AgRISTARS

8 0 - 1 0 1 1 4
SR-P9-00411
NAS9-15466 NASA CR-160548

A Joint Program for Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing

Supporting Research

November 1979

Final Report

Vol. II
Processing Techniques Development
Part 1: Crop Inventory Techniques

by **C. S. T. Daughtry and M. M. Hixson**

(E80-10114) PROCESSING TECHNIQUES DEVELOPMENT, VOLUME 2. PART 1: CROP INVENTORY TECHNIQUES Final Report, 1 Dec. 1978 - 30 Nov. 1979 (Purdue Univ.) 76 p HC A05/MF A01 CSCL 02C G3/43 N80-23741 Unclas 00114

Laboratory for Applications of Remote Sensing
Purdue University
West Lafayette, Indiana 47907

NASA

SR-P9-00411
NAS9-15466
LARS 112979

FINAL REPORT

VOL. II PROCESSING TECHNIQUES DEVELOPMENT

PART 1: CROP INVENTORY TECHNIQUES

BY

C.S.T. Daughtry and M.M. Hixson

The research reported here was initiated during the planning phases of the AgRISTARS Supporting Research Project and, although it stands on its own merit, it benefits the Supporting Research Project and became a part of those plans.

Purdue University
Laboratory for Applications of Remote Sensing
West Lafayette, Indiana 47906

November 1979

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

A. APPLICATION AND EVALUATION OF LANDSAT TRAINING, CLASSIFICATION, AND AREA ESTIMATION PROCEDURES FOR CROP INVENTORY

Marilyn M. Hixson*

1. Introduction

Accurate and timely crop production information is a critical need in today's economy. During the past decade, satellite remote sensing has been increasingly recognized as a means for crop identification and estimation of crop areas.

An extensive experiment, the Large Area Crop Inventory Experiment (LACIE), was conducted by NASA, USDA, and NOAA during 1974 through 1977 [1]. Its data analysis objective was to distinguish small grains from non-small grains using Landsat multispectral scanner (MSS) data. Several other investigations have shown that the potential also exists for identification and area estimation of corn and soybeans [2,3,4,5].

This task is the second year of a specific LARS task which resulted from a proposal in response to the Applications Notice. It is also part of the second year of effort in a larger, multi-year, multi-organizational effort to extend LACIE-like technology to crops other than the small grains. The accuracy and precision of area estimates obtained from Landsat data are affected by a combination of training, classification, and area estimation procedures used. Several types of agricultural scenes in the U.S. Corn Belt are being investigated in this task to assess scene dependent differences in optimal choices of training, classification, and area estimation procedures.

---

*Data analyses for Task 2A, Application and Evaluation of Landsat Training, Classification, and Area Estimation Procedures for Crop Inventory, were conducted by Donna Scholz, Mark Swenson, Carol Jobusch, Tsuyoshi Akiyama, and Getulio Batista. Carol Jobusch, Jeanne Etheridge, and Joan Buis aided in programming and system problems. Carol Jobusch and Mark Swenson conducted some of the statistical analyses. Many thanks are also due to Dr. Marvin Bauer, Dr. Philip Swain, Dr. Virgil Anderson, and Dr. K.C.S. Pillai who acted as consultants and advisors to the project.

2. Objectives

The overall objective of this study is to evaluate Landsat training, classification, and area estimation procedures for crop inventory. Specific objectives include:

- Assess the effect of sampling in training and classification on area estimates.
- Compare several methods for obtaining training statistics.
- Assess the ability of several classifiers to provide acreage estimates of corn and soybeans in several regions of the U.S. Corn Belt.
- Assess the potential accuracy of corn and soybean estimates as a function of growth stage, both unitemporally and multitemporally.

3. Experimental Approach

During the current contract year, four subtasks, each of which addressed several aspects of the general classification problem, were conducted. These subtasks were: (1) a study of the effects of sampling in clustering and classification, (2) a study of several alternatives in the training procedure, (3) a comparison of several classification algorithms, and (4) an assessment of the potential accuracy of corn and soybean estimates as a function of growth stage. The specific approach used in each of these subtasks will be discussed in the section addressing that objective. The experiment design permits an integrated study of sampling, training, and classification, allowing for interactions among the components of the procedure. Training method, features used in classification, and classification algorithms were varied. Effects of site location were assessed.

The data set which was used in this study was drawn from the data acquired in 1978 over the U.S. corn and soybean sites. The data obtained were from 81 sample segments located in four test areas in Iowa, Illinois, and Indiana (Figure A-1).

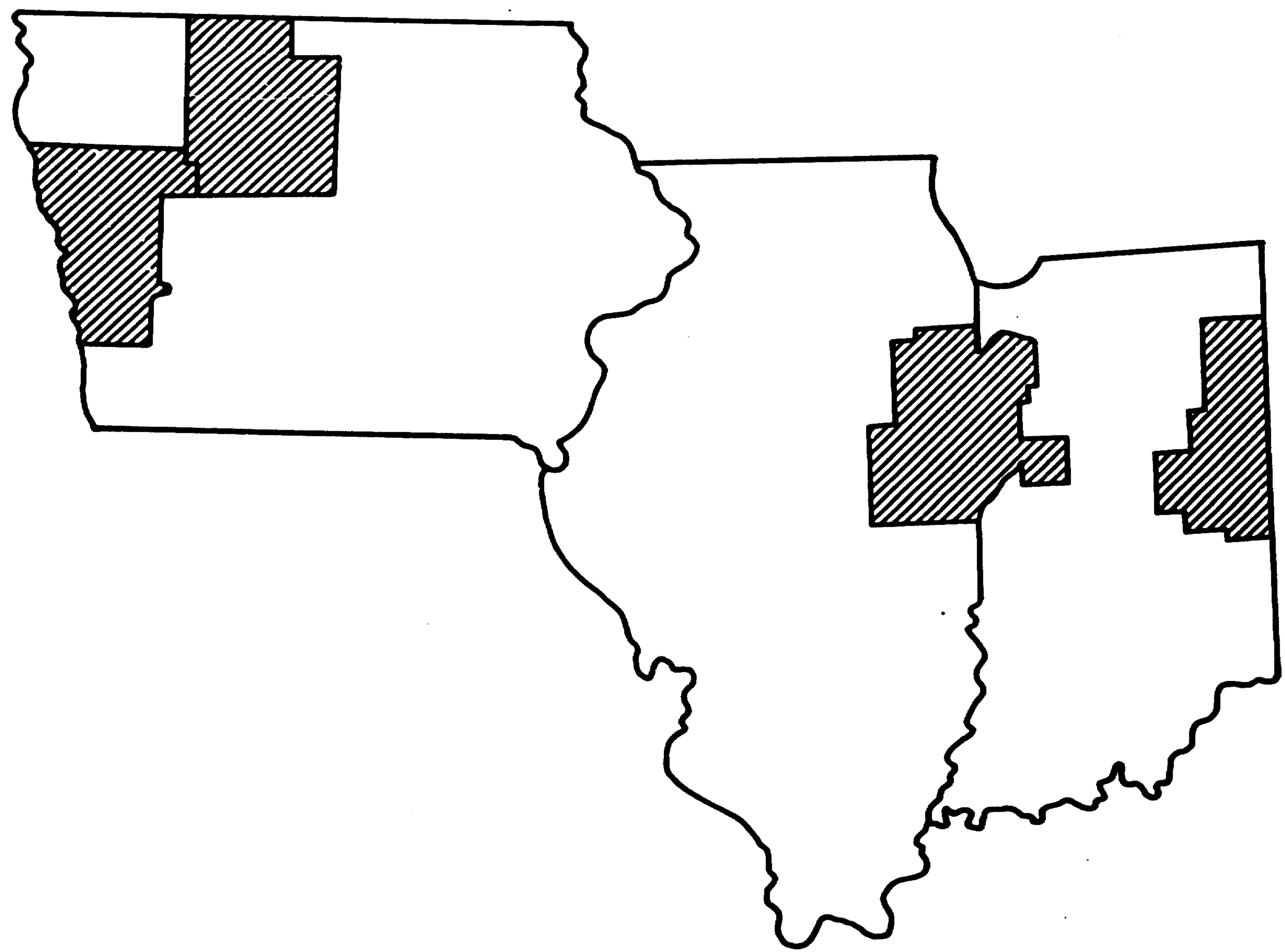


Figure A-1. Locations of the four test areas.

LACIE-type sample segments (5 x 6 nautical miles in size) were selected, generally two per county. Landsat data acquired included multitemporally registered MSS data tapes and film writer imagery (PFC Product 1) for each acquisition and segment. Color infrared prints of aerial photography with ground inventory overlays were obtained. Additional reference data were obtained for some segments in the form of labels of 418 pixels located on systematic grids in a segment. Digitized wall-to-wall inventories were obtained for some of the segments which NASA/JSC had digitized. A summary of the currently available data set is given in Table A-1.

To permit interchangeability of algorithms and approaches, a set of computer routines were written to make the LARSYS and EODLARSYS systems compatible. Routines are included for statistics conversion between formats and results conversion between formats. A description of these programs and user documentation are available on request.

A second programming effort was initiated to reduce cost and data preparation time. The objective of this effort was to program the capability for LARSYS to read either LARSYS or UNIVERSAL format data tapes. All the processors in LARSYS had previously been able to read only LARSYS format data tapes, but all data were received in UNIVERSAL format, necessitating a reformatting operation before analysis could be carried out. Now, developmental LARSYS (LSDV370) will automatically determine the format of a data tape (i.e., the format does not need to be user-specified) and will read the tape using the appropriate format statements. This programming effort was partially funded from this task.

4. Sampling Effects in Clustering and Classification

A study was conducted to investigate the "best" subset of bands for crop separability. Multitemporal data from four segments in the Corn Belt were analyzed (Table A-2). Training data were fields located on a systematic grid; labels were obtained from ground inventories. Statistics were developed by clustering all training fields of one cover type together. The best combination of four from the sixteen available channels (four dates) was selected

Table A-1. Summary of types of data available for 81 U.S. corn and soybean segments.

Test Site	State	County	Number	Landsat MSS	Segment Imagery	Aerial Photo	Ground Inventory	Pixel Labels	Digital Inventory
1	IN	Adams	832	X	X	X	X	X	X
			833	X	X	X	X		
		Allen	834	X	X	X	X		
			835	X	X	X	X		
		Blackford	838	X	X	X	X		
			839	X	X	X	X		
		Delaware	840	X	X	X	X	X	
			841	X	X	X	X		
		Henry	842	X	X	X	X	X	X
			843	X	X	X	X	X	X
		Jay	846	X	X	X	X		
			847	X	X	X	X	X	
		Madison	848	X	X	X	X	X	
			849	X	X	X	X	X	
		Randolph	852	X	X	X	X	X	X
			853	X	X	X	X	X	X
		Wayne	858	X	X	X	X		
			859	X	X	X	X		
		Wells	860	X	X	X	X	X	X
			861	X	X	X	X	X	
2	IN	Benton	836	X	X	X	X	X	
			837	X	X	X	X	X	X
		Jasper	844	X	X	X	X		
			845	X	X	X	X		
		Newton	850	X	X	X	X		
			851	X	X	X	X	X	
		Tippecanoe	854	X	X	X	X	X	
			855	X	X	X	X	X	
		Warren	856	X	X	X	X	X	
			857	X	X	X	X	X	
	IL	Champaign	820	X	X				
			821	X	X				
			822	X	X				
		Ford	823	X	X				
		Iroquois	824	X	X	X	X	X	
			825	X	X	X	X	X	
			826	X	X	X	X	X	
		Kankakee	827	X	X	X	X	X	
			828	X	X	X	X	X	
		Vermilion	829	X	X				
			830	X	X				
			831	X	X				
3	IA	Calhoun	862	X	X	X	X	X	
			863	X	X				
		Emmet	866	X	X	X	X	X	
			867	X	X	X	X	X	

Table A-1. (Cont.)

Test Site	State	County	Number	Landsat MSS	Segment Imagery	Aerial Photo	Ground Inventory	Pixel Labels	Digital Inventory
3	IA	Hamilton	868	X	X	X	X		
			869	X	X	X	X		
		Hancock	870	X	X	X	X	X	
			871	X	X	X	X	X	
		Humboldt	874	X	X	X	X	X	
			875	X	X	X	X	X	
		Kossuth	878	X	X	X	X	X	X
			879	X	X	X	X		
			139	X	X	X	X	X	
		Palo Alto	882	X	X	X	X	X	X
			883	X	X	X	X	X	
		Pocahontas	884	X	X	X	X		
			885	X	X	X	X		
		Webster	893	X	X	X	X	X	X
			894	X	X	X	X	X	
		Wright	898	X	X	X	X		
			899	X	X				
4	IA	Crawford	864	X	X	X	X	X	X
			865	X	X	X	X	X	X
		Harrison	872	X	X				
			873	X	X				
		Ida	876	X	X	X	X		
			877	X	X	X	X	X	X
		Monona	880	X	X	X	X	X	X
			881	X	X	X	X	X	X
		Pottawattamie	886	X	X	X	X	X	
			887	X	X	X	X	X	
			888	X	X	X	X	X	
		Sac	889	X	X				
			890	X	X	X	X	X	
		Shelby	891	X	X	X	X	X	X
			892	X	X	X	X	X	
		Woodbury	895	X	X	X	X	X	
			896	X	X	X	X	X	
			897	X	X	X	X	X	

Table A-2. Segments and acquisitions used in the wavelength band selection study.

Segment	Landsat Acquisition Date	Growth Stage of Corn
824 (Iroquois, IL)	6/12	emergence
	8/5	tasseling
	8/31	dent
	9/28	mature
854(Tippecanoe, IN)	6/10	emergence
	7/26	tasseling
	8/21	dough
	9/26	mature
886(Pottawattamie,IA)	6/16	emergence
	7/23	tasseling
	9/6	dent
	9/24	mature
892(Shelby,IA)	6/16	emergence
	7/23	tasseling
	8/9	blister
	9/24	mature

using the separability function in LARSYS. Channel combinations are ranked according to the average transformed divergence. A tabulation of results is in Table A-3. The first channel (.5-.6 μm) on each date was very rarely selected; the two near infrared bands were both selected with high frequency on all dates. It was discovered that of the 30 best channel combinations in four segments, neither two visible, nor two infrared channels from the same date were ever selected. Thus, either channel three (.7-.8 μm) or channel four (.8-1.1 μm), but not both, should be selected.

To decide which of the two channels should be the candidate for use, several criteria were considered. The first criterion, the channel selected most frequently for the single best combination, found channel four selected more often. Table A-3 illustrates that summed over segments, dates, and the best 30 combinations, channel four was selected more often. The final criterion was a subjective one: that channel three is in a region of rapid change in response of green vegetation and does not seem to be as reliable.

In summary, the use of all 16 channels in crop identification and classification does not seem to be necessary. Two visible channels or two near infrared channels from the same measurement date were never selected. Channels two (.6-.7 μm) and four (.8-1.1 μm) from each date appear to give a good subset to classify with or select another subset from.

A second analysis was then conducted to assess the effect of sampling in clustering and classification on classification accuracy, proportion estimates, and variance reduction factors. The sample of wavelength bands suggested in the previous analysis was evaluated, and results using a sample of data were compared with the use of all data. The study was based on two principles: (1) past studies have noted a tendency for performance to decrease as the number of wavelength bands used in classification increases and (2) it is very expensive to cluster and classify all pixels in a segment. Data were analyzed from three segments: 824 in Iroquois County, Illinois; 886 in Pottawattamie County, Iowa; and 892 in Shelby County, Iowa. Multitemporally registered data from four Landsat acquisition dates were used.

Table A-3. Number of appearances of each individual channel in the top 30 combinations.

Corn Growth Stage	Channel	Segment 824	854	886	892	Total	Rank
Emergence	1	-	2	-	5	7	13
	2	11	12	2	16	41	7
	3	18	16	7	11	52	3
	4	7	14	21	4	46	5
Tasseling	1	-	-	6	-	6	14
	2	-	4	10	6	20	10
	3	10	11	11	10	42	6
	4	11	15	19	20	65	2
Blistering to Dent	1	-	-	4	-	4	15
	2	-	8	6	-	14	12
	3	9	18	12	12	51	4
	4	21	12	18	18	69	1
Mature	1	3	-	-	-	3	16
	2	16	-	-	-	16	11
	3	8	6	1	9	24	8
	4	6	2	3	9	20	9

Three variables were investigated: sample of data used in clustering, sample of data used in classification, and number of wavelength bands used in clustering and classification. Eight treatments (a 2^3 factorial design) were applied on each of the sample segments, with segments being the random factor in the experiment design.

The general data analysis procedure which was used for the experiment was the Procedure 1 software in a LACIE-like mode. Between 40 and 60 Type 1 dots were used to seed the clustering algorithm and to label the resultant clusters. ISOCLS was used to cluster the data with a simulated single pass. The clusters were labeled using the single nearest Type 1 dot. Sum-of-densities classification was carried out on three cover types. The Type 2 dots were used to estimate a confusion matrix and compute a stratified area estimate. The variables analyzed were estimates of proportions of corn and soybeans; percent correct for corn, soybeans, and other; and variance reduction factors (R.V.) for corn and soybeans.

The dashes in Table A-4 for eight bands, 6% cluster results are indicative of a missing data problem for segment 824. This segment was primarily corn and soybeans with very few other cover types being represented in the scene. Using this set of parameters, it was not possible to find any subclasses identified as other crops, so classifications were not carried out.

Because of the missing data problem, the use of eight wavelength bands clustering a 6% sample of data could not be recommended for use. In addition, some significant factor interactions suggest that the use of a 6% cluster sample with 16 bands may also lead to different results. It is indeed possible that, although 6% vs. 100% clustering showed a significant difference, a cluster sample of a larger percent of data would be highly acceptable. This study did not pursue that possibility.

It appeared, however, that the sample of data classified did not significantly alter the resulting proportion estimates. In addition, the classification accuracy and proportion estimates using eight bands were not

Table A-4a. Proportion estimates obtained from several sampling alternatives.

Cover Type	Seg. No.	8 Bands: 6% Cluster, 6% Cla.	8 Bands: 6% Cluster, 100% Cla.	8 Bands: 100% Cluster, 6% Cla.	8 Bands: 100% Cluster, 100% Cla.	16 Bands: 6% Cluster, 6% Cla.	16 Bands: 6% Cluster, 100% Cla.	16 Bands: 100% Cluster, 6% Cla.	16 Bands: 100% Cluster, 100% Cla.
CORN	824	-	-	57.2	57.5	59.3	57.9	62.8	64.2
	886	54.2	53.6	57.2	58.3	55.2	53.4	56.5	56.5
	892	59.5	58.5	55.6	57.2	54.5	52.5	55.6	55.3
SOYBEANS	824	-	-	42.8	42.5	40.7	42.1	37.2	35.8
	886	26.3	24.9	22.9	23.1	24.6	23.9	23.5	23.7
	892	11.8	12.2	9.9	10.1	13.4	11.9	11.7	12.4

Table A4-b. Classification accuracies (percent) obtained from several sampling alternatives.

Cover Type	Seg. No.	8 Bands		16 Bands	
		6% Cluster	100% Cluster	6% Cluster	100% Cluster
CORN	824	-	90.0	90.0	86.7
	886	96.4	100.0	96.4	96.4
	892	91.2	94.1	97.1	97.1
SOYBEANS	824	-	85.7	85.7	100.0
	886	91.7	91.7	91.7	91.7
	892	85.7	100.0	85.7	100.0
OTHER	824	-	-	-	-
	886	100.0	88.9	66.7	88.9
	892	58.3	50.0	83.3	50.0

Table A-4c. Variance reduction factors (R.V.) obtained from several sampling alternatives.

Cover Type	Seg. No.	8 Bands		16 Bands	
		6% Cluster	100% Cluster	6% Cluster	100% Cluster
CORN	824	-	.436	.436	.353
	886	.163	.377	.243	.239
	892	.624	.309	.568	.503
SOYBEANS	824	-	.372	.372	.293
	886	.213	.440	.213	.300
	892	.622	.416	.672	.631

significantly different from that using all 16 bands.

5. Evaluation of Alternative Training Methods

The first investigation of training procedures was conducted using data from the CITARS project, before 1978 corn and soybean data became available [6]. These analyses used data from the Fayette County, Illinois, test site.

Several aspects of Procedure 1 were investigated and their effects on estimates were assessed. Particular items investigated included: the distance measure used in the LABEL processor, the number of pixels required per cluster class, and the number of iterations (passes) used in ISOCLS.

A study compared use of L1 and L2 distance in the LABEL processor to identify clusters with their nearest neighbor. No significant differences in estimates of corn or soybeans were found.

Another experiment compared results obtained using or deleting small cluster classes. The first method was to use all clusters large enough not to have singular covariance matrices, and the second method was to delete all clusters with fewer than 100 points. No significant differences in estimates of corn or soybeans were found. Slightly higher classification accuracies were obtained for soybeans and else when small classes were deleted, resulting in somewhat better variance reduction factors for the crops of interest.

The final analysis using data from the Fayette County site was an evaluation of the number of iterations (passes) used in ISOCLS. A four date, 16 channel clustering was carried out in two ways. The first was one iteration with no splitting of cluster classes allowed, and the second was a twenty iteration cluster with a printout of intermediate results after every five iterations. Forty Type 1 dots were input to serve as initial cluster centers; therefore, the single iteration procedure had 40 clusters. However, the twenty-pass procedure created 60 clusters, the maximum that was allowed by the user-set parameter.

The aim of the analysis was to see how well the one-pass procedure clustered the data, compared with the twenty-pass procedure. There is a very large increase in computer time needed for the twenty-pass procedure, so fewer iterations are preferable if they perform adequately.

For one, five, ten, fifteen, and twenty iterations, the computer printout contained: (1) a table of the standard deviations of each cluster for each channel, (2) a table of means of each cluster for each channel, and (3) a list of the number of points in each cluster.

Two questions were considered: (1) at what point (i.e., after how many iterations) do the standard deviations of the clusters get small or stabilize and (2) when do the cluster means stabilize.

For each channel, the three clusters with the largest standard deviations were examined. There were no real changes after five or more passes; there was, however, some tightening of clusters between one and five iterations. Next, the distributions of cluster standard deviations after one, five, and twenty iterations were examined by tabulating the number of clusters whose standard deviations were between n and n+1 for n=1,2...11. Graphs (such as Figure A-2) were drawn for band one (.5-.6 μm) on June 10 and 29, bands two (.6-.7 μm) and three (.7-.8 μm) on June 29 and July 17, and band four (.8-1.1 μm) on June 29 and August 21. The general conclusion was that the distribution of standard deviations improved very slightly with more iterations; the graphs showed very little change.

To compare distributions of cluster means, which involves dealing with a 16-dimensional measurement space, projections onto a two-dimensional space were examined; scatterplots of cluster means for one visible (.6-.7 μm) and one near infrared (.8-1.1 μm) channel for a given date were overlaid for one, five, and 20 iterations. If the 20 iteration cluster defines the measurement space, it must be concluded that the single iteration clusters cover almost all of the space.

A second study, using test segments from the Corn Belt, examined

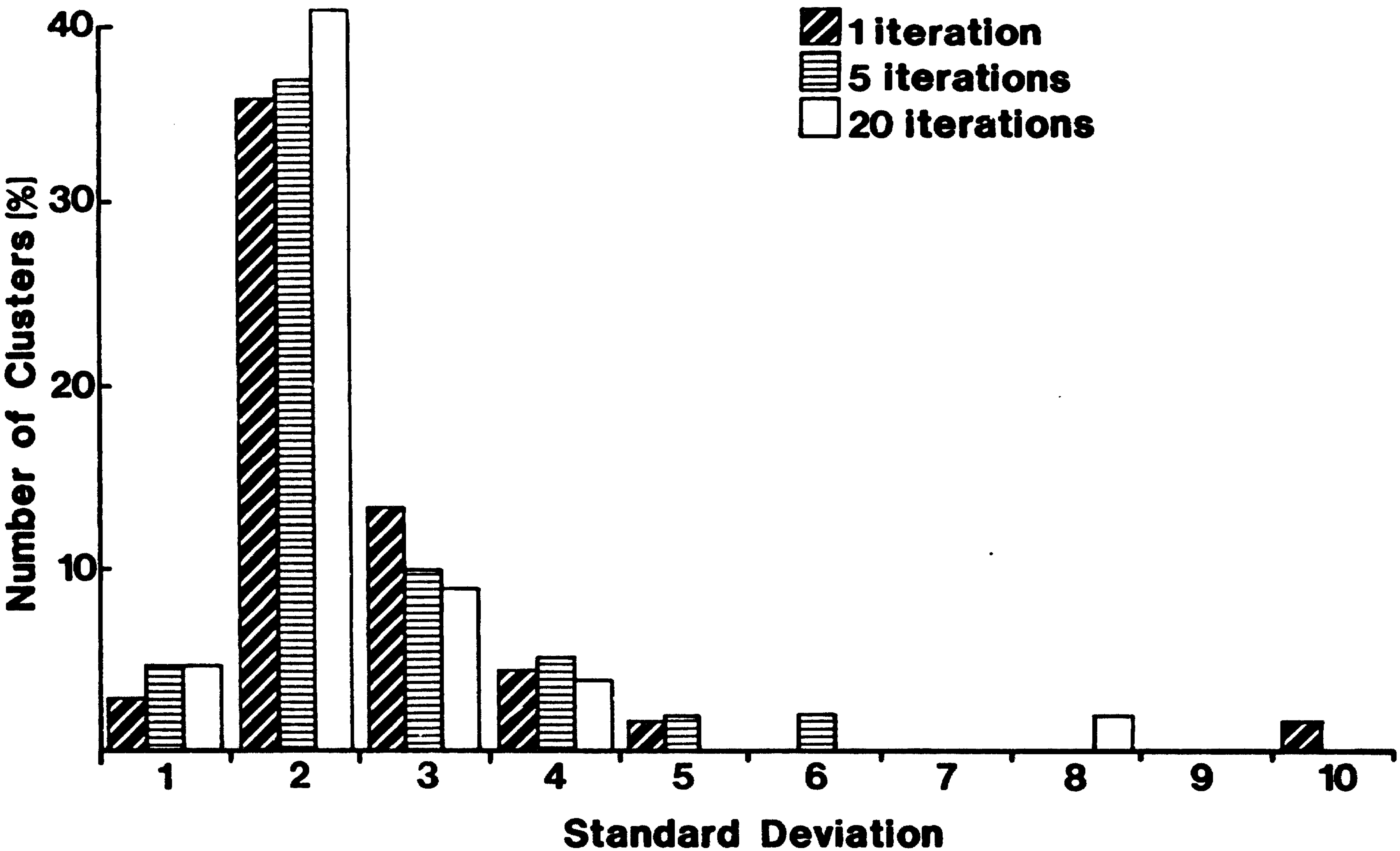


Figure A-2. Distribution of standard deviations for Band 1 (.5-.6 μm) on June 29.

procedures used in a modified supervised training approach. Four acquisitions were analyzed. These were selected one from each of four time periods which were defined based upon corn growth stage: stage 1 was preplant to eight leaves; stage 2 was ten leaves to tasseling; stage 3, tasseling to beginning dent; and stage 4, dent to maturity.

Training fields were selected on a systematic grid; all fields of one cover type (corn, soybeans, else) were clustered together, using only channels two and four from each Landsat acquisition date. Two methods for subset selection were compared. Weighted and unweighted separability measures were used to select the best four of six or eight channels for use in classification. The unweighted separability measures considered the distance between all spectral subclasses in ranking the channels; the weighted separability considered only those spectral subclasses which were of different cover types. In the majority of the cases, the same subset was selected. If a different subset was selected, the weighted method produced classification results of higher accuracy.

Another aspect of the training procedure was the number of data points used for defining each of the spectral subclasses. In general, small clusters (less than 15-20 points) were deleted or combined with other clusters. In one analysis, however, several small classes appeared to be spectrally separable from all other cover types, so classification was carried out using the small classes. Classification accuracies were lower than anticipated, so some additional analyses were conducted. It was discovered that in deleting the small clusters, performance of the classifier consistently increased. Any clusters containing few points should be carefully examined before use in analysis.

6. Comparison of the Performance of Five Classification Algorithms

6.1 Objectives

The overall objective of this study was to apply several currently available classification schemes and to evaluate their performance on several agricultural data sets. The data sets were selected to include corn, soybeans, winter wheat, and spring wheat as major crops. Classification accuracy for test fields, ease of analyst use, and computer time required were compared for the various classifiers and data sets.

6.2 Approach

Test sites were selected from three major data sets: Fayette County (south central Illinois) from the CITARS data; LACIE Phase II data from 1976 over Foster County, ND, and Grant County, Kansas; and multicrop data from 1978 over the U.S. Corn Belt: Pottawattomie (886) and Shelby (892) Counties in west central Iowa, Tippecanoe County (854) in west central Indiana, and Iroquois County (824) in east central Illinois.

The segments sample several major crops: winter wheat in Kansas: spring wheat in North Dakota: and corn and soybeans in Indiana, Illinois and Iowa. The Corn Belt segments were located in two distinct regions to sample variability in soils, climate, and agricultural practices. Both areas are intensively cropped, with corn and soybeans being the predominant agricultural crops. Ground reference data and field maps as well as cloud-free multitemporally registered digital Landsat MSS data were available over these sites.

Four acquisition dates were selected for analysis from the most cloud-free, least noisy, and best registered acquisitions which temporally sampled the crop calendar to maximize crop development differences (Table A-5). For the Corn Belt segments, an attempt was made to obtain a spring acquisition to better separate winter small grains, trees and permanent pasture from row crops. An acquisition after corn had tasseled was included to separate corn and soybeans.

Since classification costs would be too high if all 16 bands of data

Table A-5a. Multitemporal Data Set Composition for the Corn and Soybean Test Sites.

Corn Development Stage	Test Site				
	Fayette	Pottawattamie	Shelby	Tippecanoe	Iroquois
	Date of Landsat Acquisition				
Emergence	6/10	6/16	6/16	6/10	6/12
Pretassel	6/29,7/17	-	-	-	-
Tasseling	8/21	7/23	7/23	7/26	8/5
Blister	-	-	8/9	-	-
Dough	-	-	-	8/21	-
Dent	-	9/6	-	-	8/31
Mature	-	9/24	9/24	9/26	9/28

Table A-5b. Multitemporal Data Set Composition for the Spring and Winter Wheat Test Sites.

Wheat Development Stage	Test Site	
	Grant	Foster
	Date of Landsat Acquisition	
Emergence	3/13	5/26
Heading	5/15	6/30
Soft Dough	6/2	7/19
Harvest	7/8	8/24

were used, classifications were performed using four bands selected to maximize the average transformed divergence between pairs of spectral subclasses. The acquisition dates and spectral bands selected are shown in Table A-6.

Five classifiers were selected for study:

- CLASSIFYPOINTS is a per point Gaussian maximum likelihood classifier. It is a processor from LARSYS, a remote sensing data analysis system developed at LARS [7].
- CLASSIFY is a sum-of-normal-densities maximum likelihood classification rule which first assigns each pixel into an information category and then assigns the pixel to a spectral subclass within that category. It is a processor from EODLARSYS, developed at NASA, Johnson Space Center [8].
- MINIMUM DISTANCE is a linear classification rule which assigns each pixel to the class whose mean is closest in Euclidean distance [9]. It is a processor from LARSYS.
- The LAYERED classifier is a multistage decision procedure [10]. It utilizes decision tree logic with an optimum subset of features at each tree node to classify each pixel, using a Gaussian maximum likelihood decision rule. LAYERED is also a processor from LARSYS.
- ECHO (Extraction and Classification of Homogeneous Objects) utilizes both spectral and local spatial information [11]. Statistical tests are used to group data into homogeneous regions and each region is then classified using a Gaussian maximum likelihood sample classification rule. It was also developed at LARS and is part of LARSYS.

In order to insure that differences in classification accuracies were the result of classifier differences and not training methods, the same set of training statistics was used for all classifiers. Training fields were selected to represent the classes of interest. These fields were clustered to develop means and covariances defining spectral subclasses

Table A-6. Spectral Bands Used in Classification.

Test Site	Landsat Acquisition Date	Spectral Bands Selected
		(μm)
Fayette	6/10	.6-.7
	6/29	None
	7/17	.6-.7, .8-1.1
	8/21	.6-.7
Pottawattamie	6/16	.8-1.1
	7/23	.6-.7, .8-1.1
	9/6	.7-.8
	9/24	None
Shelby	6/16	.6-.7
	7/23	.8-1.1
	8/9	.8-1.1
	9/24	.8-1.1
Tippecanoe	6/10	.6-.7, .7-.8
	7/26	.8-1.1
	8/21	.7-.8
	9/26	None
Iroquois	6/12	.7-.8
	8/15	.8-1.1
	8/31	.8-1.1
	9/28	.6-.7
Grant	3/13	.8-1.1
	5/15	.6-.7
	6/12	.6-.7
	7/8	.6-.7
Foster	5/26	.7-.8
	6/30	.7-.8
	7/19	.6-.7
	8/24	.8-1.1

for each of the classes of interest. Since CLASSIFY was designed as part of an automated analysis procedure without analyst intervention, a training method using a random selection of individual pixels to define initial cluster seeds for clustering the entire area is generally used in conjunction with that algorithm (ISOCLS). Both training methods were used with CLASSIFY.

The Fayette County site had reference data over approximately 25% of its area, while reference data were available for the entire area for the other sites. These data were sampled to define training and test data. Half of the selected fields were used for training the classifiers, and the remaining half were set aside for testing the classification results. Training was based on 1.6% of the area in the Fayette site, and between 3.5 and 7.5% in the other sites.

6.3 Experimental Results

The results of this study (Table A-7) were analyzed to assess the effects of segment and classifier on classification accuracy. Segment-to-segment variability was highly significant ($p<0.01$). Segment variability was attributed to factors other than the classifier selected, including spectral data quality and characteristics of the scene.

Several factors contributed to the lower classification accuracies obtained in Fayette County: (1) the quality of multitemporal registration was only marginal, (2) the acquisitions for Fayette were not as well distributed throughout the growing season as in the other counties, and (3) less training data were available for the Fayette site, and the training data available were not as well distributed or representative as in the other counties.

Pottawattamie and Tippecanoe Counties had larger field sizes, helping to account for the relatively accurate classification. Shelby County contained more confusion crops, including sorghum and spring oats, and had

Table A-7. Comparison of Classifier Performance (Percent Correct Classification) by Test Site.

TEST SITE	CLASS	MINIMUM DISTANCE	CLASSIFY POINTS	LAYERED	ECHO	CLASSIFY Using ISOCLS Stats[1]	CLASSIFY Using LARSYS Stats[2]	TEST SITE Average
Fayette, IL								
	Corn	81.9	81.2	63.9	77.3	77.3	78.9	76.8
	Soybeans	82.0	77.0	76.8	70.7	49.7	79.0	72.5
	Other	85.5	88.6	91.3	87.8	58.8	85.6	82.9
	Overall	83.5	83.0	80.5	79.5	61.1	81.6	78.2
Pottawattamie, IA								
	Corn	98.7	97.2	95.7	98.2	93.0	98.4	96.9
	Soybeans	92.0	89.8	92.3	90.2	86.5	89.3	90.0
	Other	85.3	98.0	97.5	97.1	92.1	98.4	94.7
	Overall	94.9	94.7	94.7	95.4	90.6	95.3	94.3
Shelby, IA								
	Corn	97.1	95.1	94.5	96.1	82.8	95.9	93.6
	Soybeans	89.3	92.9	98.2	95.4	98.0	98.0	95.3
	Other	75.5	83.7	88.2	79.4	78.7	79.7	80.9
	Overall	90.0	91.7	93.3	91.5	83.9	92.1	90.4
Tippecanoe, IN								
	Corn	93.7	89.9	91.5	86.4	99.4	93.1	92.3
	Soybeans	97.6	98.2	94.9	98.0	95.1	98.4	97.0
	Other	94.3	96.7	100.0	96.7	69.9	96.7	92.4
	Overall	95.5	94.3	94.0	92.7	94.2	95.9	94.4
Iroquois, IL								
	Corn	88.1	79.5	91.0	79.3	89.9	92.8	85.1
	Soybeans	82.8	85.2	78.1	83.6	78.8	86.3	82.5
	Other	76.4	72.7	0.0	72.7	74.5	75.0	61.9
	Overall	84.9	82.1	80.5	81.2	83.6	84.2	82.8
Foster, ND								
	Small Grains	96.1	95.4	94.6	94.8	93.6	97.3	95.3
	Other	73.3	77.1	77.0	77.6	70.5	82.3	76.3
	Overall	82.7	84.7	84.3	84.8	81.3	89.3	84.5
Grant, KS								
	Small Grains	96.9	96.7	97.6	96.5	94.6	98.7	96.8
	Other	91.8	83.2	89.3	79.2	92.0	80.2	86.0
	Overall	93.1	86.5	91.4	83.5	92.6	84.8	88.6

[1]Training method generally used with CLASSIFY. Uses a random selection of individual pixels to define initial cluster seeds for clustering the entire area.

[2]Training method used with all other classifiers. Training fields were clustered to develop means and covariances to define spectral subclasses for each of the classes of interest.

smaller field sizes than the other counties. Iroquois County had very few confusion crops and was almost entirely corn and soybeans, making it difficult to obtain training for cover types other than corn and soybeans.

There was no significant difference among classifiers in percent correct classification of corn, soybeans, or other in the five Corn Belt segments. In addition, there was no significant difference in overall accuracy among classifiers for all seven segments. The sum-of-normal-densities classifier using LARSYS statistics, however, have significantly higher small grain classification accuracy (about 2% improvement).

Table A-8 shows the percent correctly classified averaged over all segments for the different cover types. The performance of the ECHO classifier was not as high as anticipated, probably due to the fact that the ECHO classifier requires the analyst to set parameters defining cell size and homogeneity factors, and the optimal settings probably were not used. Although differences were nonsignificant overall, the LARSYS training method provided a consistent improvement over the ISOCLS training method in six of the seven segments. In conclusion, given a set of training statistics capable of producing high level classification results, the choice of classification algorithm for differentiation of corn and soybeans from other cover types makes relatively little difference.

Two additional features of the classification schemes were considered: the ease of use of the classification method and the computer time required for each classifier. The classification schemes varied considerably in ease of use. In increasing order of complexity the classifiers were found to be: (1) MINIMUM DISTANCE, (2) CLASSIFYPOINTS, (3) CLASSIFY, (4) ECHO, and (5) LAYERED. The MINIMUM DISTANCE and CLASSIFYPOINTS classifiers were almost identical in ease of use.

CLASSIFY was designed as part of a total analysis scheme in which participation of the analyst is minimized in the clustering and definition of training statistics, and control is provided by a predefined set of analysis parameters. Although the classifier itself is not extremely complex, the

Table A-8. Comparison of Average Percent Correct Classification for Several Classification Approaches.

MAJOR CROPS	NO. SEGMENTS	CLASS	Classifier: MINIMUM DISTANCE	CLASSIFY POINTS	LAYERED	ECHO	CLASSIFY Using ISOCLS Stats[1]	CLASSIFY Using LARSYS Stats[2]
Corn/Soybeans	5	Corn	91.9	88.6	87.3	87.5	88.5	89.8
		Soybeans	88.7	88.6	88.1	87.6	81.6	90.2
		Other	85.4	87.9	75.4	86.7	74.8	87.1
		Overall	89.8	89.2	88.6	88.1	82.7	89.8
Small Grains	2	Small Grains	96.5	96.0	96.1	95.6	94.1	98.0
		Other	82.6	80.2	83.2	78.4	81.3	81.3
		Overall	87.9	85.6	87.8	84.2	87.0	87.0

[1] Training method generally used with CLASSIFY. Uses a random selection of individual pixels to define initial cluster seeds for clustering the entire area.

[2] Training method used with all other classifiers. Training fields were clustered to develop means and covariances to define spectral subclasses for each of the classes of interest.

training procedure typically used in this scheme involves a large number of parameters about which little is known.

ECHO utilizes both temporal and spatial information. The complexity of use for ECHO arises from the necessity of setting the parameters for cell homogeneity testing and cell size. The expertise of the analyst is essential in setting the parameters with regard to data set used. The ECHO classifier is, however, one of the few available classifiers that utilize spatial as well as spectral information in the classification process.

LAYERED implements a per point Gaussian maximum likelihood decision tree logic which requires the additional step of designing the decision tree. The decision tree is designed by obtaining class means and covariance matrices for all classes and using a feature selection algorithm to determine an optimal subset of features to be used at each node of the decision tree. No feature should be deleted which is necessary to adequately discriminate a class of interest. The decision tree is then constructed using the best features for discriminating spectral classes. This decision tree is an input to the LAYERED classifier. The time needed by the analyst to design the tree using a multitemporal or multichannel data set is related to the complexity of implementation. If many spectral classes and features are needed to characterize the scene of interest, the decision tree can become very complicated and awkward to use. This classifier is particularly well suited for use with multitemporal or multitype data sets.

The computational cost is also an important variable in selecting a classification scheme. The computer time required per square kilometer for each segment and classifier is shown in Table A-9. In order of increasing cost per square kilometer for classification, not including cost for developing training statistics, were (1) MINIMUM DISTANCE (1.7 seconds), (2) ECHO (2.3 seconds), (3) LAYERED (2.3 seconds), (4) CLASSIFYPOINTS (3.7 seconds), and (5) CLASSIFY using ISOCLS statistics (11.3 seconds).

Table A-9. Computer CPU Time (seconds per square kilometer) Used by Each Classifier.

CLASSIFIER	TEST SITE							Average
	Grant	Foster	Tippecanoe	Fayette	Pottawattamie	Shelby	Iroquois	
Minimum Distance	2.3	1.7	1.3	1.5	1.6	2.3	1.4	1.7
Classifypoints	6.1	3.5	2.9	3.6	2.7	3.7	3.6	3.7
Layered	3.5	2.4	1.8	3.1	1.7	1.7	2.0	2.3
Echo	3.9	2.3	1.9	2.0	2.0	1.8	2.3	2.3
Classify(LARSYS Stat)	5.7	3.4	3.4	2.9	3.1	3.1	5.0	3.8
Classify(P-1 Stat)	10.7	12.6	12.7	8.0	12.8	8.4	14.1	11.3

6.4 Conclusions

The results of this study show little difference in the classification accuracies achieved by the five classification algorithms which were considered. However, the results for the CLASSIFY algorithm using two different training methods did show a difference. This indicates that the major variable affecting correct classification accuracy is not the classifier, but the training method used in generating the class statistics to be used in the classification. The most important aspect of training is that all cover types in the scene must be adequately represented by a sufficient number of samples in each spectral subclass.

The ISOCLS training algorithm was a method which was designed for machine automation of a large portion of the training procedure. The statistical sampling method used for selection of training data is theoretically sound, so it is possible that the lack of analyst refinement of the training statistics is seriously limiting the performance. The clusters produced by this method are of mixed cover types which may adversely affect performance.

Additional variables of interest in the study were complexity of use of the classifier and CPU cost per classification. Among the classifiers yielding similar classification accuracies, MINIMUM DISTANCE was the easiest for the analyst to use and costs the least per classification.

In summary, the classification performance of the five classification algorithms was found to be very similar when the same training method was utilized. The results suggest that development of representative training statistics is relatively more important for obtaining accurate classifications than selection of the classification algorithm.

7. Landsat Data Acquisition Study

A study of the impact of Landsat data acquisition history on classification was initiated. Its specific objectives were:

- Assess the accuracy of early season estimates.
- Determine a minimum number and distribution of acquisitions necessary for accurate estimation of corn and soybean areas.
- Determine the gain or loss by using a subset of channels over all channels in a unitemporal as well as multitemporal mode.
- Compare minimum distance, maximum likelihood, and sum-of-densities classifications in other band/date combinations than previously assessed.

The data set analyzed consisted of eight sample segments, selected to represent a broad range of conditions found in the Corn Belt. The segments were 843 and 860 in eastern Indiana, 837 and 854 in western Indiana, 862 and 883 in north central Iowa, and 886 and 892 in west central Iowa.

A modified supervised training approach was used. After refinement of the statistics was complete, the entire segment was classified using minimum distance, maximum likelihood, and sum-of-normal-densities classifiers. One acquisition from each of the four time periods previously defined was used. Data from all possible combinations of time periods were analyzed. One visible (.6-.7 μm) and one near infrared (.8-1.1 μm) band were initially selected for the multidate analyses. A subset of four bands, selected from the available six or eight bands on the basis of the maximum transformed divergence value, was also used for classification in analyses using three or four acquisitions.

7.1 Early Season Estimate Accuracy

The accuracy of early season estimates is illustrated in Figure A-3. During the first defined time period, corn and soybeans were not spectrally separable as indicated by the low overall classification accuracy (60.0%). In the Corn Belt, however, relatively accurate identification can be made of corn and soybeans together at that time. Over the same set of segments, it was found that overall identification into two classes (corn and soybeans,

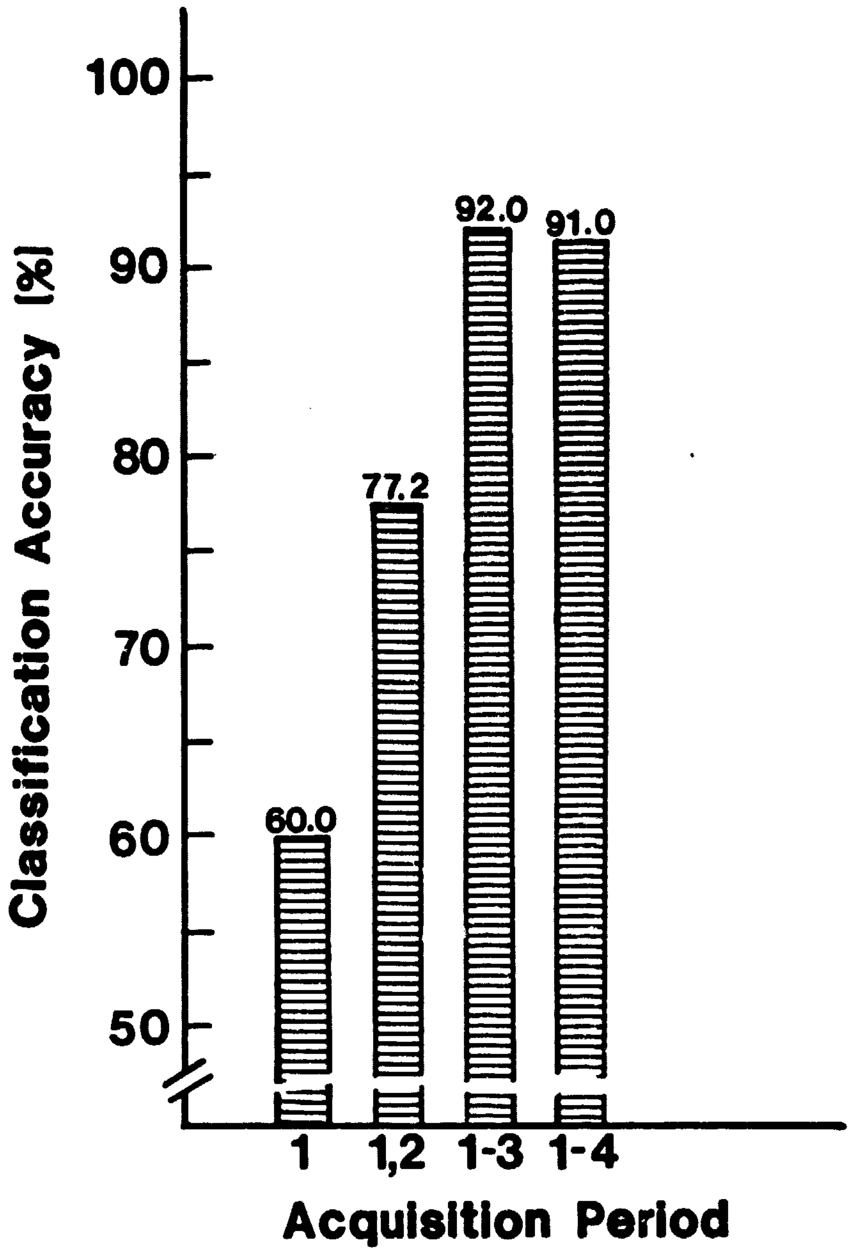


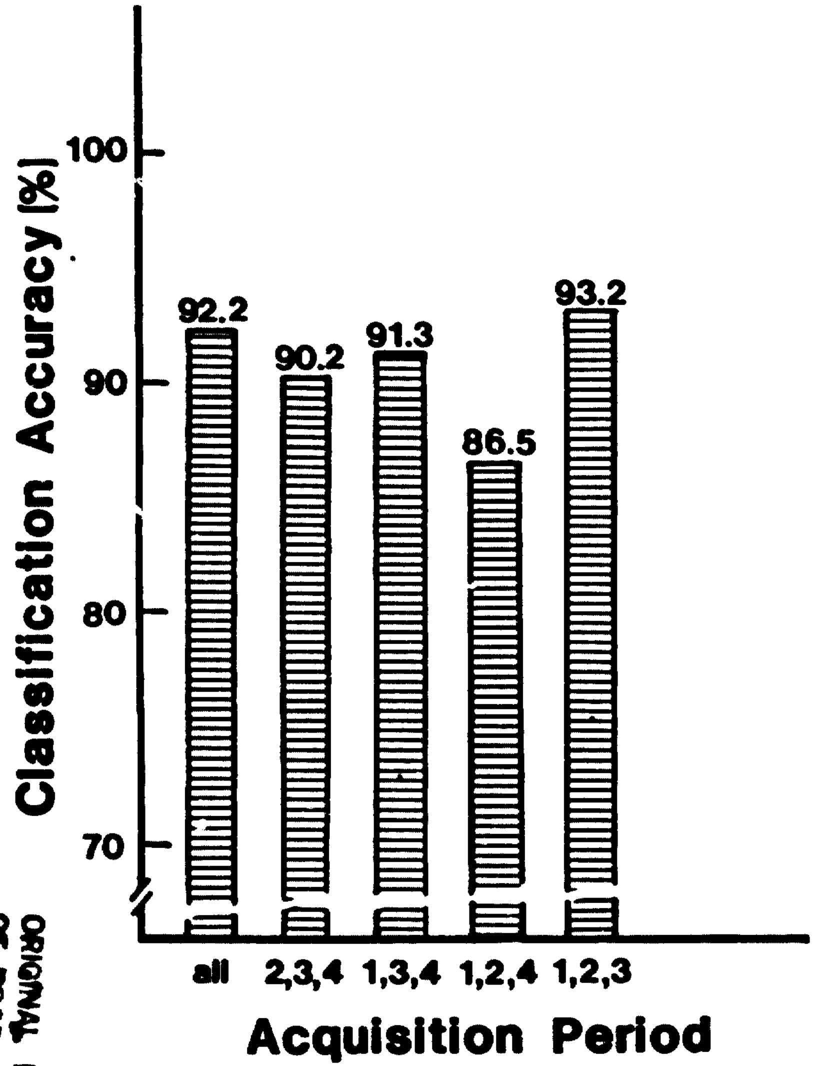
Figure A-3. Overall classification performance using cumulative spectral information with a minimum distance classifier and subsets of two, four, six, and eight channels.

else) was 92.0% correct, while the three-class classification (corn, soybeans, else) was only 60.0% correct. It is not until after the corn has tasseled (growth stage three) that consistently high classification accuracies are obtained. The classification accuracy does not improve by using later season information when the crops of interest have reached maturity.

7.2 Minimal Acquisitions Necessary

Figure A-4 illustrates the overall crop identification accuracies of classifications using two, three, and four Landsat acquisitions. A significant decrease in accuracy can be noted when the third period, tasseling to early dent, is omitted from the three date analyses. The importance of this growth stage can also be seen in examination of the two acquisition analyses; the three combinations using the third time period obtained higher overall accuracies then those without that growth stage represented. The overall accuracy of the third period alone was only 85%, illustrating that classification using the single best acquisition period is not as accurate as can be obtained using multitemporal information.

The following combinations of acquisition periods had overall accuracies which were not substantially different: 1,2,3,4; 2,3,4; 1,2,3; 1,3,4; and 1,3. These growth stage combinations had overall accuracies which varied by only 3%, and the next highest accuracy was about 3% lower than the lowest of these. It seems as though the availability of acquisitions from time periods one (about emergence) and three (after tasseling of the corn) provides a minimal set for accurate identification of corn and soybeans. No combination of acquisitions which does not include stage three gives high classification performance; a stage one acquisition appears to be less critical since growth stages two, three, and four together produce a relatively accurate estimate. The minimum number and distribution needed to obtain a good estimate of corn and soybean proportions has not yet been identified due to the lack of sufficient digitized inventories, but it is anticipated that the same pattern will hold.



ORIGINAL PAGE IS OF POOR QUALITY

Figure A-4a. Overall classification accuracies of three and four date classifications.

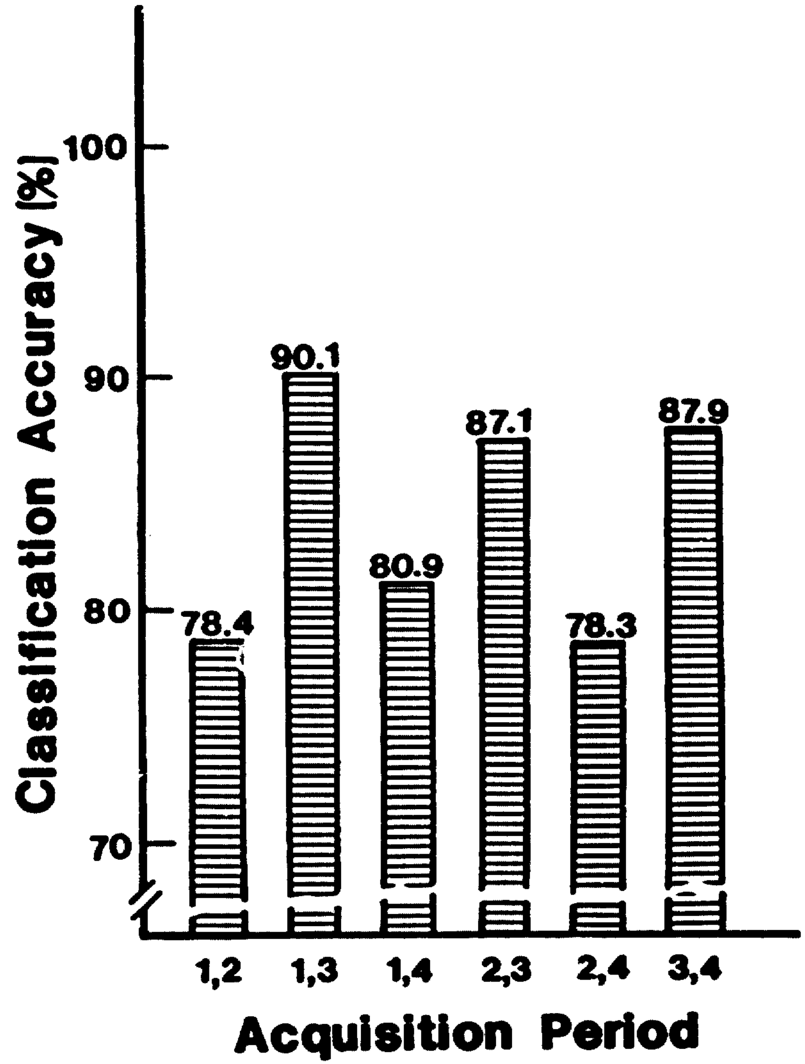


Figure A-4b. Overall classification accuracies of two date classifications.

7.3 Dimensionality Reduction

Landsat MSS channels two (.6-.7 μm) and four (.8-1.1 μm) from each acquisition (six for three date and eight for four date analyses) were compared with the best subset of four channels selected on the basis of the maximum transformed divergence value. The differences in accuracy were significant and, in general, all even channels (six or eight) gave higher classification performances than the use of a subset of four channels (Table A-10). Significant differences and the same trends held for variance reduction factors also. On the average, differences were relatively small (0-5%), but the loss in accuracy for a given segment with a particular combination of acquisitions could be quite large (one value of 10.7% was observed). In a few cases, the subset of four channels performed better. This occurrence was attributed to better defined training statistics resulting from the dimensionality reduction of the estimation problem or data problems in the bands not selected.

Single date classifications were conducted using two and four bands. Single date analyses were not conducted for growth stages one and two individually, so these two time periods were not assessed. In growth stage three, no significant differences in accuracy were found over all segments (83.1% vs. 83.0% overall accuracy). On an individual segment basis, there was a tendency for all channels to perform better (in six of eight cases). In two segments, the even channels gave higher accuracy, probably due to the misregistration of a band or noisy data in one of the wavelength bands. For growth stage four alone, the even channels gave 4% higher overall accuracy on the average, keeping this trend for four of the six available segments.

A second alternative exists for dimensionality reduction. Rather than selecting a subset of wavelength bands, a dimensionality-reduction transformation is computed using information from all of the bands. Such a transformation is one defined by the Tasseled Cap, using the first two components: greenness and brightness [12]. This analysis is in progress, but results are not yet available.

Table A-10. Overall Accuracies (percent) Obtained by the Maximum Likelihood Classifier for all Even Channels and a Subset of Channels.

Time Periods Analyzed	Averaged Over Segments: Subset	Averaged Over Segments: Even Channels	Averaged Over Segments: Difference	Maximum Difference
1,2,3	91.2	93.6	2.4	5.5
1,2,4	86.5	86.7	0.2	-2.5
1,3,4	88.2	91.6	3.4	7.6
2,3,4	85.4	90.2	4.8	10.7
1,2,3,4	89.2	92.1	1.9	9.0

7.4 Classifiers

A comparison of the minimum distance, maximum likelihood, and sum-of-densities classifiers is presented in Table A-11. Nonparametric statistical tests showed that the difference in overall classification accuracies was significant (α=.01), with the sum-of-densities classifier having the highest accuracy and the minimum distance classifier having the lowest accuracy. This pattern held for individual combinations of acquisition periods in general; in three combinations (3;1 and 3;2 and 4) minimum distance performed slightly better than maximum likelihood. Most of the performances were within about 2% for all classifiers, so classification costs (which increase in the same order performance was found to increase) should probably be considered in the choice of a classifier. The pattern of classifier performances remained fairly consistent over segments as well (Table A-12). Variance reduction factors for corn and soybeans were also analyzed, and the same pattern of performances was found.

The proportions of corn and soybeans estimated by each of the classification algorithms were compared. Averaged over dates and segments or averaged over segments alone, there was a trend in the proportions; minimum distance estimated the highest proportions for corn and soybeans, maximum likelihood was second, and sum-of-densities produced the smallest estimates of area for both cover types. The classifier producing estimates which are closest to ground inventory proportions has not been yet determined due to lack of sufficient digitized inventories.

8. Summary and Future Plans

This investigation has demonstrated that accurate identification and reliable area estimates of corn and soybeans can be made using Landsat MSS data. Some aspects of statistical sampling applied to classification have been examined, showing that wisely selected acquisitions and wavelength bands can lead to accuracies as high as the full season data set which is more costly to analyze.

Five classification algorithms were compared and little differences in

Table A-11. Overall Accuracies (percent) Obtained by the Minimum Distance, Maximum Likelihood, and Sum-Of-Densities Classifiers in Each of the Time Periods.

Time Periods Analyzed	Averaged over Segments			
	Minimum Distance	Maximum Likelihood	Sum-of-Densities	Range
3	83.1	82.9	83.4	0.5
4	72.3	72.7	74.9	2.6
1,2	77.2	77.9	79.6	2.4
1,3	86.4	85.2	87.4	2.2
1,4	77.5	78.4	81.3	3.8
2,3	85.2	86.6	87.8	2.6
2,4	78.4	78.2	79.6	1.4
3,4	85.6	86.5	88.4	2.8
1,2,3	92.0	93.6	93.9	1.9
1,2,4	85.6	86.7	87.2	1.6
1,3,4	89.6	91.6	92.7	3.1
2,3,4	88.8	90.2	91.6	2.8
1,2,3,4	91.0	92.0	93.7	2.7
Average	83.4	84.1	85.6	2.2

Table A-12. **Overall Accuracies (percent) Obtained by the Minimum Distance, Maximum Likelihood, and Sum-of-Densities Classifiers in Each of the Time Periods.**

	Averaged over Time Periods*			
Segment	Minimum Distance	Maximum Likelihood	Sum-of-Densities	Range
837	85.3	85.8	90.5	5.2
843	82.0	83.0	83.1	1.1
854	92.9	91.9	92.5	1.0
860	80.7	81.4	82.6	1.9
862	86.3	88.5	89.7	3.4
883	87.2	88.4	88.5	1.3
886	90.4	90.0	92.2	2.2
892	87.9	89.8	90.3	2.4

*Subset of channels in three and four time period combinations.

performance were observed with the training method used. Several methods for developing and refining training statistics have been examined. Further studies need to be conducted based upon the importance of the training step in obtaining good classification results.

This investigation will be continuing during the next contract year. Further studies on training unit size (fixed vs. variable) and training data selection (i.e., the use of ECHO as a training aid) will be conducted. The use of the brightness/greenness transformation will be compared with subset selection as a dimensionality reduction method.

A wider variety of segments across the U.S. Corn Belt and in the Corn Belt fringe areas will be classified. Characterization of the quality of the resulting estimates will be made based on the segment location and scene characteristics.

A study investigating sampling unit size and separation of the functions of sampling for training and sampling for area estimation is also planned.

9. References

1. MacDonald, R.B., and F.G. Hall. 1978. LACIE: An Experiment in Global Crop Forecasting. Proc. Plenary Session, The LACIE Symp., Houston, Texas, October 23-26, pp. 17-48.

2. Bauer, Marvin E., Marilyn M. Hixson, Barbara J. Davis, and Jeanne B. Etheridge. 1978. Area Estimation of Crops by Digital Analysis of Landsat Data. Photogr. Engin. 44:1033-1043.

3. Bauer, Marvin E., Jan E. Cipra, Paul E. Anuta, and Jeanne B. Etheridge. 1979. Identification and Area Estimation of Agricultural Crops by Computer Classification of Landsat MSS Data. Remote Sensing of Environ. 8:77-92.

4. Hanuschak, George, Richard Sigman, Michael Craig, Martin Ozga, Raymond Luebbe, Paul Cook, David Kleweno, and Charles Miller. 1979. Obtaining Timely Crop Area Estimates Using Ground - Gathered and Landsat Data. U.S. Dept. of Agriculture, Economics, Statistics, and Cooperatives Service. Technical Bulletin No. 1609.

5. Scholz, Donna, Nancy Fuhs, and Marilyn Hixson. 1979. An Evaluation of Several Different Classification Schemes: Their Parameters and Performance. Proc. Thirteenth Int. Symp. Remote Sens. Envir., Ann Arbor, Michigan, April 23-27.

6. Bizzell, R., F. Hall, A. Feiveson, M.E. Bauer, B. Davis, W. Malila and D. Rice. 1975. Results from the Crop Identification Technology Assessment for Remote Sensing (CITARS) Project. Proc. Tenth Int. Sym. Remote Sensing Envir., Ann Arbor, Michigan, October 6-10, pp. 1189-1196.

7. Phillips, T.L., ed. 1973. LARSYS User's Manual. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana.

8. Stewart, J., and P.J. Aucoin. 1978. Earth Observations Division Version of the Laboratory for Applications of Remote Sensing System (EOD-LARSYS) User Guide for the IBM 370/148 - Volume I. System Overview. NASA, Johnson Space Center, Houston, Texas JSC-13821.

9. Nilsson, N.J. 1965. Learning Machines: Foundations of Trainable Pattern-Classifying Systems. McGraw-Hill Book Co., New York.

10. Swain, P.H., C.L. Wu, D.A. Landgrebe, and H. Hauska. 1975. Layered Classification Techniques for Remote Sensing Applications. Proc. Earth Resources Survey Symp., Houston, Texas, June 9, Vol. I-b, pp. 1087-1097.

11. Kettig, R.L., and D.A. Landgrebe. 1976. Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects. IEEE Trans. Geos. Elect. 14:19-25.

12. Kauth, R.J. and G.S. Thomas. 1976. The Tasselled Cap--A Graphic Description of the Spectral-Temporal Development of Agricultural Crops As Seen by Landsat. Proc. Symp. Machine Proc. of Remotely Sensed Data, West Lafayette, Indiana, June 29-July 1.

B. INITIAL DEVELOPMENT OF SPECTROMET YIELD MODELS FOR CORN

C.S.T. Daughtry*

1. Introduction

As world demand for food continues to expand, increased pressures are placed on our agricultural systems to supply timely and accurate crop production information. The benefits of improved crop information include: (1) better utilization of storage, transportation and processing facilities, (2) more reliable crop production forecasts which allow decision-makers to plan policy better, and (3) increased price stability resulting from more accurate crop estimates.

Even at high levels of technology currently employed by most U.S. farmers, weather remains the most important uncontrolled variable affecting crop production and is the major cause of season-to-season variations in food production (Decker et al., 1976). During the past several decades numerous studies have attempted to develop models of the complex interactions between corn production, weather and technology. For simplicity, these studies generally considered weather and technology as independent factors in multiple-curvilinear regression models (Nelson and Dale 1978a). While these statistical models explained much of the variability in long-term crop production, they could not handle severe and unusual weather conditions or pest outbreaks (Nelson and Dale 1978b). The Thompson (1969) corn models and the wheat models of Large Area Crop Inventory Experiment (Strommen et al., 1979) are examples of statistical models.

Several alternative approaches to crop yield estimates have been developed which describe crop development and yield in physiological logic. These models are designed to simulate responses of basic plant

---

*The contributions of M.E. Bauer, D.A. Holt, C.D. Jobusch, V.J. Pollara, H.F. Reetz, C.E. Seubert and R.A. Weismiller to Task 2B, Initial Development of a Spectromet Yield Models for Corn, are gratefully acknowledged.

processes and, ultimately, yields to the environment. Some of these simulation models are too complex and detailed for large area crop yield estimations while others appear to be applicable and are currently being developed by Purdue University in conjunction with industry. Examples of complex crop simulation models are SIMED (Holt et. al., 1975) and CORN-CROPS (Reetz, 1976).

Intermediate to the classical statistical approaches and the causal physiological approaches are several models which rely on physiological logic to interpret the effects of weather on crop yields. These intermediate models tend to be less complex than physiological simulations like CORN-CROPS but more complex than LACIE's models. The Energy Crop Growth model (Dale and Hodges, 1975) and Purdue Soybean Simulator (Holt et. al 1979) are examples of approaches which seek to condense the effect of weather into a single weather index which can be related to yields.

Considerable evidence indicates that remote sensing can provide information about crop condition and thus yield potential (Bauer, 1975). If this spectral information about crops can be combined effectively with meteorological and ancillary data, then potentially much better information about crop production could be gained.

2. Objectives

The overall objective of this task represents a multiyear research effort to integrate the best mix of spectral, meteorological, and ancillary data into a crop information system for estimating crop condition and expected yield during the growing season. Specifically this task will:

- Identify important factors in determining and predicting corn yields.
- Determine how these factors can be observed or estimated from alternate sources of data.
- Define long-term data requirements for continued model development.
- Select and further develop several candidate approaches for corn yield modeling.

- Identify and obtain data required for these yield models.
- Conduct initial calibrations and tests of models using spectrometer and Landsat MSS data.

3. Description of Data

Two sources of spectral data were used in this task during the past year. Initial examination of relationships between spectral and important agronomic factors related to yield were performed using data acquired by the Exotech 20C spectrometer at the Purdue Agronomy Farm (Walburg, et al. 1979). Spectral and supporting agronomic data were acquired through the growing season on the Corn Nitrogen Fertilization Experiment of Dr. S.A. Barber. The corn in this experiment received either 0, 67, 134, or 202 kg N/hectare and had grain yields which ranged from 2910 to 8892 kg/ha (46 to 142 bushels/acre).

The other major source of spectral data was Landsat MSS data acquired over commercial corn fields in nine 5 x 6 mile segments located in six states (Figure B-1). Within each of these segments up 10 corn fields were identified and periodically observed throughout the growing season by personnel of USDA's Agricultural Stabilization and Conservation Service (ASCS) (Table B-1). These observations consisted of notes on plant height, percent soil cover, maturity stage, and recent field operations. Grain yield in each field was either estimated by the ASCS representative or acquired during an interview with the farmer. Grain yields ranged from 50 bushels per acre in Ballard, KY to 158 bushels per acre in Iroquois, IL. Data on planting dates of these fields were not obtained.

4. Results and Discussion

4.1 Factors Influencing Crop Yields and Prediction of Crop Yields

The economic end-product of crop production is often the seed which comprises about 45 percent of the above ground dry weight of corn. This accumulation of dry matter requires not only the availability of the

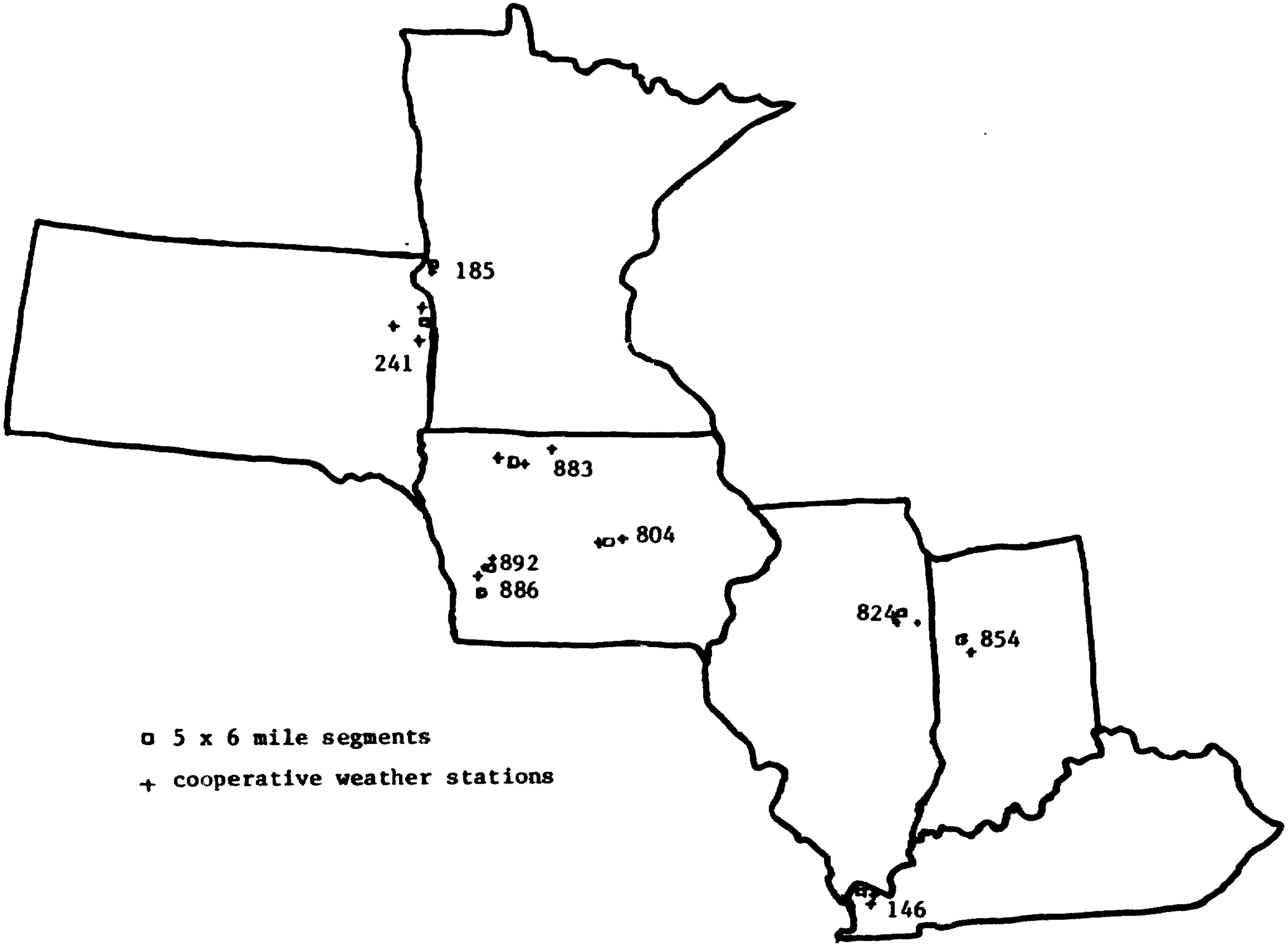


Figure B-1. Location of segments and their corresponding weather stations in 1978.

Table B-1. Dates that Landsat MSS data were acquired over corn fields which were periodically observed by ASCS personnel in 1978.

Segment No.	County, State	No. of Corn Fields	Julian Dates of Landsat Acquisitions
146	Ballard, KY	4	180, 198, 234, 270, 306
185	Traverse, MN	9	169, 187, 196, 205, 214, 223, 232, 241, 269, 287, 296
241	Deuel, SD	9	169, 187, 196, 205, 223, 232, 241, 269, 296
804	Marshall, IA	9	166, 220, 229, 247, 265, 274,292
824	Iroquois, IL	10	163, 217, 235, 243, 271, 297,306
854	Tippecanoe, IN	10	161, 197, 207, 216, 233, 243,251, 269, 305
883	Palo Alto, IA	8	186, 204, 213, 221, 258, 267,293, 303
886	Pottawatomie, IA	9	167, 186, 204, 212, 249, 258,267, 293
892	Shelby, IA	8	167, 204, 212, 221, 240, 249,266, 293

proper substrates (CO_2, H_2O, NH4+ and/or NO_3^-, and other nutrients) in the environment but also a great deal of energy which the plant derives from sunlight.

In modeling crop yields by any method, the following four types of factors influence yields:

1) crop factors - e.g., photosynthetic rate, stress tolerance, leaf area index, leaf area duration, growth rate
2) soil factors - e.g., drainage, water-holding capacity, fertility
3) management factors - e.g., planting date, weed, disease, and insect controls, cultivar selection.
4) weather factors - e.g., solar radiation, air temperature, precipitation, evaporation.

Man has exhibited varying degrees of control over the first three of these factors, but weather over which he has the least control remains the most important factor influencing year to year variations in crop production.

If weather is truly the most important factor controlling crop yields, how can the effects of weather on crop response (yield) be quantified? Reviews of research on environmental and physiological aspects of crop yield have identified and generally attempted to quantify optimum conditions for assimulation processes, growth, development, and ultimately yields for various crops (Eastin, 1969; Pierre et al. 1966). Rather than discuss how physical measures of the environment influences crop response, the reader is referred to any of several review on crop physiology and yields (Kramer, 1969; Hill et al., 1978; Decker et al., 1976; Eastin, 1969).

Of the various physical measurements of the environment, temperature, moisture and solar radiation are most frequently used to estimate crop yields. Researchers have used various experimental techniques to relate hourly, daily, weekly or monthly means of temperature, moisture (precipitation or soil moisture) and/or solar radiation to yields. Some have used selected weather variables from the entire growing season (Thompson, 1969) while other have preferred to identify physiologically important periods during which they felt crops were most sensitive to the effects of weather (Leeper et al., 1974;

Dale and Hodges, 1975; Nelson and Dale, 1978). While these fitted parameters may be associated with reasonable proportions of the variance in fitted crop yield series, the predictive equations generally explain disappointingly little of the crop yield variance in independent tests.

In addition to these yield models with empirical functions of weather variables, crop yields have also been estimated from within season sampling of crop dry matter and stand parameters. These methods use the crop as an integrator of weather effects, and then measure various plant characteristics at specific development stages which are related to grain yields. Prior to harvest estimates of crop yields by USDA-ESCS are based on similar techniques. These methods tend to become more accurate as crop maturity and harvest approaches.

4.2 Data Requirements and Sources of Data

Data requirements for crop model development vary greatly depending on the specific type of model employed. I have chosen to limit this discussion to those yield models which employ weather data (physical measures of the environment) directly or indirectly to estimate other quantities or which use remotely sensed measures of plant condition.

The most commonly recorded physical measures of the environment are daily maximum and minimum air temperatures and daily total precipitation. Less common measurements include solar radiation, evaporation, wind travel, soil temperatures and soil moisture on daily and in some cases hourly basis. These data are frequently used in crop models either by design or necessity since other data are available only in special instances.

Variability of precipitation patterns in time and space makes precipitation both the most important and most error-prone in any water budget or weather and crop yield study. The standard 8-inch precipitation gauge of the National Weather Service stations samples only 3.2×10^{-6} hectare and it is commonly used to represent county-size areas. The space-time variability of precipitation patterns in Illinois (Huff, 1971) probably represent the magnitude of variability in precipitation to be expected in other areas of the Corn Belt. Thus more than one precipitation station in close proximity to or within each 5 x 6 mile segment is desirable.

While average rainfall is more frequently used to identify the moisture situation in county or state corn yield studies, soil moisture in the root zone is more meaningful for crop growth studies. Much rainfall may run off, percolate through the soil profile or otherwise become unavailable to plant roots. This has been recognized, but the great variability of soils and sampling problems in measurement of soil moisture make it difficult to establish a representative and homogeneous series of soil moisture data. Several soil moisture estimating methods have been developed. Shaw (1963) described a method for estimating soil moisture in well drained soils and Stuff and Dale (1978) developed a method for poorly drained soils. Both appear to work reasonably well for their particular areas and soils.

Other commonly measured weather variables tend to be more conservative elements (or less time-space varying) than precipitation (Dale and Hodges, 1975). Thus one station per segment or county should be adequate for air temperature, solar radiation and pan evaporation.

In addition to these environmental measurements, information is also needed on the crop itself for yield model development. Each model has different requirements and one data set cannot satisfy all of them. A minimum set of observations about the crop in each location is desirable. This data set should include the following:

1. <u>one time per season</u>
   - planting date
   - harvest date
   - yield
   - cultivar or hybrid planted
   - fertility program, especially amount of N applied
   - row width

2. periodic observations at 7-14 day intervals during the growing season
   - maturity stage
   - plant height
   - field operations
   - crop condition (weeds, disease, hail, etc.)
   - irrigation times and amounts

3. additional data - for more detailed studies
   - soil type and drainage
   - percent soil cover
   - soil moisture
   - harvest losses in field
   - biomass
   - leaf area index

Since crop response to weather may differ from year to year, a homogeneous series of crop and weather factors are required for continued model development.

4.3 Approaches for Crop Yield Modeling

A conceptual framework of a large area crop information system has evolved during this task. This framework provided overall mathematical expressions for computing production estimates. Crop production was separated into its components, and major tasks which must be accomplished to arrive at a production forecast were identified. The kinds of information that must flow to each component and the potential sources of such information were listed.

Crop production consists of a yield component and an acreage component. The acreage of a crop can be estimated by ground surveys or as in the Large Area Crop Inventory Experiment (LACIE) by the use of Landsat MSS data. Yield of a crop may be computed as the product of four general factors as follows:

Yield = Yield Potential * Weather Factor * Episode Factor
* Management Factor *

where,

Yield Potential represents the yield that would be obtained on a given area with its particular soil conditions if the yield were not limited by weather, episodes of diseases and insects, or management conditions that were peculiar to that particular year.

Weather Factor is a number between 0 and 1 representing the limitations imposed on yield by weather conditions prevailing during that season.

Episode Factor represents a number between 0 and 1 representing the limitations placed upon yield by infestations of diseases or insects or by catastrophic weather conditions, such as hail, floods, or high winds.

Management Factor is a number representing the average impact of management decisions made in that particular area which causes the general level of management to differ from other years.

These four factors and acreages which when multiplied together can provide a crop production estimate. Accurate estimates of each component are required to achieve an accurate forecast. Obtaining an accurate estimate of each of these components is a separate project and these projects may serve as the basis for organizing a crop production forecasting system. This task (Initial Development of Spectromet Corn Yield Model) has focused on how remote sensing technology can provide information on "yield potential" (e.g., soil productivity) and "weather factor" (e.g., crop development and condition).

Yield Potential

Yield potential as defined earlier in this section can be estimated either indirectly from historical or directly from soil productivity indicies. Indirect estimates of yield potential can be derived as follows:

$$\text{Yield Potential} = \frac{\text{Historical Yield}}{\text{Weather Factor} * \text{Episode Factor} * \text{Management Factor}}$$

This estimate of yield potential for a particular area can be expected to remain rather constant from year to year. Long-term changes in yield potential are expected as new technologies are adopted or as soil productivity changes causing general trends in yields for an area. This approach to potential yield requires several years of data on yields, weather, management and episodes for each area in question.

Alternatively yield potential could be estimated directly from soil productivity indices by using existing soil surveys or potentially from remotely sensed information. Soils differ in their inherent capability to produce crops. Although proper management in some cases can compensate for deficiencies in native productivity of soils, differences in crop yields which are related to soil characteristics do occur.

Soil texture and organic matter content are important components in assessing native soil productivity. Soil drainage classes which are indirectly related to soil texture and organic matter content are identifiable from Landsat MSS data. Thus, potentially Landsat MSS data could be used to estimate soil productivity based on soil drainage classes.

Corn yield potential was estimated for soils in Tippecanoe (segment 854) and selected areas in Jasper Counties in Indiana by the methods of Walker (1976). Multivariate regression analyses of these data sets using yield potential as the dependent variable and soil spectral classes from Landsat MSS data as the independent variables were performed. Only 17 per cent of the variation in yield potential was associated with the spectral classes of these soils. Inclusion of indicator variables for texture in the regression model, along with the spectral class information, accounted for about 68 percent of the variation in yield potential. However, correlations of soil particle size (texture) with spectral response data has not been very high (Montgomery et al., 1976). Further research into methods of directly assessing yield potential with remotely sensed data is planned.

Weather Factor

Limitations imposed on crop yields by weather conditions have been depicted with varying degrees of success by several different mathematical models. The three basic types of models include:

1) Simulation or causal models which describe crop performances as a series of functions with daily solar radiation, air temperature, and moisture. Simulation models are broadly applicable, require short historical data bases for development, and can provide local detail. Examples of simulation models are SIMED (Holt et al., 1975) and CORN-CROPS (Reetz, 1976).

2) Statistical or correlative models which are equations with statistically-derived coefficients that represent the relationship between weekly or monthly mean weather and crop performance. These have been used successfully in LACIE. They are generally useful for crop reporting district (CRD) or larger areas and require long historical data bases to derive their coefficients. (Strommen et al. 1979, Thompson, 1969).

3) Hybrid models which seek to combine some of the best features of both simulation and statistical models by condensing the effects of weather on crops into a single weather index which can be related to yield (Holt et al., 1979; Nelson and Dale, 1978).

Each of these basic model types has potential to utilize spectrally-derived information. For example, in simulation models this information may be used as independent verification of model estimates of crop biomass, maturity stage, and/or yields. Since statistical models require coefficients derived from several years of homogeneous data sets (including yield, weather, and spectral data) which may not be available, the use of spectral data as an integral part of a statistical model is probably not possible. An example of an alternative approach would be to estimate with spectral data one of the variables in a statistical model and then substitute this spectrally-derived variable (when available) into the model. Hybrid models possibly can use both of the above approaches.

4.4 Initial Calibrations of Models

Initially these models will be calibrated and tested without the use of spectral data to establish their baseline performance in a bootstrap approach. The models will be calibrated using historical county average yields from USDA-ESCS, but will be tested using average yields in 10 corn fields per segment in the county in 1978. This step has been delayed because of difficulties encountered in acquiring historical meteorological data, but should proceed rapidly now that meteorological data for the first ten segments has been received.

After modification to include spectrally-derived information these models will be tested, if possible. Because long term data sets exist for corn yields and weather variables but not for spectral data, complete sets of test data exist only for selected sites in 1978 and possibly 1979. This lack of data will hamper conventional statistical tests of these model's performance with and without spectral data. By normalizing for soil productivity and substituting locations for years, some inferences about model performance possibly can be made. More years of complete data sets (yield, spectral, meteorological, and ancillary data) are required for adequate evaluation of these models.

A first step toward incorporating spectral data into any of these models requires an understanding of the spectral characteristics of corn canopies. Task 1A (Experiment Design and Data Analysis) examined spectrometer data acquired at Purdue Agronomy Farm in 1978. These data were analyzed to determine the basic spectral characteristics of corn and to assess how agronomic treatments affect these spectral characteristics. An expansion on these analyses used spectral data representing the four Landsat MSS bands to predict leaf area index (LAI) (Figure B-2) and percent soil cover (Figure B-3).

These two pieces of information about crop condition may be used, for example, to calculate intercepted solar radiation for the Energy-Crop-Growth (ECG) Model (Dale and Hodges, 1975). The solar radiation intercepted by a corn canopy was estimated as a function of leaf area index (Figure B-4A) and total solar radiation incident on a horizontal surface. This provides a continuous LAI weighting of solar radiation within the season. Leaf area index is estimated from Figure B-4B which represents seasonal values of LAI for different populations of corn plants. These LAI values are "visually-smoothed" averages from several researchers. Actual LAI for fields may very greatly due to different planting dates, hybrids, stresses, and row spacings. An estimate of intercepted solar radiation based on spectral derived LAI or soil cover percentages should more accurately depict conditions in the field. The corn cultural practices experiment of 1979 (see Volume 1, Task 1B) should be an excellent data set with its three plant populations and three planting dates to test this concept.

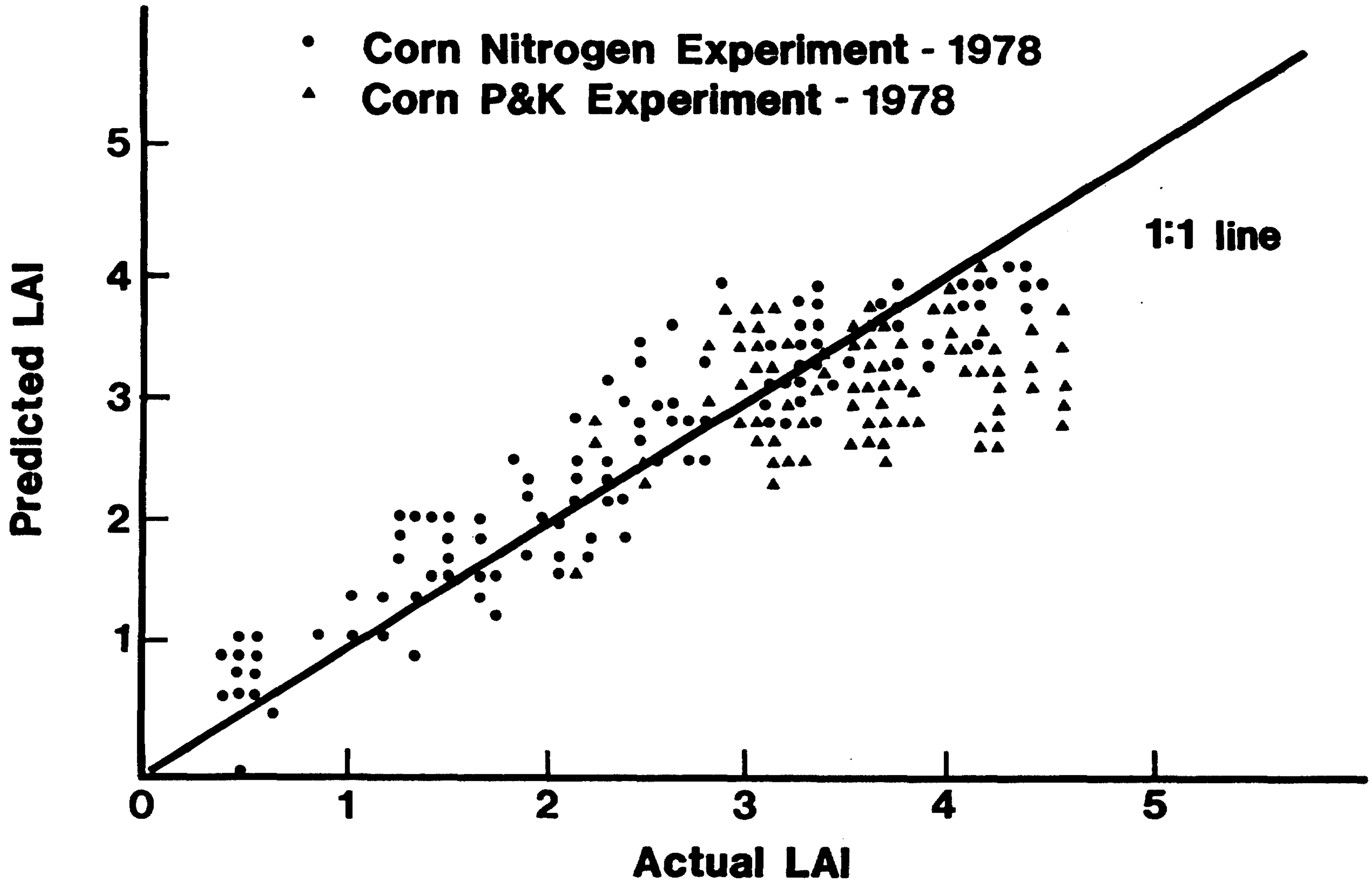


Figure B-2. A comparison of measured leaf area index (LAI) and LAI predicted from spectral data in the four Landsat MSS bands for two experiments at Purdue Agronomy Farm in 1978. The coefficients of the regression equation were derived with data from the Corn Nitrogen Experiment and were plotted with data from both experiments. $\widehat{LAI} = 0.523 - 0.953 * B50 + 0.399 * B60 + 0.154 * B70 + 0.380 * B80$.

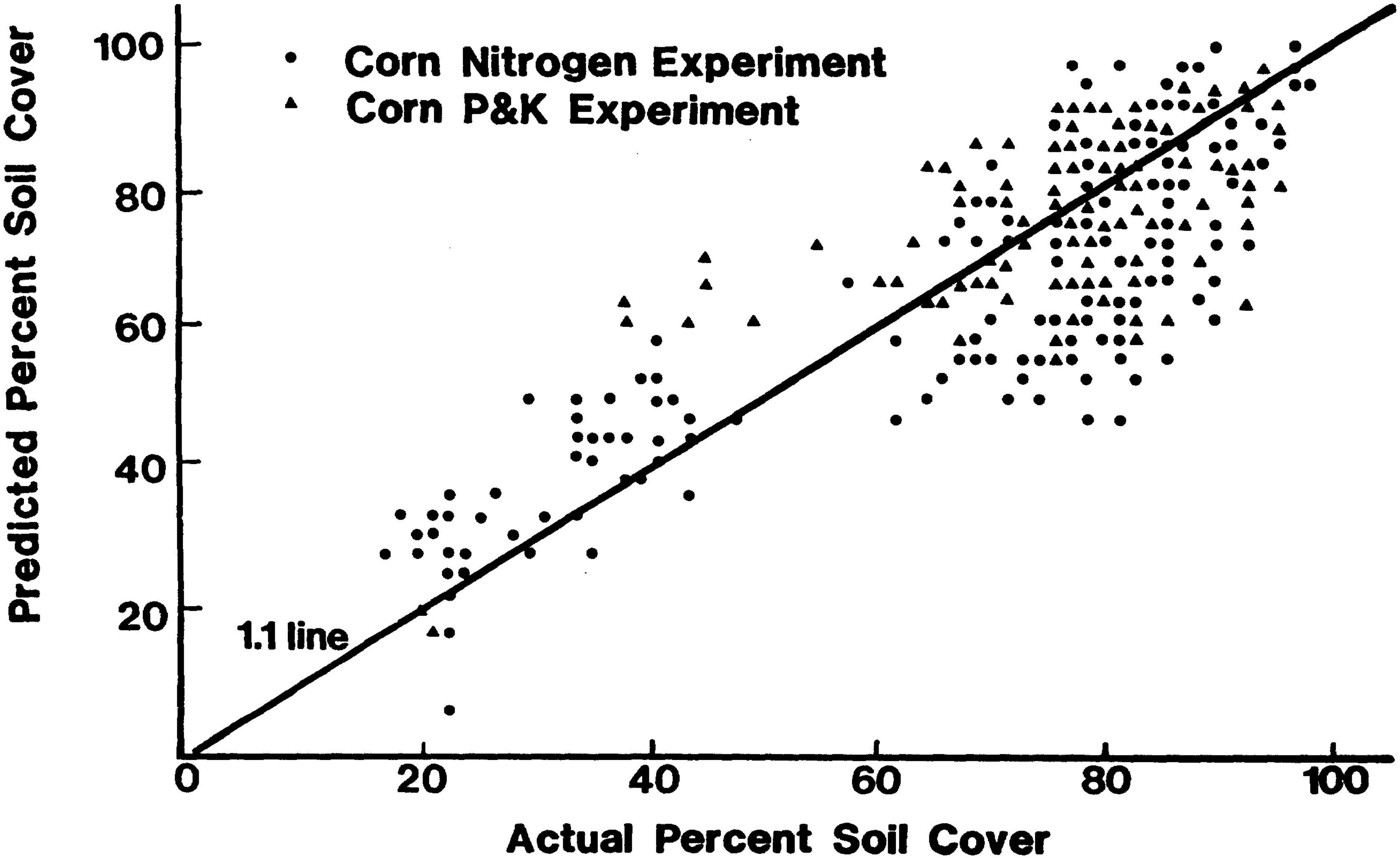


Figure B-3. A comparison of measured soil cover and percent soil cover predicted from spectral data in the four Landsat MSS bands for two experiments at Purdue Agronomy Farm in 1978. The coefficients of the regression equation were derived with data from the Corn Nitrogen Experiment only and were plotted with data from both experiments.

$\widehat{Cover} = 30.9 - 32.2 * B50 + 16.4 * B60 + 5.1 * B70 - 0.09 * B80$

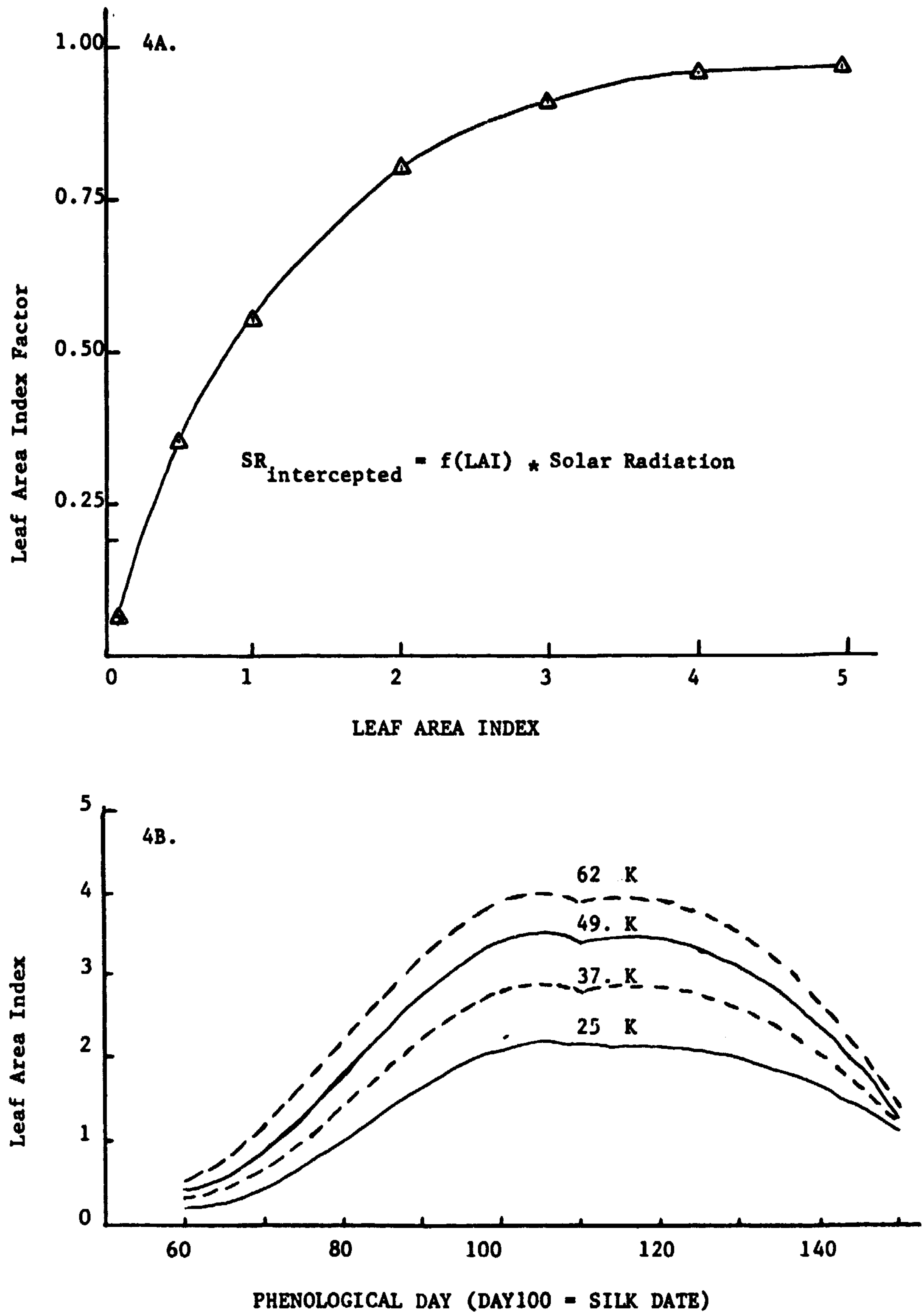


Figure B-4. The solar radiation intercepted($SR_{Int.}$) by a corn canopy was estimated as a function of leaf area index and total solar radiation(SR) received on a horizontal surface. The average seasonal leaf area index curves (B-4B) were visually smoothed from experimental data for 25,000, 37,000, 49,000, and 62,000 corn plants per hectare (from Dale and Hodges, 1975).

Regardless of which crop model is employed, its spatial resolution is limited by the distribution of weather stations. The best estimate of yield that can be expected from any of these models is the mean of a region. If there exists considerable variation in yields within a region due to, for example, soil fertility then these models are not likely to estimate yields very precisely or accurately at the local level. Spectral data, on the other hand, is limited by the spatial resolution of the sensor which 0.45 ha for Landsat MSS and 4 m^2 of Exotech 20C spectrometer at 10 m above the soil.

Figure B-5 and B-6 illustrate the departures of individual plot yields from mean yield due to nitrogen fertility and how some of this variation about the mean is associated with two spectral variables such as the ratio of reflectances in 0.8-1.1 and 0.6-0.7 μm bands and the greenness transformation. These relationships appear to be rather stable for 4 to 6 weeks during the tassling and grain filling periods of corn (Table B-2). From this limited data set it appears that this period occurs at or shortly after the time when the maximum IR/red ratio of corn is reached (Figure B-6). Together Figures B-5, B-6, and B-7 represent a potential method, not only to adjust yield predictions from meteorological models, but also to identify the time interval when remotely-sensed data are most highly correlated with corn yields.

Extension of these simple concepts developed from spectrometer data gathered at an agricultural experiment station to Landsat MSS data acquired over commercial fields represented quantum leaps in scene complexity and potential sources of unaccounted for variability. Initial examinations of the Landsat MSS data from selected corn fields indicated that maximum Kauth Greenness occurred at or shortly after tasseling (Figures B-8 and B-9) as expected from spectrometer data (Figure B-7).

Figures B-8 and B-9 represent typical fields of corn in Pottawatomi County, Iowa and Tippecanoe County, Indiana and have basically similar shapes. The abrupt changes in greenness over a two day period are data from consecutive day passes with Landsat MSS. The influence of the

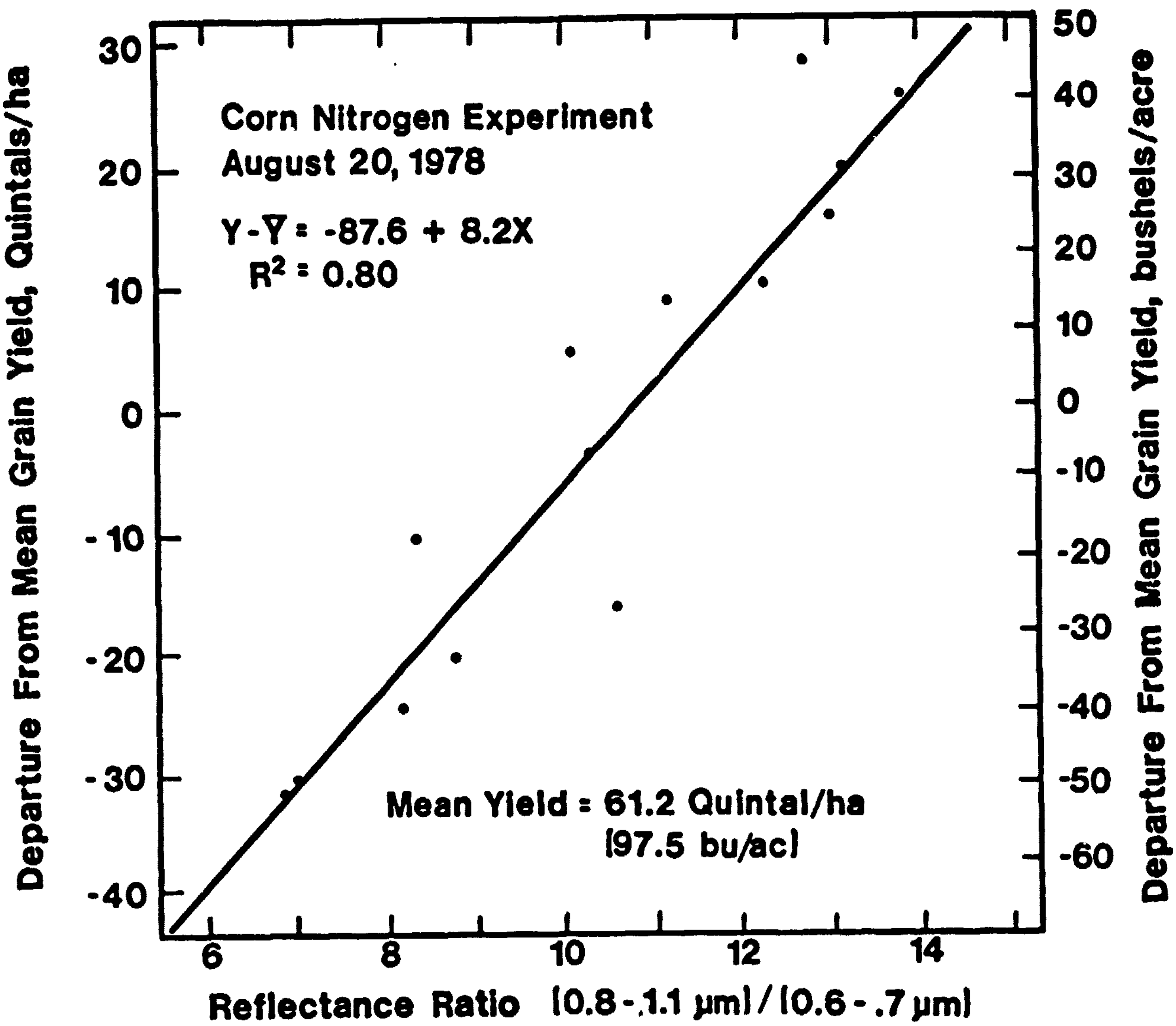


Figure B-5. Association of the ratio of reflectances in the near infrared (0.8 - 1.1 μm) band and the red (0.6 - 0.7 μm) band with departures from mean grain yield for the Corn Nitrogen Experiment in 1978.

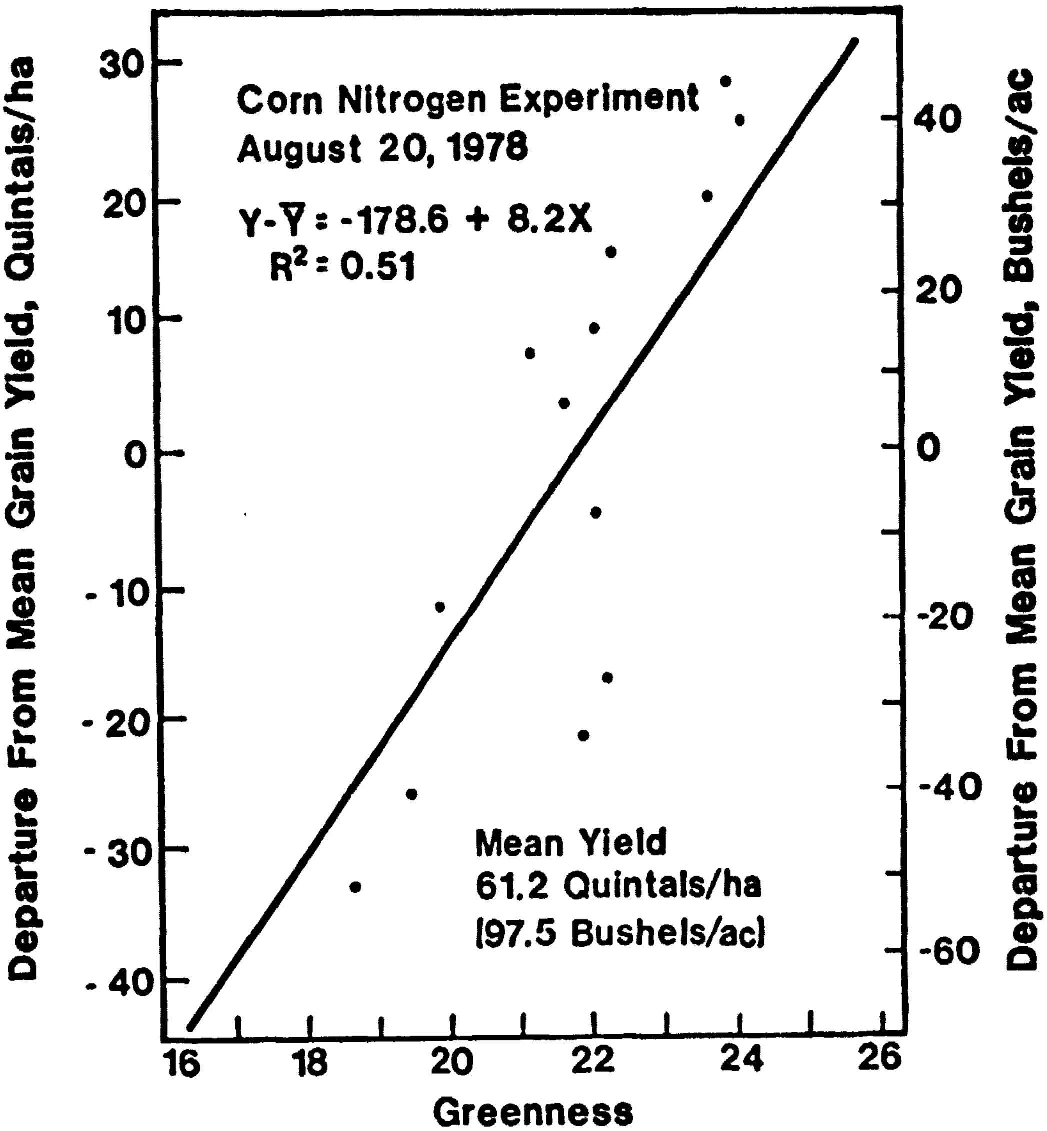


Figure B-6. Association of the greenness transformation of spectrometer data with departures from mean grain yield for the Corn Nitrogen Experiment in 1978.

Table B-2. Variation in corn grain yields associated Kauth Greenness and infrared (0.8-1.1 μm) to red (0.6-0.7 μm) ratio at several dates during the growing season for the Corn Nitrogen Experiment in 1978.

Date	Maturity Stage[1/]		Greenness[2/]	IR/Red
			---- R^2 ----	
June 28, 29	1.5	6-leaf	0.38	0.02
July 5	2.0	8-leaf	.50	.47
July 6	2.0	8-leaf	.21	.34
July 15	2.3	10-leaf	.38	.63
July 28	3.5	14-leaf	.45	.71
Aug 3	5.9	silk	.28	.64
Aug 16	-	blister	.42	.75
Aug 20	6.3	milk	.51	.80
Aug 31	7.0	dough	.55	.73
Sept 15	8.0	begin dent	.28	.55
Sept 23	9.0	full dent	.15	.32

1/ Hanway, J.J. (1966)

2/ Greenness = - 0.489*B50 - 0.612*B60 + 0,173*B70 + 0.595*B80

where: B50 = 0.5 - 0.6 μm wavelength band reflectance
B60 = 0.6 - 0.7 μm wavelength band reflectance
B70 = 0.7 - 0.8 μm wavelength band reflectance
B80 = 0.8 - 1.1 μm wavelength band reflectance

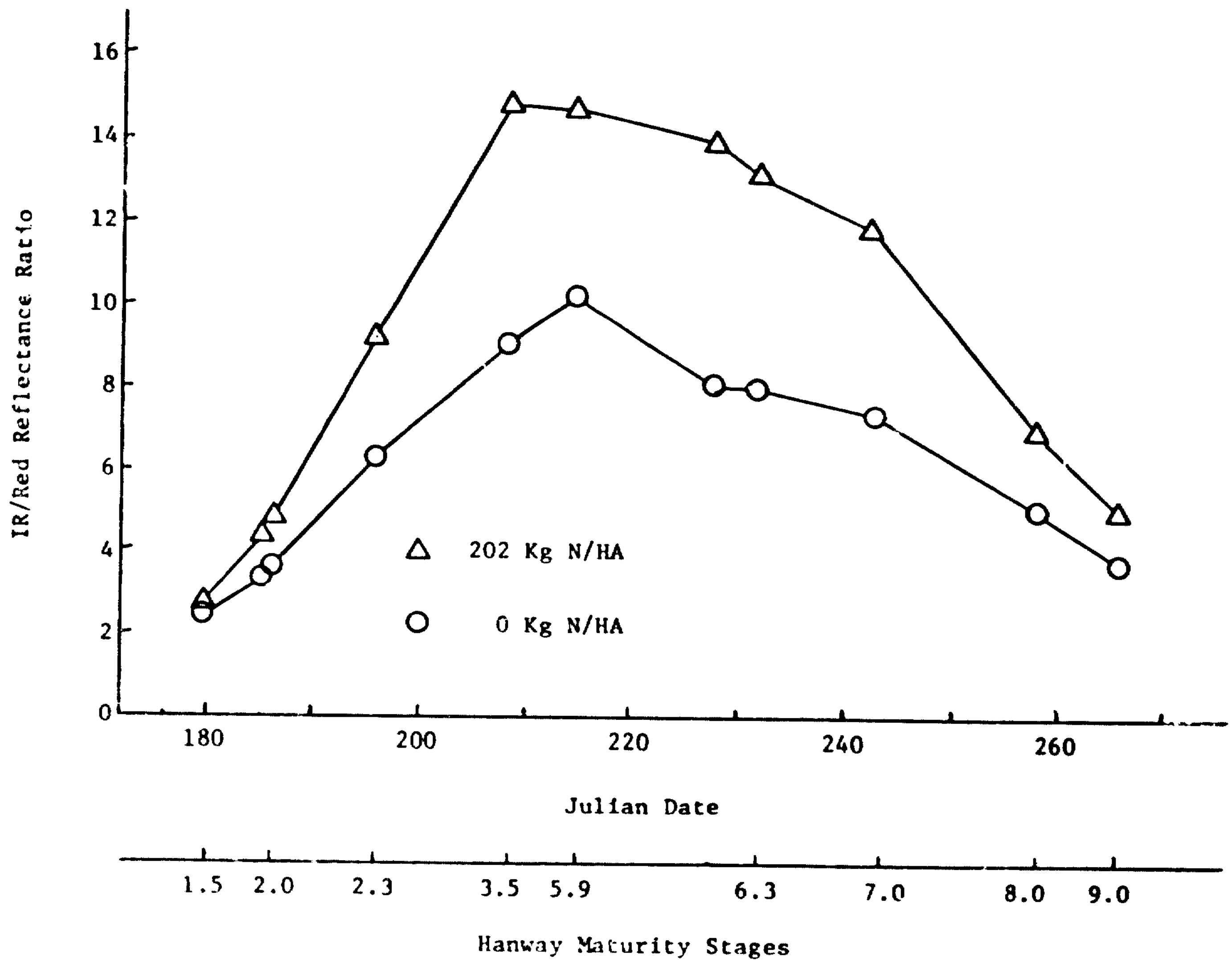


Figure B-7. Seasonal changes in ratio of reflectances in a near infrared (0.8 - 1.1 μm) band and a red (0.6 - 0.7 μm) band for the Corn Nitrogen Experiment in 1978. Note that the maximum reflectance ratio occurs near time of tasseling (Maturity Stage 5). Only high and low N treatments are shown for clarity.

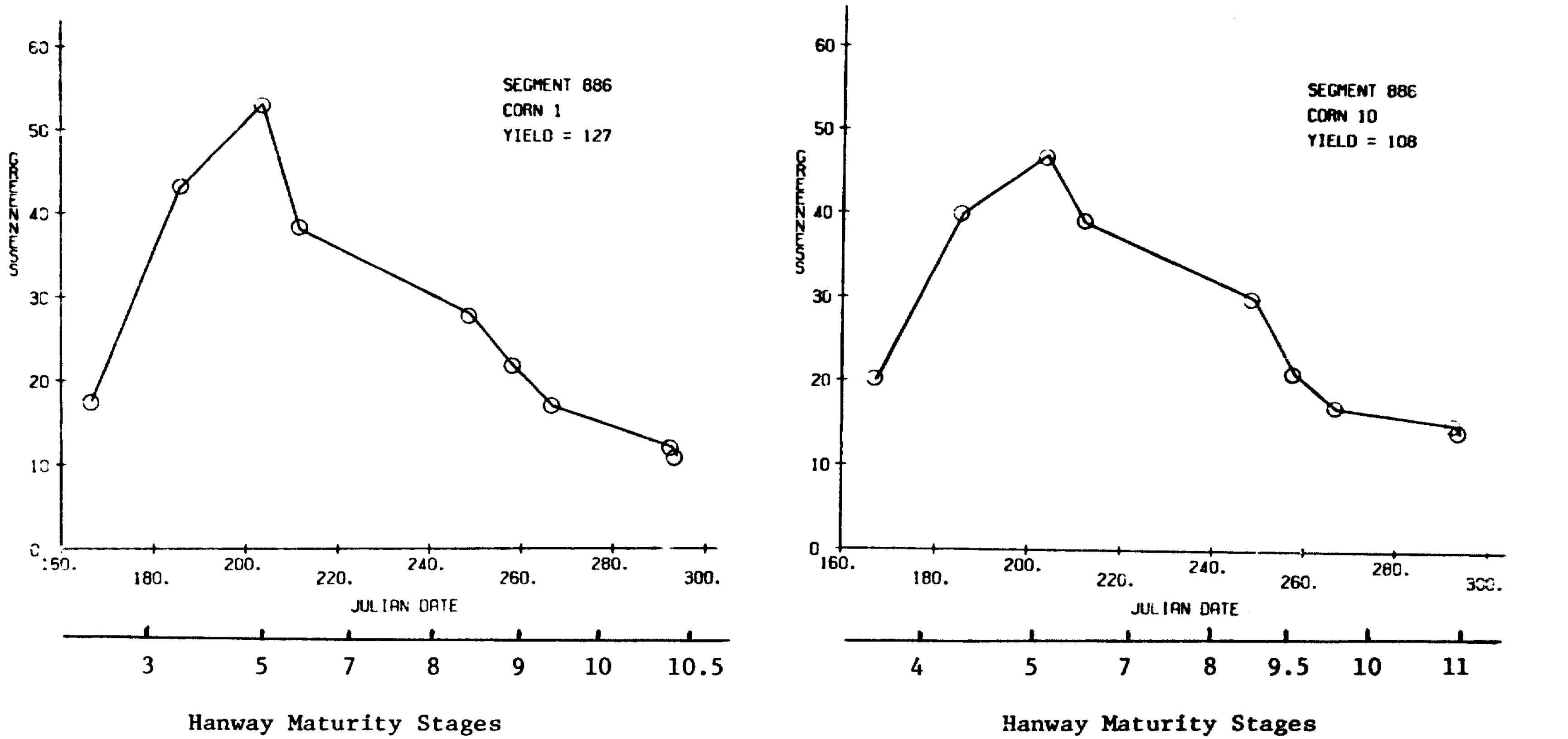


Figure B-8. Seasonal changes in Kauth Greenness transformation of Landsat MSS data acquired over two corn fields in Pottawatomie, Iowa in 1978. Yields are in bushels per acre. Maximum Greenness occurs near tasseling/silking (Maturity Stage 5).

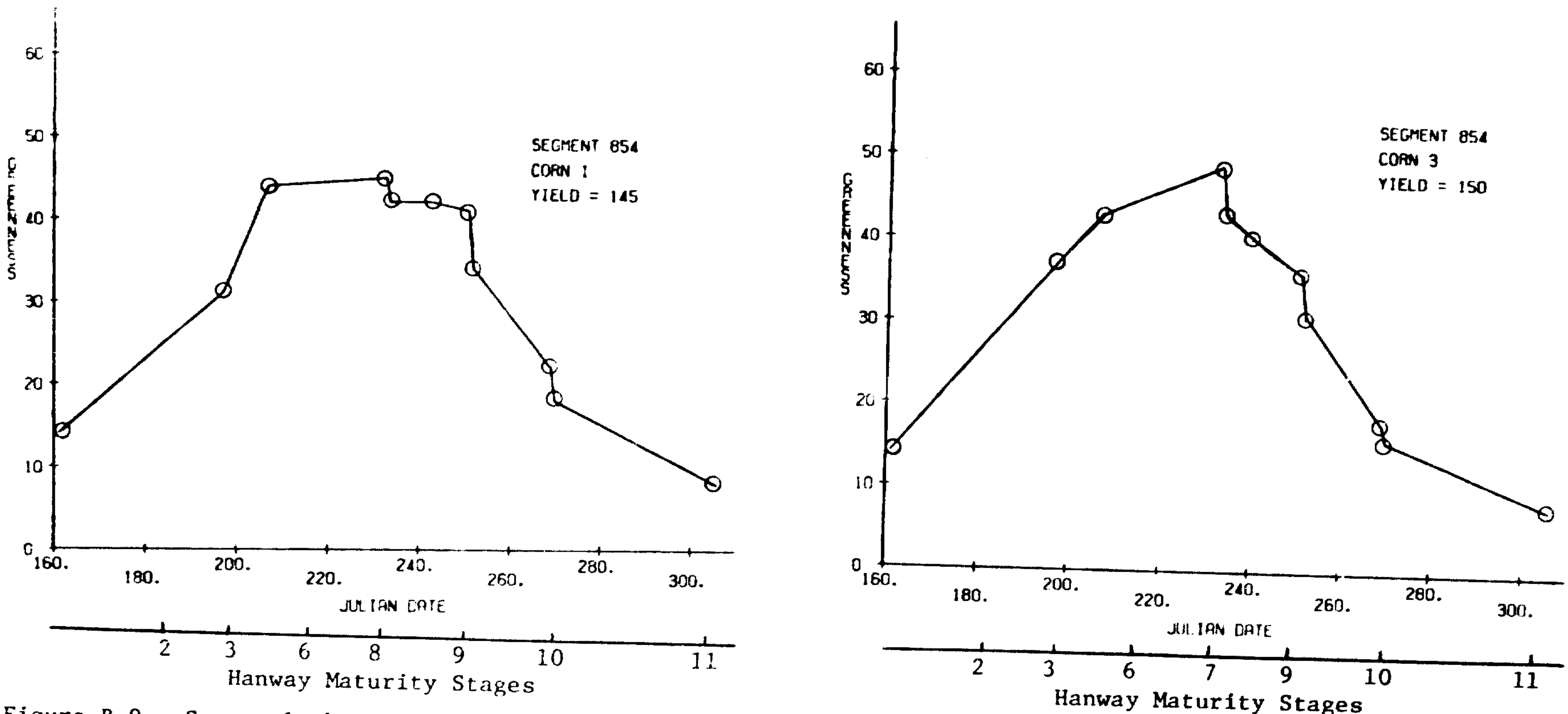


Figure B-9. Seasonal changes in Kauth Greenness transformation of Landsat MSS data acquired over two corn fields in Tippecanoe County, Indiana in 1978. Yields are in bushels per acre. Maximum Greenness occurs near tasseling/silking (Maturity Stage 5).

atmosphere on spectral response was not considered and may account for some of the abrupt changes in greenness over 9 to 18 day periods.

Correlations of Greenness and IR/Red ration with yields are greatest near tasseling (Table B-3). Preliminary indications are that simple correlations of Landsat MSS data and two transformations with departures from mean yield for each segment will not be sufficient to explain the variation in yields observed in individual fields (Table B-3). Additional research is in progress to examine these relationships fully. Alternative approaches which will use spectral data indirectly to estimate yields are also being pursued.

Table B-3. Correlations of corn grain yields of individual fields with Kauth Greenness and infrared (0.8-1.1 μm) to red (0.6-0.7 μm) ratio of Landsat MSS data at specific maturity stages in 1978.

Maturity[1/] Stage	Number of Fields	Yield		Residual Yield	
		Greenness[2/]	IR/Red	Greenness	IR/Red
<3	23	0.52	0.53	0.17	0.11
3-4	17	.27	.28	.29	.36
4-5	29	.34	.53	.00	-.05
5-6	26	.68	.85	.47	.54
6-7	56	.68	.66	.02	-.01
7-8	31	.44	.55	-.11	.01
8-9	15	.59	.60	.55	.38
9-11	111	.26	.19	.08	.11
>11	65	-.57	-.56	-.06	-.03

[1/] Hanway, J.J. (1966)

[2/] Greenness = 0.283*MSS4 - 0.660*MSS5 + 0.557*MSS6 + 0.388*MSS7 + 32

where: MSS4 = Landsat MSS radiance in 0.5-0.6 μm band
MSS5 = Landsat MSS radiance in 0.6-0.7 μm band
MSS6 = Landsat MSS radiance in 0.7-0.8 μm band
MSS7 = Landsat MSS radiance in 0.8-1.1 μm band

[3/] Residual Yield is the difference between individual field yields within a segment and the mean yield for that segment.

5. References

1. Bauer, M.E. 1975. The role of remote sensing in determining the distribution and yield of crops. Adv. Agron. 27:271-304.

2. Dale, R.F. and H.F. Hodges 1975. Weather and corn yield study for Tippecanoe County, Indiana. Final Report Environmental Data Service, NOAA Grant NG-44-72. Purdue Univ., W. Lafayette IN. 85p.

3. Decker, Wayne L., et al., 1976. Climate and Food: Climatic Fluctuation and U.S. Agricultural Production. Committee on Climate and Weather Fluctuations and Agricultural Production. National Academy of Sciences, Washington, D.C. 212p.

4. Eastin, J.D. (Chairman) 1969. Physiological Aspects of Crop Yield. American Society of Agronomy.

5. Hanway, J.J. 1966. How a corn plant develops. Iowa State Univ. Special Publication No. 48. Cooperative Extension Service, Ames, IA.

6. Hill, J.D., K.E. Henderson, and M.W. Seeley (eds) 1978. Proceeding of the 1978 LACIE Corn-Soybean Seminar. NASA/Johnson Space Center, Houston, TX. JSC-13744.

7. Holt, D.A., F.J. Bula, G.E. Miles, M.M. Schreiber, and R.M. Peart. 1975. Environmental physiology, modeling, and simulation of alfalfa growth. I. Conceptual development of SIMED. Purdue Agr. Exp. Sta. Bull. 907. 26p.

8. Holt, D.A., C.S.T. Daughtry, H.F. Reetz, and S.E. Hollinger. 1979. Separating soil and weather effects in large area yield prediction. Agron. Abstr. 71:12.

9. Huff, F.A. 1971. Sampling errors in measurement of mean precipitations. J. Appl. Meteor. 9:35-44.

10. Kramer, P.J. 1969. Plant and Soil Water Relationships: A Modern Synthesis. McGraw-Hill, New York. 481p.

11. Leeper, R.A., E.C.A. Runge, and W.M. Walker. 1974. Effect of plant-available stored soil moisture on corn yields, Agron. J. 66:723-733.

12. Montgomery, O.L. 1976. An investigation of the relationship between spectral reflectance and chemical, physical, and genetic characteristics of soils. Ph.D. Thesis. Purdue Univ.

13. Nelson, W.L. and R.F. Dale. 1978a. Effect of trend or technology variables and record period on prediction of corn yields with weather variables. J. Applied Meteorol, 17:926-933.

14. Nelson, W.L. and R.F. Dale 1978b. A methodology for testing the accuracy of yield predictions from weather-yield regression models for corn. Agron. J. 70:734-740.

15. Pierre, W.H., D.Kirkham, J. Pesek, R. Shaw (eds.) 1966. Plant Environment and Efficient Water Use. American Society of Agronomy. Madison, Wisconsin. 295p.

16. Reetz, H.F. Jr., 1976. CORN-CROPS: A physiology-based simulation of the corn crop. Ph.D. Thesis. Purdue University, West Lafayette, IN. 159p.

17. Seeley, M.W. 1977. The development and applications of biological time scale models for field corn hybrids in Nebraska. Ph.D. Thesis. Univ. of Nebraska, Lincoln, NE.

18. Shaw, R.H. 1963. Estimation of soil moisture under corn. Iowa Agric. Expt. Sta. Res. Bul. 520:968-980.

19. Storie, R.E. 1950. Rating soils for agricultural, forest and grazing use. Trans. IV Amsterdam Int. Congr. of Soil Sci. 1:336-339.

20. Strommen, N.D., C.M. Sakamoto, S.K. LeDuc and D.E. Umberger. 1979. Development of LACIE CCEA-I weather/wheat yield models. pp. 99-108. (In) A technical description of the Large Area Crop Inventory Experiment (LACIE). The LACIE Symposium. Oct. 1978. NASA/Johnson Space Center, Houston, TX JSC 16015.

21. Stuff, R.G., and R.F. Dale. 1978. A soil moisture budget model accounting for shallow water table influences. Soil Sci. Soc. Am. J. 42:637-643.

22. Thompson, L.M. 1969. Weather and technology in the production of corn in the U.S. Corn Belt. Agron. J. 61:453-456.

23. Walker, C. 1976. A model to estimate corn yields for Indiana. M.S. Thesis. Purdue Univ. West Lafayette, IN.

24. Walburg, G., C.S.T. Daughtry, M.E. Bauer, J.C. Kollenkark, and S.A. Barber. 1979. Maize canopy reflectance as influenced by nitrogen nutrition, Agron. Abstr. 71:17.