LARS Contract Report 113078

Annual Technical Summary Report

NASA Contract NAS9-15466
December 1, 1977 — November 30, 1978

D. A. Landgrebe, Purdue University
Principal Investigator

J. D. Erickson, NASA/JSC
Technical Monitor

Submitted by
Laboratory for Applications of Remote Sensing
Purdue University West Lafayette, Indiana
1978

STAR INFORMATION FORM

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No
113078		

4. Title and Subtitle	5. Report Date
Technical Summary	November 1978
	6. Performing Organization Code
7. Author(s) D. A. Landgrebe and Staff	8. Performing Organization Report No 113078
9. Performing Organization Name and Address Laboratory for Applications of Remote Sensing Purdue University West Lafayette, Indiana 47906	10. Work Unit No. 11. Contract or Grant No. NAS9-15466
12. Sponsoring Agency Name and Address J. D. Erickson/SF3 NASA/Johnson Space Center Houston, Texas 77058	13. Type of Report and Period Covered Technical Summary 12/1/77 – 11/30/78 14. Sponsoring Agency Code

15. Supplementary Notes

D. A. Landgrebe was LARS principal investigator.

16. Abstract

This Final Technical Summary summarizes work done and results obtained in a broad set of remote sensing research studies. These include studies of spectral characteristics of crops and soils; the status of a field research data bank and a software data handling and analysis system for it; specification of a standardized multispectral field data acquisition system; studies of machine implemented training sample labeling methods; scene stratification and area estimation procedures for future crop inventory systems; a software system for studying the optimality and effectiveness of various sets of multispectral scanner parameters; the construction of multitype data sets involving the combination of Landsat data with synthetic aperature radar data and map-derived data; an assessment of the methods used by various countries of North and South America and Asia for acquiring, analyzing and reporting crop production statistics; and the status of a multilocation, multiterminal computer processing system for supporting remote sensing research. More complete reports on the studies are contained in the five volume final report of the contract and in various technical reports, which are listed in the technical summary.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Spectral characteristics, training sample labeling, scene stratification, area estimation, optimal scanner parameters, multitype data, crop production statistics.	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	89	

*For sale by the National Technical Information Service, Springfield, Virginia 22161

NASA JSC

Preface

This Final Technical Summary Report summarizes the work done and results obtained under contract NAS9-15466 during the period from December 1, 1977 to November 30, 1978. The complete report of this work is contained in the contract Final Report published in five volumes as follows:

Volume		
1	Agricultural Scene Understanding	115 pgs.
2	Multispectral Scanner System Parameter Study and Analysis Software System Description	138 pgs.
3	Processing Techniques Development	176 pgs.
4	Assessment of Methods of Acquiring, Analyzing, and Reporting Crop Production Statistics	102 pgs.
5	Computer Processing Support	160 pgs.

Further details on the results in several cases are in Technical Reports which were published at appropriate times during the contract. These are listed in Table 1 immediately following the Table of Contents.

Copies of these reports may be obtained from the National Technical Information Service, U. S. Department of Commerce, Springfield, VA 22161. In some cases they may also be obtained from Purdue University via the Interlibrary Loan System or directly from Purdue/LARS.

Table of Contents

1. Agricultural Scene Understanding

1A. Analysis of Agronomic and Spectral Data for Physical Understanding

1B. Field Measurements Data Management

1C. Multicrop Field Research

1D. Determining the Climatic and Genetic Effects on the Relationships between Multispectral Reflectance and Physical-Chemical Properties of Soils

2. Processing Techniques Development

2A. Application of Statistical Pattern Recognition to Image Interpretation

2B. Application and Evaluation of Landsat Training, Classification, and Area Estimation Procedures for Crop Inventory

2C1. Multispectral Scanner System Parameter Analysis System

2C2. Multisensor Multidate Spatial Feature Matching, Correlation, Registering, Resampling and Information Extraction

3. Assessment of Methods of Acquiring, Analyzing and Reporting Crop Production Statistics

4. Computer Processing Support

TECHNICAL REPORTS AND PAPERS - CONTRACT NAS9-15466

061578 "A Parametric Multiclass Bayes Error Estimator for the Multispectral Scanner Spatial Model Performance Evaluation" by B. G. Mobasseri, C. D. McGillem, and P. E. Anuta.

061778 "Analytical Techniques for the Study of Some Parameters of Multispectral Scanner Systems for Remote Sensing" by E. R. Wiswell and G. R. Cooper.

TECHNICAL REPORTS AND PAPERS - PREVIOUS SR&T CONTRACTS

110477 "Analysis of the Effects of Interpolation and Enhancement of Landsat-1 Data on Classification Accuracy" by N. Chu, C. McGillem and P. Anuta. Contract No. NAS9-14970.

110577 "The LARSYS Educational Package: Instructors Notes for Use with the Data 100" by J. C. Lindenlaub and J. D. Russell. Contract No. NAS9-14970.

110677 "Demonstration of LARSYS on a Data 100 Terminal - Student's Notes" by John Lindenlaub and Staff. Contract No. NAS9-14970.

110777 "Data 100 Remote Terminal - A Hands-On Experience - Student's Notes" by J. D. Russell. Contract No. NAS9-14970.

083077 "ECHO User's Guide" by J. L. Kast, P. H. Swain, B. J. Davis, and P. W. Spencer. Contract No. NAS9-14970.

011678 "Optimum Filter for Minimization of Image Registration Error Variance" by Clare D. McGillem and Martin Svedlow. Contract No. NAS9-14016.

030178 "Bayesian Classification in a Time-varying Environment" by Philip H. Swain. Contract No. NAS9-14970.

051578 "The Development of a Spectral/Spatial Classifier for Earth Observational Data" by D. A. Landgrebe. Grant NGL 15-005-112, Contract No. NAS9-14016, and Contract No. NAS9-14970.

OTHER REPORTS

Marvin E. Bauer, Michael C. McEwen, William A. Malila, and J. C. Harlan, 1978. "Design, Implementation, and Results of LACIE Field Research". In Final Report, Large Area Crop Inventory Experiment (also in Proceedings of LACIE Symposium). NASA/Johnson Space Center. 85 pages. Contract No. NAS9-15466.

Marilyn M. Hixson, Larry L. Biehl, and Marvin E. Bauer. "Crop Spectra from LACIE Field Measurements" LARS Contract Report 011578 and LACIE Report 00469. March 1978. 157 pages. NAS9-14970 and NAS9-15466.

Larry L. Biehl. "LACIE Field Measurements, Data Library Catalog, Volume III, 1976-77 Crop Year" October 1978. Published by NASA/JSC. Contract No. NAS9-15466.

OTHER REPORTS (Con't)

Marvin E. Bauer. "Spectral Characteristics and Remote Sensing of Corn and Soybeans." Presentation at Corn and Soybean Seminar, NASA/Johnson Space Center, Houston, Texas, May 1-3, 1978. (Written paper also prepared and published by NASA/JSC). Contract No. NAS9-15466.

Task 1. Agricultural Scene Understanding

Results of four investigations, all related to increasing our understanding cf agricultural crops and soils and their spectral responses are described. The Analysis of Agronomic-Spectral Data section describes the results of analyses of LACIE Field Research Data, including the relationships of agronomic and reflectance characteristics of wheat canopies, effects of cultural and environmental factors on reflectance properties of wheat, and discrimination of wheat and other crops as a function of wavelength band selection and acquisition date. The Field Measurements Data Management section describes the field research data base developed at LARS, as well as the development of graphical and statistical analysis software data processing software, and distribution of data. The Multicrop Supporting Field Research section describes the measurements made in 1978 of the spectral characteristics of corn and soybeans and the development of a multispectral data acquisition system for field research. The last section describes the objectives, experimental approach, and initial results of a study of the relationships between the reflectance and physical-chemical properties of over 250 different soils.

1A. Analysis of Agronomic and Spectral Data for Physical Understanding

Crop identification and area estimation promises to be one of the major applications of remote sensing and the Large Area Crop Inventory Experiment (LACIE) has pushed the technology to near operational use for wheat. Remote sensing also offers great potential for obtaining accurate and timely information about the condition and yield of crops.

To fully realize the potential of remote sensing for crop identification, condition assessment, and yield prediction it is important to understand and quantify the relation of agronomic characteristics of crops to their multispectral reflectance properties. For example, it is essential to know in which regions of the spectrum information relating to variations in crop parameters is contained. This information is necessary for the optimum use of current Landsat technology, as well as for the design and development of future remote sensing systems.

Differences among crop species and dynamic changes due to growth, development, stress, and varying cultural practices cause differences in the reflectance spectra of crops. Many of the factors affecting the reflectance properties of plant leaves have been identified and investigated utilizing laboratory measurements. The relationships of physical-biological parameters such as chlorophyll concentration, water content and leaf morphology to reflectance, transmittance and absorption have been well-established for leaves.

Knowledge of the reflectance characteristics of single leaves is basic to understanding the reflectance properties of crop canopies in the field, but cannot be applied directly since there are significant differences in the spectra of single leaves and canopies. The reflectance characteristics of canopies are considerably more complex than those of single leaves because there are many more interacting variables in canopies. Some of the more important agronomic parameters influencing the reflectance of field-grown canopies are: leaf area index, biomass, leaf angle, soil cover percentage, soil color, and leaf color. Differences in these parameters are caused by variations in many cultural and environmental factors, including planting date, cultivar, seeding rate, fertilization, soil moisture, and temperature. Solar elevation and azimuth angle and the view angle and direction of the sensor also affect the measured reflectance of crops and soils.

The spectral and agronomic measurements which have been acquired during the three years of the LACIE field research program are being analyzed to provide an understanding of the relationship of reflectance to the biological and physical characteristics of crops and soils. The primary data being analyzed are the spectrometer data acquired by the truck- and helicopter-borne systems. These data are particularly useful because the spectral data were acquired in 0.01 μm wavelength intervals and are calibrated in terms of bidirection reflectance factor. Having the entire spectrum from 0.4 to 2.4 μm permits simulation of the response in any specified waveband. In other words, the analysis is not restricted to a fixed set of bands such as Landsat MSS or one of the aircraft scanner systems. Calibration of the data permits valid comparisons to be made among different dates, locations, and sensors.

The overall objective of the analyses conducted this year has been to quantitatively determine the spectral-temporal characteristics of wheat, small grains, and other agricultural crops. The specific objectives of the analyses reported here were:

1. To determine the relationship of agronomic variables such as biomass and leaf area index to the multispectral reflectance of spring wheat and evaluate the potential for predicting the agronomic characteristics of wheat canopies from reflectance measurements.

2. To determine the effects of cultural and environmental factors on the spectral response of spring wheat.

3. To determine the discriminability of wheat and other crops as a function of acquisition date and spectral band selection.

1. Prediction of Agronomic Characteristics of Spring Wheat Canopies

Spectral and agronomic measurements made in 1976 at the Williston, North Dakota Agriculture Experiment Station of spring wheat canopies were analyzed to determine the relation of agronomic properties of crop canopies to their spectral reflectance. One of the primary factors found to affect the reflectance of canopies was the amount of vegetation present (Figure 1A-1). Strong relationships between spectral response and percent soil cover, leaf area index, biomass and plant water content were found (Figure 1A-2). The relationship, however, is influenced by crop maturity. The best time period for assessing these canopy variables is from the tillering to heading stages of development. Prior to tillering the spectral response is strongly dominated by the soil background and, as the crop begins to ripen, the spectral sensitivity to measures such as leaf area index, biomass, and plant water content decreases.

In each wavelength region, the correlation of the thematic mapper band with crop canopy variables was greater than that of the corresponding Landsat MSS band (Table 1A-1). Regression equations developed to explain the variation in crop canopy variables showed that the 2.08-2.35 μm wavelength band was the single most important band in explaining the variation in fresh biomass, dry biomass, and plant water content; whereas, the near infrared band (0.76-0.90 μm) explained the most variation in leaf area index and percent soil cover. The results demonstrate the importance of collecting spectral information in the middle infrared wavelength region, as well as the visible and near infrared, for crop assessments.

The R^2 values for comparisons of measured and predicted canopy variables ranged from 0.80 to 0.91 when three or more spectral bands were included, indicating the potential for using remotely sensed spectral measurements to characterize the status of crops (Figure 1A-3). Analyses showed that the best four thematic mapper bands could estimate crop canopy variables more accurately than the four Landsat bands (Table 1A-2). The difference is attributed to the narrower and more optimum placement of the thematic mapper bands in relation to the spectral characteristics of vegetation.

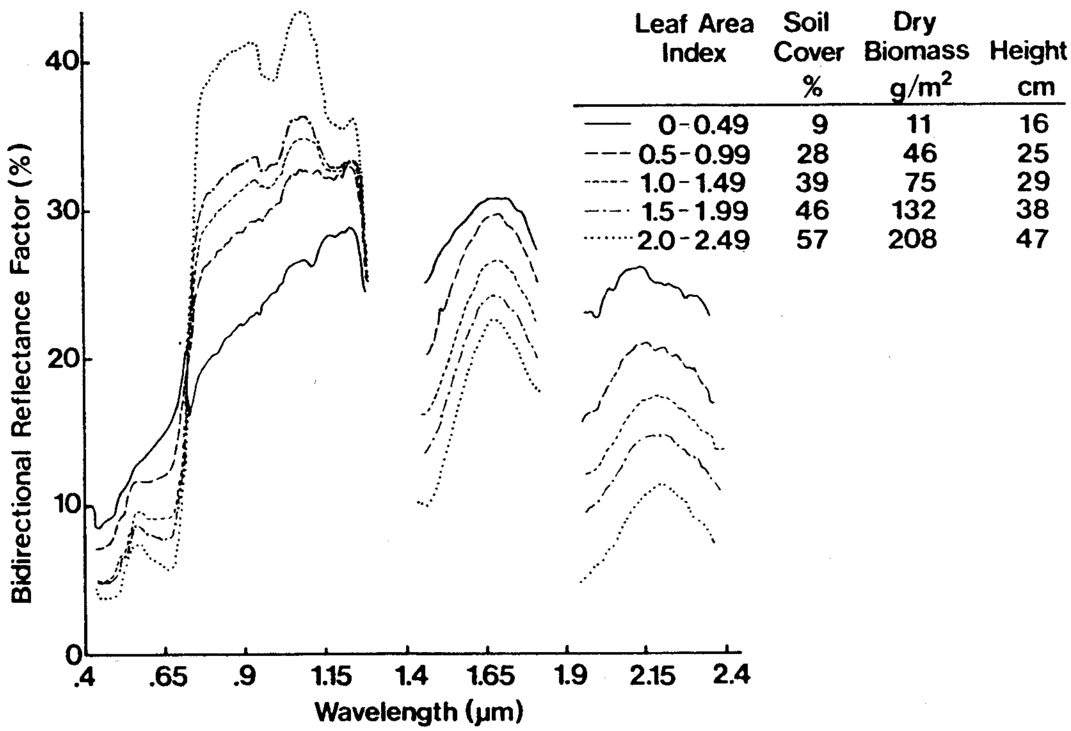


Figure 1A-1. Effect of leaf area index, percent soil cover, dry biomass, and plant height on the spectral reflectance of spring wheat during the period between tillering and the beginning of heading, when the maximum green leaf area is reached. Data were acquired at Williston, North Dakota on May 28 - June 18, 1976 and include plots with different soil moisture levels, planting dates, nitrogen fertilization and cultivars.

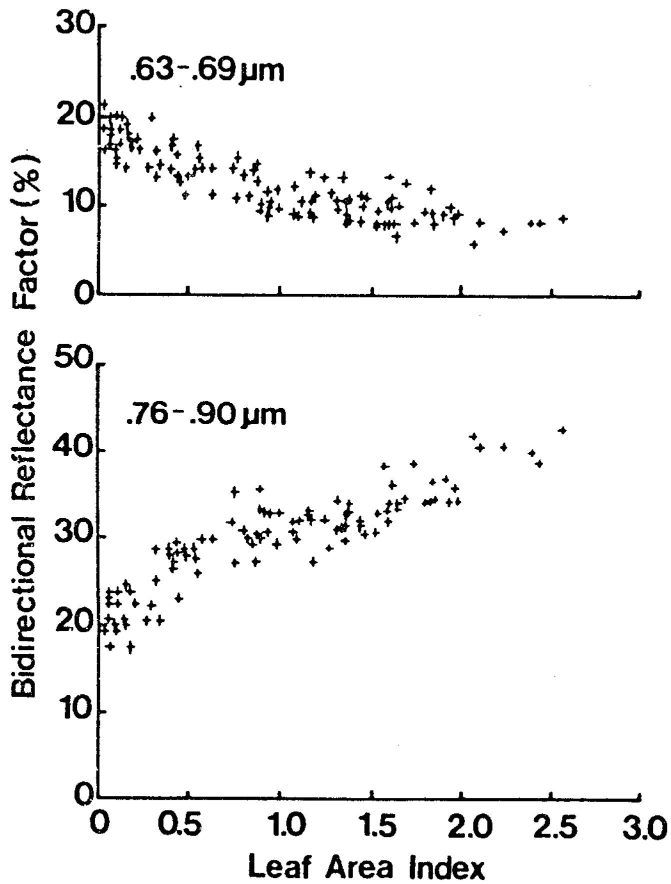


Figure 1A-2 Relationship of leaf area index to reflectances in the chlorophyll absorption (0.63-0.69 µm) and near infrared (0.76-0.90 µm) regions. Measurements for seedling through flowering stages of maturity are included for 16 treatments representing different levels of soil moisture availability, planting dates, nitrogen fertilization, and cultivar.

Table 1A-1. The linear correlations of reflectances in the thematic mapper and Landsat MSS wavelength bands with percent soil cover, leaf area index, fresh and dry biomass, and plant water content.

Wavelength Band (μm)	Percent Soil Cover	Leaf Area Index	Fresh Biomass	Dry Biomass	Plant Water Content
Thematic Mapper					
0.45-0.52	-0.82	-0.79	-0.75	-0.69	-0.76
0.52-0.60	-0.82	-0.78	-0.81	-0.77	-0.82
0.63-0.69	-0.91	-0.86	-0.80	-0.73	-0.81
0.76-0.90	0.93	0.92	0.76	0.67	0.79
1.55-1.75	-0.85	-0.80	-0.83	-0.79	-0.84
2.08-2.35	-0.91	-0.85	-0.86	-0.81	-0.86
Landsat MSS					
0.5-0.6	-0.82	-0.79	-0.81	-0.76	-0.81
0.6-0.7	-0.90	-0.85	-0.81	-0.74	-0.82
0.7-0.8	0.84	0.84	0.57	0.46	0.60
0.8-1.1	0.91	0.90	0.77	0.68	0.79

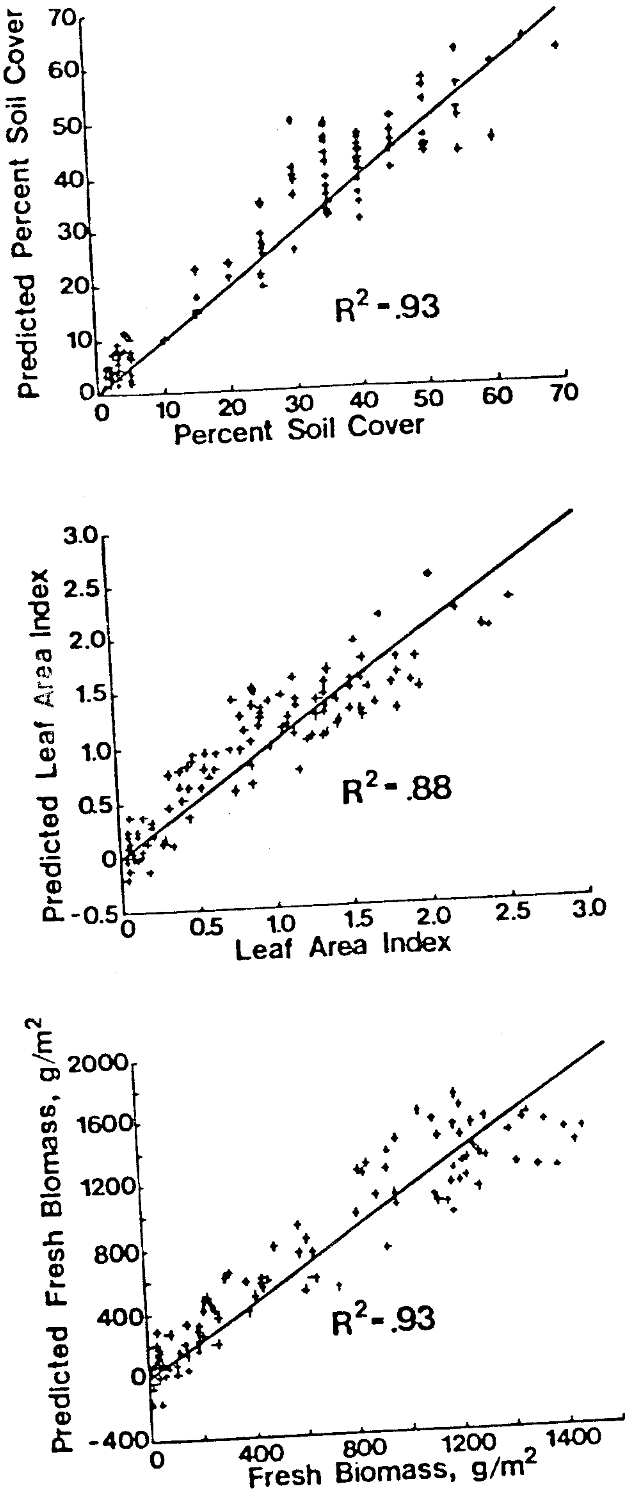


Figure 1A-3. Comparison of measured and predicted percent soil cover, leaf area index, and fresh biomass of spring wheat canopies.

Table 1A-2. The R^2 values for predictions of percent soil cover, leaf area index, fresh and dry biomass, and plant water content with four Landsat MSS bands, the best four thematic mapper bands, and the six thematic mapper bands.

Wavelength Bands	Percent Soil Cover	Leaf Area Index	Fresh Biomass	Dry Biomass	Plant Water Content
Landsat MSS Bands	0.91	0.86	0.86	0.84	0.85
Best Four Thematic Mapper Bands	0.93	0.88	0.88	0.84	0.88
Six Thematic Mapper Bands	0.93	0.88	0.91	0.88	0.90

The strong relationship between spectral reflectance and different crop canopy variables illustrates the potential to monitor crop growth and predict yield. Future research needs to investigate the amount of variation induced by different agronomic treatment factors on spectral reflectance and whether important treatment factors are spectrally separable. The type of prediction equations developed in this investigation need to be extended to several years of data, then used to estimate independent data sets.

2. Effects of Cultural and Environmental Factors on the Reflectance Characteristics of Spring Wheat Canopies

The crop canopy is a dynamic entity influenced by many cultural and environmental factors. Therefore, it is important to quantify and understand the sources of variation in spectral measurements of crops. Some of the variation may be associated with important agronomic factors which it may be desirable to inventory or monitor (for example, dryland vs. irrigated wheat). On the other hand it is also important to know the magnitude of variation associated with a factor such as cultivar which we would most likely not want to identify or monitor.

Examples of spectra acquired in 1976 at the Williston, North Dakota, Agriculture Experiment Station are shown in Figure 1A-4 to illustrate some of the effects of agronomic treatments on the spectral response of spring wheat. Further experiments using an improved experimental design were conducted in 1977 to determine the effects of several agronomic treatments (soil moisture availability, planting date, nitrogen fertilization, and cultivar) on the reflectance of spring wheat canopies. The results of the 1977 investigation are summarized in this section.

Early in the growing season planting date is the primary agronomic factor influencing the reflectance of spring wheat, accounting for about 50 percent of the variation in reflectances. The spectral differences were primarily due to differences in the amount of vegetation present. Later in the season at the heading to ripening stages, the level of soil moisture becomes the most important factor. Wheat grown on land with higher levels of available soil moisture had a greater percent soil cover,

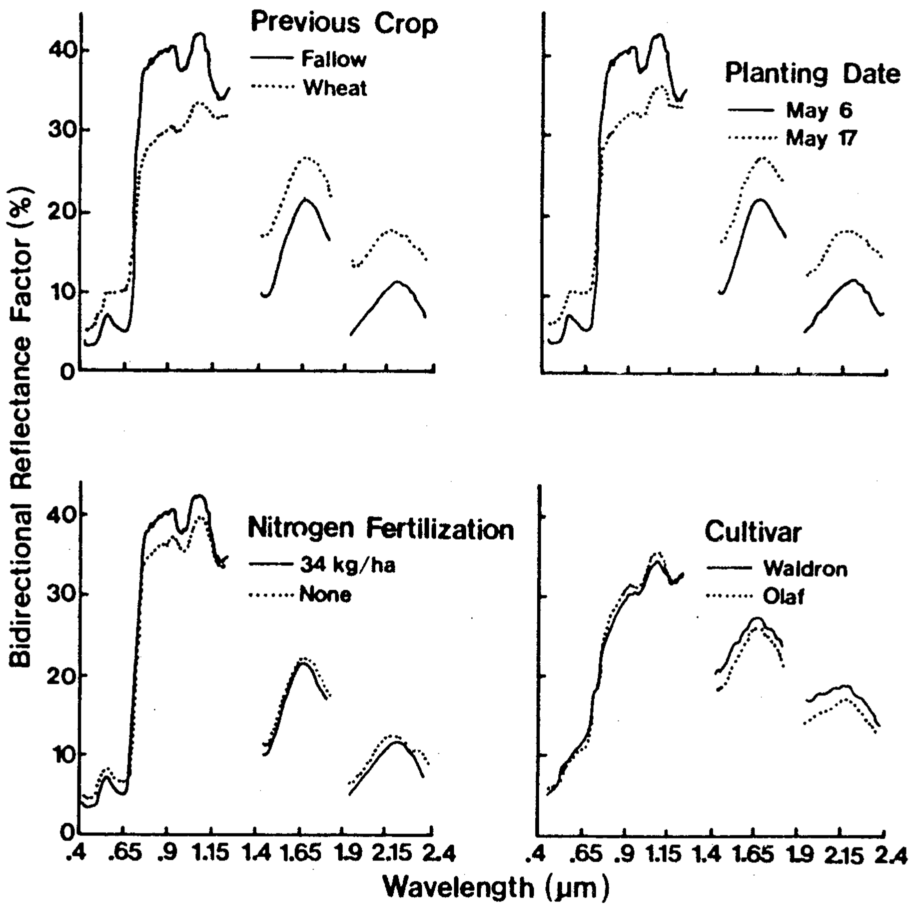


Figure 1A-4. Effects of agronomic treatments on the spectral reflectance of spring wheat. Spectra were measured on June 18, 1976, during the stem extension stage of development, except for the spectra of cultivars which were measured on July 16 after heading.

leaf area index, and biomass causing increased near infrared and reduced visible reflectance. In these experiments cultivar and nitrogen fertilization had relatively little effect on the spectral response of spring wheat. Again, examination of the agronomic data shows that these two treatments had little influence on the growth and development of the spring wheat canopies. The primary difference in the two cultivars was in plant height, rather than in leaf area or biomass. And, since the soils of this area of North Dakota are relatively high in nitrogen supplying capacity, the addition of nitrogen fertilizer had only minor effects of the growth of wheat.

3. Assessment of Crop Discriminability as a Function of Crop Maturity Stage and Wavelength Band Selection

The application of remote sensing to crop inventories relies upon the ability to detect and identify the crops of interest from the spectral data. Unless correct crop identification can be consistently made, the goal of accurately predicting crop acreage and production cannot be achieved.

Discriminant analyses of LACIE Field Measurements data acquired by the helicopter spectrometer over the intensive test sites at Finney County, Kansas in 1975 and Williams County, North Dakota in 1976 were performed. The overall objective of these analyses was to assess the potential for crop discrimination. The specific objectives were: (1) to determine which cover types may be confused, (2) to determine at what growth stages separability is maximized, and (3) to determine which regions of the spectrum are best for crop discriminability.

North Dakota spring wheat classification accuracy was 90-95 percent during the middle part of the growing season and 80-85 percent early and late in the season. Spring wheat was confused with pasture by the Landsat MSS bands when only single date information was used. If multitemporal information was used, however, the two cover types were separable and classification accuracy increased. The training method was found to have a greater effect on classification results than amount of training.

In Kansas, classification accuracies of 80-90 percent were obtained. Winter wheat was occasionally confused with alfalfa and was confused with fallow land early in the season before sufficient soil cover had been reached for the wheat to appear characteristically green. Corn and sorghum could not be identified until much later in the season due to their planting date.

In both the North Dakota and Kansas analyses classification performance for the thematic mapper bands was higher than for the Landsat MSS bands (Figure 1A-5). The improved classification accuracy of the thematic mapper bands is attributed to the narrower and more optimum placement, particularly inclusion of middle infrared wavelengths, of the thematic mapper bands with respect to the spectral characteristics of vegetation.

4. Summary and Conclusions

Analyses of the LACIE Field Research data are providing new knowledge and understanding of the spectral characteristics of wheat and the biological-physical factors affecting spectral response. For example, strong relationships have been found between reflectance and percent soil cover, leaf area index, biomass, and plant water content of spring wheat canopies. These are fundamental measures of crop vigor which can be used in crop growth and yield prediction models. In relating agronomic and spectral characteristics of wheat it was found that a middle infrared wavelength band, 2.08-2.35 μm, is most important in explaining variation in biomass and plant water content, while a near infrared band, 0.76-0.90 μm, accounts for the most variation in percent soil cover and leaf area index.

Analyses of the effects of several agronomic practices on the reflectance characteristics of spring wheat showed that early in the growing season planting date is the primary factor causing variation in the spectral response. Later in the season during heading to ripening stages, the level of available soil moisture at the beginning of the season explained the most variation in spectral response. In evaluating sensor characteristics it was found that the reflective wavelength bands proposed for the thematic mapper are more strongly related to and better predictors of

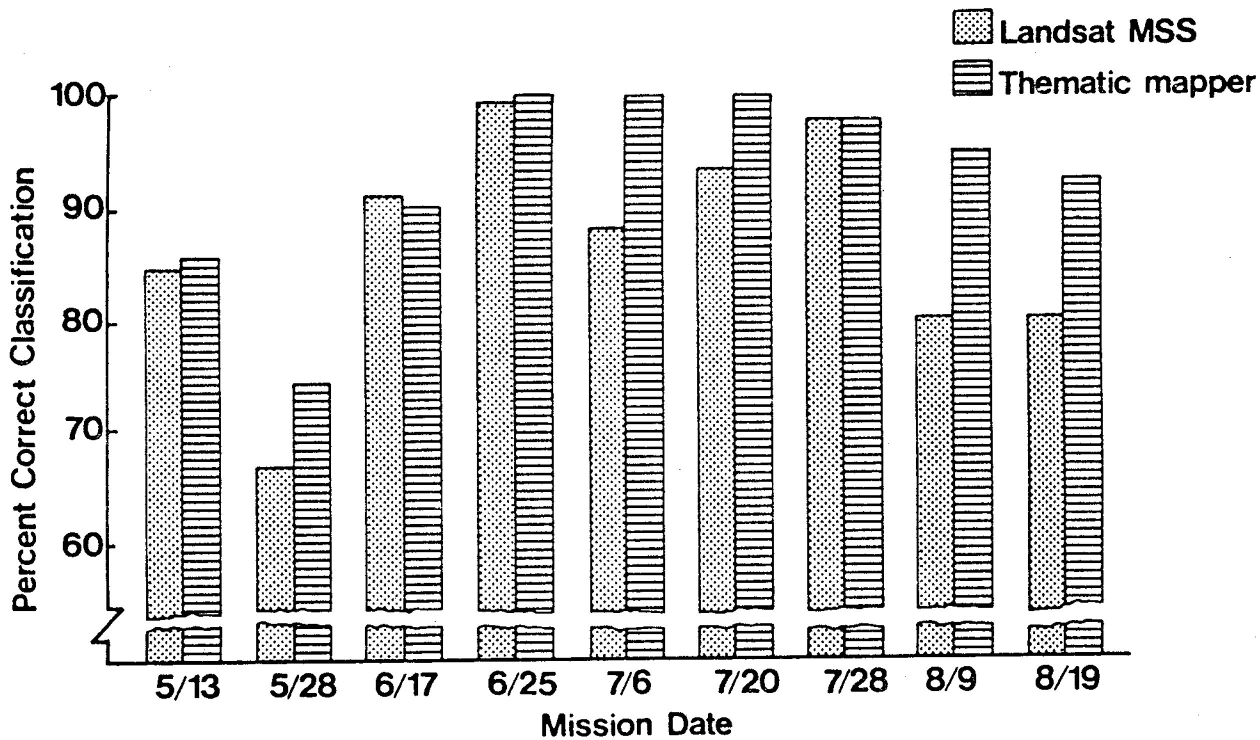


Figure 1A-5. Comparison of Landsat MSS and thematic mapper spectral bands for spring wheat identification (Williams County, North Dakota, 1976).

the canopy variables than the Landsat MSS bands. Thematic mapper bands also are better for crop identification. The improved performance of the thematic mapper bands is attributed to their narrower width, and location with regard to the spectral characteristics and vegetation and the addition of bands from the middle infrared region.

Task 1B. Field Measurements Data Management

The development of the field research data library at Purdue/LARS was initiated in the fall of 1974 by NASA/Johnson Space Center with the cooperation of the United States Department of Agriculture as a part of the Large Area Crop Inventory Experiment (LACIE). The purpose for developing the data base is to provide fully annotated and calibrated multitemporal sets of spectral, agronomic, and meteorological data for agricultural remote sensing researchers. Spectral, agronomic, and meteorological measurements over primarily wheat have been made on three LACIE test sites in Kansas, North Dakota, and South Dakota for three years. During this past year the data library was expanded to include data collected for corn and soybean experiments in Indiana.

Milestones reached during this year have been: (1) completion of the 1976-77 data processing, (2) development of graphical and statistical analysis software, (3) development of documented data processing software, and (4) distribution of data to researchers representing university, government and commercial research organizations.

A report entitled "Crop Spectra from LACIE Field Measurements", LARS Technical Report 011578, was prepared and distributed this past year. This document contains examples of the spectrometer data in the form of spectral reflectance curves illustrating the major sources of spectral variation of wheat and related crops. The examples include variations in wheat spectra due to differences in maturity, biomass, soil color, and soil moisture, as well as comparisons of spectra of wheat and other crops.

1. Field Research Data Library and Distribution

The general organization of the field research data library is illustrated in Figure 1B-1. The data in the library include spectral measurements (multispectral scanner and spectrometer), agronomic measurements, meteorological measurements, photography, mission logs and data verification reports. The data formats available to researchers are digital tape, film, and data listings.

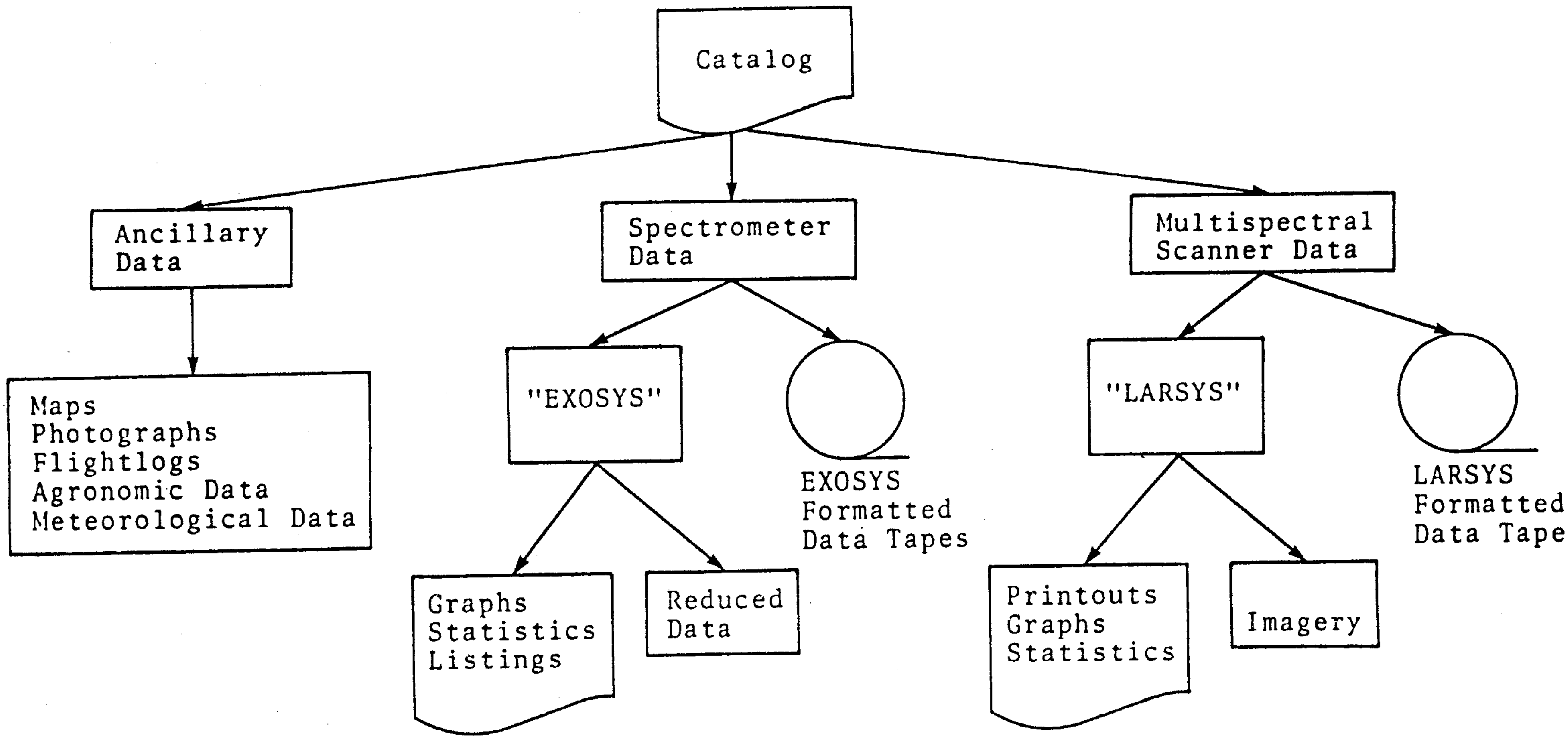


Figure 1B-1. Organization of Field Research Data Library. EXOSYS and LARSYS are Purdue/LARS software systems to analyze spectrometer and multispectral scanner data.

The spectrometer data are processed into a common format (spectral bidirectional reflectance factor) in order to make meaningful comparisons of the data acquired by the different sensors at different times and locations. The spectrometer data tapes contain the bidirectional reflectance factor measurements along with the corresponding agronomic and meteorological measurements. The data are stored on the tape in LARS spectrometer data storage tape format.

The multispectral scanner data stored on the LARSYS Version 3 formatted tapes are approximately linearly related to scene radiance. The information is available for the researcher to calibrate the scanner data to in-band bidirectional reflectance factor if he wants to. Multispectral scanner data are also available in universal format.

In the past twelve months, 45 aircraft scanner runs and 30,000 spectrometer runs have been processed and made available in the field research data library. The status of data processing is summarized in Table 1B-1. Processing of the remaining 1976-77 crop year data along with data from a sun angle/view angle/view direction experiment were completed earlier this year. Processing of the 1977-78 data was begun this fall. The Field Research Data Library Catalogs summarizing all available data have been prepared for each crop year during which data were collected. Discussions with the past and present users of the catalog and data have provided direction for future updates of the catalog.

Six institutions, representing university, government, and commercial research organizations, requested and were provided LACIE field research data this year. In addition all the data are routinely available to researchers at Purdue/LARS and investigators at NASA/Johnson Space Center have direct access to the digital data via the remote terminal to the LARS computer located there. Table 1B-2 summarizes the data distribution during the past year.

Table 1B-1. Summary of LACIE Field Measurements data in the Field Research Library by year, instrument, and data type.

Instrument/Data Type	Crop Year 1975	1976	1977
Landsat MSS			
Whole Frame CCT (Frames)	20	62	34
Aircraft Scanner (Dates/Runs)	19/149	16/97	7/35
Helicopter Mounted Field Spectrometer (Dates/Runs)			
Field Averages	19/2,343	27/2,193	21/1873
Individual Scans	19/29,579	27/38,476	21/34,060
Truck Mounted Field Spectrometer (Dates/Runs)			
FSAS	6/65	23/322	16/426
Exotech 20C	24/1599	18/2542	21/1075
Exotech 20D	45/645	- - - -	- - - -

Table 1B-2. Recipients of Field Research data during 1977-78.

Institution	Data Requested and Received
Enviromental Research Institute of Michigan (ERIM)	Spectrometer data tapes, ground photos, flight logs, field maps.
Texas A&M University, Remote Sensing Center	Spectrometer data tapes, ground photos flight logs, field maps
NASA, Goddard Space Flight Center	Spectrometer data tapes, Landsat MSS data tapes, aerial photography, ground photos, field maps.
General Electric Corporation (contract with NASA/GSFC)	Spectrometer data tapes, agronomic and meteorological measurements, field maps, ground photos, aerial photography.
General Electric Corporation	Landsat MSS data tapes, agronomic measurements, Landsat color composite imagery.
USDA, Agriculture Research Service, Weslaco, Texas	Helicopter (FSS) photography, agronomic measurements, meteorological measurements, field maps.
University of South Florida, Department of Mathematics	Spectrometer data tapes.

2. Development of Graphics, Statistics and Data Processing Software

During the LACIE Field Measurements project over 100,000 observations of spectrometer data were collected. Additional observations were collected this year in preparation for Multicrop studies. These increasingly large numbers of field measurements type observations have brought about a need for more powerful research tools to verify and analyze the data efficiently and effectively. During this past year, several new analysis software tools were developed and implemented. These include new graphics capabilities and additional statistical capabilities (such as clustering) for spectrometer data. Also, an updated version of SPSS (Statistical Package for the Social Sciences) having statistical analysis capabilities was obtained and made available to researchers on the Purdue/LARS computer.

The software system implemented on the Purdue/LARS computer to access the spectrometer data and associated agronomic and meteorological measurements stored on tape is called EXOSYS. The version which contains the new graphics and statistics capabilities is termed EXOSYSDV. The new analysis tools for researchers include capabilities to:

- Review spectrometer data quickly on a CRT terminal
- Plot data on high resolution Varian printer/plotter in addition to line printer
- Plot, print or punch numerical agronomic and meteorological data stored on tape
- Fit polynomial curves of one to ten degrees to data and print regression coefficients and correlation coefficients
- Apply arithmetic transformations to data
- Cluster spectrometer data with one to 30 user specified bands and print tables with summaries of results

For the graphical software routines, EXOSYSDV makes use of the Graphics Compatibility System (GCS). The package of GCS software routines is very extensive for graphical purposes; the use of the GCS software routines to do the graphical plotting instead of writing completely new software reduced the implementation time for the above mentioned tools from years to months.

Examples of some of the graphical output are illustrated in Figures 1B-2 and 1B-3 Other graphical options available to the researcher using EXOSYSDV include polar coordinate system graphs, natural logarithmic and common logarithmic (base ten) formatted graphs, spline curve fitting, and other scaling and axes variations.

The clustering capability implemented in EXOSYSDV is the same cluster algorithm used in LARSYS (unsupervsied with Euclidean distance measure), except that the initial cluster centers are placed along the principal eigenvector. Also, the spectral information and some of the agronomic measurements stored on the spectrometer tapes can be punched and used in other statistical analysis packages such as SPSS or in user written Fortran programs.

Documentation of the new capabilities in EXOSYSDV is the next step to be developed to make the system more useful to researchers both at Purdue/LARS and to researchers at LARS remote computer terminal sites. Initial development of documentation was completed this past year. Development of new analysis tools and documentation will continue so that researchers are able to make the fullest use of the data in the field research data library.

In addition to the graphics and statistics software additions, the software for processing spectrometer data has been improved. The software for reformatting the Exotech 20C data was revised to make it more efficient, saving both computer and personnel time. The original reformatting system was designed to handle a few hundred measurements per year, compared to 2,000 to 3,000 observations currently collected each year. At the same time additional items of agronomic and measurement data were added to the header record. Capabilities for updating the spectrometer data tapes have been added and documentation describing the reformatting systems written.

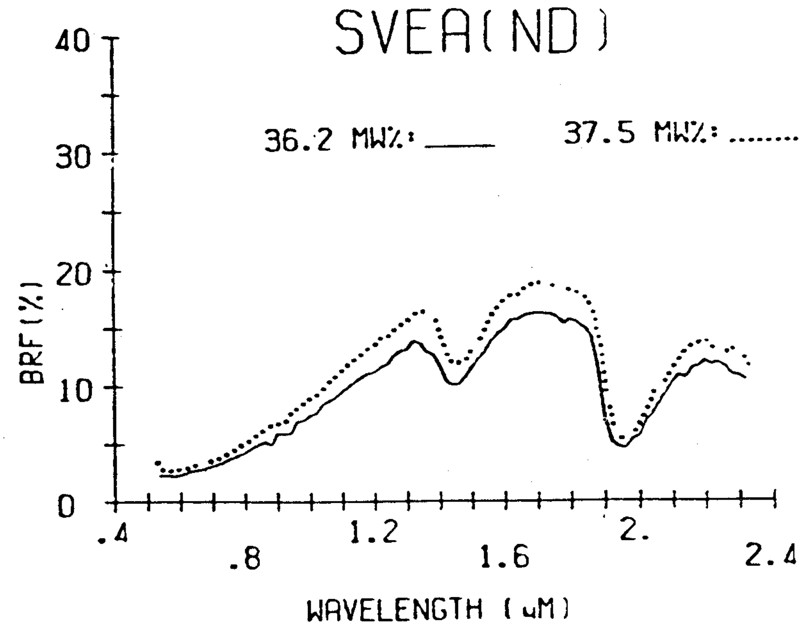


Figure 1B-2. Example graph from EXOSYSDV of spectral response as a function of wavelength.

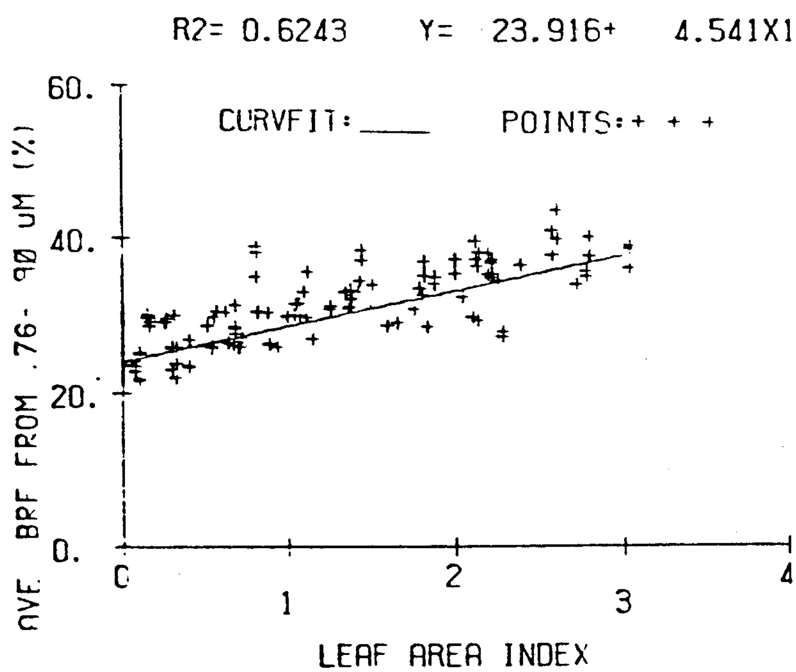


Figure 1B-3. **Example of polynomial curve fitting using EXOSYDV software.**

Task 1C. Multicrop Field Research

Although the LACIE Field Measurements project acquired a large quantity of spectral and agronomic data and analyses of it are providing quantitative information about the spectral characteristics of crops, the data are limited to crops grown in the Great Plains region of the United States. As we look ahead to the development of a global food and fiber information system utilizing remote sensing techniques, it is important to begin to conduct the field research required to more fully understand the spectral characteristics of crops other than wheat such as corn, soybeans, rice, cotton, and rangeland.

As part of the LACIE Transition and Multicrop program initiated in 1978, field research for corn and soybeans was initiated. The overall objectives of the multicrop supporting field research are:

- To increase the understanding of the basic radiation patterns of agricultural crops and their soil backgrounds under normal and stressed conditions.
- To assess the capability of current and planned satellite sensor systems to capture available useful spectral information.
- To determine the potential incremental improvement in performance of improved future systems.

The multicrop field research at Purdue/LARS during this contract year consisted of two components. The first was the acquisition of data describing the spectral properties of corn and soybeans under normal and stressed conditions. The second component of the project was development of specifications for a new approach to spectral measurements for field research.

1. Measurements of Spectral Characteristics of Corn and Soybeans

The potential of remote sensing to provide information on the condition and yields of crops has not been adequately explored and developed, especially for corn and soybeans. This task represents the initial phase of a multiyear research effort to evaluate and understand information about corn and soybean crops contained in various types of remotely sensed data. The results of these carefully controlled field experiments can provide the foundation on which larger experiments involving Landsat data should be built.

The overall objectives of this research task which include data acquisition and analysis are:

1. To determine the reflectance characteristics of corn and soybeans as a function of maturity stage and amount of vegetation present.
2. To examine the effects of moisture and nutrient stresses on the reflectance and radiant temperatures of corn and soybeans.
3. To determine the effect of important cultural/management practices on the reflectance of soybeans.
4. To develop and assess various methods of incorporating spectral, meteorological, and ancillary data into models of crop condition assessment and yield prediction.

Activities this year toward fulfillment of these objectives include collection of spectral, agronomic, and meteorological data at the Purdue University Agronomy Farm for five experiments with two different spectrometer systems. The experiments were: (1) corn moisture stress, (2) nitrogen fertilization of corn, (3) and (4) phosphorous and potassium fertilization of corn and soybeans, and (5) cultural practices of soybeans. The experiments provide both normal and stressed conditions of corn and soybeans. Spectral and agronomic data were collected at approximately weekly intervals throughout the growing season.

Spectral measurements were made with the Exotech 20C high resolution spectrometer system and the Exotech 100 Landsat-band radiometer system (Figure 1C-1). The Exotech 20C field system collected data over the 0.4 to 2.4 μm range. The Exotech 100 field system which is a prototype of a

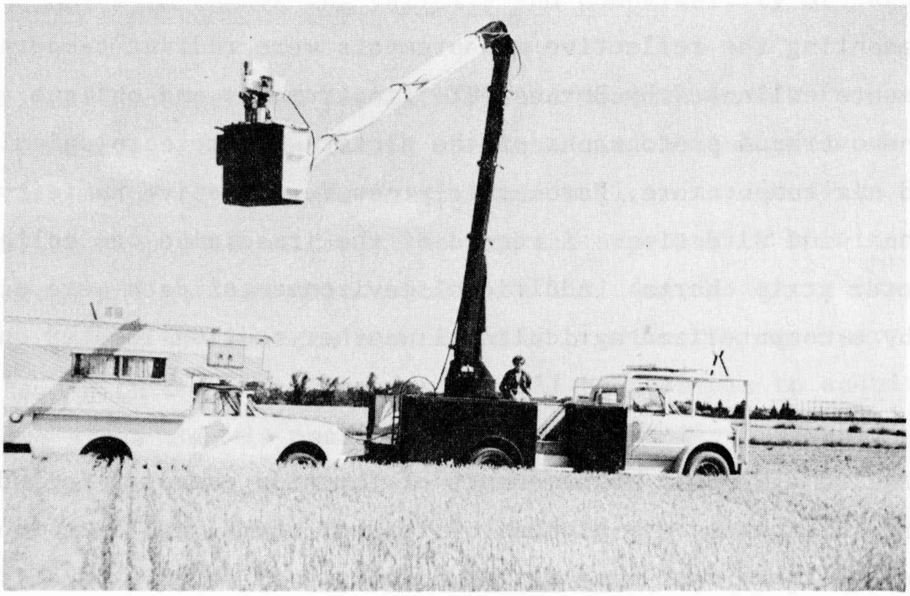


Figure 1C-1. Spectroradiometer (top) and multiband radiometer (bottom) systems used for measuring spectral characteristics of corn and soybeans.

more advanced multiband radiometer system (see Section 2), collected data in the four Landsat MSS bands.

Augmenting the reflective measurements were radiant temperature measurements collected by Barnes PRT-5 instruments and oblique ground level and overhead photographs of the plots. The meteorological data included air temperature, barometric pressure, relative humidity, wind speed, and wind direction. A record of the irradiance was collected on pyranometer strip charts. Additional environmental data were acquired hourly by a computerized agricultural weather station located on the Agronomy Farm.

Detailed agronomic measurements of the crop canopies included leaf area index, fresh biomass, dry biomass of leaves, stems, and ears (pods for soybeans), heights, maturity stage, percent soil cover, and soil moisture. Leaf samples were collected and dried for nutrient analysis. Grain yields were measured at harvest.

Data will be analyzed to determine the relationship of spectral response and maturity stage, biomass, leaf area index, height, percent soil cover, degree of stress severity and grain yield.

2. Specification of a Standardized Multispectral Data Acquisition System for Field Research

To develop the full potential of multispectral data acquired from satellites, increased knowledge and understanding of the spectral characteristics of specific earth features is required. Knowledge of the relationships between the spectral characteristics and important parameters of earth surface features can best be obtained by carefully controlled studies over areas, fields, or plots where complete data describing the condition of targets is attainable and where frequent, timely spectral measurements can be obtained. The currently available instrumentation systems are either inadequate or too costly to obtain these data. Additionally, there is a critical need for standardized acquisition and calibration procedures to ensure the validity and comparability of data.

The primary sensors used for LACIE Field Measurements were spectrometers capable of producing high-resolution spectra. In the future a new approach to the collection of field measurements data will be needed since it will not be feasible simply to multiply the current approach by the increased number of crops and regions which should be included in future experiments. Multiband radiomter systems can economically provide the necessary spectral measurements. With these instruments it will be possible to acquire measurements at more sites than is possible with the currently available high-spectral resolution spectrometer systems. And, it is more observations of crops and soils under a wide variety of conditions (not detailed spectral measurements of a limited number of locations and crop conditions) that is needed to increase our understanding of the spectral characteristics of agricultural scenes. There will be a continuing need for the high resolution spectrometer systems to be utilized in field research, but less complex systems are also required which can be used to make large numbers of measurements at many sites economically and accurately.

The overall, long-term objective of this project is to develop a multispectral data acquisition system which will improve and advance the capability for remote sensing field research. The specific objectives are to:

1. Specify, develop and test the prototype of a radiometric instrument system.
2. Develop and document calibration, measurement and operation procedures.
3. Develop software for data handling capability.

The radiometric instrument will be a multiband radiomter with up to 8 bands between 0.4 and 2.4 μm. The data acquisition system will record data from the multiband radiometer, a radiation thermometer, and ancillary sources. The radiometer and data handling systems will be adaptable to helicopter, truck, or tripod platforms.

The general characteristics of the system are that it will be: (1) comparatively inexpensive to acquire, maintain, and operate; (2) simple to operate and calibrate; (3) complete with data handling hardware and software; and (4) well-documented for use by researchers.

The specific results of the project will be: (1) a multiband radiometer; (2) a data recording and handling system; (3) a machine independent software package; and (4) a comprehensive system manual documenting the design and use of the above. The most significant result will be the establishment of the capability for researchers to acquire the data to effectively investigate relationships between the physical-biological and the multispectral characteristics of crops, soils, forests, water, and geological features.

The instrument system will be a prototype of an economical, standardized system which can be utilized by many researchers to obtain large numbers of accurate, calibrated spectral measurements. As such it is a key element in improving and advancing the capability for field research in remote sensing. The general specifications of the radiometer module are:

- Rugged, lightweight, and capable of operation from a variety of platforms
- Battery operated
- Relatively low cost
- Eight spectral bands from 0.4-12.5 micrometers (including thematic mapper bands)
- Optical filters to define each spectral band
- Optical units to be exchangeable to obtain the spectral pass bands desired by the researcher

- Selectable field of view (by lens substitution)
- All radiometric data simultaneously available in parallel format

The data handling module will convert the signals from the radiometer to digital format and store them with digital ancillary data (i.e., date, time, observation number) in a removable solid state module having a data retention battery. The data handling module will also print data, compute reflectance or radiance, and serve as the interface to enter data to digital computers and peripherals. The general specifications of the data handling module are:

- Rugged, lightweight, and portable
- Battery operated
- Relatively low cost
- Parallel multichannel analog input port compatible with analog outputs of radiometer
- A simultaneous sample and hold circuit for each detector channel (or equivalent)
- Remote activation of the acquisition sequence enabling synchronous operation with boresighted cameras operated by intervalometers
- Provision to enter date, time and observation number
- Digital hard copy printer to tabulate data values of each radiometer channel, plus ancillary data
- Digital data output port for RS 232-C and "12-bit parallel" with request to send/clear to send lines
- Auxilliary analog input ports to accomodate signals from other sensors
- Provision for reflectance factor or radiance computations

In order for researchers to acquire accurate and comparable measurements it will be important to fully document and describe the use of the instrument system. An educational package including audio-tutorial materials is being planned. It would provide researchers with the theoretical and practical knowledge to gather, process, and analyze high quality spectral data. Major topics to be treated are:

- Measurement and calibration fundamentals and procedures, including important parameters of the measurement situation and procedures for data evaluation and comparison
- Specifications for the radiometer and instrument system, system tests, and typical results
- Methods of positioning the instrument, including suggested hardware and mounting techniques for operation on a helicopter, mobile tower, and tripod platforms
- Data handling and data processing procedures including a machine independent computer program to organize and present the data
- Case studies of typical experiments including experimental design, measurement and data handling procedures, and analysis techiques and results

During the past six months, the general characteristics of the radiometer and data acquisition system were developed and discussed with active remote sensing researchers at other institutions with regard to their experimental needs and data handling preferences and capabilities. Potential vendors were contacted for their reaction to the specifications. Vendors were encouraged to suggest alternate means to obtain a practical, effective, commercially available hardware system. Following development of the final specifications, a formal request for quotation will be prepared at the beginning of the next contract year for submission to qualified vendors. The goal will be to ensure that each subsystem will be available for purchase on a continuing basis from a reliable supplier.

During the next year the following tasks and accomplishments leading to commercial availability and use of the instruments in 1980-81 are planned:

- A review panel will be established to review the development, procurement and testing of the system
- A prototype system will be acquired and tested by Purdue/LARS
- Based on tests of the prototype system, final specifications for production units will be prepared

- A system manual will be prepared
- An experimenter's manual will be prepared
- Software for storage and use of radiometric and ancillary data will be developed

Task 1D. Determining the Climatic and Genetic Effects on the Relationships between Multispectral Reflectance and Physical-Chemical Properties of Soils

Although a large body of knowledge has been accumulated about the physical and chemical characteristics of soils as they are influenced by the soil forming factors of climate, parent material, relief, biological activity, and time, there is only limited knowledge of how these factors relate to the reflected radiation from surface soils. Earlier studies have shown that information about the spectral properties of soils may be useful in their identification and characterization.

If present satellite sensors and the improved sensor systems planned for future satellites are to be used most effectively in the preparation of land use capability maps and soil productivity ratings as these relate to crop production, it is crucial to define quantitatively the soil variables related to productivity which can be measured by or correlated with multispectral radiation from the surface soil. This research seeks to contribute significantly to the understanding of the multispectral reflectance of soils as it relates to climate, physical and chemical properties of soils, engineering aspects of soil use and potential agricultural productivity.

1. Objectives

The general objective is to define quantitatively the relationships between soil reflectance and physiochemical properties of soils of significance to agriculture and engineering. Selection of soil samples with a wide range of important soil characteristics by statistical stratification of continental United State climatic zones permitted the evaluation of climatic and genetic effects on the relationships between multispectral reflectance and these soil properties. A further objective is to define the relationships sufficiently to design further research to quantify the contributions which different soil components make to the multispectral characteristics of specific soils.

The ultimate objective of this research approach is to provide a body of knowledge and interpretive skills which will render remote multispectral sensing a valuable tool for mapping soils, determining land use capabilities and soil productivity ratings, identifying crops and predicting crop yields.

2. Experimental Approach

A total of 250 soils, representing a statistical sampling of the more than 10,000 soil series in the United States were selected for this investigation. Selections were made from a list of the 1377 Benchmark soil series representing those soils with a large geographic extent and whose broad range of characteristics renders these soils so widely applicable for study.

Stratification of soil sampling was based on series type location within climatic zones. Climatic strata included the frigid, mesic, thermic, and hyperthermic soil temperature regimes as defined by the U.S. Soil Taxonomy as well as the perhumid, humid, subhumid, semiarid, and arid moisture regions as identified by Thornthwaite's 1948 Moisture Index. A random selection procedure was used within each stratified climatic zone to select a number of soils series approximately in proportion to the geographic extent of that region. Considerations were also made to include soils which represent the major parent material categories and the ten soil orders of the U.S. Soil Taxonomy.

The Soil Survey Investigations Division of the Soil Conservation Service (USDA) cooperated with LARS in the collection of field samples from 39 states. Duplicate field samples were collected for all Benchmark soil series requested: one sample from a site near the type location for the current official series, and one sample from a site located from one to 32 kilometers from the first site and in a different mapping delineation. Soil Conservation Service field survey personnel were responsible for sample collection of Benchmark soils in their locality. Of the original list of 250 Benchmark soils requested, the Soil Conservation Service has collected, properly identified, and forwarded 247 Benchmark soils consisting of 484 duplicate soil samples to LARS. This excellent response of almost 99 percent of the requested samples forms an outstanding collection of soil samples for detailed chemical, physical, and spectral analysis. All samples conform to the central concept of each individual soil series as nearly as each soil could be identified and mapped by an experienced soil surveyor in the field.

A portable spectroradiometer, the Exotech Model 20C was used in an indoor configuration with the bidirectional reflectance factor reflectometer in order to obtain soil spectral readings in the 0.52-2.32 μm wavelength range. Illumination of the 10 cm diameter sample surface was accomplished by a 1000 watt tungsten iodine lamp with transfer optics. Data were recorded on magnetic tape for subsequent reformatting into a readily accessible data logging system.

Facilities of the Purdue University Agronomy Department were used to perform a wide range of physical and chemical measurements such as texture, organic matter content, cation exchange capacity, iron oxide content and others. Measurements of liquid limit, plastic limit, plasticity index, activity, liquidity index, volumetric shrinkage, and compression index were determined in the Department of Geosciences.

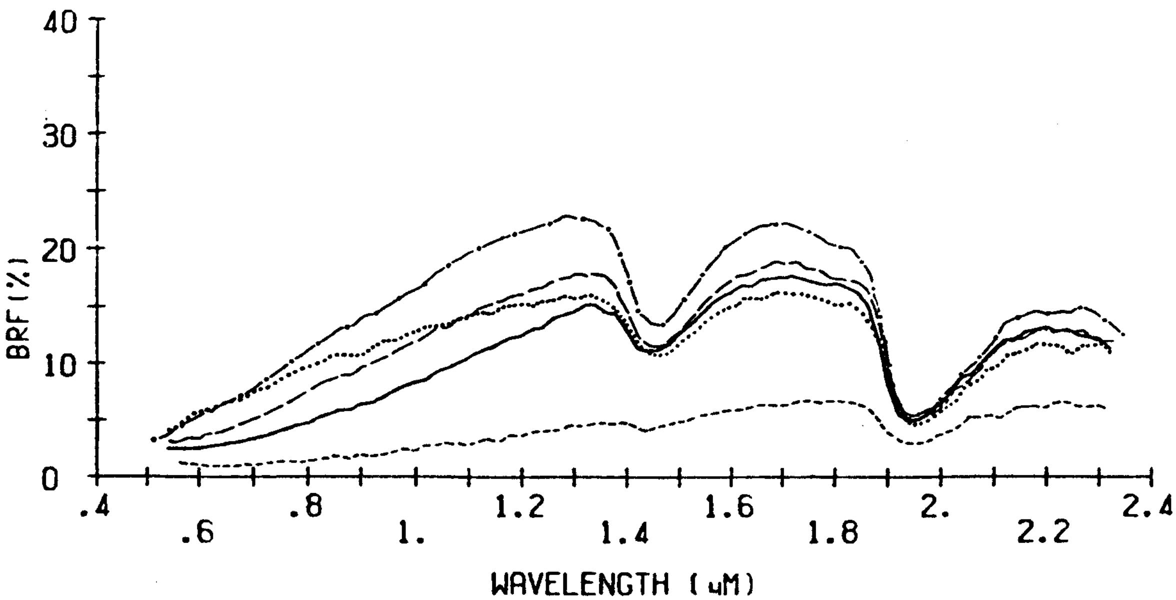
An identification record containing complete soil taxonomic information along with site characteristics and results of laboratory measurements was established and implemented for storage and rapid retrieval as part of the EXOSYS software package. The record consists of 100 items of information about each soil sample and is arranged on seven computer cards.

The need to provide an equipotential moisture condition for spectroradiometric measurements of soil samples led to the development of an asbestos tension table apparatus large enough to secure up to 28 sample holders at one time. Two of these tension tables were used at each spectral measurement session in order to equilibrate 56 soil samples at one-tenth bar moisture tension. Results using check samples verified the accuracy of this method to maintain a constant moisture equilibrium independent of experimental setup or soil sample preparation. As a result, soil spectral curves for all check samples were virtually identical.

3. Results

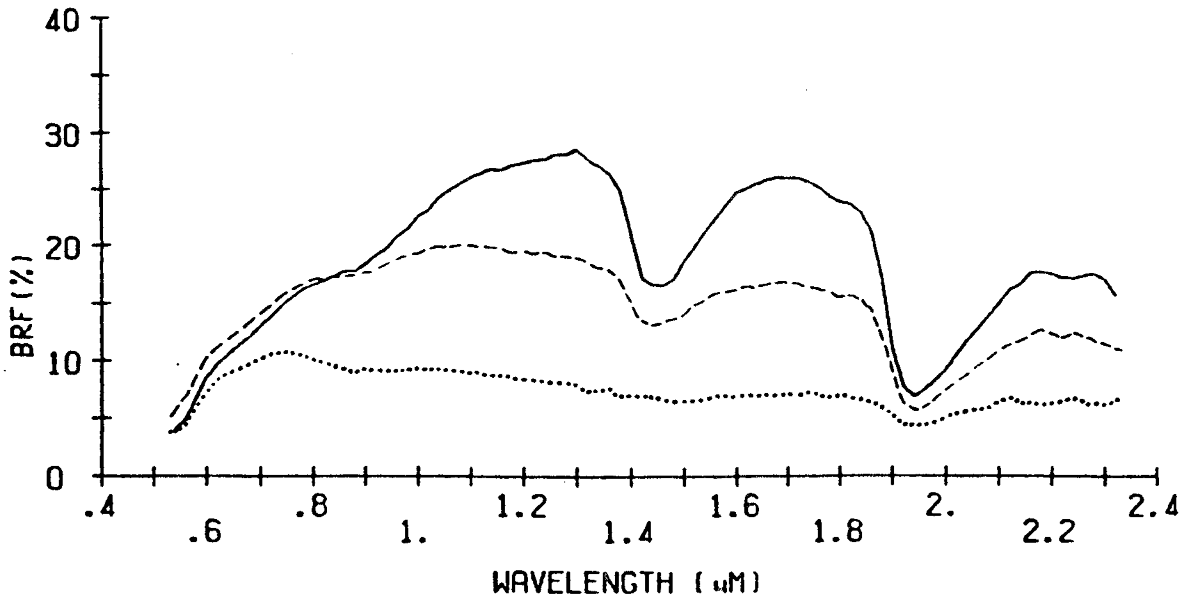
The data analysis phase of the project has started very recently, therefore only examples of the data and preliminary results can be included in this report. Comparison of soil spectral curves from four soils on which wheat is commonly grown in the Great Plains region of the United States is shown in Figure 1D-1. All four of these soils belong to the Mollisol soil order of soils with nearly black, organic-rich surface horizons and high base supply of available plant nutrients. These four prairie-formed soils are contrasted with a non-mineral organic soil from Florida belonging to the Histosol soil order. Organic matter contents of all of these soils along with Munsell color notation and texture class are as shown. The strong influence of organic matter content on soil spectral response is evident, as is the diverse nature of soil reflectance even among soils of the same soil taxonomic order.

Reliance on a subjective evaluation of soil spectral response in the visible wavelength region using a standard color chart falls short of adequately describing the spectral nature of soils. A good example can be seen in Figure 1D-2 in which two soils with the same Munsell color notation and which upon verification appeared very similar in regard to color and aggregation, exhibited strikingly different spectral response. Differences in texture and soil mineralogy are probable causes of such contrasting soil reflectance between soils of similar appearance. It is interesting to note that both the Inceptisol (a soil with weakly differentiated horizons) from Oklahoma and the Oxisol (a highly weathered intertropical region soil) from Parana State, Brazil are extensive soils upon which wheat is grown. In all three curves shown, including the Alfisol (a medium base status forest region soil), the prominent iron absorption band at 0.9 μm can be seen to affect the general trend of the spectral curves.



----	Terra Ceia (Florida) 7.5 YR 2/0 Organic	Histosol Black 80.0% O.M.
···	Richfield (Kansas) 10 YR 3/2 Silt Loam	Mollisol V. Dk. Gr. Brown 2.0% O.M.
-·-	Newtonia (Oklahoma) 7.5 YR 3/2 Silt Loam	Mollisol Dk. Brown 2.1% O.M.
——	Jansen (Nebraska) 5 YR 3/1 Loam	Mollisol V. Dk. Gray 2.7% O.M.
——	Svea (N. Dakota) 7.5 YR 2/0	Mollisol 4.8% O.M.

Figure 1D-1. Spectral response of five soils high in organic matter content.

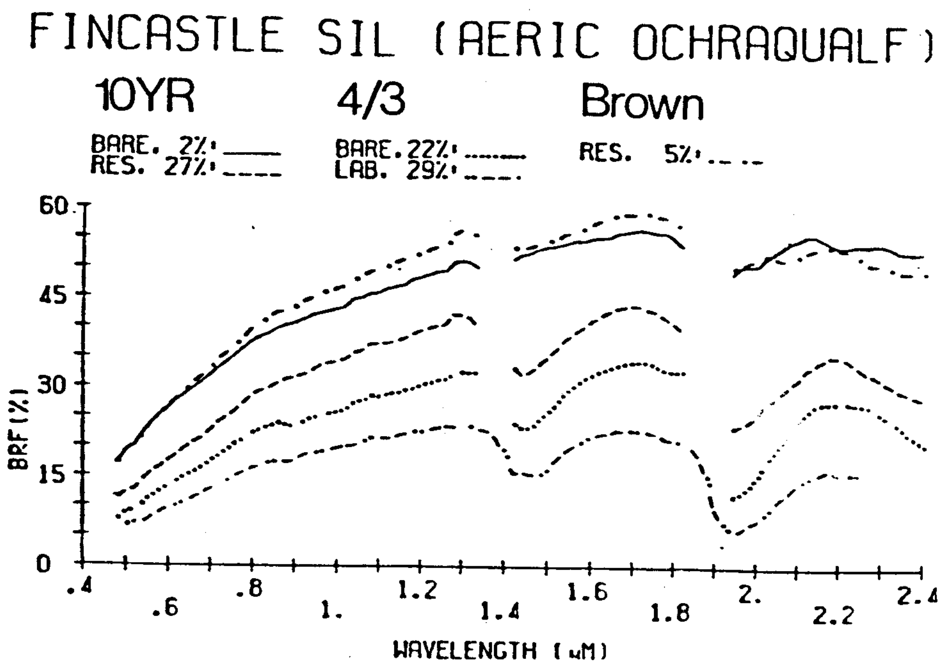


....	Londrina (Brazil) 2.5 YR 3/6 Clay	Oxisol Dk. Red 2.3% O.M.
----	Ontonagon (Michigan) 2.5 YR 4/4 Silty Clay	Alfisol Red Brown 4.4% O.M.
——	Dill (Oklahoma) 2.5 YR 3/6 V. Fine Sandy Loam	Inseptisol Dk. Red 0.3% O.M.

Figure 1D-2. Spectral response of three reddish hue soils.

Field characterization of soil spectral response has shown that individual soils have a wide range of spectral reflectance characteristics defined by specified conditions of irradiation and viewing and varying according to naturally occurring states of surface roughness, residue amount, and green vegetative cover. Comparison of field measured soil conditions with laboratory measured prepared soil samples indicates that similar spectral trends can be expected for given soils (Figure 1D-3).

During the next several months a complete statistical analysis of the data will be conducted to relate the physical-chemical properties of the sample soils to their reflectance characteristics. Several approaches are anticipated: (1) analysis of the variability of individual soil properties relative to the dupliate samples of a givensoil series, (2) analysis of the degree to which soil series can be uniquely distinguished for individual properties, and (3) analysis of the possibility of separating Benchmark soils on the basis of their reflectance characteristics.



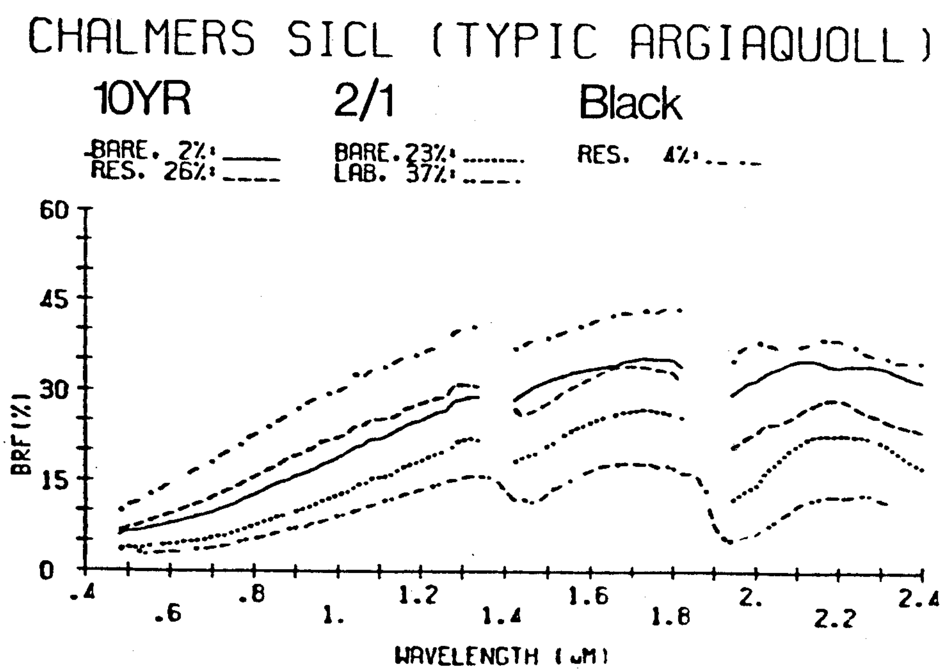


Figure 1D-3. Spectral response of two soils as measured in the field and in the laboratory. % = moisture content; RES = corn residue cover of 35%; BARE = residue-free field plots; LAB = Laboratory data.

Task 2. Processing Techniques Development

Task 2A. Application of Statistical Pattern Recognition to Image Interpretation

Analysis of remotely sensed agricultural crop survey data by pattern recognition algorithms requires the availability of training samples (data of known classification). In large-scale Landsat crop surveys, training samples cannot be acquired solely by ground observations, due either to cost considerations or to inaccessibility of the survey site. For both of these reasons, the labeling of training samples based on interpretation of the Landsat data and associated ancillary data has been utilized in LACIE, in which the manual image interpretation process has been supported by meteorological data and historical agronomic data. Although the performance of the analyst-interpreters (AIs) in LACIE has apparently been adequate to support the project goals, it is widely recognized that the labeling process, implemented in this manner, involves a great deal of subjective judgement, and hence the accuracy and precision of the results can vary greatly from one AI to the next. The overall objective of this task has been to investigate ways to upgrade the objectivity and reliability of the image labeling process. The basic approach proposed involved introduction of quantitative methods, often related to pattern recognition, in place of subjective judgement wherever possible. At the outset, it was hoped that it might be possible to develop a completely machine-implemented labeling method.

To proceduralize the labeling process, a questionnaire containing segment-related and pixel-related questions was formulated by a team at JSC to lead the AI systematically through the available image data and supporting data [1]. The supporting data included, in addition to meteorological data and historical agronomic data, a set of spectrally normalized imagery [2] and temporal greenness/brightness trajectories for each pixel to be labeled [3].

The results of using the questionnaire suggested a means for further reducing the role of the AI in the labeling process. The AI labeling of pixels as "small grains" was observed to be very highly correlated with the AI's evaluation of the temporal greenness/brightness trajectory as being typical for small grains. It had been suggested [4] that this could be quantified by taking the inner product of the first greenness image eigenvector with the temporal green number vector for the pixel to be labeled. We found that the results would be more consistent if the inner product were normalized by the 2-norm of green number vector.

Table 2A-1 compares the results obtained by labeling by computer using the inner product described above to labeling produced by three analysts using the full questionnaire. The data base consisted of seven 1976 blind-site segments in Kansas. Interestingly, the mean accuracy for the computer labeling was higher than that for any one AI, and the standard deviation was lower.

Table 2A-1. Comparison of Analyst and Computer Labeling Accuracies.

Segment	Labeling Accuracies (%)			
	Analyst			Computer
	A1	A2	A3	
1163	84.5	87.9	76.9	91.2
1165	98.4	100	79.7	95.3
1852	89.7	92.4	89.4	84.8
1855	59.4	82.8	82.8	78.1
1857	85.2	78.7	78.5	77.0
1860	65.5	62.1	66.1	87.9
1865	82.4	85.3	69.4	89.0
mean	80.7	84.2	77.5	86.2
std. dev.	13.6	11.9	7.9	6.7

It may be concluded that the temporal greenness feature can be used to automatically discriminate small grains from other classes of ground cover with accuracy rivaling that obtainable by analyst-interpreters. It is important to note, however, that the results obtained to date depend on the availiability of strategically timed data acquisitions and were obtained over a relatively limited geographical area. The potential benefits to be reaped by such an approach, measured by reduced manual labor and increased consistency of the results, appear to warrant more extensive testing.

References

1. M.D. Pore, "Label Identification from Statistical Tabulation (LIST)," Technical Memorandum LEC-10825, Lockheed Electronics Company, Houston, Texas, October 1977.

2. G. L. Kraus, "Description of the Kraus Product Imaging Technique," Internal Memorandum, Johnson Space Center, July 5, 1977.

3. R. J. Kauth and G. S. Thomas, "The Tasselled Cap - A Graphic Description of the Spectral-Temporal Development of Agricultural Crops as Seen by Landsat," Proc. Symp. Machine Processing of Remotely Sensed Data, Purdue University, June 1976.

4. R. A. Abotteen and M. D. Pore, "Development of Semi-automated, Operational Label Identification from Statistical Tabulation (LIST)," Technical Memorandum LEC-11457, Lockheed Electronics Company, Houston, Texas, February 1978.

2B. Application and Evaluation of Landsat Training, Classification, and Area Estimation Procedures for Crop Inventory

The Large Area Crop Inventory Experiment (LACIE) has demonstrated that accurate area estimates of wheat can be successfully made from classifications of Landsat MSS data. Recently, there has been increasing emphasis on making production estimates for crops other than wheat; in particular, corn and soybeans are the two crops of immediate interest. Since the LACIE system was designed, new information has been acquired on scene stratification, training sample selection, classification algorithms, and area estimation methods. This research task will build upon these recent advancements in the technology to develop and improve future crop inventory systems, particularly for corn and soybeans.

The activities of this task during the past contract year have been in the areas of development of an experiment design, stratification and sample selection for Multicrop research, recommendations for reference data acquisition, and evaluation of the LACIE Procedure 1 (P-1) for the corn/soybeans crop identification problem. Future research will include investigations of recent developments in training area selection and training, classification, and area estimation procedures using the Multicrop sample segments.

1. Stratification and Sample Selection for Multicrop Experiments

The purpose of the stratification and sampling effort was to identify the locations of the sample segments for 1978-79 Multicrop experiments to support:

- Development and evaluation of procedures for using LACIE and more advanced technologies for the classification of corn and soybeans.

- Identification of factors likely to affect classification performance such as field size, climatic conditions, and confusion crops present.

Three variables were used in the stratification: production of corn and production of soybeans (normalized by agricultural area) and average farm size (Figure 2B-1). The first two variables were selected to define strata homogeneous with respect to the relative importance of corn and soybeans in the agricultural scene. The average farm size was selected to represent problems which might be encountered in Landsat data classifications with different field sizes.

Eight strata covering 14 states in the U.S. corn and soybean producing region were determined. The large farm, highest production stratum (stratum 8) is geographically located at the center of the Corn Belt. Strata 7, 6, and 4 are located around its perimeter outward according to decreased production. In these strata of large farms, corn and soybeans are of approximately equal importance. Stratum 5 is located geographically apart from the other strata with large farms. This stratum, in which soybeans have a greater importance than corn, is located in the Mississippi River Valley where the climate and soils are more suited to soybeans than to corn. Stratum 3, the small farm stratum with the greatest production of corn and soybeans, is located primarily in eastern Indiana and western Ohio where the cropland is productive, but the terrain is rolling. The lesser production small farm strata (strata 1 and 2) are centered about this area on the outskirts of stratum 3.

Two types of samples were selected. Low density segments were distributed throughout corn and soybeans producing areas to sample all variations of conditions which could affect classification accuracy and to more completely represent conditions which might be found in other countries. Twenty low density segments were selected in each of the eight strata (Figure 2B-2). Eighty high density segments were located in four test sites spread throughout the Corn Belt (Figure 2B-3) to sample the varied climatic conditions, soil types, crop distributions, and field sizes which are present there.

2. Recommendations for Data Acquisition

A new data set required to support classification research for corn and soybeans was to be acquired by NASA/JSC. In order to insure that this data

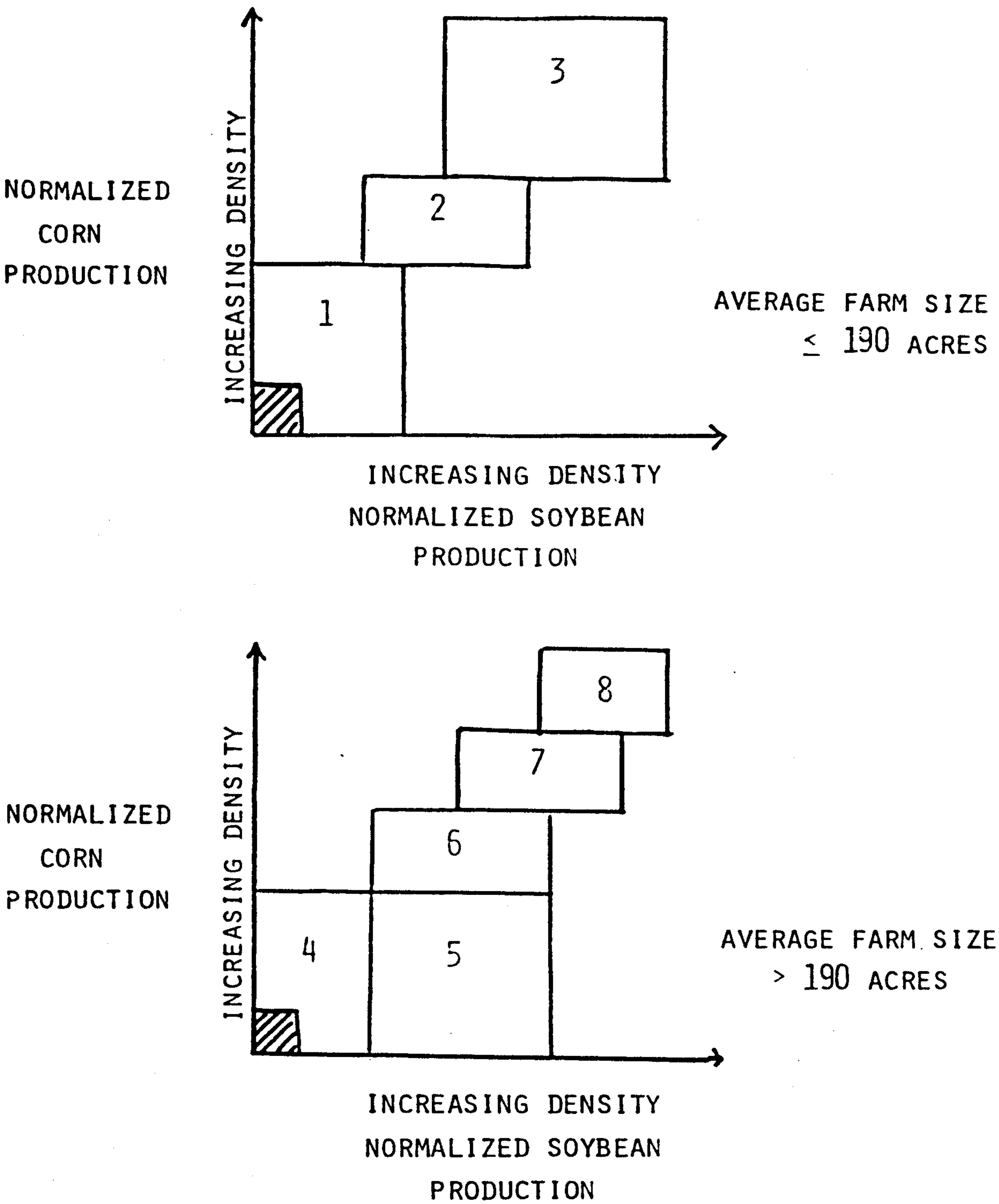


Figure 2B-1. Schematic diagram illustrating the determination of strata for Multicrop experiments based on normalized production of corn and soybeans and average farm size.

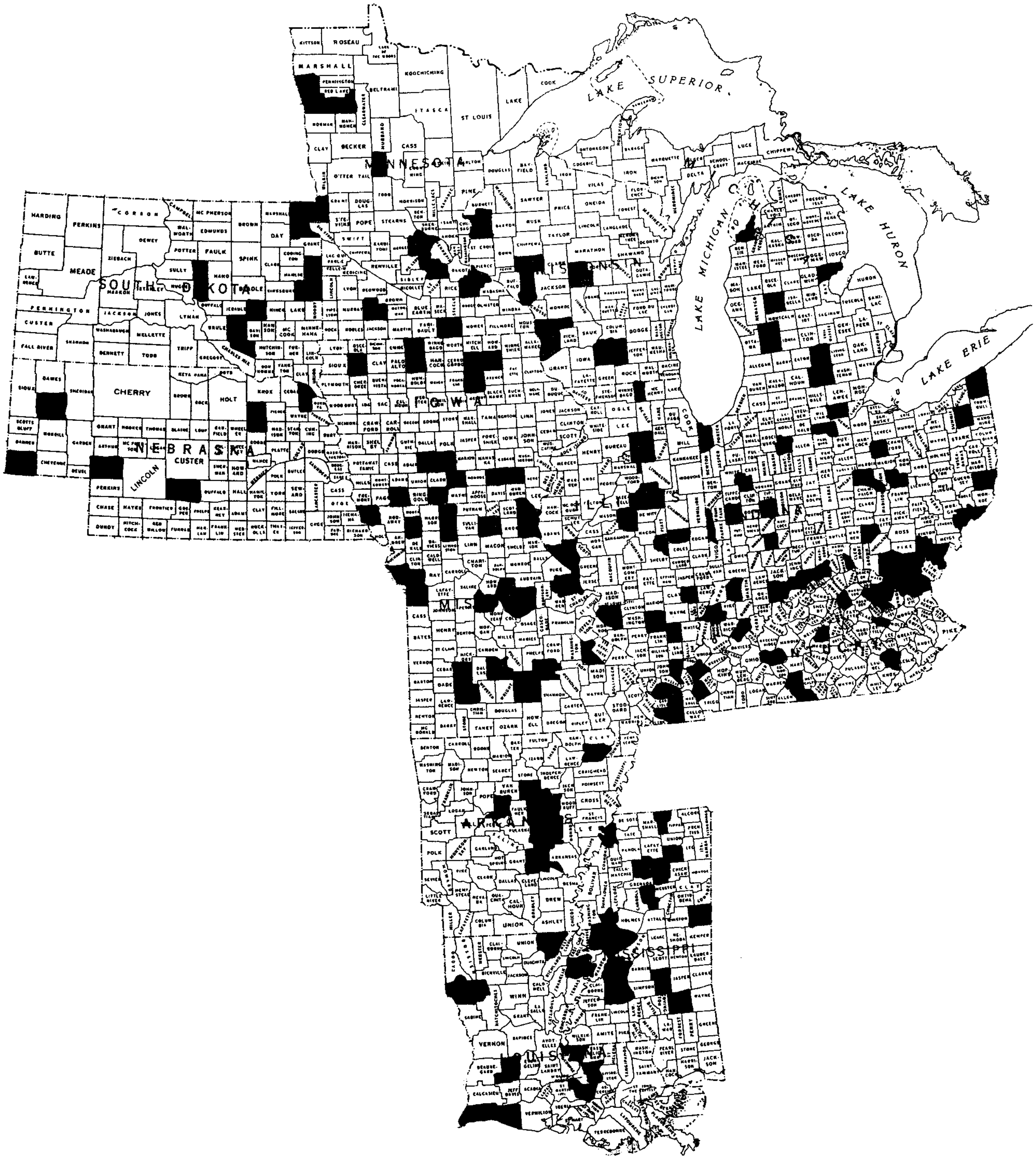


Figure 2B-2. Locations of counties receiving low density sample segments. Twenty segments were located in each of eight strata in corn and soybean producing regions to sample the variability present in confusion crops, climatic conditions, and agronomic practices.

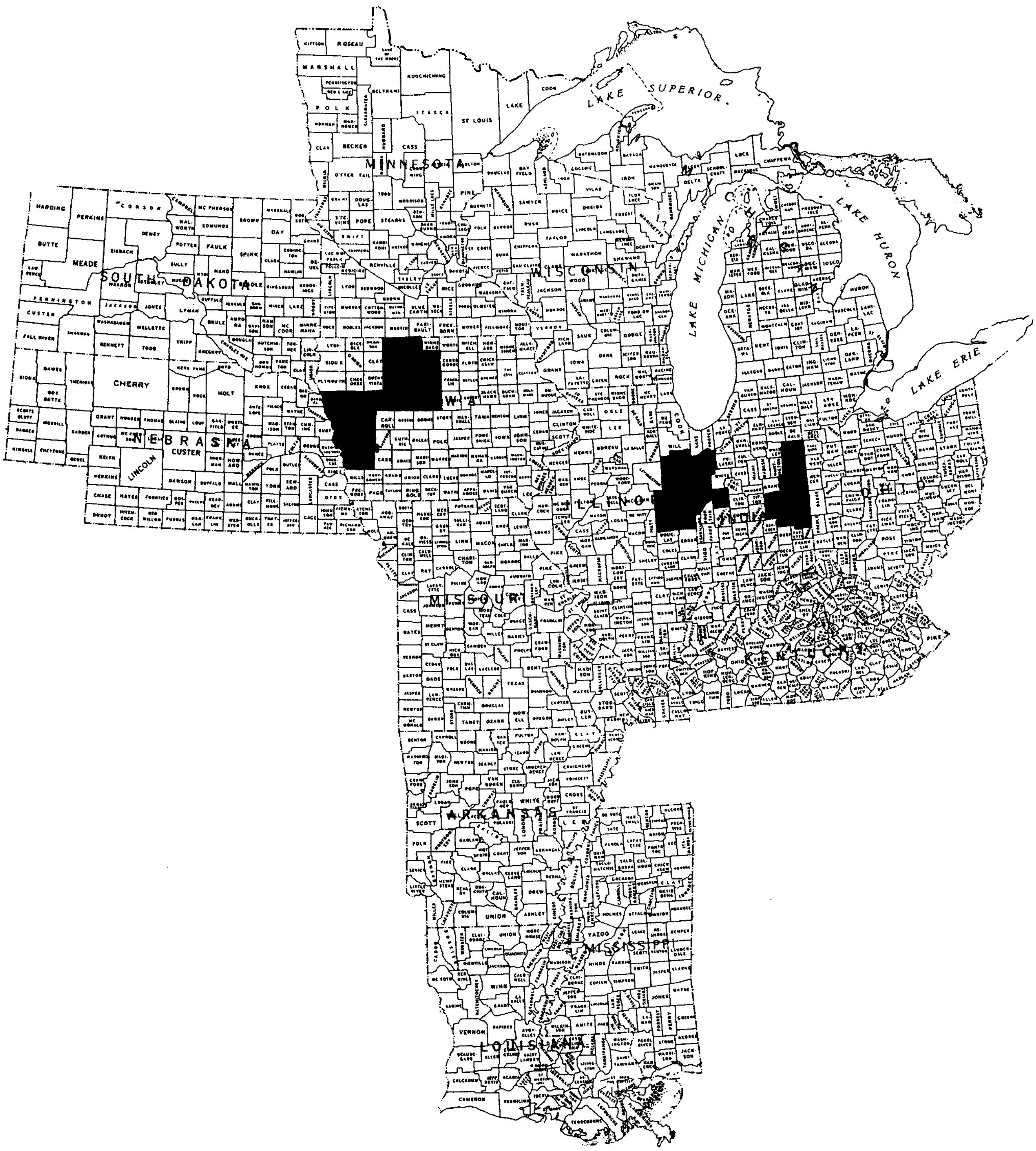


Figure 2B-3. Locations of four high density test sites for study of Landsat training, classification, and area estimation procedures for corn and soybeans. Approximately 20 sample segments were acquired in each test site.

set would meet our research objectives, recommendations were made by Purdue/LARS to NASA/JSC in the areas of crop inventory, periodic observations, and acquisition of aerial photography.

3. Evaluation of Procedure 1 for Corn and Soybeans

The analysis procedure known at Procedure 1 (P-1) was developed for use during the Large Area Crop Inventory Experiment. The procedure encompasses the areas of training, classification, and area estimation and emphasizes the use of multitemporal information. In order to allow for extension of the LACIE procedures into foreign countries, ground reference data were not used for training, but analysts labeled training data by image interpretation. This procedure significantly reduced analyst time, allowing the analyst to concentrate on just the labeling operation. Use of training data selected from a random grid reduced bias from the use of analyst-selected training fields.

In P-1, a clustering algorithm is used to statistically define the training classes. Type 1 dots are used as starting vectors for the clustering algorithm and are also used to label the resulting clusters. Following a sum of densities classification, the stratified area estimate is computed using the Type 2, "bias correction", dots to make unbiased proportion estimates.

In an assessment of the P-1 technology for corn and soybeans, this task analyzed data on corn and soybeans which were acquired during the CITARS project. A key aspect of the approach was that ground truth or photointerpreted labels were used rather than analyst labels for both Type 1 (cluster seeds and labeling) and Type 2 ("bias correction") dots. This permitted evaluation of the analysis procedure itself rather than the image interpretation accuracy.

Progress has been made in identifying areas of difficulty in the classification of corn and soybeans. Early in the investigation, several general issues in crop inventory were identified. Two such issues are: (1) in corn and soybean production areas, the practice of double cropping, particularly soybeans following winter wheat, is becoming increasingly important and (2) cloud cover, a greater problem in the Corn Belt than in the U.S. Great Plains, has potential impact on the handling of designated unidentifiable (DU) areas.

There was a significant amount of variability among stratified area estimates made using dots from different systematic grids. Empirical studies showed that dot distributions have a significant effect on proportion estimates.

Since the scene in the Corn Belt may be more complex than scenes of primarily wheat, more starting dots are believed to be needed to create a sufficient number of clusters to represent the scene. An increased number of Type 2 dots would result in a further variance reduction of the area estimates. The LACIE Transition analysis procedures call for a minimum of 40 Type 1 and 60 Type 2 dots rather than the 30 and 40 required in LACIE Phase III. It is possible that even more dots might result in significantly more accurate and precise area estimates. It appears, however, that a judicious choice of dots and a good selection of clustering and classification parameters will provide a greater improvement in results than merely selecting more dots.

4. Future Plans

Continuation and completion of the analyses described here, as well as additional analyses, are planned in the new SR&T contract to LARS. Multicrop sample segments located in Indiana, Illinois, and Iowa will be used to investigate training, classification, and area estimation procedures. In the area of sampling, the variability induced by training and/or classifying a sample of data rather than the entire segment will be evaluated. Methods of sampling from spectral space to obtain training data will be investigated. Several clustering and classification methods, including ECHO which uses spatial as well as spectral information, will be compared. The separation of the functions of sampling for training and sampling for area estimation will be investigated.

Task 2C1. Multispectral Scanner System Parameter Analysis System

The design of a multispectral scanner is a very complex matter; many different, interacting factors must be properly taken into account. Currently operational systems such as MSS have been designed primarily using subjective judgements based upon experience with experimental data. In designing a scanner the use of empirical methods, at least in part, is essential. Each of the large collections of scenes which a given scanner will be used upon is a very complex information source; not enough is known to make a simple (or even a complex) model of it by which to make the design of a scanner a simple straightforward exercise of a mathematical procedure.

And yet, more is known than when MSS was designed, and it is important to be able to carry out future designs (and indeed to decide whether future designs are needed) on a more objective basis than in the past. Thus the purpose of the present work is the development of appropriate mathematical design machinery within a theoretical framework to allow: (a) formulation of an optimum multispectral scanner system according to defined conditions of optimality and (b) an ability for convenient manipulation of candidate system parameters so as to permit comparison of the theoretically optimum designs with that of practical approximations to it.

In order to deal with the complexity of the design situation, the first step is to determine a suitable set of parameters which adequately characterize it but which is not so large as to be unmanageable. It has been observed [1] that there are five major categories of parameters which are significant to the representation of information in remotely sensed data. They are:

1. The spatial sampling scheme
2. The spectral sampling scheme
3. The signal-to-noise ratio
4. The ancillary data type and amount
5. The informational classes desired

Furthermore the effect of these parameters are interrelated to one another Thus, it is necessary to have present in the design machinery, some means for evaluating the impact of change in parameter values in each of these five categories.

Such a scanner design tool has been assembled in the form of a software package for a general purpose computer. Each of the parts of this package, called Unified Scanner Analysis Package (USAP) has been carefully devised and the theory related to it fully documented [2,3,4,5].

USAP itself gives explicit concern for the spatial, spectral and noise characteristics of the systems; the other two parameter groups are treated by means input data selection. A diagram of the system is shown in Figure 2C1-1. USAP is composed of two distinct subsystems. The spatial aspect of it contains (a) a data spatial correlation analyzer, (b) a scanner IFOV model and (c) a random noise model. The spectral portions are capable of producing an optimum spectral representation by modeling the scene as a random process as a function of wavelength followed by the determination of optimum generalized spectral basis functions. Conventional spectral bands can also be generated.

Two different data bases are used in the system. The spectral techniques require field spectral data because of their spectral detail while the spatial techniques require multispectral scanner generated data, aircraft and/or satellite. The system performance, defined in terms of the classification accuracy, is evaluated by two parametric algorithms. Following is a further description of the various parts of the system.

1. Spatial Characteristics

Two of the major spatial variables which bear upon class separability are the degree of correlation between pixels and the shape of the point spread functions (PSF). If adjacent pixels are correlated with one another it has been long observed experimentally that classification accuracy is diminished. In this work the relationship between these variables has been worked out [2] and the ability to study them experimentally has been incorporated into USAP.

Figure 2C1-1. Block Diagram of the Unified Scanner Analysis Package (USAP).

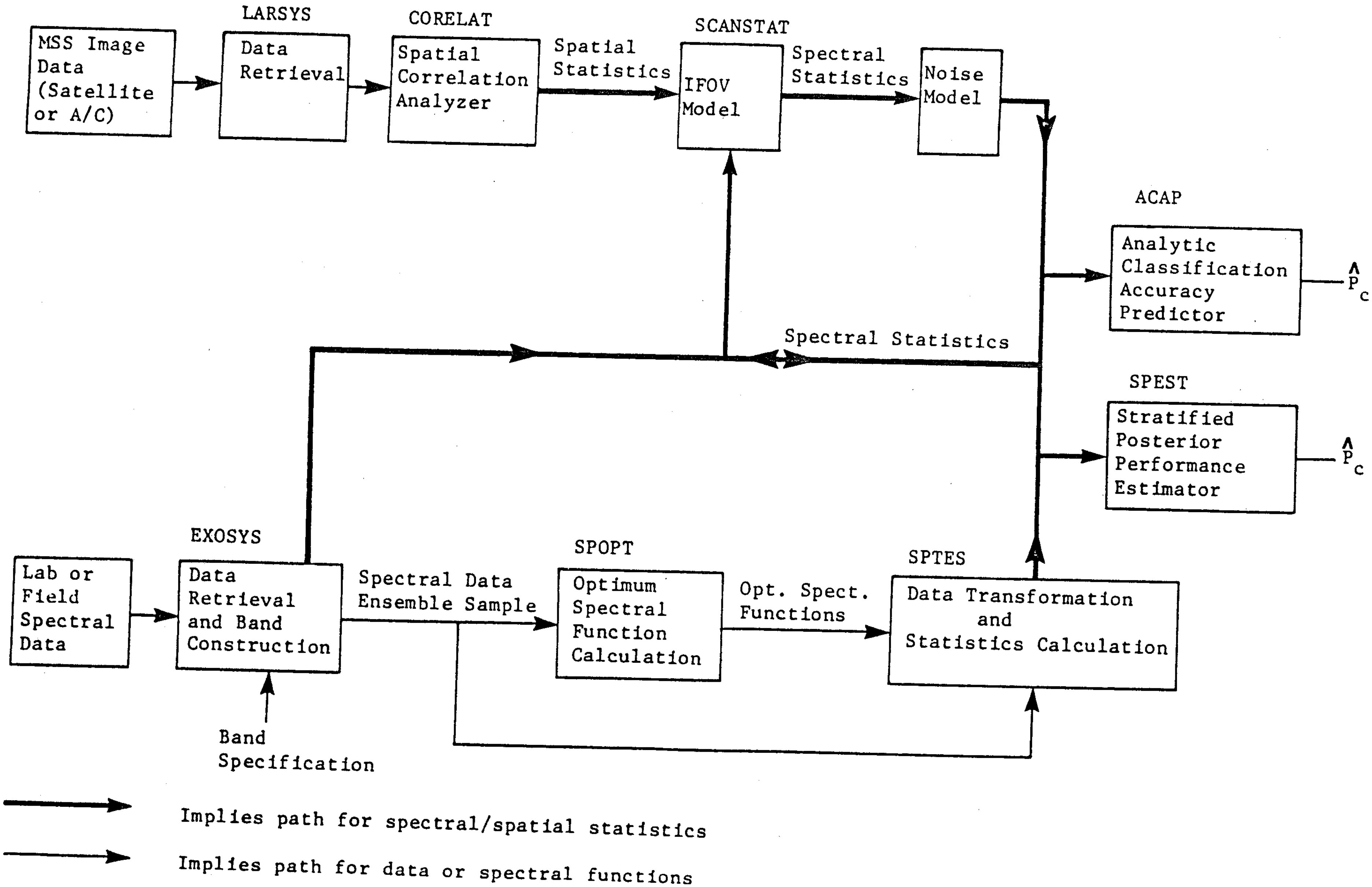


Figure 2C1-2 shows the result of having varied the adjacent sample correlation (ASC) on some hypothetical data. The graphs show classification accuracy vs. Instantaneous Field of View (IFOV). USAP has the capability to simulate PSF's of the IFOV which are either rectangular or Gausian in shape.

2. Random Noise

Additive random noise, which enters at various points in a scanner system, can degrade the overall system performance substantially. The noise can be classified into two broad categories: external and internal. A substantial source of external noise is atmospheric, due to absorption (e.g., water vapor) and scattering; the detector and quantization noise comprise a major component of internal noise. In USAP the impact which various noise conditions have can be simulated by modifying the class statistics accordingly. Figure 2C1-3 shows the result of several signal-to-noise ratios (SNR) on classification accuracy for the hypothetical data of Figure 2C1-2.

3. Spectral Characteristics

Current remote sensing schemes rely heavily upon the spectral detail present in the data to distinguish between classes. However, not much is known about where the limits of this ability to discriminate between classes are relative to the amount spectral detail preserved in the data collection process. Furthermore, because of the limited knowledge in this area, objective procedures for designing the spectral bands of a multispectral scanner are not available.

A major factor to take into account in a satellite scanner design is that the scanner must operate effectively over all types of land scenes at all times of the year for a wide variety of applications. As a result the system cannot be designed to provide for maximal discrimination between a given set of classes. Rather an optimal design is one which most completely represents the entire ensemble of spectral functions. Furthermore for a truely optimal design the representation must be efficient in the sense of achieving the minimum expected error in the smallest possible number of terms (features).

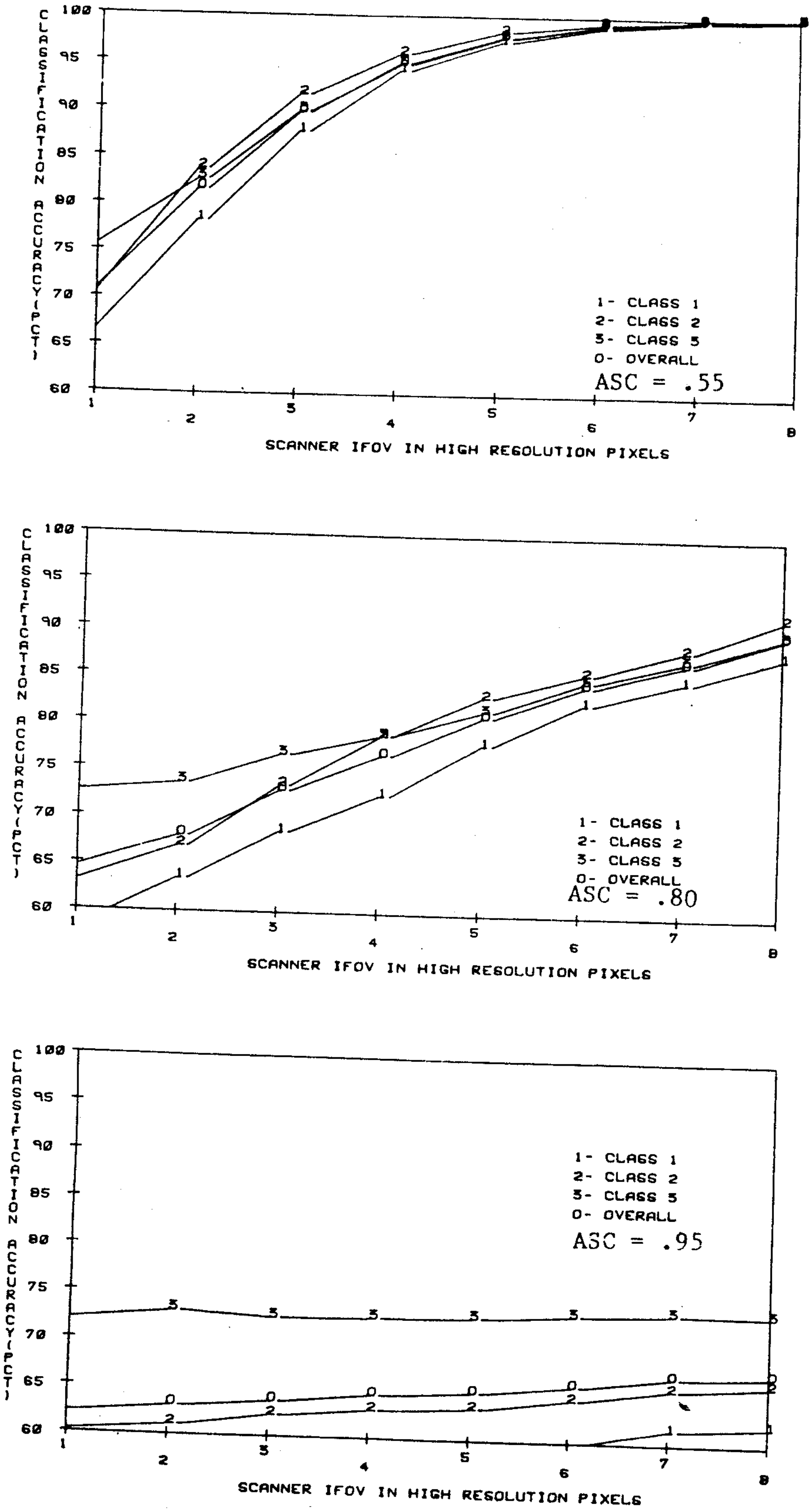


Figure 2C1-2. The result of varying the Adjacent Sample Correlation (ASC). Note that the classification accuracy improves as the IFOV is increased but this effect is very dependent on ASC.

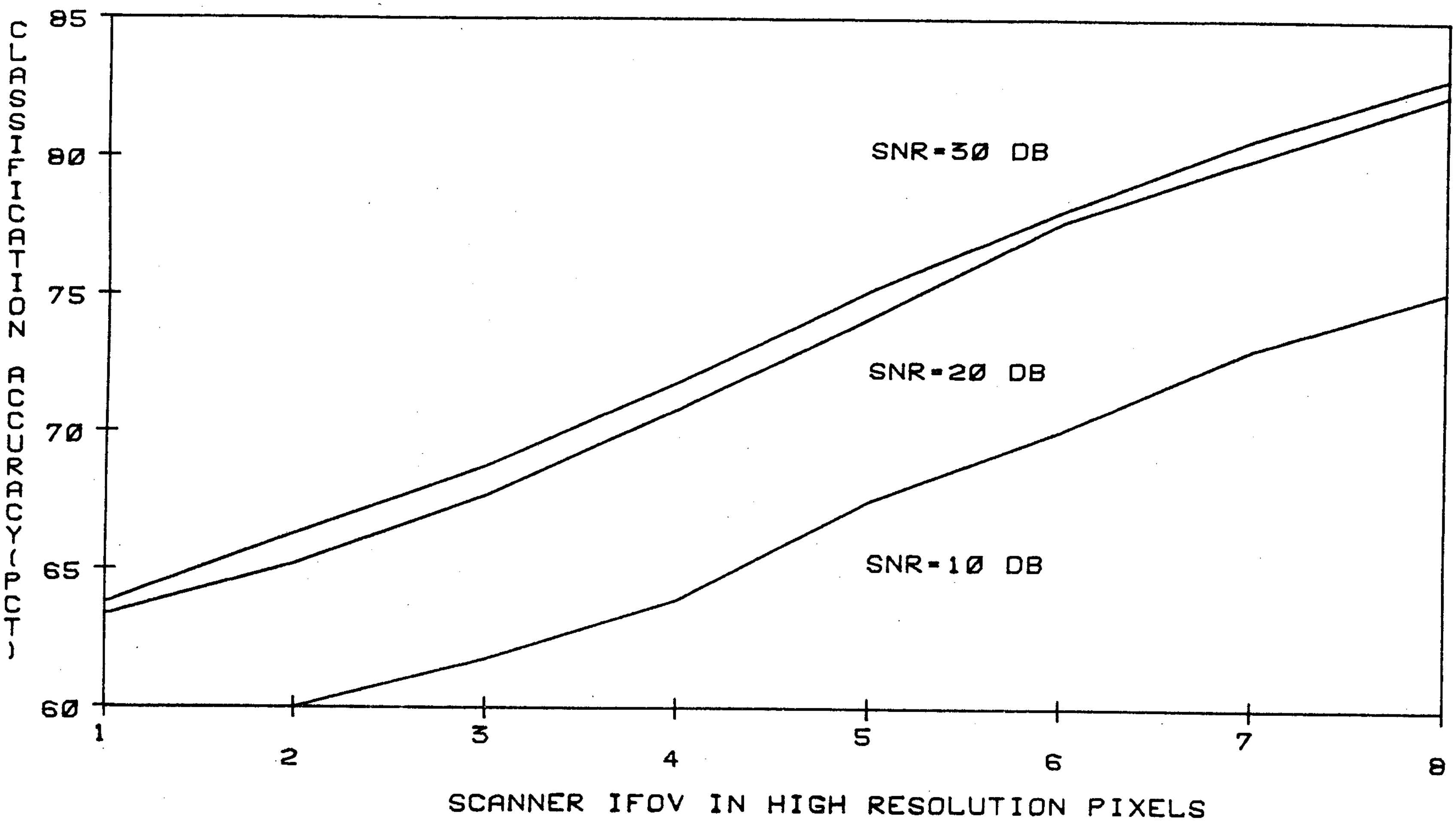


Figure 2C1-3. Overall Output Classification Accuracy Variation with Noise and IFOV.

In our development of the optimal representation scheme [3], the restriction that spectral features must occur in discrete, non-overlapping bands or even that the spectral sensitivity must always be positive has also been removed. Thus the spectral features (they cannot be called spectral bands because of the above generality) so derived would not be realizable in an actual scanner; rather they serve by providing an upper bound on attainable separability against which practical bands can be compared. USAP provides the capability not only of determining these optimal features and estimating the resulting probability of correct classification, but also to compare their performance with that from any arbitrary, practical set of features one might want. Also incorporated is the ability to weight any wavelengths more or less heavily than others.

Shown in Figure 2C1-4 are the magnitudes of the first four optimal features for an example data set. In this case the weighting function has been set to zero in the water absorption bands around 1.4 and 1.9 μm; it was set to 1.0 at all other wavelengths. The data set used consists of about 1200 representative spectra from a Williston, North Dakota area collected in May 1977. The spectra cover the range of $0.40 \leq \lambda \leq 2.4$ μm and three classes, wheat, fallow, and pasture where represented. The magnitude of the optimal features are shown because this tends to indicate the relative importance of the various wavelengths.

In this case the first feature has the general shape of the spectral response of bare soil, due no doubt to the fact that two of the three classes, fallow and wheat, would be predominantly bare soil as a result of the early season collection date. In the second feature the importance of relatively broad individual spectral regions (bands) begins to emerge; specifically the regions $0.7 \leq \lambda \leq 1.3$ and $2.0 \leq \lambda \leq 2.4$ are suggested by this optimal feature.

The higher order optimal features derived by this means tend to represent increasingly fine (high spectral resolution) detail. Figure 2C1-5 shows the tenth optimal feature for this data set, for example.

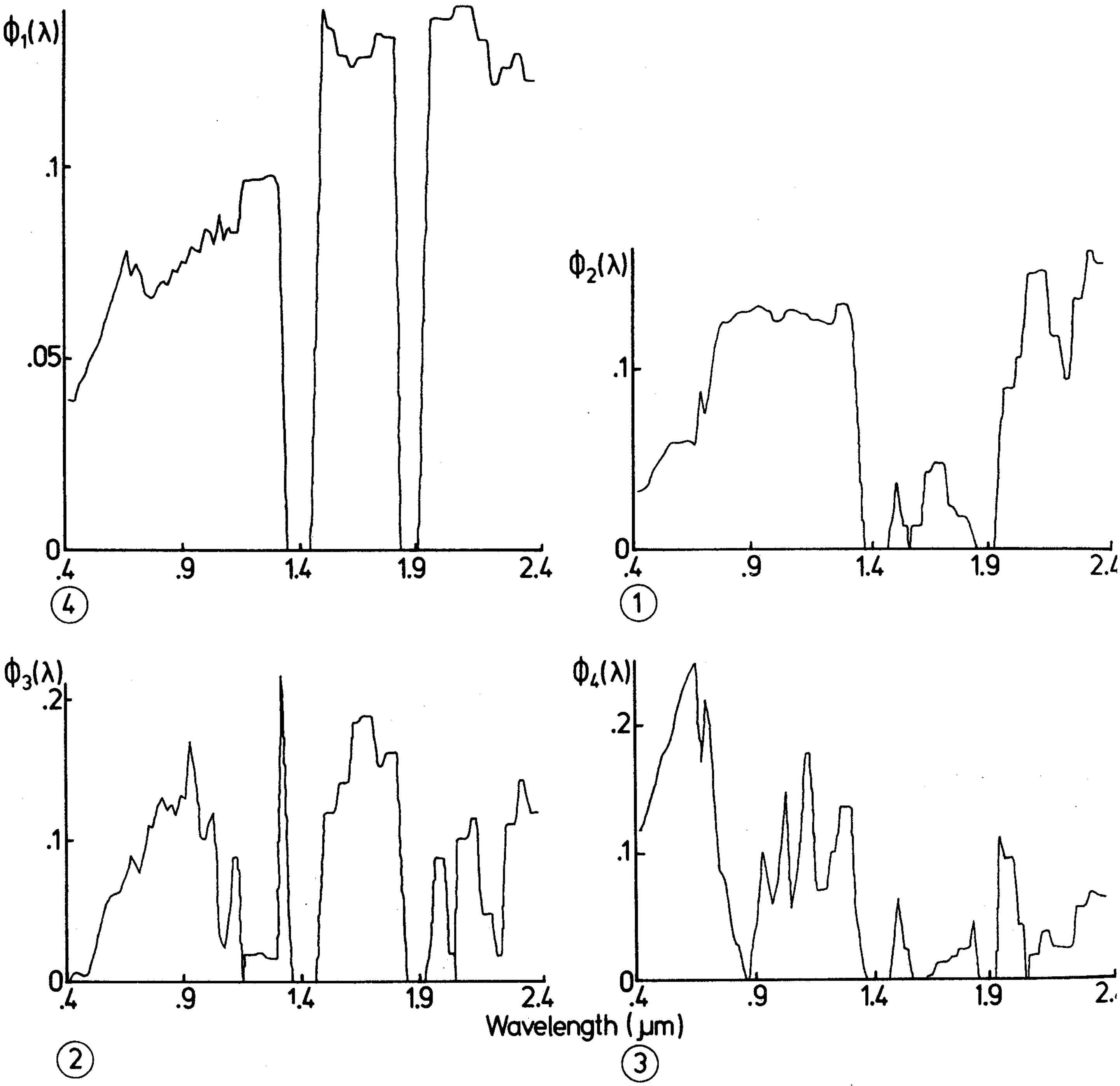


Figure 2C1-4. The magnitude of the first four optimal spectral basis functions for a data set from the Northern vs. Great Plains in early season (May). The data set contained spectra from wheat, fallow and pasture classes. A weighting function set to zero in the water absorption bands near 1.4 and 1.9 μm was used. The encircled number shows the order of selection by a feature selection algorithm for discriminating between the three classes.

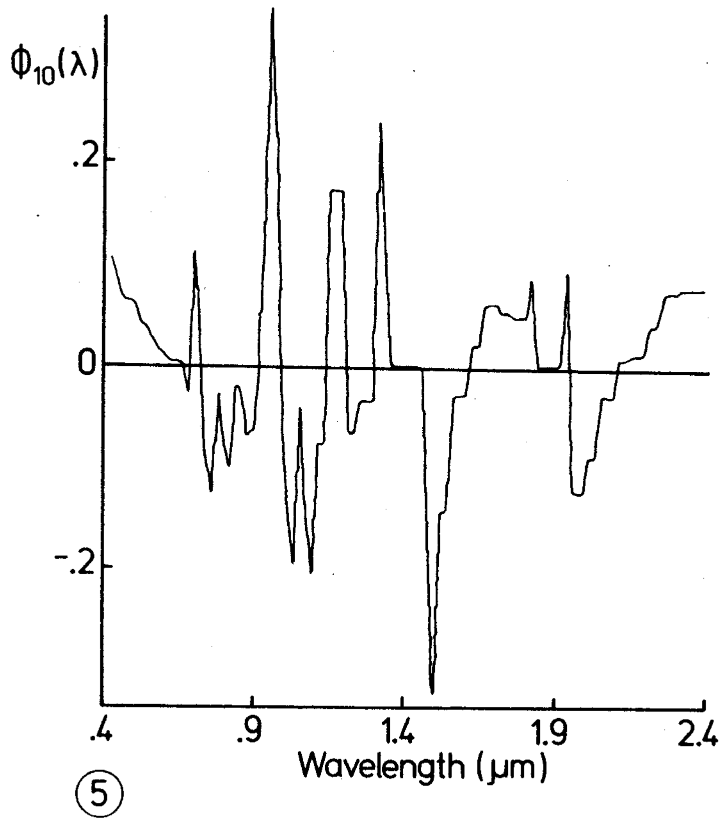


Figure 2C1-5. The tenth optimal spectral basis function for three classes of early season Northern vs. Great Plains data. Higher order basis functions tend to concentrate on higher spectral resolution characteristic of the spectra. Note that this feature was selected fifth out of the first ten in terms of its ability for discriminating between wheat, fallow, and pasture.

In addition to subjectively evaluating the optimal spectral functions in this fashion quantitative means for evaluation provide additional insight. Specifically, feature selection, representation error estimation, and probability of correct classification error estimation are useful for this purpose. These optimal features were subjected to the feature selection processor of LARSYS to determine in what order they would be selected for purposes of discriminating between the three classes. The encircled numbers near each graph of Figures 2C1-4 and 2C1-5 give this order when the first ten optimal features are selected in this fashion, i.e., $\phi_2(\lambda)$ was selected as the most important single feature, $\phi_3(\lambda)$, second, etc.

The plot in Figure 2C1-6 shows the manner in which representation error is reduced as features are added while Figure 2C1-7 shows the relation of this error to the probability of correct classification. For reference purposes, the $\log_{10}$ MSRE for both the MSS bands and those of Thematic Mapper which are in this interval would be 3.37 and 2.17, respectively, on this scale.

These results cannot be regarded as definitive as the use of additional data sets at different locations, time of the year and for different applications must be used. However, some aspects of them were unexpected and could have very significant meaning. For example the fact that P_{cc} is still increasing after ten terms and the facts the higher ordered features, representing finer spectral detail, contribute significantly to improved discrimination have not to our knowledge been reported before.

4. Probability Error Estimation

As seen in Figure 2C1-1, two different algorithms have been incorporated into USAP for estimating the probability of error, P_e, or its compliment the probability of correct classification, P_{cc}, for a given set of statistics. What is required is a parametric classification accuracy estimator for a multiclass, multidimensional Gausian Bayes classifier. Two different techniques are used to insure efficient evaluation regardless of the number of features or classes.

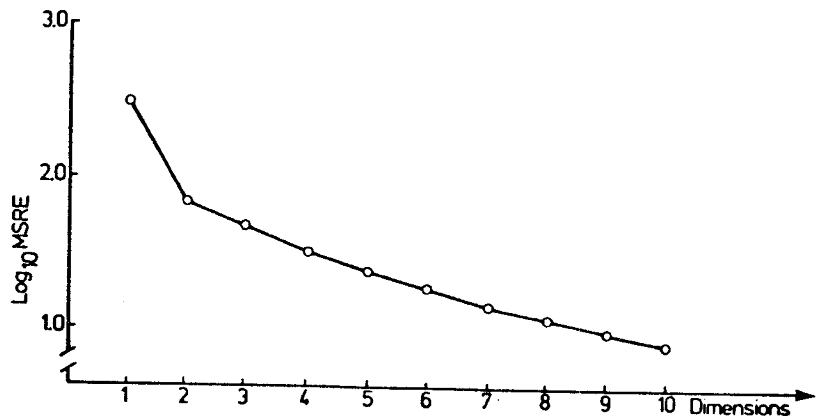


Figure 2C1-6. Log mean square representation error vs. the number of dimensions.

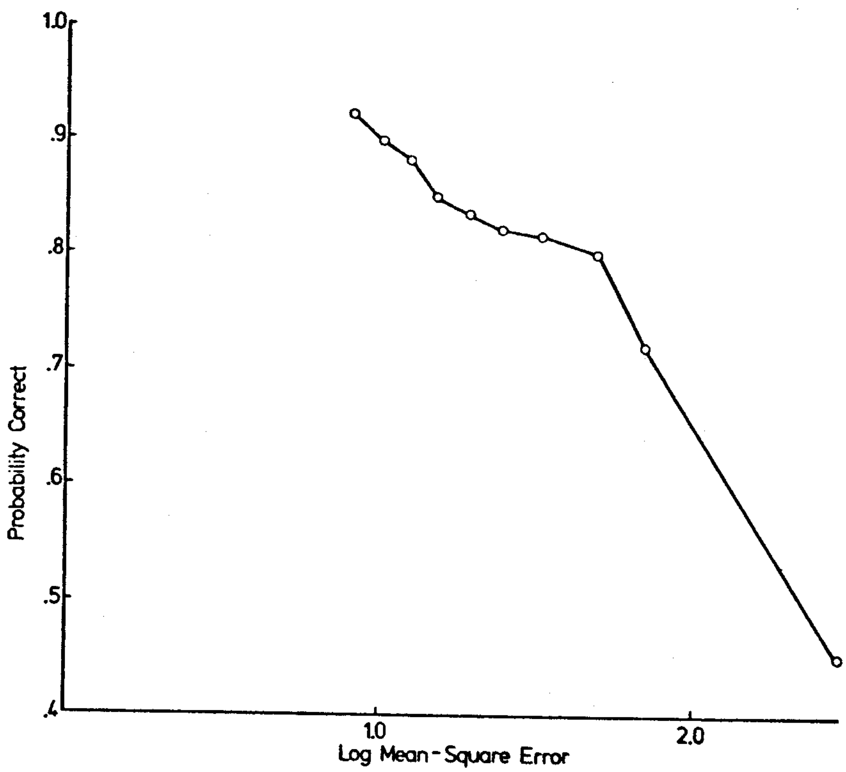


Figure 2C1-7. Probability of correct classification vs. representation error for the optimal basis functions.

ACAP (for Analytical Classification Accuracy Prediction) [2] uses an approximation scheme for directly evaluating the integral of the class conditional density functions within each decision region. It is especially efficient for lower dimensionality cases. Table 2C1-1 shows a comparison of ACAP predicted accuracy compared to actual classification results for simulated multispectral data derived from actual multispectral statistics.

Table 2C1-1. ACAP and Classification Results Comparison for Simulated Data

CLASS	% CORRECTLY CLASSIFIED	ACAP %	DIFFERENCE %
Bare Soil	77.8	78.3	0.5
Corn	91.2	91.0	-0.2
Pasture	95.3	95.1	-0.2
Wheat	94.2	93.9	-0.2
Overall	89.6	89.6	0.0

SPEST (for Stratified Posterior Performance Estimator) uses a Monte Carlo approach [4]. Table 2C1-2 provides results of a test of the accuracy of this estimator for dimensionality from one to ten. The three class statistics have idealized mean and covariance matrices so that it is easy to determine the correct accuracy by direct calculation.

Table 2C1-2. Test of the SPEST Error Estimate

No. of Features	Correct Estimate P_{c_1}	P_{c_2}	P_{c_3}	P_c	SPEST Estimate $\hat{P}_{c1}$	$\hat{P}_{c2}$	$\hat{P}_{c3}$	$\hat{P}_c$	Difference $\hat{P}_c - P_c$
1	0.6915	0.3829	0.6915	0.5886	0.6859	0.3793	0.7001	0.5884	-.0002
2	0.7602	0.5205	0.7602	0.6803	0.7671	0.5116	0.7700	0.6829	.0026
3	0.8068	0.6135	0.8068	0.7423	0.8037	0.6202	0.8081	0.7440	.0017
4	0.8413	0.6827	0.8413	0.7885	0.8283	0.6852	0.8550	0.7895	.0010
5	0.8682	0.7364	0.8682	0.8243	0.8642	0.7425	0.8703	0.8256	.0013
6	0.8897	0.7793	0.8897	0.8529	0.8767	0.7939	0.8787	0.8498	-.0031
7	0.9071	0.8141	0.9071	0.8761	0.8993	0.8242	0.9065	0.8766	.0005
8	0.9214	0.8427	0.9214	0.8951	0.9129	0.8472	0.9240	0.8947	-.0004
9	0.9332	0.8664	0.9332	0.9109	0.9193	0.8809	0.9360	0.9120	.0011
10	0.9431	0.8862	0.9431	0.9241	0.9209	0.9012	0.9481	0.9234	-.0007

References

1. Swain, P.H. and Davis, S.M., Editors, Remote Sensing: The Quantitative Approach, Section 7-3, McGraw-Hill International Book Company, 1978.

2. Mobasseri, B.G., "A Parametric Multiclass Bayes Error Estimator for the Multispectral Scanner Spatial Model Performance Evaluation," Ph.D. Thesis TH29340. LARS Technical Report 061578, 1978.

3. Wiersma, D.J. Ph.D. Thesis to be published.

4. Whitsitt, S. J., "Error Estimation and Separability Measure in Feature Selection for Multiclass Pattern Recognition." Ph.D. Thesis TH28539. LARS Technical Report 082377. 1977.

5. Moore, D.S., Whitsitt, S. J. and Landgrebe, D. A., "Variance Comparisons for Unbiased Estimators of Probability of Correct Classification" IEEE Transactions on Information Theory, Vol. IT-27, No. 1, January 1976.

Task 2C2. Multisensor Multidate Spatial Feature Matching, Correlation, Registering, Resampling and Information Extraction

This task was formulated to seek answers to the problems of data merging and information extraction using multiple remote sensing and ancillary data types. The specific remote sensing data type considered in this contract year were synthetic aperture radar (SAR) and Landsat data. Methods of merging map data and remote sensing data are also considered. Interest is growing in the remote sensing community in the utility of radar imagery as an addition to Landsat data. The tasks were oriented toward determining the spatial and spectral characteristics of SAR data and definition of merging system parameters.

Airborne synthetic aperture radar image data over a site near Phoenix, Arizona flown on June 17, 1977 was studied. The problem of registering this data type with Landsat imagery was considered. Due to Landsat data availability problems time coincident data was not obtained and an October 1972 Landsat data set was used. Field patterns and roads were generally the same permitting registration analysis to proceed. Control points were manually obtained from the SAR and Landsat data and several distortion functions were evaluated. Table 2C2-1 presents the results of this analysis. For the block of data analyzed an affine representation was adequate to achieve sub-pixel registration. The grid size used was 25 meters. Cubic convolution was used to resample the 57 by 79 meter Landsat data to the output grid size.

Having achieved an accurate registration an analysis was performed to determine the potential for automatic registration of SAR and MSS data. Correlation tests were run on several blocks of the registered data. Results indicated that the correlation between SAR and Landsat is very low. Gradient enhancement was performed and the enhanced images were correlated with further negative results. Figure 2C2-1 contains the Landsat and SAR images along with the gradient enhancements. Similar structure is clearly evident and further study is recommended to seek a method which will enable automatic matching of the SAR and MSS data types.

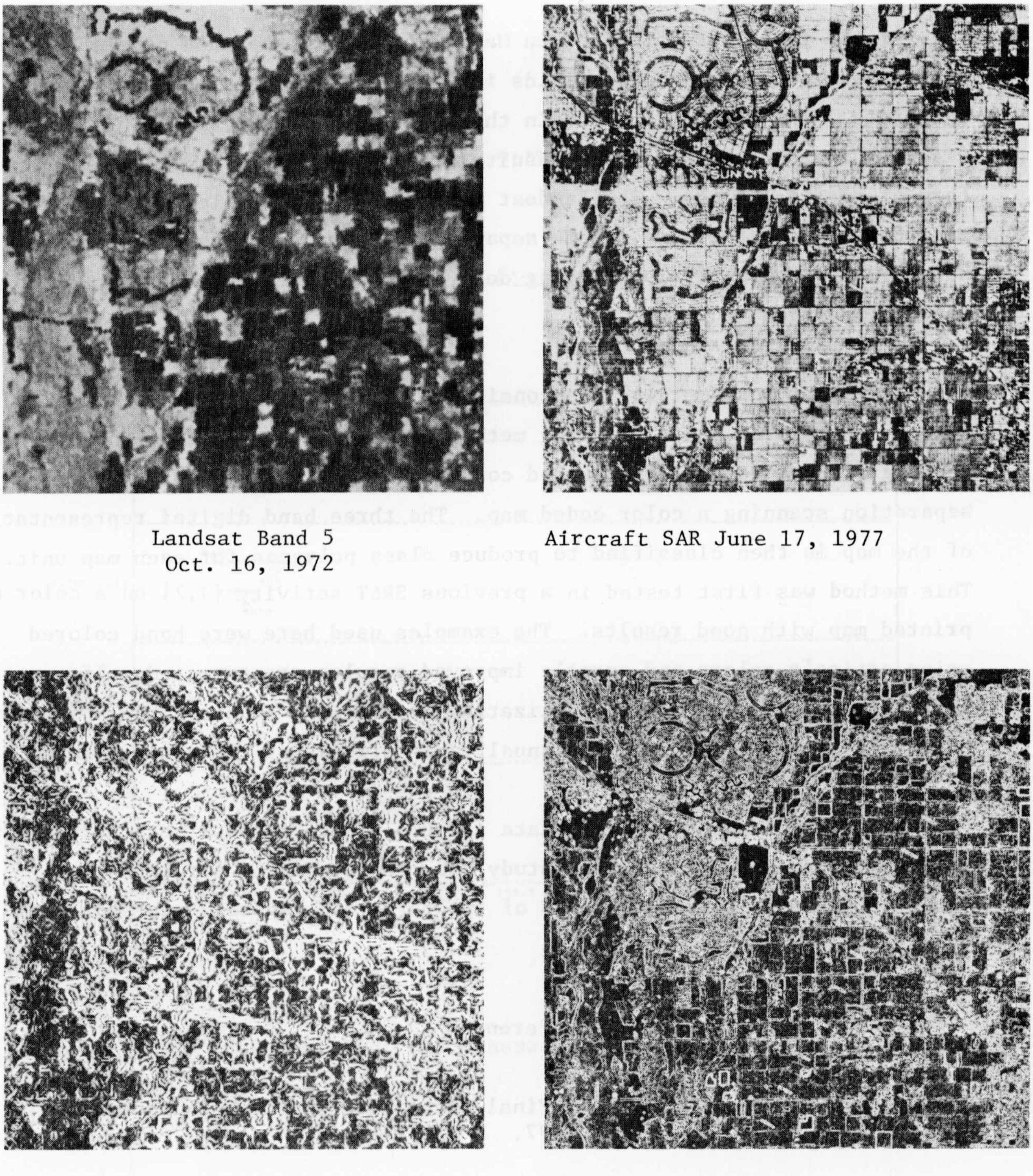
Table 2C2-1. Evaluation of Phoenix Overlay Models

Distortion Model	# of Terms	# of Points	SAR Line R.M.S.	SAR Column R.M.S.	Landsat Line R.M.S.	Landsat Column R.M.S.
			Residuals in Reference Frame			
Affine						
SPSS/CDC	3	17	3.89	3.91	0.91	0.66
SPSS/IBM	3	17	3.89	3.91	0.91	0.66
LARS/IBM	3	17	3.53	3.58	0.90	0.81
Biquadratic						
SPSS/CDC	6	17	3.02	3.37	0.67	0.67
SPSS/IBM	6	17	3.02	3.37	0.67	0.67
LARS/IBM	6	17	2.72	2.92	0.54	0.66
Bicubic						
SPSS/CDC	10	17	2.53	1.54	0.21	0.65
SPSS/IBM	10	17	2.53	1.54	0.21	0.65
Biquintic						
SPSS/CDC	10	17	---	---	0.28	0.62
SPSS/IBM	10	17	3.15	0.07	0.28	0.62

LANDSAT Grid Size - 76.2M. x 61.10M.

SAR Grid Size - 14.8M. x 14.2M.

Registered Grid Size - 25M. x 25M.



Landsat Band 5
Oct. 16, 1972

Aircraft SAR June 17, 1977

Mag. of Gradient of
Landsat Band 5

Mag. of Gradient of SAR

Figure 2C2-1. Original and Gradient Images of Landsat and SAR Data from Phoenix, AZ site. 25 M pixels.

Ground truth was collected in March of 1978 in the area covered by the SAR and the contents of many fields in June 1977 was obtained from farm operators. The SAR reflectance in these fields was analyzed using the LARSYS STATISTICS processor and these results are presented in Figure 2C2-2 in the form of spectral plots. The Landsat responses are not relevant here since the data date is not the same. Some separation between classes is observed; however, further analysis is being delayed until the time coincident Landsat data is received.

Ancillary data merging was considered from the standpoint of map digitization. A colored map digitizing method was further evaluated; however, no results are available. The method consists of photographing and color separation scanning a color coded map. The three band digital representation of the map is then classified to produce class polygons for each map unit. This method was first tested in a previous SR&T activity [1,2] on a color dot printed map with good results. The examples used here were hand colored using artist's colors and greatly improved results are expected. If successful an alternate ancillary data digitization method will be available which may be an attrative alternative to manual table digitizing for complex maps.

In addition to the Phoenix data set two other SAR data sets were analyzed. These were generated in another study and made available for this task. Table 2C2-2 described the details of the three SAR Landsat data sets now available in the LARS data base.

Reference

1. Landgrebe, D. A. and Staff, Final Technical Report, NASA Contract NAS9-14970, Volume 2, May 1977.

2. Chu, N.Y. and P. E. Anuta, Automatic Color Map Digitization by Spectral Classification to be published by Photogrammetric Engineering and Remote Sensing Journal.

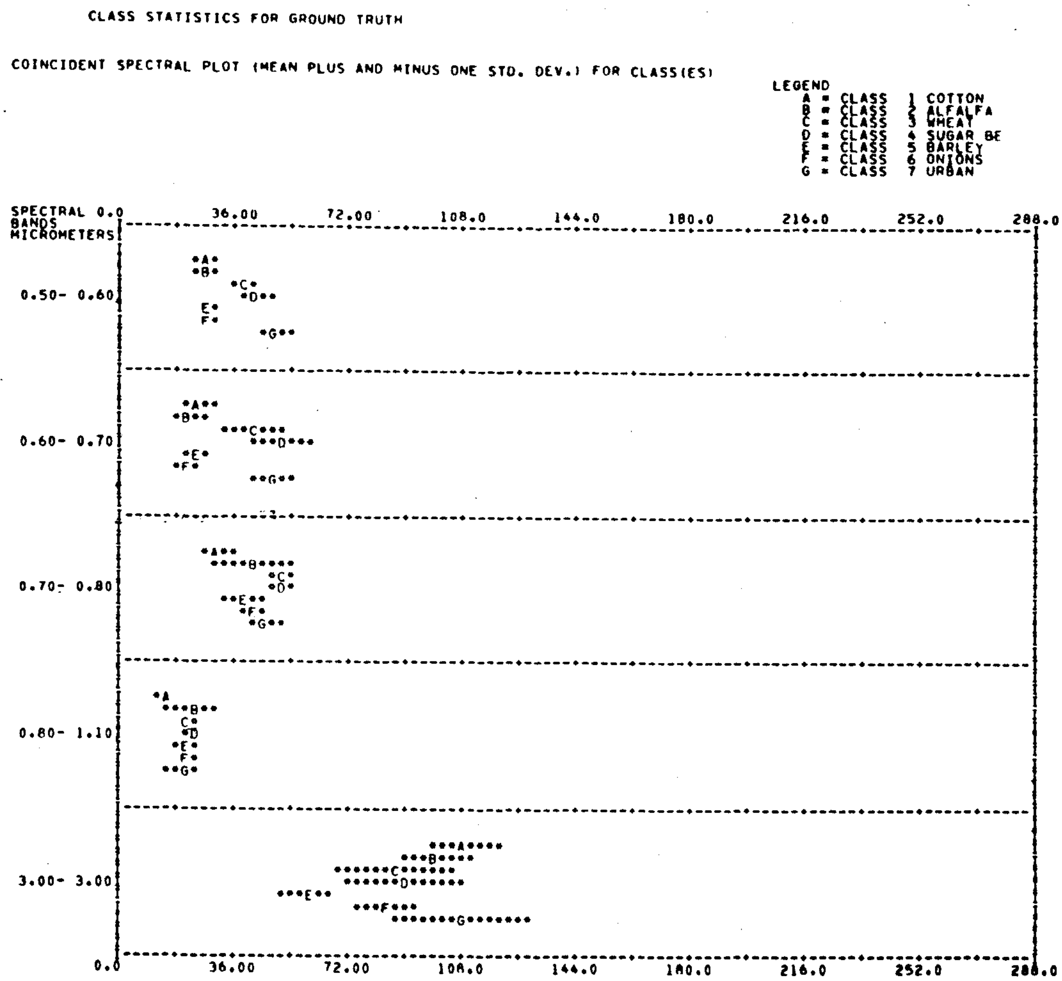


Figure 2C2-2. Spectral Plot for Classes (Phoenix, AZ).

Table 2C2-2. Merged SAR/Landsat Data Set Description.

Data Set No.	Site Identifier	Date of SAR FLIGHT	Landsat Frame/Date	LARS Data Set No.	Number of Lines	No. of Samples/Line	Pixel Size	No. of Channels	Tape No.	File No.
1	Salisbury, Maryland	August 22, 1976	2579-14535 August 23, 1976	76016404	2700	1906	25.4 x 25.4 M	5	3620	1
2	Cambridge, Maryland	August 22, 1976	2579-14535 August 23, 1976	76016413	681	598	25.4 x 25.4 M	7	3692	1
3	Phoenix, Arizona	June 17, 1977	1085-17330* October 16, 1972	72069110*	512	512	25x25M	7	160*	1*

* will change when 1977 Landsat data is received.

Task 3. Assessment of Methods of Acquiring, Analyzing and Reporting Crop Production Statistics

The objective of this study was to describe and document the current methodologies for obtaining, analyzing and reporting crop production statistics in Argentina, Canada, India, the Soviet Union and the United States. Each country uses the same general methodology for each of the major crops within that country. Although this project considered crop statistics in general, major attention was given to wheat statistics methodologies.

Of the five major wheat-producing countries examined, most wheat area estimates are made by subjective or nonprobability methods (Figure 3-1). The United States relies substantially on area frame sampling. Objective methods for determining areas in wheat are used in the other countries to a very limited degree.

Country	Subjective Methods	Objective Methods
Argentina	Inspectors (Interviews with farmers)	Very limited use (Buenos Aires Province only)
Canada	Mail surveys Agricultural census - enumeration every 10 years	Agriculture Enumerative Survey (experimental) Farm Expenditure Survey (initiated in 1977 in prairie provinces)
India	Land revenue officers total enumeration	Investigators (limited area)
Soviet Union	Total enumeration on state and collective farms (97%)	Sample surveys on private lands (3%)
United States	Mail surveys	Trained enumerators (area frame sampling)

Figure 3-1. Summary of methods used to estimate wheat areas

Wheat yield estimates are not readily available on a regular basis to the public in most of the major wheat-producing countries. Where yield estimates are reported, most statistics are derived from the subjective methods (Figure 3-2). Of the five countries examined, the United States relies most on objective yield surveys, and India uses crop cutting surveys.

Country	Subjective Methods	Objective Methods
Argentina	- Biweekly reports of inspectors - Interviews with farmers, grain merchants, harvest crews	None
Canada	- Mail surveys	None
India	- None	Investigators (Crop cutting surveys)
Soviet Union	- No official forecast made	None
United States	- Mail surveys	Trained enumerators (Objective yield surveys)

Figure 3-2. Summary of methods used to estimate wheat yields.

The reporting of wheat statistics varies significantly among the five countries studied. In general, the public reporting on a regular basis of wheat area, predicted yields and production is extremely limited (Figure 3-3). The two extremes are represented by the Soviet Union and the United States. The Soviet Union regularly reports to the public the area planted in wheat as the growing season progresses. However, the only public reporting of yield and production is released as historical data many months after harvest has been completed. The United States issues on a regular basis throughout the growing season public reports on area estimates and predicted yields and production.

Country	Month											
	J	F	M	A	M	J	J	A	S	O	N	D
Argentina												
Area						x	x		x			
Yield	x		x									x
Production	x		x									x
Canada												
Area	x			x		x		x			x	
Yield	x								x	x	x	
Production	x								x	x	x	
India												
Area			x	x							x	
Yield			x	x								
Production												
Soviet Union												
Area				x	x	x	x	x	x	x	x	
Yield												
Production												
United States												
Area	x					x	x	x	x	x	x	x
Yield						x	x	x	x	x	x	x
Production						x	x	x	x	x	x	x

Figure 3-3. Comparison of schedules for reporting wheat statistics by Argentina, Canada, India, USSR and USA.

In order to formulate meaningful summary statements from this study, the authors felt the need to express two assumptions:

- More accurate, timely statistics on current and predicted world wheat area, yield and production will be beneficial to society through
  - stabilization of prices
  - more effective production planning
  - more effective distribution
- Current and projected advances in data acquisition, data analysis and information dissemination technology suggest that a significant improvement can be made during the next decade in a global information system for wheat.

With these assumptions in mind, the following summary statements of weaknesses of the present methodologies suggest the critical need for and feasibility of an improved global information system for wheat:

1. There is no standardized, global system for acquiring, analyzing and reporting wheat production statistics.
2. Among the major wheat-producing countries there is no common rationale for reporting wheat production statistics publicly.
3. Under current methods of reporting, it is not possible to determine quantitatively the statistical reliability of the global estimates of wheat area, yield and production.
4. Current methods of making wheat production estimates in several major wheat-producing countries are subject to gross error.
5. The U.S. Department of Agriculture relies substantially on objective yield data to predict wheat production at the state level; to predict national production, subjective adjustements are made in the data prior to release of the periodic crop reports.

Task 4. Computer Processing Support

Since 1972, NASA/JSC's Earth Observations Division has supported and utilized a computer terminal connected by a communications link to the Purdue/LARS computer facility. Between 1972 and 1977, this terminal served as the nucleus of a technology transfer effort, blending the terminal hardware and the LARSYS analysis software with specially designed training materials and user-oriented operating procedures.

During late 1977, the decision was made to upgrade the LARS software/hardware facilities at JSC to provide increased data processing capabilities and terminal support in order to:

* Improve the capability for two-way techniques exchange,
* Relieve RT&E computational constraints by augmenting on-site computing as new processing capabilities were implemented on the LARS computer,
* Reduce total costs by improving the efficiency of computer operations,
* Maintain valuable computational facilities supporting LARS and JSC research and development, and
* Increase for both organizations access to useful resources (data, technology, processing systems, hardware, etc.).

This decision was viewed as a first step in the development of a shared research data processing system for the research community supporting the Earth Observations Division. It is anticipated that a data processing system, shared by the key, geographically dispersed, research groups supporting EOD would potentially provide:

* Reduced costs through centralization of hardware, software and certain computer support personnel,
* Immediate access to software (packages, new analysis techniques, utilities, etc.) and data bases by the entire research community,
* A modular, baseline analysis system as a framework for new techniques development, alternative techniques comparisons, and a standard for techniques delivery,

* The opportunity to effect rapid technology exchange, and
* The opportunity for construction of a single program of computer-aided training materials, useful at all sites.

1. Accomplishments

Efforts during the past year have resulted in accomplishments in the areas of:

Remote User Support

New Hardware Capabilities

New Software Capabilities

Data Availability

Information Exchange

Expanded User Community

a. Remote User Support

Support of remote users includes maintaining the necessary hardware and software at the LARS computer facility to support the IBM 2780 and Data 100 remote job entry stations, providing computer services, and providing system consultation. During the first half of the year, the magnitude of this task increased dramatically. Figure 4-1 indicates the number of JSC ID's on the system by month during 1978, a three-fold increase occurred over the period. For each ID on the system, disk storage space, virtual machine configuration, library tapes and other personnel computing resources are assigned and changed in response to user requests.

Figure 4-2 presents the total usage of computer by JSC personnel and by all users. It should be noted that the decision to limit JSC usage to 75 hours per month was made by managers at JSC during the peak month of August. Computer usage for the year averaged 58 hours of CPU time per month as compared to 10 hours of CPU per month during the previous contract year.

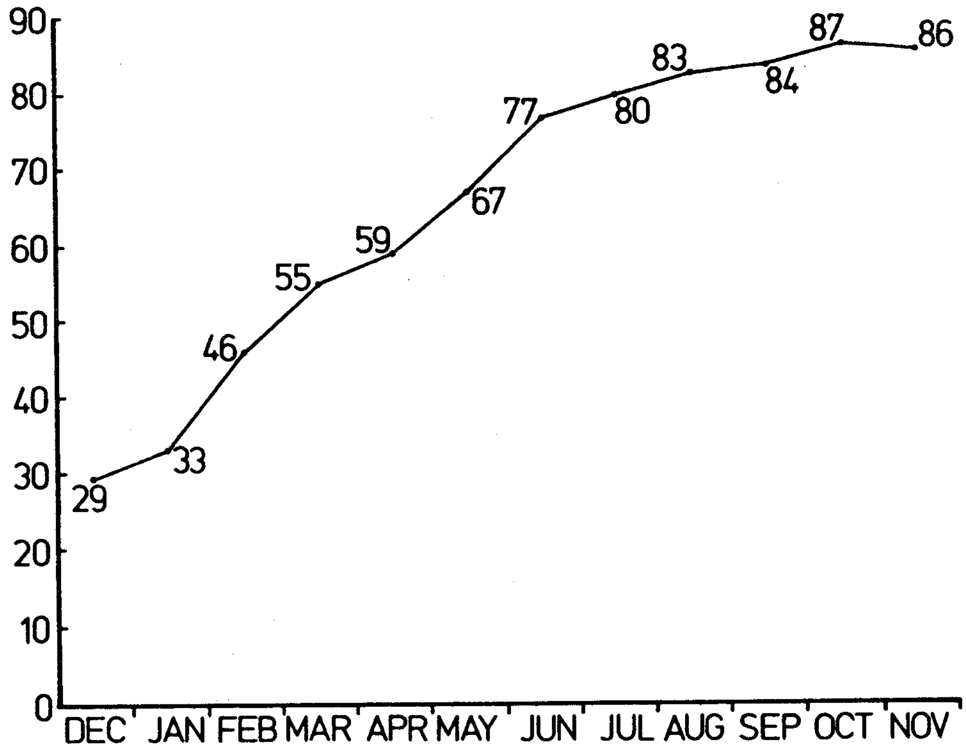


Figure 4-1. JSC user ID's on the LARS system during Contract Year 1978.

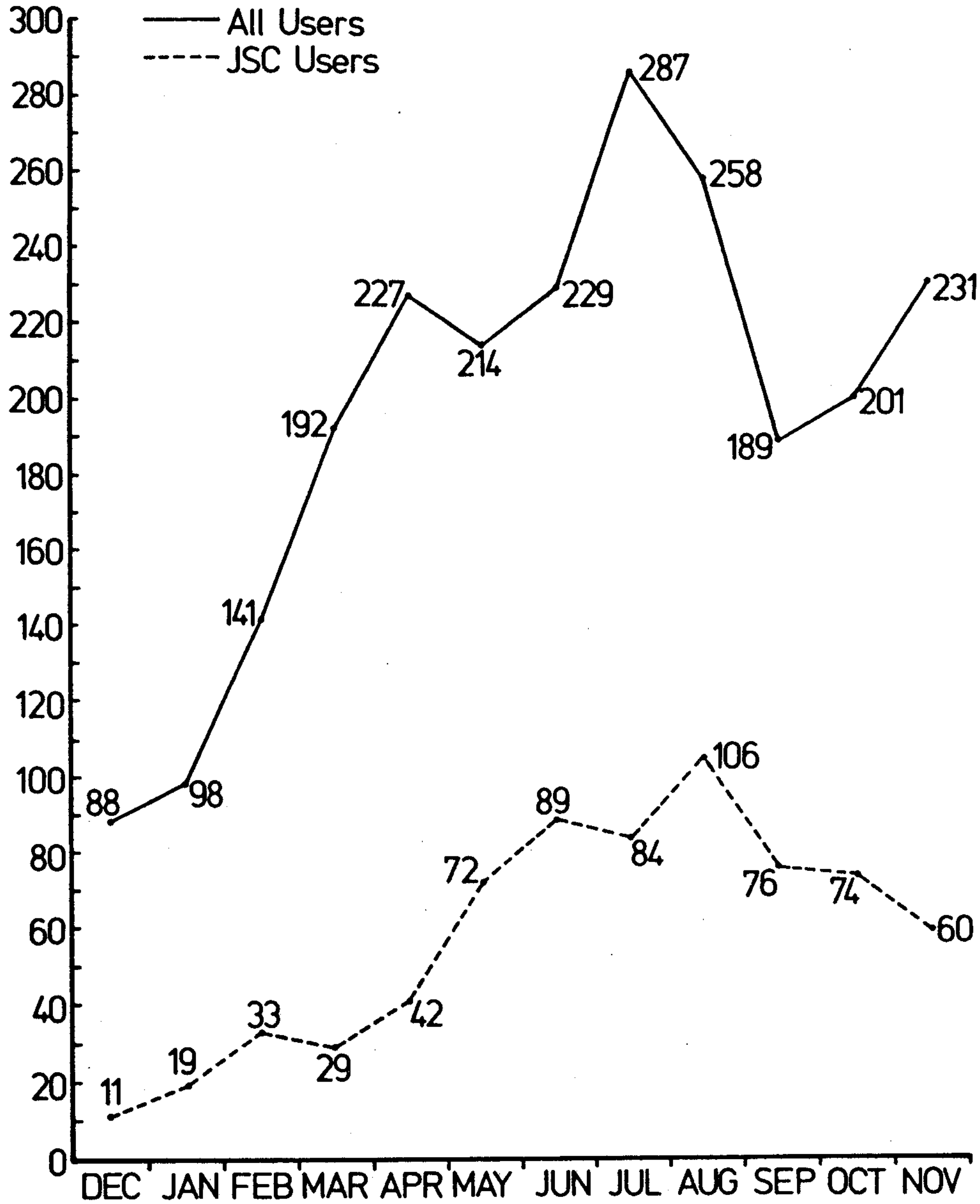


Figure 4-2. CPU hours used per month during Contract Year 1978.

With the three-fold increase in users came a large increase in demand for consulting services. To satisfy the need for consultants, a consulting team was established at LARS and publicized at JSC. Consulting activities consume a substantial portion of the personnel resources devoted to the project. In addition to consulting activities by telephone, bimonthly consulting visits are made to JSC. These visits typically are a week in duration and frequently include presentations and demonstrations of various aspects of the computer system as well as consulting with users on a one-to-one basis. The increase in computer users and their usage to a large extent reflects the success of the consulting activity.

b. New Hardware Capabilities

Several significant hardware improvements at JSC helped to support expanded usage of the LARS system during 1978:

* Installation of a second remote job entry station
* Additional keyboard terminal access
* New tape transfer capability

1) Installation of a second Remote Job Entry Station

During 1978 a Data 100 remote job entry station accessing the Purdue/LARS system was installed. this terminal utilized faster protocol, faster I/O devices, and a faster communications link than the older IBM 2780 terminal, greatly enhancing the amount of I/O available to JSC users.

2) Additional Keyboard Terminals

During June the number of terminal ports available to JSC users was increased from 3 to 8. Four of these additional ports to the Purdue facility support dial-up terminal access. This allows individual users to access the LARS system from their offices and also provides users external to NASA and Purdue (e.g., Texas A&M University) with a means of access to the system.

3) New Tape Transfer Capability

Prior to this contract year, data stored on magnetic tape, such as Landsat data for LACIE segments, had to be shipped to LARS via the U.S. Postal System. A JSC user wishing access to a tape located at JSC, but to be analyzed on the LARS system, had to wait from three weeks to a month for that tape to become available at Purdue. In order to reduce this wait time, the Data 100 remote job entry station installed at JSC included a tape drive. Software to support tape transfers between the two sites has been designed, implemented, tested and place in operational use.

The computational hardware available at Purdue/LARS is summarized in Figure 4-3.

A survey of the remote terminal hardware market was conducted to identify a potential replacement for the aging IBM 2780 terminal hardware resident at JSC. The survey was conducted jointly by personnel at JSC and LARS with LARS providing an estimate of the impact of each of the alternatives on the Purdue system. A plan for the replacement for the 2780 has been established by personnel at JSC and is being supported by Purdue.

c. New Software Capabilities

In preparation for using the Purdue system for research, test, evaluation, and exchange of newly developed analysis algorithms and procedures, the LARSYS Procedure 1 analysis system was implemented in a modular form, to serve as a baseline. LARSYS Procedure 2 software is currently being developed on the LARS system in a form compatible with the Procedure 1 software.

In order to support the particular batch machine requirements of JSC users, three CMS/370 batch machines were designed and implemented.

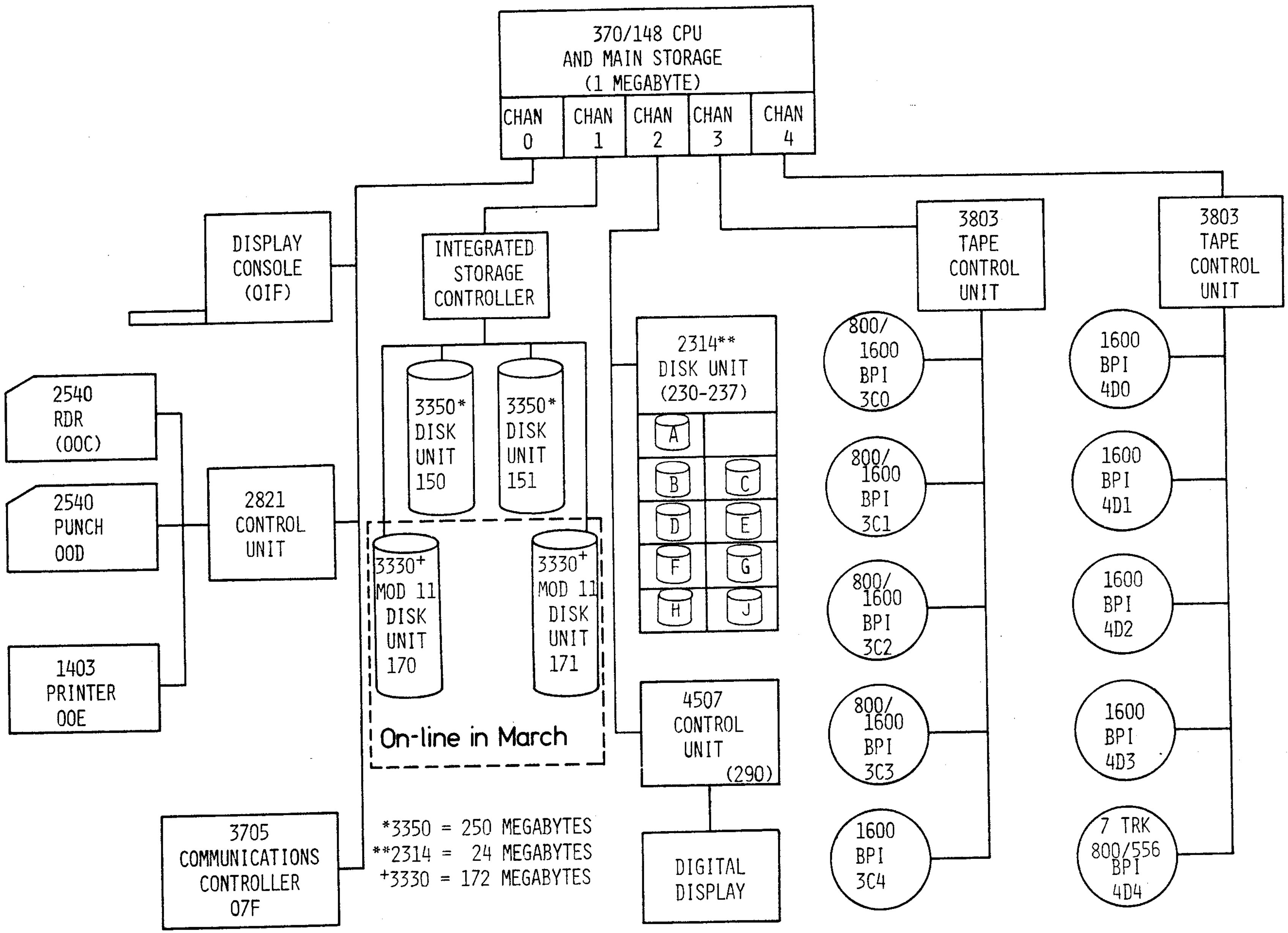


Figure 4.3 Purdue/LARS Computer Hardware.

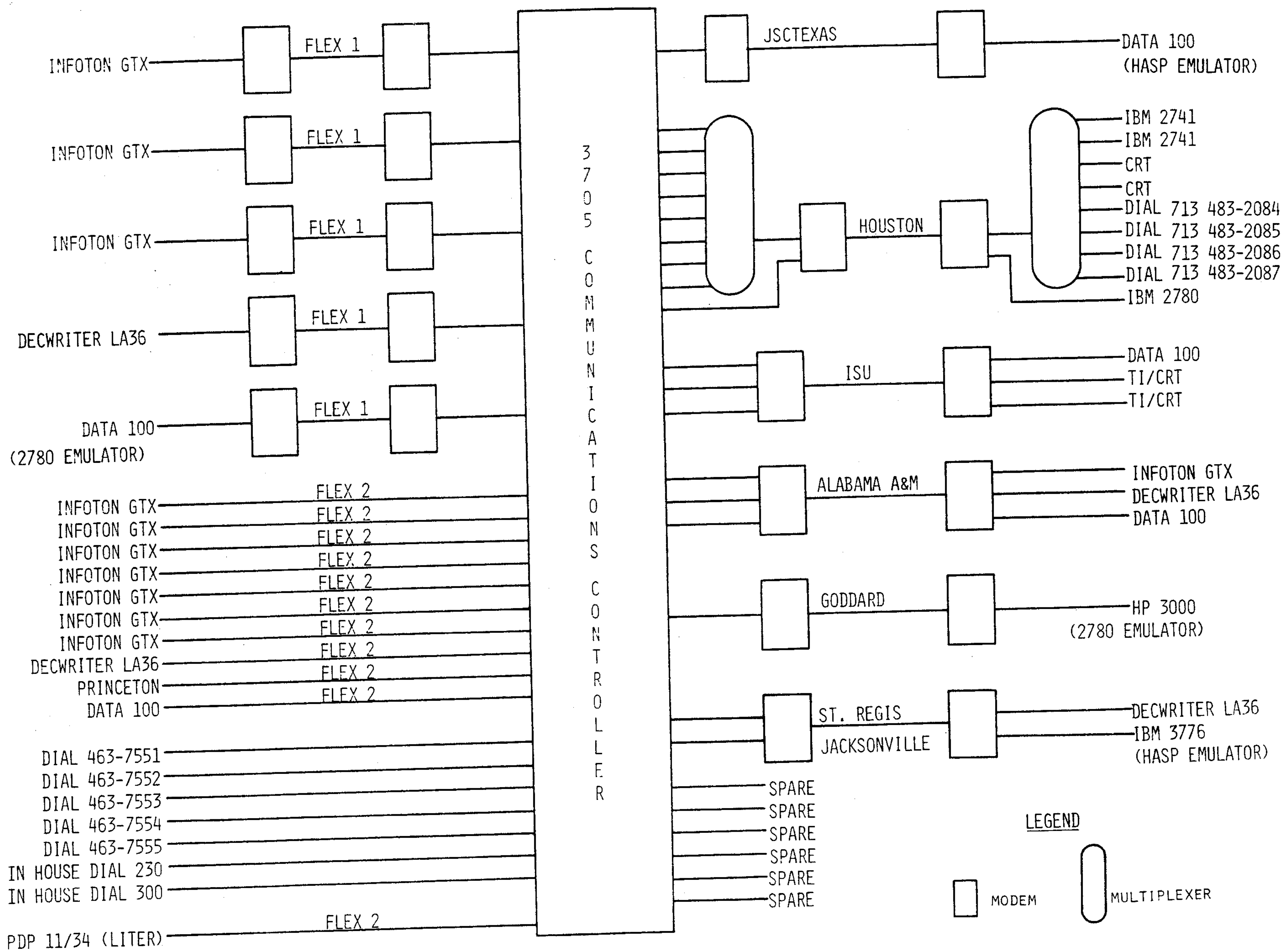


Figure 4-3. (Continued)

A number of statistical analysis capabilities were necessary for evaluation of results of demonstrations and experiments related to the development of remote sensing technology. For a number of years, LARS has had the Statistical Package for the Social Sciences (SPSS), and the Biomedical Statistics Package (BMD) on the system in an unsupported mode. JSC users have requested that acquisition of the International Mathematical and Statistical Library (IMSL) and Statistical Analysis System (SAS) be investigated. IMSL has been acquired and made available on the system. There are plans to acquire SAS as soon as a CMS/370 compatible version becomes available (circa May 1979) and to establish a rate in order to provide SPSS in a "supported" mode.

Figure 4-4 summarizes the software available on the Purdue system.

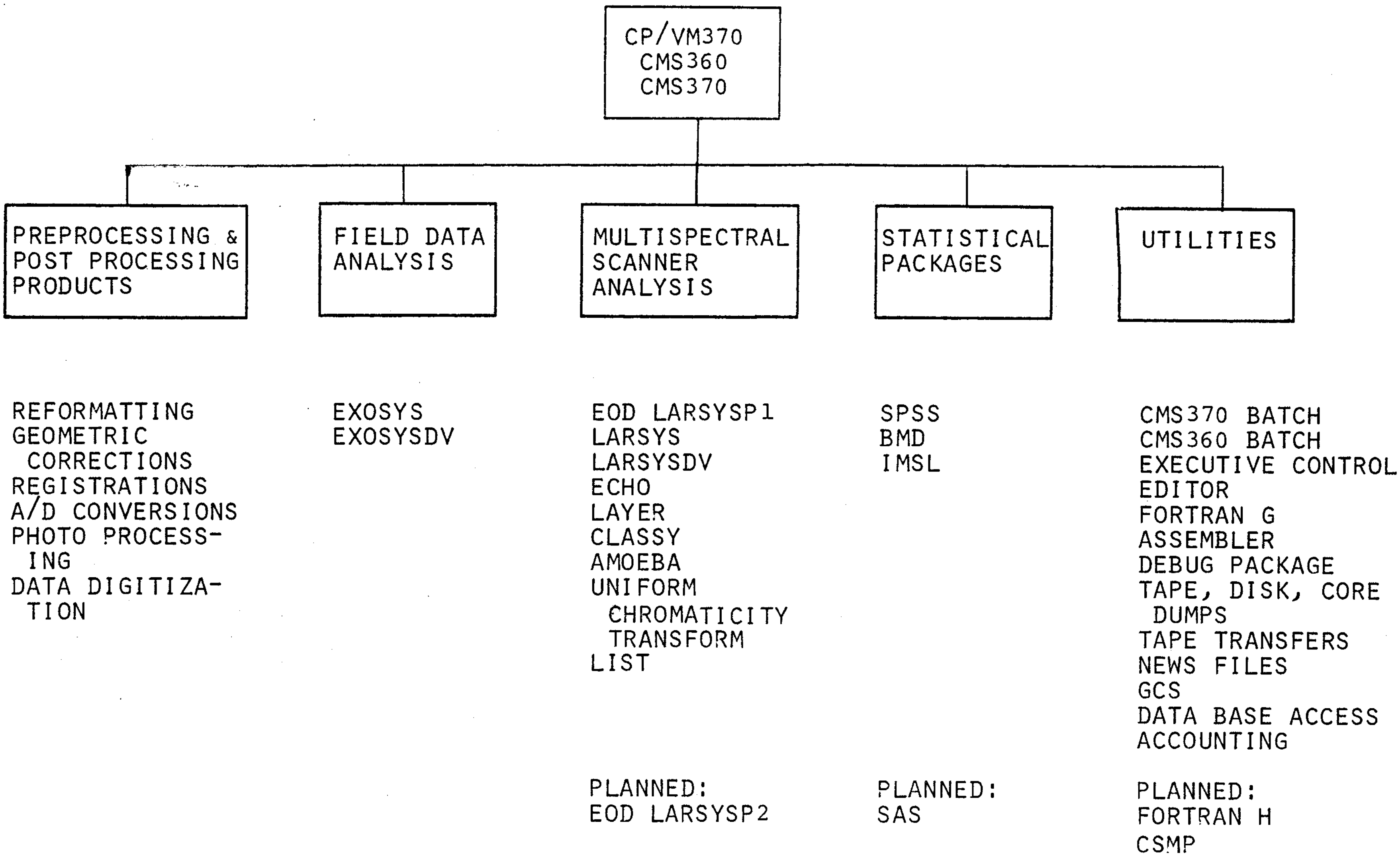
d. Expanded Accessability

The facilities at LARS and at JSC for dial-up access to the Purdue/LARS system have made possible access of that system by a large portion of the research community. During 1978 individuals from Purdue/LARS, NASA/JSC, Lockheed Electronics Corporation, IBM, Mitre, Texas A&M University, and Ft. Lewis College, Colorado have made use of the system from terminals located at JSC or Purdue, or from dial-up terminals at their own institutions. In addition, documentation on how to use the LARS system has been distributed to the University of California at Berkeley and Environmental Research Institute of Michigan (ERIM). A short course on how to use the LARS system is to be presented during December at ERIM by Lockheed personnel.

e. Information Exchange

To support transfer of capabilities and techniques which have been placed on the Purdue computer for use by members of the research community, several information exchange activities have taken place. During February a CP/CMS370 short course was presented by LARS personnel in order to inform new JSC users of the characteristics and capabilities available on the LARS

Figure 4-4. SOFTWARE AVAILABLE ON THE LARS COMPUTER



system. During March a short course was presented at LARS on the characteristics, syntax, and use of the LARSYS Procedure 1 analysis system which was being implemented on the LARS machine. During October a seminar was presented by LARS personnel to members of the research community at Houston interested in the use of the field data analysis system (EXOSYSDV). Additional information exchange activities are planned for 1979 including a second CP/CMS 370 short course. Such activities are a significant element necessary for conducting a coordinated program involving a number of geographically dispersed research groups.

f. Expanded User Community

As use of the Purdue system by the research community has expanded, the need to make information about new capabilities and developments has also grown. To help meet that need, new users have been added to the distribution list of SCANLINES, LARS' monthly newsletter. In addition, a special information system, SRTNEWS, has been made available over the terminal.

2. Data Availability

To support research, development, test, and evaluation of analysis techniques and procedures, a number of substantial data bases are desirable. During 1978 the LACIE Phase I, II and III digital (Landsat) data bases were installed at LARS, joining the LACIE field measurements data base which has traditionally been housed at Purdue. A data base directory, data base search capability, and tape indexing capability are under development to aid users of these data bases.

It is anticipated that LACIE Transition Year and multicrop data will be shipped to LARS during 1979. Digital ground truth is also expected during the coming year.

3. Recommendations

The remote terminal concept has proven to be successful for the transfer of proven image analysis techniques. During 1978 much has been done to establish the raw materials for a data processing facility to be shared by the Earth Observations Divison's research community for research, development, test, evaluation and interchange of data processing techniques and procedures.

Some of the potential benefits of the shared system are being realized, e.g., by researchers at LARS who are accessing and using the JSC LARSYS Procedure 1 analysis system and at JSC by researchers utilizing the field data analysis package (EXOSYS).

It is recommended that improved means of access to the shared system be investigated and established where warranted for some of the key research organizations which do not currently have remote job entry stations accessing Purdue. More powerful terminal configurations should also be investigated. In particular, intelligent remote terminals and remote interactive color devices should be considered.

Uniform documentation and algorithm delivery standards should be established and implemented to aid in the exchange of promising analysis techniques and procedures. A detailed, computer-aided training program should be formulated by Purdue/LARS and NASA/JSC personnel as an introduction to use of the shared system.

LARS currently rents its computational hardware on a yearly basis because funding for that hardware cannot be guaranteed for a longer period. If funding could be guaranteed for four years or longer, alternatives for more cost-effective computational hardware would exist. Should funding be sufficient, it might be possible to purchase a machine with capabilities significantly greater than those of the IBM 370/148 for less than the four year rental of the 148.

To help accomplish the above goals, it is suggested that a steering committee be established to help plan and critique the evolution of a shared SR&T data processing system.