies, breaking vectors into direction and magnitude is a meaningful way of looking at them. It appears that vectors of the same direction represent a common stage of plant canopy development while vector magnitude varies with the underlying soil brightness. Figure E-6 shows model results for various degrees of plant cover and three soil brightnesses. The radial pattern apparent in this figure is the reason for saying that Landsat agricultural data may be characteristically radial.

We can make vector magnitude and direction the visually independent features in our imagery in either of two ways. We could transform to polar coordinates, Greenness Angle and Brightness Radius, and proceed to map these variables into a fundamentally rectangular color space, such as L,C,D space. We could also map the data space linearly into a color space which is fundamentally radial. The latter approach has been pursued because we would like to use the L^*,a^*,b^* Uniform Color Space and it is a radial color space. Our scheme is to map the apparent data origin (not the zero vector) to the origin of color space ($L^* = -16$, $a^* = 0$, $b^* = 0$) and align Brightness with L^*. This preserves the radial distinction in the data, as indicated in Figure E-7. The Greenness direction can be rotated to any hue we find convenient. In the imagery currently being processed we have tried placing Greenness along the b^* axis as well as in the direction of the green primary.

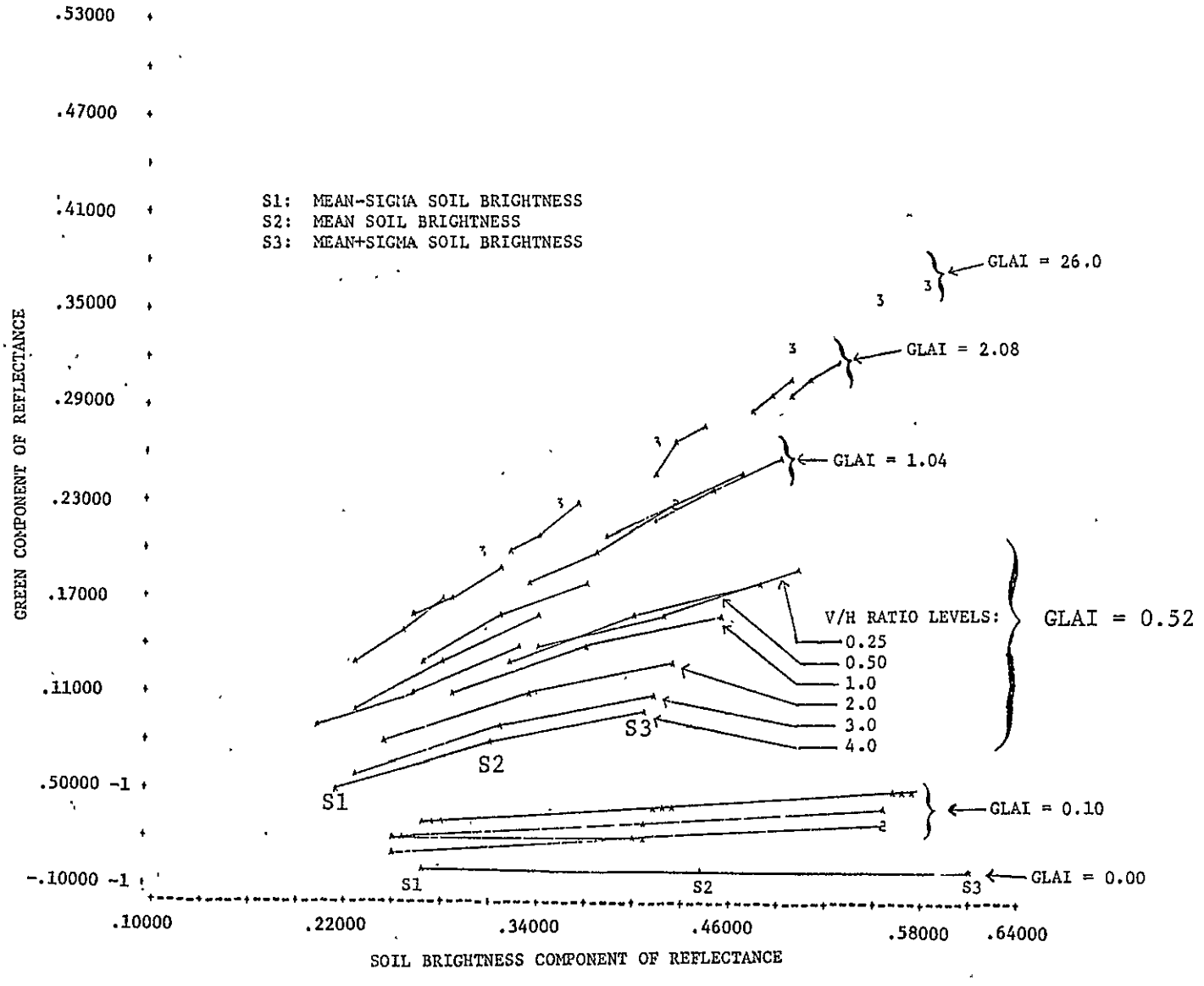


FIGURE E-6. TASSELLED-CAP TRANSFORMATION OF SIMULATED WHEAT REFLECTANCES, ILLUSTRATING SOIL BRIGHTNESS EFFECTS

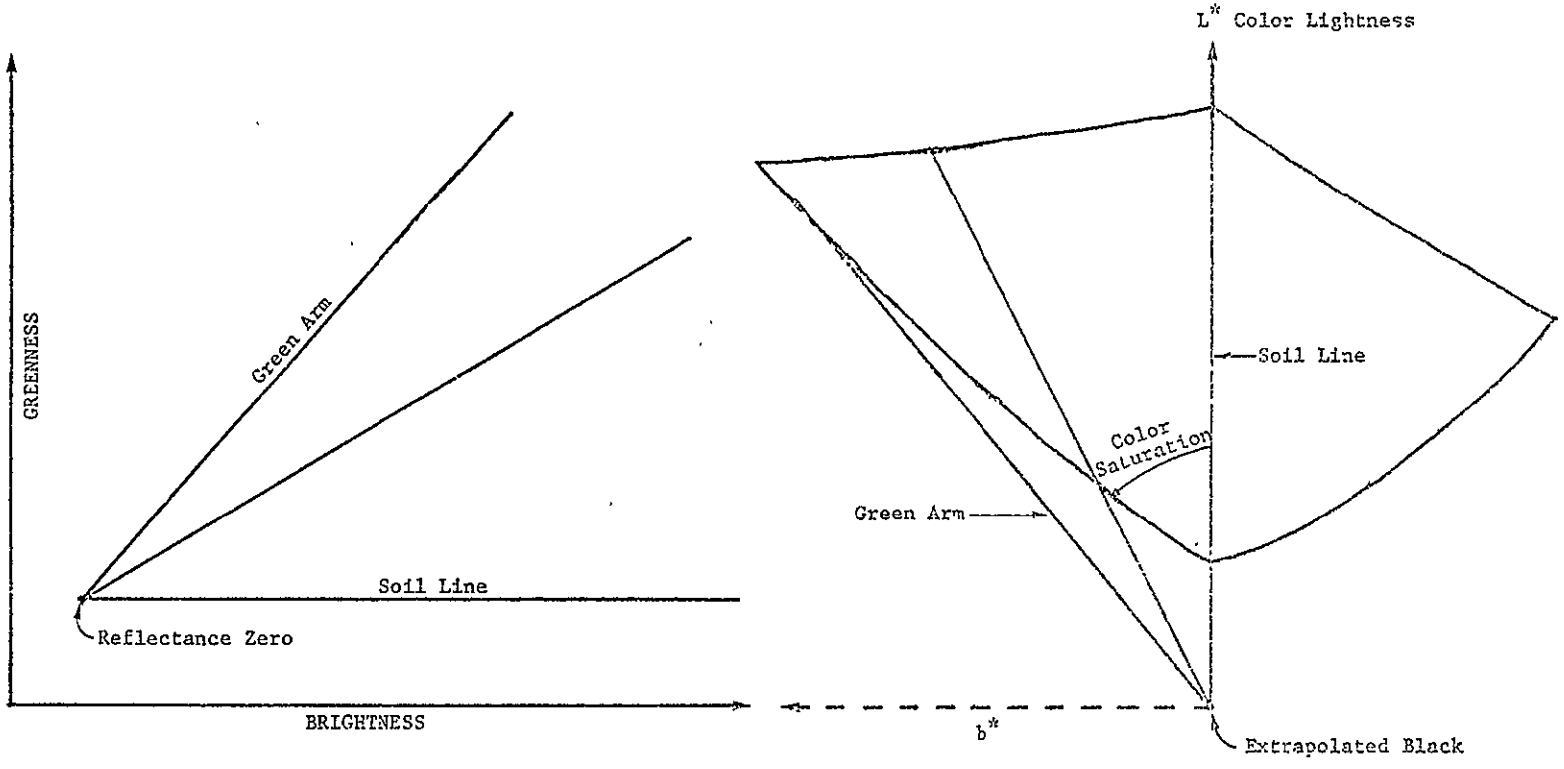


FIGURE E-7. ADVANCED IMAGE PRODUCT

APPENDIX F
SIGNATURE CHARACTERIZATION STUDIES

This appendix contains a broad description of a method for extracting multitemporal, multisegment signatures from training data. As of now, the method has not been evaluated. Section F.1 outlines the method. Section F.2 gives some theory on estimation with missing data which might lend this method some plausibility. Section F.3 gives the formula, required by the method, for computing conditional expectations when the underlying joint density is a mixture density.

F.1 MULTITEMPORAL, MULTISEGMENT SIGNATURES

A method is outlined for extracting signatures from multitemporal, multisegment training data. We suppose that for a fixed object class, say wheat, there is data from T passes with $T \approx 20$. For each of the S segments, $S \approx 40$, we have data for about four of the passes. If we think of feature space as being L, $L \approx 2$, spectral features over all T passes, then the data is incomplete in that we observe about 4 x L components out of a total of approximately $p = L \cdot T \approx 40$. We think of each data vector as being of length $L \cdot T$ with real values for certain components and blanks for the others. The number of incomplete data vectors from each segment is approximately 50 (e.g., 50 wheat blobs).

Thus the data set may be characterized as follows: We have about 2000 incomplete data vectors of length 20L. The missing components are approximately the same within each segment, but may vary from segment to segment. The number of missing components is about 80% of the data vector length.

We would like to obtain a p-variate density function which characterizes the set of the training data vectors. A reasonable form for the

density is a convex mixture of p-variate Gaussian densities, i.e., a density f of the form

$$f = \sum_{i=1}^{r} \lambda_i f_i$$

where $\lambda_i \geq 0$, $\sum_{i=1}^{r} \lambda_i = 1$, and f_i are p-variate Gaussian densities; with r, the λ_i, and the parameters of the f_i unknown. Rassbach et al [7] have developed a procedure for doing this when the data vectors are complete. The computer program implementation is called CLASSY.

We describe an iterative procedure for extending CLASSY to the case of incomplete data vectors. A flow chart of the procedure is given in Figure F-1. The first step is to fill in the missing data values. To do this we use the method of profiles (see Section 4.4 and Appendix A). Loosely speaking, this method fits a curve based on empirical data to the observed data values and permits extrapolation and interpolation. The next step is to employ CLASSY on the completed data vectors to obtain a signature density function

$$f = \sum \lambda_i f_i$$

Next is a decision as to whether to stop the iteration. The criterion for stopping might be a fixed number of iterations or might be based on the change in f after successive iterations.

If the iteration continues, the next step is to fill in the incomplete data vectors using the method of conditional expectation relative to f. By that we mean the following. Let Z be a random variable having density f. Let V be the random variable determined by a particular (observed) combination of components of Z, and Y the random variable determined by the remaining (not observed) components of Z. Then we have the conditional expectation $E_{Y|V}(v)$ of Y given V = v. Now for a

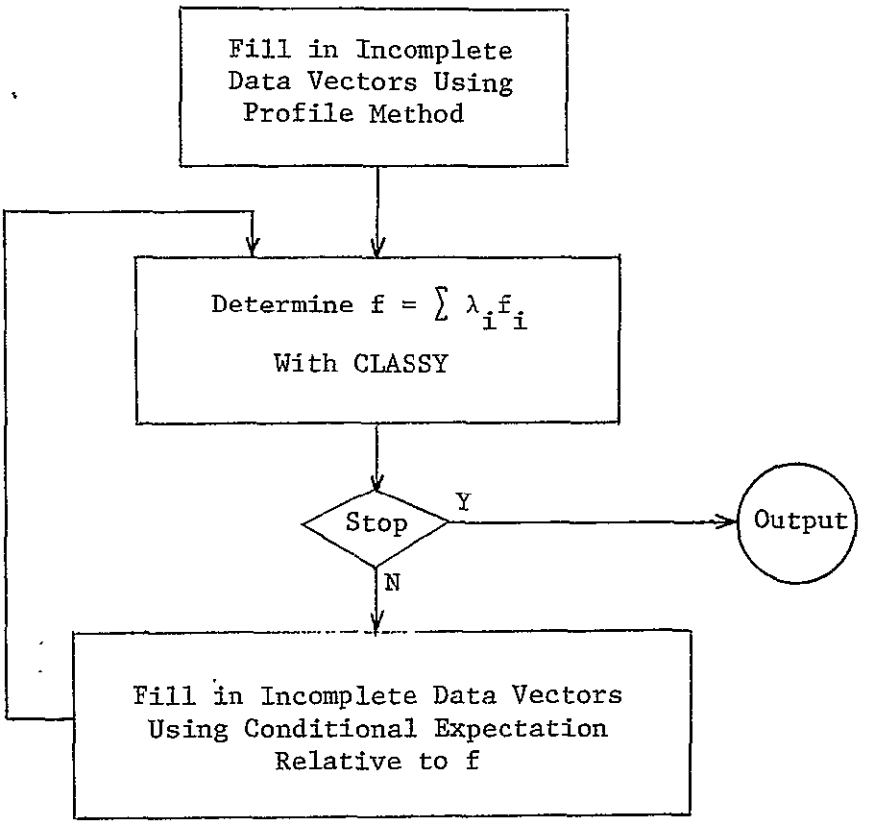


FIGURE F-1. FLOW CHART FOR SIGNATURE EXTRACTION PROCEDURE

specific incomplete data vector z_o, let v_o denote the vector of observed components. Let the random variable V be associated with this set of components and let the random variable Y be associated with the remaining components. Then the missing components are filled in with the values

$$E_{Y|V}(v_o).$$

It would be interesting to try this procedure on a small scale using one spectral feature, say Greenness, and a limited number of passes, say 10. The following remarks may be helpful:

1. Our concept is that the data vectors are actually blob means.
2. The signatures will be both wheat and non-wheat data separately clustered.
3. There does not appear to be any conceptual difficulty in incorporating collateral features into this signature extraction procedure.
4. The problem of estimation with incomplete data has been considered by Boullion [31] et al. They consider the case of a sample from a multivariate Gaussian.
5. There is concern that the covariance matrix obtained after the first step might be ill-conditioned. It may be necessary to add a diagonal matrix to it to overcome this difficulty.
6. An excellent presentation of conditional expectation for Gaussian distributions is in Anderson [58]; note especially Theorem 2.5.1, p. 29.

F.2 ESTIMATION WITH INCOMPLETE DATA

We supply here some background to lend plausibility to the procedure described above. Let X denote a data matrix of dimensions p x N with the columns representing the data vectors. We suppose that some of the entries of X are missing. For each missing entry we substitute a real variable. If r of the entries of X are missing, then we denote the vector of variables representing them by

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix}$$

Thus X is a matrix valued function of the vector x,

$$X = X(x).$$

Suppose that the columns of X are observations from a p-variate Gaussian distribution with unknown mean vector μ and unknown covariance matrix M. We want to find the maximum likelihood estimate of the triplet x, μ, M.

Let $\bar{\mu}(x)$ denote average column vector of X and let $A = A(x)$ denote the matrix obtained from X by subtracting $\bar{\mu}(x)$ from each column of X. Set

$$M(x) = \frac{1}{N} A(x)A^t(x)$$

and let

$$|M(x)|$$

denote the determinant of the matrix M(x). Then the following theorem may be proved by modifying the proof of Theorem 3.2.1 in Anderson [ ].

Theorem 1: (x^*,μ^*,M^*) is a maximum likelihood estimate iff

$$\begin{aligned} &\text{i} \quad x^* \text{ is optimal for} \\ &\qquad \min_x |M(x)| \qquad\qquad (F\text{-}1) \\ &\text{ii} \quad \mu^* = \bar{\mu}(x^*) \\ &\text{iii} \quad M^* = M(x^*) \end{aligned}$$

In order to describe an iterative procedure for optimizing (F-1), we require some notation. Each column of X contains observed components and missing (not observed) components. If the mean μ and covariance matrix M were known, we could fill in the missing components with their conditional expectations given the values of the observed components. If we do this for each column of X, we will have filled in all the components of x. We call the values of x obtained via this process the conditional expectation of x given the observed entries X_o of X relative to μ and M and denote it by

$$E[x|X_o;\mu,M]$$

For convenience we define g(x) by

$$g(x) = E[x|X_o;\bar{\mu}(x),M(x)]$$

Consider the following iterative procedure. Choose $x^{(0)}$ arbitrarily.

$$\begin{aligned} x^{(1)} &= g(x^{(0)}) \\ x^{(2)} &= g(x^{(1)}) \\ &\cdots \\ x^{(n)} &= g(x^{(n-1)}) \end{aligned}$$

It can be shown that

$$|M(x^{(n-1)})| \geq |M(x^{(n)})|$$

with equality iff

$$x^{(n)} = x^{(n-1)}$$

If $x^{(n)}$ converges to x^*, then

$$x^* = g(x^*)$$

In fact, the limit of any convergent subsequence of the $x^{(n)}$ will be a solution of the equation $x = g(x)$. If (x^*,μ^*,M^*) is a maximum likelihood estimate, then

$$x^* = g(x^*), \quad \mu^* = \overline{\mu}(x^*), \quad M^* = M(x^*)$$

From these conclusions, we obtain the following theorem.

Theorem 2: If there is unique x^* for which

$$x^* = g(x^*)$$

then $x^*,\overline{\mu}(x^*),M(x^*)$ is the unique maximum likelihood estimate and the $x^{(n)}$ in the iterative procedure will converge to x^*.

The following question has not been answered: What condition does $X(x)$ have to satisfy to insure a unique solution to the equation $x = g(x)$?

F.3 COMPUTATION OF CONDITIONAL EXPECTATION

At each step in the iteration of the procedure of Section F.1, the CLASSY stage provides an estimate of the signature density f which is a convex mixture of Gaussian densities, i.e.,

$$f = \sum_{1}^{n} \lambda_i f_i$$

where $\lambda_i \geq 0$, $\sum_i \lambda_i = 1$, and the f_i are Gaussian densities. We give here the formula required for computing the conditional expectation of the missing variables given the observed variables relative to f. Let $Z = \begin{pmatrix} Y \\ V \end{pmatrix}$ be a random variable with density f (Y represents the missing components of a particular column of the data matrix and V represents the observed components), and let

$$Z_i = \begin{pmatrix} Y_i \\ V_i \end{pmatrix}, \quad 1 \leq i \leq n$$

be a random variable with density f_i. We want to relate the conditional expectation

$$E(Y|V = v) \equiv h(v)$$

to the conditional expectations

$$E(Y_i|V_i = v) \equiv h_i(v)$$

To this end, let

$$f_V(v) \equiv \int f(y,v)dy$$

denote the marginal distribution of V and let

$$f_{V_i}(v) \qquad 1 \leq i \leq n$$

denote the marginal distribution of V_i. Then

$$h(v) = \int y \frac{f(y,v)}{f_V(v)} dy$$

But

$$f_V(v) = \sum \lambda_i f_{V_i}(v)$$

and

$$\int y\, f(y,v)dy = \sum\lambda_i \int y\, f_i(y,v)dy$$

$$= \sum\lambda_i f_{V_i}(v) \int \frac{y f_i(y,v)dy}{f_{V_i}(v)}$$

$$= \sum\lambda_i f_{V_i}(v) h_i(v)$$

Let

$$\eta_i = \frac{\lambda_i f_{V_i}(v)}{\sum\lambda_i f_{V_i}(v)}$$

Then

$$h(v) = \sum\eta_i h_i(v)$$

Thus $h(v)$ can be obtained from the $h_i(v)$ and the marginals $f_{V_i}(v)$.

APPENDIX G
SPRING WHEAT/BARLEY LABELER

G.1 BACKGROUND

The problem of separating spring wheat from other spring small grains, of which barley if by far the most prevalent, has emerged as one of the key problems in obtaining accurate estimates of spring wheat production [1]. In response to that need, research was begun at ERIM to devise a spring wheat/barley labeling technique. This work resulted in a first generation machine labeler described in Section G.2 below and in Reference 3. Evaluation of the deficiencies found in the first generation led to an approach and methodology for a refined labeler based upon a more sophisticated utilization of agronomic understanding and Landsat information content. Progress achieved in this refinement is presented in subsequent sections of this Appendix.

G.2 FIRST-GENERATION LABELER

G.2.1 DESCRIPTION

The machine technique assumes prior analyst-interpreter involvement of two types: labeling of spring small grains targets and determination of suitability of the segment for the machine technique. A segment's suitability is determined by the number and timing of acquisitions in the growing season for spring small grains, as will be described later.

Data passed to the machine are assumed to be normalized by a series of preprocessing steps comprised of satellite calibration correction, cosine sun angle correction, flagging of bad data, clouds, etc. [36], and haze correction [37].

Once targets have been identified as spring small grains and acquisition histories have been checked, the machine carries out the the spring wheat/barley separation in two steps: estimation of crop calendar shift and assignment of labels.

1. Estimation of Crop Calendar Shift: The first step in the machine process involves a technique by which signal variability due to differences in stage of development (as expressed in spectral appearance) is minimized. This technique, described in Reference 3 , also allows more accurate selection of critical acquisitions for the label assignment step.

Crop calendar shift estimation utilizes a mathematical representation (profile) of Tasseled-Cap Greenness development with a fixed time axis. Data for a particular target are shifted along the time axis until the best fit of data to the reference profile is obtained. The fixed time axis is then used as the new time axis for the target observations, and the difference between the original acquisition day axis and this fixed axis is the estimated crop calendar shift. The process requires that at least three acquisitions, 18 or more days apart, be available in the course of the spring small grains development cycle (approximately 90 days in length).

2. Label Assignment: Those targets for which a crop calendar shift were successfully estimated are passed on to the second phase of the process--label assignment. The labeling logic is designed to exploit observed differences in the temporal-spectral development of spring wheat and barley in Tasseled-Cap Brightness-Greenness space. Figures G-1 and G-2 illustrate examples of spectral trajectories for the two crops. As the crops green up, they move along a "green arm"

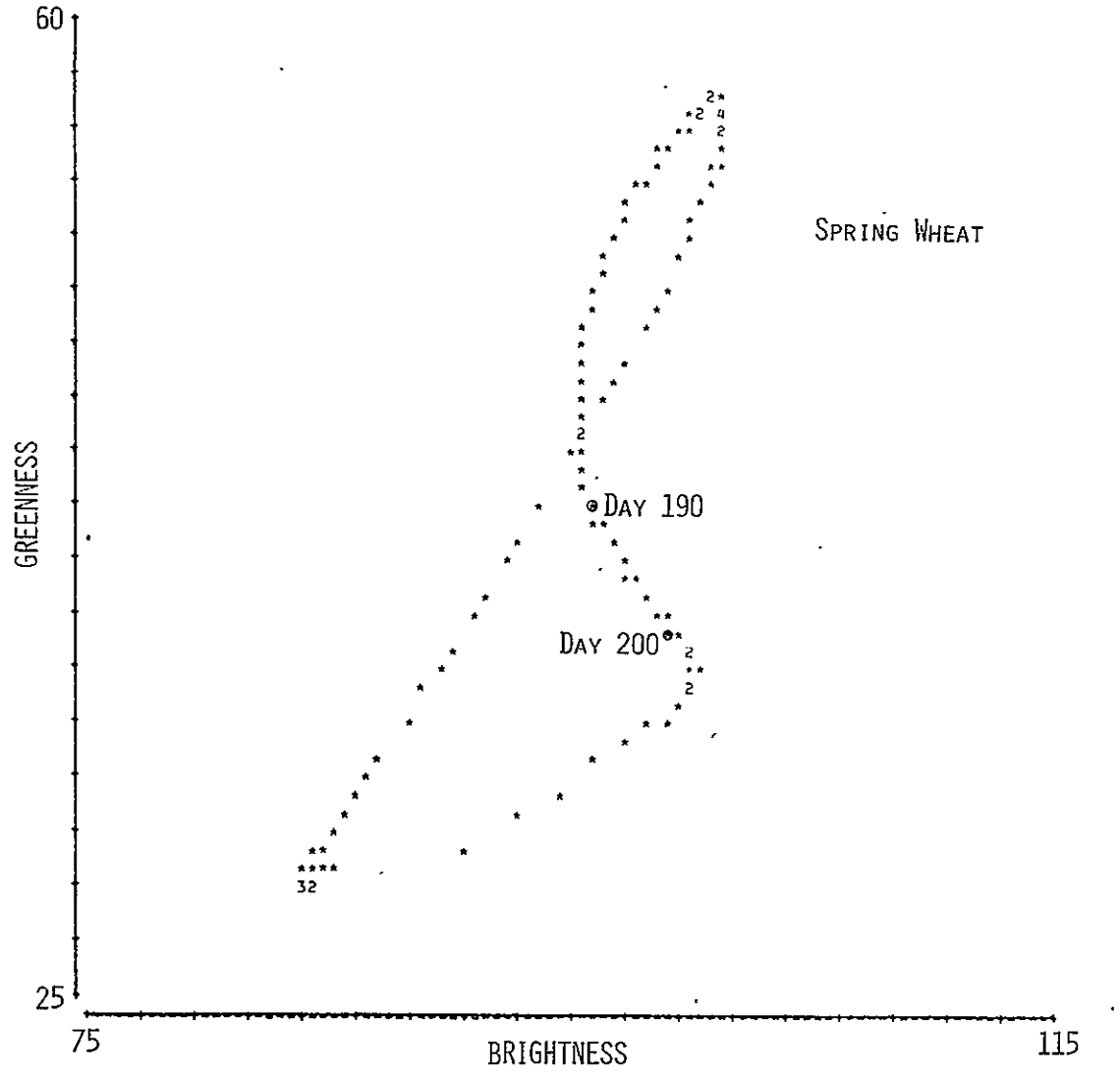


FIGURE G-1. EXAMPLE SPRING WHEAT TRAJECTORY - SMOOTHED CURVE OF MEANS

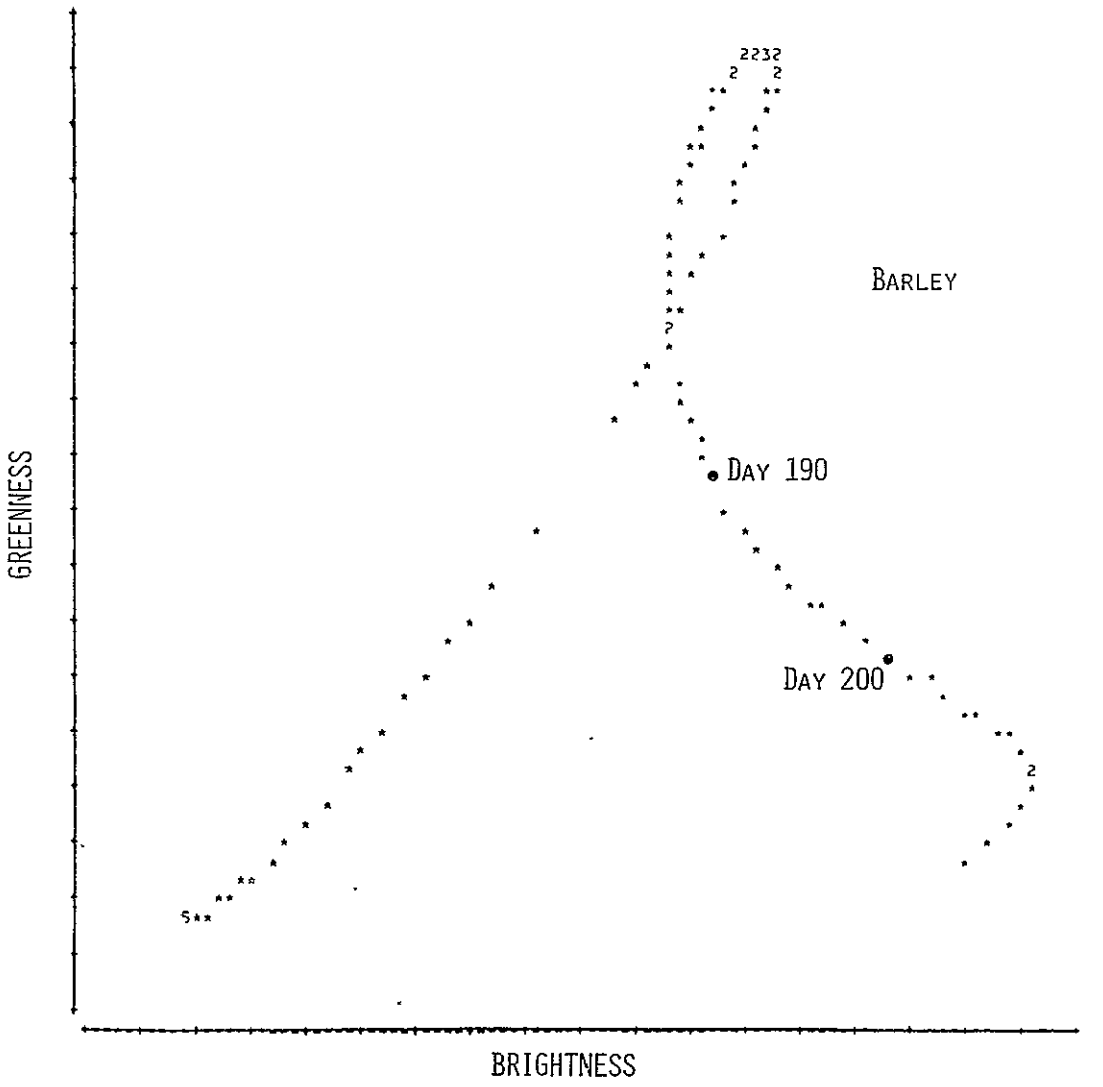


FIGURE G-2. EXAMPLE BARLEY TRAJECTORY - SMOOTHED CURVE OF MEANS

to some maximum Greenness. Then they typically loop back toward the Greenness axis and show a reduction in Greenness accompanied by increasing Brightness. This movement away from the green arm corresponds to the ripening of the crops. It is during the course of ripening, and particularly around the dough stage of wheat development (around 11.2 on the modified Feekes scale), that barley tends to be found farther from the green arm than spring wheat.

A reference line was defined which is perpendicular to the observed path of the two crops away from the green arm, and the distance from this reference line is used to discriminate between the two crops. Reference 3 gives further details. In order to be labeled, targets must have an acquisition in an 18-day range between shifted days 186 and 203.

G.2.2 TEST RESULTS

The first generation machine labeler was tested on 28 LACIE blind sites from the Northern Great Plains spanning two crop years (Phases 2 and 3). The results, described in detail previously [3], indicated that the technique worked well in the area encompassing the segments used for development of the logic, but poorly in areas separated by some distance from those segments. This stratification is illustrated in Figure G-3.

Two phenomena were consistently observed in the segments for which poor results were obtained. First, the errors were always errors of omission for spring wheat, i.e., most of the spring wheat targets were mislabeled as barley. In fact, both crops had distance measures significantly greater than expected on any given day in the time period of expected separability. Second, the segments all fell in areas likely to experience drought conditions.

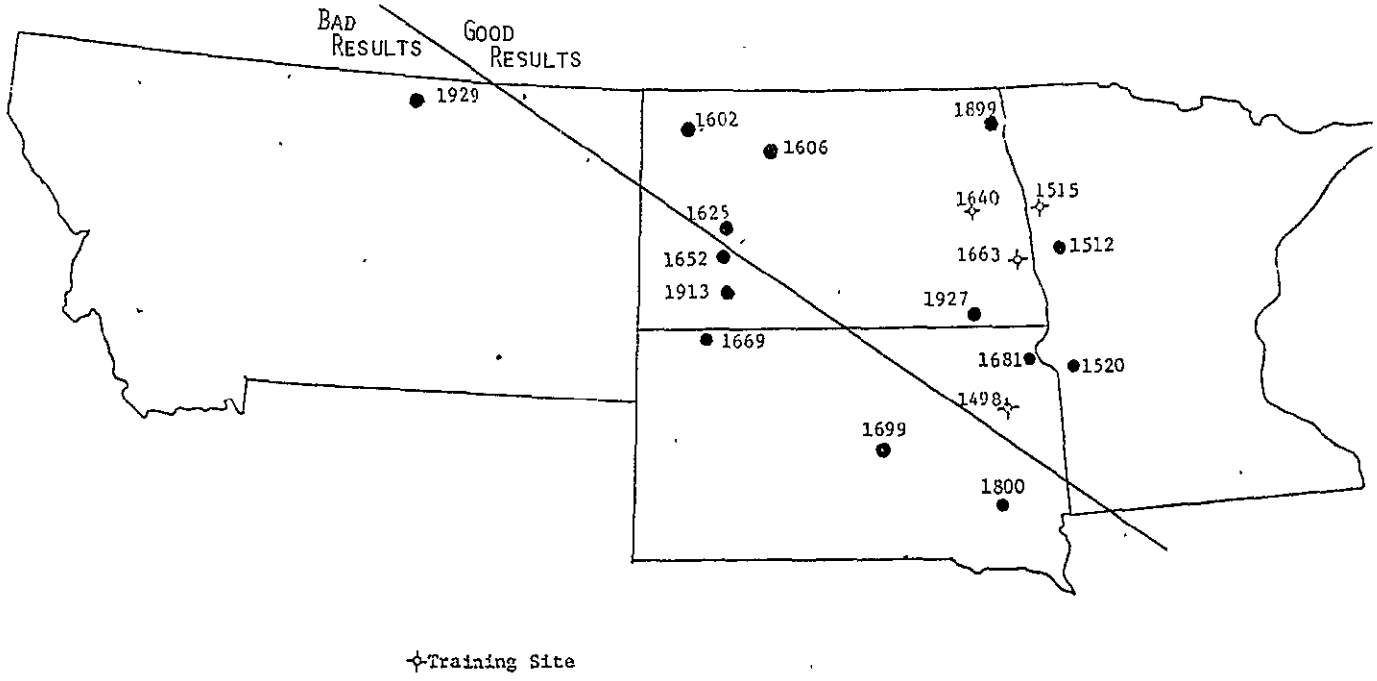


FIGURE G-3. ORIGINAL LABELER TEST RESULTS - DECEMBER 1978

As a result of these tests, an effort was undertaken to revise the labeler so as to increase its geographic range of effectiveness.

G.3 APPROACH TO REFINEMENT

G.3.1 PHILOSOPHY

Although the entities to which labels are assigned are picture elements or groups of picture elements, the actual entities of interest are the populations of plants whose spectral properties are represented in the pixels. If any labeling technique is to be successful, it must be the case that the primary influences on pixel spectral values are the populations of plants themselves. When external effects such as sun angle and haze are eliminated or minimized, we must assume that biological factors are the main spectral drivers. Given this assumption, it is clear that labeling techniques must be firmly grounded in biological and physical phenomena, in the behavior of and influences on the plant populations themselves.

G.3.2 STEPS

In order to insure that the results of the labeling technique revision were biologically sound, a sequence of steps was followed, starting with the plants themselves and working out towards the sensor.

1. Develop Hypotheses: In order to provide an initial focus for the research, a set of hypotheses must be formulated to answer the question of interest. These represent not a set of absolute constraints, but rather a set of potentially useful directions to pursue.

2. Identify Pertinent Physiological Relationships, Effects: Using results of previous agronomic research, crop characteristics or responses to given conditions which are likely to affect spectral response can be identified. These characteristics and responses on the level of individual plants or small groups of plants are the driving factors in determining the spectral characteristics of the plant population as represented in the picture elements.

3. Model Canopy Reflectance: Utilization of canopy reflectance models allows the spectral effects of relevant physiological or environmental changes to be predicted and studied. Models allow a stricter control, and at the same time a broader range of conditions than might be available in field data.

4. Analyze Field Measurements: Physiological relationships and their spectral effects can be verified through the use of field reflectance data. These data provide the crucial link between simulation and the real world, and also between the ground and the satellite. Field data also aid in understanding the range of natural variability of spectral appearance and influential biological and physical factors.

5. Analyze Landsat Data: Only after the previous steps have been completed does satellite data enter into the process. At this point, hypotheses have been supported or refuted at several levels, and the results expected in the Landsat data are already well understood. Thus, this step is a means of confirming expectations rather than one of finding a way to do the task at hand. The way has already been defined. The Landsat data will only confirm its reliability or indicate a need to return to an earlier stage in the process for additional work.

G.4 RESULTS OF REFINEMENT

G.4.1 HYPOTHESES

Based on the distinct geographic separation of "good" and "bad" results (Figure G-3), and the correlation of "bad" results with high drought-susceptibility areas, it was likely that moisture stress was altering crop signatures and thereby causing spring wheat to look like barley. The geographical separation could also suggest differences in soil brightness. One of these factors, or a combination of them, were thought likely to be responsible for the labeling problems encountered.

G.4.2 PHYSIOLOGICAL RELATIONSHIPS AND EFFECTS

Since soil brightness per se has no physiological effect on the plants, and associated soil properties would primarily be reflected in moisture availability, only moisture stress effects were studied in this stage of the labeler revision effort.

The clear indication from a fairly wide range of research related to the impact of inadequate moisture on small grains is that effects vary considerably as a function of the stage of crop development at the time of stress initiation and the duration of the stress. In areas such as those where poor labeler results were obtained, prolonged moisture stress is likely. The effects of prolonged stress on small grains plants may include reduction in plant height and number of tillers per plant, thinner and smaller leaves which may roll or wilt, particularly in the middle portion of the day, reduction in the number of plants per acre, and increased rate of plant development. These effects stem from two major factors: a need to reduce surface area from which moisture may be lost through evapotranspiration and reduced survival rates for the plant populations.

G.4.3 REFLECTANCE MODELING

Approach: Reflectance modeling results were obtained using a three-layer version of the Suits' canopy reflectance model [21,59]. A set of "normal" parameters based in field measurements of wheat plants were modified to simulate the physiological effects of moisture stress. These modifications included changes in leaf orientation, leaf color, plant height and density, and relative depths and densities of canopy layers. Parameters were determined for seven stages of development for both the normal and stressed conditions, as indicated in Table G-1. Both canopies at all stages of development were combined with three different soil spectra corresponding to dark, medium, and bright soils [60]. Landsat band reflectances were computed utilizing sensor spectral response functions, and these reflectance data were then transformed into a Tasseled-Cap-like projection.

Results: The trends apparent in the model data indicate significant and distinguishable effects of moisture stress and soil Brightness on canopy reflectance. As illustrated for normal canopies in Figures G-4 and G-5, soil Brightness exerts little if any significant influence on Greenness, but considerable influence, as expected, on Tasseled-Cap Brightness. The effect on Brightness is most pronounced at the tails of the profile, where canopy closure is at a minimum, resulting in maximum soil exposure. Indeed, soil effects are all but eliminated when canopy closure is at a maximum.

Figure G-6 illustrates for the normal canopy the distance measure used in labeling, plotted over time. The descending portion corresponds to a change from bare soil to vegetated soil, while the flat middle portion corresponds to the greening-up of the crop (movement along the green arm). The later ascending portion represents the ripening of the crop, and includes the time period in which spring wheat and barley are most distinguishable (Section G.2.1).

TABLE G-1. STAGES OF DEVELOPMENT MODELED IN CANOPY REFLECTANCE ANALYSIS

Stage #	Name	Approx. Day of Year	Characteristics
1	Emergent	131	Detectable Greenness
2	Jointing	158	Stem elongation occurring, considerable green vegetation
3	Boot	167	Just prior to heading, increasing component of mature green parts
4	Post-head	185	Heads present, mature green component dominant, lower leaves dead, entire plant not yet turning
5	Senescing	194	Entire plant turning yellow/brown
6	Ripe	203	No green matter remaining, plants dead
7	Harvested	221	Stubble only

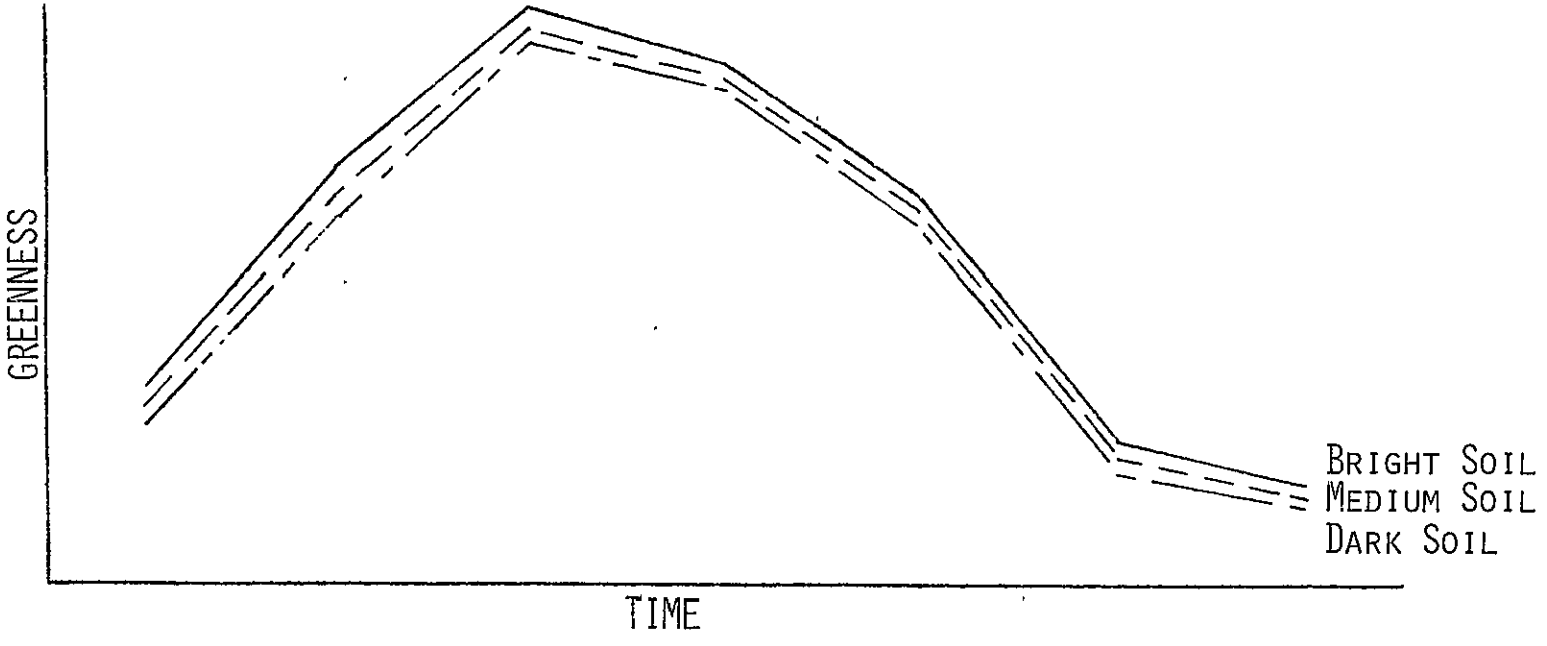


FIGURE G-4. CANOPY REFLECTANCE MODELING RESULTS, SOIL BRIGHTNESS EFFECTS ON GREENNESS - NORMAL CANOPY

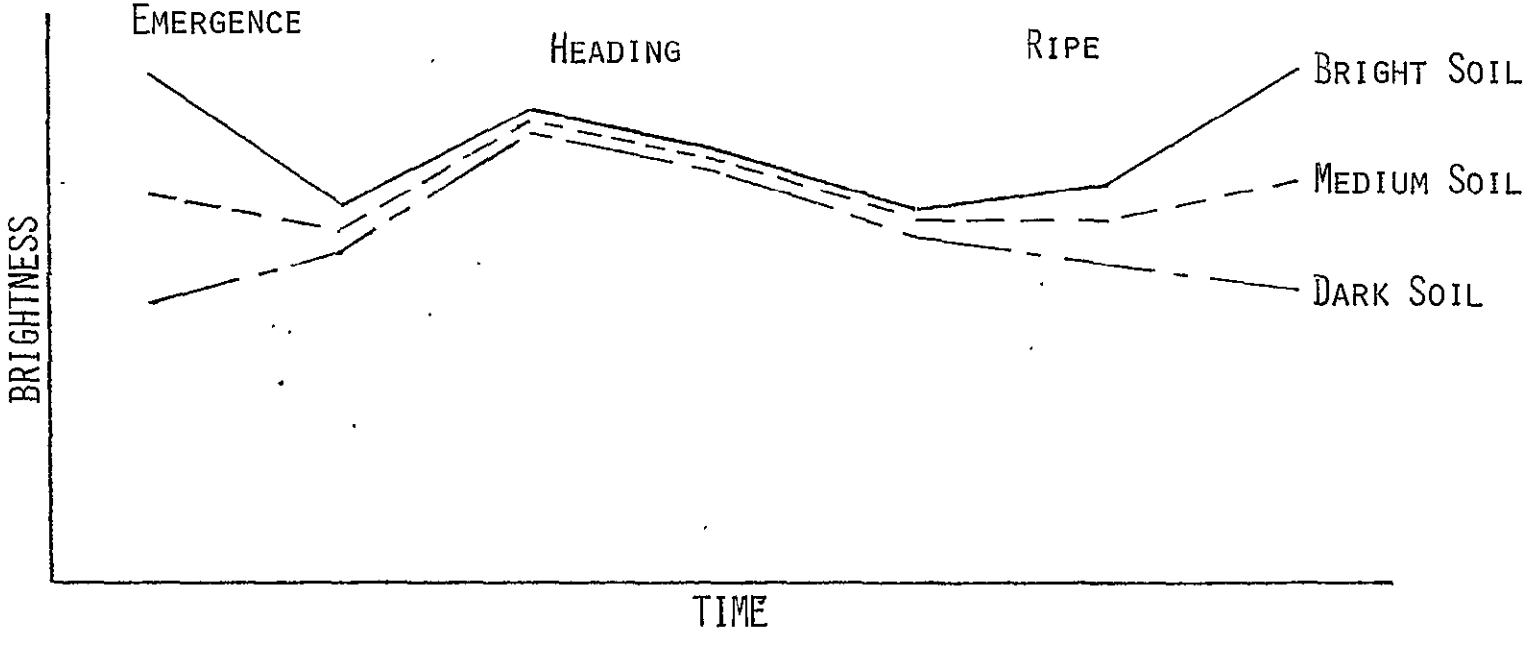


FIGURE G-5. CANOPY REFLECTANCE MODELING RESULTS, SOIL BRIGHTNESS EFFECTS ON BRIGHTNESS - NORMAL CANOPY

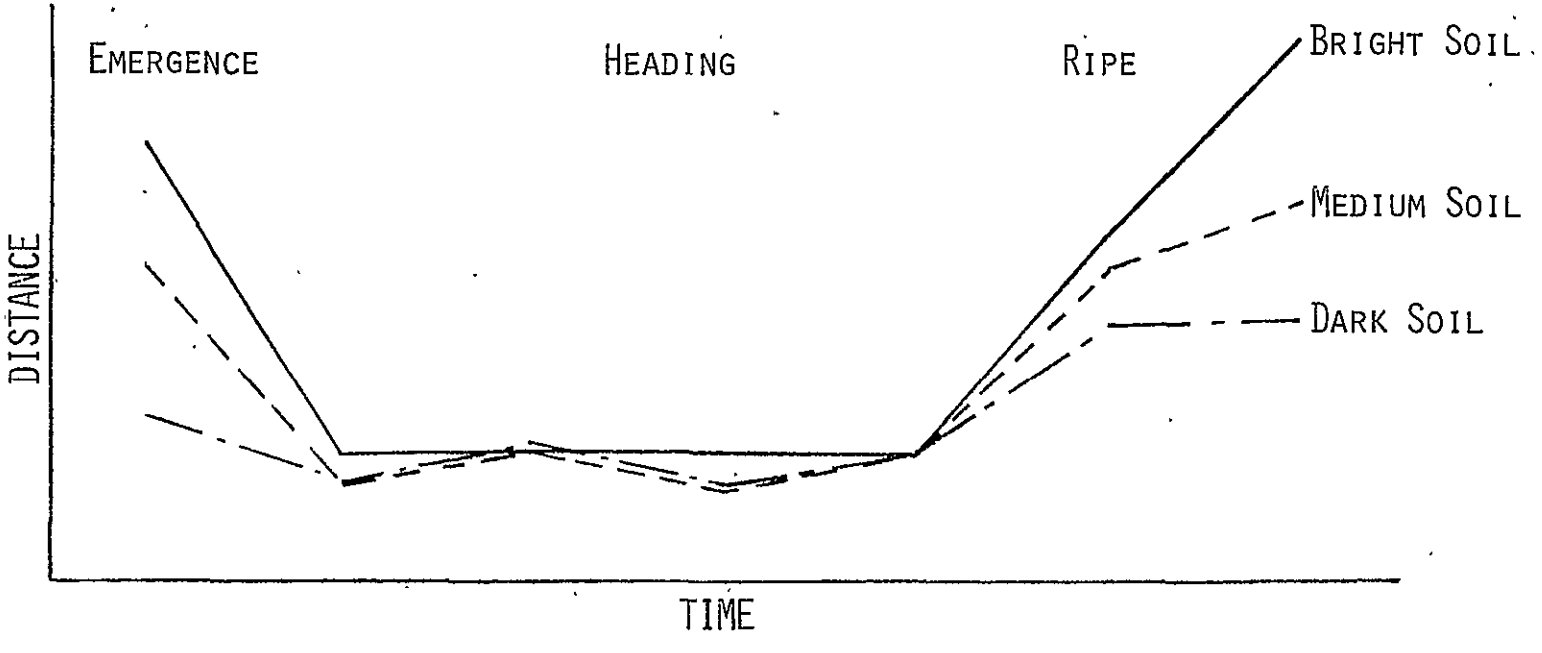


FIGURE G-6. CANOPY REFLECTANCE MODELING RESULTS, SOIL BRIGHTNESS EFFECTS ON DISTANCE MEASURE - NORMAL CANOPY

Since the distance measure is a linear combination of Greenness and Brightness, changes in those two components as a result of soil brightness variation or any other factor will be reflected in changes in the distance profile. Thus, soil brightness effects on the distance measure are similar to those on Brightness, occurring primarily at the times when canopy closure is low. The primary effects of bright soil seem to be an increased slope in the non-level portions of the distance profile and a positive offset of all distance values.

While Greenness is little affected by soil brightness, it is significantly impacted by moisture stress, as illustrated in Figure G-7. A considerable effect was also seen in Brightness, as demonstrated by comparison of Figure G-8 to Figure G-5; but here the effect is primarily an enhancement of soil influences resulting from reduced canopy closure. As a result, the impact of moisture stress on Brightness cannot be predicted without knowledge of the soil brightness.

Again, the effect of moisture stress on the distance profile is a combination of Greenness and Brightness effects, as illustrated by comparison of Figure G-9 to Figure G-6. As in the Brightness profile, soil brightness effects are enhanced with the more open stressed canopy. The overall reduction in Greenness values is translated into a general increase in the distance measure, and a reduction or near elimination of the level middle portion of the profile. Of greatest probable significance to labeling is the fact that the ascending portion of the distance profile, the portion containing the time period of maximum spring wheat/barley separability, comes sooner than it does for the normal canopy.

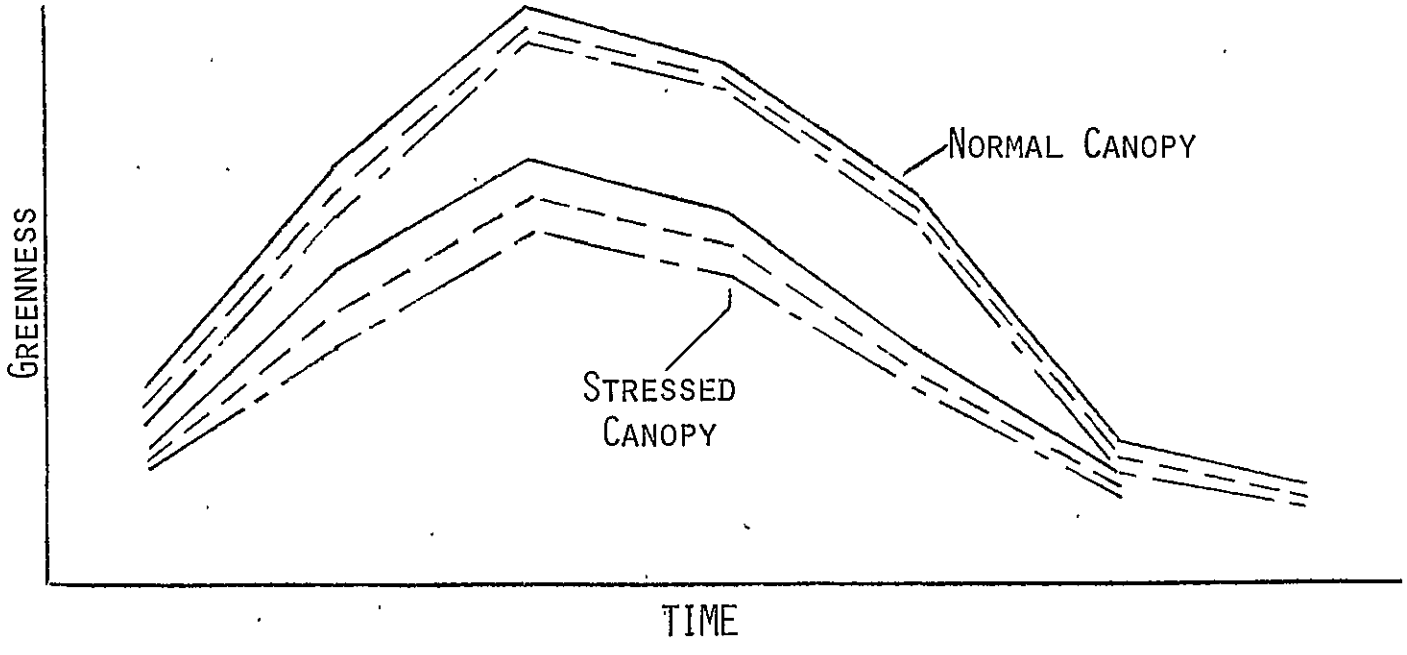


FIGURE G-7. CANOPY REFLECTANCE MODELING RESULTS,
MOISTURE STRESS EFFECTS ON GREENNESS

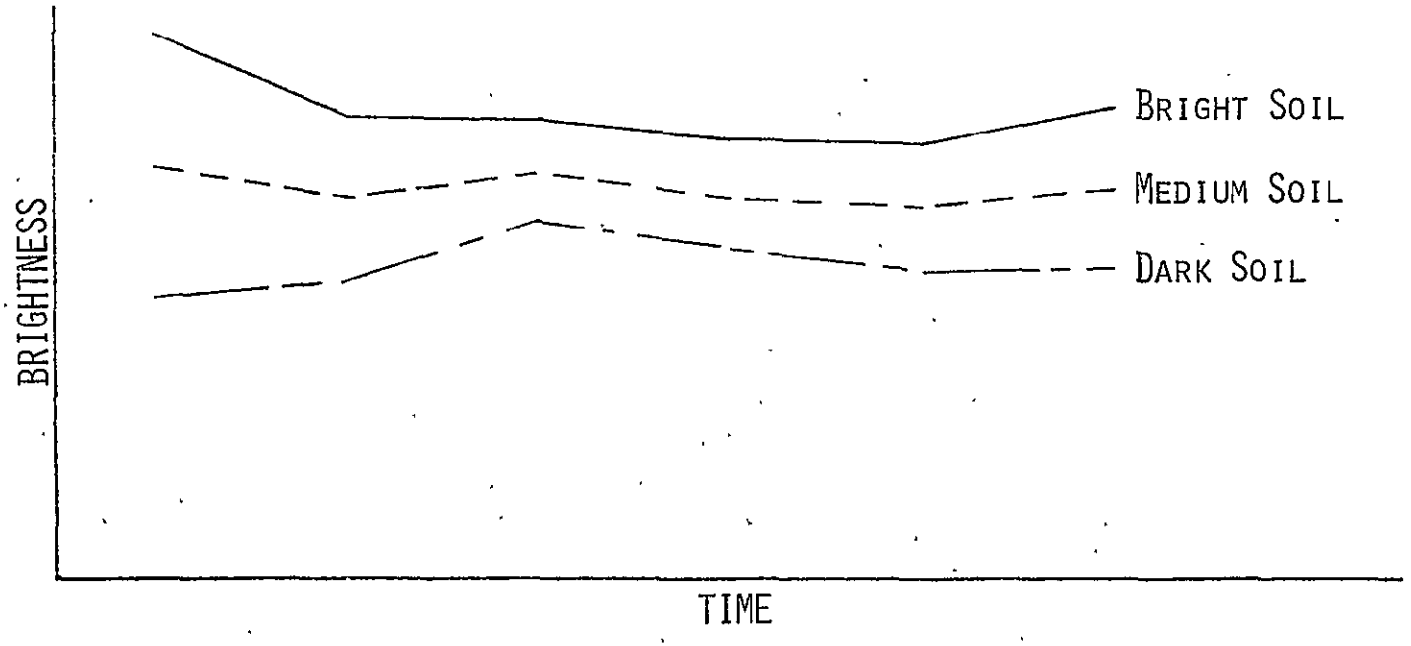


FIGURE G-8. CANOPY REFLECTANCE MODELING RESULTS, SOIL BRIGHTNESS EFFECTS ON BRIGHTNESS - STRESSED CANOPY

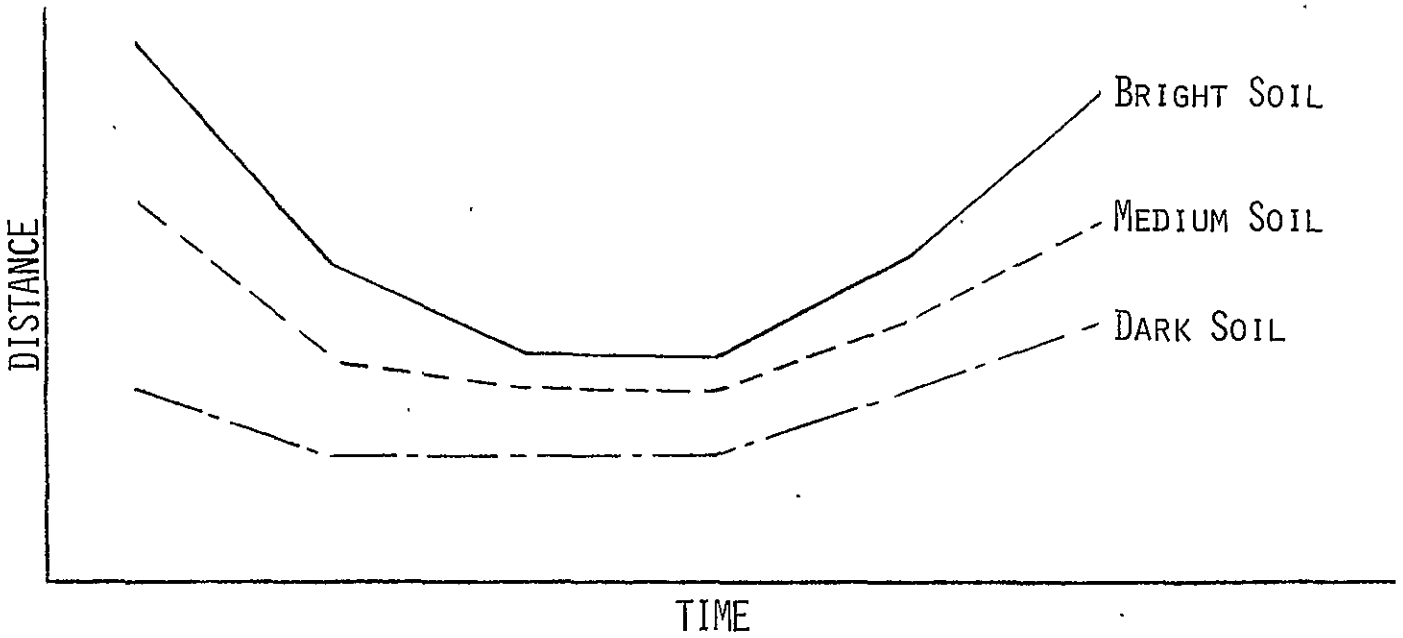


FIGURE G-9. CANOPY REFLECTANCE MODELING RESULTS, SOIL BRIGHTNESS EFFECTS ON DISTANCE MEASURE - STRESSED CANOPY

Summary: The model data thus suggest that both moisture stress and soil brightness exert detectable influences on canopy reflectance and, perhaps most importantly, on the distance profile in the time period of separability. Neither hypothesis can be rejected as a possible cause for the experimental results obtained previously.

Finally, the two factors should be distinguishable and detectable by separate processes. Moisture stress should be expressed in a reduced Greenness peak, while soil brightness should be expressed in early season Brightness values.

G.4.4 FIELD MEASUREMENTS

In this study, field measurements were used primarily in a support role. Field data were used to provide normal canopy inputs for the reflectance model. In addition, on-site observations of field experiments being conducted under the direction of Dr. Ray Jackson at the U.S. Water Conservation Laboratory, Phoenix, Arizona, served to substantiate the physiological effects of moisture stress reported by agronomic researchers. These on-site observations thus aided in the understanding of the physiological phenomena involved, and their spectral impacts.

G.4.5 LANDSAT DATA ANALYSIS

Data: The data used in verifying the expected impact of moisture stress and soil brightness variations and revising the labeler to accommodate those impacts consisted of observations of small grains targets from seven LACIE blind sites, a subset of the original test set (see Section 1.2). Table G-2 provides additional information on the segments, and Figure G-10 shows their locations.

TABLE G-2. SEGMENTS USED IN LANDSAT DATA ANALYSIS

Segment #	Location
1498	Codington County, South Dakota
1515	Norman County, Minnesota
1640	Barnes County, North Dakota
1663	Richland County, North Dakota
1669	Perkins County, South Dakota
1800	McCook County, South Dakota
1929	Blaine County, Montana

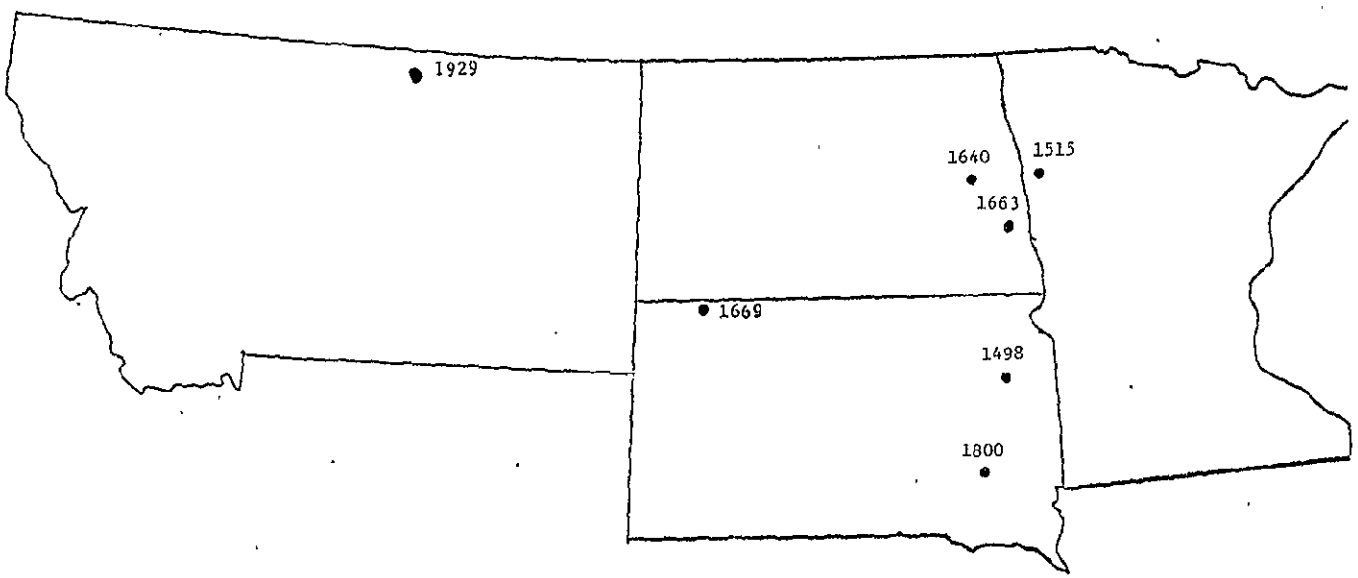


FIGURE G-10. SEGMENTS USED IN LANDSAT DATA ANALYSIS

Expected Results: Based on the previous steps in the research process, the following relationships were expected to be observed in the Landsat data.

Moisture stress of a prolonged nature should be detectable in a reduced peak in the Greenness profile for small grains in a segment. As a result of the impact of moisture stress on the distance profile, the time period during which spring wheat and barley are distinguishable should occur earlier than normal.

Soil brightness variations should be manifested in early season Tasseled-Cap brightness values (i.e., before significant growth of green vegetation has occurred). An increase in the slope of the distance profile during the time period of separability should also be observed and as a result a more steeply-sloped decision line would be expected. In addition, moisture stress and bright soils should interact by enhancing the soil effects, resulting in higher distance values during the time period of separability.

Test Results: Two iterations of data analysis were carried out, intended to provide a rough-cut revision of the original labeling logic, followed by a more precise revision.

The first iteration utilized data from six segments. Crop calendar shift was estimated using a two-step refinement of the original technique. After a shift estimate was obtained using a standard profile, a new profile was computed using all small grains pixels (shifted) in the segment. This segment-specific profile was then used to refine the first shift. While the difference between data shifted with the segment-specific profile and data shifted with the standard profile is far less dramatic than that between shifted and unshifted data, the segment-specific profile shift did tend to produce smoother distributions over time. In addition, fitting a profile to segment data provided a means of obtaining a segment-level peak Greenness estimate.

The apparent time period for separability of wheat from barley (using the Greenness-Brightness distance) was determined through a combination of quantitative and qualitative methods. Optimum linear discriminants were computed for each 3-day interval from peak Greenness through a time at or near harvest. These results were then combined with visual evaluation of the distance measure plotted against shifted day of year, and used to produce estimates of the day range of separability. Optimum linear discriminants were again computed, this time for the defined day range as well as for an 18-day interval comprising a subset or expansion of the defined range.

Figures G-11 through G-14 illustrate the results of this first iteration, and indicate support for the hypothesized relationships. Segments with low green peaks (1669 and 1929) have earlier day ranges of separability. These same segments also had very bright soil, and this was reflected both in the starting distance value for the linear discriminant and the slope of the decision line.

Based on the sparse data available, a preliminary labeling logic revision was devised, and tested on the six training segments. Table G-3 illustrates the labeling steps used, while Table G-4 and Figure G-15 show the results. Noteable in the results is the fact that labeling performance on problem segments was distinctly improved, while labeling performance on good segments was not degraded.

The second iteration of Landsat data analysis was intended to bring in a significantly larger data base from which relationships could be more precisely defined. Data processing and availability problems precluded this approach, however, and only one additional segment was added.

Nevertheless, the data from the seven segments were reprocessed, this time using the new profile model (Appendix A) in the segment-specific shifting procedure. In addition, linear estimates of the

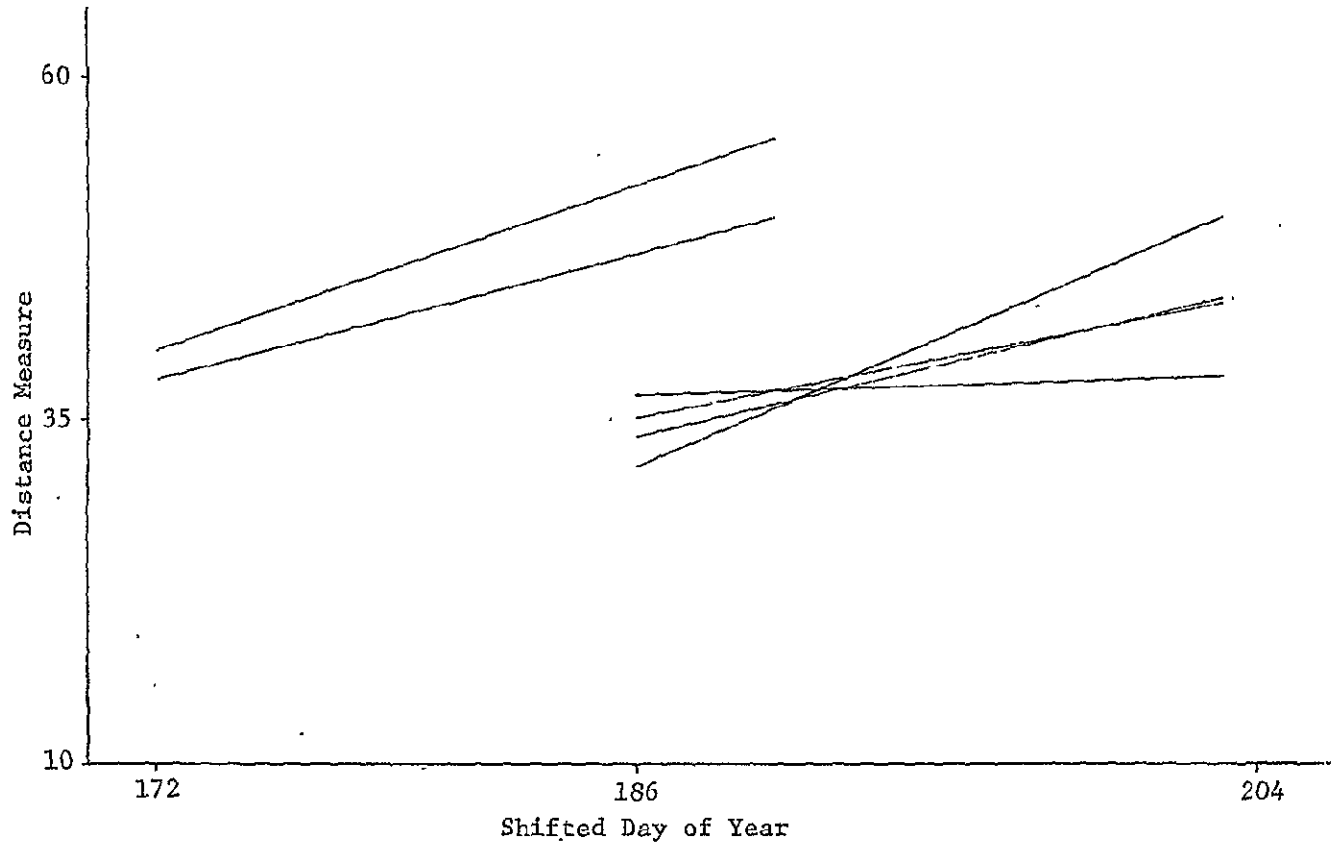


FIGURE G-11. OPTIMUM DECISION LINES - 1st ITERATION

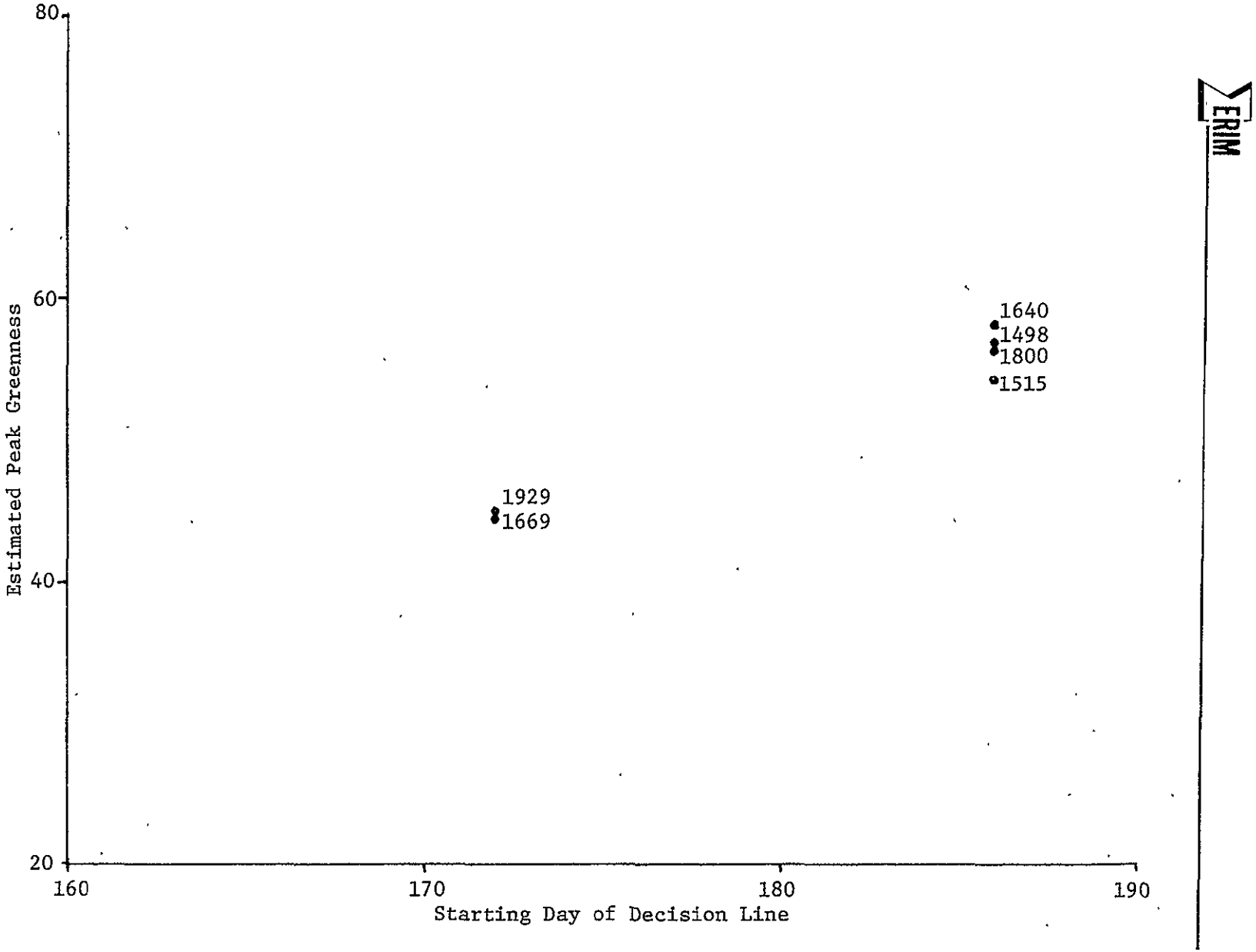


FIGURE G-12. CORRELATION OF ESTIMATED PEAK GREENNESS WITH FIRST DAY OF SEPARABILITY - 1st ITERATION

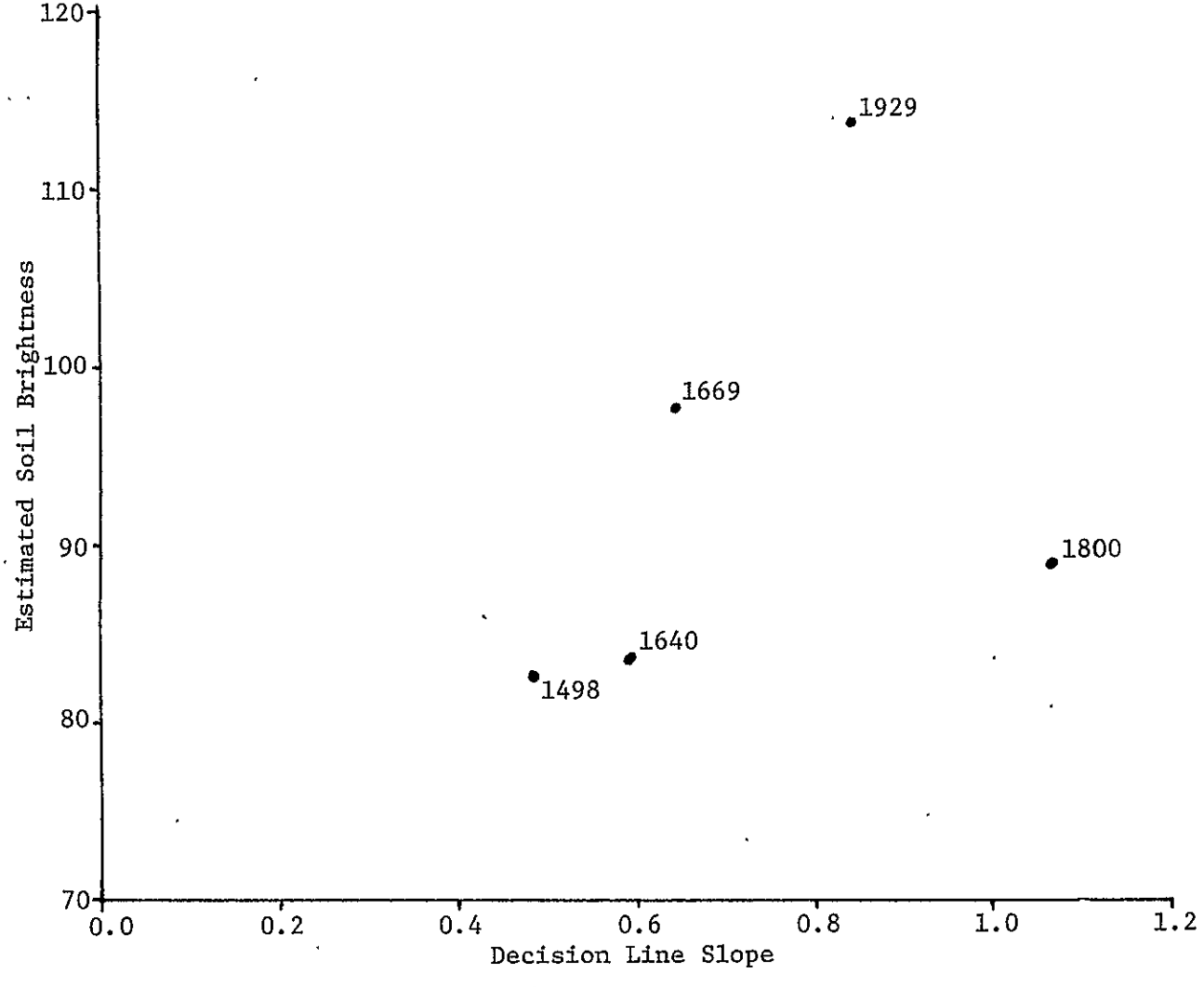


FIGURE G-13. CORRELATION OF ESTIMATED SOIL BRIGHTNESS WITH DECISION LINE SLOPE - 1st ITERATION

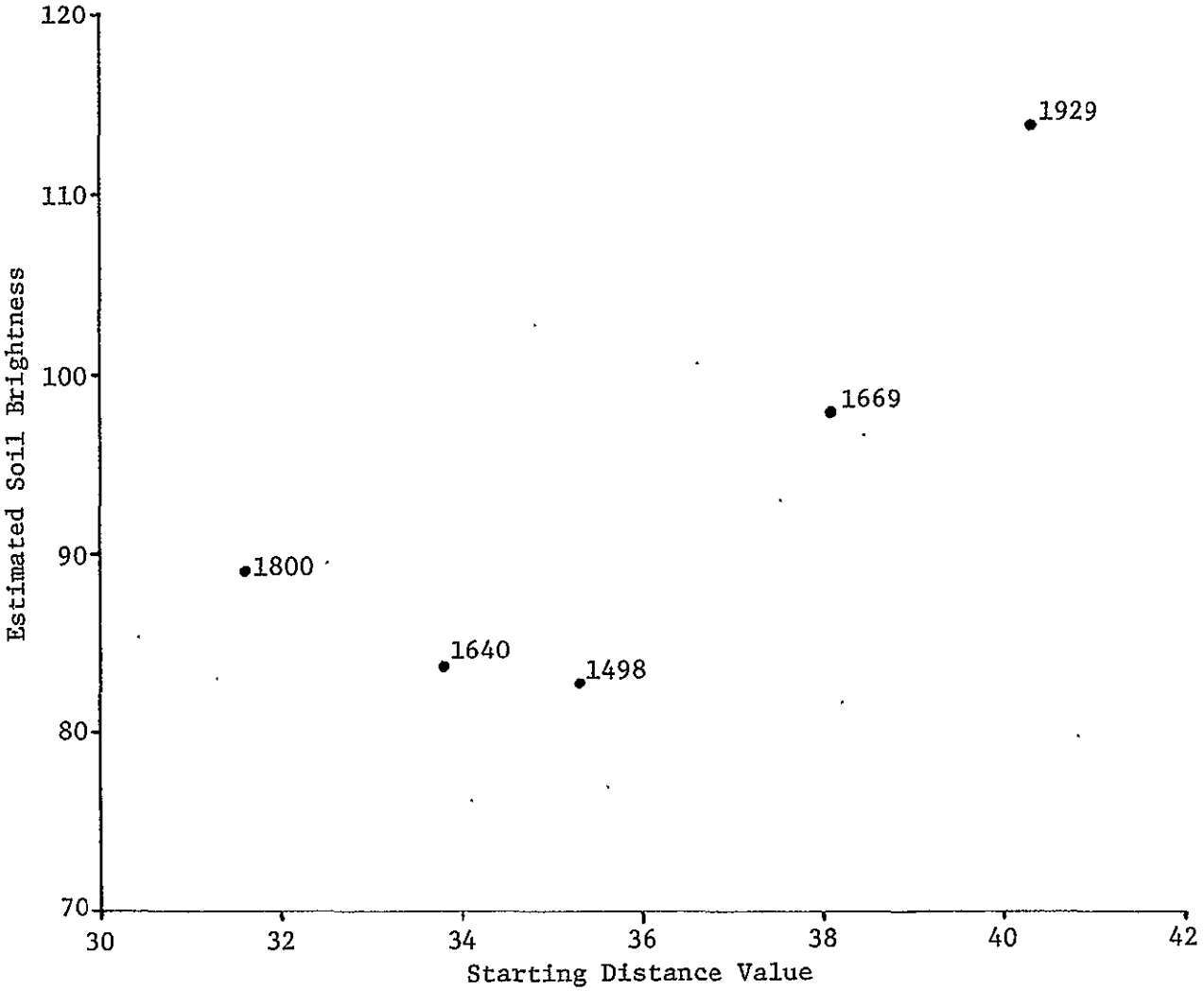


FIGURE G-14. CORRELATION OF ESTIMATED SOIL BRIGHTNESS WITH STARTING DISTANCE VALUE - 1st ITERATION

TABLE G-3. STEPS USED IN APPLICATION OF PRELIMINARY LABELING LOGIC

Data Set: Small grains pixels (identified from ground truth tapes). Data correction and normalization, and Tasseled-Cap transformation applied.

Step	Description
1	Divide segments into two bins based on estimated peak Greenness
2	Utilize linear regression estimate of soil brightness/decision line slope relationship to estimate decision line slope. Default slope is .53 (if no soil brightness data available)
3	Compute decision lines: low peak bin – starting day = 172 y-intercept = 37.25 - slope * 170.1 (indicates standard "pivot point" from which slope can be calculated) high peak bin – starting day = 186 y-intercept = 37.25 - slope * 191.1
4	Assign labels

TABLE G-4. DETAILED LABELING RESULTS OF LANDSAT DATA ANALYSIS - 1st ITERATION

Segment	Ground Truth	Original Labeling Procedure		Optimum Linear Discriminant		Revised Labeling Procedure (Prelim.)	
		SPW/OSMG	% Correct	SPW/OSMG	% Correct	SPW/OSMG	% Correct
1498	Spring Wheat	806/239	77.1	846/220	79.4	802/264	75.2
	Barley	114/310	73.1	134/283	67.9	116/301	72.2
1515	Spring Wheat	2597/449	85.3	2805/551	83.6	2795/561	83.3
	Barley	421/2033	82.8	737/2022	73.3	801/1958	71.0
1640	Spring Wheat	4593/971	82.5	4493/1077	80.7	4601/969	82.6
	Barley	909/1363	60.0	770/1497	66.0	822/1445	63.7
1669	Spring Wheat	11/384	2.8	231/51	81.9	242/40	85.8
	Barley	10/261	96.3	75/102	57.6	81/96	54.2
1800	Spring Wheat	41/34	54.7	55/20	73.3	44/31	58.7
	Barley	133/871	86.8	315/679	68.3	278/716	72.0
1929	Spring Wheat	4/1346	0.3	678/312	68.5	662/328	66.9
	Barley	1/545	99.8	147/232	61.2	139/240	63.3

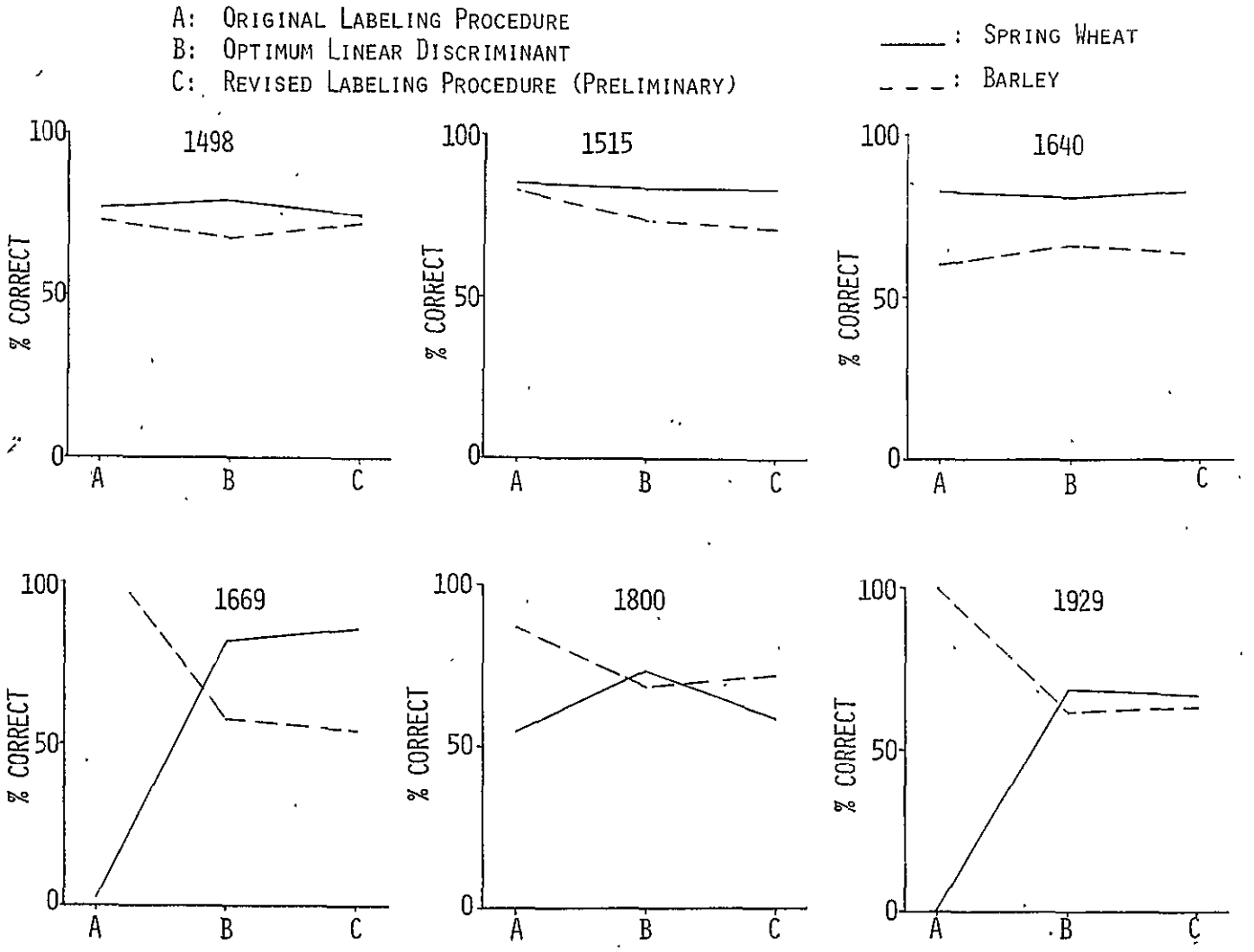


FIGURE G-15. RESULTS OF LABELER TEST ON TRAINING DATA - 1st ITERATION

new model parameters were obtained for each of the small grains pixels in the segments (about 33000 total). These pixel-specific profiles provided a means of eliminating pixels with distinctly non-small-grain-like trajectories (ground truth discrepancies, etc.) and so resulted in a purer data base of small grains observations. About 4000 pixels were eliminated by this process. In addition, the pixel-specific parameter estimates made possible an analysis, not yet completed, of pixel spectral behavior as a function of estimated Greenness peak.

As in the first iteration, day ranges of separability were defined for good small grains pixels, using quantitative and qualitative analysis techniques, and optimum linear discriminants were computed in those ranges and in corresponding 18-day intervals.

These decision lines are illustrated in Figure G-16, and show much the same trends as did the first set. The particular trends of interest are illustrated in Figures G-17 through G-19.

The one discrepancy between the two iterations involves the soil brightness/decision line slope relationship. While clear in the first iteration, this relationship is not apparent in the second. Small changes in slope had little impact on labeling accuracies in either case, suggesting that this particular element of the labeling logic is less important than, for example, the time period in which separability occurs. Nonetheless model results and our understanding of the physical processes involved suggest that the relationship should be a factor, and further work aimed at more conclusively supporting or rejecting the hypothesis is desirable.

G.4.6 PROCEDURAL SPECIFICATION OF REFINED LABELER

Although precise tuning of the labeling logic was not possible due to data problems, the procedural steps in the refined small grains labeler have been defined, at least in a preliminary context. The

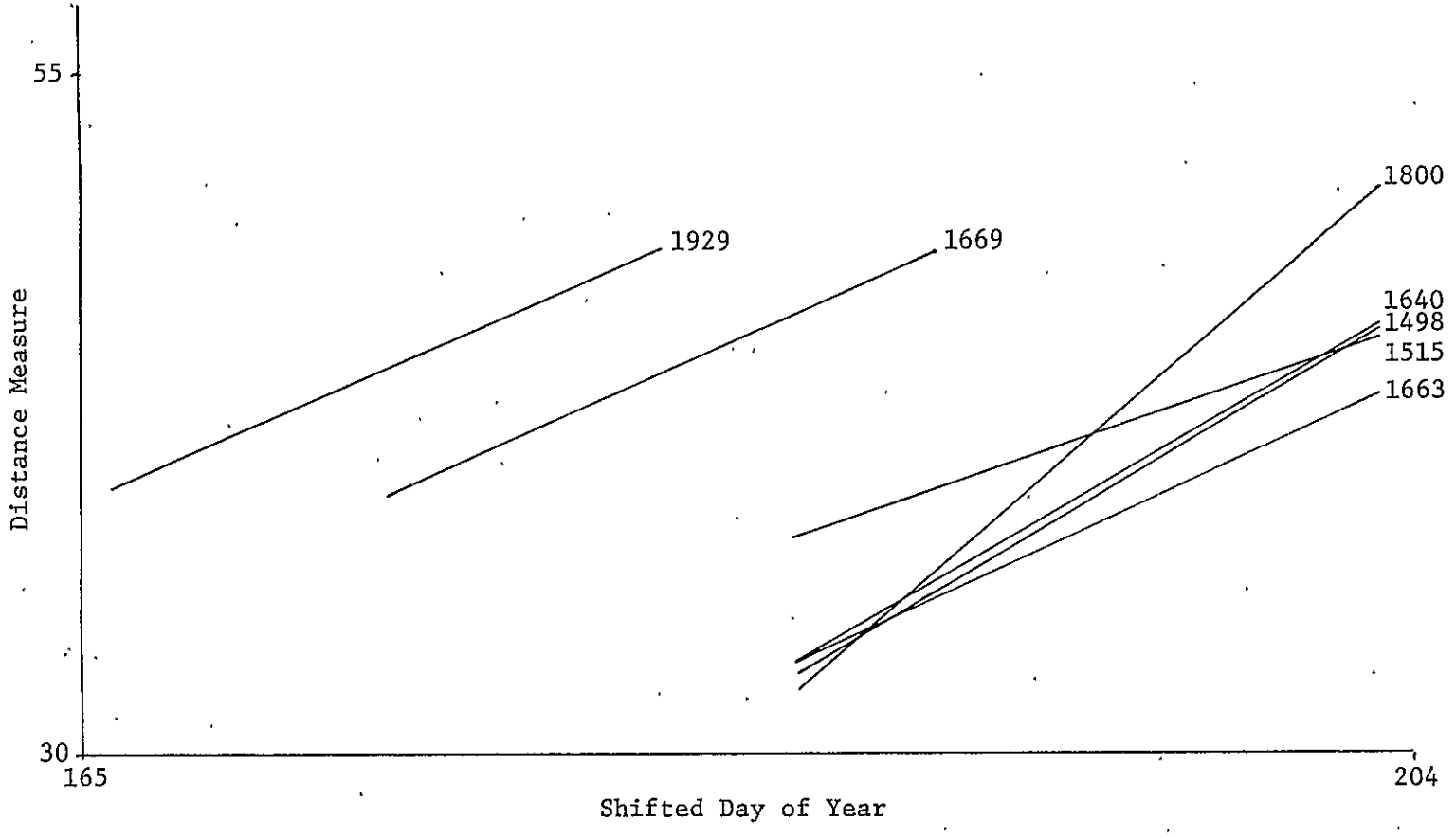


FIGURE G-16. OPTIMUM DECISION LINES - 2nd ITERATION

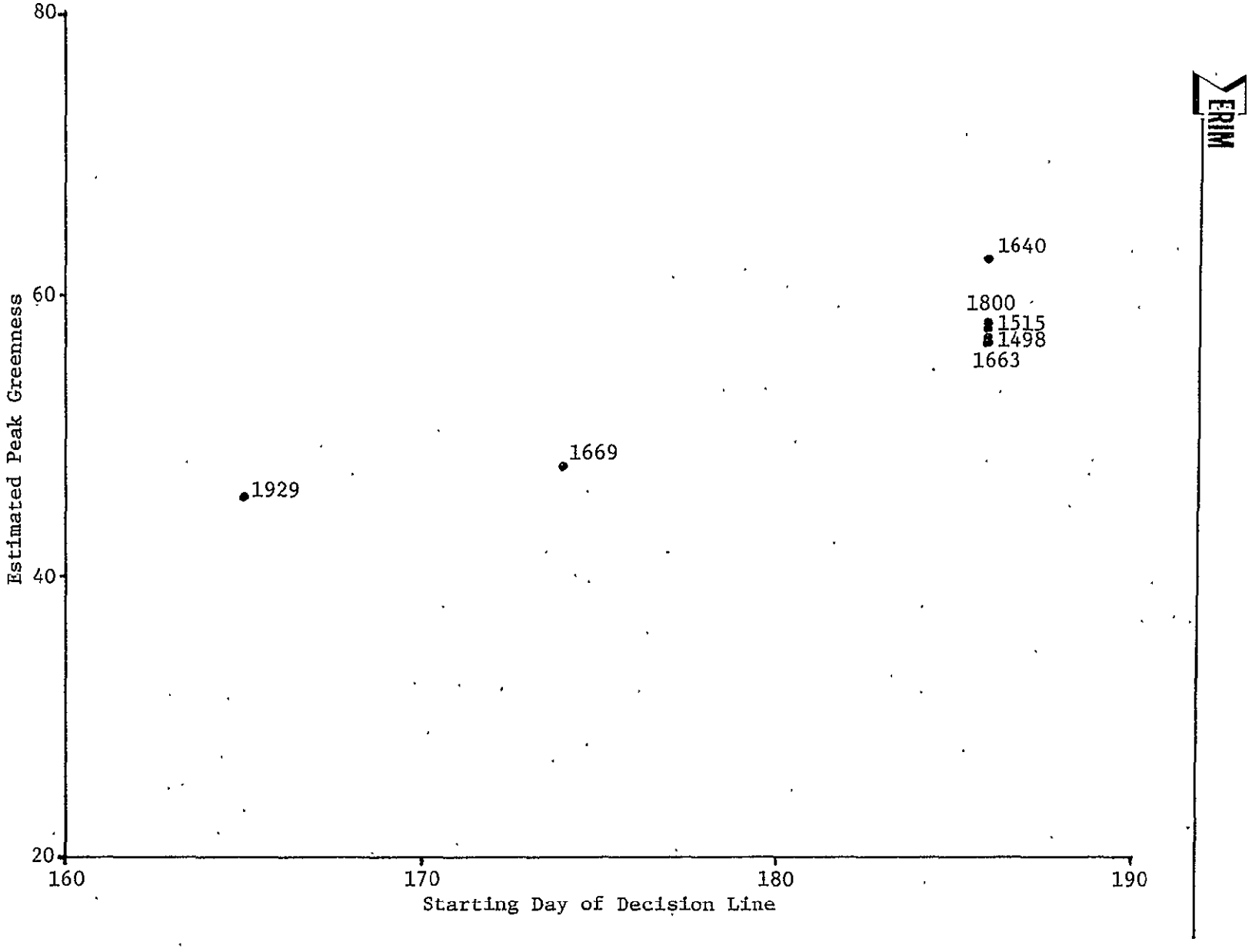


FIGURE G-17. CORRELATION OF ESTIMATED PEAK GREENNESS WITH FIRST DAY OF SEPARABILITY - 2nd ITERATION

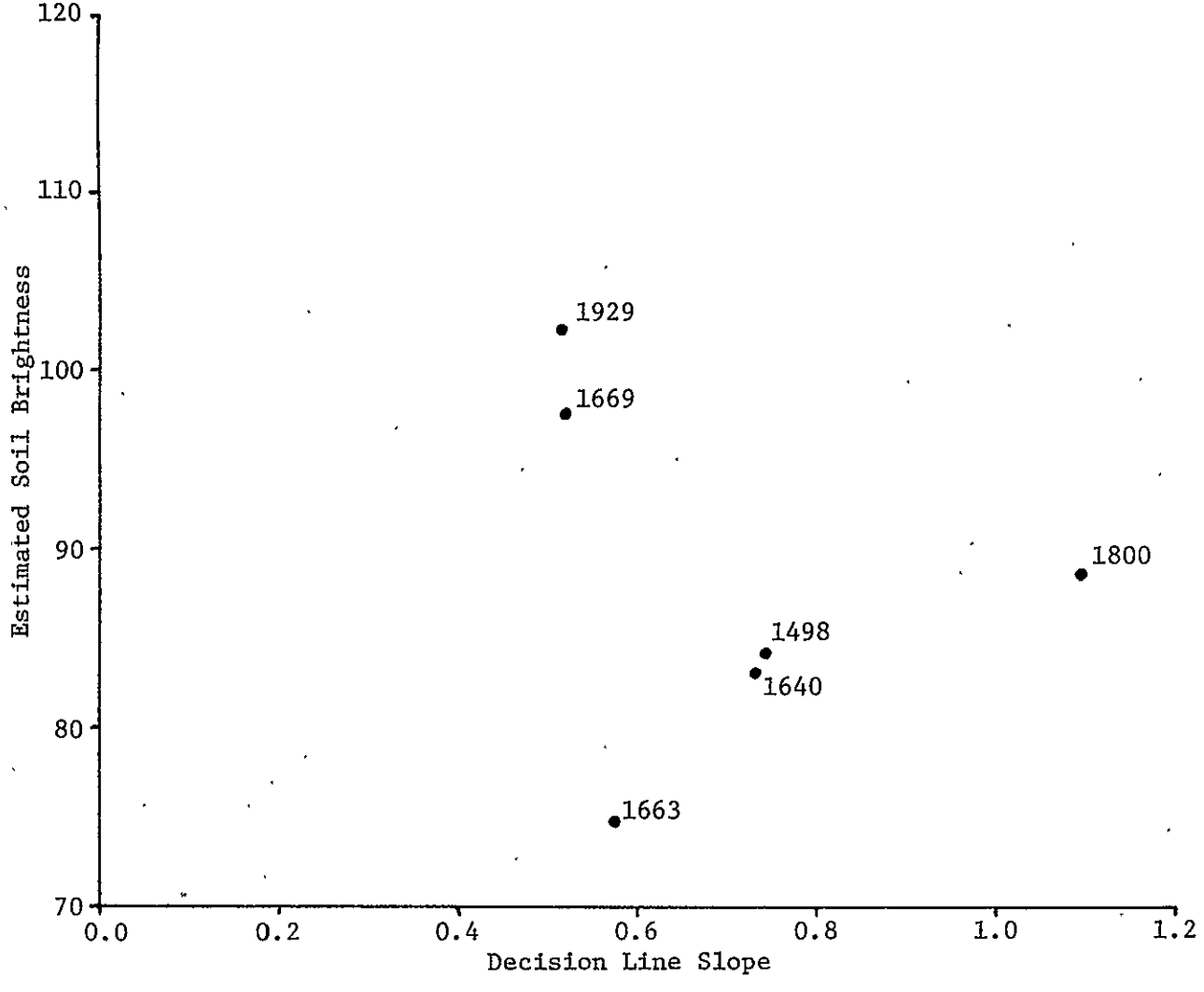


FIGURE G-18. CORRELATION OF ESTIMATED SOIL BRIGHTNESS WITH DECISION LINE SLOPE - 2nd ITERATION

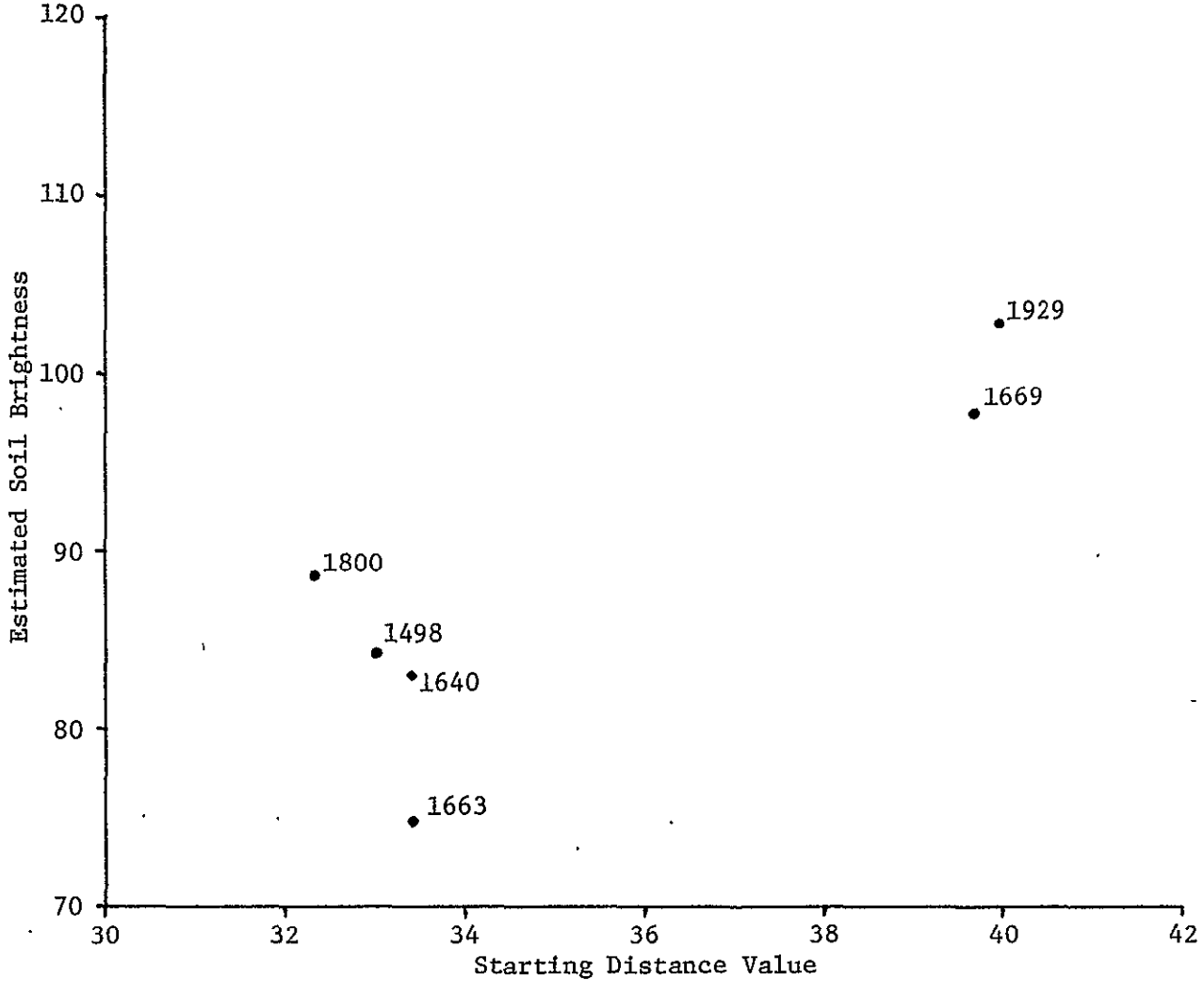


FIGURE G-19. CORRELATION OF ESTIMATED SOIL BRIGHTNESS WITH STARTING DISTANCE VALUE - 2nd ITERATION

functions resulting from analysis of a more extensive data set would almost certainly vary from those here defined, but the forms of the preliminary functions do have strong support in physical understanding. Thus while the labeling logic as specified may fail to do justice to the concepts on which it is based, it does provide a benchmark, and a means by which the underlying concepts may be both understood and incorporated into test and evaluation exercises.

The refined labeling logic allows for three modifications to the orientation of the decision line to accommodate moisture stress and soil brightness effects. Of these three, two have been specified. The third, modification of the decision line slope in response to soil brightness variations, could not be supported with the available data. Thus in this procedural specification, a standard slope of 0.61 is used. This is the slope of the decision line in the original labeler, and represents something of a median value for slopes computed in the second iteration of Landsat data analysis.

The relationship between estimated peak Greenness and the start of separability is defined as illustrated in Figure G-20.

The data points support but do not prove the validity of the described function. There is, however, physical basis for the relationship depicted. Above a Greenness peak representing the low end of "normal" moisture conditions one would expect to find crops of increasing leaf area or, qualitatively, "lushness", perhaps as a result of increasing abundance of moisture (up to the point at which the over-abundance of moisture itself became stressful). It would not be expected, however, that these crops would exhibit a pronounced lengthening of the growing cycle, since climatic constraints quickly come into play. Normal and above-normal crops should develop at the same rate, and so have the same time period for separability.

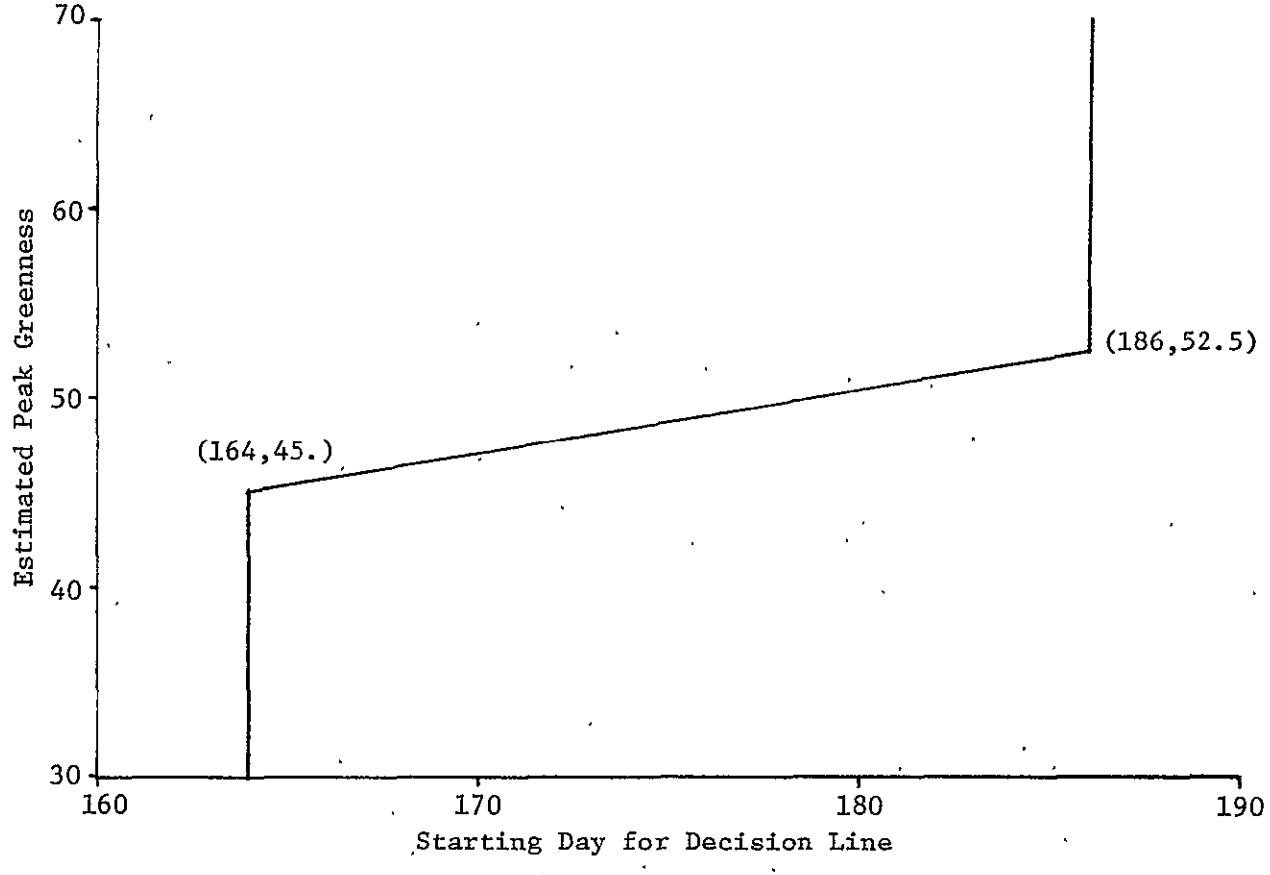


FIGURE G-20. FUNCTION USED TO PREDICT STARTING DAY OF SEPARABILITY FROM ESTIMATED PEAK GREENNESS

The impact of moisture conditions on start of separability is felt at below-normal moisture levels where stress triggers an increase in rate of development. This is reflected in the sloping portion of the function. There must also, however, be some limit on how much development can be speeded up, or how much stress can be tolerated. This is reflected in the vertical portion of the function at extremely low peak values.

The relationship between soil brightness and the initial distance value of the decision line is defined as shown in Figure G-21. Modeling results indicated that this relationship only comes into play in stressed (more open) canopies. A peak value of 52.5 is defined as being the minimum peak of a normal canopy. For segments with a lower peak estimate than 52.5, the function is utilized.

Again, support can be found both in the data and in physical understanding, but precise definition with the sparse data was not possible. The vertical portion of the function reflects the fact that some limit is imposed on how small the distance value can become. Since label assignment takes place after the crops have begun to move off the green arm, they must be at a distance from the reference line greater than that of the green arm itself. Even a very dark soil would not be expected to shift the data past the green arm. The actual value of the vertical portion (Distance = 33.0) is simply based on the apparent clustering of the darker-soil segments around this value.

Given these functional relationships, the procedure for labeling spring small grains is as follows:

1. Distinguish spring small grains from non-small grains by some unspecified technique.

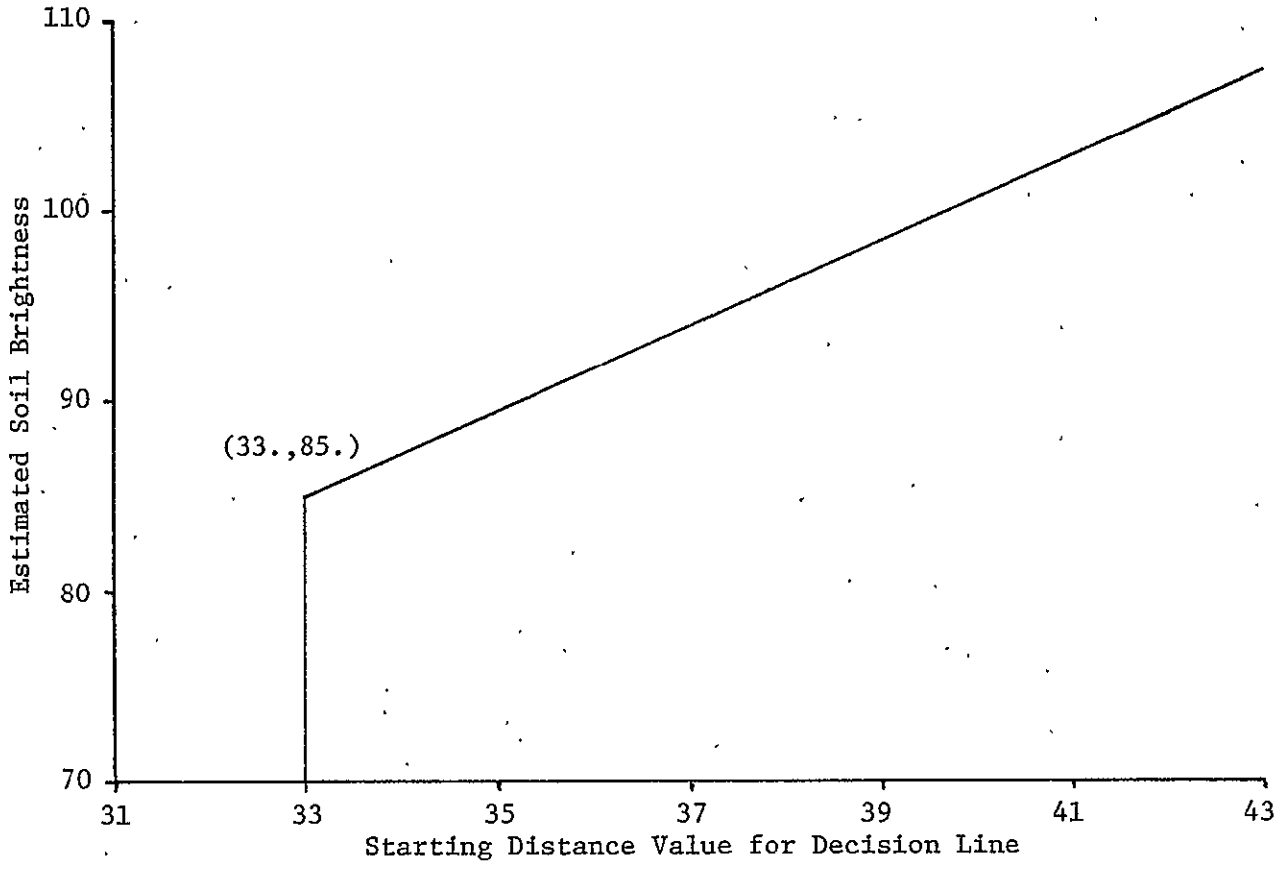


FIGURE G-21. FUNCTION USED TO PREDICT STARTING DISTANCE VALUE FOR DECISION LINE FROM ESTIMATED SOIL BRIGHTNESS

2. If not already done, apply data normalization techniques comprised of:
   a. satellite calibration
   b. cosine sun angle correction
   c. bad data, cloud, etc. detection (ERIM's SCREEN)
   d. haze correction (ERIM's spatially-varying XSTAR)

3. Carry out manual screening of segments:
   a. Acquisition history requirements:
      1) At least three acquisitions (at least 18 days apart) in the spring small grains growing season (typically May through July)
      2) One acquisition around the time of turning (typically mid-July)
   b. Presence of unusual phenomena - clearly anomalous conditions such as wide-spread abandonment, pre-maturity harvesting, etc.
   c. If acquisition history is inadequate or unusual phenomena are observed, spring wheat/barley labeling is not attempted.

4. Estimate crop calendar shift using the two-stage segment-specific technique:
   a. Choose that shift which maximizes the cross-correlation term:

$$R = \frac{2}{1 + \dfrac{\sum F_{i+\tau}^{2} * \sum G_i^{2}}{\left(\sum F_{i+\tau} G_i\right)^{2}}}$$

where F_i = reference profile function value
G_i = data value
τ = shift value

b. Make a first estimate using a reference profile of the form:

$$F(t) = at^b e^{ct^2}$$

where $F(t)$ = Greenness - 25.*

t = shifted day of year - 125.

$a = 0.65163$

$b = 1.2957$

$c = -0.52415 \times 10^{-3}$

The profile is calculated for offset days 1 through 120. In addition, a 30-day tail is added to the early side of the profile, with a constant value equal to the value at offset day 1.

c. Make a second estimate using a profile of the form:

$$F(t) = \begin{cases} ae^{b_1(t-t_p)^2} & ; t < t_p \\ ae^{b_2(t-t_p)^2} & ; t \geq t_p \end{cases}$$

where $F(t)$ = Greenness - 25.*

t = shifted day of year

t_p = shifted day of peak = 160.

a, b_1, b_2 = parameters

The parameters a, b_1, and b_2 are estimated for each segment using linear regression on small-grain blob means after the first shift has been applied; note that

$a + 25.$ = estimated peak Greenness

*This parameter reflects the use of data that have been Tasseled-Cap transformed and to which an offset of 32 counts has been added to eliminate negative numbers.

5. Compute segment-level mean Brightness in shifted days 110-120 for all available small grains pixels (this is soil Brightness diagnostic).

6. Modify the decision line, based on estimated segment-level peak Greenness and soil Brightness, in the manner specified in Figures G-20 and G-21.

   Default values are:

   a. Slope = 0.61
   b. Starting Day = 186
   c. Starting Distance = 33.

7. Assign labels: barley if above line, wheat if below.

APPENDIX H
DESCRIPTION OF DEVELOPMENT OF THE LANDSAT 3 TO LANDSAT 2 CALIBRATION CORRECTION

In order to carry out analysis procedures on Landsat data in such a way that the procedures act in a consistent manner regardless of which Landsat collects the data, a normalizing transformation must be carried out so that the calibration of counts to received signal level is consistent. At ERIM, work was previously completed and reported [38] for adjusting the calibration of Landsat 1 and 2 data to a standard. The standard was that calibration used for Landsat 2 segment data used within LACIE during its first three years. This section discusses the work recently completed to determine and test a similar normalizing transformation for Landsat 3 data.

The criterion used as a goal in developing the transformation was that the behavior of algorithms SCREEN and XSTAR [36, 37] on the resulting transformed data must be equivalent to their behavior on Landsat 2 data. For this to happen, the processed Landsat 3 data would need to exhibit a stable location in spectral space for the overall pattern of scene pixels (e.g., a stable Tasseled Cap [12]), and the location of the pattern would be the same as for Landsat 2 data.

H.1 DEVELOPMENT

The first step in developing the transform was to prepare a data set consisting of acquisition pairs in which one of each pair is Landsat 2 data, and the other is Landsat 3 data over the same site, exactly nine days earlier or later (since the satellite orbits are nine days apart, and consistent position within a full frame was desired).

The available data consisted of Landsat acquisitions, obtained from Landsats 2 and 3 throughout the 1978 spring wheat growing season, over 13 5x6-miles sites in the Northern U.S. Great Plains. From this base, 31 acquisition pairs were prepared. Care was taken to avoid acquisitions with a significant amount of clouds or water, and to avoid pairs for which the haze levels or scene content were extreme.

All acquisitions in the 31 pairs were processed using algorithms SCREEN and PFEAT. SCREEN, which recognizes and flags cloud, shadow, water, and unusual pixels, was applied to the Landsat 3 data even though it was believed that the calibration was not optimum for SCREEN. This was done since a visual screening eliminated acquisitions that had many such pixels, and since SCREEN still succeeded in flagging some unwanted points and false flagging presented no major difficulty.

Algorithm PFEAT processed the remaining unflagged pixels and extracted features used in subsequent analysis. The features consisted of a "green arm mean" and a "soil arm mean", each of which contain four components. These features, defined below, previously have been found to be relatively stable within the "Tasseled Cap" (the pattern of agricultural pixels plotted in Landsat signal space), as long as the scene contains crops at a variety of growth stages (even if the portion of the scene at a given growth stage changes). To calculate these features, one applies first the cosine sun angle correction, and next the Tasseled Cap transformation. The features then are computed as shown in Table H-1. Once the two feature vectors were extracted by PFEAT, they were averaged to form a single scene diagnostic for each acquisition. Independently with each Landsat band, a regression was carried out of the scene diagnostic for Landsat 3 acquisition with the diagnostic for the paired Landsat 2 acquisitions.

TABLE H-1. FEATURES COMPUTED BY PFEAT

Component	"Soil Arm Mean"	"Green Arm Mean"
1	Mean Brightness for pixels in soil window (Greenness between +10 and -10).	Brightness value associated with a point at fifth percentile of D (= .68* Brightness - .73* Greenness) and mean of E (=.73* Brightness + .68* Greenness) for pixels below fifth percentile of D.
2	Fifth percentile of Greenness for all pixels.	Greenness value associated with the above point.
3	Fifth percentile of Yellow for all pixels.	Mean of Yellow for pixels below fifth percentile of D.
4	Mean of Nonsuch for pixels in soil window.	Mean of Nonsuch for pixels below fifth percentile of D.

In examining the results of this initial regression, it was found that the residuals of ten of the pairs were unusually large. These pairs were excluded from the data set in preparing the final transformation. But reasons for the excessive residuals were sought. A scatter plot of the average Tasseled-Cap Yellow component (which has been shown [12] to be related to the haze level in the scene) indicated that for at least half of the high-residual pairs, the haze level was not the same, and therefore those pairs were measures as much of haze difference as of calibration difference. Those pairs not explained by haze condition are suspected to have some other but still undetermined scene effects that caused confounding signal differences between acquisitions in a pair.

The remaining 21 pairs were used to determine a final transformation. Two models for the transformation were considered: the first was the simple multiplicative and additive transformation in each band that had been successfully used for other satellite calibration correction transformations; the second was pure multiplicative in each band. In examining the regression results, it was found that the additive constants were relatively small in the first model, and furthermore that there was no important difference in the goodness of fit in any band, and so the pure multiplicative model was accepted.

The transformation determined by this effort is given below:

$$\vec{x}' = \begin{bmatrix} 1.1371 & 0 & 0 & 0 \\ 0 & 1.1725 & 0 & 0 \\ 0 & 0 & 1.2470 & 0 \\ 0 & 0 & 0 & 1.1260 \end{bmatrix} \vec{x} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

where $\vec{x}$ is the untransformed Landsat 3 signal vector

$\vec{x}'$ is the transformed signal vector.

The following three steps were taken to determine how well this transform works.

First, corrected Landsat 3 acquisitions were processed using the SCREEN algorithm. The resulting data was mapped and examined for incorrect screening. Bodies of water flagged by SCREEN were found to have the same general size and shape as the same bodies of water flagged by SCREEN in Landsat 2 acquisitions over the same site. Indications of incorrect operation of SCREEN were sought, including whether especially light or dark fields had an increased tendency to be flagged on Landsat 3 data. It appeard, at least in the Northern U.S. Great Plains, that SCREEN operated normally on Landsat 3 data transformed by the above equation.

Second, Landsat 2 and corrected Landsat 3 acquisitions were processed using SCREEN, and scatter plots were made of pairs of Tasseled Cap components. A subjective visual examination of this data indicated that with the correction, the position of features within the Tasseled Cap pattern (such as the "soil line") was relatively stable between the two satellites.

And third, an objective examination of the performance of XSTAR on calibration corrected Landsat 3 data was carried out. Several consecutive day Landsat 3 acquisition pairs were identified and processed through calibration correction, SCREEN, XSTAR, and Tasseled Cap. As indicated above, the data patterns looked right qualitatively. The average difference (RMS error) between data on consecutive days was computed on a pixel-by-pixel basis. Since these same measures were previously used in the original XSTAR evaluation [3], the expected range was known. The values computed fall within the previously established normal range, indicating that XSTAR was working as well as it normally does.

Given the above three indications, we feel that the transformation has proven in its self reasonably well, at least in the region for which Landsat 3 data was available in this study.

APPENDIX I

DESCRIPTION OF TOLERANCE BLOCK STRATIFICATION

This Appendix motivates and describes a study whose purpose was to find improved statistical methods of spectral stratification in the context of Procedure M, a system for estimating the acreage of an agricultural crop, such as wheat, from digitized Landsat data [3]. The development of this procedure was stimulated and supported by the Large Area Crop Inventory Experiment (LACIE). Results are presented in Section 5.2.

Procedure M as applied to wheat recognition stratifies spectral data by:

1. Clustering the pixels into field-like groups called "quasi-fields" that are homogeneous spectrally and spatially.
2. Clustering the quasi-fields spectrally into strata.

Step 2, the clustering of quasi-fields into strata, is designed to separate wheat from non-wheat strata and thereby achieve a sampling efficiency.

The grouping of pixels into quasi-fields has been largely successful. Figure I-1 is a histogram of the percent wheat in quasi-field interiors. (The interiors consist of pixels faced on all four sides by pixels from the same quasi-field). This histogram was compiled over all quasi-fields that have interiors from 12 Kansas segments, three acquisitions each. Most of the quasi-fields have less than 10% or more than 90% wheat. Between 10% and 90% wheat, there is only a small scattering of quasi-fields.

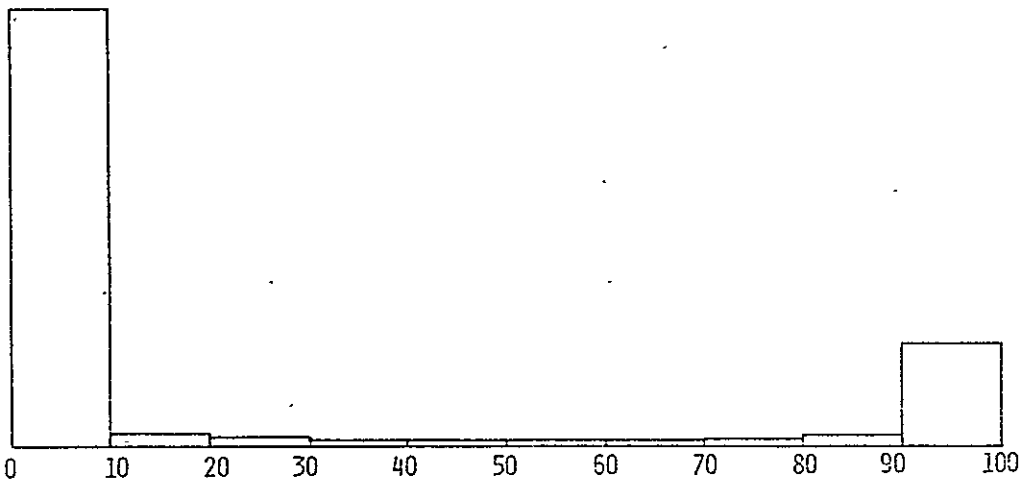


FIGURE I-1. HISTOGRAM OF PERCENT WHEAT IN QUASI-FIELD INTERIORS FOR 12 KANSAS SEGMENTS

The picture would not be as pretty if we included edge pixels (i.e., those that are not interior) in the quasi-fields but we would not expect it to be. Edge pixels are often crossed by field boundaries and are the ones that suffer most from misregistration.

The corresponding histogram for strata (Figure I-2) shows some mixing of wheat and non-wheat quasi-fields. To make this histogram comparable to the other, the stratum count is weighted by the number of quasi-fields in each stratum. Also for comparability, the histogram is based on quasi-field interior ground truth. So whatever fuzziness is in this histogram is not caused by edge pixels.

A big group of non-wheat quasi-fields are put together into relatively pure strata. The group is not as big as in the quasi-field histogram, for when we compare the two figures, we see that some of the 0 to 10 percent quasi-fields in the quasi-field histogram have spilled over into the 10 to 20 and 20 to 30 percent bins in the stratum histogram. Similarly, the stack of wheat quasi-fields is spread out into the 80 to 90 and the 70 to 80 bins.

The stratification was carried out by our unsupervised clustering algorithm BCLUST [9]. The question we are considering is whether stratification can be improved by a better clustering algorithm.

One problem with BCLUST is its tendency to produce a few large clusters and many small ones. Figure I-3 shows a typical distribution of pixels in a 40-cluster stratification. We try to sample in proportion to the size of the strata because this is the best rule when the stratum wheat proportions are unknown. But in the BCLUST stratification, the big strata are multiple sampled and many small strata are unsampled. Leaving the small strata out would create a bias, so we combine the zero-allocation strata into one wastebasket stratum and sample from it proportional to size. (But we require at least one quasi-field in the sample.) We cannot expect that this wastebasket stratum will be pure, so the sampling from it is inefficient.

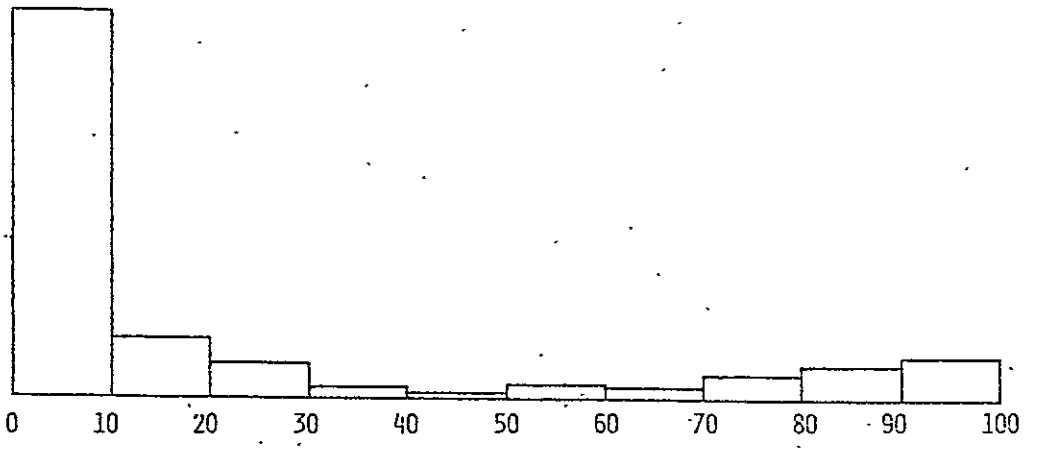


FIGURE I-2. HISTOGRAM OF PERCENT WHEAT IN SPECTRAL STRATA FOR 12 KANSAS SEGMENTS

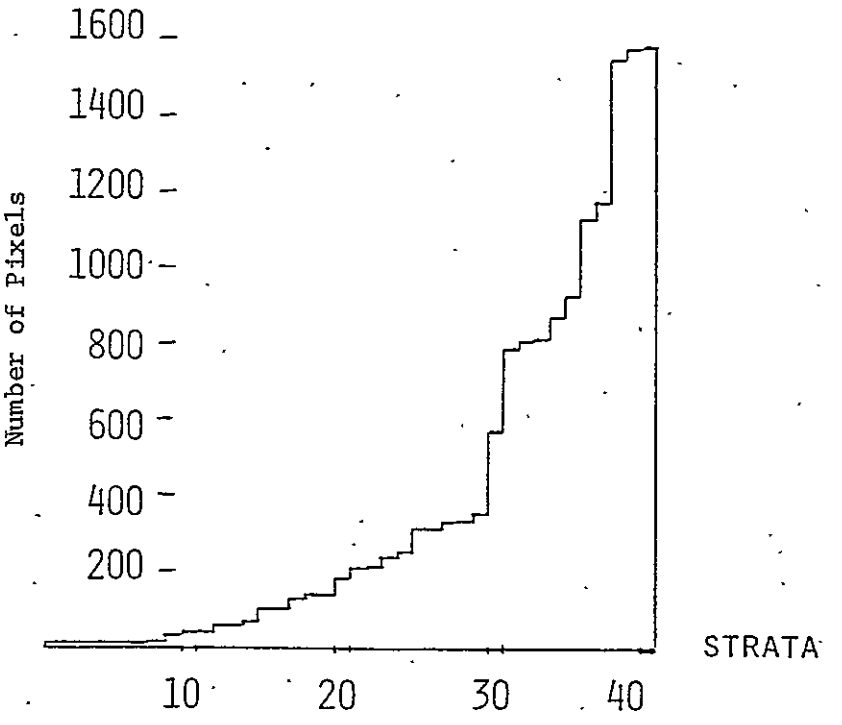


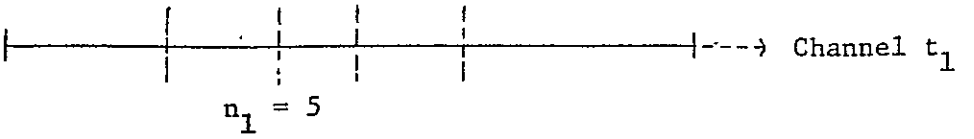
FIGURE I-3. PIXEL DISTRIBUTION FOR BCLUST STRATA SEGMENT 1165

The large strata do not have sampling problems if they truly separate wheat from non-wheat. But if they are so large that they mix up the wheat and non-wheat quasi-fields then it would be better to divide them further into small strata, more localized spectrally and more homogeneous with respect to crop type.

A good clustering algorithm that produced more uniformly-sized strata might improve on the stratification performance of BCLUST.

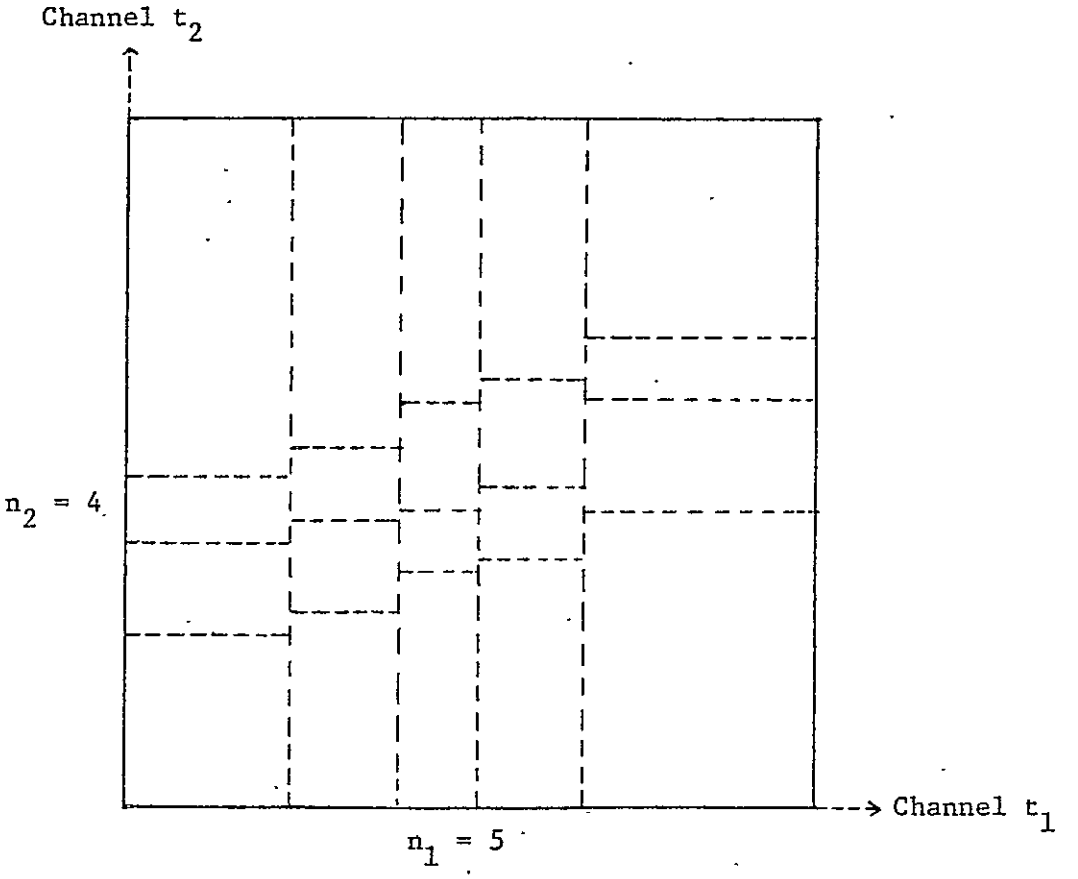
An approach to defining a clustering algorithm producing equal-sized clusters is the use of tolerance blocks. "Tolerance blocks" are equally-populated regions of spectral space constructed as follows. We decide on a small number of channels, $t_1,\ldots t_k$, to generate the blocks. We consider the first channel t_1 and order all the quasi-fields according to this channel. We separate this ordered group of quasi-fields into n_1 equal-sized subgroups--equal in the sense of having approximately the same number of pixels (Figure I-4). Then we consider each subgroup in turn, order it according to our next channel t_2, and divide it into n_2 smaller subgroups (Figure I-5). We can now consider each one of the smaller subgroups, order it according to our third channel t_3, and divide it into n_3 still smaller subgroups. We keep this up for all the generating channels specified. The final subgroups are the tolerance blocks, n_1, $n_2,\ldots n_k$ in all.

Not all channels need be included in this process. If the same set of channels is used in a different order, the tolerance blocks produced are not necessarily the same. (The results, however, were very similar in our tests.) When channel t_2 is used to divide the first set of subgroups, the points of division will, in general, be different from subgroup to subgroup (column to column in Figure I-5). Because we don't cut any quasi-fields in half, but rather assign them to one subgroup or another, the equality of the pixel size of the subgroups can only be approximate.



(The cuts in Channel t_1 separate the quasi-fields into five regions of nearly equal pixel size.)

FIGURE I-4. FIRST CUT TO CREATE TOLERANCE BLOCKS



(The columns are equal-sized groups of quasi-fields separated by cuts in Channel t_1. The rectangles are equal-sized groups of quasi-fields separated by cuts in Channel t_2.)

FIGURE I-5. FIRST AND SECOND CUTS TO CREATE TOLERANCE BLOCKS

The blocks are spectrally homogeneous with respect to the generating channels. How homogeneous they are depends on the number of divisions in each channel. But because the number of blocks is the product of the number of divisions, the number of divisions in each channel must be small if we are to end up with a reasonably small number of blocks. So spectral homogeneity of tolerance blocks is limited in two ways. Some channels are left out of the block construction and those that are represented may have coarse divisions.

In order to achieve a greater spectral homogeneity, we defined a second tolerance block algorithm that uses all the spectral channels in the clustering process. The tolerance block means are used as seeds distributed like a network throughout spectral space. Around the seeds, clusters are formed by ordinary spectral clustering using a distance function. Although a subset of channels may have been used to create the blocks, all channels are used to compute the block means and carry out the clustering. We hoped to combine in one algorithm the virtues of uniformly-sized clusters and spectral homogeneity.

How well the tolerance block algorithms have succeeded in equalizing the clusters can be seen in Figure I-6, a comparison of distributions of strata sizes produced by the three algorithms. BCLUST has a very uneven distribution as we have seen. Many clusters have only a very small number of pixels. When the tolerance blocks themselves are used as clusters, the distribution is very even. When the tolerance blocks are used as seeds, the distribution is less even than for the blocks but considerably more even than for BCLUST.

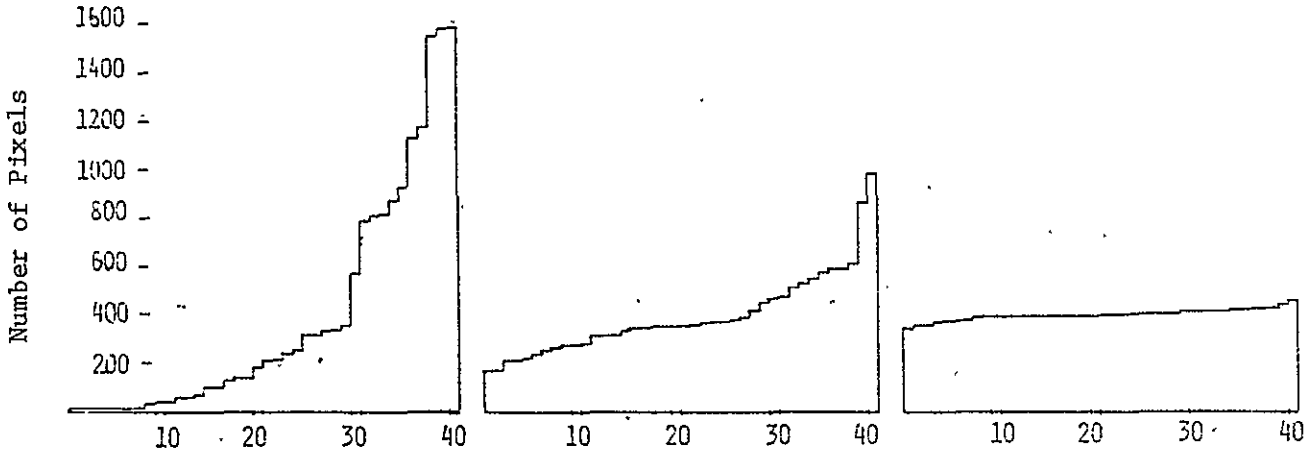


FIGURE I-6. PIXEL DISTRIBUTIONS FOR THREE CLUSTERING ALGORITHMS SEGMENT 1165

APPENDIX J
ERROR MODEL STUDIES

This appendix contains investigations of the effect of labeling error on the error in the proportion estimate. Section J.1 is an analysis of the effect of labeling error on the bias of the estimate in a simplified situation. Section J.2 presents a model for predicting bias and variance of Procedure M stratified sampling estimates when there are labeling errors. Classification errors are assumed constant in this model. In Section J.3, the results are extended to the case of distributed labeling errors. In both J.2 and J.3 the simplifying assumption is made that blobs are sampled and estimates made without regard to the number of pixels in the blob. This assumption permits a manageable expression for variance. Finally, Section J.4 relates labeling error to achieving some level of performance, as is exemplified by the 90/90 criterion.

J.1 LABELING ERROR AND ESTIMATION BIAS: SIMPLE MODEL

A simple model is described and a corresponding estimation problem analyzed which dramatizes the effect of labeling error on proportion estimation.

We assume a random sample from an infinite population that contains two classes of objects, wheat and other. The proportion p of wheat is unknown. We don't observe the class to which the points in the sample belong, but the points are labeled (by analyst-interpreters or any other method) and we observe the labels only. It is assumed that the labeling process introduces error as follows. If a sample point is in class W (wheat), α is the probability that it is correctly labeled. If a point is in class O (other), β is the probability it is incorrectly labeled. The only knowledge we have about α is that it lies in some subinterval of the unit interval, $\alpha_1 \leq \alpha \leq \alpha_2$; similarly β lies in some interval, $\beta_1 \leq \beta \leq \beta_2$. The problem is to estimate p.

In formal terms, let

$$X_1, \ldots, X_N$$

be N independent observations of a random variable X with

$$P_r(X=1) = p\alpha + (1-p)\beta$$

$$P_r(X=0) = 1 - P_r(X=1).$$

The parameters α, β, and p lie in the intervals, $\alpha_1 \leqq \alpha \leqq \alpha_2$, $\beta_1 \leqq \beta \leqq \beta_2$, $0 \leqq p \leqq 1$. The problem is to estimate p.

Let

$$\rho = p\alpha + (1-p)\beta$$

Then $\sum_1^N X_i$ has a binomial distribution with parameters (N,ρ). If we knew ρ exactly we could not determine p exactly. All we could say was that p was in a certain range, this range depending upon the ranges of α and β. Thus we have an identifiability problem and p is not estimable in that an unbiased estimate of p does not exist.

Figure J-1 illustrates the range of p as a function of ρ for $0.65 \leqq \alpha \leqq 0.90$, $0.05 \leqq \beta \leqq 0.25$. The range of ρ is $0.05 \leqq \rho \leqq 0.90$.

Now let $\hat{p}$ be an estimate of p and let $E(\hat{p}|p,\alpha,\beta)$ denote the expected value of $\hat{p}$ given p, α, and β. Then for fixed ρ_o and all triplets p,α,β for which $\rho_o = (\alpha-\beta)p+\beta$

$$E(\hat{p}|p,\alpha,\beta)$$

is constant. Thus for this set

$$\max_{\alpha,\beta,p} \left| E(\hat{p}|p,\alpha,\beta) - p \right|$$

$$\rho_o = (\alpha-\beta)p + \beta$$

$$\rho = (\alpha-\beta)p+\beta$$

$$0.65 \leq \alpha \leq 0.90$$

$$0.75 \leq \beta \leq 0.95$$

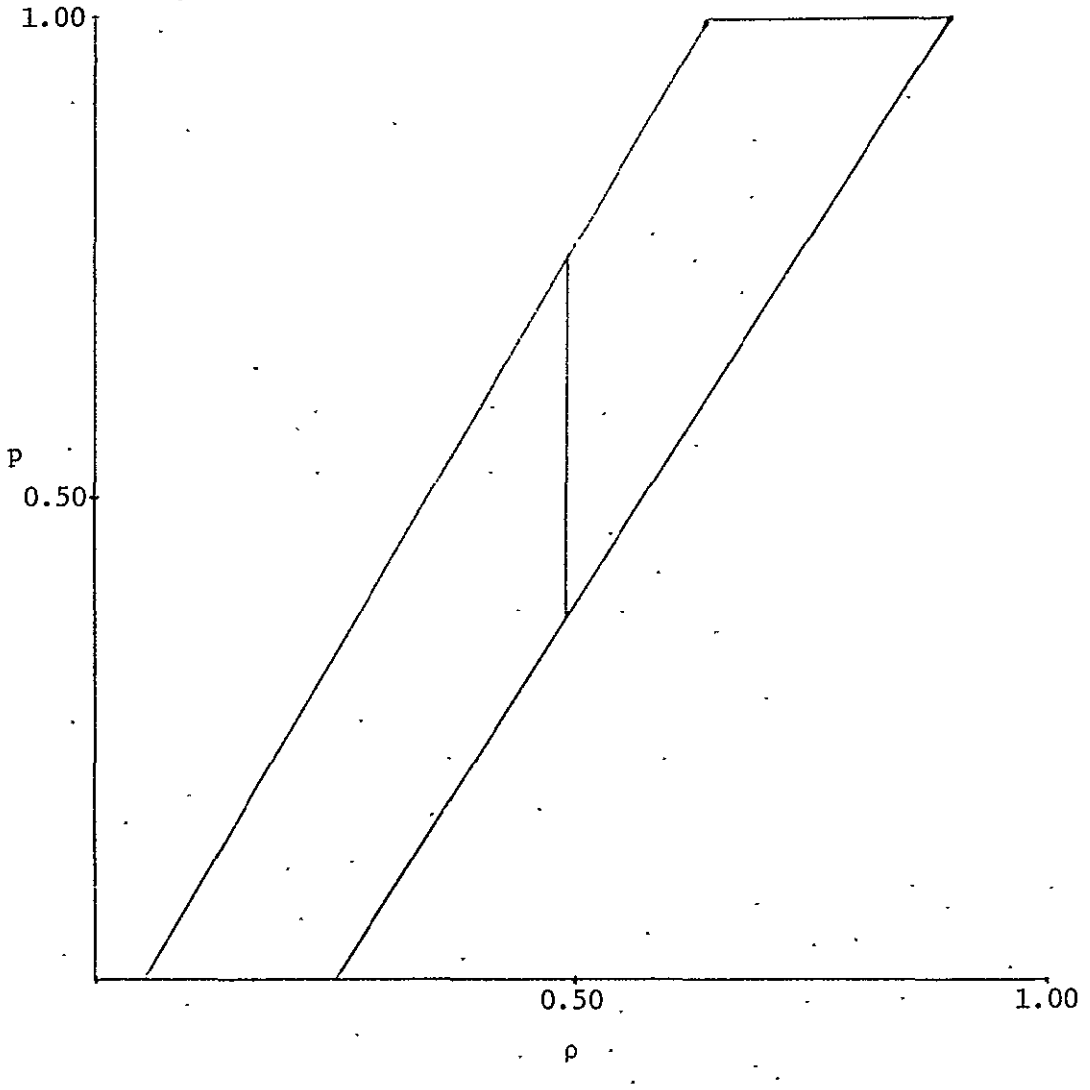


FIGURE J-1. RANGE OF p AS A FUNCTION OF ρ

is at least half the range of p in this set. For $\rho_o = 0.50$, the range of p (indicated by the vertical segment in Figure J-1) is from approximately 0.38 to 0.75. This means that for any sample size and any estimator the absolute value of the bias will be at least 0.18 for some p on the range 0.38 to 0.75. Nor will any sampling scheme improve the situation; only additional prior information about α, β, and p will help.

J.2 A MODEL FOR PREDICTING THE BIAS AND VARIANCE OF PROCEDURE M STRATIFIED SAMPLE ESTIMATES WHEN THERE ARE LABELING ERRORS

Procedure M clusters blobs into spectral strata in order to gain the efficiency of stratified sampling in forming crop estimates. We have measured the success of this clustering by comparing two variances: the variance of the stratified sample estimate and that of the unstratified sample estimate. The biases are not considered in this evaluation because they can be shown to be zero. The ratio of the two variances is a convenient way to express the comparison; it provides a single number to represent the performance of a particular clustering algorithm or parameter setting.

Our first expression for this "reduction of variance factor" [42] was

$$\frac{\Sigma\, n_i\, p_i(1-p_i)}{np(1-p)} \qquad \text{(J-1)}$$

where

n_i is the number of pixels in stratum i,

n is the total number of Σn_i of pixels in the segment,

p_i is the true proportion of wheat in stratum,

p is the true proportion $\Sigma n_i p_i / n$ of wheat in the segment.

This expression is based on the assumptions that sampling from the strata is binomial with replacement and that sample blobs are allocated to strata in proportion to the number of pixels in each stratum.

In the tolerance block study [40], the second assumption was abandoned as unrealistic, and an "m-sample reduction of variance factor" was used. This ratio of variance is computed by allocating a sample of size m as nearly as possible proportional to the size of the strata, but subject to the necessity that the allocation be an integer. The ratio turns out to be

$$\frac{\sum_{\text{all strata } i} \left(\frac{n_i}{n}\right)^2 \frac{p_i(1-p_i)}{a_i}}{\frac{p(1-p)}{m}} \qquad \text{(J-2)}$$

where a_i is the number of sample blobs allocated to stratum i,

m is the size Σa_i of the sample.

This measure, too, suffers from the unrealistic assumption that the blobs are sampled with replacement. The actual sampling technique is unbiased [3] but with a variance that is exceedingly difficult to calculate. A model is necessary and it can be made more realistic by sampling without replacement -- a model we should have used in the tolerance block study. The simplifying assumption of the model is that we are sampling blobs and making estimates without regard to the number of pixels in the blob.

The number of wheat blobs in the sample from stratum i then has the hypergeometric distribution with variance

$$a_i p_i(1-p_i)\left(\frac{b_i-a_i}{b_i-1}\right) \qquad \text{(J-3)}$$

where b_i is the number of blobs in stratum i. The reduction of variance factor becomes

$$\frac{\sum_{\text{all strata } i} \left(\frac{n_i}{n}\right)^2 \frac{p_i(1-p_i)}{a_i}\left(\frac{b_i-a_i}{b_i-1}\right)}{\frac{p(1-p)}{m}\left(\frac{b-m}{b-1}\right)} \qquad \text{(J-4)}$$

where b is the total number Σb_i of blobs in the segment.

When we extend the model to include labeling error, we introduce the possibility of bias. We therefore have two performance measures to consider, bias and variance, and a ratio of variances is no longer especially convenient.

A simple labeling error model, widely used at JSC and ERIM, is of the following pattern:

The Probability of Choosing	When We Truly Have Wheat	Other
Wheat	α	β
Other	$1-\alpha$	$1-\beta$

(J-5)

The probability that a blob from stratum i will be labeled wheat is

$$q_i = \alpha p_i + \beta(1-p_i) = (\alpha-\beta)p_i + \beta$$

This fact allows us to obtain the bias of the wheat estimate whether the model or the Procedure M sampling technique is used. For completeness, we present the derivation here.

The wheat estimate for stratum i is

$$w_i = \frac{\Sigma \lambda_j x_j}{\Sigma \lambda_j}$$

where x_j is the labeled percent wheat of blob j in the sample,

λ_j is the number of pixels in blob j if we use the Procedure M sampling scheme,

λ_j is 1 if we use the model.

The expected value of w_i is

$$\frac{\Sigma \lambda_j Ex_j}{\Sigma \lambda_j} = Ex_i = q_i \qquad \text{(J-6)}$$

The wheat estimate for the segment is

$$\frac{\Sigma n_i w_i}{\Sigma n_i}$$

It's expectation is

$$\frac{\Sigma n_i q_i}{\Sigma n_i} = \frac{\Sigma n_i[(\alpha-\beta)p_i + \beta]}{\Sigma n_i}$$

$$= (\alpha-\beta)\frac{\Sigma n_i p_i}{\Sigma n_i} + \beta$$

$$= (\alpha-\beta)p + \beta$$

The bias is this expression minus p

$$= (\alpha-\beta)p + \beta - p = (\alpha-\beta-1)p + \beta \qquad \text{(J-7)}$$

If we suppose a model where α and β are different for different strata, the bias is

$$\frac{\Sigma n_i[(\alpha_i-\beta_i-1)p_i + \beta_i]}{\Sigma n_i} \qquad \text{(J-8)}$$

The variance of the wheat estimate is obtained from the variances of the stratum estimates. We approximate each stratum estimate by using the hypergeometric model (sampling without replacement). We let W be the number of true wheat blobs in the stratum sample and L, the number of blobs labeled wheat. We are assuming that W is hypergeometric and therefore has mean $a_i p_i$ and variance

$$a_i p_i (1-p_i)\left(\frac{b_i-a_i}{b_i-1}\right)$$

We obtain the variance of L from the general theorem

$$\text{Var } L = E\,\text{Var}(L|W) + \text{Var } E(L|W) \qquad \text{(J-9)}$$

To begin with, we will evaluate the first term $E\,\text{Var}(L|W)$. When W is given, L is the sum of two independent random variables: the number of wheat blobs labeled wheat and the number of non-wheat blobs labeled wheat. The first is binomial (W,α), and the second binomial (a_i-W,β). The variance of the sum L is the sum of the two variances

$$\alpha(1-\alpha)W + \beta(1-\beta)(a_i-W)$$

The expectation of this term with respect to W is

$$\alpha(1-\alpha)a_ip_i + \beta(1-\beta)(a_i-a_ip_i)$$

$$= a_i[\alpha(1-\alpha)p_i + \beta(1-\beta)(1-p_i)]$$

After algebra, this expression can also be written

$$a_i[\beta-\beta^2 + (\alpha-\beta)(1-\alpha-\beta)p_i]$$

The second term $\text{Var } E(L|W)$, obtained as before by considering L as the sum of two binomials,

$$= \text{Var}[\alpha W + \beta(a_i-W)]$$

$$= \text{Var}[(\alpha-\beta)W + \beta a_i]$$

The constant term does not affect the variance and can be omitted. Hence,

$$\text{Var } E(L|W) = \text{Var}(\alpha-\beta)W = (\alpha-\beta)^2\text{Var } W$$

$$= a_i(\alpha-\beta)^2\, p_i(1-p_i)\left(\frac{b_i-a_i}{b_i-1}\right)$$

The variance of L is the sum of the two terms. What we really want is the variance of L/a_i, the wheat estimate for stratum i based on the labeled blobs.

This is

$$\frac{\text{Var } L}{a_i^2} = \frac{\beta-\beta^2 + (\alpha-\beta)(1-\alpha-\beta)p_i + (\alpha-\beta)^2 p_i(1-p_i)\left(\frac{b_i-a_i}{b_i-1}\right)}{a_i} \tag{J-10}$$

$$= \frac{(\alpha-\alpha^2)p_i + (\beta-\beta^2)\ (1-p_i) + (\alpha-\beta)^2 p_i(1-p_i)\left(\frac{b_i-a_i}{b_i-1}\right)}{a_i}$$

The variance of the segment estimate is the sum of these stratum variances weighted by $(n_i/n)^2$. The variance of the unstratified segment estimate is a single term analogous to (J-10) with p, b and m replacing p_i, b_i, and a_i. If α and β vary from stratum to stratum, then α_i and β_i replace α and β in (J-10).

Incidentally, theorem (J-9) may be used to verify that the variance of the with-replacement estimate is

$$[(\alpha-\beta)p_i - 2\beta(\alpha-\beta)p_i + (\alpha-\beta)^2 p_i^2 + \beta-\beta^2]/a_i \tag{J-11}$$

The result can be checked directly because the number L of blobs labeled wheat now has a binomial (a_i, q_i) distribution with variance $a_i q_i(1-q_i)$ which agrees with (J-11) when q_i is expanded.

J.3 BIAS AND VARIANCE WHEN LABELING ERRORS ARE DISTRIBUTED

In Section J.2, above, a model was developed for the bias and variance of the Procedure M estimate under labeling error. The result was that the expected value of the wheat estimate is

$$(\alpha-\beta)p + \beta, \tag{J-12}$$

the bias is

$$(\alpha-\beta-1)p + \beta, \tag{J-13}$$

and the variance is

$$\frac{(\alpha-\alpha^2)p + (\beta-\beta^2)(1-p) + (\alpha-\beta)^2 p(1-p)\left(\frac{b-a}{b-1}\right)}{a} \qquad \text{(J-14)}$$

where

α = probability of labeling a blob wheat given that it is wheat

β = probability of labeling a blob wheat given that it is not wheat

p = proportion of wheat in a stratum or segment

b = number of blobs in the stratum or segment

a = number of blobs in the sample from stratum or segment

The model was based on the assumption of a fixed α and β. Here we apply the model when α and β have a known distribution.

Let $E\alpha$, $E\beta$, var α and var β denote the expected value of α and of β and the variance of α and of β, respectively. Let $\text{cov}(\alpha,\beta)$ be the covariance of α and β. $E\alpha^2$ can either be thought of as the expectation of α^2 or as var $\alpha + (E\alpha)^2$, and similarly for $E\beta^2$. $E\alpha\beta$ is the expectation of $\alpha\beta$ or equivalently, $\text{cov}(\alpha,\beta) + (E\alpha)(E\beta)$.

To get the bias of the wheat estimate $\hat{p}$ we use the conditional expectation formula

$$\text{bias} = E(\hat{p}-p) = E\alpha\beta\, E[(\hat{p}-p)\,|\,\alpha,\beta] \qquad \text{(J-15)}$$

Now $E[(\hat{p}-p)\,|\,\alpha,\beta]$ is given by (J-13). Hence

$$\begin{aligned} E(\hat{p}-p) &= E[(\alpha-\beta-1)p + \beta] \\ &= (E\alpha-E\beta-1)p + E\beta \end{aligned} \qquad \text{(J-16)}$$

To get the variance of $\hat{p}$ we use the conditional variance formula

$$\text{var}\ \hat{p} = \text{var}\,\alpha\beta\, E(\hat{p}\,|\,\alpha,\beta) + E\,\alpha\beta\ \text{var}\ (\hat{p}\,|\,\alpha\beta) \qquad \text{(J-17)}$$

We will evaluate (J-17) one term at a time. The first term is, using (J-12),

$$\begin{aligned} &\text{var}_{\alpha\beta}\,[(\alpha-\beta)p + \beta] \\ &\quad = \text{var}\,[p\alpha + (1-p)\beta] \\ &\quad = p^2\,\text{var}\,\alpha + (1-p)^2\,\text{var}\,\beta + 2p(1-p)\,\text{cov}\,(\alpha,\beta) \end{aligned} \qquad \text{(J-18)}$$

The second term is, using (J-14),

$$E_{\alpha\beta}\,\frac{(\alpha-\alpha^2)p + (\beta-\beta^2)(1-p) + (\alpha-\beta)^2 p(1-p)\left(\frac{b-a}{b-1}\right)}{a}$$

$$= \frac{(E\alpha-E\alpha^2)p + (E\beta-E\beta^2)(1-p) + (E\alpha^2-2E\alpha\beta+E\beta^2)p(1-p)\left(\frac{b-a}{b-1}\right)}{a} \qquad \text{(J-19)}$$

The variance of $\hat{p}$ is the sum of (J-18) and (J-19).

Independence of α and β were not assumed because there is generally a relationship between Type I and Type II errors in detection problems. If the system is adjusted to make α as large as possible, then β is also increased, and if we try to reduce β, we may end up reducing α.

The detection component in the wheat estimate is a human being who may be sensitive to pressures to estimate a reasonable proportion of wheat for the segment. It may be, therefore, that if his α is low, he may tilt his judgment to increase β to make the overall proportion acceptable. Thus, an inverse, rather than a direct relationship between α and β could apply in this case.

The distribution of α and β will be difficult to estimate with confidence. If we guess the distribution, then the expected bias becomes a simple sum (J-16) of our guesses. The distribution depends on the prejudices and pressures affecting the AI's which are difficult to quantify. It would vary considerably from AI to AI and from situation to situation. Our experiment on three AI's would be a very small sample from which to estimate this distribution.

J.4 LABELING ERROR IMPACT ON THE 90/90 CRITERION

It was shown in Sections J.2 and J.3 that the bias of the Procedure M target crop proportion estimate in the presence of labeling error is given by the expression

$$b = \alpha p + \beta(1-p) - p \qquad \text{(J-20)}$$

where b is the bias,

p is the true proportion of the target crop,

α is the accuracy of labeling that target,

and β is the inaccuracy in labeling other (errors of commission).

Equation (J-20) can be written

$$b = (\alpha-\beta-1)p + \beta \qquad \text{(J-21)}$$

expressing the bias b as a linear function of p in the parameters α and β, with slope $\alpha-\beta-1$ and y intercept β.

Figure J-2 illustrates the parametric equations over the dynamic range of α, β, and p:

$$0.0 \leq \alpha \leq 1.0$$
$$0.0 \leq \beta \leq 1.0$$
$$0.0 \leq p \leq 1.0$$

Note that an unbiased estimate of p can be achieved by compensating errors of omission and commission relative to the proportion p. In the extreme, errors of total commission (β=1.0) result in an unbiased estimate only if p=1.0, that is, no other exists in the population; and errors of total omission (α=0.0) result in unbiased estimates only if p=0.0, that is, no target crop exists in the population. Similarly, a commission error of .5 and omission error of .5 is unbiased only if the target crop is 50% of the population.

It was shown in Section J.1 that without knowledge of the range of α and β and their distribution over that range, that an unbiased

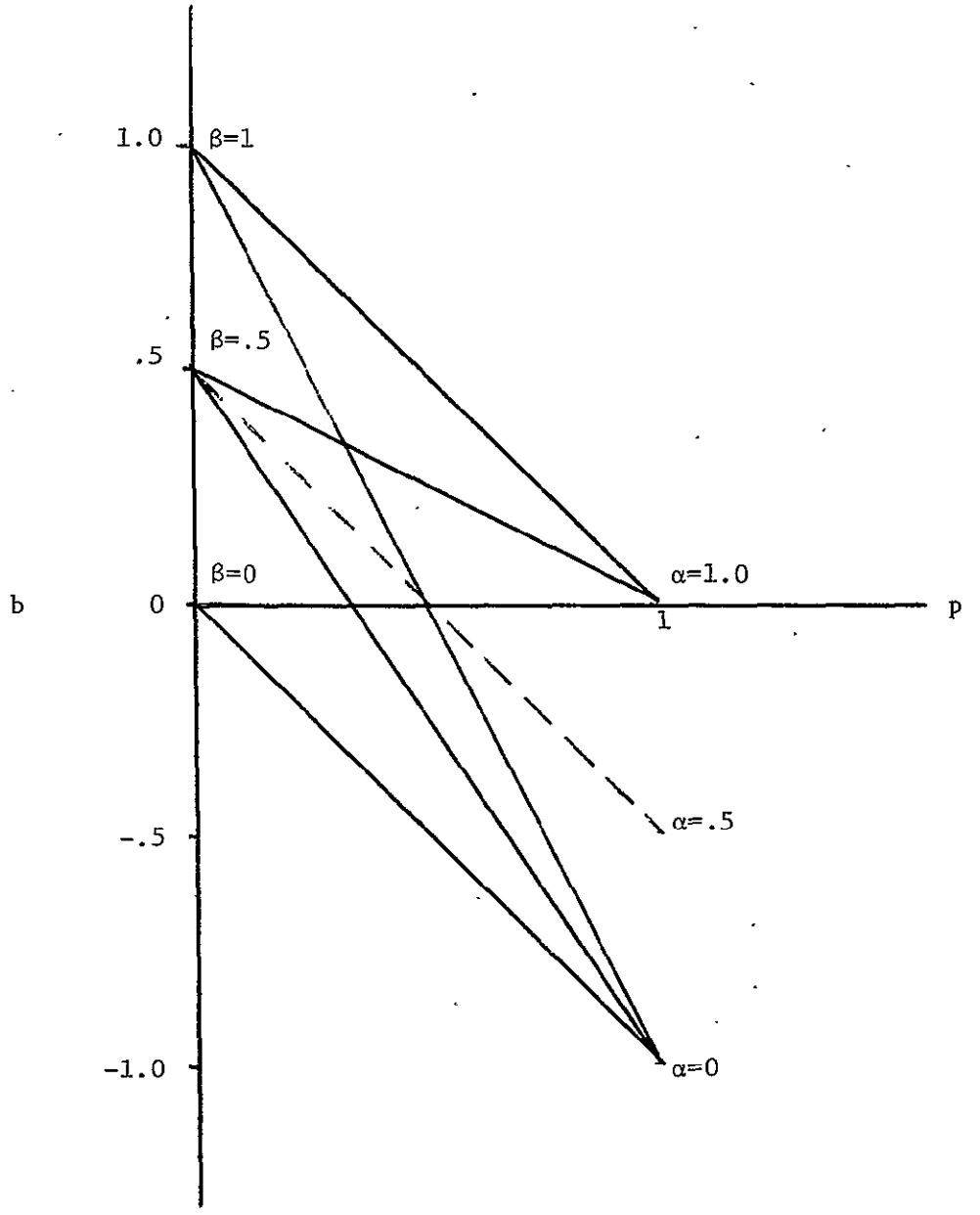


FIGURE J-2. BIAS AS A LINEAR FUNCTION OF TARGET CROP PROPORTION IN PARAMETERS α AND β

estimate of p does not exist. Here we will examine what limited knowledge of these parameters gained by experience implies with respect to the 90/90 criterion.

Evaluations of analyst labeling error in LACIE suggest that the following are reasonable ranges of α and β in labeling wheat:

$$.6 < \alpha < .8$$
$$.0 < \beta < .1$$

Though little is known about the distribution of error over this range, for purpose of illustration we will assume the error can be described by a normal distribution or other definable function.

Figure J-3 illustrates the range of bias for a uniformly distributed and bounded error. The maximum positive bias in this range is .1 when α=.8 and β=.1, occurring when p=.0. The maximum negative bias occurs when p=1.0, α=.6, and β=.0. If p exceeds .33, an unbiased estimate is not achieved in this range of α and β. In addition, the target crop is always underestimated. Note that if a system is biased, one way to improve overall accuracy is by introducing compensating error in labeling.

Further, restricting p to the typical range

$$.2 \leq p \leq .5$$

we find the bias limited to the range illustrated in Figure J-4. Note that an overestimate will occur only if p < .33 and errors of commission approach maximum, while errors of omission approach minimum. Associated with the points (A,B,C,D) are the relative error rates (.2, -.4, -.1, -.4). If the stated range of performance were the absolute range, the 90/90 criteria would never be achieved for p=.5.

Over the stated range, 90% accuracy is achieved wherever

$$\left|\frac{b}{p}\right| \leq .1 \qquad \text{(J-22)}$$

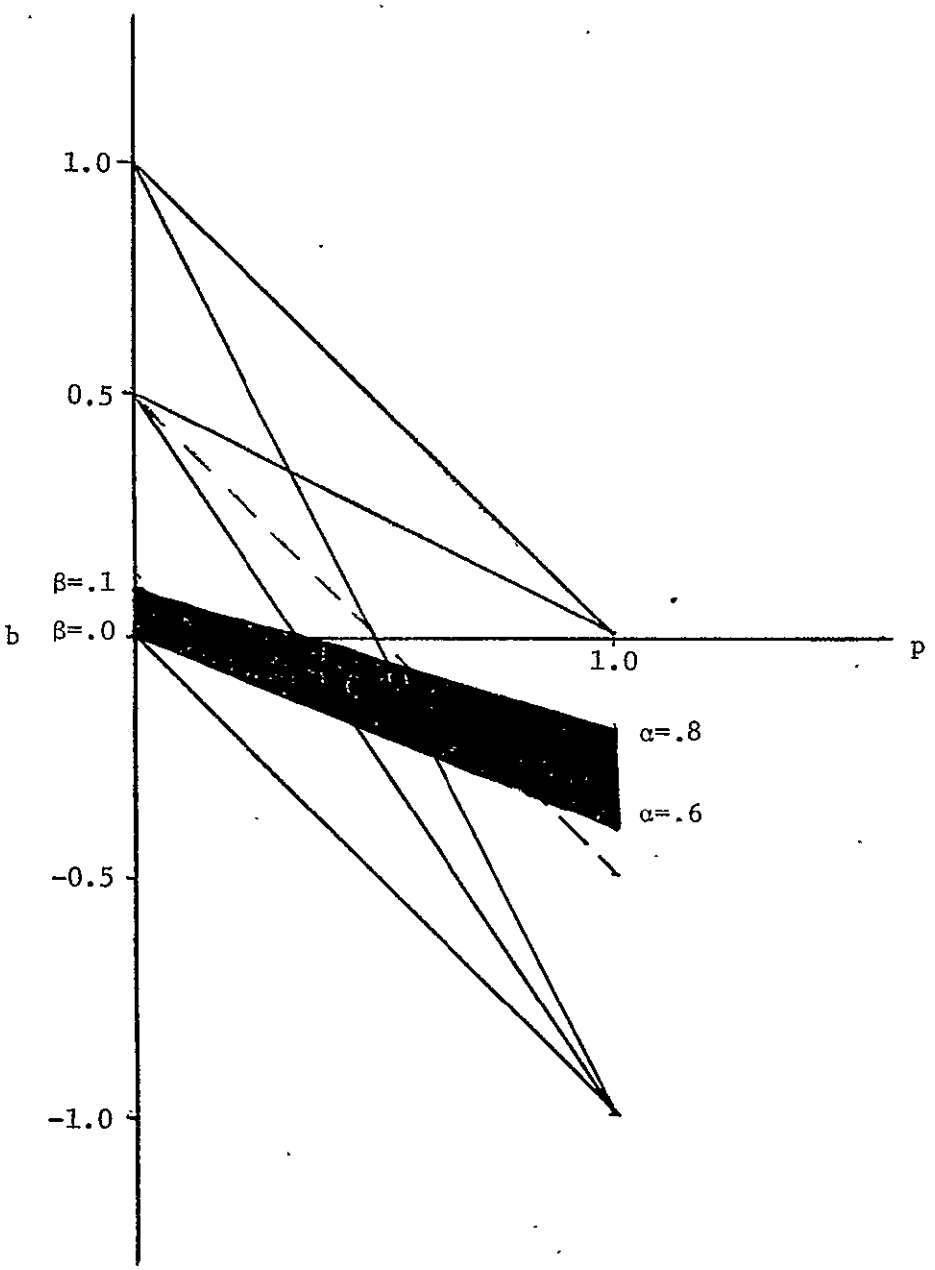


FIGURE J-3. RANGE OF BIAS FOR

$.6 \leq \alpha \leq .8$

$0 \leq \beta \leq .1$

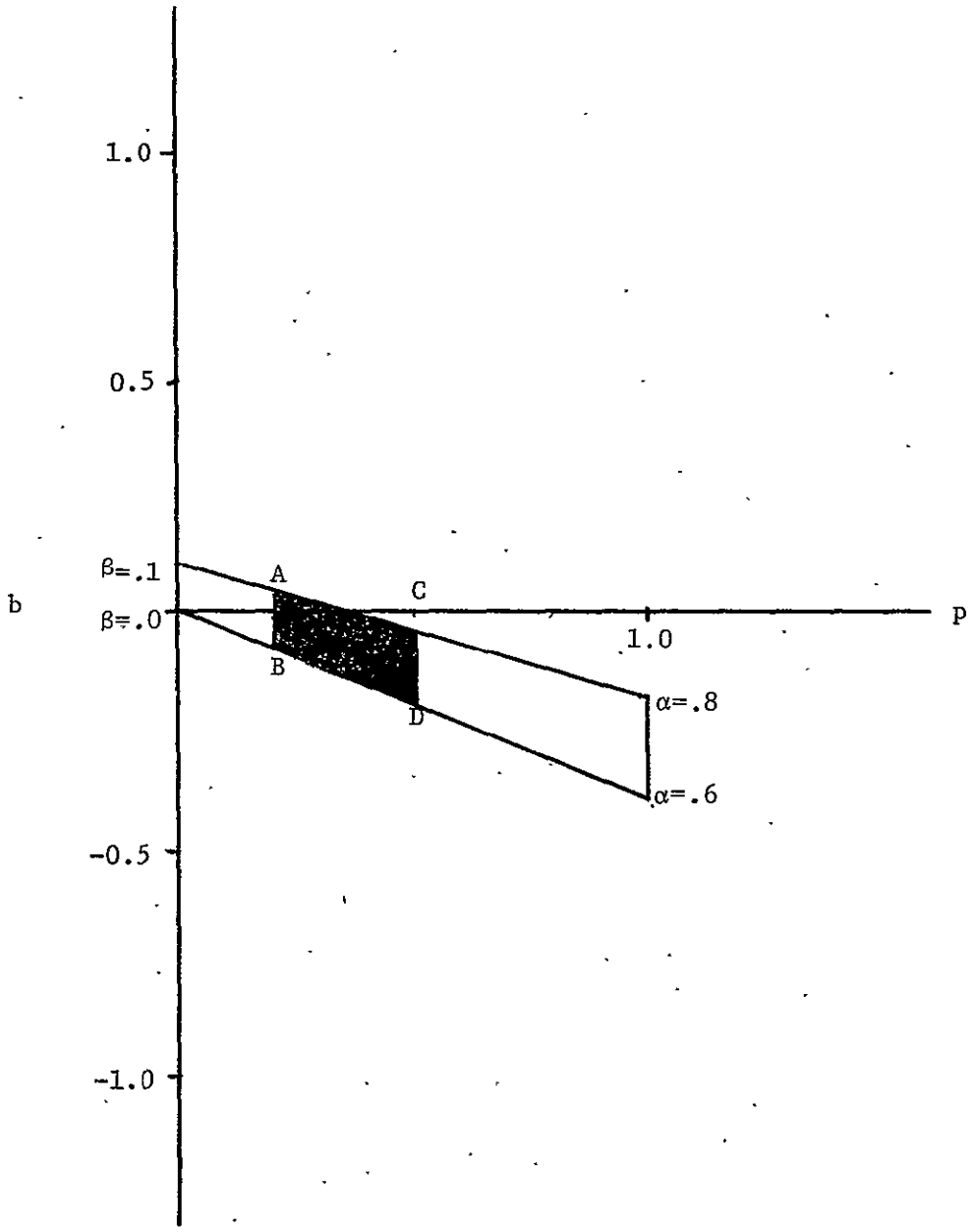


FIGURE J-4. RANGE OF BIAS FOR

$.6 \leq \alpha \leq .8$

$.0 \leq \beta \leq .1$

$.2 \leq p \leq .5$

Using Equation (J-21), 90% is achieved where:

$$(\alpha-\beta-.9)p + \beta > 0 \text{ and } (\alpha-\beta-1.1)p + \beta < 0 \qquad \text{(J-23)}$$

From (J-22), this is necessarily contained in the range

$$-.1p \leq b \leq .1p \qquad \text{(J-24)}$$

Figure J-5 illustrates Equation J-24 (the dashed lines) and Equation J-23 (the shaded area). If one associates a distribution of labeling error the efficiency of the system can be computed. For example, if the specified range totally bounded a uniform distribution of error, then the efficiency of achieving 90% accuracy over the range $.2 \leq p \leq .5$ is the ratio of the area shaded in Figure J-5 to the area shaded in Figure J-4. At a given p, the efficiency is the ratio of the respective ranges of b at that p in the two figures. For example, the efficiency of achieving 90% accuracy at p=.5 is 0%. On the other hand, if the distribution of labeling error were normally distributed, the efficiency of this system at a 90% accuracy level is the area under the normal curve at a p within the range specified by Equation J-24. This is illustrated in Figure J-6.

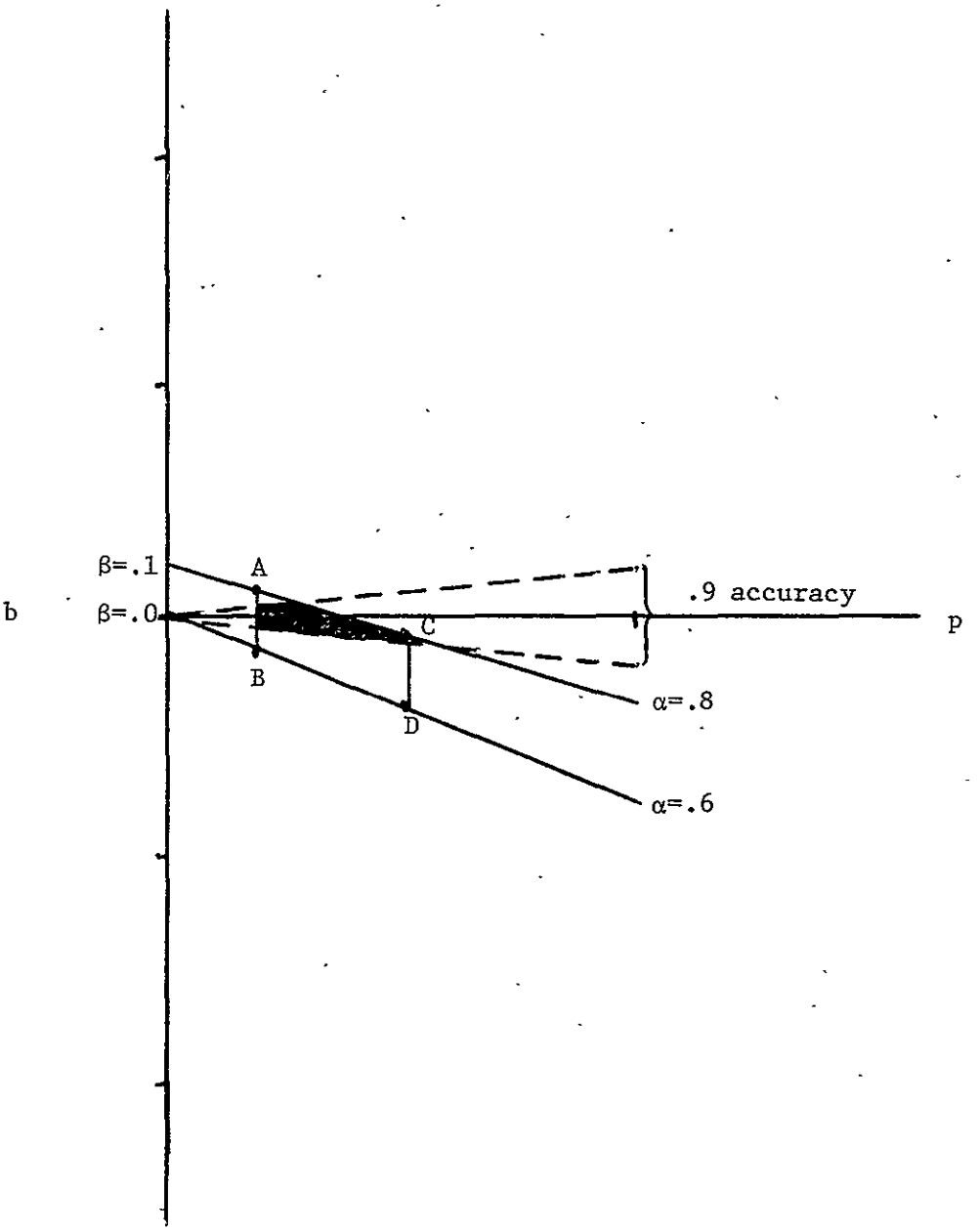


FIGURE J-5. RANGE OF 90% ACCURACY FOR

$.6 \leq \alpha \leq .8$

$.0 \leq \beta \leq .1$

$.2 \leq p \leq .5$

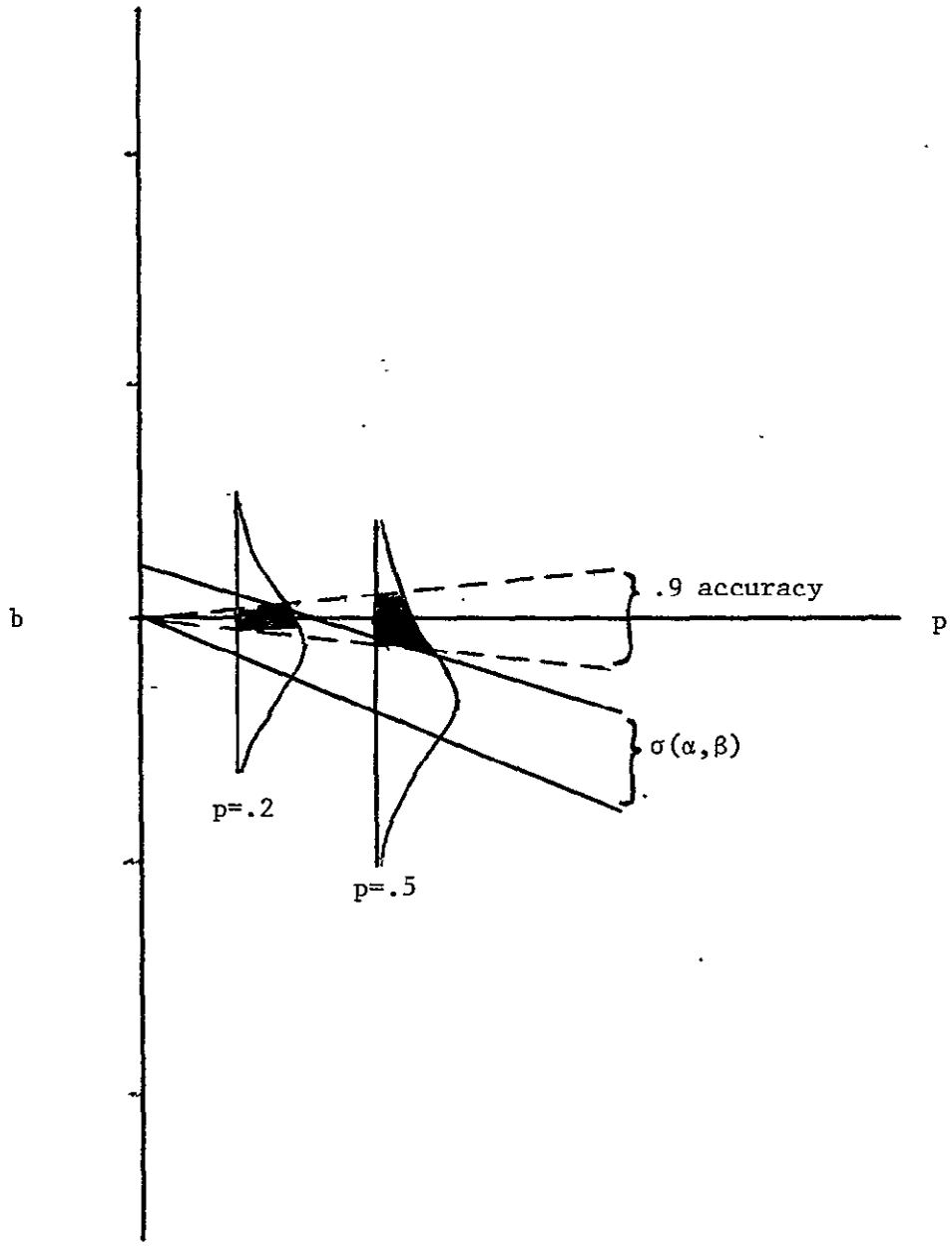


FIGURE J-6. MEETING 90/90 WHEN

$.6 < \alpha < .8$

$.0 < \beta < .1$

$.2 < p < .6$

$\sigma(\alpha,\beta) = (.1,.05)$

APPENDIX K

INFORMATION THEORETIC MEASURES OF AGRICULTURAL INVENTORY SYSTEM PERFORMANCE

Multispectral remote sensing systems are being used for agricultural inventory applications. Their overall performance usually is measured by their accuracy in estimating crop acreages (or crop proportions in sample areas) or by their marginal probability of correct classification of known scene elements. Intermediate stages of some area estimation procedures involve stratification and sampling of subareas (notably, LACIE's Procedure 1 and ERIM's Procedure M).

The variance reduction factor is currently used to measure the performance of such stratification procedures. This appendix introduces and develops alternative performance measures or figures of merit that should be of value, particularly when several classes or crops are involved.* These new measures are based on information theory concepts. Application of information theory concepts to other and more general aspects of area estimation problems may also prove productive and is briefly discussed.

K.1 INFORMATION THEORY CONCEPTS

Consider the simple communication channel portrayed in Figure K-1. The input has m possible states and the output n possible states; in general, $m \neq n$. A transfer of information occurs when the output state is influenced by the input state. The transfer is perfect when the output always has a 1:1 correspondence with the input or, in other words, when there is no uncertainty about the input given a particular value of the output. Noise in the communication channel can introduce errors or uncertainty about the input state for any given output state.

*These concepts were first reported at the June 1979 Quarterly Review under this contract, as documented in ERIM 132400-24-P, Analysis of Scanner Data for Crop Inventories, Environmental Research Institute of Michigan, Ann Arbor, MI, August 1979.

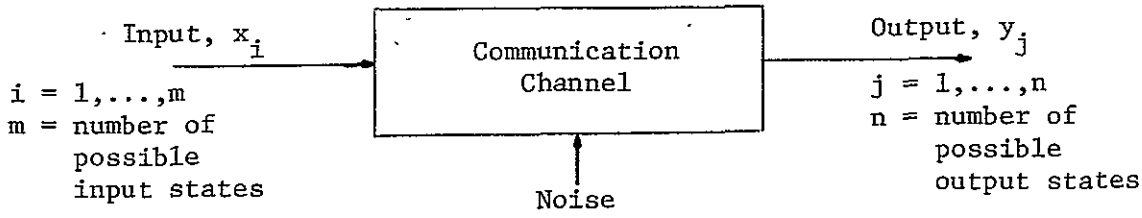


FIGURE K-1. DIAGRAM OF A COMMUNICATION CHANNEL

K.1.1 SELF INFORMATION

To quantify information transfer processes, information theorists, beginning with Claude Shannon, have defined an information measure. The self information or information associated with the known occurrence of a state x_i, which occurs with probability $P(x_i)$, is defined to be:

$$I(x_i) = \log_a \left(\frac{1}{P(x_i)}\right) = -\log_a P(x_i) \qquad \text{(K-1)}$$

where the base "a" usually equals two, in which case the units of information are in "bits".* This quantity also may be interpreted as the uncertainty in x_i or the (a priori) information needed to make x_i certain. The more rare the state, the greater is the information associated with its occurrence. A useful characteristic of this measure is that it is additive, e.g., the total information represented by two independent events is the sum of their self-information values.

K.1.2 ENTROPY

Next, let us consider the average amount of information associated with repeated observations of the event. This quantity is called the entropy or average uncertainty of x; it is the average (a priori) information needed to make x certain. The entropy over the ensemble $X = \{x_i\}$ is

$$H(X) = -\sum_{i=1}^{m} P(x_i) \log P(x_i) = \sum_{i=1}^{m} P(x_i) \log \left(\frac{1}{P(x_i)}\right) \qquad \text{(K-2)}$$

If the probabilities $P(x_i)$ are equal, then $P(x_i) = \frac{1}{m}$ and H is a maximum:

$$H_{max} = \log m \qquad \text{(K-3)}$$

*If the base a = 10, the units of information are in "hartleys" and, if a = e, the units are in "nats".

The minimum value for H is zero and results when a single state occurs with probability one, i.e., there is no uncertainty about the input state.

Since the value of entropy depends on the probability distribution of the input states, a relative measure is sometimes defined. The relative entropy is normalized by the maximum entropy possible for the given number of input states, i.e.,

$$H_R = \frac{H}{H_{max}} \tag{K-4}$$

K.1.3 JOINT INFORMATION

The information content of the joint occurrence of two events depends on their joint probability of occurrence, $P(x_i,y_j)$:

$$I(x_i,y_j) = -\log P(x_i,y_j)$$

But, since

$$P(x_i,y_j) = P(x_i)\ P(y_j|x_i) \tag{K-5a}$$

$$= P(y_j)\ P(x_i|y_j) \tag{K-5b}$$

we have

$$I(x_i,y_j) = -\log[P(x_i)\ P(y_j|x_i)]$$

$$I(x_i,y_j) = -\log P(x_i) - \log P(y_j|x_i)$$

Thus,

$$I(x_i,y_j) = I(x_i) + I(y_j|x_i). \tag{K-6a}$$

Similarly,

$$I(x_i,y_j) = I(y_j) + I(x_i|y_j). \tag{K-6b}$$

K.1.4 INFORMATION EXCHANGE

If x_i and y_j are independent, then

$$P(y_j|x_i) = P(y_j)$$

and

$$I(y_j|x_i) = I(y_j)$$

giving

$$I(x_i,y_j) = I(x_i) + I(y_j)$$

This is the maximum value of $I(x_i,y_j)$. However, from the standpoint of a communication channel it represents no information transfer from input to output.

Maximal information transfer occurs when the input x_i and output y_j are perfectly correlated, i.e., when

$$P(y_j|x_i) = 1$$

$$P(x_i|y_j) = 1$$

and

$$P(x_i,y_j) = P(x_i) = P(y_j)$$

giving

$$I(x_i,y_j) = I(x_i) = I(y_j)$$

Let us then define I_n, the next exchange or transfer of information through the channel, to be:

$$I_n(x_i,y_j) = \text{Max}\{I(x_i,y_j)\} - I(x_i,y_j)$$

or

$$I_n(x_i,y_j) = I(x_i) + I(y_j) - I(x_i,y_j) \qquad \text{(K-7)}$$

Other expressions can be obtained by substituting Equations K-6b and K-6a into Equation K-7:

$$I_n(x_i,y_j) = I(x_i) - I(x_i|y_j) \quad \text{(K-8a)}$$

$$= I(y_j) - I(y_j|x_i) \quad \text{(K-8b)}$$

From Equation K-8a we see that the joint occurrence provides all the information associated with x_i, less that lost due to confusion or the a posteriori uncertainty about which x was sent given that a particular y was received. Again, if x_i and y_j are perfectly correlated, then $I_n(x_i,y_j) = I(x_i)$ and the net information exchanged then is the entire input information. At the other extreme, if y_j is independent of x_i, then $I_n(x_i,y_j) = 0$, the output bears no relationship to the input, and the net exchange is zero.

By utilizing Equation K-8a and the definition of self information, we obtain the following:

$$I_n(x_i,y_j) = \log P(x_i|y_j) - \log P(x_i) \quad \text{(K-9)}$$

and

$$I_n(x_i,y_j) = \log\left(\frac{P(x_i|y_j)}{P(x_i)}\right) \quad \text{(K-10)}$$

K.1.5 MUTUAL INFORMATION

The definition for entropy H(X) in Equation K-2 is the expected value of the self information $I(x_i)$. Similarly, we can define the mutual information to be the expected net information or average information exchanged:

$$I(X;Y) = E\{I_n(x_i,y_j)\} \quad \text{(K-11)}$$

Thus,

$$I(X;Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} P(x_i,y_j) \log\left(\frac{P(x_i|y_j)}{P(x_i)}\right) \qquad \text{(K-12)}$$

Other forms are, patterned after Equation K-10,

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \qquad \text{(K-13)}$$

and, patterned after Equation K-8,

$$I(X;Y) = H(X) - H(X|Y)$$
$$((X;Y) = H(Y) - H(Y|X) \qquad \text{(K-14)}$$

The interrelationships between the quantities in Equations K-13 and K-14 are diagrammed in Figure K-2. If we consider x_i to be the input and y_j the output of the communications channel in Figure K-1, then:

- $I(X;Y)$ is the average information exchanged or the mutual information,
- $H(X)$ is the entropy or average self information of the input, also referred to as the a priori entropy,
- $H(X|Y)$ is the equivocation or average loss of information about the input given the output of the noisy communications channel, also referred to as the a posteriori entropy or average uncertainty about the input given the output,
- $H(Y|X)$ is the average error in the output information given the input, and
- $H(X,Y)$ is the total entropy or average information associated with the channel.

The maximum possible value of $I(X;Y)$ is called the channel capacity, C.

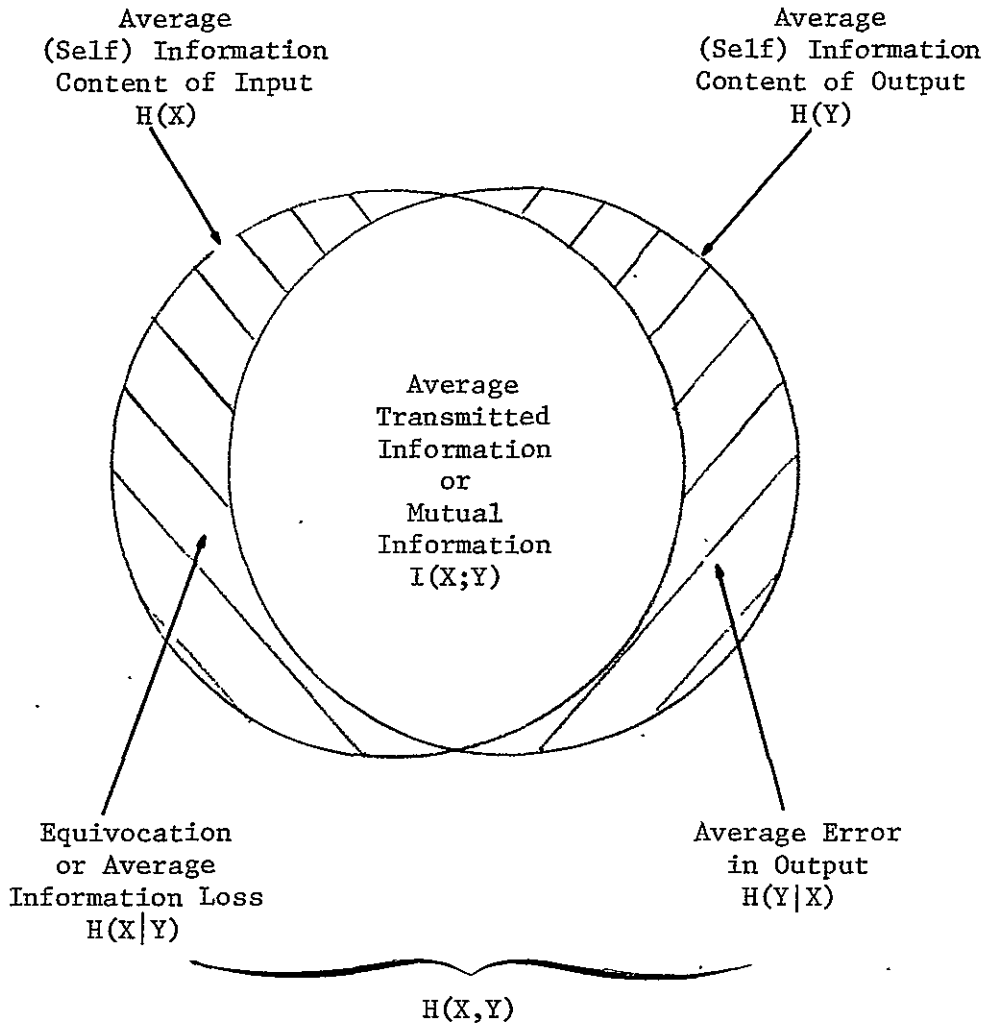


FIGURE K-2. DIAGRAMMATIC REPRESENTATION OF AVERAGE INFORMATION RELATIONSHIPS FOR TWO VARIABLES

The relationships discussed herein assume that selection of any given input state is not influenced by any previous state selection, i.e., that the source does not possess memory. If this is not true, more involved relationships apply.

K.2 LINKING INFORMATION THEORY CONCEPTS TO AREA ESTIMATION SYSTEMS EMPLOYING REMOTELY SENSED DATA

Let us begin by adapting the communications channel concept of Figure K-1 to elements of systems that employ and extract information from remotely sensed data for use in area estimation. To do so, interpret the communications channel itself as consisting of the information extraction and processing operations which are conducted to derive or estimate various characteristics of the observed scenes. For any desired output quantity, there are appropriate input characteristics and processing operations, as shown in Figure K-3. The desired output may vary, from a single crop proportion estimate over a worldwide production area or subunit, on one hand, to a detailed pixel-by-pixel or hectare-by-hectare map of crop production in selected agricultural areas on the other hand.

The underlying idea proposed here is to use information theory approaches to measure system performance by finding the appropriate pair of input-output quantities and considering the intermediate processing to be a communications channel. The less closely the derived output matches the input characteristic being estimated, the "noisier" is the processing channel. Information theory provides quantitative measures of this degradation or failure of the estimation system to duplicate the true input information.

Sections K.3 and K.4 discuss the problem of measuring the performance of a stratification procedure, whether it be the spatial-spectral stratification performed by the BLOB algorithm, the spectral stratification performed by BCLUST, or the classifier stratification and adjustment

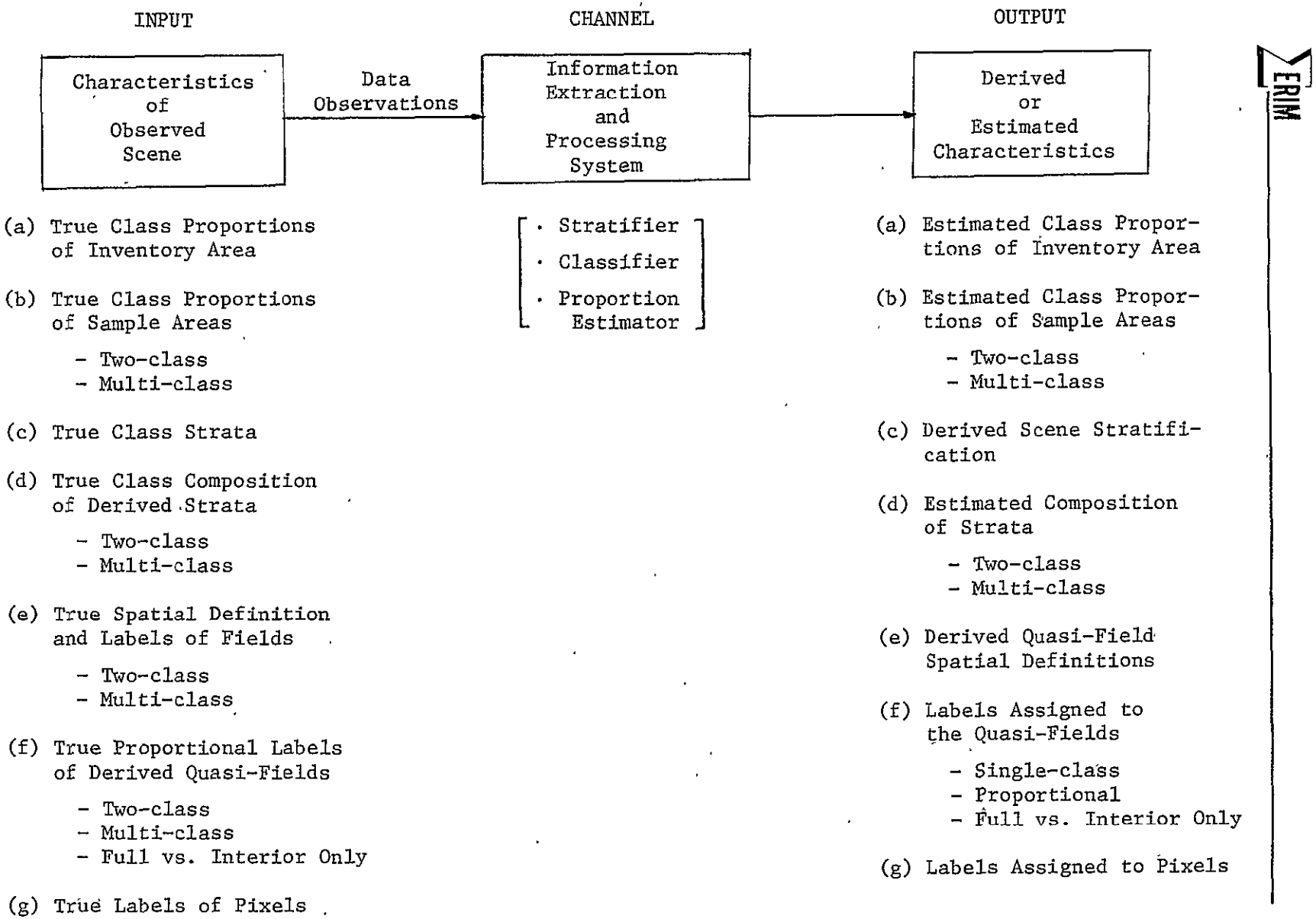


FIGURE K-3. COMMUNICATIONS-CHANNEL VIEW OF INFORMATION EXTRACTION AND AREA ESTIMATION SYSTEMS USING REMOTELY SENSED DATA

performed in Procedure 1. In order to compute information theoretic quantities, it is necessary to substitute stratum proportions for state probabilities in the formulas.

As a digression to indicate relative magnitudes of information quantities, let us consider a 5x6-mile LACIE segment which contains 22,932 Landsat pixels. In principle, each pixel could have a unique spectral vector and represent a unique scene class. In that case, the maximum entropy would be:

$$H_{max} = \log_2 (22,932) = 14.485 \text{ bits/pixel}$$

At the other extreme, if the segment's pixels were either wheat or non-wheat, the maximum entropy would be:

$$H_{max} = \log_2 (2) = 1 \text{ bit/pixel}$$

Between these extremes, we have roughly 1000 quasi-fields per segment, giving

$$H_{max} = \log_2 (1024) = 10 \text{ bits/pixel}$$

and 32 to 128 spectral strata, yielding

$$H_{max} = \log_2 (128) = 7 \text{ bits/pixel}$$

and

$$H_{max} = \log_2 (32) = 5 \text{ bits/pixel}$$

K.3 DERIVATION OF A FIGURE OF MERIT FOR STRATIFICATION

Consider the problem of evaluating the performance of the BLOB algorithm in delineating quasi-fields in the image data. The output characteristic here is the resultant grouping or stratification of pixels into quasi-fields. The input characteristic is the true labeling

of the pixels in "ground-truth" data associated with each pixel. Ideally, one would want pure or single-class quasi-fields to be defined. The extent to which the stratification algorithm achieves this purity is reflected in the mutual information of the above-described input and output.

As was shown in the preceding section, the mutual information $I(X;Y)$ depends on the average self information or entropy $H(X)$ of the input. This entropy will vary from scene to scene and segment to segment. Mutual information may be the appropriate measure for multi-segment or multi-area analyses. For other purposes, such as within segment comparison of processing techniques, however, normalization may be desirable. Therefore, the following figure of merit is defined:*

$$M_X = \frac{I(X;Y)}{H(X)} = 1 - \frac{H(X|Y)}{H(X)} \qquad \text{(K-15)}$$

$H(X|Y)$ measures the average loss of information about X in passing through the system.

If the stratification gives no information about the classes being observed, i.e., there is a total information loss, then $H(X|Y) = H(X)$ and $M_X = 0$. On the other hand, if X is determinable exactly from the stratification Y, there is no transmission loss and no remaining uncertainty so $H(X|Y) = 0$ and $M_X = 1$. Thus, one interpretation of M_X is as the transmittance factor for input information through the processing system.

In expanded form, the figure of merit is:

$$M_X = 1 - \frac{\sum_{j=1}^{n} P(y_j) \sum_{i=1}^{m} P(x_i|y_j) \log[P(x_i|y_j)]}{\sum_{i=1}^{m} P(x_i) \log[P(x_i)]} \qquad \text{(K-16)}$$

*Figures of merit based on other normalization factors are discussed in Section K.4.

A two-class version of Equation K-16 is helpful for many of our current studies. For convenience, it is expressed here in terms of the wheat and nonwheat classes:

$$M_X = 1 - \frac{\sum_{j=1}^{n} \left(\frac{n_j}{N}\right) \left[p_{W_j} \log p_{W_j} + p_{\overline{W}_j} \log p_{\overline{W}_j} \right]}{P_W \log P_W + P_{\overline{W}} \log P_{\overline{W}}} \qquad \text{(K-17)}$$

where

n_j = number of pixels in Stratum j,

N = total number of pixels being considered,

p_{W_j} = proportion of wheat in Stratum j,

$p_{\overline{W}_j}$ = proportion of nonwheat in Stratum j,

P_W = overall proportion of wheat,

and

$P_{\overline{W}}$ = overall proportion of nonwheat.

Calculation of the figure of merit M_X was incorporated into one of the programs used to evaluate the two-class performance of the BLOB algorithm in our Fall 1978 test of Procedure M [3]. Table K-1 lists results obtained for 27 segments, along with the corresponding variance reduction factors and average blob purity percentages. Values are included separately for full big blobs and for their interior pixels.

Figure K-4 presents a scatter diagram of variance reduction factor vs. figure of merit values for blob interiors. Note the high degree of correlation between them, except for a few outliers. These outliers are

TABLE K-1. RESULTS OBTAINED IN 1978 TEST OF PROCEDURE M

		FULL "BIG" BLOBS			BLOB INTERIORS		
0. CASE#	1. SEGMENT	2. RV,FULL	3. FOM,FULL	4. PUR,FULL	5. RV,INT	6. FOM,INT	7. PUR,INT
1	1104.0	.30800	.73000	97.210	.10100	.90900	98.250
2	1498.0	.33300	.63200	90.270	.13600	.83200	95.270
3	1512.0	.37000	.58100	88.710	.12000	.86500	96.000
4	1513.0	.21500	.75700	94.140	.22000 -1	.97900	99.370
5	1515.0	.27900	.69300	90.770	.64000 -1	.94500	97.620
6	1520.0	.39600	.60100	91.090	.13700	.84100	96.110
7	1602.0	.47300	.49400	83.720	.18800	.80800	93.920
8	1606.0	.35200	.61600	88.890	.10700	.85700	96.110
9	1614.0	.53800	.38800	81.230	.25300	.71600	91.070
10	1625.0	.57800	.40600	83.820	.32100	.76600	92.990
11	1633.0	.25100	.69600	91.660	.76000 -1	.91200	97.190
12	1637.0	.34100	.60600	89.270	.10200	.88600	96.410
13	1640.0	.35600	.58800	87.550	.11400	.86400	95.980
14	1652.0	.53600	.41000	80.450	.35300	.62000	85.390
15	1662.0	.39100	.55900	86.310	.16100	.85000	94.410
16	1663.0	.22700	.73200	91.920	.26000 -1	.96800	98.770
17	1669.0	.51900	.54300	92.590	.44200	.77500	95.720
18	1681.0	.28500	.67100	90.910	.71000 -1	.90900	97.170
19	1699.0	.27700	.68500	93.660	.82000 -1	.90300	97.640
20	1800.0	.37100	.58400	89.970	.10900	.85600	96.410
21	1803.0	.50100	.60700	98.680	.38900	.91800	98.960
22	1805.0	.52700	.46700	90.010	.33300	.68900	93.950
23	1811.0	.34900	.75800	97.930	.11500	.92500	98.720
24	1899.0	.26400	.69700	91.290	.43000 -1	.95800	98.440
25	1913.0	.43900	.40700	92.150	.20600	.78300	96.700
26	1927.0	.29600	.65500	91.240	.99000 -1	.87200	96.640
27	1929.0	.74200	.23900	67.760	.42000	.59600	82.310

Notes: RV = Reduction of Variance Factor
FOM = Figure of Merit, M_X
PUR = Average Purity (Percent)

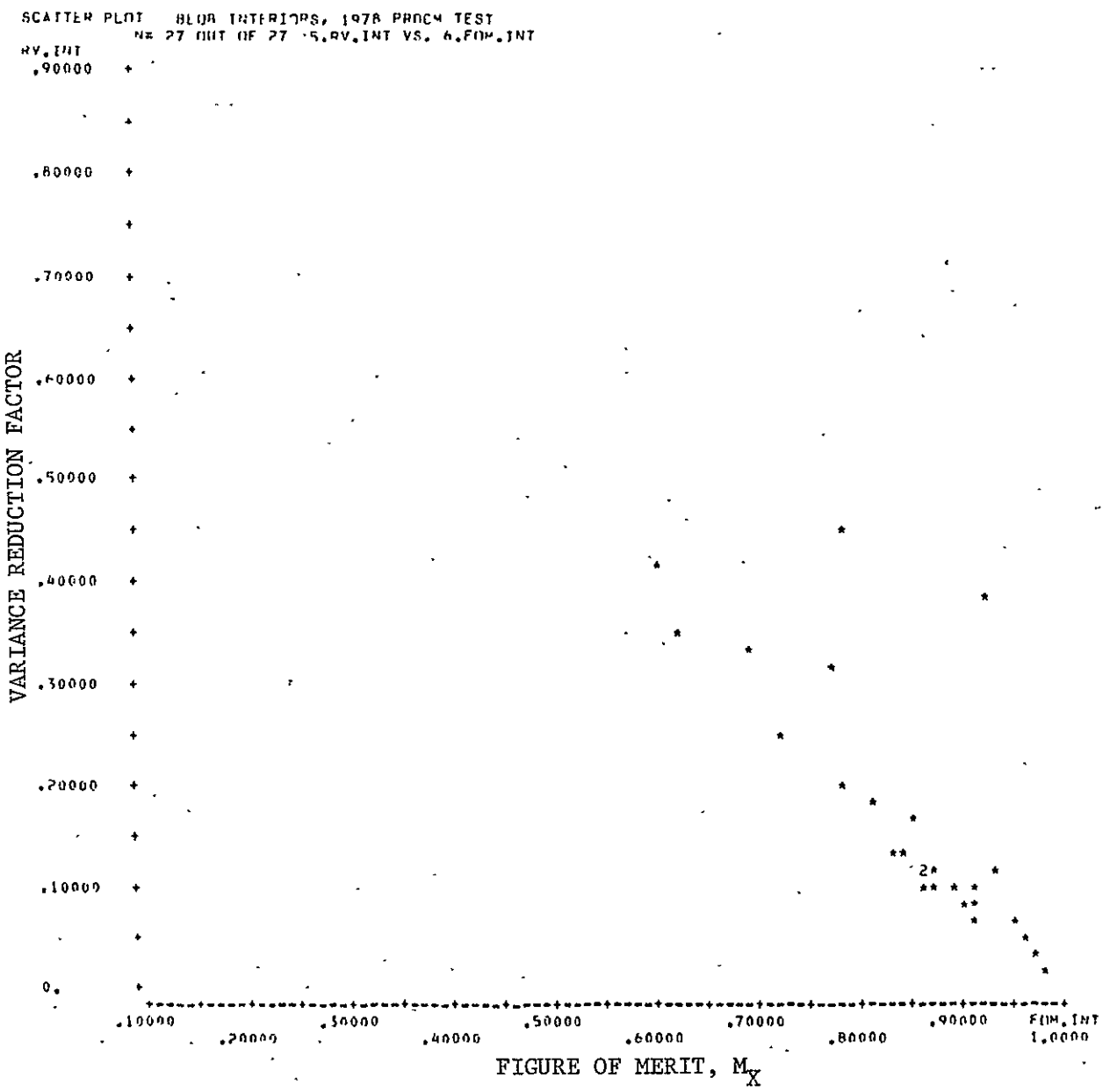


FIGURE K-4. COMPARISON OF VARIANCE REDUCTION FACTOR AND FIGURE OF MERIT FOR BLOB INTERIORS

segments with low proportions of spring small grain, in which case the entropy-based figure of merit appears to give more consistent results. Comparable correlations with average blob purity also were observed.

K.4 ADDITIONAL STRATIFICATION PERFORMANCE MEASURES

In stratification, the true class associated with each output stratum is not of importance; only the stratum purity is of concern. For example, whenever all members of stratum y_j are of the same input class, $P(x_i|y_j)$ is unity for one of the i's, say i = I, giving $\log P(x_I|y_j) = \log (1) = 0$, so $H(x_I|y_j) = 0$ irrespective of the value of $P(x_i,y_j)$. For each other $i \neq I$, $P(x_i,y_j) = 0$ and $H(x_i|y_j) = 0$, so $H(X|y_j) = \sum_i H(x_i|y_j) = 0$. Thus with pure strata, H(X|Y), the equivocation or average expected loss of information, is zero. As a result, M_X is a maximum, $M_X = 1$, because the integrity of the input classes has been preserved by the stratification procedure. There is, however, one important characteristic that is not measured by M_X.

The unmeasured characteristic is the number of strata defined at the output. As noted above, when all strata are pure, H(X|Y) is zero and $M_X = 1$; this is irrespective of the number or size of the strata. An extreme case would be if each scene element or pixel were assigned to a unique stratum and each and every pixel were pure, i.e., belonged to only one of the input classes; here again M_X would be unity, an undesirable value for a performance measure in this situation.

The information content (self information) of the output is measured by the output entropy H(Y) which <u>does</u> depend on the number of output strata. This suggests a normalization by H(Y) instead of H(X), giving rise to:

$$M_Y = \frac{I(X;Y)}{H(Y)} = \frac{H(Y) - H(Y|X)}{H(Y)} = 1 - \frac{H(Y|X)}{H(Y)} \qquad \text{(K-18)}$$

Parts (a) and (b) of Figure K-5 illustrate the relationships between quantities used to define the figures of merit M_X and M_Y. Referring to the diagrams in Figure K-5, it is obvious that M_Y (as well as the other figures of merit) has a maximum value of unity and a minimum value of zero. To better understand the maximum of M_Y, consider the following version:

$$M_Y = \frac{H(X) - H(X|Y)}{H(Y)} \tag{K-19}$$

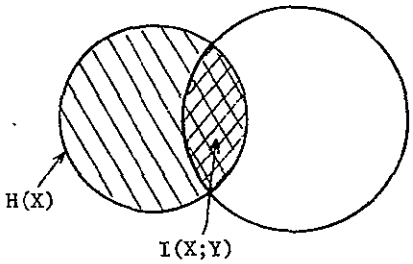
If all output strata are pure, $H(X|Y) = 0$, resulting in $M_Y = \frac{H(X)}{H(Y)}$. This quantity is less than unity, except when $H(Y) = H(X)$, i.e., when these pure output strata exactly match the input classes in number and size. The minimum value, $M_Y = 0$, results when the output Y is independent of the input X and $H(Y|X) = H(Y)$ (See Equation (K-18)).

Weighted-sum normalization factors were suggested as possibilities (See Figure K-5(d)), arising out of a concern that M_Y might impart too severe a penalty on a system, as a function of the number of output strata produced. Figure K-6 illustrates with solid lines the relationships for two input classes of equal size and 2^n output classes of equal size. A pronounced decrease in the M_Y curve is also present as the average purity of output strata decreases, as shown for 95% and 90% purity strata. Since M_X remains constant at the value for two output strata, division by a weighted sum of H(X) and H(Y) would decrease the rate of falloff from that shown in Figure K-6. However, a suitable criterion for selecting the weighting factor ω has not yet been found.

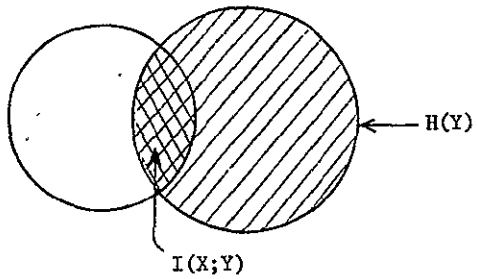
The final normalization factor considered herein is H(X,Y):

$$M_{X,Y} = \frac{I(X;Y)}{H(X,Y)} = \frac{H(X) - H(X|Y)}{H(Y) + H(X|Y)} \tag{K-20a}$$

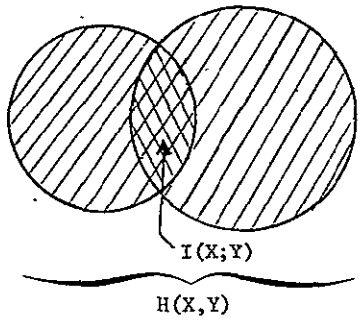
$$M_{X,Y} = \frac{H(Y) - H(Y|X)}{H(X) + H(Y|X)} \tag{K-20b}$$



(a) INPUT ENTROPY $\Longrightarrow$ $M_X = \frac{I(X;Y)}{H(X)} = 1 - \frac{H(X|Y)}{H(X)}$



(b) OUTPUT ENTROPY $\Longrightarrow$ $M_Y = \frac{I(X;Y)}{H(Y)} = 1 - \frac{H(Y|X)}{H(Y)}$



(c) TOTAL ENTROPY $\Longrightarrow$ $M_{X,Y} = \frac{I(X;Y)}{H(X,Y)} = \frac{H(X) - H(X|Y)}{H(Y) + H(X|Y)}$

[NO DIAGRAM]

(d) WEIGHTED SUM OF ENTROPIES $\Longrightarrow$ $M_{X,Y,\omega} = \frac{I(X;Y)}{\omega H(X) + (1-\omega)H(Y)}$

FIGURE K-5. ILLUSTRATION OF VARIOUS NORMALIZATION FACTORS

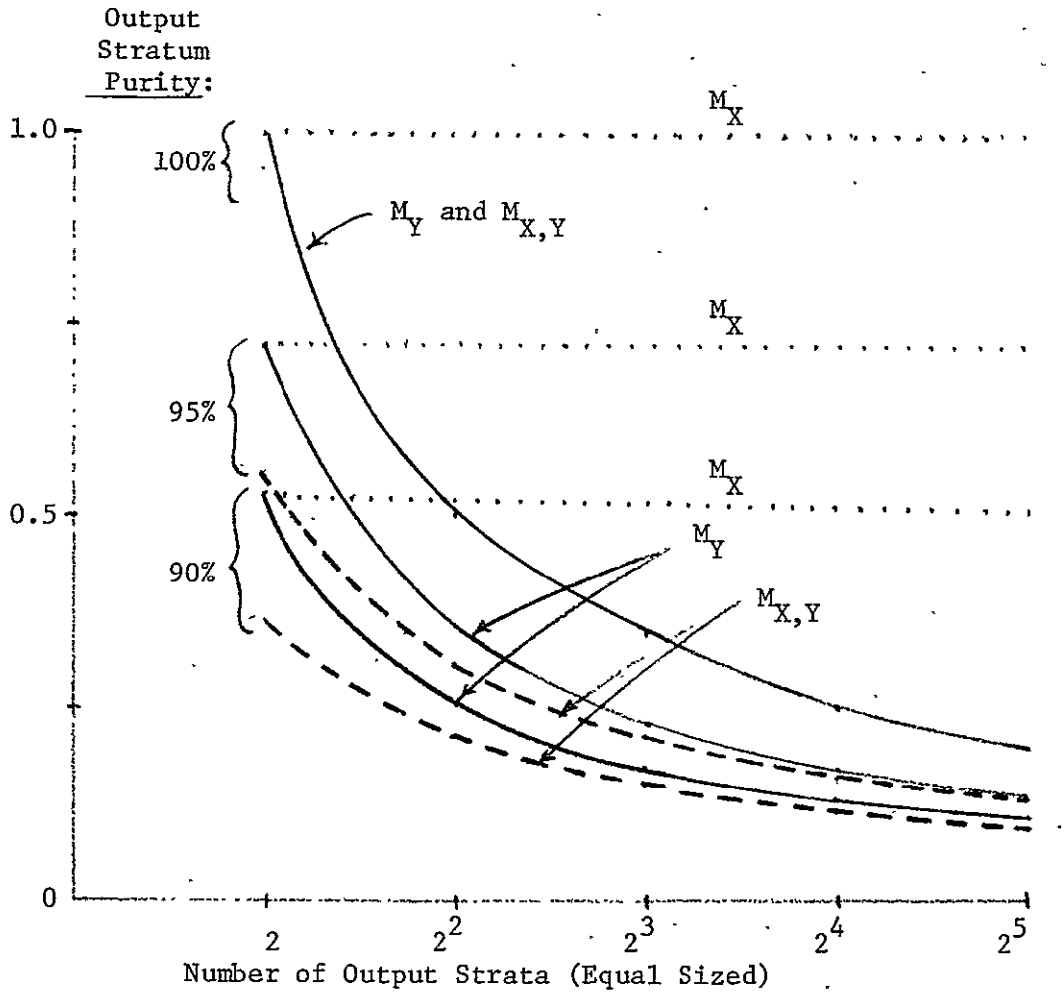


FIGURE K-6. COMPARISON OF INFORMATION THEORETIC FIGURES OF MERIT FOR STRATIFICATIONS OF TWO EQUALLY LIKELY INPUT CLASSES

or

$$M_{X,Y} = \frac{I(X;Y)}{H(X) + H(Y|X)} \qquad \text{(K-20c)}$$

An appealing feature of this figure of merit is the symmetry exhibited by Equations K-20a and K-20b. Upon comparing Equation K-20a to Equation K-19, one also sees that $M_{X,Y}$ decreases more rapidly with increasing output entropy than does M_Y, whenever $H(X|Y) \neq 0$; see dashed lines in Figure K-6. Both numerator and denominator are affected by the information loss. When $H(X|Y) = 0$ and there is no information loss, $M_{X,Y} = \frac{H(X)}{H(Y)} = M_Y$ and prior comments apply.

Another insight into this last figure of merit is facilitated by Equation K-20c. The denominator represents the sum of two terms, the entropy or self information of the input and the error associated with the output stratification, i.e., the uncertainty beyond that associated with X. This is the total entropy of the processing channel and equals the input entropy only when the error is zero.

The relationships in Equation K-20 also have implications for sampling requirements to identify or determine which input class(es) should be associated with each output stratum, e.g., for proportion estimation. If the strata were pure and the label determination perfect, only one sample per stratum would be required. The amount of information (in bits) acquired from each sample would differ, however, according to the size of its stratum, but summed together they would represent the output entropy H(Y). Yet as can be seen in Equation K-20a, there is an additional term paired with H(Y) in the denominator, i.e., $H(X|Y)$ which is the equivocation or average expected loss of information. $H(X|Y)$ reflects the impurity of the output strata. One way to recover this lost information would be by additional sampling so that proportional labels could be assigned to some strata. $H(X|Y)$ would appear to represent the needed information content of additional samples, but relationships between entropy and sampling have not yet been established.

To summarize this section, it appears that a figure of merit for evaluating the performance of stratifiers should include the output entropy in the normalization factor so as to penalize a stratifier with too many strata. Several candidates were discussed, with $M_{X,Y}$, which normalizes by the total entropy, having several desirable features. Testing and evaluation should be conducted to make a choice. Also, relationships between sampling and information measures need study and development.

K.5 SUMMARY AND DISCUSSION

The concept of relating information theory concepts to evaluation of area estimation systems employing processed remotely sensed data has been developed and presented. New figures of merit have been defined for comparing the relative performances of various stratification, clustering, and classification operations.

One figure of merit M_X measures the transmittance of scene information by the processing system to the user. Example use in Blob performance evaluation was presented, with preliminary indications of more consistency than the variance reduction factor -- more analysis is required to be definitive.

The approach may be even more appropriate for the evaluation of spectral stratification techniques, since spatial entities (and, therefore, the input entropy, H(X)) will remain constant from technique to technique. However, use of the $M_{X,Y}$ figure of merit would be advantageous since it would penalize a stratifier with a large number of strata.

Information theory measures may well not be the complete answer, but they appear to have potential. Therefore, it is recommended that their utility in measuring system and component performance be investigated and evaluated as supplements to current measures, such as variance reduction factor, proportion bias, and probability of correct classification. The figures of merit defined here provides a starting point; additional uses or measures may result from analysis of their use, advantages, and disadvantages.

APPENDIX L

DESCRIPTION OF DATA BASE FOR SMALL GRAINS

The purpose of this section is to extend the data base discussion begun in Section 7, and specifically to present the segments selected, to describe the processing algorithms used, and to indicate the form of the data. Much of the discussion of Section 7 is repeated to make this appendix self-contained.

The data base consists of 67, 5x6-mi segments located throughout the United States Great Plains during the 1976 and 1977 winter and spring wheat growing seasons. Each segment consists of data from all available useable acquisitions (average 8) of Landsat 1 and Landsat 2, merged together with wall-to-wall ground truth inventory data supplied by USDA and prepared by JSC/LEC. Important pixel-by-pixel information generated during processing, such as cloud/water/shadow identifications or blob numbers, were retained with each segment.

In order to support the variety of tasks depending on this data base, segments were selected after initial screening into two categories, A and B. The principal differences between the two categories were in the segment selection procedures and the state to which the segments were processed. For some applications it is appropriate to use Category A segments for development and training, and Category B segments for test and evaluation.

The segment selection procedures used were as follows. First the 171 segments available to us were screened to eliminate those whose acquisition history, ground truth quality, and data quality are inadequate. When carried out, 107 segments remained for possible selection. Next, Category B segments were selected by randomly picking a quota (1/3) of segments within each APU* and each year, subject to the constraint

* Agrophysical Unit, as used within LACIE.

that at least one segment is selected from each APU. Since there is data from two growing seasons, for which different conditions prevailed, the APU's were considered distinct from one growing season to the next when establishing this random draw. By this procedure, 38 segments were selected as a uniform draw that tended to cover the variability of the region in a carefully drawn statistical sample.

The remaining segments were considered in picking Category A segments. Stricter minimum limits on acquisition history and data quality were observed, so that selected segments were well suited for development work. However, the selections were more qualitative, taking into account desire for more acquisitions versus desire for a nice spread to cover variability. Even though selecting Category A segments first would have resulted in more and nicer segments, it was necessary to select them second in order to maintain the statistical integrity of the Category B sample.

The segments selected as described above are presented in Tables L-1 and L-2. Location, strata, and other information is include as well.

A set of flow diagrams that summarized the processing carried out on each segment is given in Figure L-1. Category B segments are processed using all steps except SUPERB, STRIP and COMPRS. A one-paragraph description of each processing step and related comments is given below.

CONVRT performs a reformatting of Accuracy Assessment 1-channel ground truth data in 2x3 subpixel form into standard pixel form with 7 channels (summary code + 6 subpixel codes for each pixel).

MERGE reads ground truth data and several single-acquisition Landsat files and writes a single file containing merged data.

SCREEN flags data as clouds, water, shadows, wild, etc, if not within a predefined envelope for agricultural data. One channel is added for each acquisition.

TABLE L-1. CATEGORY A SITES IN THE SMALL GRAINS DATA BASE

Site	State	County	APU	Phase	Latitude	Longitude	Number of Acquisitions
1005	CO	Cheynne	10	III	38.49	102.20	9
1035	KS	Ford	11	II	37.44	99.58	14
1041	KS	Meade	8	II	37.19	100.16	10
1059	TX	Ochiltree	4	III	36.15	100.52	7?
1099	CO	Baca	9	III	37.25	102.18	10
1158	KS	Washington	12	III	39.47	97.06	7
1175	KS	Sedgwick	7	III	37.46	97.32	6
1178	KS	Bourbon	7	II	37.43	94.59	7
1506	CO	Sedgwich	10	III	40.51	102.31	11
1512	MN	Clag	15	III	47.01	96.22	5
1523	MN	Cokkin	20	III	46.31	96.25	7
1566	NB	Kimball	15	III	41.02	103.43	10
1586	NB	Perkins	103	III	40.46	101.20	13
1606	ND	Ward	19	III	48.16	101.22	4
1619	ND	Grand Forks	20	III	48.04	97.30	6
1637	ND	Stutsman	21	II	47.15	99.19	6
1640	ND	Barnes	19	III	46.55	97.51	9
1642	ND	Cass	20	II	46.44	97.34	9
1645	ND	Traill	20	II	47.33	96.56	9
1648	ND	Recoman	19	III	46.04	103.06	5
1652	ND	Stark	20	III	46.56	102.50	5
1662	ND	Ransom	19	II	46.24	98.02	9
1681	SD	Roberts	21	III	45.32	96.49	8
1807	SD	Bonhomme	18	III	43.03	97.57	6
1852	KS	Lane	11	II	38.30	100.26	12
1857	KS	Grant	9	II	37.41	101.08	15
1860	KS	Hodgeman	11	II	38.08	99.42	13
1865	KS	Stevens	14	II	37.12	101.09	11
1887	KS	Russel	11	II	39.07	98.55	7

TABLE L-2. CATEGORY B SITES IN THE SMALL GRAINS DATA BASE

Site	State	County	APU	Phase	Latitude	Longitude	Number of Acquisitions
1000	CO	Logan	10	III	40.34	102.54	13
1008	CO	Kit Carson	10	III	39.33	102.19	5
1049	OK	Texas	9	III	36.46	101.20	9
1060	TX	Sherman	9	III	36.22	101.41	4
1102	MT	Yellowstone	104	III	45.57	108.20	8
1154	KS	Mitchell	12	II	39.25	98.00	10
1163	KS	Coffey	8	II	38.15	95.38	5
1165	KS	Lina	14	II	38.10	94.53	12
1167	KS	Shawnee	14	II	39.04	95.43	6
1169	KS	Commanche	60	II	37.00	99.16	6
1179	KS	Butler	7	II	37.35	97.00	4
1181	KS	Cowley	7	II	37.17	96.51	4
1242	OK	Canadian	7	III	35.27	98.50	6
1355	OK	Beaver	7	III	36.35	100.00	7
1489	SD	Walworth	19	III	45.29	99.42	3
1498	SD	Codington	16	III	44.57	97.02	8
1513	MN	Kittson	15	III	48.52	97.06	4
1521	MN	Grant	20	III	46.05	96.01	5
1560	NB	Banner	10	III	41.29	104.00	10
1568	NB	Sheridan	15	III	42.21	102.41	7
1576	NB	Lancaster	14	III	40.52	96.50	7
1602	ND	Mountain	21	III	48.21	102.25	6
1614	ND	Pierce	19	II	48.31	100.10	5
1618	ND	Grand Forks	20	II	47.44	97.32	10
1661	ND	McIntosh	21	II	46.16	99.45	6
1663	ND	Richland	20	III	46.23	96.44	11
1675	SD	McPherson	21	III	45.56	99.13	5

TABLE L-2. CATEGORY B SITES IN THE SMALL GRAINS DATA BASE (Cont'd)

Site	State	County	APU	Phase	Latitude	Longitude	Number of Acquisitions
1677	SD	Spinks	19	III	45.04	98.06	7
1699	SD	Hyde	16	III	44.28	99.27	10
1739	MT	Teton	104	III	47.45	111.30	10
1800	SD	Milook	19	III	43.41	97.24	7
1851	KS	Graham	9	II	39.33	99.57	11
1861	KS	Kearny	9	II	38.12	101.25	13
1886	KS	Rush	8	II	38.30	99.20	9
1891	KS	Reno	12	II	37.58	98.25	13
1913	ND	Hettinger	21	III	46.33	102.47	7

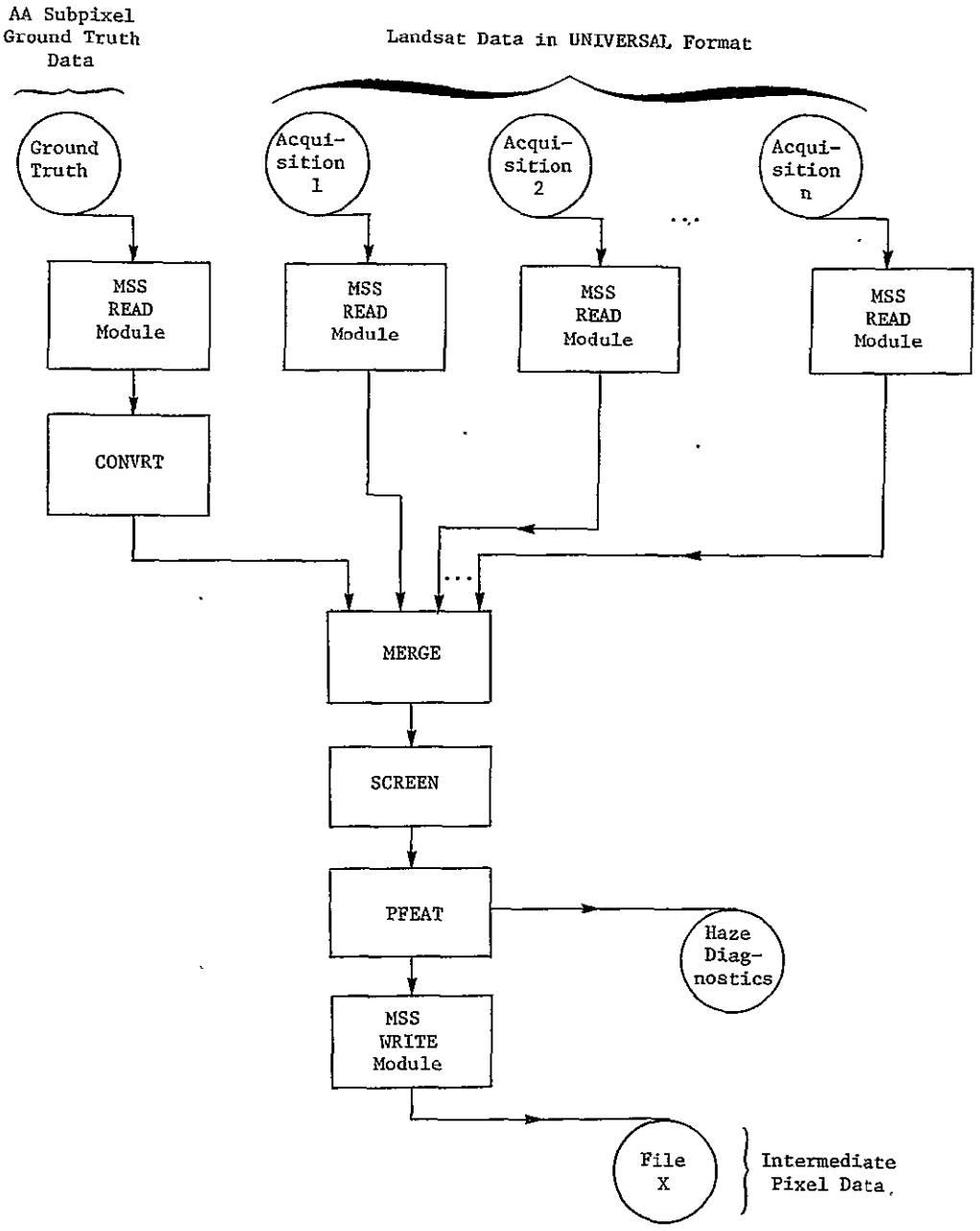


FIGURE L-1. FLOW OF PROCESSING CARRIED OUT FOR EACH SEGMENT

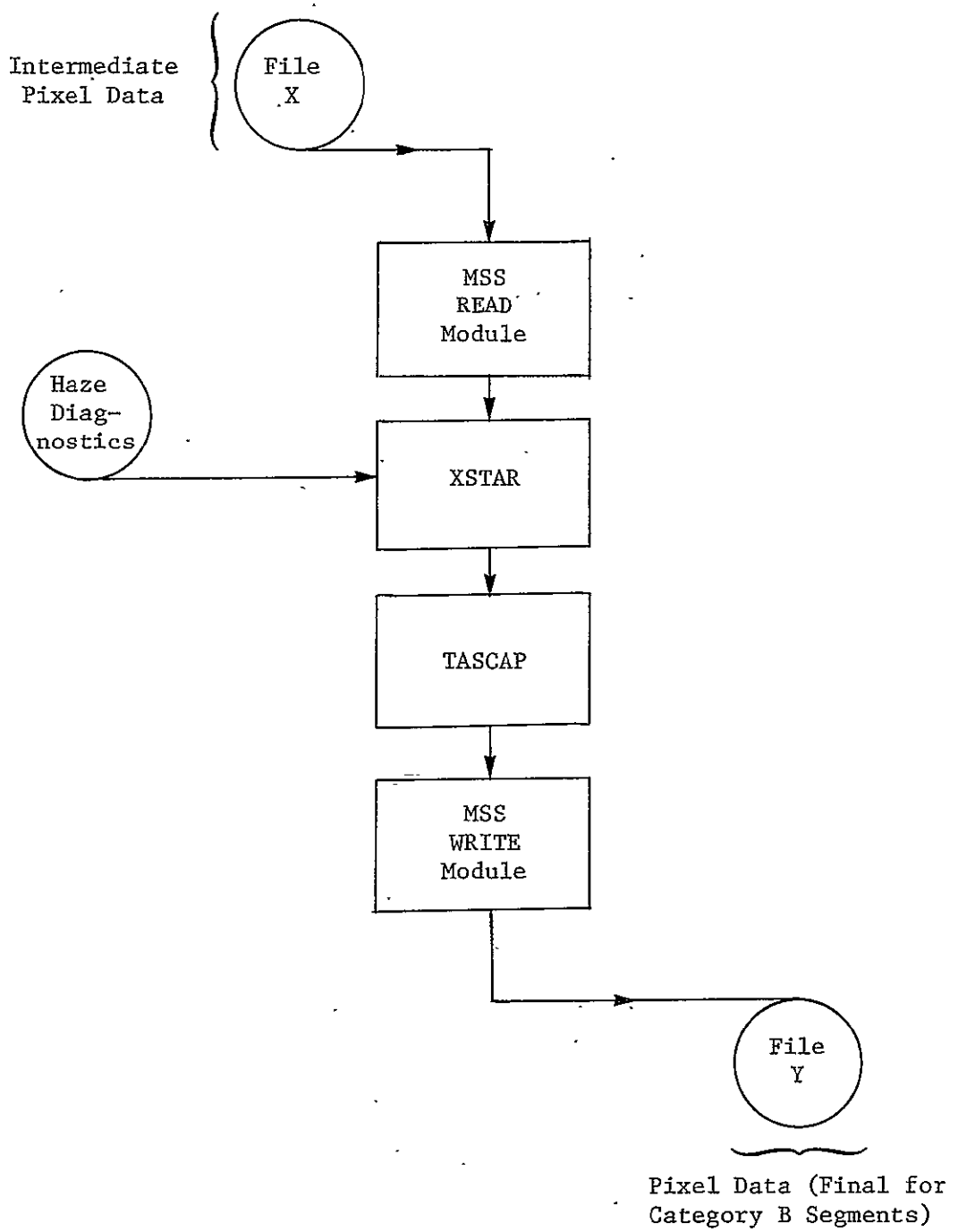


FIGURE L-1. FLOW OF PROCESSING CARRIED OUT FOR EACH SEGMENT (Cont'd)

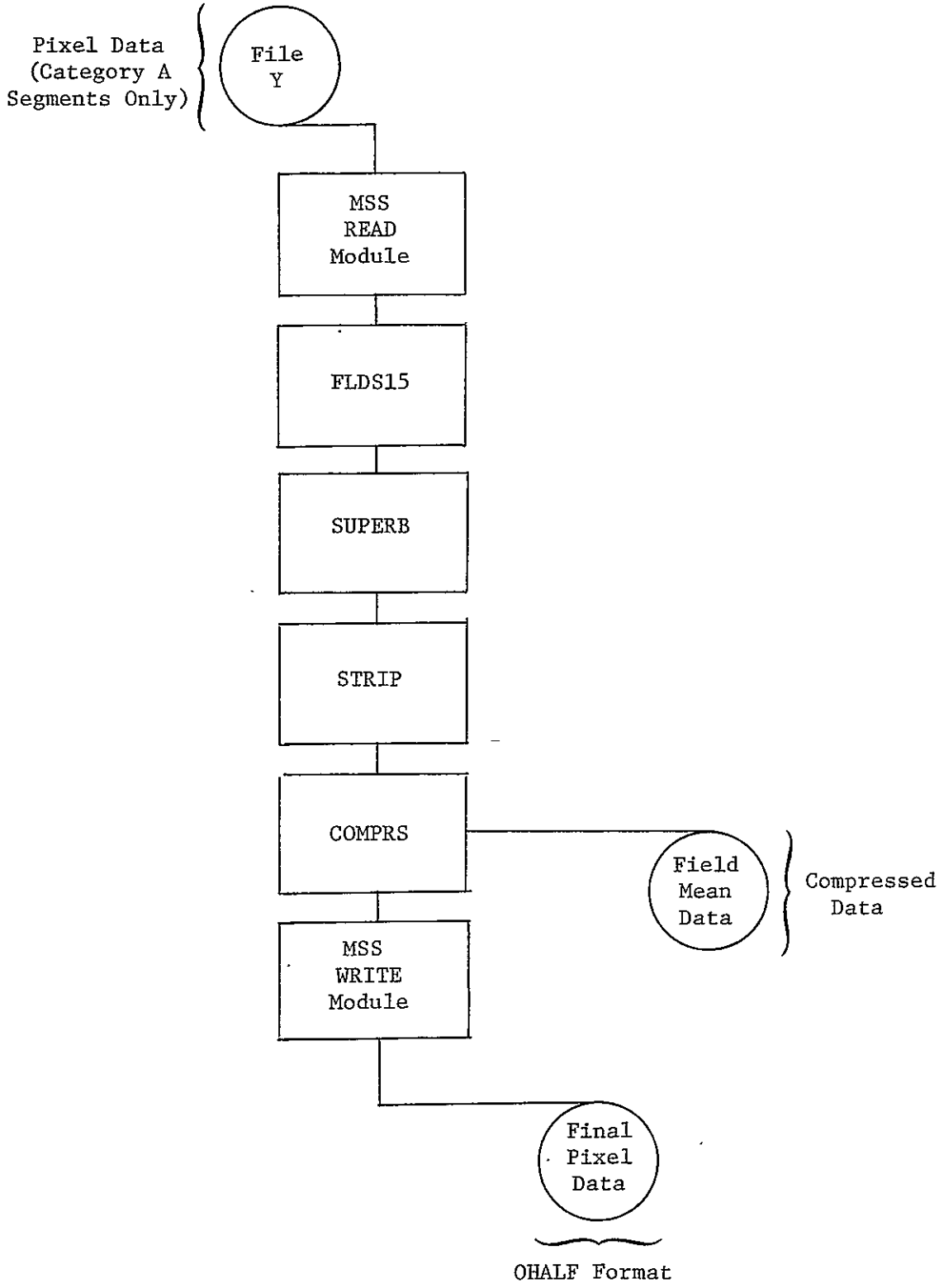


FIGURE L-1. FLOW OF PROCESSING CARRIED OUT FOR EACH SEGMENT (Cont'd)

PFEAT calculates features (haze diagnostics) in support of a spatially varying XSTAR algorithm.

XSTAR carries out a spatially varying haze correction on each acquisition. It also applies sun angle (cosine) and satellite calibration corrections to the data.

TASCAP performs a linear transformation on the data that for each acquisition replaces the four Landsat bands with four Tasseled Cap features.

FLDS15 modifies ground truth codes referring to special fields, when necessary, so that they reflect the correct ground truth class of the fields.

SUPERB, described in Appendix D, was used to identify quasi-fields that are spectrally homogeneous and spatially contiguous field patterns that obey existing ground truth boundaries.

STRIP identifies pixels on quasi-field edges. A pixel is flagged as edge if any of its four strong neighbors belong to a different blob.

COMPRS computes and outputs spectral and spatial mean values for each quasi-field, as a compressed data set. This compressed data set can be used economically for development tasks that require many passes through the data.

For each segment, two output data products are made part of the data base. These are pixel data and compressed data (field means). Category B segments are not processed through BLOB or COMPRS, and do not have compressed data available. For these segments, the final pixel data used in the data base is that resulting from XSTAR and TASCAP. It is in the same form as pixel data for Category A segments, except that channels reserved for BLOB and STRIP results are zeroed.

The pixel data for a segment consists of 22,932 pixels organized into 117 scan lines. It contains (16+5n) channels, where n is the number of Landsat acquisitions processed for the segment. The channel assignments, detailed in Table L-3, were designed to allow a variable number of acquisitions yet provide consistent locations for most channels.

Due to the number of processing steps that transform the pixel data, it was necessary to avoid the possibility that significance error would creep in when numerical values are converted to integer form between steps. Thus, pixel values for some channels are maintained in floating point form, both during processing and when stored in an external format.

For this reason, and because the largest number of channels we used for one data set is 91, the data was stored in an ERIM-defined format, rather than in a more widely known format such as UNIVERSAL. The name of the format used for pixel data is OHALF.

The second data product used in the data base is compressed data consisting of blob means produced by routine COMPRS. This product consists of one record for each blob, containing the channels described in Table L-4. The channel values within each record are stored in 4-byte binary floating point form.

TABLE L-3. CHANNEL ASSIGNMENTS FOR PIXEL DATA USED IN THE SMALL GRAINS DATA BASE

Channel	Explanation
1	Ground truth summary code (zero or equal to unanimous subpixel codes)
2-7	Subpixel ground truth codes in the order: upper-left, middle-left, lower-left, upper-right, middle-right, lower-right)
8-9	Blob number (is 256* chan 8 + chan 9)
10	Blob ground truth code (or zero if ground truth is ambiguous or unknown)
11	Strip channel (1 if pixel is on edge of blob, 0 otherwise)
12-16	Reserved for future processing results
17	Tasseled Cap Brightness component, acquisition 1
18	Tasseled Cap Greenness component, acquisition 1
19	Tasseled Cap Yellowstuff component, acquisition 1
20	Tasseled Cap Nonesuch component, acquisition 1
(13+4i) thru (16+4i)	Tasseled Cap components for acquisition i in the order given above for i=1.
(16+4n+i)	Screen channel for acquisition i, given that there are n acquisitions. Zero means pixel is good in acquisition i.

TABLE L-4. CHANNEL ASSIGNMENTS FOR COMPRESSED (FIELD MEAN) DATA USED IN THE SMALL GRAINS DATA BASE

Channel	Explanation
1*	Segment number
2*	A unique job identifier
3	Blob number
4	Blob type. If 1.0, blob has interior pixels. If 2.0, blob has only edge pixels.
5	Total number of pixels in blob
6	Number of interior pixels
7	Number of pixels used to form the mean (same as Channel 5 if blob has only edge pixels; same as Channel 6 if blob has some interior pixels)
8	Blob ground truth code
9	Midrange of line number coordinates for this blob
10	Midrange of point number coordinates
11-(10+n)	Acquisition date given as number of days since 31 December
(11+n)-(10+2n)	Spectral mean vector

* Same for all blobs

APPENDIX M

NONPARAMETRIC CROP PROPORTION ESTIMATION

Cover and Hart [61] proved an asymptotic property of the nearest neighbor decision rule which raises the question of how this nonparametric classification technique can be extended to a nonparametric crop proportion estimation technique. This section explores relationships between classification and proportion estimation and proposes a proportion estimation procedure based on nearest neighbor classification.

In order to reduce notation, we will restrict the discussion to the two-class problem. The reader can think of them as wheat and nonwheat. Generalization to the multiclass problem offers no conceptual difficulty. Accordingly, let us assume a population with a multivariate mixture density

$$f(x) = pf_1(x) + (1-p)f_0(x)$$

It is convenient to think of the two-dimensional case. Further, we assume a training sample of N points $x_1, x_2, \ldots, x_N$ together with the true classification λ_i of the point x_i, $1 \leq i \leq N$; i.e.,

$$\lambda_i = 1 \text{ if } x_i \text{ is from class 1}$$
$$\lambda_i = 0 \text{ if } x_i \text{ is from class 0}$$

The Nearest Neighbor Classifier

The nearest neighbor (NN) rule classifies a sample point x as being from class 1 if the point closest (Euclidean distance) to x from among the x_i is in class 1, and is classified as from class 0, otherwise. The probability of misclassification for this classification scheme is denoted by R_N, and R is defined by

$$R = \lim_{N \to \infty} R_N$$

Now suppose p, f_1, and f_o were known. Then one could use the optimum Bayes classification rule to minimize the probability of misclassification. Let this probability be denoted by R^*. The Cover and Hart result is that under very general conditions

$$R^* \leq R \leq 2R^*(1-R^*)$$

(Of course the left side inequality is an immediate consequence of the optimality of the Bayes classifier.)

Classifiers and Proportion Estimation

We digress here to discuss the relationship between classification and proportion estimation. Let us suppose we have a classifier such that the probability of correctly classifying a point in class 1 is α and the probability of correctly classifying a point in class 0 is δ. If α and δ are known, then the minimum variance unbiased estimate $\hat{p}$ of the proportion p is

$$\hat{p} = \frac{\tilde{p} - (1-\delta)}{\alpha+\delta-1} \quad \dagger$$

where $\tilde{p}$ denotes the relative number of sample points classified into class 1.

If α and δ are unknown, then, as discussed in Section 6, no unbiased estimate of p exists and the mean square error of any estimate of p is bounded below, independently of sample size, with the bound depending on the ranges assumed for α and δ. Thus the extension of a classifier into a satisfactory proportion estimator depends upon obtaining adequate information about α and δ.

†The reader will recognize the NASA (Feiveson) bias correction factor.

Extension of Nearest Neighbor Classifier to Proportion Estimator

In order to obtain estimates $\hat{\alpha}$ and $\hat{\delta}$ of the conditional probabilities of correct classifications α and δ of the nearest neighbor classifier, we proceed as follows. For each i, $1 \leq i \leq N$, let $\mu_i = 1$ if the point in the sequence $x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_N$ closest to x_i is in class 1, and let $\mu_i = 0$ otherwise. Then set

$$\hat{\alpha} = \left(\sum_{\substack{i \\ \lambda_i=1}} \mu_i\right) \Big/ \sum_{\substack{i \\ \lambda_i=1}} \lambda_i$$

and

$$\hat{\delta} = \left[\sum_{\substack{i \\ \lambda_i=0}} (1 - \mu_i)\right] \Big/ \sum_{\substack{i \\ \lambda_i=0}} (1 - \lambda_i)$$

It appears reasonable to expect $\hat{\alpha}$ and $\hat{\delta}$ to converge (in some sense) to α and δ with increasing N.

Now let $y_1, y_2, \ldots, y_M$ be a sample of data points without true class information available. We refer to the x_i's as the training data, and the y_i's as the test data. We would like to estimate p using both the training data and the test data. Let ν_j, $1 \leq j \leq M$, denote the class assigned to y_j by the nearest neighbor rule. For estimating p, we have the information contained in the x_i with their true classifications λ_i, the y_j with their assigned classifications ν_j, and the estimated conditional probabilities $\hat{\alpha}$ and $\hat{\delta}$ of correct classification of the NN classifier.

If we had the training points x_i with true class given, test points y_j, a classifier of the test points, and conditional probabilities α and δ of the classifier, then we could readily write down the likelihood for the M+N observations

$$(x_1,\lambda_1),\ (x_2,\lambda_2),\ldots,(x_N,\lambda_N);\ (y_1,\nu_1),(y_2,\nu_2)\ldots(y_M,\nu_M)$$

assuming all these observations were independent. The nearest neighbor classifier introduces dependencies between the (y_j,ν_j)'s and the (x_i,λ_i)'s. Furthermore, we only have estimates of α and δ available. Nevertheless, it seems reasonable to proceed to estimate p by using the maximum likelihood procedure for independent observations and exact values of α and δ instead of estimates.

Accordingly, we define the estimate $\hat{p}$ of p to be the value of p which maximizes the function L(p) given by

$$L(p) = p^{N_1}(1-p)^{N_0}[(\hat{\alpha} + \hat{\delta} - 1)p + (1 - \hat{\delta})]^{M_1}[\hat{\delta} - (\hat{\alpha} + \hat{\delta} - 1)p]^{M_0}$$

where

$$N_1 = \sum_{i=1}^{N} \lambda_i$$

$$N_0 = \sum_{i=1}^{N} (1 - \lambda_i) = N - N_1$$

$$M_1 = \sum_{j=1}^{M} \nu_j$$

$$M_0 = \sum_{j=1}^{M} (1 - \nu_j) = M - M_1$$

Now the function

$$\ln L(p) = N_1 \ln p + N_0 \ln(1-p) + M_1 \ln[\alpha+\delta-1)p + (1-\delta)] + M_0 \ln[\delta-(\alpha+\delta-1)p]$$

is concave (its second derivative with respect to p is negative) so that there is a unique maximum. Also, concavity facilitates numerical computation of $\hat{p}$.

As of now we do not know if $\hat{p}$ possesses any useful properties. For example, what about mean-square error? Let

$$\varepsilon^2_{N,M} = E(\hat{p}-p)^2$$

Let $\tilde{p}$ denote the minimum variance unbiased estimate of p based only on the N training data points and their associated classes. Let Var_N denote the variance of $\tilde{p}$. If $\alpha + \delta \neq 1$, is it true that

$$\overline{\lim_{\substack{N\to\infty \\ M=KN}}} \frac{\varepsilon^2_{N,M}}{\mathrm{Var}_N} < 1 \quad ?$$

where $\overline{\lim}$ denotes limit superior, and K is any fixed integer. In words: Does the test data help to estimate p in the large sample case?

We now consider the form the analog of the Cover and Hart result for $\hat{p}$ might take. Suppose $f_1(x)$ and $f_0(x)$ are known. Then one can, in principle, obtain the maximum likelihood estimate of p based on the observations (x_i,u_i) and the y_j. Let this estimate be denoted by $\bar{p}$ and let $\bar{\varepsilon}^2_{N,M}$ be the mean-square error of this estimate. An analog of the Cover and Hart result would specify an upper bound for

$$\overline{\lim_{\substack{N\to\infty \\ M=KN}}} \frac{\varepsilon^2_{N,M}}{\bar{\varepsilon}^2_{N,M}}$$

APPENDIX N

STATIC STRATIFIED SAMPLING ALLOCATION TECHNIQUES

N.1 INTRODUCTION, ASSUMPTIONS, AND NOTATION

In this Appendix, we assume that there are M strata which are defined only by spectral/temporal variables. Thus all segments would have comparable strata. We also assume that the sampling targets are the same size and the ratio of the sample size to the population size is so small that we can approximate the sampling model with the model which samples with replacement.

We denote the strata by S_i for i=1,...,M; the size of the strata by N_i, $N_\cdot = \sum_{i=1}^{M} N_i$, and the strata sample sizes by n_i.

N.2 SAMPLE ALLOCATION WHEN LABELS ARE CORRECT AND EACH STRATUM HAS A PRIOR DISTRIBUTION FOR P_i

The Bayesian view of stratified sampling with prior distributions for the strata proportions P_i is as follows: The proportion of wheat in stratum i is a random variable with an unknown distribution τ_i. Mother nature randomly chooses P_i according to τ_i then chooses X_i according to a binomial (n_i, P_i) distribution, where n_i is the sample size of stratum i.

We assume that we have a prior distribution for each stratum. These priors were obtained from past observations or from the experimenter's intuition. The priors and outcomes of an experiment yield posterior distributions. If the priors are "good" then the posterior becomes the best estimate of the unknown τ_i's. The posterior contains all information from the prior and the experimental data. We wish to use the priors to allocate samples to the strata in order to minimize expected variance of the segment proportion estimate.

If the proportion of wheat labels in stratum i is P_i and the sample size is n_i, then the distribution of X_i, the number of wheat labels in the sample, has density

$$f_i(x|P_i) = \binom{n_i}{x} P_i^x (1-P_i)^{n_i-x} \quad x=0,1,2,\ldots,n_i, \quad 0 \leq P_i \leq 1$$

If the density of the i^{th} prior is $g_i(P)$, then the density of the joint distribution is

$$h_i(X,P) = f_i(X|P)g_i(P)$$

and the density of the marginal distribution of X_i is

$$f_i(X) = \int_0^1 h(x,P)dP$$

The posterior distribution of P_i given $X_i=x_i$ is

$$g_i(P\ X_i) = \frac{h_i(X,P)}{f_i(X)}$$

Often for convenience, it is assumed that

$$g_i(P) = \frac{(r+s+1)!}{r!s!} P^r(1-P)^s \quad 0 \leq P \leq 1,\ r \geq 0,\ s \geq 0$$

This is called a beta (r,s) density. In this case $g_i(P|x_i)$ is a beta $(r+X_i,\ s+n_i-x_i)$ density.

We estimate the segment proportion

$$\overline{P} = \sum_{i=1}^{M} \frac{N_i}{N} \frac{X_i}{n_i}$$

And the expected variance of $\overline{P}$ is

$$V(\bar{P}) = \sum_{i=1}^{M} \left(\frac{N_i}{N_.}\right)^2 \int_o^1 \frac{P(1-P)}{n_i} g_i(P)dP$$

Thus, the allocation which minimizes $V(\bar{P})$ with a fixed total sample size is to choose n_i proportional to

$$\left(\frac{N_i}{N}\right)\sqrt{\int_o^1 P(1-P)g_i(P)dP}^{\,1}$$

Example 1: Suppose that there are four strata, total sample size of 30. We assume that we have no prior information about τ_1, τ_2, τ_3, and τ_4, so we set our prior densities as:

$$g_i(P) = 1 \text{ if } 0 \leq P \leq 1$$
$$= 0 \text{ otherwise}$$

Suppose that N_1=250, N_2=150, N_3=50, N_4=50, and N=500. In order to minimize the expected variance we much have.

$$n_i \propto \frac{N_i}{N_.} \int_o^1 P(1-P)dp = \frac{N_i}{3N_.} \propto N_i$$

which is consistent with [1]. In this case, it is possible to allocate exactly proportional $\tau = n_.$, thus n_1=15, n_2=9, n_3=3, and n_4=3. The table gives the number of wheat labels observed.

i	N_i	n_i	X_i
1	250	15	1
2	150	9	2
3	50	3	2
4	50	3	3
	500	30	

We estimate P with

$$\bar{P} = \frac{250}{500}\left(\frac{1}{15}\right) + \frac{150}{500}\left(\frac{2}{9}\right) + \frac{50}{500}\left(\frac{2}{3}\right) + \frac{50}{500}\left(\frac{3}{3}\right) = \frac{12}{45}$$

and obtain the posteriors as described above, namely:

$$g_1(P|X_1=1) = 240P(1-P)^{14}$$
$$g_2(P|X_2=2) = 360P^2(1-P)^7$$
$$g_3(P|X_3=2) = 12P^2(1-P)$$
$$g_4(P|X_4=3) = 4P^3$$

The graph of these posteriors is given in Figure N-1.

Example 2: Suppose that we have the same four strata definitions with another segment and we use the posteriors from Example 1 (view the segment in Example 1 as a training segment). In order to allocate a sample of size 50 to these four strata, we give

$$N_i, \left(\frac{N_i}{N}\right)\sqrt{\int_o^1 P(1-P)g_i(P)dP} \quad \text{and } n_i$$

as follows.

i	N_i	$\frac{N_i}{N}\sqrt{\int_o^1 P(1-P)g_i(P)dP}$	n_i	
1	200	$\left(\frac{200}{500}\right)\sqrt{\frac{5}{51}} = .12524$	16	(16.52)
2	150	$\left(\frac{150}{500}\right)\sqrt{\frac{2}{11}} = .12792$	17	(16.87)
3	100	$\left(\frac{100}{500}\right)\sqrt{\frac{1}{5}} = .08944$	12	(11.796)
4	50	$\left(\frac{50}{500}\right)\sqrt{\frac{2}{15}} = .036514$	5	(4.81)

In the last section, when we assumed that the labels were all correctly classified we could update our prior every time we sampled. In

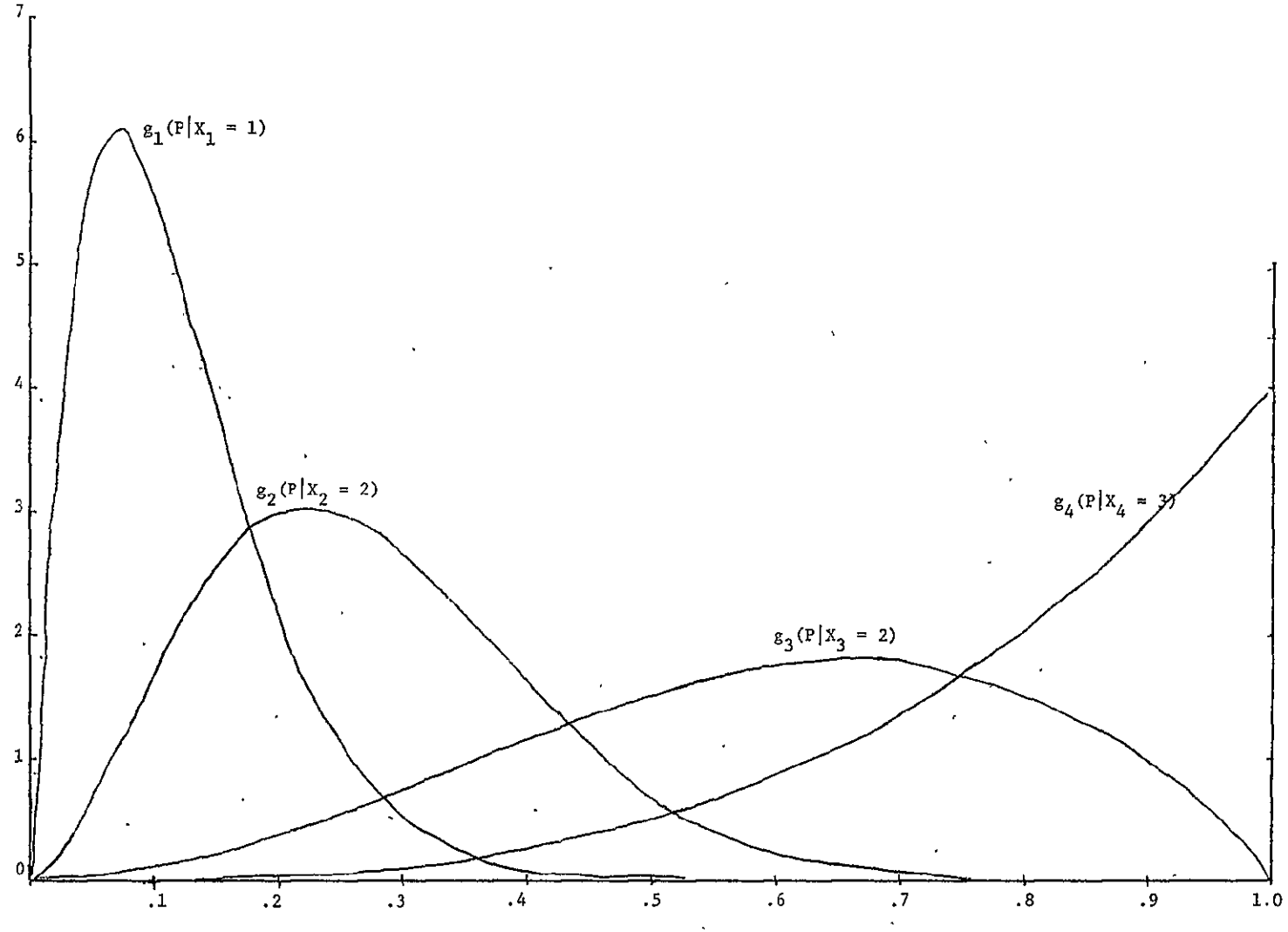


FIGURE N-1. ILLUSTRATION OF POSTERIOR PROBABILITY DENSITIES

this section, we assume that grain targets from stratum i are classified as grain with probability α_i and non-grain targets are classified as non-grain with probability β_i. Unless we have ground truth for all of the targets in the sample, we cannot update our prior.

We first assume that for some training segments we have ground truth for our sample so that we can construct a prior. On the first segment we set $g_i(P,\alpha,\beta) = 1$, $(P,\alpha,\beta)\varepsilon[0,1]^3$, $i=1,2,\ldots,M$. Let n_i denote the i^{th} sample size, W_i denote the grain targets which are labeled grain, Y_i' denote the grain targets which are labeled non-grain, W_i' denote the non-grain targets which are labeled grain, and Y_i denote the non-grain targets which are labeled non-grain. The density of (W_i, Y_i', W_i', Y_i) given (P_i, α_i, β_i) is

$$f_i(W,Y',W',Y|P_i,\alpha_i,\beta_i) = \binom{N_i}{W\ Y'\ W'\ Y} P_i^{W+Y'}\alpha_i^{W}(1-\alpha_i)^{Y'}(1-P_i)^{W'+Y}(1-\beta_i)^{W'}\beta_i^{Y}$$

where $W+Y'+W'+Y = n_i$.

The joint density of $(W_i, Y_i', W_i', Y_i, P_i, \alpha_i, \beta_i)$ is

$$h_i(W,Y',W',Y,P,\alpha,\beta) = g_i(P,\alpha,\beta)f_i(W,Y',W',Y|P,\alpha,\beta)$$

The marginal density of (W_i, Y_i', W_i', Y_i) is

$$f_i(W,Y',W',Y) = \binom{N_i}{W\ Y'\ W'\ Y} \frac{W!Y'!W'!Y!}{(N_i+1)!(W+Y'+1)(W'+Y+1)}$$

The computations are shown in Table N-1.

The posterior of (P_i, α_i, β_i) given (W_i, Y_i', W_i', Y_i) is

TABLE N-1. COMPUTATION OF THE MARGINAL DENSITY OF (W_i, Y_i', W_i', Y_i)

$$f_i(W,Y',W',Y)$$

$$= \int_0^1\int_0^1\int_0^1 \binom{n_i}{W\ Y'\ W'\ Y} P^{W+Y'}\alpha^W(1-\alpha)^{Y'}(1-P)^{W'+Y}(1-\beta)^{W'}\beta^Y\, dP\, d\alpha\, d\beta$$

$$= \binom{n_i}{W\ Y'\ W'\ Y} \int_0^1\int_0^1 \alpha^W(1-\alpha)^{Y'}(1-\beta)^{W'}\beta^Y \left(\frac{(W+Y')!(W'+Y)!}{(n_i+1)!}\right)\int_0^1\left(\frac{(n_i+1)!}{(W+Y')!(W'+Y)!}\ P^{W+Y'}(1-P)^{W'+Y}dP\right) d\alpha\, d\beta$$

$$= \binom{n_i}{W\ Y'\ W'\ Y} \frac{(W+Y')!(W'+Y)!}{(n_i+1)!}\int_0^1\int_0^1 \alpha^W(1-\alpha)^{Y'}(1-\beta)^{W'}\beta^Y\, d\alpha\, d\beta$$

$$= \binom{n_i}{W\ Y'\ W'\ Y} \frac{(W+Y')!(W'+Y)!}{(n_i+1)!}\int_0^1 (1-\beta)^{W'}\beta^Y\, \frac{W!Y'!}{(W+Y'+1)!}\int_0^1 \frac{(W+Y'+1)!}{W!Y'!}\alpha^W(1-\alpha)^{Y'}d\alpha\ \ d\beta$$

$$= \binom{n_i}{W\ Y'\ W'\ Y} \frac{W!Y'!(W'+Y)!}{(n_i+1)!(W+Y'+1)}\int_0^1 (1-\beta)^{W'}\beta^Y\, d\beta$$

$$= \binom{n_i}{W\ Y'\ W'\ Y} \frac{W!Y'!(W'+Y)!}{(n_i+1)!(W+Y'+1)}\left(\frac{W'!Y!}{(W'+Y+1)!}\int_0^1 \frac{(W'+Y+1)!}{W'!Y!}(1-\beta)^{W'}\beta^Y\, d\beta\right)$$

$$= \binom{n_i}{W\ Y'\ W'\ Y} \frac{W!Y'!W'!Y!}{(n_i+1)!(W+Y'+1)(W'+Y+1)}$$

$$g_i(P,\alpha,\beta \mid W,Y',W',Y) = \frac{h_i(W,Y',W',Y,P,\alpha,\beta)}{f_i(W,Y',W',Y)}$$

$$= \frac{(n_i+1)!\,(W+Y'+1)\,(W'+Y+1)}{W!\,Y'!\,W'!\,Y!}\, P^{W+Y'}\alpha^{W}(1-\alpha)^{Y'}(1-P)^{W+Y}(1-\beta)^{W'}\beta^{Y}$$

This posterior could be updated by still another experiment with labels and ground truth using similar methods as given above.

In an experiment without ground truth, we only observe W_i+W_i' and Y_i+Y_i', that is we only observe the number of grain labels and non-grain labels. Thus, it is not reasonable to update the priors. The variance of

$$\overline{P}_i = \frac{W_i+W_i'}{n_i}$$

given (P,α,β) is

$$V(\overline{P}_i) = \frac{[P\alpha+(1-P)(1-\beta)]\,[P(1-\alpha)+(1-P)\beta]}{n_i}$$

Thus the variance of

$$\overline{P} = \sum_{i=1}^{n_i} \frac{N_i}{N_\cdot}\,\overline{P}_i$$

given (P,α,β) is

$$V(\overline{P}) = \sum_{i=1}^{n_i} \left(\frac{N_i}{N}\right)^2 V(\overline{P}_i)$$

The stratified sampling allocation which minimizes $EV(\overline{P})$ with a fixed total sample size is to choose n_i proportional to

$$\frac{N_i}{N}\sqrt{\int_0^1\int_0^1\int_0^1 [P\alpha+(1-P)(1-\beta)]\,[P(1-\alpha)+(1-P)\beta]\,g_i(P,\alpha,\beta)}$$

where the priors are the posterior from the training experiment.

N.4 USE OF BOUNDS ON P_i, α_i, AND β_i IN STRATIFIED SAMPLE ALLOCATION

We assume that we have obtained from training segments the following bounds:

$$\alpha_i \geq A_i > \frac{1}{2}, \quad \beta_i \geq B_i > \frac{1}{2}$$

and case (i) $P_i \geq C_{L_i} > \frac{1}{2}$ or

(ii) $P_i \leq C_{U_i} < \frac{1}{2}$

This gives the following results:

$$V(\overline{P}_i) = \frac{[P_i\alpha_i + (1-P_i)(1-\beta_i)]\ [P_i(1-\alpha_i) + (1-P_i)\beta_i]}{n_i}$$

$$= \frac{P_i^2\alpha_i(1-\alpha_i) + P_i(1-P_i)(\alpha_i\beta_i + (1-\alpha_i)(1-\beta_i)) + (1-P_i)^2\beta_i(1-\beta_i)}{n_i}$$

In case (i) we have

$$V(\overline{P}_i) \leq \frac{A_i(1-A_i) + C_{L_i}(1-C_{L_i})(3-A_i-B_i) + (1-C_{L_i})^2B_i(1-B_i)}{n_i}$$

and in case (ii)

$$V(\overline{P}_i) \leq \frac{C_{U_I}^2A_i(1-A_i) + C_{U_i}(1-C_{U_i})(3-A_i-B_i) + B_i(1-B_i)}{n_i}$$

In either case, denote these upper bounds on $V(\overline{P}_i)$ as D_i^2/n_i. Thus, the variance of

$$\overline{P} = \sum_{i=1}^{m} \frac{N_i}{N} \overline{P}_i$$

is as follows

$$V(P) = \sum_{i=1}^{M} \left(\frac{N_i}{N_.}\right)^2 V(\bar{P}_i) \leq \sum_{i=1}^{M} \left(\frac{N_i}{N_.}\right)^2 \frac{D_i}{n_i}^2$$

In this case we allocate the samples to the strata proportional to

$$\frac{N_i}{N} D_i$$

Example 3. Suppose we have four strata with $\alpha_1 \geq .8$, $\alpha_2 \geq .9$, $\alpha_3 \geq .7$. $\alpha_4 \geq .9$, $\beta_1 \geq .9$, $\beta_2 \geq .95$, $\beta_3 \geq .8$, $\beta_4 \geq .99$, $P_1 \leq .1$, $P_2 \leq .4$, $P_3 \leq .6$, $P_4 \leq .8$, $N_1=250$, $N_2=200$, $N_3=150$, and $N_4=100$. Values for N_i, A_i, B_i, C_i, and D_i are:

i	N_i	A_i	B_i	C_{L_i}	C_{U_i}	D_i	$N_i D_i$
1	250	.8	.9	-	.1	$\sqrt{.2086}$	114.182
2	200	.9	.95	-	.4	$\sqrt{.3379}$	116.258
3	150	.7	.8	.6	-	$\sqrt{.5956}$	115.763
4	100	.9	.99	.8	-	$\sqrt{.267996}$	51.768

Thus, if we have a total sample size of 50, then

$n_1 = 14$ (14.35)
$n_2 = 15$ (14.61)
$n_3 = 15$ (14.54)
$n_4 = 6$ (6.50)

REFERENCES

1. The LACIE Symposium: Proceedings of the Plenary Session (JSC-14551) and Independent Peer Evaluation of the Large Area Crop Inventory Experiment (JSC-14550), NASA Johnson Space Center, Houston, TX, October 1978.

2. The LACIE Symposium: "The USDA Applications Test System", Technical Sessions, Vol. II (JSC-16015), July 1979, pp. 1067-1125.

3. Cicone, R., E. Crist, R. Kauth, P. Lambeck, W. Malila, and W. Richardson, Development of Procedure M for Multicrop Inventory, With Tests of a Spring Wheat Configuration, Final Report 132400-16-F, Environmental Research Institute of Michigan, Ann Arbor, MI, March 1979.

4. Bizzell, R., A. Feiveson, F. Hall, M. Bauer, B. Davis, W. Malila, and D. Rice, Crop Identification Technology Assessment for Remote Sensing (CITARS), Vol. X: Interpretation of Results, JSC-09393, NASA Johnson Space Center, Houston, TX, December 1975.

5. Myers, R. M. and S. G. Wheeler, Fisher Linear Classification and Cross-Validation Bias Correction Comparison With the LACIE Procedure 1 Small Grains Estimation, IBM Report RES 23-70, International Business Machines Corporation, Federal Systems Division, Houston, TX, August 1979.

6. Peters, B. C., Jr. and H. F. Walker, "An Iterative Procedure for Obtaining Maximum Likelihood Estimates of the Parameters of a Normal Distribution", SIAM J. Appl. Math., Vol. 35, No. 2, September 1978, pp. 362-378.

7. Lennington, R. K. and M. E. Rassbach, Mathematical Description and Program Documentation for CLASSY, An Adaptive Maximum Likelihood Clustering Method, Technical Memorandum LEC-12177, Lockheed Electronics Co., April 1979.

8. Peters, Charles, "Spectral/Spatial Models for Agricultural Landsat Data", NAS-9-15543, Mod. 2S, Quarterly SR&T Review, September 11, 1979.

9. Kauth, R. J. and W. Richardson, Procedure B: A Multispectral Training Selection and Proportion Estimation Procedure for Processing Landsat Agricultural Data, Technical Report 122700-31-F, Environmental Research Institute of Michigan, Ann Arbor, MI, November 1977.

REFERENCES (Cont'd)

10. LACIE Phase III Accuracy Assessment, Final Report, LACIE-00478, JSC-13766, NASA Johnson Space Center, Houston, TX, August 1979.

11. Balon, R. J. and R. C. Cicone, Uniform Color Space Analysis of LACIE Image Products, Technical Memorandum 132400-10-R, Environmental Research Institute of Michigan, Ann Arbor, MI, May 1979.

12. Kauth, R. J. and G. S. Thomas, "The Tasseled-Cap -- A Graphic Description of the Spectral-Temporal Development of Agricultural Crops As Seen by Landsat", Machine Processing of Remotely Sensed Data, Symposium Proceedings, Purdue/LARS, West Lafayette, IN, 1976.

13. Kauth, R., P. Lambeck, W. Richardson, G. Thomas, and A. Pentland, "Feature Extraction Applied to Agricultural Crops as Seen by Landsat", Proceedings of the Technical Sessions, Vol. II, The LACIE Symposium, JSC 16015, NASA Johnson Space Center, Houston, TX, July 1979, pp. 705-721.

14. Pore, M. D. and R. A. Abotten, "A Programmed Labeling Approach to Image Interpretation", Proceedings of the Technical Sessions, Vol. II, The LACIE Symposium, NASA Johnson Space Center, Houston, TX, July 1979, pp. 923-936.

15. Thomson, D. R. and O. A. Wehmanen, "Application of Landsat Digital Data for Monitoring Drought", Proceedings of the Technical Sessions, Vol. I, The LACIE Symposium, July 1979, pp. 431-438.

16. Badhwar, G. D., "Small Grains" presented at NASA Supporting Research and Technology Quarterly Review, NASA Johnson Space Center, Houston, TX, September 1978.

17. Hay, C. M., R. W. Thomas, and A. S. Benson, Development of Techniques for Producing Static Strata Maps and Development of Photointerpretation Methods Based on Multitemporal Landsat Data; Final Report. NASA Contract NAS9-14565, Principal Investigator, R. N. Colwell. University of California, Space Sciences Laboratory, Series 19, Issue 1, Berkeley, CA, August 1977.

18. Malila, W. A. and E. P. Crist, "Spectral Separability of Spring Wheat from Other Spring Small Grains", presented at NASA SR&T Quarterly Review, NASA Johnson Space Center, Houston, TX, September 1978; (Documented in ERIM Report 132400-12-P).

REFERENCES (Cont'd)

19. Kauth, R. P. Lambeck, W. Richardson, G. Thomas, and A. Pentland, "Feature Extraction Applied to Agricultural Crops as Seen by Landsat", Proceedings of the Technical Sessions, Vol. II, The LACIE Symposium, JSC-16015, NASA Johnson Space Center, Houston, TX, July 1979, pp. 705-721.

20. Ritchie, J. T,, U.S.D.A., Grassland, Soil and Water Research Laboratory, Temple, TX (private communications, March and May 1979).

21. Suits, G. H., "The Calculation of the Directional Reflectance of a Vegetative Canopy", Remote Sensing of Environment, Vol. 2: 117-125, 1972.

22. Turner, R. E., Radiative Transfer in Real Atmospheres, Final Report 190100-24-T, Environmental Research Institute of Michigan, Ann Arbor, MI, July 1974.

23. Kaneko, T., Analysis of Methods of Generating Bias and Scale Factors for PFC Guns, IBM Memo RES 20-53-134, August 12, 1976.

24. Kraus, G., Another Method for Computing PFC Scaling Factors, NASA Memo TF3/76-630, September 9, 1976.

25. Hocutt, W. T., An Improved Method of Calculating Scale Factors and Bias for Color Film Generation, LEC-8301, June 1976.

26. Cate, R. B., D. E. Phinney, M. C. Kinsler, M. L. Sestak, T. Hodges, and J. Dishler, Color Analysis of Landsat Digital Data Using a Cubic Color Model Based on Relative Energies, Lockheed Electronics Co. (to be published in early 1980)

27. Cicone, R. and R. Balon, "Multicrop Labeling Aids", Quarterly Progress Report on Analysis of Scanner Data for Crop Inventories, ERIM 132400-12-P, Environmental Research Institute of Michigan, Ann Arbor, MI, June-September 1978.

28. Walker, M. M., "Principal Components Colour Display of ERTS Imagery", Proceedings of the Second Canadian Symposium on Remote Sensing, Guelph, Ontario, Canada, 1974.

29. Juday, Richard D., Colorimetric Principles as Applied to Multichannel Imagery, NASA Johnson Space Center, Houston, TX, Revision B, July 15, 1978.

REFERENCES (Cont'd)

30. Lambeck, P. F., Signature Extension Preprocessing for Landsat MSS Data, Final Report 122700-32-F, Environmental Research Institute of Michigan, Ann Arbor, MI, November 1977.

31. Boullion, T. L., B. S. Duran, and P. L. Odell, Estimation and Classification With Incomplete Data, Annual Report, University of Texas at Dallas, JSC-09703, June 1, 1974 - May 31, 1975.

32. Helmer, D. C., C. V. Nazare, R. W. Payne, and W. L. West, III, Report on the Analyst Test Using the Blob Labeling Procedures, Lockheed Electronics Corp., AD 63-1827-4845-33, May 17, 1979.

33. Wheeler, S. G., "Procedure 1A Development", presented at NASA SR&T Quarterly Review, NASA Johnson Space Center, Houston, TX, September 10, 1979.

34. Malila, W. A., R. C. Cicone, and J. M. Gleason, Wheat Signature Modeling and Analysis for Improved Training Statistics, Final Report 109600-66-F, Environmental Research Institute of Michigan, Ann Arbor, MI, May 1976.

35. Malila, W. A., J. M. Gleason, and R. C. Cicone, Atmospheric Modeling Related to Thematic Mapper Scan Geometry, Final Report 119300-5-F, Environmental Research Institute of Michigan, Ann Arbor, MI, April 1976.

36. Kauth, R. J., P. F. Lambeck, and G. S. Thomas, "SCREEN, A Procedure for Automatically Detecting Garbled Data, Clouds, Snow, Cloud Shadows, and Water in Landsat MSS Data", New Technology Report, Environmental Research Institute of Michigan, Ann Arbor, MI, November 10, 1977.

37. Lambeck, P. F., "Spatially Varying XSTAR Haze Correction", New Technology Report, Environmental Research Institute of Michigan, Ann Arbor, MI, March 30, 1979.

38. Kauth, R. J. and J. F. Hemdal, "Landsat 1 to Landsat 2 Data Conversion", New Technology Report, Environmental Research Institute of Michigan, Ann Arbor, MI, November 21, 1977.

39. Cochran, William G., Sampling Techniques, Third Edition, John Wiley & Sons, NY, 1977, pp. 89-90.

REFERENCES (Cont'd)

40. Richardson, W., A Study of the Tolerance Block Approach to Spectral Stratification, Technical Memorandum 132400-25-R, Environmental Research Institute of Michigan, Ann Arbor, MI, August 1979.

41. Lennington, K., "PIA Unsupervised Procedure Development", presented at NASA SR&T Quarterly Review, NASA Johnson Space Center, Houston, TX, LEC/SSD REF 642-7701, September 1979.

42. Kauth, R. J. and W. Richardson, "Task 1, Multisegment Training", Analysis of Scanner Data for Crop Inventories, Third Quarterly Report on NASA Contract NAS9-15476, ERIM Report 132400-12-P, Environmental Research Institute of Michigan, Ann Arbor, MI, September 1978, pp. 1-5.

43. Pont, W. Frank, "Appendix B: Effects of Random Stratification", from Technical Memorandum 132400-25-R, Environmental Research Institute of Michigan, Ann Arbor, MI, August 1979, pp. 49-55.

44. Kauth, R. J., A. P. Pentland, and G. S. Thomas, "BLOB, An Unsupervised Clustering Approach to Spatial Preprocessing of MSS Imagery", Eleventh International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, MI, Vol. 2, 1977, pp. 1309-1317.

45. Kuretz, C. A. and C. M. Hay, "AI Procedures for Episodal Events", presented at NASA SR&T Quarterly Review, NASA Johnson Space Center, Houston, TX, June 1979.

46. Badhwar, G. D., "Multitemporal Classification Scheme for Corn", presented at NASA SR&T Quarterly Review, NASA Johnson Space Center, Houston, TX, June 1979.

47. International Mathematical & Statistical Libraries, Inc., Reference Manual, LIB-007, January 1979 (Revised).

48. Jackson, R. D. and P. Pinter, U.S. Water Conservation Laboratory, Phoenix, Arizona, private communication, March 1979.

49. Malila, W., J. Gleason, F. Sadowski, R. Cicone, and E. Crist, "Applications of Modeling to Analysis and Processing of Landsat Data", Proceedings of Twelfth International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, MI, April 1978.

REFERENCES (Cont'd)

50. Nalepka, R. F., J. E. Colwell, and D. P. Rice, Forecasts of Winter Wheat and Production Using Landsat Data, Final Report 114800-38-F, Environmental Research Institute of Michigan, Ann Arbor, MI, December 1977.

51. Malila, W. A., J. M. Gleason, and R. C. Cicone, "Multispectral System Analysis Through Modeling and Simulation", Proceedings of Eleventh International Symposium on Remote Sensing of Environment, Environmental Research Institute of Michigan, Ann Arbor, MI, April 25-29, 1977.

52. Rouse, J., R. Haas, J. Schell, and D. Deering, "Monitoring Vegetation Systems in the Great Plains", Third ERTS Symposium, NASA/SP-351, I:309-317, 1973.

53. Wyszecki, G. and W. S. Stiles, Color Science, John Wiley & Sons, New York, 1967.

54. Kaneko, Toyohisa and Sharon Tinkler, LACIE Film Study, Generation of Color Infrared Films, 78-FM-8, JSC-13847, NASA Johnson Space Center, Houston, TX, February 1978.

55. Coberly, W. A., Jack Tubbs, and P. L. Odell, The Influence of False Color Infrared Film on Training Errors in LACIE, University of Texas at Dallas, Annual Report, June 1, 1975 - August 31, 1976.

56. Austin, W. W., PFC Film Products for LACIE Segments, Memorandum to L. M. Flores, NASA, December 6, 1976.

57. Malila, W. A. and J. M. Gleason, Investigations of Spectral Separability of Small Grains, Early Season Wheat Detection, and Multicrop Inventory Planning, Final Report 122700-34-F, Environmental Research Institute of Michigan, Ann Arbor, MI, November 1977.

58. Anderson, T. W., An Introduction to Multivariate Statistical Analysis, John Wiley & Sons, N.Y., 1958.

59. Suits, G. H., "The Cause of Azimuthal Variations in Directional Reflectance of Vegetative Canopies", Remote Sensing of Environment, Vol. 2: 175-182, 1972.

REFERENCES (Cont'd)

60. Condit, H. R., "The Spectral Reflectance of American Soils", Photogrammetric Engineering, Vol. XXXVI, No. 9, September 1970, pp. 955-966.

61. Cover, T. M. and P. E. Hart, "Nearest Neighbor Pattern Classification", IEEE Trans. IT, Vol. IT-13, No. 1, January 1967.


FEB 1980
RECEIVED
NASA STI FACILITY
ACCESS DEPT.