T-1314/4
MA-129TA
LARS Contract Report 112677

Final Report

Agricultural Scene Understanding

by M. E. Bauer
L. F. Silva
R. M. Hoffer
M. F. Baumgardner

November 1977

Principal Investigator
D. A. Landgrebe

Prepared for
National Aeronautics and Space Administration
Johnson Space Center
Earth Observation Division
Houston, Texas 77058
Contract No. NAS9-14970, Task I
Technical Monitor: J. D. Erickson/SF3

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

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
112677		

4. Title and Subtitle	5. Report Date
Agricultural Scene Understanding	November 1977
	6. Performing Organization Code
7. Author(s) Marvin E. Bauer, LeRoy Silva, Roger M. Hoffer, and Marion F. Baumgardner	8. Performing Organization Report No. LARS Contract Report 112677
	10. Work Unit No.
9. Performing Organization Name and Address Laboratory for Applications of Remote Sensing, Purdue University, 1220 Potter Drive, West Lafayette, Indiana 47906	11. Contract or Grant No. NAS9-14970
	13. Type of Report and Period Covered T-1314/4 Final Report
12. Sponsoring Agency Name and Address J.D. Erickson/SF3, NASA/Johnson Space Center, Houston, Texas 77058	14. Sponsoring Agency Code

15. Supplementary Notes

D. A. Landgrebe was LARS principal investigator.

16. Abstract

Results of four investigations, all related to agricultural remote sensing are described. The four tasks are: (A) LACIE Field Measurements, (B) Thermal Band Canopy Modeling, (C) Forestry Applications Project, and (D) Soil Classification and Survey.

A. The LACIE Field Measurements project report describes the rationale for the experiment, the data acquisition, processing, and storage/retrieval by LARS. Results of the sensor correlation and data verification studies are discussed, along with the rationale and procedures for calibration of reflectance measurements. Analytical results of initial analyses relating spectral and agronomic measurements are described. The report concludes with recommendations for future field measurements investigations.

B. The Thermal Band Canopy Modeling results demonstrate the relationship between geometric parameters of wheat canopies, environmental variables, and radiance temperature.

C. The Forestry Application Project report describes investigations of (1) the acceptability of Landsat area estimates as inputs to forest inventory, (2) definition of an efficient and cost effective method of developing optimal Landsat training statistics for mapping forest cover, and (3) a comparison of five different classification techniques in terms of cost, accuracy, and output products.

D. The Soil Classification and Survey report describes the results of (1) field experiments relating spectral reflectance measurements to dark and light soils at two surface moisture levels and two amounts of surface residue, and (2) classification for soil survey of multiple dates of Landsat data covering the same scene.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Agricultural Remote Sensing, Earth Resources, LACIE, Field Measurements, Thermal Measurements, Soil Classification, Forestry, Landsat	

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

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

Table of Contents

List of Figures

List of Tables

1. AGRICULTURAL SCENE UNDERSTANDING

Task 1, Agricultural Scene Understanding, consists of four sub-tasks which have the common objective of increasing our understanding of the energy-matter interactions in agricultural scenes. The tasks are: (A) LACIE Field Measurements, (B) Thermal Band Canopy Modeling, (C) Forestry Applications Project, and (D) Soil Classification and Survey.

A. LACIE Field Measurements

I. Introduction

Major advancements have been made in recent years in the capability to acquire, process, and interpret remotely sensed multispectral measurements of the energy reflected and emitted from crops, soils, and other earth surface features. With the initiation of experiments such as the Large Area Crop Inventory Experiment (LACIE), the technology is moving rapidly toward operational applications. There is, however, a continuing need for quantitative studies of the multispectral characteristics of crops and soils if further advancements in the technology are to be made. In the past, many such studies were made in the laboratory because of a lack of instrumentation suitable for field studies. However, the applicability of such studies is generally limited. The development of sensor systems capable of collecting high quality spectral measurements under field conditions has made it possible to pursue investigations which would not have been possible a few years ago.

A major effort was initiated in the fall of 1974 by NASA/JSC with the cooperation of the USDA to acquire fully annotated and calibrated multitemporal sets of spectral measurements and supporting agronomic and meteorological data. Spectral, agronomic, and meteorological measurements have been made on three LACIE test sites in Kansas, South Dakota, and North Dakota for three years. The remote sensing measurements include data acquired by three truck-mounted spectrometers, a helicopter-borne spectrometer, two airborne multispectral scanners, and the Landsat-1 and -2 multispectral scanners. These data are supplemented by an extensive set of agronomic and meteorological data acquired during each remote sensing data collection mission. The data collection program is illustrated and summarized in Figure A-1.

The LACIE Field Measurements data form one of the most complete and best documented data sets ever acquired for remote sensing research. Thus, they are well-suited to serve as a data base for research to (1) quantitatively determine the relationship of spectral to agronomic characteristics of crops, (2) define future sensor systems, and (3) develop advanced data analysis techniques. The data base is undoubtedly the largest of its type now available for research purposes. It is unique in its comprehensiveness in terms of sensors and missions over the same sites throughout the growing season. The calibration of all multispectral data to a common standard is also unique. Finally, the kind and quantity of supporting agronomic and meteorological data is impressive compared to most remote sensing experiments.

During the past 18 months, several key milestones were reached. The first of these was completion of data processing for the 1974-75 and 1975-76 data. These data, including over 75,000 individual spectra, are available from the data library. Significant improvements were made in the second and third year data acquisition and processing procedures; and, additional data evaluation and verification steps were implemented. Lastly, data were distributed to researchers at five different institutions.

At the beginning of the project, Purdue/LARS was requested to provide the technical leadership and coordination for the project, as well as assume major responsibilities for the acquisition, processing, distribution, and analysis of the data. This report summarizes our activities and results of the past 18 months (June 1976 through November 1977) since the final report for the previous contract was prepared.

The second section of the report summarizes our activities in acquiring, processing, and distributing data. The third section discusses calibration, verification, and correlation of the spectral measurements--a fundamental aspect of the entire effort. The fourth section describes the results of analyses of the data performed thus far. The concluding section presents recommendations for future field measurements work.

II. Data Acquisition, Processing, and Distribution

This section describes the field measurements data acquisition, processing, and library functions performed by LARS during this contract.

a. Data Acquisition

Data acquisition activities during the past 18 months have included two summers of data collection in Williams County, North Dakota. The sites of data collection were the North Dakota State University Agricultural Experiment Station near Williston, the calibration location at the intensive test site (ITS #1960), and a commercial wheat field used for acquiring canopy modeling data. The activities were a continuation of those begun during 1975.

At the agricultural experiment station, spectral bidirectional reflectance factor, radiant temperature, agronomic parameters, and meteorological conditions were measured for experiment designs containing 60 plots in 1976 and 70 plots in 1977. The treatments represented major agronomic factors affecting the growth, development, and yield of spring wheat and other small grains.

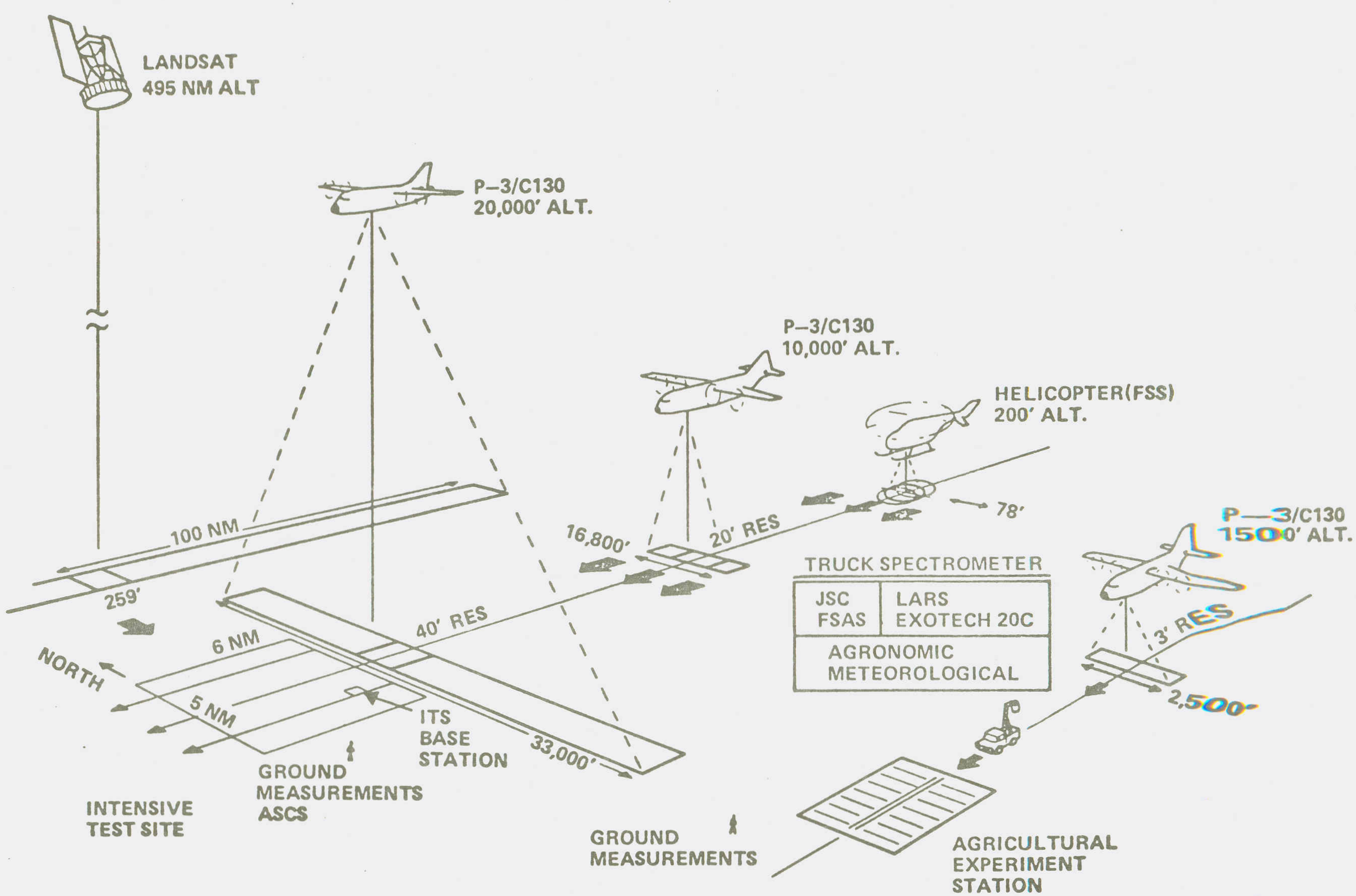


Figure A-1. Illustration of LACIE field measurements data acquisition.

At the calibration location in the intensive test site, spectral bidirectional reflectance factor data were collected over gray panels used to calibrate the FSS (helicopter-borne spectrometer) and aircraft multispectral scanner (MSS) data. Also during 1977, additional measurements were collected in parallel with the FSS and the NASA truck spectrometer system (FSAS) to determine how well the data from the three instruments were correlated.

At the modeling field, Landsat band reflectance factor measurements, photographic, and agronomic measurements were collected. The data is to be used to help develop and test canopy models. In addition, during 1976, spectral bidirectional reflectance factor and radiant temperature were measured at view angles distributed over the complete hemisphere for several sun angles during the day.

Additional details describing the data acquisition are provided in the LACIE Field Measurement Project Plans [1, 2].

1. 1976 Data Collection

Agriculture Experiment Station. Data were collected over 60 spring wheat and other small grain plots. The layout of the experimental plots is illustrated in Figure A-2. There are three overall experiments included in the plot design - small grain, planting date, and seeding rate. Some data, when time allowed, were collected for a fourth experiment - other crops. The other crops were in separated locations at the experiment station. The dates of data collection over the plots and corresponding wheat growth stages are illustrated in Figure A-3.

The spectral bidirectional reflectance factor data were collected by the Purdue/LARS Exotech 20C field spectrometer system (Figure A-6). Problems were experienced with the portable motor-generator set which supplied AC power. Frequency variations caused the data signals to be wow modulated. An algorithm was implemented which corrected the data signals absolutely, i.e., no assumptions were made. The correction used to eliminate the tape recorder wow worked very well. The results indicate that the error due to the wow modulation was reduced to practically zero. A report has been prepared which describes the algorithms and the results of the correction.

Meteorological measurements, including air temperature, barometric pressure, relative humidity, wind speed and direction, and total irradiance, were recorded continuously during each day that spectral data were collected (Figure A-6).

Detailed agronomic measurements were made for each plot at the time of spectral data collections. These measurements included: growth stage; canopy height; percent ground cover; percent green, yellow, and brown leaves; stems per meter row; leaf area index; fresh biomass; dry biomass of leaves, stems, and heads; and surface soil moisture and condition. Grain yield was measured at harvest. The agronomic data are supplemented with vertical and oblique views of each plot at the time of data collection.

1976 Williston, North Dakota Agriculture Experiment Station

Remote Sensing Experiments

Treatment Descriptions

Soil Moisture	Planting Date	Variety	Nitrogen	Seeding Rate
M_1 Wheat in 1975	D_1 May 6	V_1 Waldron	N_0 None	R_1 30 lbs/acre
M_2 Fallow in 1975	D_2 May 17	V_2 Olaf (Semi Dwarf)	N_1 30 lbs/acre	R_2 60 lbs/acre
				R_3 90 lbs/acre

Figure A -2. Experimental design and treatment description of remote sensing experiments at Williston Agriculture Experiment Station, 1976.

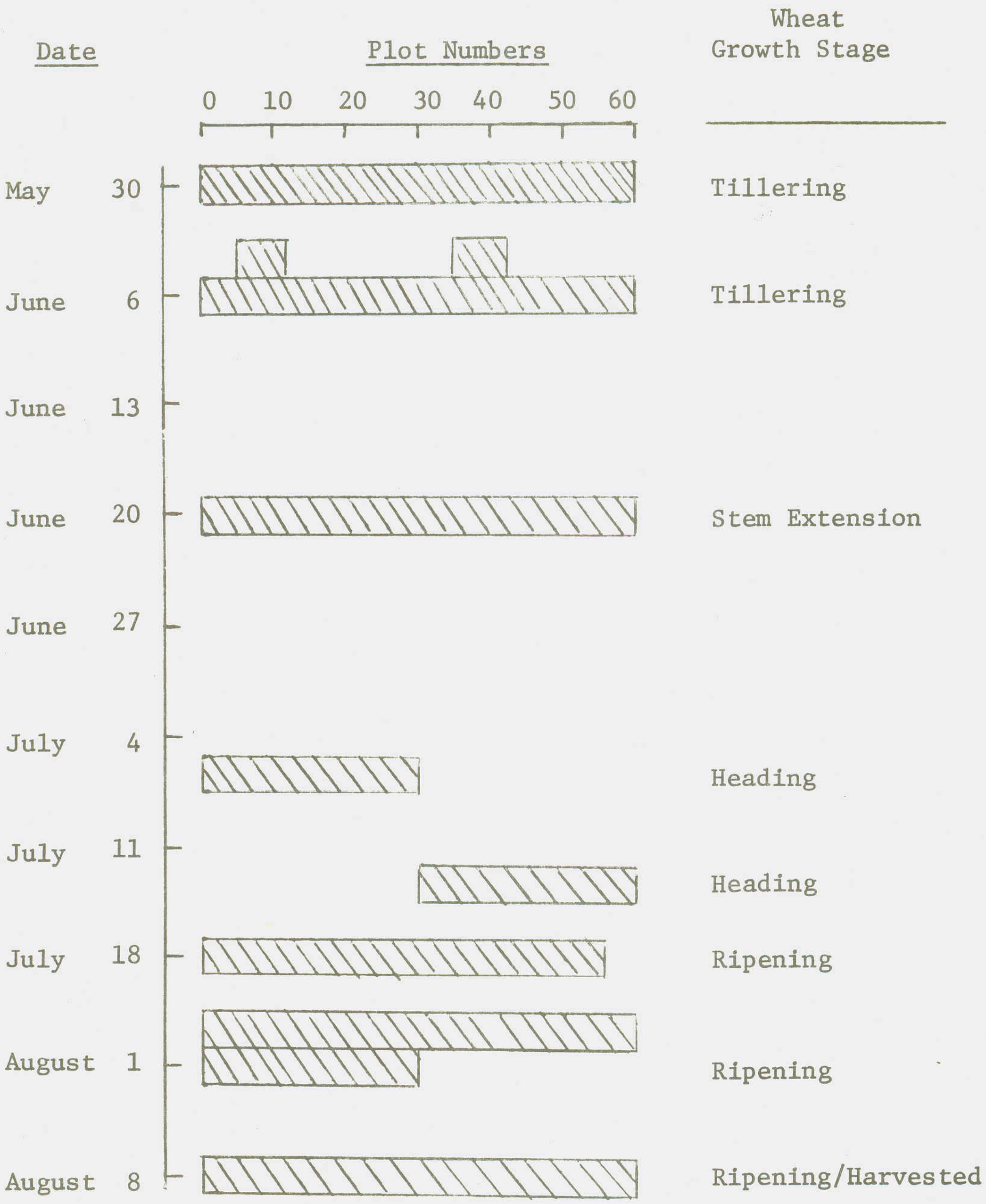


Figure A-3. Summary of data collection dates, plots measured, and spring wheat growth stages at the Williston Agriculture Experiment Station during 1976.

Calibration Measurements. Spectral bidirectional reflectance factor measurements of gray panels were collected at the intensive test site three times (May 28, June 25, and July 28) during the summer in support of the FSS and aircraft MSS data calibration. Measurements were collected over the brightest gray panel (panel 1) having a nominal reflectance of 55% which is used to calibrate the FSS. Measurements were also collected over the four darker gray panels which along with the bright panel can be used to calibrate the aircraft scanner data. These measurements have also been used for instrument verification and instrument correlation of the FSS, FSAS, and Exotech 20C.

Canopy Modeling Measurements. Canopy modeling data including Landsat band reflectance factor measurements, profile and angular photographs, and agronomic measurements were collected four times during the summer (June 19 and 21, July 17 and 31). The data collected and procedures used were formalized in conjunction with Colorado State University and the Environmental Research Institute of Michigan. The data were collected over four separate locations within the modeling field to sample the field.

The Landsat band radiometer data were collected six to eight different times during the day. The photographic data included angular photos at 10° intervals from 0° to 50° zenith angles both with the row and across the row. Photographs were also collected illustrating the profile of individual wheat plants. The agronomic data included plant counts, plant height, leaf area index, biomass, plant condition, head area, and stem area.

An experiment was also designed to collect angular spectrometer data. Data were collected for the experiment only when time, resources, and weather conditions were adequate. In other words, the agriculture experiment station plots were the first priority. Angular measurements were collected three times during the summer (June 21, July 17 and 31). The data set consists of spectral measurements over the range of 0.4 to 2.4 μm made at approximately hour intervals at five zenith and eight azimuth angles.

2. 1977 Data Collection

Correlation Experiments. Data were collected in conjunction with the NASA truck interferometer system (FSAS) and the NASA helicopter-borne spectrometer (FSS) to determine how well the data from the separate systems compare. The Purdue/LARS Exotech 20C spectrometer system was driven to Finney County, Kansas during mid-May. Poor weather conditions postponed the experiment until May 22. Data were collected with the FSAS on May 22, 23 and 24. The FSS didn't collect data with the trucks since it was in South Dakota during that week.

The experiment, designed in cooperation with personnel from NASA/JSC, included measurements of the five gray panels, the new light gray panel used to calibrate the FSS, a green color panel and measurements to determine the extent of reflective interactions.

Unfortunately, the data collected by the FSAS system was later lost. To complete the experiment, the FSAS was driven to Williams County, North Dakota during mid-July where the Purdue/LARS field spectrometer system was stationed. A complete correlation data set was collected by the truck-mounted systems on July 13 (see Figure A-7). Data for the correlation studies were collected with the FSS on July 15 and August 4. These data along with the Purdue/LARS Exotech 20C data collected in late May in Finney County, Kansas will be used to verify the correlation of the different instrument systems and determine if any problems exist.

Agriculture Experiment Station. Data were collected over 70 spring wheat and other small grain plots using the Exotech 20C field system and an Exotech 100 (Landsat band radiometer) field system. A new experimental design and layout of treatments was used this year which made collection of data more efficient and will improve the statistical soundness of the data (see Figure A-4). The key aspect is that, for example, if only a portion (16 or 32) of the plots can be measured in a day, they are a complete statistical unit, i.e., replicated treatments. The small grains treatments were expanded and given increased emphasis in the new design.

Six sets of data were collected over the plots (see Figure A-5). Excessive cloud cover during the first half of June caused a three week break in data collection. The Exotech 20C field system (see Figure A-6) collected high wavelength resolution data over as many plots as possible once during the day. The Exotech 100 field system (see Figure A-6) was built during early June. The system was designed to collect lower spectral resolution data over all plots several times during the day to study diurnal reflectance changes. The system consists of a boom and platform mounted on the top of a van, the Exotech 100 instrument, a hard copy data logger, and two operators. One operator, on top of the van, levels the instrument and verifies the target location. The other operator records the data and drives the van. During operation the Exotech 100 radiometer is positioned four meters above the ground and 3.5 meters away from the van. The system collected a set of data over all 70 plots in an hour. The measurements were repeated every 90 minutes. The system performed very well and it is recommended that data collection systems of this type be used in the future to increase the number of treatments and/or replications which can be measured compared to high spectral resolution truck-mounted systems.

The same agronomic and meteorological measurements were made in 1977 as in 1976.

Calibration Measurements. Spectral bidirectional reflectance factor measurements of the canvas calibration panels were collected at the intensive test site on June 29, July 17, and August 4 in support of the FSS and aircraft MSS data calibration. Measurements were collected using the Indoor Exotech 20C [3] on September 14 at Purdue/LARS. With the Indoor Exotech 20C, the spectral bidirectional reflectance factor of the helicopter calibration panel (gray panel 1) was measured by comparing the gray panel directly with pressed barium sulfate as opposed to an intermediate standard of painted

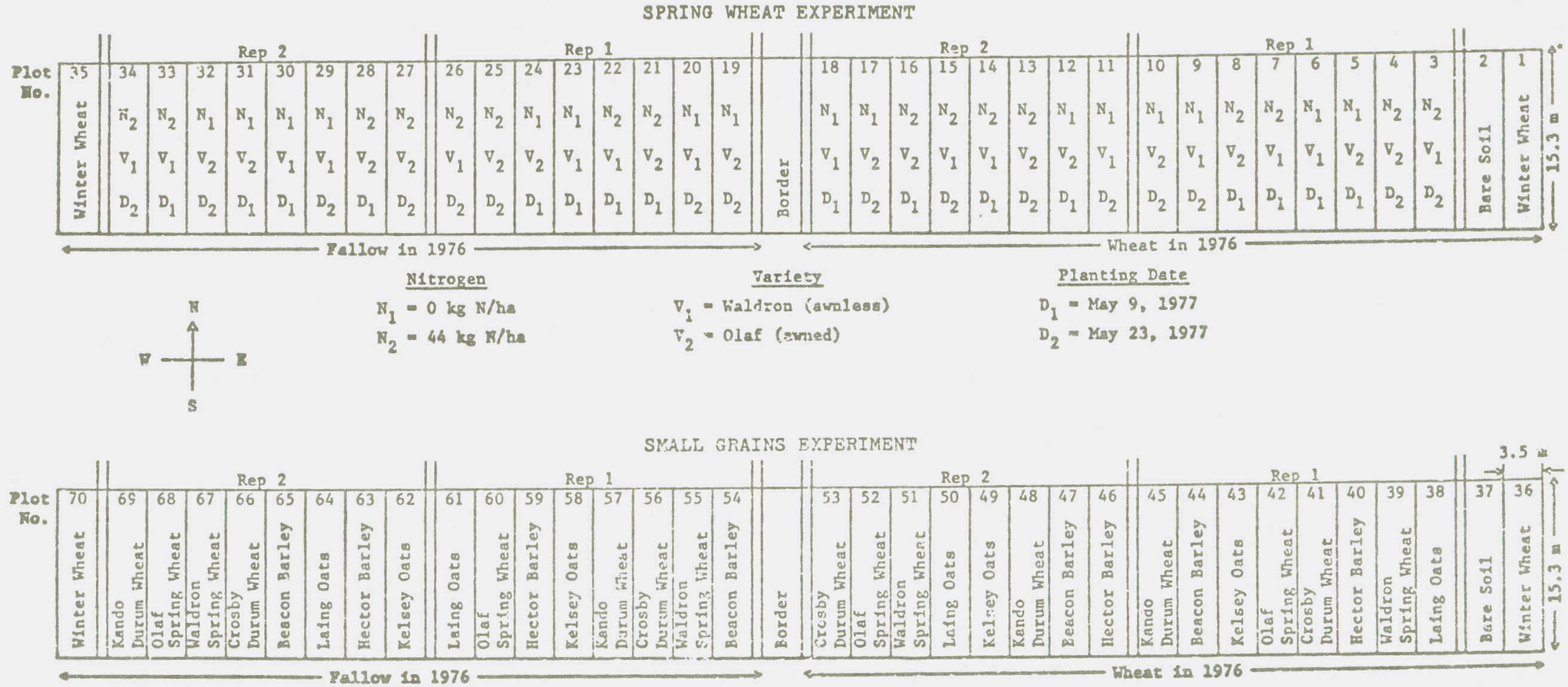


Figure A-4. Spring wheat and small grains experiments at Williston Agriculture Experiment Station, 1977.

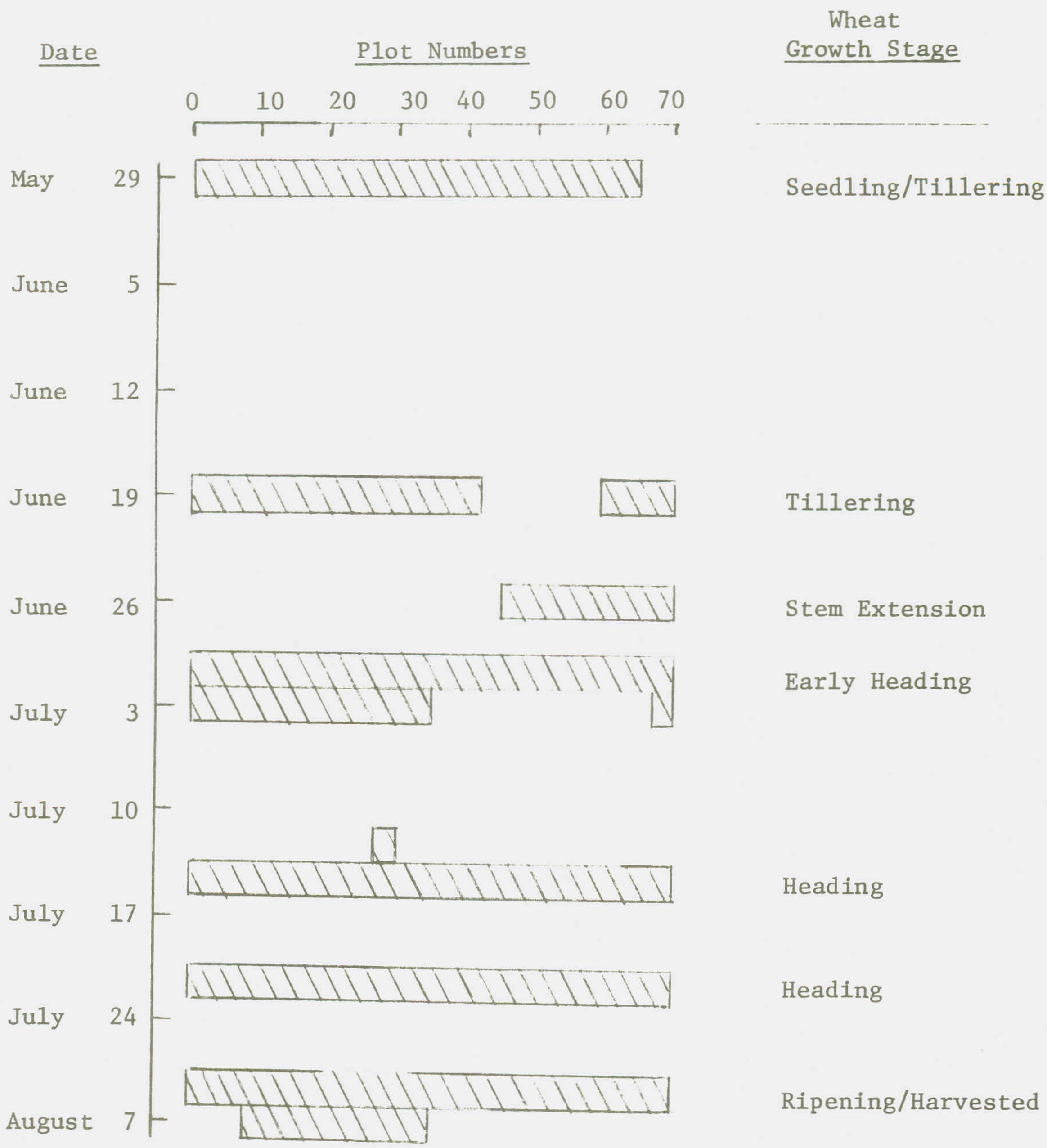
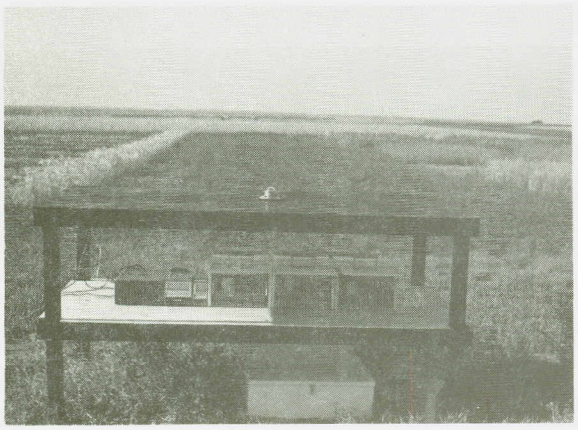


Figure A-5. Summary of data collection dates, plots measured, and spring wheat growth stages at the Williston Agriculture Experimental Station during 1977.

Purdue/LARS Exotech 20C Field System.

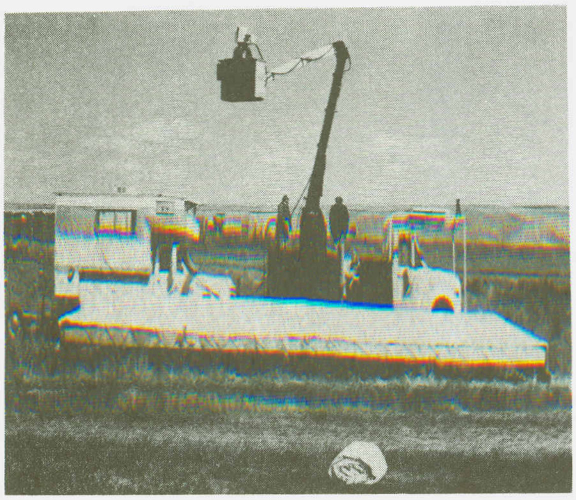
Purdue/LARS Exotech 100 Field System.



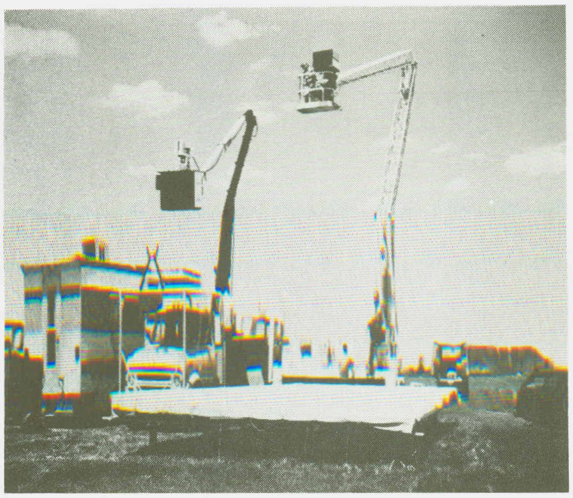
Weather station near the experimental plots.

Portion of the experimental plots illustrating variations in maturity, height, and biomass.

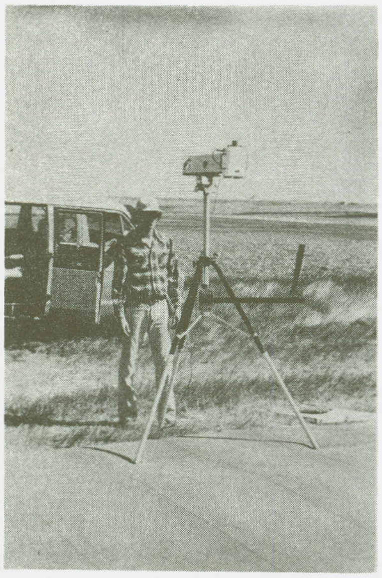
Figure A-6. Illustrations of data collection systems and experimental plots at Williston, North Dakota Agriculture Experiment Station.



Purdue/LARS Exotech 20C field system at calibration site in Williams County, North Dakota.



Purdue/LARS Exotech 20C field system and NASA/JSC FSAS at calibration site in Finney County, Kansas for correlation experiments.



Purdue/LARS Exotech 100 tripod-mounted field system at calibration site in Williams County, North Dakota.

Purdue/LARS Exotech 100 tripod-mounted field system at modeling field in Williams County, North Dakota.

Figure A-7. Illustrations of data collection systems at the calibration site and modeling field in Finney County, Kansas and Williams County, North Dakota.

barium sulfate. Additional aspects of the calibration procedures and data acquisition are described in the Calibration and Verification section.

Canopy Modeling Measurements. At the modeling field, radiometric, photographic, and agronomic data were collected in support of the wheat canopy modeling studies (see Figure A-7). The modeling data collection activities differed from past years in that a fairly complete set of data were collected over four plots at the agriculture experiment station and only a partial set of data were collected over a large commercial field. Usually modeling data are collected only over a large commercial field, but the commercial fields of spring wheat in the area had been planted two to three weeks earlier than normal and were therefore past the tillering stage before our data collection began.

Following discussions with Mr. Malila of ERIM, it was decided that the first priority was to collect modeling data over four plots at the experiment station in order to obtain data for investigating the detection threshold of wheat. Modeling data were collected in a commercial field when time and weather allowed. The dates on which modeling data were collected on the four plots are: June 1, 18, 19, 23; July 3, 4, 5, 7, 14, 20, 28; and August 8. Data were collected at the commercial field on July 19, 22, 23, 25; and July 18.

The modeling data include Landsat band reflectance factor measurements, profile and angular photographs, agronomic measurements, plant counts, plant height, leaf area index, biomass, plant condition, head area, and stem area.

b. Data Processing and Distribution

The data being processed for the field measurement project at Purdue/LARS are of two basic types, multispectral scanner data and spectrometer data. The data are from three sites: Finney County, Kansas; Williams County, North Dakota; and Hand County, South Dakota. The multispectral scanner data includes that from the 4-channel Landsat MSS, the 11-channel MMS on the NASA-P3 aircraft, and the 24-channel MSS on the NASA-C130 aircraft. The spectrometer data includes that collected by the NASA/ERL Exotech 20D field system, the NASA/JSC Field Spectra Acquisition System (FSAS), the NASA/JSC Field Spectrometer System (FSS), the Purdue/LARS Exotech 20C field system, and the Purdue/LARS Exotech 100 field system.

The data is processed into comparable formats in order to make meaningful comparisons of the data acquired by the different sensors at different times and locations. The spectrometer tapes (EXOSYS format) contain the spectral bidirectional reflectance factor measurements along with the corresponding agronomic and meteorological measurements. Processing of the Exotech 20D, Exotech 20C, and FSS spectrometer data includes keypunching the associated agronomic and meteorological data which is then integrated with the spectral bidirectional reflectance factor data. The FSAS spectrometer data are converted from the NASA tape format into EXOSYS format; i.e., NASA/JSC includes the agronomic and meteorological measurements with the spectral data before it is sent to Purdue/LARS.

Processing scanner data is a straight conversion from the original format (either Landsat bulk or Universal format) to LARSYS Version 3 format. The Exotech 100 bidirectional reflectance factor data is presently stored as hard copy, listings, or cards.

In the past 18 months, 116 Landsat frames, 108 aircraft runs, and 85,000 spectrometer runs have been reformatted and entered into the field measurements data library. The data processing/reformatting status is summarized in Table A-1. In addition, four dates and two dates of 1975 Landsat-2 data collected over the Finney County, Kansas, and Williams County, North Dakota, intensive test sites, respectively, were registered.

The general organization of the LACIE field measurements library is illustrated in Figure A-8. The data in the library includes multispectral scanner data, spectrometer data, photography both aerial and ground, agronomic observations, meteorological measurements, flight logs, and verification reports. The data formats are either tape, film, or hard copy.

All of the spectrometer data indicated in Table A-1 as being in the library, plus flight logs, mission reports, and photography have been distributed to researchers at Texas A&M University and the Environmental Research Institute of Michigan. Also, data have been provided, upon approval from NASA/JSC, to Goddard Institute of Space Science and the NASA/Goddard Space Flight Center.

A Field Measurements Data Library Catalog providing information on what data are available was prepared for present and future users of the data [4]. The catalog is divided into separate volumes - one for each crop year during which data were collected. Each volume lists the data by site, date, and instrument type. Brief descriptions of the data are also included.

A document providing examples of the field measurements data (Landsat, aircraft, FSS, Exotech 20C, Exotech 100, and ground observations) was prepared in a limited quantity. Two copies were provided to NASA/JSC; parts of one were later sent to NASA/Headquarters. An additional report containing examples of the spectrometer data in the form of spectral reflectance curves is currently being prepared. This document is being prepared to illustrate the major sources of variation in the reflectance of wheat and related crops and will contain labeled examples showing variation 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. An introductory text will explain and illustrate how the data have been acquired.

III. Spectral Data Calibration, Verification, and Correlation

a. Calibration Techniques

A key objective of the LACIE Field Measurements program has been the acquisition of calibrated spectral bidirectional reflectance factor data.

Table A-1. Field Measurements Data Processing/Reformatting Status as of November 9, 1977.

Instrument/Data Type	1974-75 Data		1975-76 Data		1976-77 Data		
	Completed & In Library	In Processing	Completed & In Library	In Processing	Completed & In Library	In Processing	At JSC
Landsat MSS							
Whole Frame CCT (Frames)	20	0	62	0	34	0	*N.A.
Aircraft Scanner (Dates/Runs)	19/149	0	16/97	0	7/35	0	10/
Helicopter Mounted Field Spectrometer (Dates/Runs)							
Field Averages	19/2,343	0	27/2,193	0	5/333	9/	7/
Individual Scans	19/29,579	0	27/38,476	0	5/8539	9/	7/
Truck Mounted Field Spectrometer (Dates/Runs)							
FSAS	6/65	0	23/322	0	5/116		8/
Exotech 20C	20/1,129	+4/526	14/1,356	+4/1,244	0	20/953	N.A.
Exotech 20D	45/645	0	-	-	-	-	-

* Not Applicable

\+ Modeling field data which will be processed when resources allow

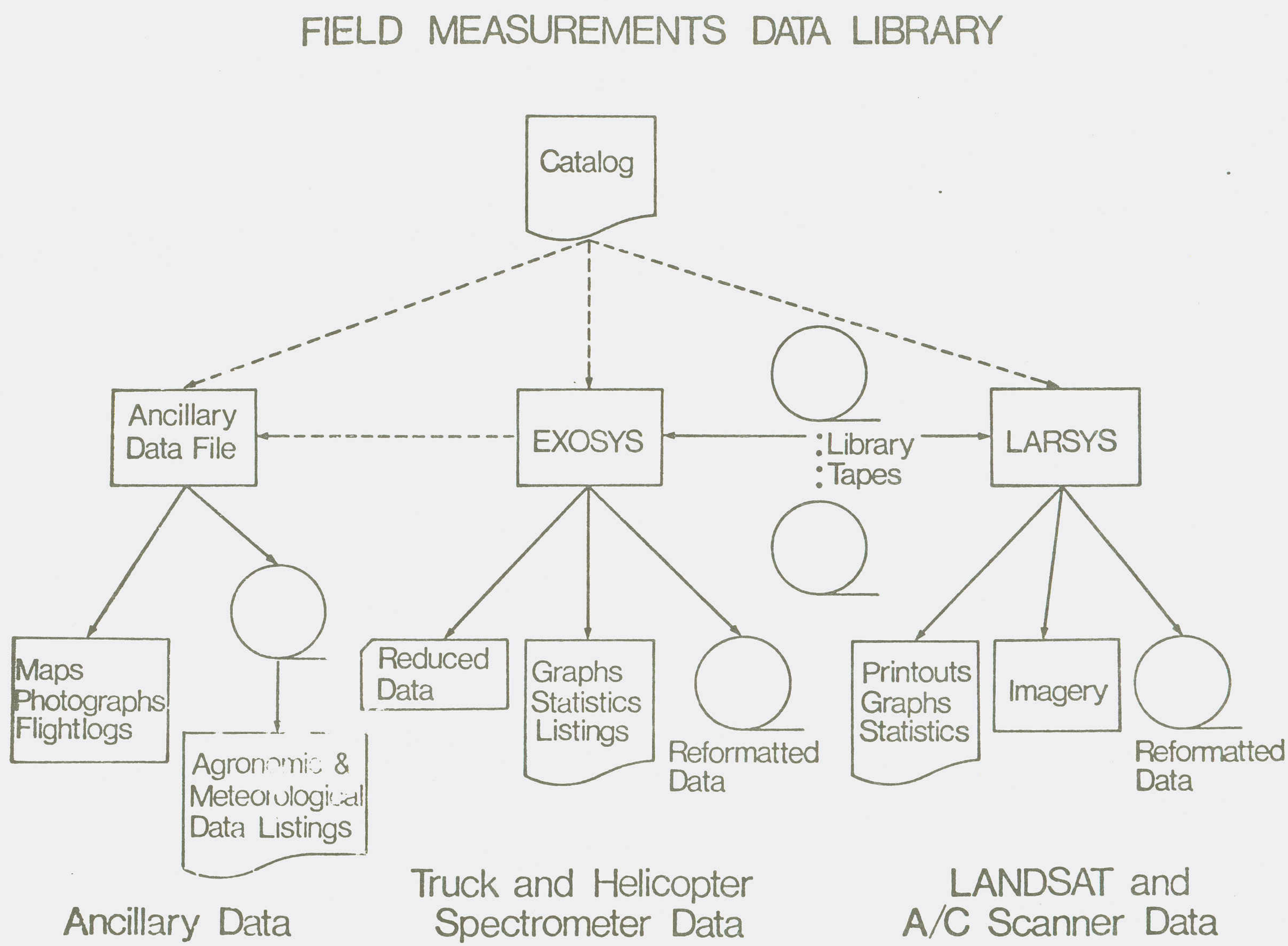


Figure A-8. Organization of LACIE Field Measurements Data Library.

Calibrated data are required in order to: (1) facilitate comparisons of data from different sensor systems and (2) compare and relate spectral measurements made at one time or location to those made at other times or locations.

Three major sensor types were used to gather spectral data: truck-mounted spectrometers were used on agricultural experiment station plots; a helicopter-borne spectrometer was used for agricultural fields; and an air-borne scanner was used on agricultural fields and experiment station plots. The data from these instruments were related through the use of five 20x40 foot gray canvas panels and related to the reflectance of pressed barium sulfate. Figure A-9 shows the initial laboratory comparison of the painted barium sulfate reference panels to pressed barium sulfate. The painted reference panels serve as reference surfaces for the subjects measured by the truck-mounted spectrometers. The truck-mounted spectrometers also compare the painted reference surfaces with the canvas gray panels which are used as reference surfaces by the helicopter and airplane-mounted instruments. Since the airborne instruments measure the reflectances of fields having many Landsat pixels, these data may be used to calibrate Landsat data to reflectance units.

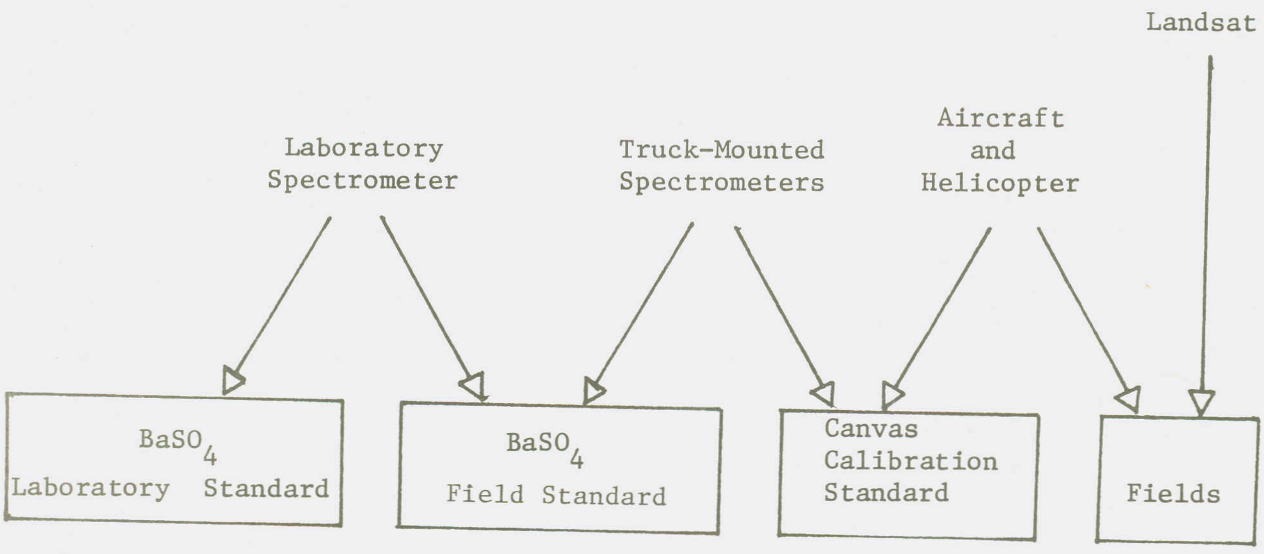


Figure A-9. Calibration method for LACIE field measurements data.

The truck-mounted spectrometer relied on frequent reference to painted barium sulfate panels to determine the spectral bidirectional reflectance factor of the subject. The helicopter spectrometer used frequent reference to the gray panel having a nominal reflectance of 55%. The airborne scanner was calibrated by reference to the five gray canvas panels. Calibration and correlation procedures are described in detail in the Project Plan [2].

Significant characteristics of the procedure described in Figure A-9 are: (1) all spectral reflectance factor data are related to pressed barium sulfate and (2) each instrument periodically measured the reflectance of the 20x40 foot gray panels. These procedures enabled the comparison of instrument performance under nearly ideal conditions and provided a basis to evaluate system performance as described in IIb, IIIb, and IIIc of this report.

Key milestones in the development of calibration techniques to date are:

(1) Documentation and implementation of standardized procedures for field measurements involving truck and helicopter spectrometer data and aircraft scanner data. November 1974. [1].

(2) Improvement of calibration procedures for helicopter data. January 1976 and April 1976. [5,6].

(3) Development of the hardware and procedures to enable indoor measurement of the reflectance of painted reference panels and determination of the bidirectional reflectance properties of the painted and canvas reference panels. March 1975. [3].

(4) Application of field measurements data to calibrate aircraft scanner and Landsat MSS data. July 1977. [7].

(5) Side-by-side instrument comparison studies using canvas gray panels by the JSC FSAS, Purdue Exotech 20C, JSC FSS, and NASA MSS. July 1977. [8].

The important task of bringing the instruments together for comparative tests was attempted throughout the project, but was not accomplished until May and July 1977. The data is being processed for analysis. The results of this analysis should provide some definitive answers relative to calibration procedures.

Fundamentals of the Calibration Procedures. The essential field calibration procedure consists of the comparison of the response of the instrument viewing the subject to the response of the instrument viewing a level reference surface. The term bidirectional reflectance factor has been used to describe the measurement: one direction being associated with the viewing angle (usually 0° from normal) and the other direction being the solar zenith angle.

For the natural situation, however, some of the light incident upon the subject and scene is due to scattered sunlight. The effects of this non-directional component (skylight) of the incident radiation have been assumed to be small with the basis for this assumption being the qualitative consideration that the skylight component is small compared to the direct component and that the reflectance of the subject surfaces is approximately lambertian. The alternative of measuring the diffuse component for the reference and the target has been avoided due to the added uncertainty associated with the additional measurements and required computations.

The uncertainties involved in the two measurement situations were analyzed. During the past quarter an adequate treatment of the analysis was made possible by: (1) sufficient field experience to provide quantitative estimates of the uncertainties involved in the measurements and (2) recent work which quantifies the limit of error associated with the assumption that hemispherically irradiated agricultural scenes will appear to be lambertian [9].

The technique of directly computing the reflectance factor is analyzed with the assumption that the skylight is uniformly distributed and that the reference surface is lambertian. It has been shown that:

$$R_F(\theta,\phi;0,0) = R(\theta,\phi,0,0) \quad [1 + K_1 \cdot K_2 \;(\theta,\phi)]$$

where R_F is the reflectance factor with respect to the reference surface measured without regard for the skylight; R is the reflectance factor of the subject measured if the sky were black (i.e., no scattering); $K_2(\theta,\phi)$ is the ratio of the diffuse to the total irradiance at the time of the measurement; and K_1 is the amount which the bidirectional reflectance distribution function (BRDF) of the subject (viewed along the normal) differs from the BRDF of a lambertian reflector having the same normal - hemispherical reflectance.

The subtractive technique of measurement of bidirectional reflectance factor is described as:

$$R_s(\theta,\phi;0,0) = \frac{L'_t(\theta,\phi) - L'_{d,t}}{L'_s(\theta,\phi) - L'_{d,s}} = \frac{L'_{D,t}(\theta,\phi)}{L'_{D,s}(\theta,\phi)}$$

where L'_t is the response of the instrument to the total irradiance and $L'_{d,t}$ is the response of the instrument to the shadowed target, and $L'_{D,t}$ is the computed response of the target to the direct component of the irradiance. The s subscript is associated with measurements of the reflectance surface.

To provide a basis for comparison, an "ideal" measurement was also treated. This corresponds to a hypothetical measurement under black-sky conditions:

$$R(\theta,\phi;\ 0,0) \;= \frac{L'_{D,t}(\theta,\phi)}{L'_{D,s}(\theta,\phi)}$$

Analysis of the limit of uncertainty associated with these three techniques yields the following limits of relative uncertainty:

Ideal

$$\frac{\Delta R}{R} = \left[\frac{1}{R} + 1\right] \cdot \frac{NE\Delta L}{L'_{D,s}} + \tan\theta\Delta\theta$$

where R is the true reflectance factor, $NE\Delta L/L'_{D,s}$ is the ratio of the noise equivalent radiance of the instrument to the direct component of the radiance of the reference surface and the $\tan\theta\Delta\theta$ term derives from the uncertainty involved in orienting the reference surface to a level position.

Subtractive

$$\frac{\Delta R}{R} = \left[\frac{1}{R} + 1\right] \cdot \frac{2NE\Delta L}{L'_{D,s}} + \tan\theta\Delta\theta + \frac{\Delta L_\tau}{L'_{d,t}} + \frac{\Delta L_\sigma}{L'_{d,s}}$$

where $\Delta L_\tau/L'_{d,t}$ is the relative uncertainty associated with shadowing the target and $\Delta L_\sigma/L'_{d,s}$ is the relative uncertainty associated with shadowing the standard.

Direct

$$\frac{\Delta R}{R} = K_1 \cdot K_2 + (1-K_2) \cdot \left[\frac{1}{(1 - K_1 \cdot K_2)\, R} + 1\right] \cdot \frac{NE\Delta L}{L'_{D,s}} + \tan\theta\Delta\theta$$

where K_1 and K_2 have been described previously. It can be seen that, if K_1 and K_2 are 0, the direct technique passes to the ideal technique. If K_1 is 0, the direct technique is superior to the ideal technique in that the additional radiant power flux supplied by the diffuse component of the irradiance will improve the relative uncertainty by providing a greater signal-to-noise ratio.

Using published and measured values for K_1 and K_2 and typical limits of relative uncertainty associated with field measurements, the limits of relative uncertainty for each technique have been evaluated and graphed in Figures A-10 and A-11. The values for K_2 were chosen to represent moderate and very noticeable haze conditions. Otherwise, it is assumed that a typical instrument is making reasonable measurements under worst case conditions. K_1 was selected to be positive since that represents the worst case condition for the direct technique.

Figures A-10 and A-11 show that the direct method is superior to the subtractive method in all respects over the range of reflectances encompassed by environmental scenes and that for moderate haze conditions there is

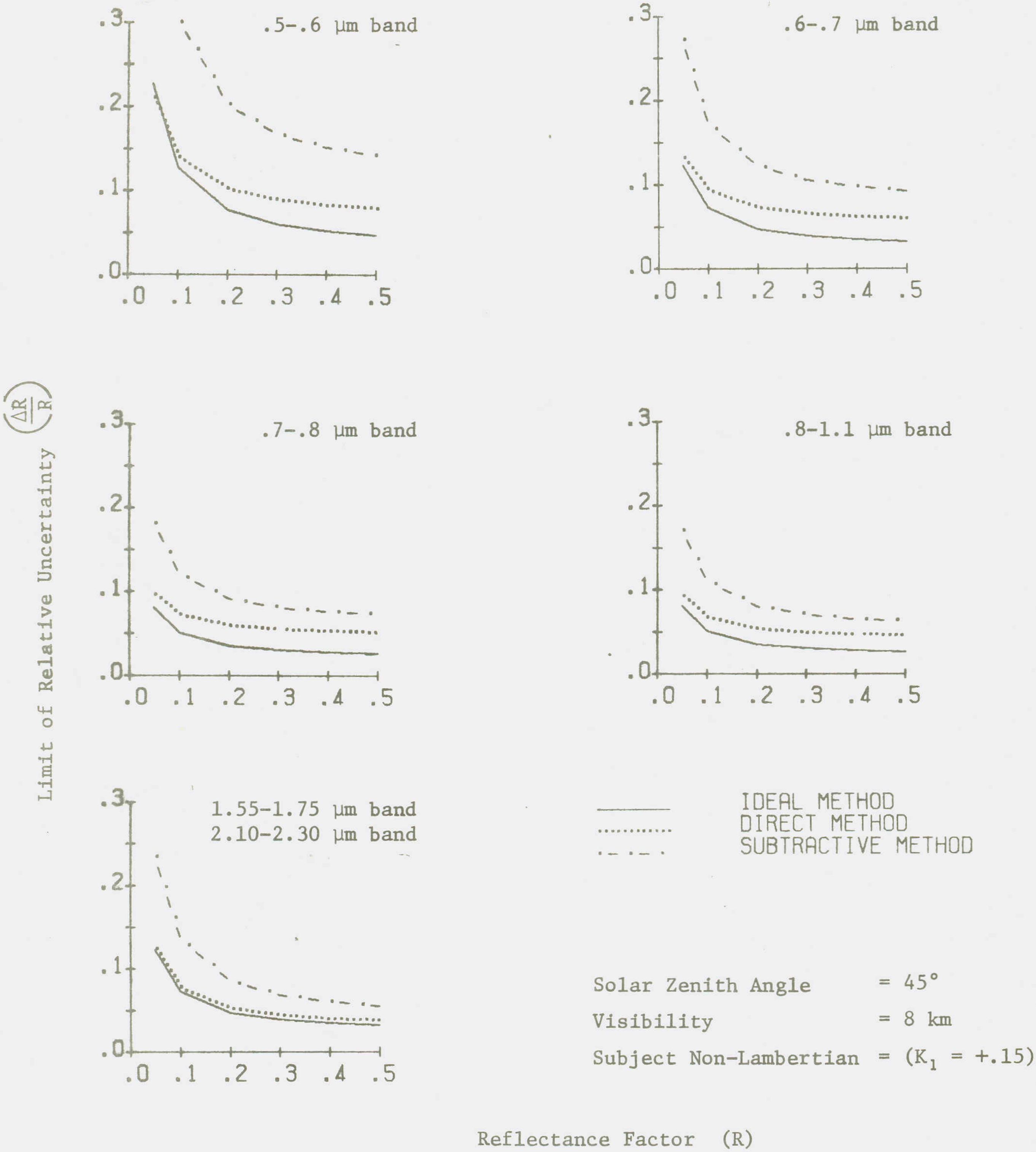


Figure A -10. Comparison of limit of relative uncertainty vs reflectance factor for three measurement methods under very hazy conditions.

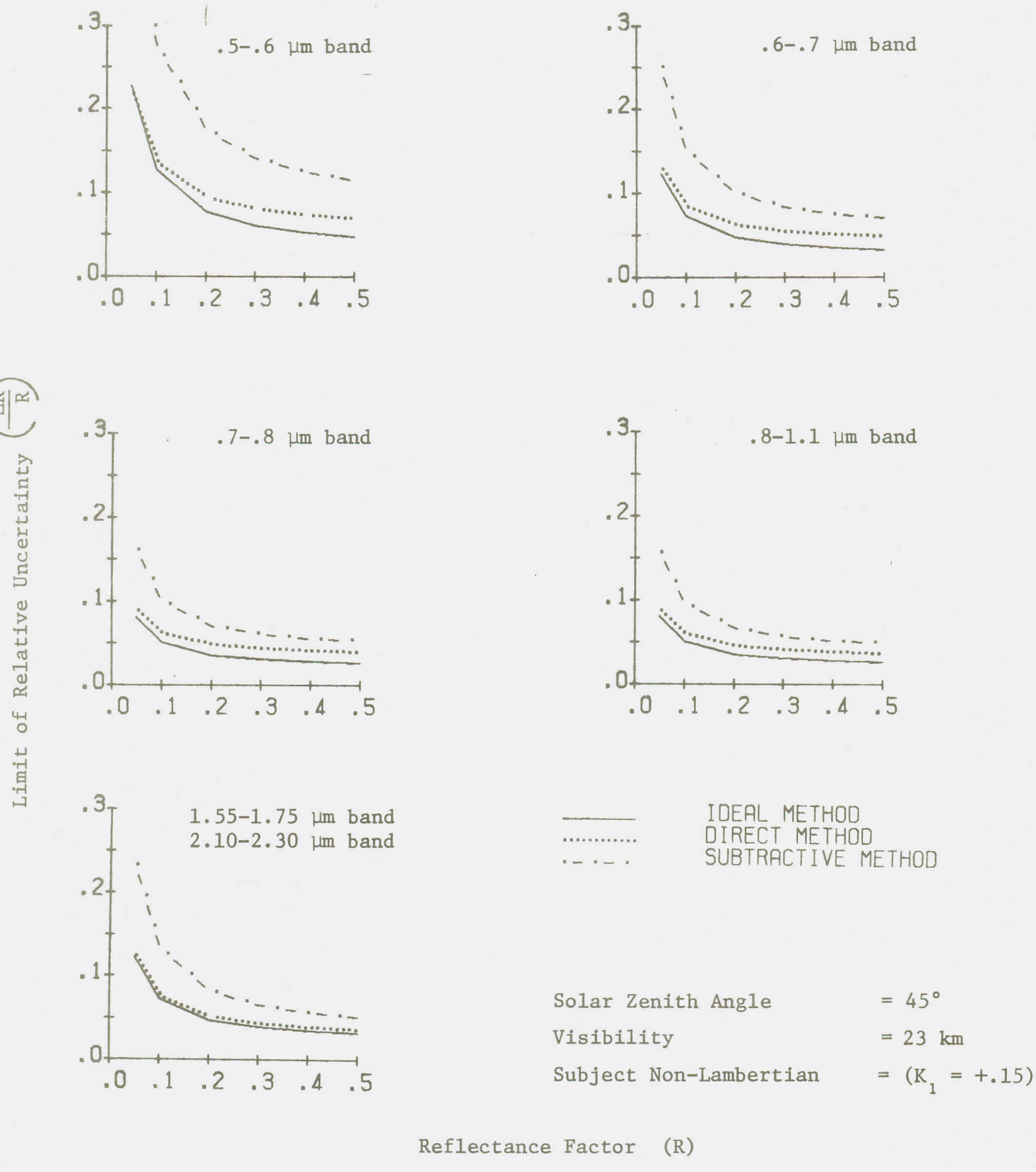


Figure A -11. Comparison of the limit of relative uncertainty vs reflectance factor for three measurement methods under moderate hazy conditions.

scarcely any difference between the limits of relative uncertainty associated with the direct technique and the ideal technique.

b. Field Measurements Data Verification

An important part of the field measurements project is the acquisition of high quality, calibrated spectral measurements. To a large degree, this depends on having timely and quantitative methods available for determining data characteristics. This information can be used (1) for identifying sensor deficiencies which can be corrected and (2) by data analysts in selecting data for analysis and in interpreting analysis results.

In April 1976 a document describing the technical recommendations of Purdue/LARS for data quality evaluation and verification for the spectrometer and aircraft scanner systems was prepared under NASA contract NAS9-14016. The recommendations pointed out the importance of the verification tasks which had been implemented since the beginning of the field measurements project and included tasks which NASA/JSC and Purdue/LARS needed to implement.

Due to limited resources, not all of the recommendations could be implemented and some of those that were could not be implemented as quickly as desired. The recommendations for aircraft and spectrometer verification tasks are summarized below along with dates indicating when the tasks were implemented.

- Aircraft Multispectral Scanner Data

    1. A record of total irradiance at the aircraft calibration site as a function of time to determine the illumination conditions. [NASA, June 1976]

    2. Histograms of data to determine detector operation, bit dropping or favoritism, dynamic range, sensitivity, and saturation. [NASA and Purdue, May-June 1976]

    3. Implement a set of machine programmable quantitative measures to be used during the reformatting process to determine the quality of the data. [Not implemented]

- FSS Data

    1. A record of the total irradiance at the helicopter calibration site as a function of time to determine the illumination conditions. [NASA, June 1976]

    2. A calibration uniformity test. After a cosine correction for sun angle the calibration spectra should be nearly identical. [Purdue, January 1977]

    3. Histogram tests for the processed data to indicate proper A-D

conversion. These tests would be similar to those described for the scanner data. [Not implemented]

4. A system performance test utilizing measurements of a series of gray-scale calibration panels. [NASA and Purdue, May-July 1977]

- Truck-Mounted Spectrometer Data (Exotech 20C, Exotech 20D, and FSAS)

    1. Operational procedures which include frequent checks of alignment of field of view. [Purdue Exotech 20C, Nov. 1974; ERL Exotech 20D, July 1975; JSC FSAS, October 1976]

    2. Calibration procedures using large barium sulfate field standards [Purdue Exotech 20C, November 1974; ERL Exotech 20D, March 1975]

    3. A record of total irradiance at the site as a function of time. [Purdue Exotech 20C, May 1975; JSC FSAS, June 1976]

    4. A calibration uniformity test similar to the one described above for the S-191 data [Purdue Exotech 20C, November 1974]

    5. Histogram tests for verifying analog to digital conversion similar to the ones described for the scanner data. [Purdue Exotech 20C, 1974]

    6. A system performance test including measurements of gray-scale calibration panels. Spectra would be examined for continuity, form, and magnitude. [Purdue Exotech 20C, November 1974; ERL Exotech 20D, March 1975; JSC FSAS, October 1976]

The results of the verification tasks completed are available to the users of the data. Results of the FSS calibration uniformity test are given below.

Results of FSS Calibration Uniformity Test. Each of the calibration panel runs for each date was normalized with respect to sun angle (see equation A-1) and plotted.

$$CV_t = V_t/\cos\theta_t \tag{A-1}$$

where

V_t = FSS response to calibration panel at time t,

CV_t = Sun angle adjusted, FSS response to calibration panel at time t,

θ_t = solar zenith angle at time t.

(See Figures A-12 and A-13 for examples). Ideally the sun angle adjusted response curves would fall on the same curve. This assumes that the calibration panel is a lambertian reflector, which is a good assumption down to 50° off normal. The situations which will cause the curves to be quite different are:

1. Data being collected on a cloudy day.

2. Not filling instrument field of view with the calibration panels.

3. Instrument instabilities.

4. Shadowing of the calibration panel.

It should be understood that a review of these curves does not necessarily give the whole story about all of the data collected on that day. However, a review of the plots does point out particular flightlines which the analyst may want to pay closer attention to. The values given with the plots (coefficients of variation) attempt to place some quantitative value on the variation of the curves so that dates can be compared. A summary of the average coefficients of variation is given in Table A-2. The average coefficient of variation is not computed with band centers 1.425, 1.475, 1.825, 1.875, and 1.925 μm since they represent regions of strong water absorption.

The results in Table A-2 indicate that there is significantly less variation on the adjusted calibration responses for the 1975-76 crop year than for the 1974-75 crop year.

c. Spectrometer Correlation Studies

The truck spectrometers including the NASA/ERL Exotech 20D, the NASA/JSC FSAS, and the Purdue/LARS Exotech 20C have collected data over five gray panels which are deployed at the intensive test sites for the aircraft and helicopter systems. The data collected by the spectrometers over the gray panels were used to compare the spectral measurements of the different spectrometers. The assumption made is that the reflectance of the gray canvas panels does not change appreciably with time during a crop year. Experience has shown that the reflectance of the panels changes very little with time. The bright gray panel used as the FSS calibration panel changes the most (2-3%) during the crop year because of dust collecting on the surface as the helicopter hovers over it.

Plots of the spectral measurements are shown in Figures A-14 to A-18 for gray panels 1 thru 5, respectively. Gray panel 1 is the brightest and gray panel 5 is the darkest. Each plot represents the average of all spectral measurements by a given spectrometer for a given year.

Analysis of the data shown in Figures A-14 to A-18 indicate that there is a significant variation in the different spectrometer measurements of panel 1. The variation may be caused in part by a variation in the

SITE WILLIAMS COUNTY N.D. LATITUDE 48.310N LONGITUDE 103.335W

SYMBOL	FLT. LINE	CAL TIME	IFOV	ZENITH	COSINE ZENITH
1	1	18.21.09	1.25	37.3	0.795
2	2	18.39.23	1.25	36.7	0.801
3	3	19.03.15	1.25	36.6	0.803
4	4	19.22.21	1.25	37.0	0.799

AVERAGE COEFFICIENT OF VARIATION (COEF. VAR.) IS .06

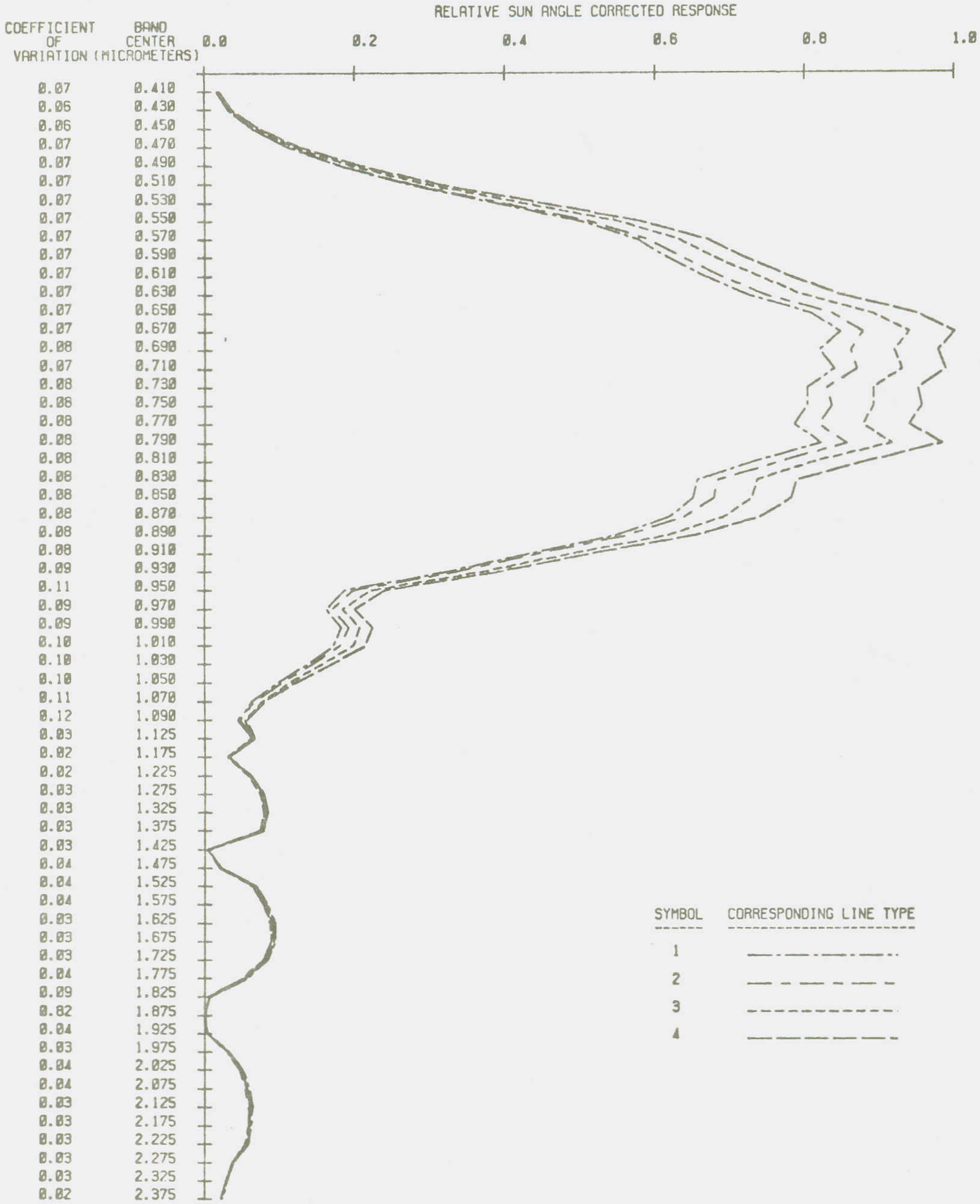


Figure A-12. Sun angle adjusted FSS values for four observations of the gray calibration panel collected on 8/23/75.

SITE FINNEY CO. 2, KANSAS LATITUDE 38.148N LONGITUDE 100.665W

SYMBOL	FLT. LINE	CAL TIME	IFOV	ZENITH	COSINE ZENITH
1	8	14.57.51	1.25	52.8	0.604
2	9	15.20.54	1.25	48.4	0.664
3	10	15.42.44	1.25	44.2	0.717

AVERAGE COEFFICIENT OF VARIATION (COEF. VAR.) IS .03

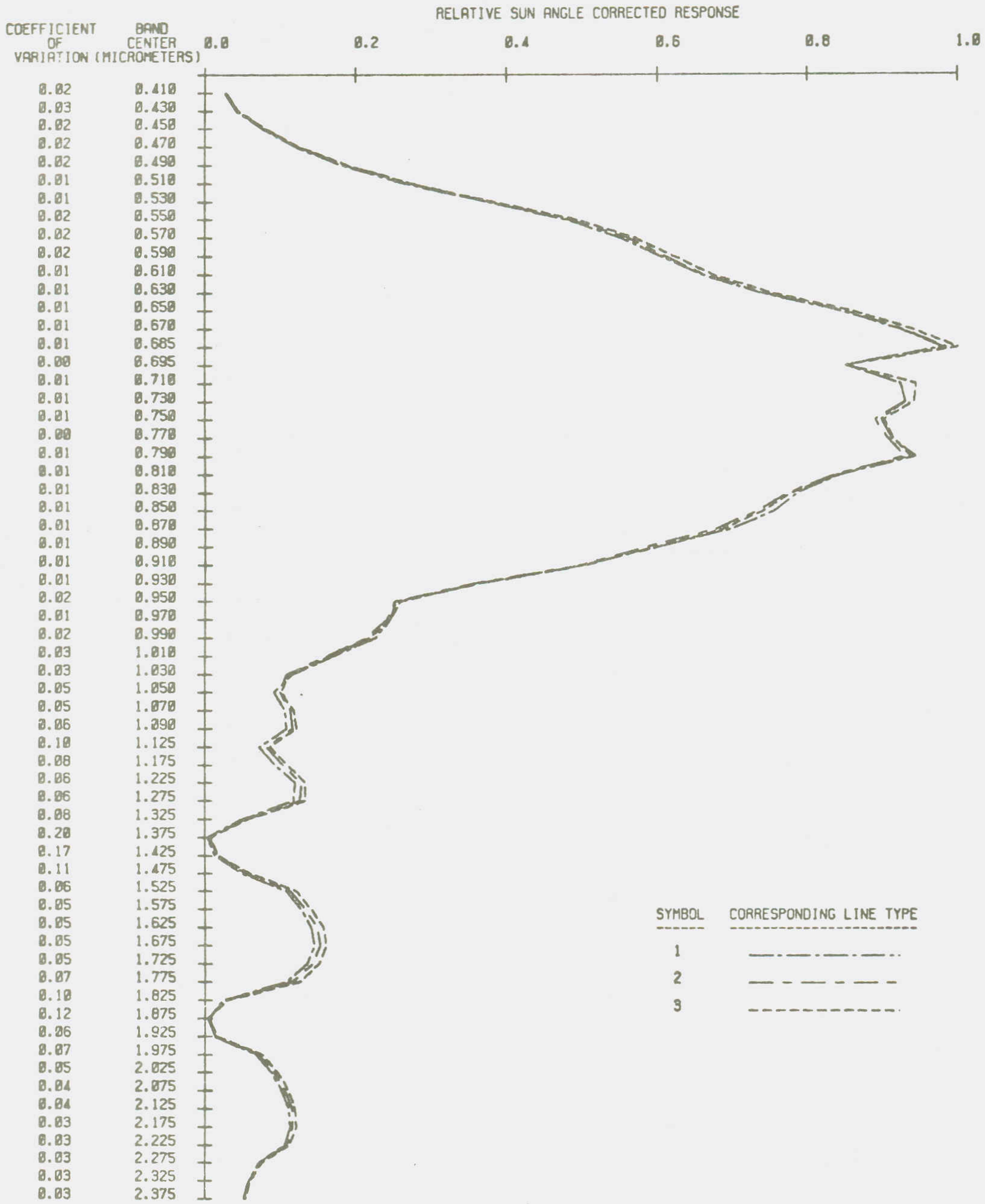
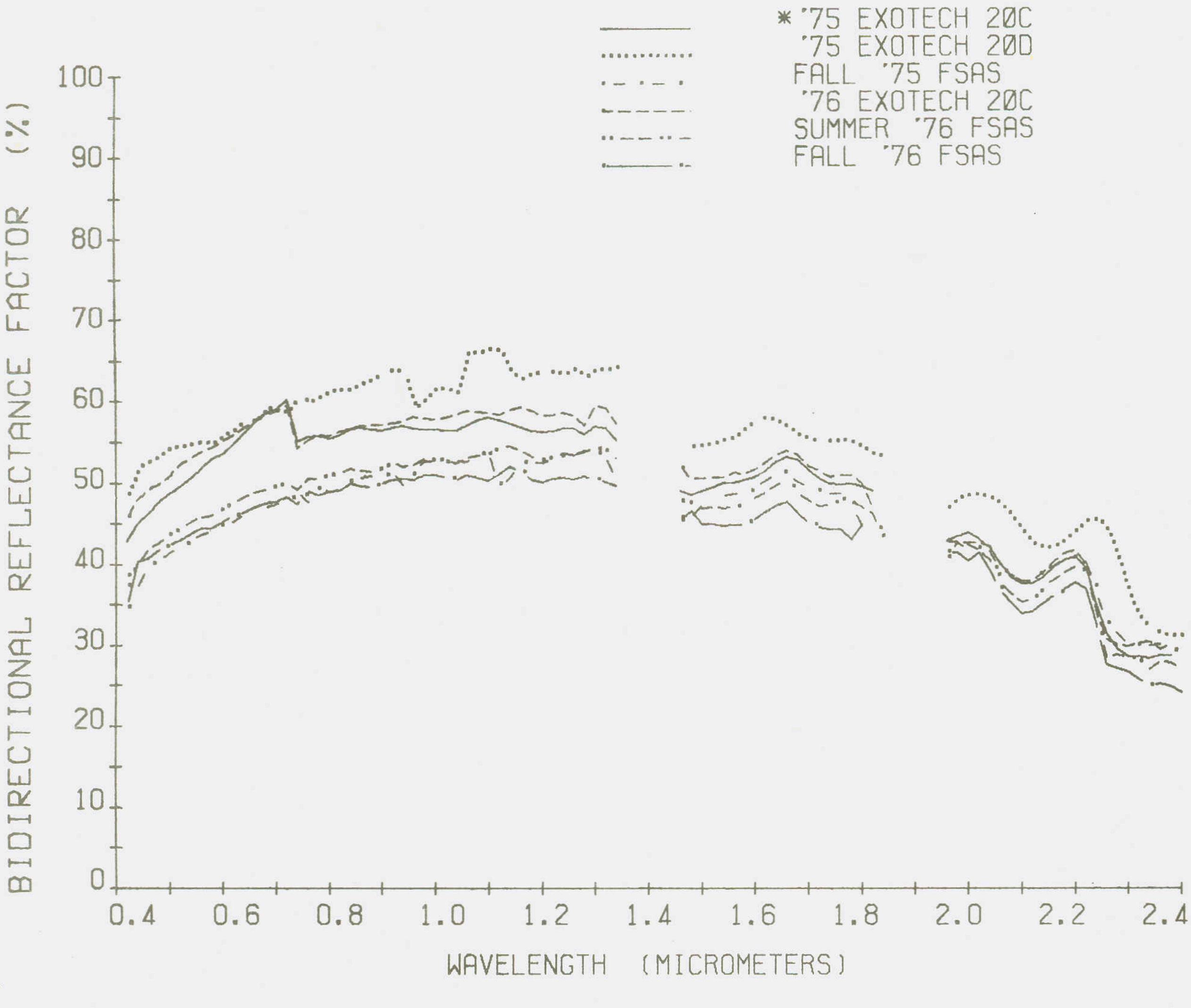


Figure A -13. Sun angle adjusted FSS values for three observations of the gray calibration panel collected on 5/6/76.

Table A-2. Summary of FSS calibration panel response variations.

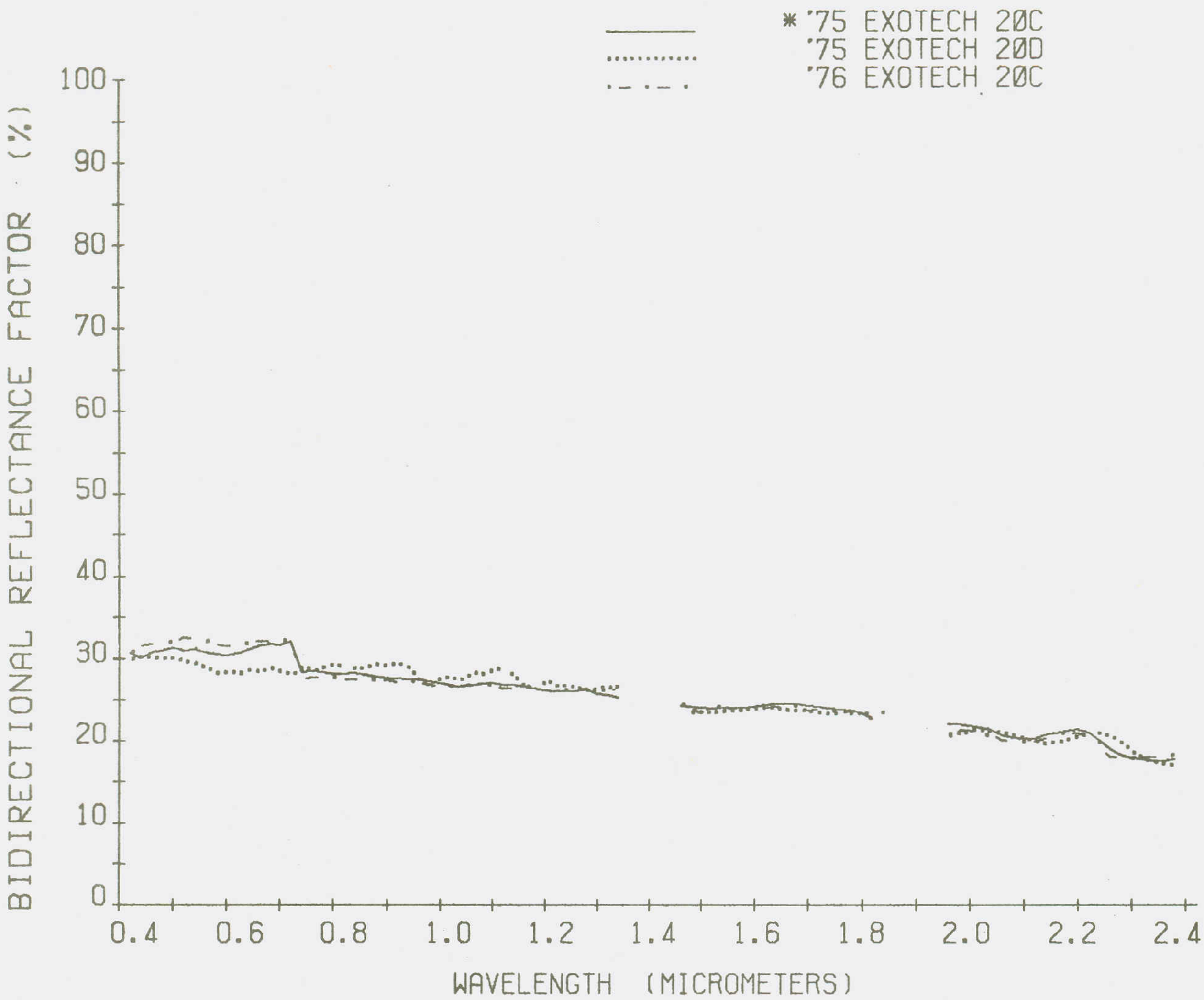
Finney Co., Kansas		Williams Co., N. Dakota		Hand Co., S. Dak.	
Date	Ave. Coef.* Variation	Date	Ave. Coef. Variation	Date	Ave. Coef. Variation
		1974-75 Crop Year			
11/05/74	.13	6/5/75	.09		
3/20/75	.09	6/22	.08		
4/8	.09	7/10	.09		
5/14	.03	7/18	.06	Not	
5/21	.12	7/27	.02		
6/2	.04	8/5	.04	Applicable	
6/9	.10	8/15	.04		
6/17	.09	8/23	.06		
6/26	.04	9/1	.08		
7/6	.09				
		1975-76 Crop Year			
9/16/75	.05	5/13/76	.01	10/15/75	.06
10/3	.24	5/28	.01	10/30	-
10/21	.04	6/17	.01	11/5	.04
11/11	.03	6/25	.01	5/11/76	.03
3/18/76	.02	7/6	.03	6/1	.03
3/31	.05	7/20	.03	6/19	.05
4/18	.02	7/28	.03	7/8	.04
5/6	.03	8/9	.02	7/31	.03
6/12	.06	8/19	.08		
6/30	.07				
		1976-77 Crop Year			
9/28/76	.01	5/8/77	.04	9/21/76	.02
10/15	.03	5/24	.03	10/21	.03
11/3	.03			4/21/77	.04
3/8/77	.02			5/10	.04
5/3	.02			6/16	.05

* Average coefficient of variation (standard deviation/mean) for all FSS bands except those at 1.425, 1.475, 1.825, 1.875, and 1.925 μm.



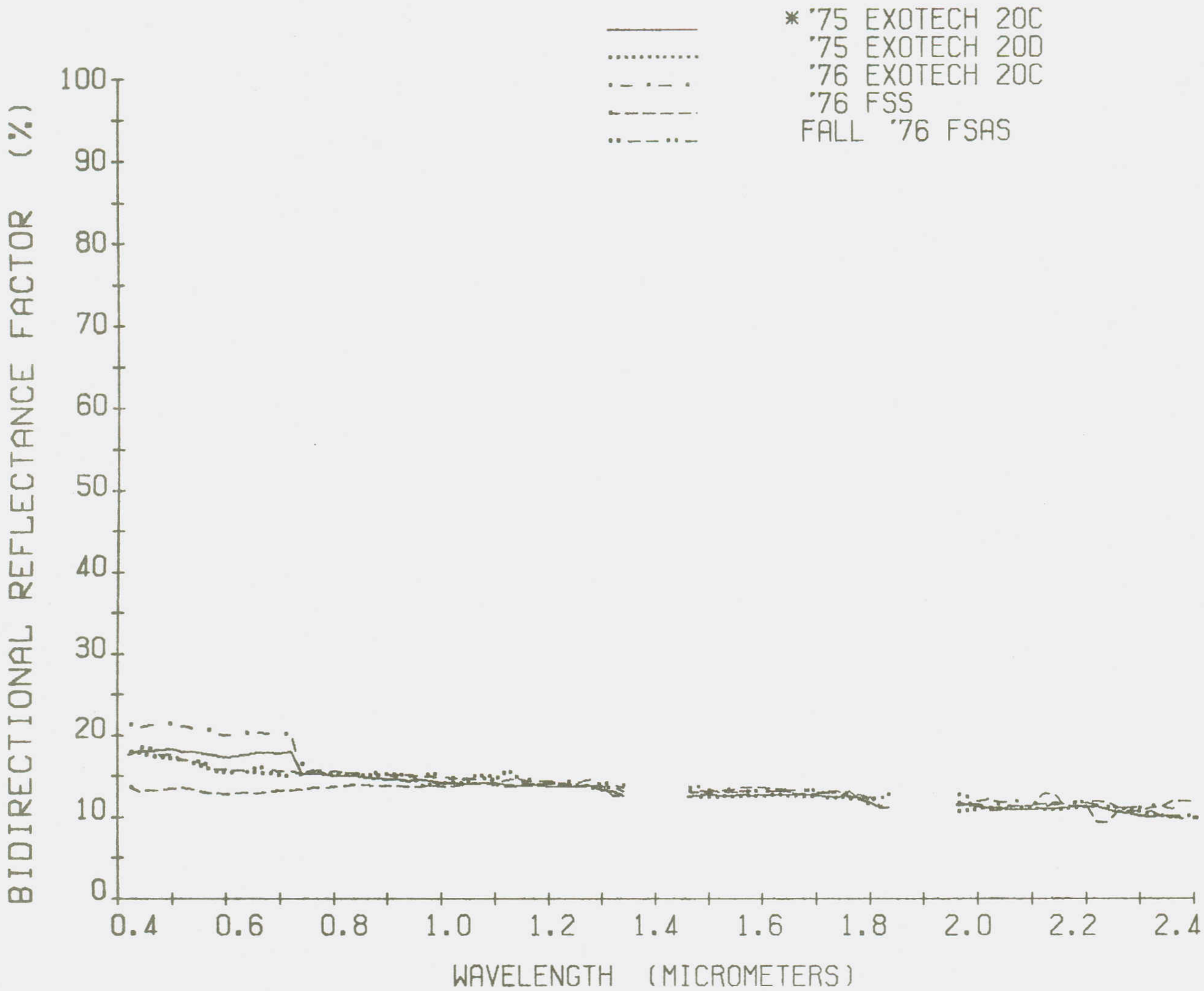
*AVERAGE OF PANEL-1 RUNS COLLECTED BY INSTRUMENT FOR SEASON.

Figure A-14. Summary of spectrometer measurements of gray panel 1.



*AVERAGE OF PANEL-2 RUNS COLLECTED BY INSTRUMENT FOR SEASON.

Figure A -15. Summary of spectrometer measurements of gray panel 2.



*AVERAGE OF PANEL-3 RUNS COLLECTED BY INSTRUMENT FOR SEASON.

Figure A-16. Summary of spectrometer measurements of gray panel 3.

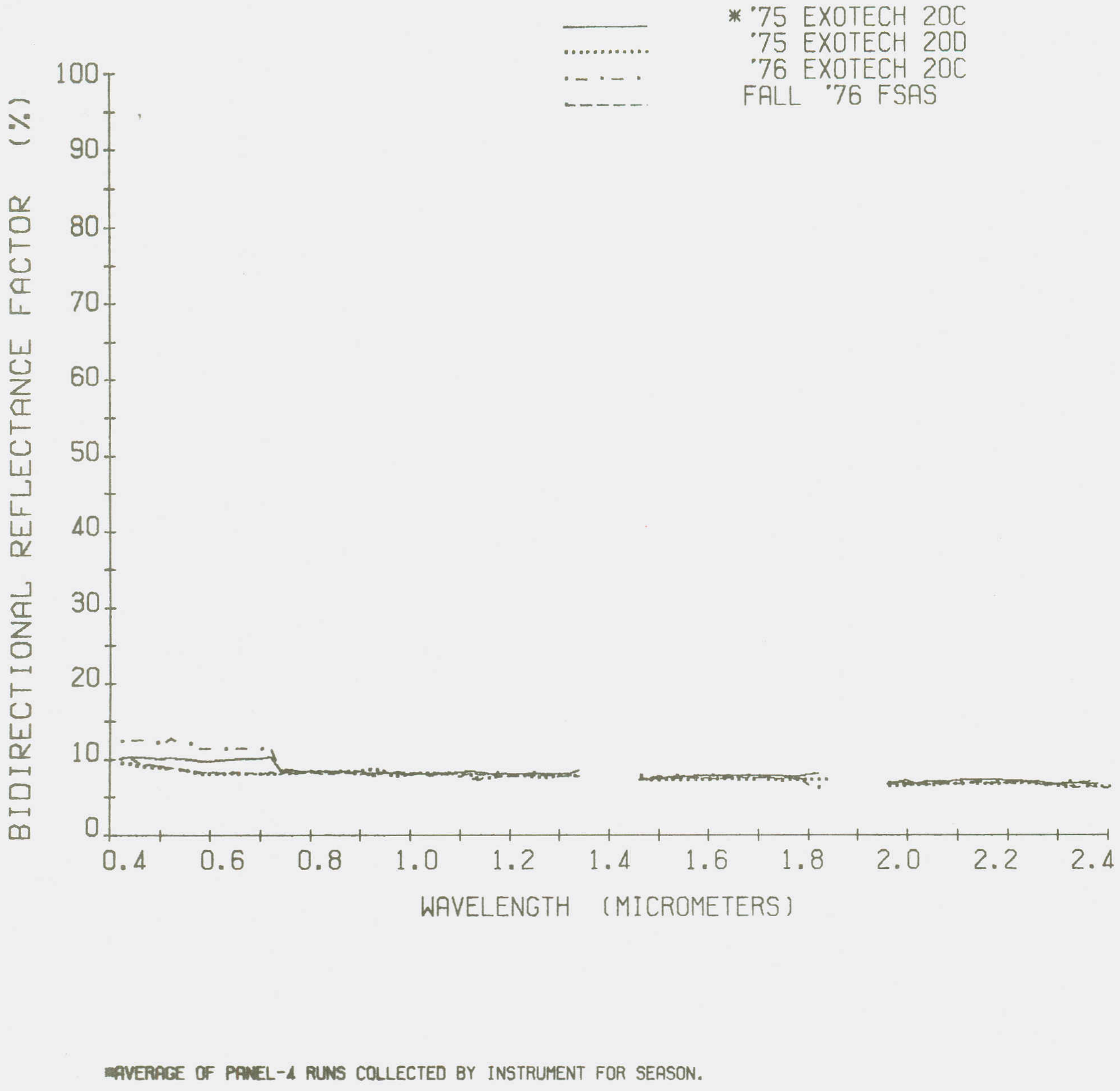
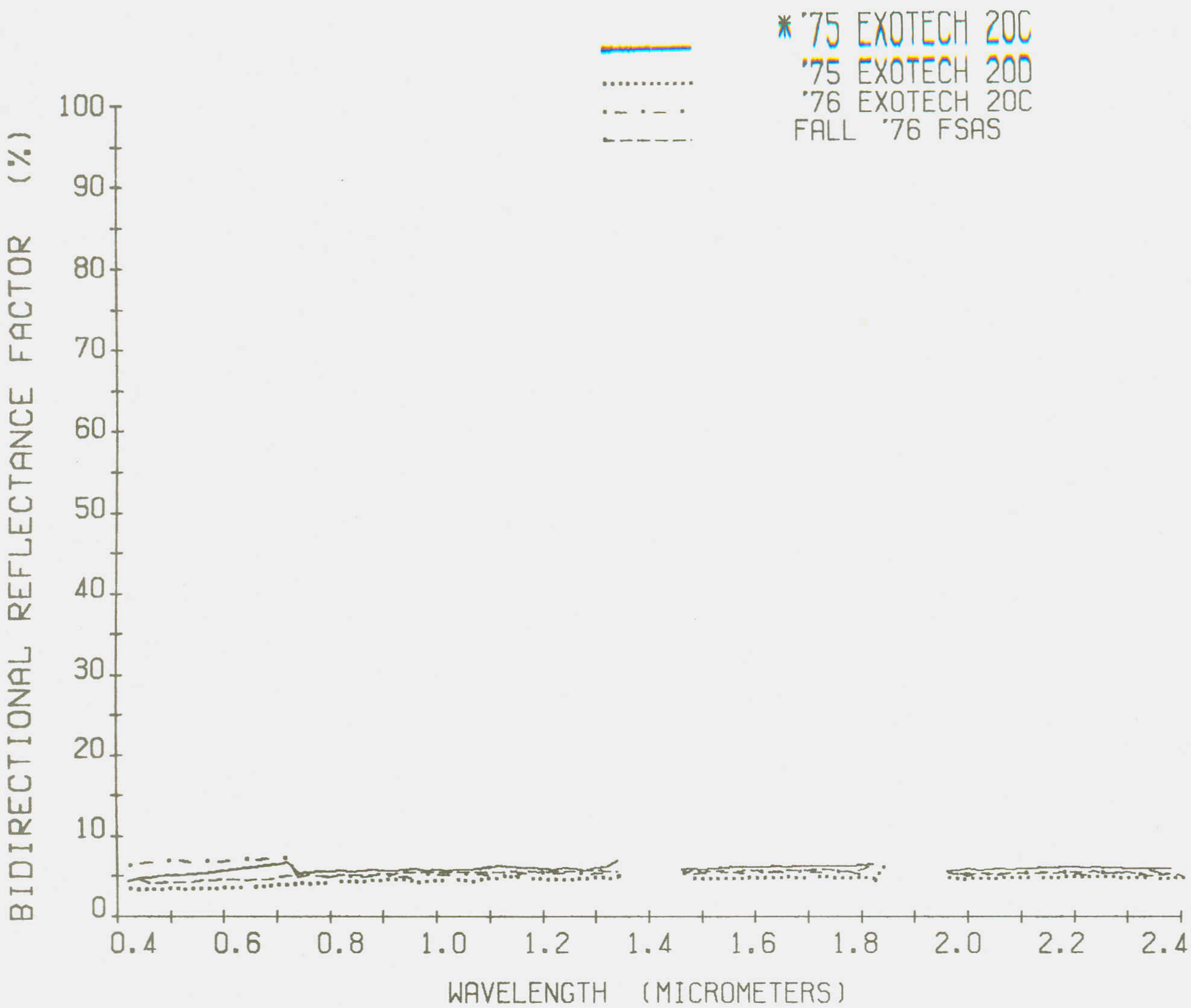


Figure A-17. Summary of spectrometer measurements of gray panel 4.



*AVERAGE OF PANEL-5 RUNS COLLECTED BY INSTRUMENT FOR SEASON.

Figure A-18. Summary of spectrometer measurements of gray panel 5.

actual reflectance of the panel, since the panel is washed about every six months. Some of the instruments measured the panel before it was washed and some after it was washed. Also, the margin of error associated with the 1975 Exotech 20D measurements is probably larger than that for the other instruments, since an optical misalignment occurred during the period of their data collection at Garden City in 1975. A review of the gray panel data indicates that the misalignment may have occurred around June 1, since before this date the gray panel measurements are consistent.

The analysis of the gray panel data also indicates that there is an offset in the spectral measurements made by the Purdue/LARS Exotech 20C field system at .73 μm. This is the wavelength where the silicon detector/filterwheel combination ends and the lead sulfide detector/filterwheel combination begins. The cause was found to be a combination of filterwheel deterioration and a nonlinearity in the electronics processing of the signal from the silicon detector which affected the .60-.73 μm wavelength region. Both of these problems were corrected for the 1977 data collection year. Figure A-14 also indicates a systematic difference between the FSAS and Exotech 20C measurements of gray panel 1.

Correlation of the data is summarized in Figures A-19 and A-20 and the linear regression analyses given in Table A-3. The correlation plots and the linear regression analyses include data for six wavelengths selected from the wavelength spectrum being measured. If the data from the instruments were correlated perfectly, the points would fall on a straight line ($r^2 = 1$ for regression analysis), have a slope of one ($c_1 = 1$), and pass through the origin ($c_o = 0$).

The information indicates that the gray panel data from the instruments are highly correlated. The r^2 values are high-for both the FSAS versus Exotech 20C comparison (.998-1.000) and the Exotech 20D versus Exotech 20C comparison (.999-1.000). The gray panel spectrometer data is the most closely correlated in the 0-30 percent (BRF) range, i.e., gray panels 2 thru 5.

The analysis of the correlation of the spectrometer data collected for the field measurements project is not complete. The task of developing and following procedures to calibrate and correlate the data from different spectrometers in a field environment is being accomplished and for the first time quantitative information is available.

The analysis results so far in comparing four spectrometers of two different types (filterwheel and interferometer), collecting data in locations hundreds of miles apart (Kansas and North Dakota) and at different times of the year are very encouraging. The reflectance data for the four gray panels used to calibrate the aircraft scanner data has a correlation better than .99 and agrees to within 4% of value. The reflectance data for panel 1 is within 10% of value for the FSAS and Exotech 20C data at two different sites with the panel on two different platforms. It is expected that the cause(s) of the systematic differences will be determined during

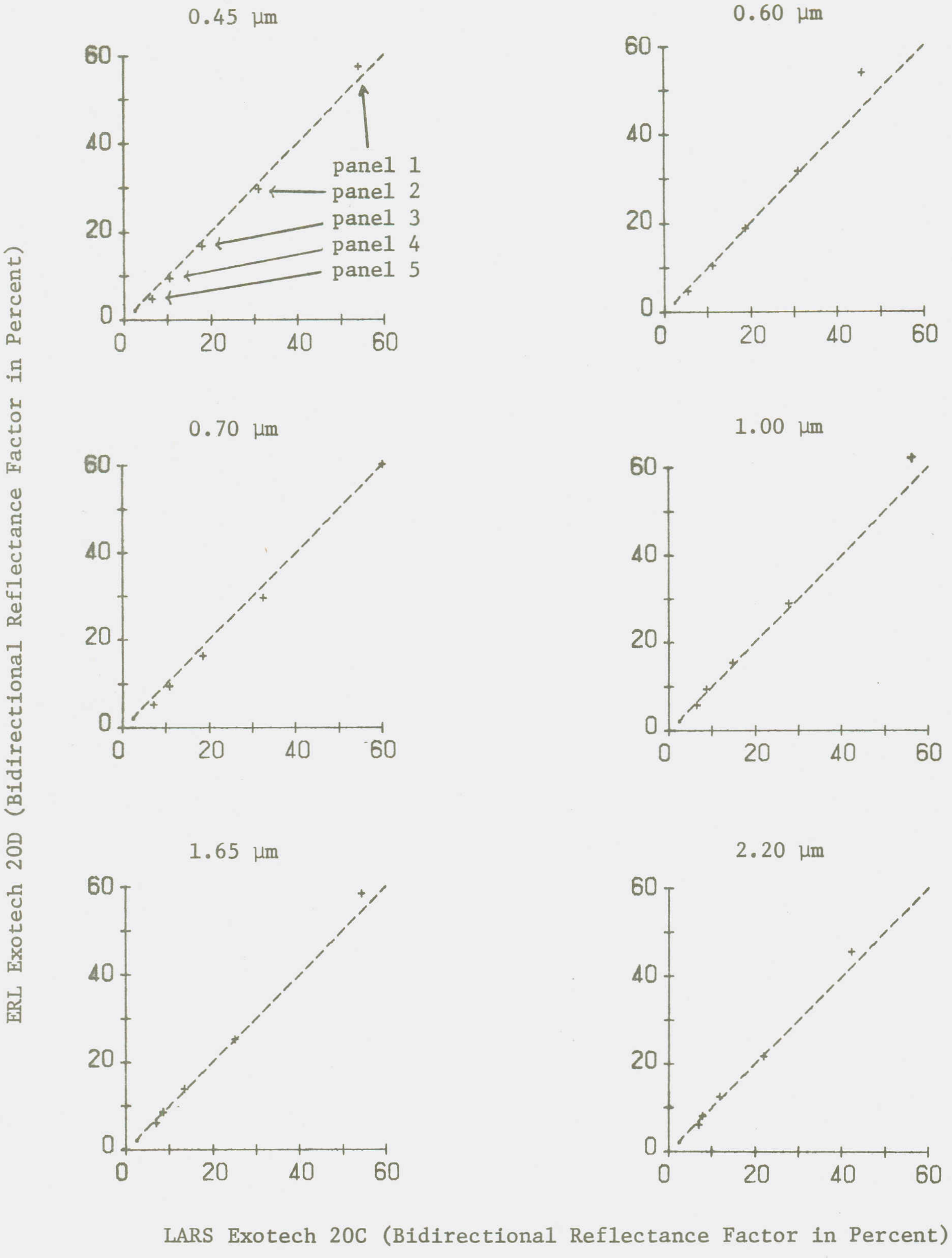


Figure A-19. Correlation plots of 1975 gray panel spectrometer data collected by NASA/ERL Exotech 20D and Purdue/LARS Exotech 20C field systems.

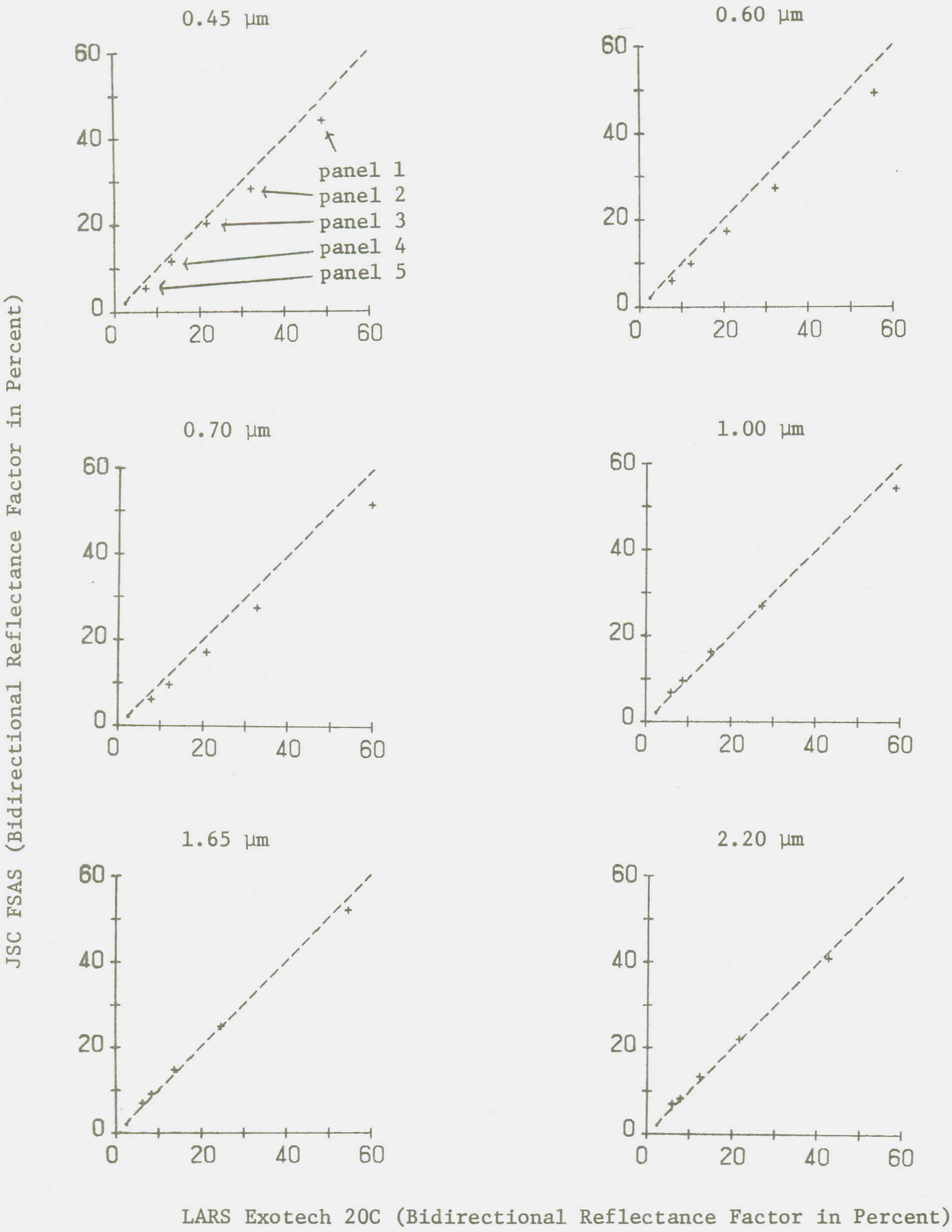


Figure A-20. Correlation plots of 1976 gray panel spectrometer data collected by NASA/JSC FSAS and Purdue/LARS Exotech 20C field systems.

Table A-3. Linear regression analyses of spectrometer gray panel data.

+BRF Range (percent)	Regression Coefficients*	Wavelength (μm) 0.45	0.60	0.70	1.0	1.65	2.20
		1975 Data					
		ERL Exotech 20D vs Purdue/LARS Exotech 20C					
0-60	r^2	0.992	0.998	0.999	1.000	0.999	0.998
	c_0	-3.6	-3.2	3.1	-1.7	-1.8	-1.7
	c_1	1.199	1.087	1.032	1.107	1.089	1.096
0-30	r^2	1.000	1.000	0.999	0.998	0.997	0.994
	c_0	-1.7	-2.0	-2.1	-1.3	-1.0	-0.8
	c_1	1.050	0.997	0.961	1.072	1.022	1.005
		1976 Data					
		NASA/JSC FSAS vs Purdue/LARS Exotech 20C					
0-60	r^2	0.998	0.999	1.000	1.000	1.000	0.999
	c_0	-1.8	-2.1	-2.0	1.2	0.8	0.6
	c_1	0.916	0.897	0.884	0.897	0.924	0.942
0-30	r^2	0.994	1.000	1.000	0.999	0.999	0.999
	c_0	-1.5	-1.5	-1.5	-0.7	0.3	0.1
	c_1	0.898	0.854	0.853	0.942	0.966	0.987

*Regression Equation: $y = c_0 + c_1 x$

+Bidirectional Reflectance Factor

the next contract year when the data collected during May and July of 1977 in the correlation experiments have been studied.

IV. Data Analysis

The spectral and agronomic measurements which have been acquired during the three years of the LACIE Field Measurements program are being analyzed to provide a quantitative understanding of the relationship of reflectance to the biological and physical characteristics of crops and soils. Knowledge of how important agronomic factors affect reflectance is necessary for optimal use of the current Landsat technology as well as for design and development of future remote sensing systems.

The primary data analyzed to date are the spectrometer data acquired by the truck- and helicopter-borne systems. These data are particularly useful because the spectral data are acquired in very small wavelength bands and are calibrated in terms of bidirectional reflectance. The narrow wavebands (i.e., complete spectra) permit 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 dates, locations, and sensors. Additional advantages of these data compared to Landsat are: no boundary pixels, higher spatial resolution, higher signal-to-noise ratio, and a more complete sampling of the crop through the growing season.

The analyses reported here were initiated in the spring of 1977 when data sets and resources for analysis became available. The analysis of much of the data has not been completed; future results may confirm or contradict results obtained to date, possibly altering conclusions which may be drawn from the analyses. In many ways, this section should be viewed as a status or preliminary report of our results rather than a final report. Continuation and completion of the analyses described here, as well as additional analyses are planned in the new SR&T contract to LARS.

a. Objectives

The overall objective of this task is to quantitatively determine the spectral-temporal characteristics of wheat, small grains, and confusion crops. The specific objectives are:

(1) To determine the relationship of agronomic variables such as biomass and leaf area index to multispectral reflectance.

(2) To determine the effects of cultural and environmental variables on the spectral response of wheat.

(3) To determine optimal times and wavelengths for discriminating between wheat and other crops.

(4) To determine optimal times and wavelengths for discriminating wheat from other small grains.

b. Approach

The data used to date for analysis were collected during the first two years of the field measurements program at the agriculture experiment stations (AES) in Garden City, Kansas, and Williston, North Dakota, and at the intensive test sites (ITS) in Finney County, Kansas, and Williams County, North Dakota. The general approach taken has been to analyze band means for the Landsat MSS and proposed thematic mapper bands.

Plots of the reflectance data were made to verify data quality and to qualitatively assess the information contained in the data. Regression and correlation were used to relate biological and physical parameters such as leaf area index, biomass, percent ground cover, height, and maturity stage to spectral response. Analysis of variance and covariance and discriminant analysis were performed to determine the threshold of detection and separability of wheat from other cover types. Details of the specific approach to each objective will be discussed in the results section.

c. Results and Discussion

The objectives of this task were pursued in two areas: spring wheat in North Dakota and winter wheat in Kansas. In each location, both FSS data from the intensive test site and truck-mounted spectrometer data from the experiment station were analyzed. The results will be discussed by location and objective.

1. Spring Wheat in North Dakota

This section describes the analyses which have been performed using data from both the agriculture experiment station and the intensive test site in North Dakota. For each analysis objective, details of the approach and results are discussed.

Relationship of Agronomic Variables to Reflectance. This objective has been investigated with the 1975 and 1976 AES data and, to a lesser extent, with the 1975 ITS data. Knowledge of the relationship of agronomic factors and reflectance is important in the crop identification problem for determining when wheat is detectable and when it can be distinguished from other crops. The ability to accurately assess a crop's condition and to predict crop yields by remote sensing techniques will depend on how well variation in key agronomic factors can be explained by the crop's reflectance. The general approach taken to this problem has been:

(1) Correlation of agronomic variables, squares of agronomic variables, reflectance in Landsat MSS and proposed thematic mapper bands, and ratios of reflectance values.

(2) Polynomial regression fitting reflectance as a function of one agronomic factor.

(3) Regressions predicting the value of an agronomic variable using the reflectances in either the Landsat MSS or proposed thematic mapper bands.

Extensive sets of ground observations were acquired at the Williston experiment station including leaf area index, percent ground cover, maturity stage, fresh and dry biomass, height, leaves per plant, stems per meter row, and green leaf weight.

First, agronomic variables were correlated with one another. The results of this study can be useful in several ways including reducing the number of factors which must be measured in the field in future experiments. The correlations presented in Tables A-4 and A-5 were calculated for the 1975 and 1976 experiment station data, respectively.

Growth stage and plant height are very highly correlated; growth stage is also highly correlated with both fresh and dry biomass and with percent green leaves. Percent ground cover is highly correlated with number of stems, leaf area index, and fresh biomass.

In general, the correlations were higher for the 1976 data. More confidence can be placed in the accuracy of the measurements made in 1976 due to the experience gained in 1975. For this reason, results from correlations of agronomic variables with reflectance will be discussed for 1976 data only.

By correlating the agronomic variables available in the two years with spectral reflectance, it was found that many agronomic factors are highly correlated with spectral response at some wavelengths: height and percent ground cover in the visible, leaf area index in the near infrared, and leaf area index and biomass in the middle infrared (Tables A-6 and A-7). More than 80% of the variation in leaf area index (leaf area per unit ground area) is accounted for by reflectance in the chlorophyll absorption region (0.63-0.69 μm) and more than 90% by reflectance in a near infrared region (0.76-0.90 μm). Correlations of reflectance with leaf area index, plant height, percent ground cover, and growth stage generally were improved by using one of the narrower bands of the proposed thematic mapper rather than a Landsat MSS band.

Tables A-6 and A-7 show the linear (Pearson) correlations of reflectance in the Landsat MSS and thematic mapper bands with several agronomic variables measured at the experiment station in 1976. Two correlations are presented in the tables: one is the correlation calculated using all data collected during the year and the other is calculated using only the data collected through wheat heading. The latter correlation is higher for many agronomic variables. The relationship between leaf area index and reflectance is approximately linear (Figures A-21 and A-22), so the two correlations are quite similar in magnitude. Reflectance of wheat throughout

Table A - 4. Linear correlations of agronomic variables (Williston, 1975).

	Growth Stage	% Cover	Height	Stems/ Meter	Leaf area index	Dry Biomass
Growth Stage	1.00*	.55	.88	.14	.31	.85
	1.00	.08	.81	.02	-.28	.86
% Cover		1.00	.70	.39	.76	.52
		1.00	.40	.42	.61	.08
Height			1.00	.12	.49	.85
			1.00	.07	.12	.81
Stems/Meter				1.00	.56	.09
				1.00	.55	.04
Leaf area index					1.00	.41
					1.00	-.03
Dry Biomass						1.00
						1.00

* Upper number is for data from seedling through heading stage of maturity; lower number is seedling through ripe stage of maturity.

Table A - 5. Linear correlations of agronomic variables (Williston, 1976).

	Growth stage	% Cover	Height	Stems/ meter	Leaf area index	Fresh bio- mass	Dry bio- mass	% Green
Growth stage	1.00*	.65	.96	.24	.58	.91	.94	-.92
	1.00	.48	.97	.14	.34	.76	.82	-.84
% Cover		1.00	.74	.78	.89	.80	.68	-.52
		1.00	.65	.74	.84	.74	.45	-.35
Height			1.00	.39	.66	.94	.95	-.92
			1.00	.28	.45	.91	.95	-.82
Stems/meter				1.00	.83	.53	.37	-.19
				1.00	.81	.50	.31	-.07
Leaf area index					1.00	.79	.62	-.46
					1.00	.65	.37	-.09
Fresh biomass						1.00	.96	-.84
						1.00	.81	-.64
Dry biomass							1.00	-.91
							1.00	-.80
% Green								1.00
								1.00

* Upper number is for data from seedling through heading stage of maturity; lower number is seedling through ripe stage of maturity.

Table A - 6. Linear correlations of agronomic factors and reflectance (Williston, 1976).

Wavelength Band	Growth Stage	Ground Cover	Height	Leaf area index	Fresh Biomass	Dry Biomass	Percent Green Leaves
.5 - .6	-.72*	-.82	-.77	-.79	-.41	-.74	.62
	-.26	-.72	-.52	-.75	-.60	-.23	.26
.6 - .7	-.69	-.90	-.75	-.85	-.43	-.71	.58
	-.03	-.66	-.40	-.82	-.43	.05	.13
.7 - .8	.40	.84	.51	.84	.21	.43	-.30
	.24	.74	.38	.79	.45	.17	-.16
.8 - 1.1	.61	.91	.70	.90	.33	.65	-.52
	.37	.84	.57	.86	.65	.29	-.32

* Top number is with data through heading; lower number is with all growth stages.

Table A - 7. Linear correlations of agronomic factors and reflectance (Williston, 1976).

Wavelength Band	Growth Stage	Ground Cover	Height	Leaf area index	Fresh Biomass	Dry Biomass	Percent Green Leaves
.45 - .52	-.64*	-.82	-.71	-.79	-.34	-.67	.55
	-.27	-.74	-.51	-.75	-.59	-.26	.24
.52 - .60	-.73	-.82	-.78	-.79	-.42	-.75	.63
	-.26	-.71	-.52	-.75	-.60	-.22	.25
.63 - .69	-.68	-.90	-.74	-.86	-.43	-.70	.56
	-.02	-.66	-.39	-.83	-.41	.07	.12
.76 - .90	.61	.93	.70	.92	.35	.64	-.50
	.31	.85	.53	.89	.61	.22	-.26
1.55 - 1.75	-.76	-.85	-.79	-.79	-.51	-.76	.64
	-.37	-.77	-.61	-.75	-.67	-.36	.31
2.08 - 2.35	-.77	-.91	-.82	-.85	-.48	-.77	.65
	-.49	-.86	-.68	-.80	-.76	-.50	.38

* Top number is with data through heading; lower number is with all growth stages.

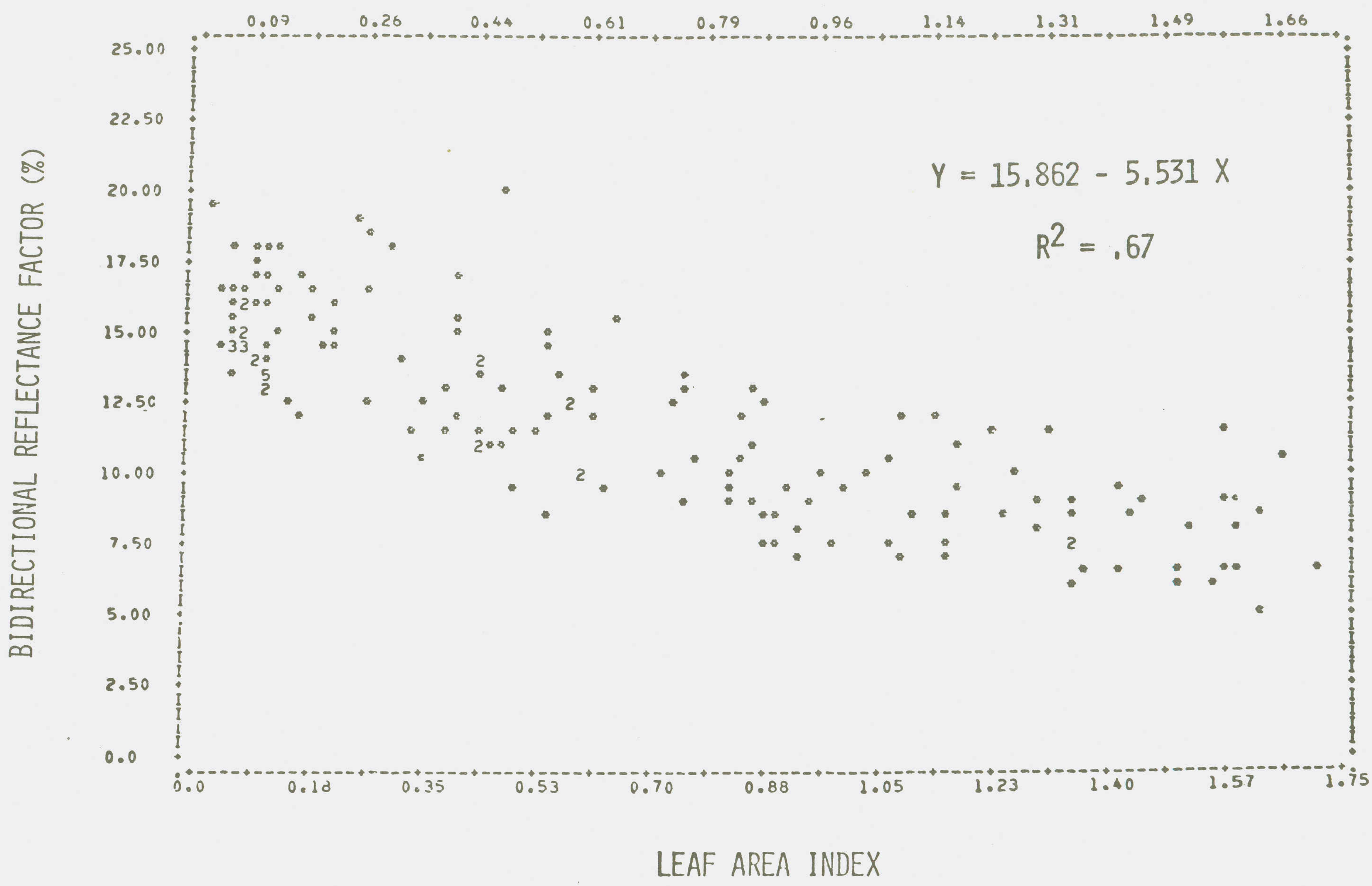


Figure A-21. Reflectance in the 0.63 to 0.69 μm band as a function of leaf area index (Williston, 1976).

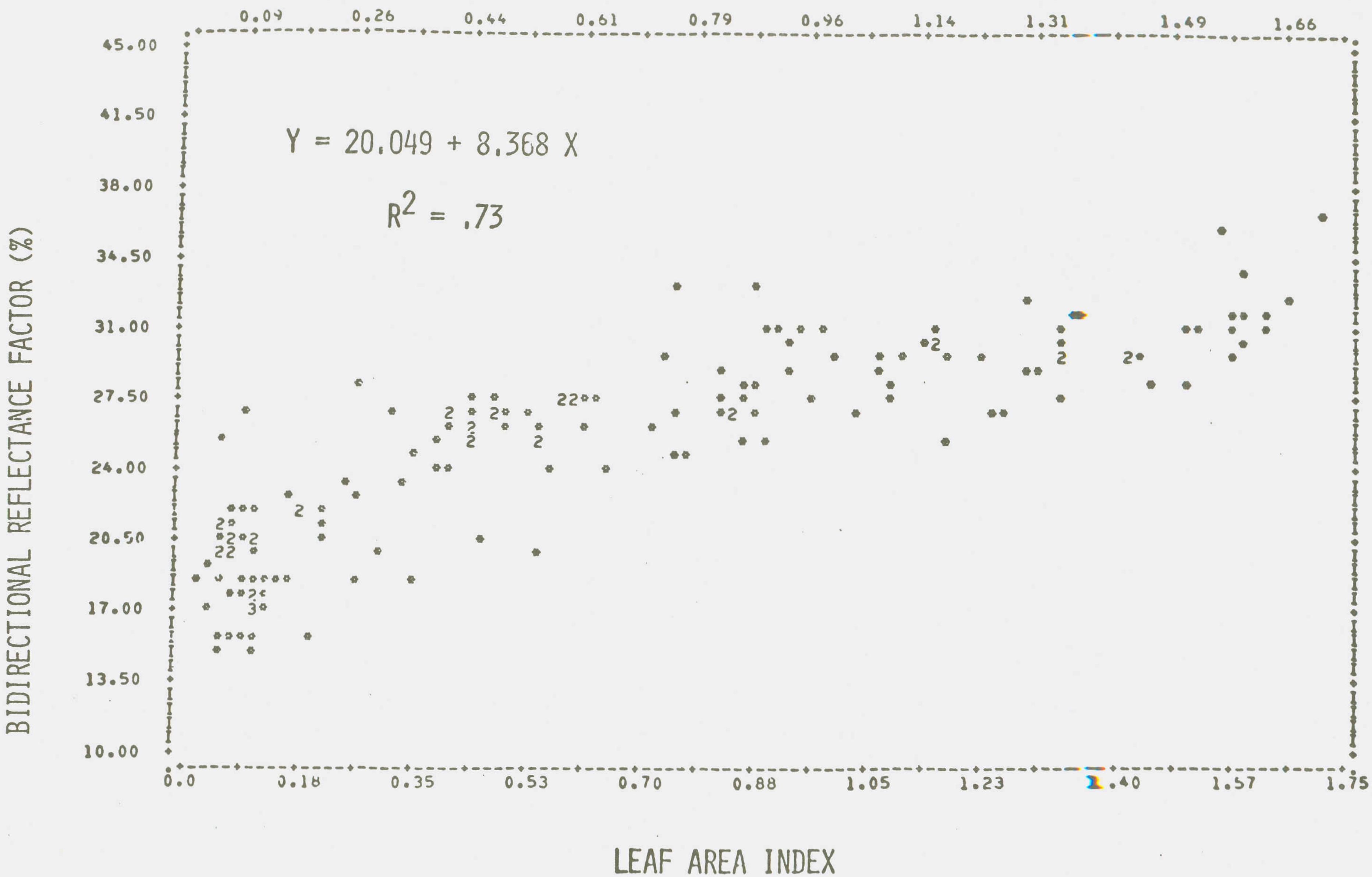


Figure A-22. Reflectance in the 0.76 to 0.90 μm band as a function of leaf area index (Williston, 1976).

its growing season typically is curvilinear in nature with the minimum (or maximum, depending on wavelength region) occurring just prior to grain ripening when the wheat has maximum leaf area and greatest plant height (Figures A -23 and A-24). This quadratic relationship accounts for the substantial increase in correlation of reflectance with growth stage using only the data acquired through heading. The change in the correlations of dry biomass with reflectance are due not to a quadratic relationship but to scatter in the data at the end of the season.

In determining polynomial regression equations to fit reflectance as a function of one agronomic factor, it was found that the relationship for most variables was linear during the measured time interval. In general, a quadratic term in the agronomic variable did not add significantly to the fit of the regression. The major exception to this was the relationship of date or growth stage to spectral reflectance which has been illustrated in Figures A -23 and A-24.

Finally, regression equations were derived to predict the value of an agronomic variable from reflectance values in one or more wavelength bands. All possible regressions were run to select an "optimal" subset of wavelength bands for prediction of each agronomic variable. In the prediction of percent ground cover, four variables appeared to be the best subset size for prediction and the four bands giving the highest R^2 (.91) were .45-.52, .52-.60, .76-.90, and 2.08-2.35. To predict leaf area index, however, the inclusion of only four variables yields a slightly biased regression equation, requiring five variables to achieve unbiasedness. If four variables are selected, they are .63-.69, .76-.90, 1.55-1.75, and 2.08-2.35. If five variables are selected, they are .52-.60, .63-.69, .76-.90, 1.55-1.75, and 2.08-2.35. This analysis demonstrates the importance of having information from the middle infrared region of the spectrum to predict agronomic factors. An analysis to compare these results with those which can be obtained using the four Landsat bands is in progress.

A similar study to relate agronomic factors and reflectance is being conducted using the 1975 data from the Williams County ITS. Only percent ground cover, growth stage, and height were measured. Some preliminary results indicate that the correlations between these variables and reflectance is not as high as with the AES data. Analysis of this data is still in progress.

Further studies to relate agronomic factors and reflectance are being carried out using multi-year data from the experiment station. AES data from 1977 and ITS data from both 1976 and 1977 will be analyzed. Investigation of the correlation of agronomic variables with ratios and other transformations of reflectance values is in progress.

Effects of Cultural and Environmental Variables. Only 1975 data from the Williston experiment station have been used to date to investigate effects of cultural and environmental variables on the spectral response of spring wheat. The results must be considered only indicative of effects present due to the confounding effect of time of day which was found to be an important source of variation in initial studies.

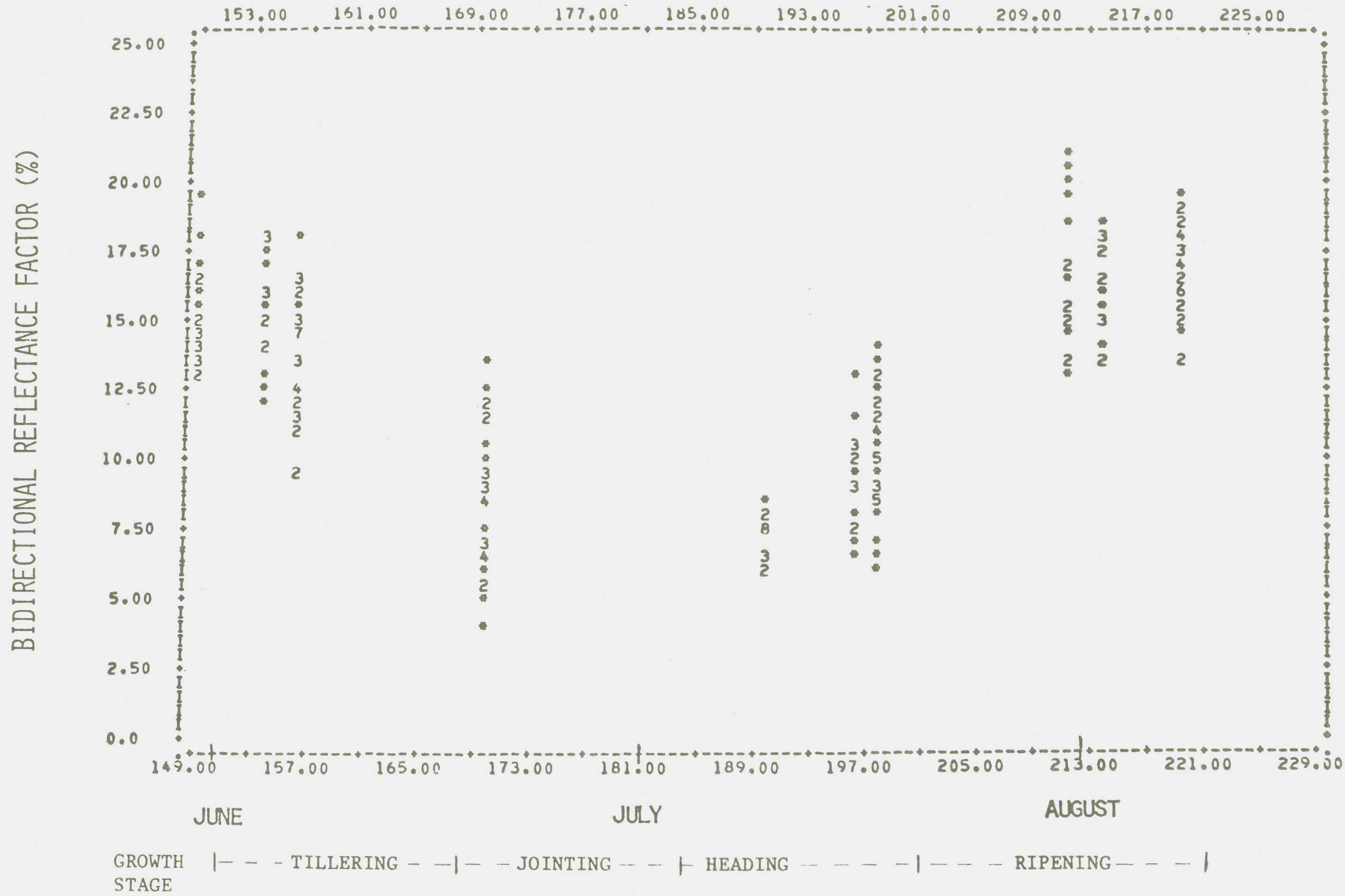


Figure A-23. Reflectance in the 0.63 to 0.69 μm band as a function of date (Williston, 1976).

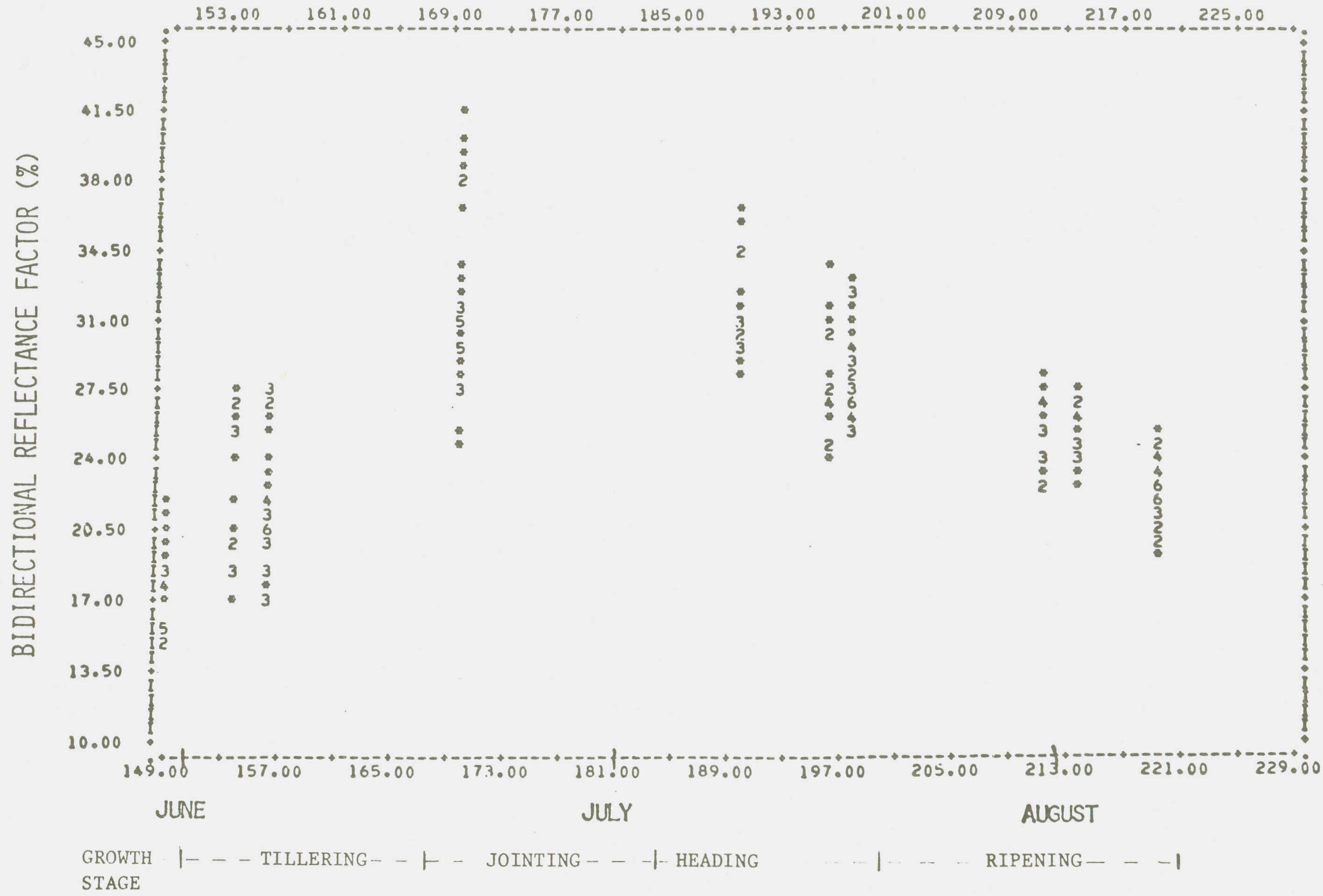


Figure A-24. Reflectance in the 0.76 to 0.90 μm band as a function of date (Williston, 1976).

There appeared to be a significant difference on all dates and at all wavelengths between wheat planted on recrop and fallow land; however, this factor is completely confounded with measurement order and blocks. Differences in variety appeared in the near and middle infrared regions of the spectrum. The effect of nitrogen was highly significant on July 10, when the wheat was heading. Effects of planting date seem to be significant at some wavelengths throughout the season.

Analyses of data acquired in two other years are being performed for confirmation of these results. In addition, any information acquired in 1977 about time of day effects will increase confidence in results of this nature.

Discriminability of Spring Wheat from Other Cover Types. FSS data acquired on the ITS in 1975 have been analyzed to determine the discriminability of spring wheat from other cover types. Three cover types had sufficient data available to permit statistical analysis: spring wheat, pasture, and fallow. The discriminability of these three cover types was assessed in first the univariate (single bands) and then the multivariate (multiband) case.

A univariate analysis of variance was performed on spectral reflectance for each Landsat MSS and proposed thematic mapper band for the nine measurement dates. Cover type produced a significant effect at the 1% level for nearly all times and wavelength bands indicating that spring wheat, pasture, and fallow could be easily identified at particular growth stages.

Subsequently, univariate multicomparison analysis were carried out to determine which cover types were distinguishable at particular maturity stages and wavelengths. Newman-Keuls range tests at the 5% significance level were used for testing. Results indicated that wheat and fallow were not separable at any wavelengths in June as the wheat still resembled bare soil. By July 10, all three cover types were separable in the near infrared wavelengths. Later in July, all three were differentiable at most wavelengths. In mid to late August, wheat and pasture looked alike in the middle infrared, but were separable elsewhere. These results are illustrated in Figures A-25 to A-28.

Discriminant analysis was performed to use multiband information in assessing the separability of these three cover types. The percent correct identification of wheat obtained with this data set probably is an upper bound for that obtainable at the same growth stage with satellite data. The method used (linear discriminant functions with three classes) should provide a lower bound on discriminability for this data set which could be improved by using higher order discriminant functions and more spectral subclasses. Prior probabilities were chosen proportional to the amount of training data in each category.

The first analysis performed used all the data for deriving the discriminant functions and the same set of data for assessing the discriminability (Figure A-29). The percentage of wheat scans which were correctly

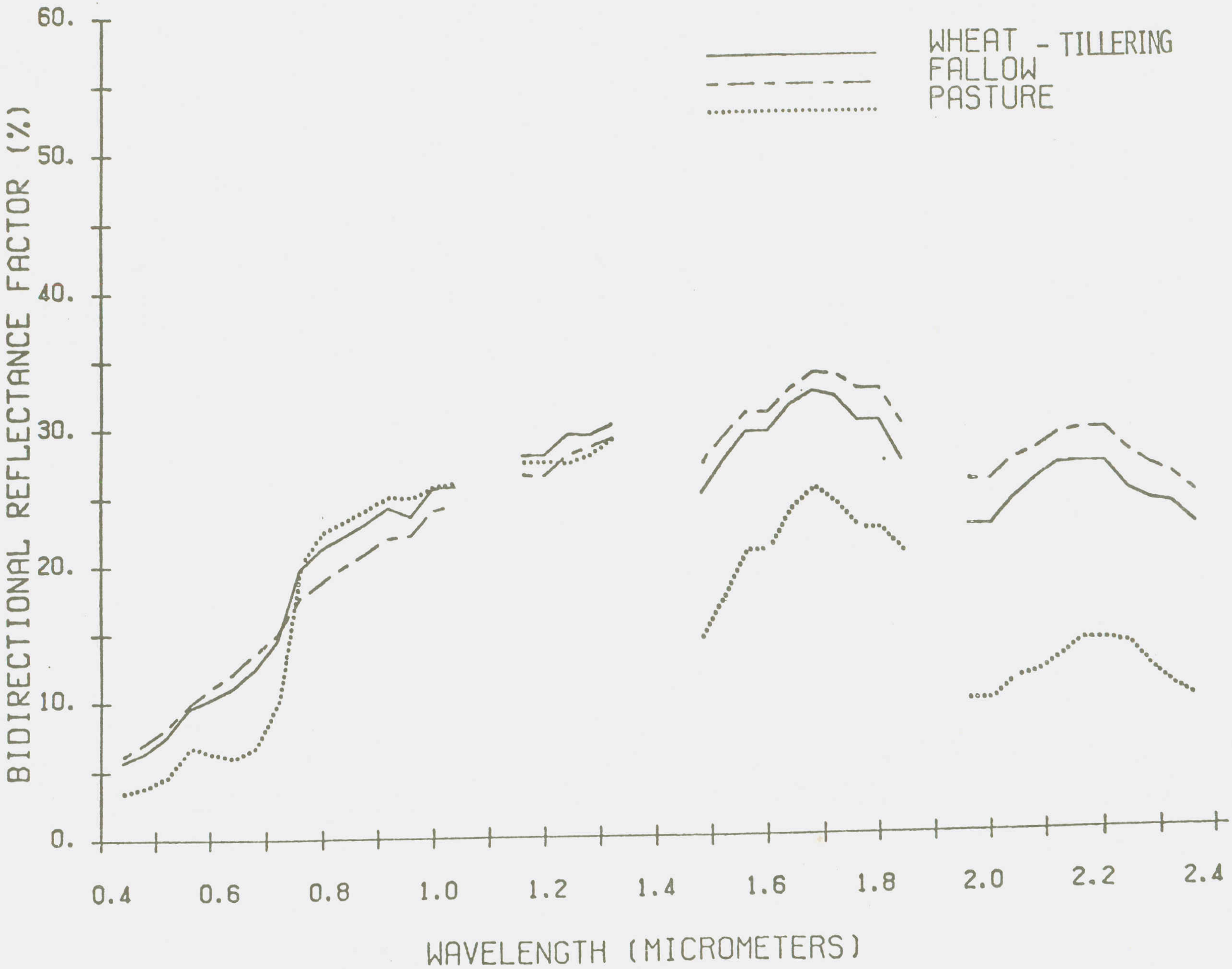


Figure A-25. Comparison of reflectance of spring wheat, pasture, and fallow (Williams County, North Dakota; June 22, 1975).

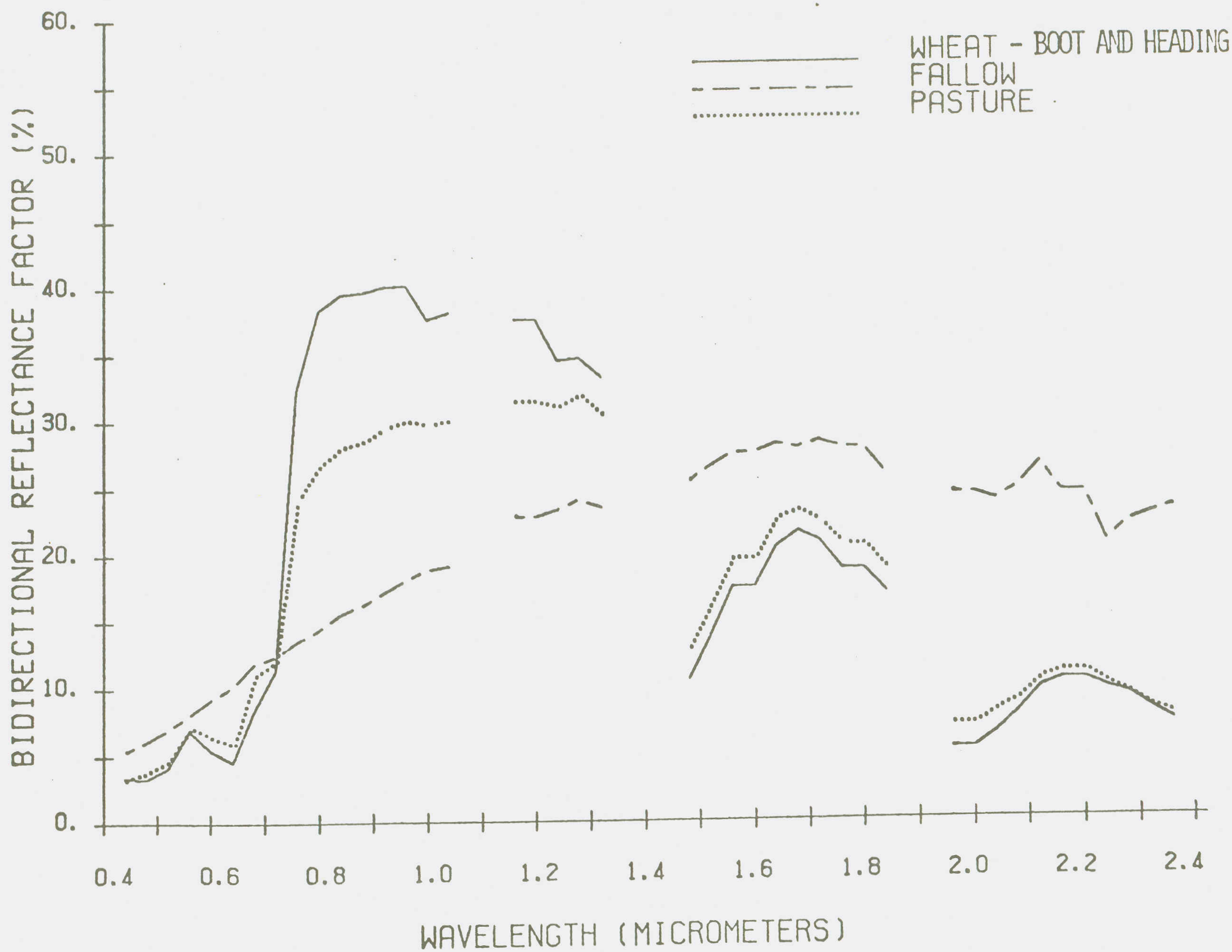


Figure A -26. Comparison of reflectance of spring wheat, pasture, and fallow (Williams County, North Dakota; July 10, 1975).

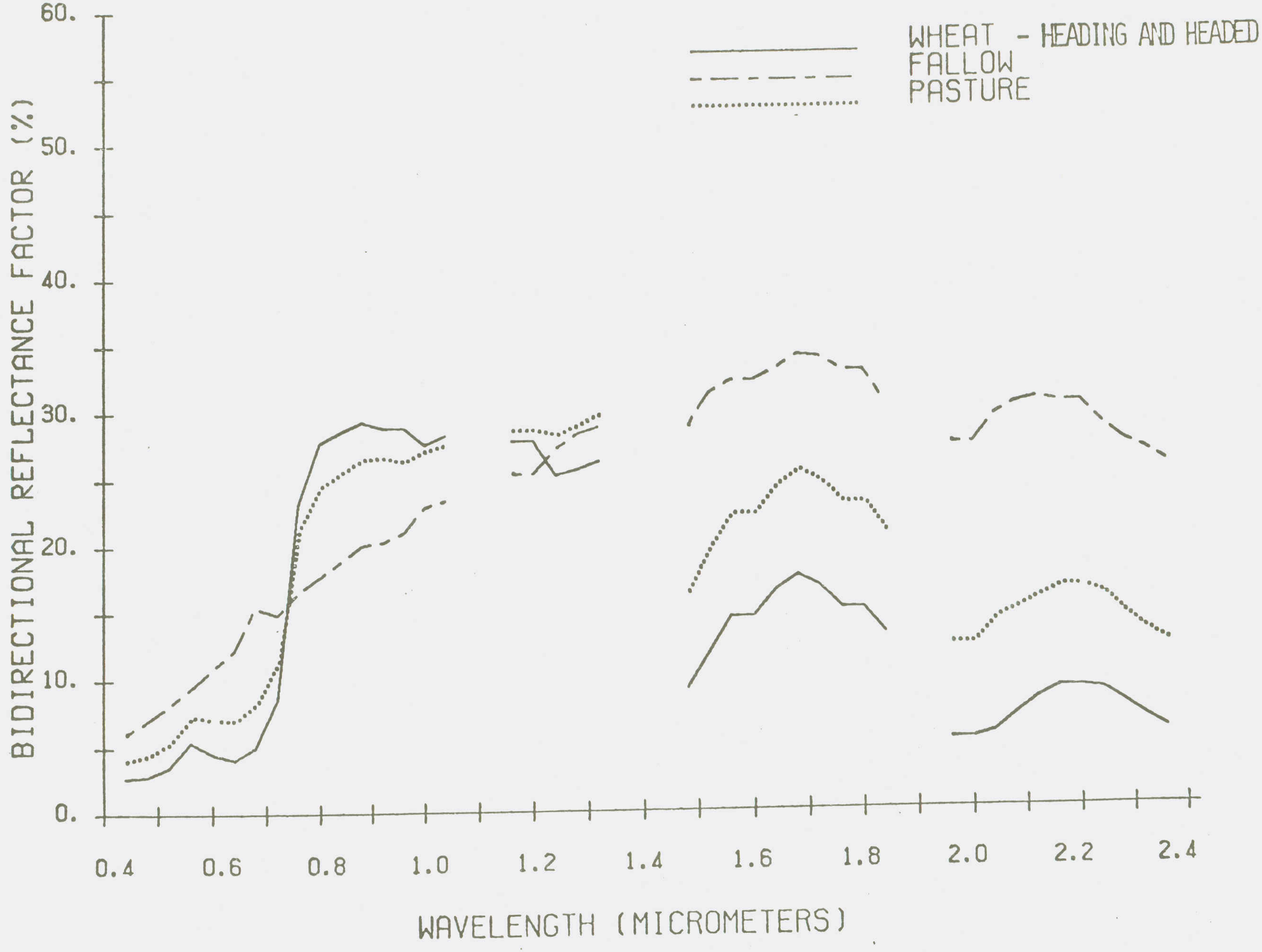


Figure A-27. Comparison of reflectance of spring wheat, pasture, and fallow (Williams County, North Dakota; July 18, 1975).

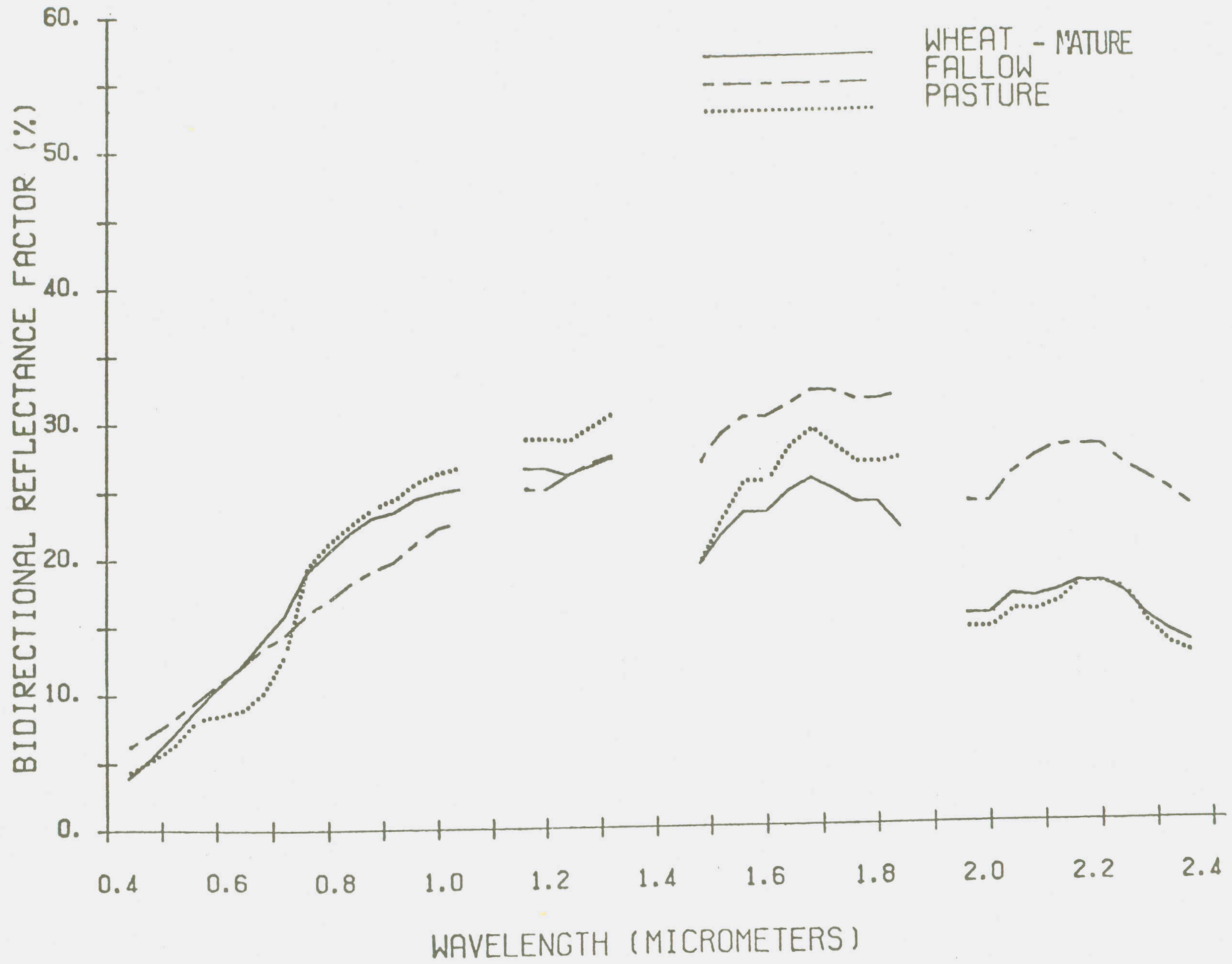


Figure A-28. Comparison of reflectance of spring wheat, pasture, and fallow (Williams County, North Dakota; August 23, 1975).

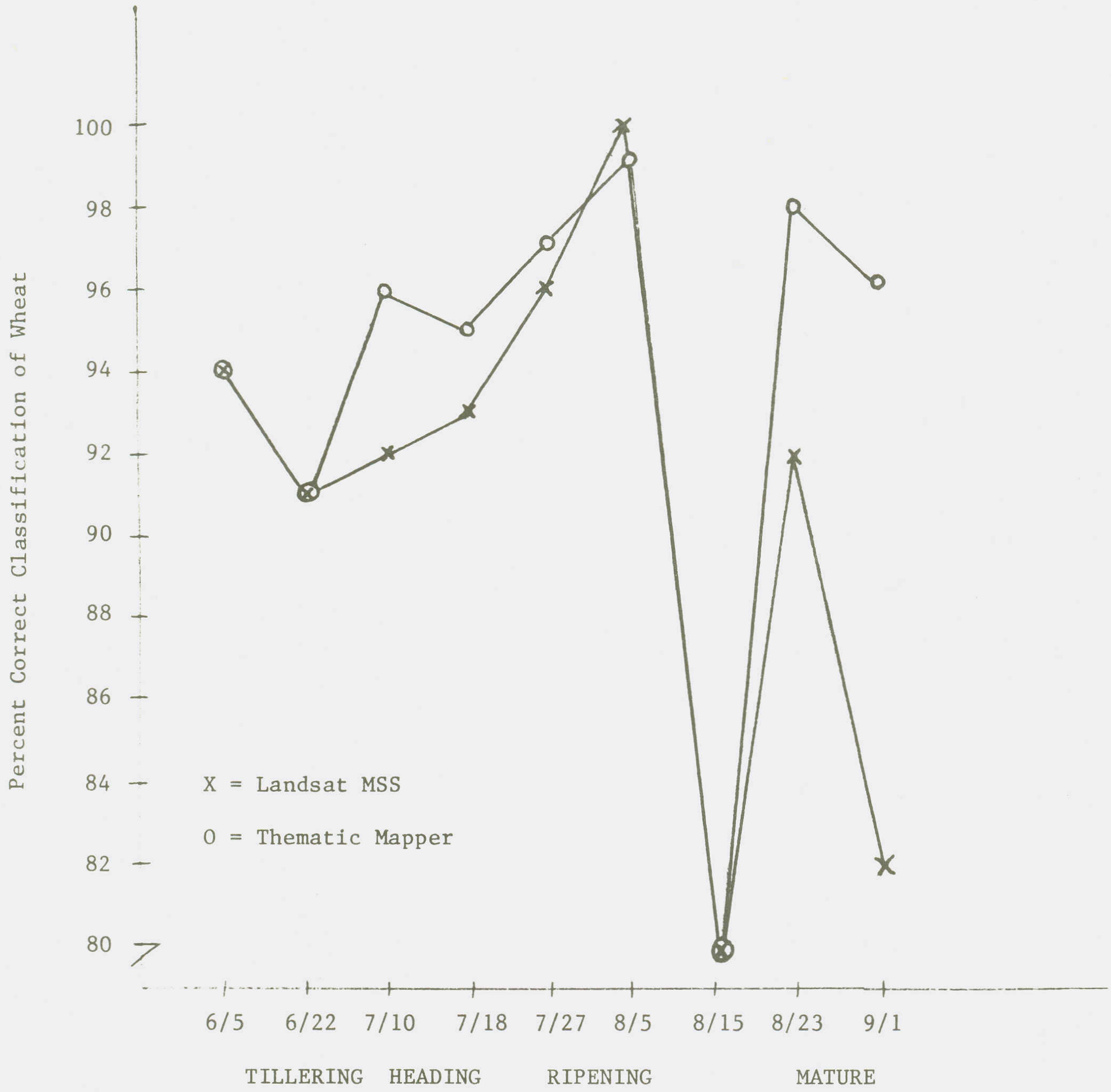


Figure A-29. Discriminability of wheat on nine dates using Landsat MSS and proposed thematic mapper bands.

classified generally increased with maturity stage until the wheat reached the ripe stage for both the Landsat MSS and thematic mapper bands. The high percentage of wheat correctly identified on August 5 is considered to be an artifact of the small data set acquired on that date. On August 15, only 80% of the wheat was correctly classified. This low figure was due to the wide range of variability present in the wheat at this time. The six proposed thematic mapper bands performed about the same as the Landsat MSS bands for wheat identification until late in the season when the thematic mapper bands appeared to have greater power for discrimination. The reason for this can be observed in Figure A-28; separability of wheat from fallow and pasture at this time is achieved in the middle infrared region of the spectrum which is not available in the current Landsat satellite. On the average by this method, over 90% of the wheat scans were correctly classified on each date.

On June 5 and 22, although most wheat is correctly identified, many fallow scans are incorrectly classified as wheat. The proposed thematic mapper bands exhibit a distinct advantage over the Landsat MSS bands in avoiding errors of commission (other cover types being incorrectly identified as wheat). From July on, pasture and fallow are each identified with high accuracy (85-100%) by the thematic mapper bands while the corresponding accuracies with Landsat MSS bands are only 64-98%.

Discriminant analyses were also performed using disjoint training and test sets and several training methods. Some results (Table A-8) show reasonable accuracies are achieved by three methods of training: a 50% sample of fields, a 25% sample of fields, and a systematic sample of 25% of scans. Systematic sampling showed generally higher accuracies than sampling by fields and larger samples generally achieved higher accuracies.

Good discriminability can be achieved among these cover types at many individual dates, but if spectral information is available at more than one maturity stage of the wheat, even better results can be obtained (Table A-9). June 22 and August 15 together do not provide sufficient information for discriminability as some information during the middle of the growing season is required. Most other combinations of two or three measurement dates provide accurate identification of wheat, with pasture and fallow generally being identified with 95-99% accuracy.

This objective cannot be pursued with data from the Williston AES, but analyses are planned and underway to investigate the discriminability of spring wheat from other crops on the intensive test site in both 1976 and 1977.

Discriminability of Spring Wheat from Other Small Grains. The data currently available for analysis were inadequate for drawing definite conclusions concerning the separability of spring wheat from other small grains although some indications of separability could be obtained.

In 1975 on the intensive test site, only two fields of rye were available on all dates and one field of oats was measured on two dates.

Table A - 8. Proportion of wheat scans correctly identified by several training methods.

Mission Date	Method 50% sample of fields	25% sample of fields	25% sample of scans
6/5	79*	98	98
	80	85	97
6/22	80	62	84
	84	65	87
7/10	96	87	97
	98	93	97
7/18	98	94	93
	98	99	95
7/27	99	83	94
	98	91	96
8/5	99	96	100
	97	98	100
8/15	64	77	75
	68	79	78
8/23	88	89	88
	89	94	98

* Top number is using Landsat MSS bands; lower number is using proposed thematic mapper bands.

Table A-9. Discriminability of spring wheat, fallow, and pasture using information from two or three growth stages.

Dates	Percent Wheat Correct	
	Landsat MSS	Thematic Mapper
6/22 and 8/15	80	84
7/10 and 8/23	96	99
6/5, 7/27, and 8/15	97	99
6/22, 7/18, and 8/23	98	99

Table A-10. Correlations of agronomic factors and reflectance (Garden City, 1975).

Wavelength Band	Data To Ripening Used		All Data Used	
	Height	Ground Cover	Height	Ground Cover
.5-.6	-.80	-.83	-.47	-.61
.6-.7	-.80	-.87	-.38	-.55
.7-.8	.29	.57	-.03	.27
.8-1.1	.42	.68	.05	.35
.45-.52	-.80	-.87	-.50	-.66
.52-.6	-.80	-.82	-.47	-.61
.63-.69	-.79	-.88	-.36	-.55
.76-.9	.48	.72	.10	.40
1.55-1.75	-.82	-.76	-.71	-.70
2.08-2.35	-.80	-.87	-.68	-.79

These data are insufficient for a sound statistical assessment of discriminability, although qualitative evaluations of the data showed that some discriminability may be possible if there are differences in maturity stage.

One problem encountered in analysis of the AES data was that differences between small grains were confounded with effects due to the time of day measurements were acquired. The results of an initial investigation on spring wheat clearly indicated that the time interval required to measure all plots of an experiment is an important source of variation and should be minimized. Knowledge was gained which enabled a better experimental design for the third year. Time elapsed during measurement of an experiment was minimized and data were acquired several different times during the day on each plot which permits a covariance analysis.

2. Winter Wheat in Kansas

This section describes the analyses which have been performed on the agriculture experiment station and intensive test site data from Kansas. For each objective, the specific approach and results are discussed.

Relationship of Agronomic Variables to Reflectance. Two agronomic variables, height and percent ground cover, were measured at the Garden City experiment station in 1975. These two variables were highly correlated (R = .86) using only green data, but high correlations should not occur if all data throughout the growing season are used and measurements are acquired properly.

Correlations of height and percent ground cover with reflectance in each of the Landsat MSS and proposed thematic mapper bands were computed (Table A-10). Both variables show high correlations with reflectance in the visible and middle infrared wavelengths, with correlations being higher when only data through heading were used.

In deriving polynomial regression equations to fit reflectance as a function of one agronomic factor, it was found that the relationships were approximately linear in the measured time interval. Illustrations of reflectance in the visible and near infrared regions as a function of percent ground cover are given in Figures A-30 and A-31. Date or growth stage is related to reflectance in a quadratic manner as illustrated in Figures A-32 and A-33.

Finally, regression equations were determined to predict the value of an agronomic variable from reflectance in one or more wavelength bands. All possible regressions were run to select an "optimal" subset of wavelength bands for prediction of each agronomic variable. Two thematic mapper bands gave a higher R^2 for prediction of both height and percent ground cover than did all four Landsat MSS bands. This improvement appears to be due to the presence of a middle infrared band which is always selected as important. The R^2 values are generally near maximum for about four thematic mapper bands and the regression equation becomes relatively unbiased for either four or five bands.

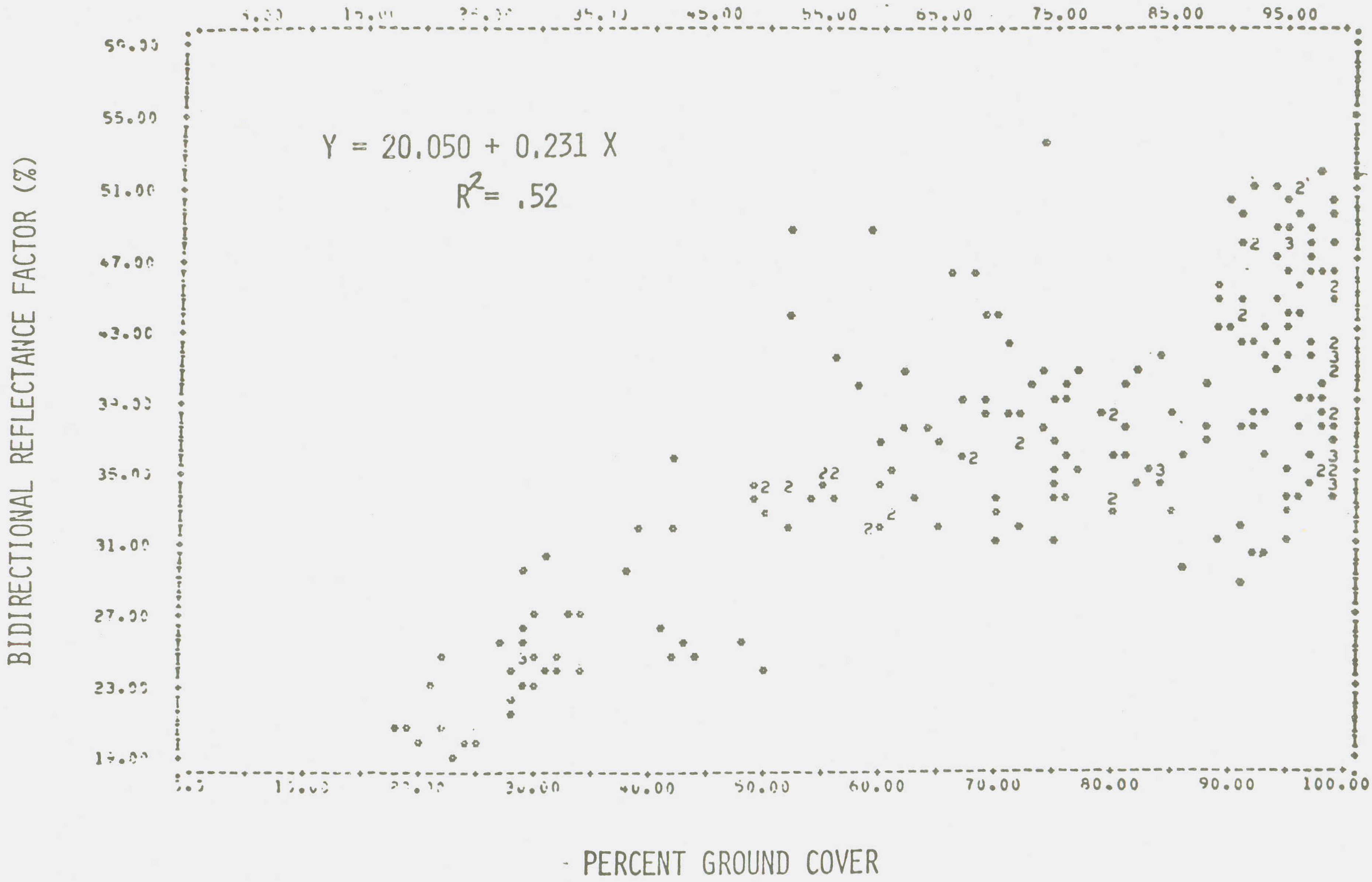


Figure **A-30.** Reflectance in the 0.63 to 0.69 μm band as a function of percent ground cover (Garden City, 1975).

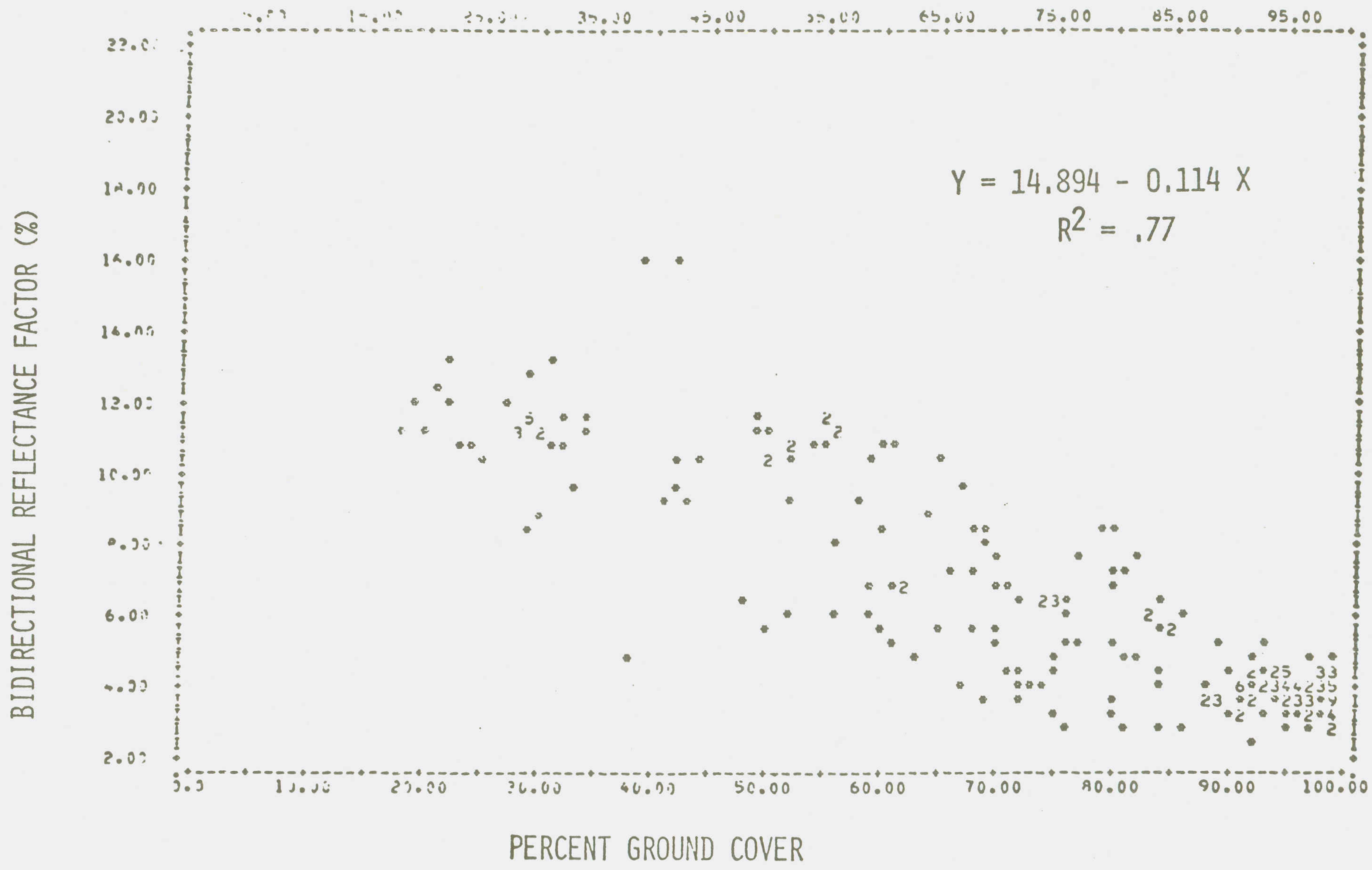


Figure A -31. Reflectance in the 0.76 to 0.90 μm band as a function of percent ground cover (Garden City, 1975).

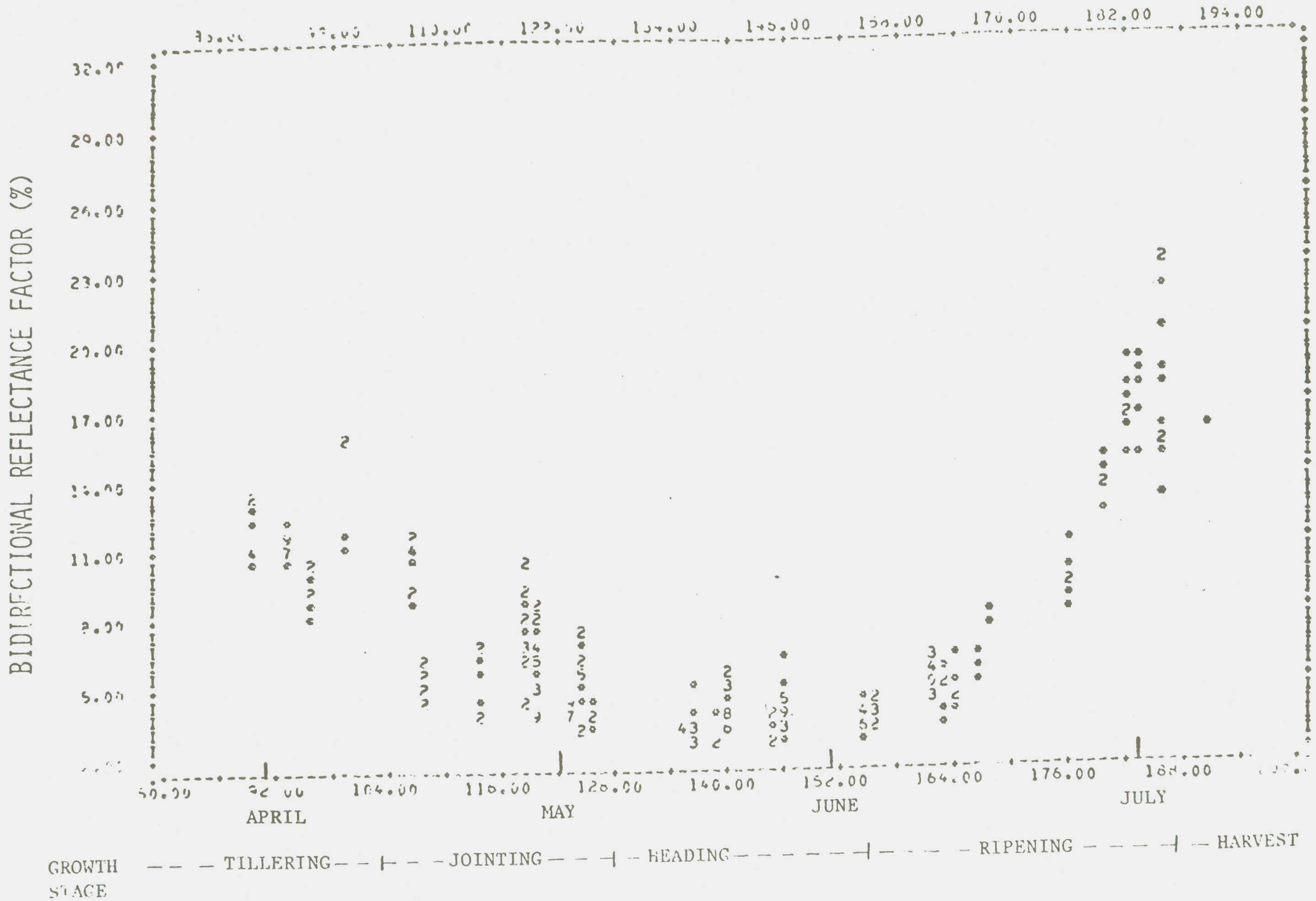


Figure A-32. Reflectance in the 0.63 to 0.69 μm band as a function of date (Garden City, 1975).

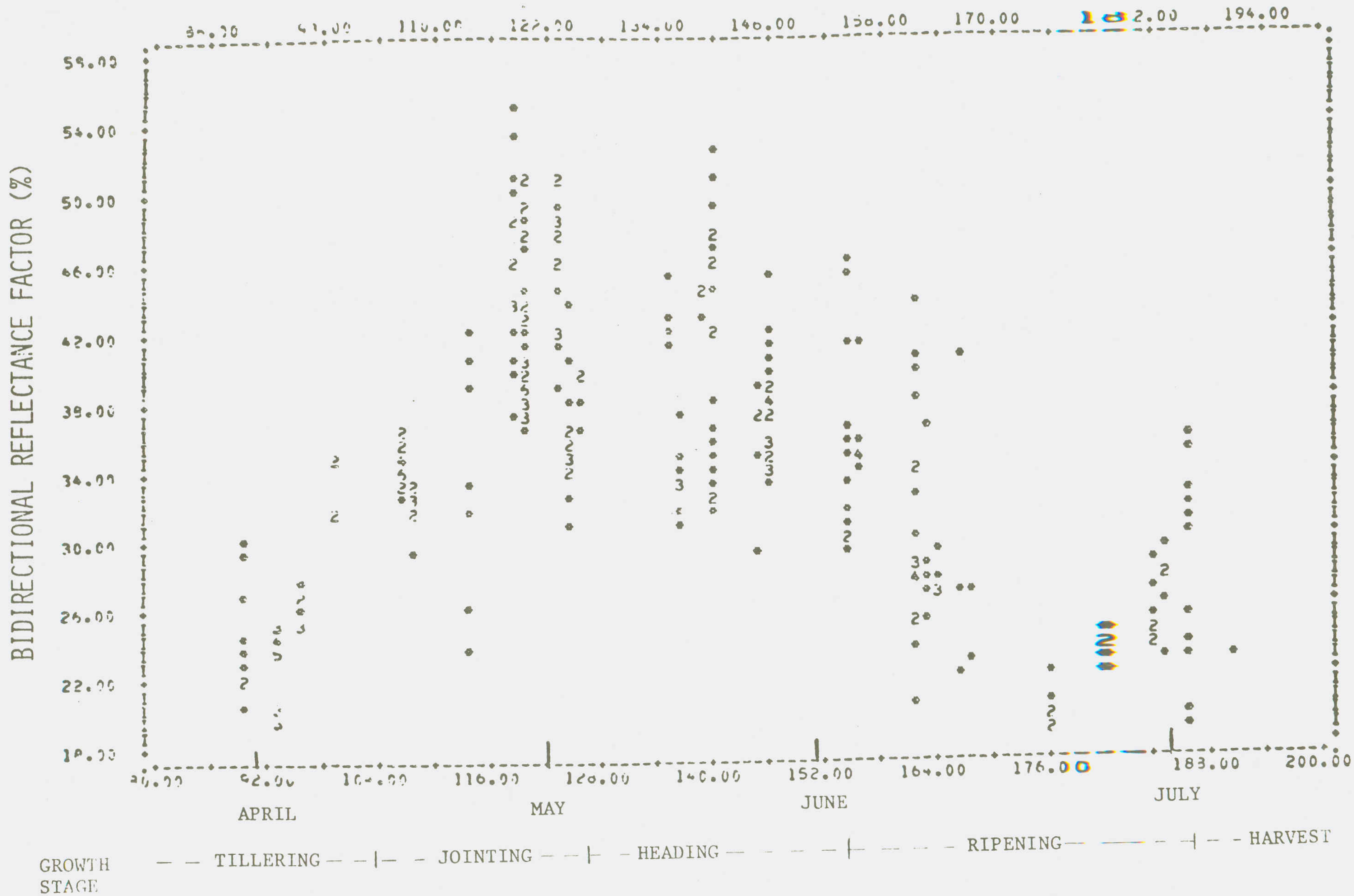
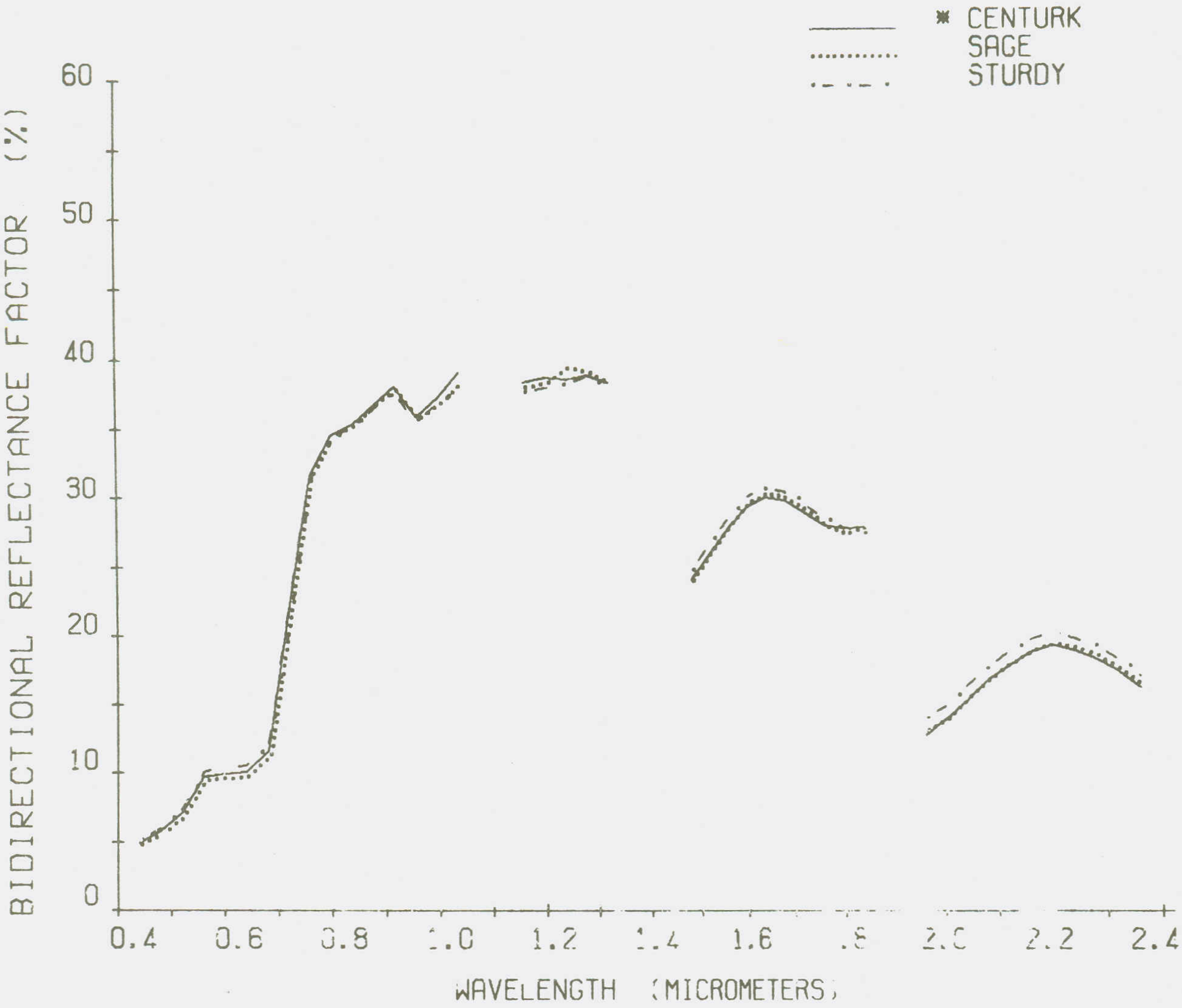
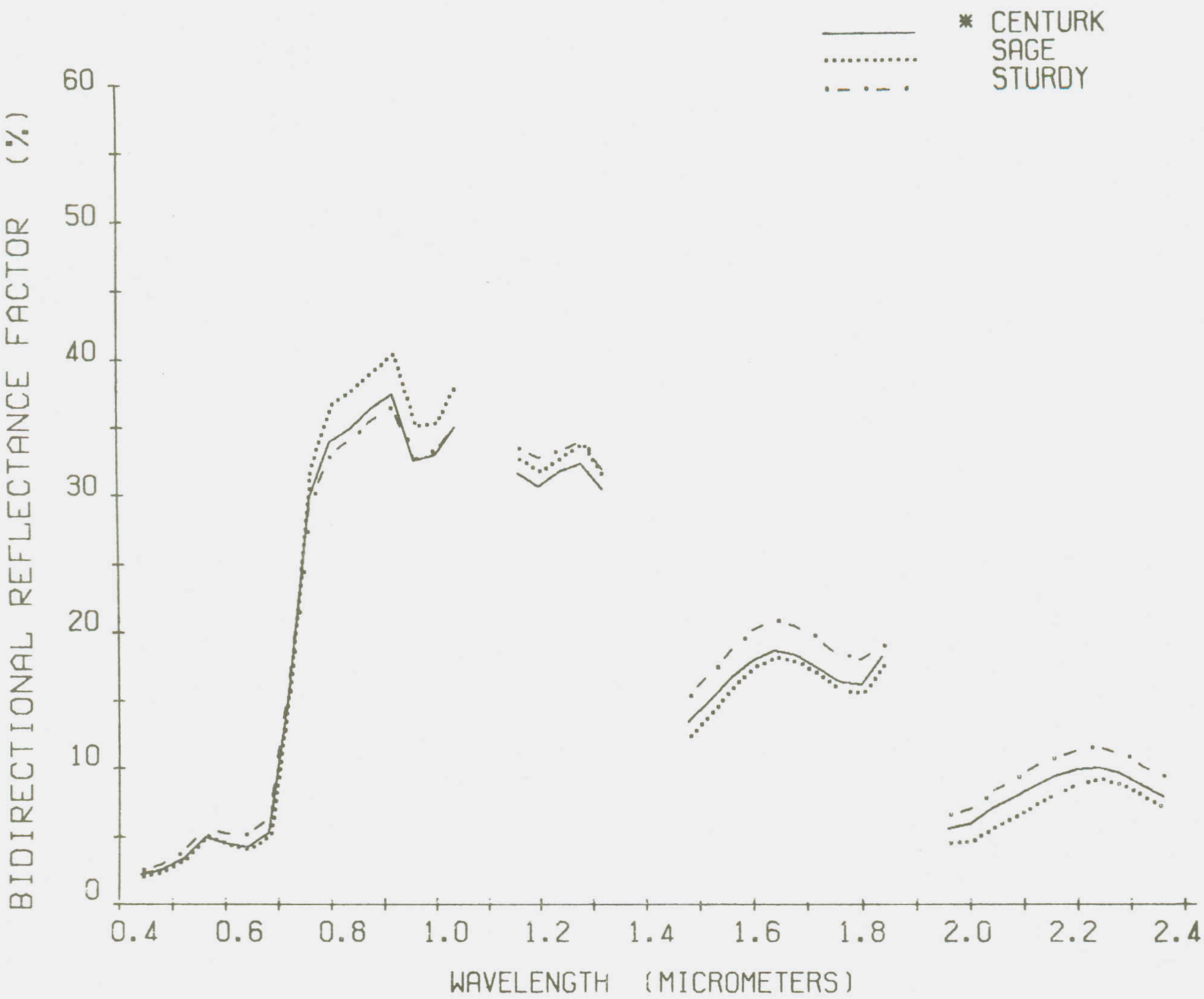


Figure A-33. Reflectance in the 0.76 to 0.90 μm band as a function of date (Garden City, 1975).



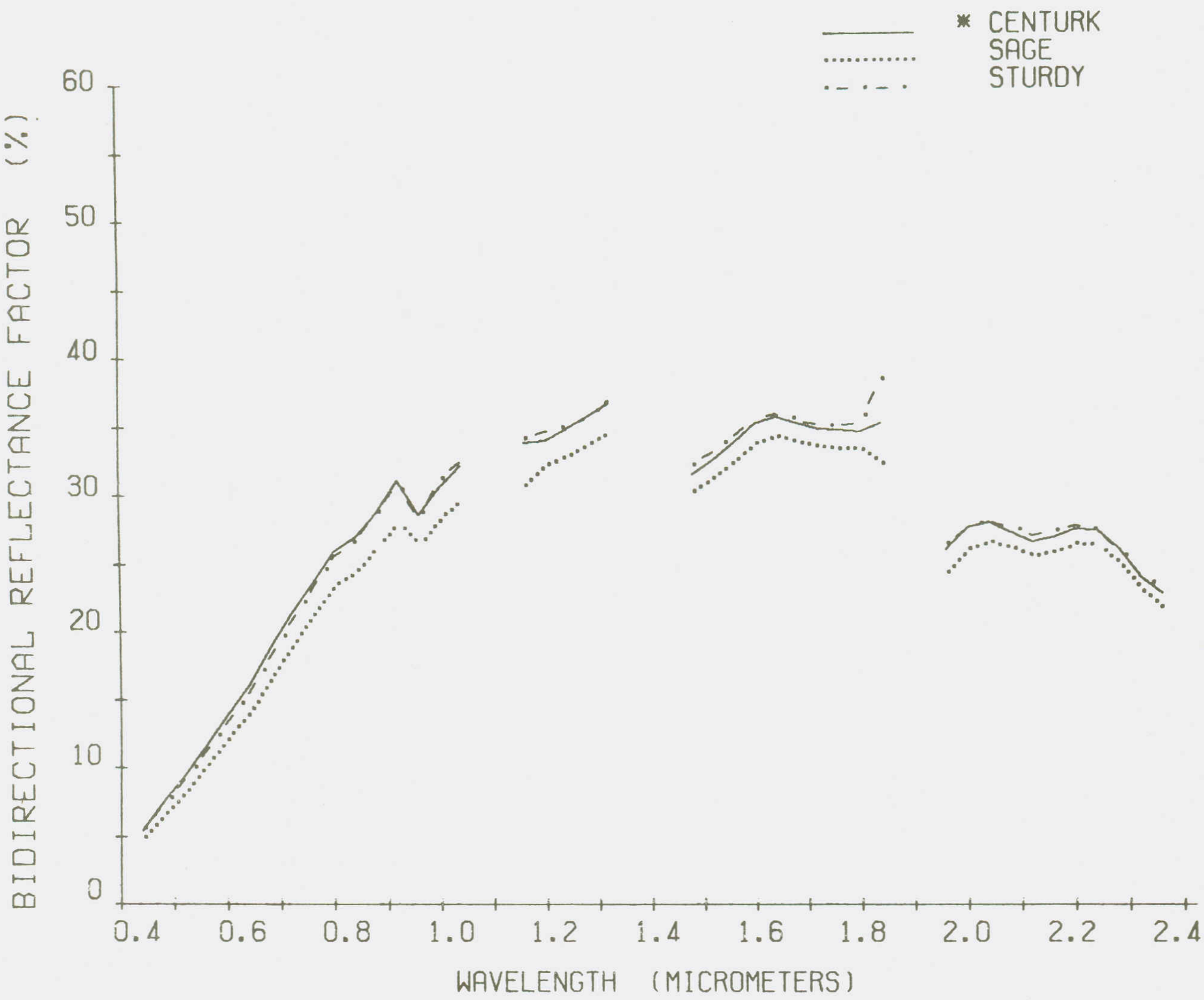
* AVERAGES OF 6 PLOTS.

Figure A-34. Reflectance of several varieties of winter wheat (Finney County, Kansas; April 16, 1975).



* AVERAGES OF 3, 2, AND 4 PLOTS, RESPECTIVELY.

Figure A -35 . Reflectance of several varieties of winter wheat (Finney County, Kansas; May 20, 1975).



* AVERAGES OF 2, 3, AND 3 PLOTS, RESPECTIVELY.

Figure A-36 . Reflectance of several varieties of winter wheat (Finney County, Kansas; July 4, 1975).

Analysis is currently underway for the 1976 experiment station data. No analyses to relate reflectance with agronomic factors have yet been conducted with intensive test site data.

Effects of Cultural and Environmental Variables. Statistical evaluation of the effects of cultural and environmental variables was not feasible due to the incomplete designs obtained and the confounding effect of time of day. Qualitative evaluations were performed for several variables such as wheat variety, nitrogen level, and residue management. Figures A-34 to A-36 show that there is little difference in the spectral response of three varieties of wheat throughout the growing season.

Discriminability of Winter Wheat from Other Cover Types. FSS data acquired in 1975 over the Finney County ITS was analyzed to investigate the separability of wheat from other crops. All results obtained are considered very preliminary and are not presented for that reason. A greater understanding of this problem should be achieved through analysis of the 1976 and 1977 FSS data which are believed to be of higher quality and over a more representative test site.

Discriminability of Spring Wheat from Other Small Grains. Only data from the Garden City experiment station supported this analysis objective as no small grains were measured on the ITS in either 1975 or 1976. Again, due to time of day confounding and incomplete data sets, no quantitative statistical evaluations were carried out. Even qualitative indications must be regarded as very preliminary due to the difficulties mentioned above and the limited sample. It appears that early planted wheat and rye may be identified at some times due to differences in growth stage, but that other small grains are fairly similar to one another.

d. Plans for Future Analysis

The results which have been obtained to date in the LACIE Field Measurements data analysis task must be considered preliminary. Before evaluation of the results from the project can be made, analysis should be completed on the data from all three years of the measurement program. In particular, from the knowledge gained to date, the following analyses have been planned:

(1) Conduct similar analyses to those described here on the complete data sets for three years from the experiment stations and intensive test sites.

(2) Study in greater depth the effect of time of day on spectral response.

(3) Investigate the use of basis functions to represent and discriminate spectra.

Analyses of the three years of data acquired will continue to investigate the problems of relationship of agronomic variables to reflectance, differences in reflectance due to cultural and environmental variables, and the detection and discriminability of wheat from other small grains and cover types. Statistical techniques including analysis of variance, regression, correlation, and discriminant analysis will be used. The purpose of these continued analyses is to confirm the results which have been obtained to date and to seek an understanding of how spectral response changes between locations and between years in the same location.

Secondly, an analysis will be made of the data acquired during 1977 by the Exotech Model 100 at the Williston experiment station. Analysis of this data should provide an understanding of how reflectance varies according to the time of day and should enable a more accurate analysis to be made of data acquired by the field spectrometer systems where a substantial time interval may elapse during the course of measurement acquisition.

Investigation of the use of basis functions to represent and discriminate spectra is already underway. The basis function approach shows much promise theoretically in that it permits the use of a greater portion of the available spectral information. The approach needs to be tested in practice, however, and physical interpretations for the basis vectors defined.

V. Conclusions and Recommendations

The LACIE Field Measurements project, directed by Purdue/LARS, has successfully acquired a substantial amount of high quality spectral measurements during the past three years. A majority of the data have been processed and provided to investigators who are now using it in research programs. The spectral data are supported by numerous and, in many ways, quantitative agronomic and meteorological measurements and observations. Together the spectral, agronomic, and meteorological data form what is undoubtedly the most comprehensive data set available for agricultural remote sensing research. Analysis of these data are now well underway and the results should help provide a basis for the continued development of remote sensing technology.

One of the key aspects of the LACIE Field Measurements spectral data is that they are radiometrically calibrated. Calibration enables valid comparisons of measurements from different dates, sensors, and/or locations. The procedures for field calibration of spectrometer data have been tested and refined during this period and the knowledge gained can now be applied in future investigations. A large amount of information on how to successfully obtain reliable spectral measurements was acquired and is being documented.

A system encompassing all phases of experiment planning, data acquisition, data processing, and data storage and retrieval has been developed and successfully operated. This capability will be valuable for future research programs. The experience and knowledge gained from the LACIE Field Measurements project can be the foundation for the field measurements research required for the development of a global food and fiber information system.

Although the project acquired a large number of spectra, the sample of crop, soil, and weather conditions over which the data were acquired was small even for wheat. Each of the three years in each site was different in terms of the weather and the crop response to it. However, the crop cannot be treated as a constant even if the weather did not vary significantly from year to year. Changing economic conditions and advancements in agricultural technology will bring changes in crop and soil management (e.g., minimum tillage) and even the crop itself (e.g., introduction of semi-dwarf varieties of wheat). Measurements of wheat and its confusion crops should be continued over additional years if the full potential of the current efforts is to be achieved.

And, as we look ahead to the development of a global food and fiber information system utilizing remote sensing techniques it is critical to begin to make field measurements for crops other than wheat. One of the lessons which should come from the LACIE Field Measurements project, is the importance of establishing viable field measurements data acquisition, processing, and analysis efforts before results upon which to base the design of a large-scale effort are needed. Five years for planning experiments and acquiring and analyzing data would not be too long since at least three years of data should be available for analysis to obtain the information needed for system design.

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, processing, and analysis of field measurements data will be needed since it will not be feasible to simply expand upon the LACIE approach as might be indicated by the increased number of crops which should be included in future efforts.

An approach similar to the one tested by LARS during the past summer is recommended. Use of a simpler instrumentation system (i.e., multiband radiometer) will make the collection, processing, and analysis of spectral measurements more economical and will allow more measurements to be made. Such instruments could be utilized at more sites than was possible with the currently available high spectral resolution spectrometer systems. And, it is more observations of crops and soils under a variety of different conditions (not detailed measurements of a limited number of locations and crop conditions) that are really needed to increase our understanding of the spectral characteristics of agricultural scenes. There will be a continuing need for the high resolution spectrometer system to be utilized in field research, but greater emphasis should be placed on systems which can be used to economically and accurately make large numbers of measurements.

Analysis of the LACIE Field Measurements data, although not completed, is beginning to provide new knowledge and understanding of the spectral characterisitcs of wheat and the biological and physical factors affecting spectral response. For example, the relation of spectral reflectance to leaf area index and biomass has been quantified. This is important to understanding when in its development wheat can be identified using remotely sensed data and what the condition and potential yield of the crop may be. Analysis of the data should be pursued with resources commensurate with the amount of data collected if the potential benefits of the LACIE Field Measurements project are to be fully realized.

VI. References

1. LACIE Field Measurements Project Plan. 1974-75 and 1975-76.

2. LACIE Field Measurements Project Plan. 1976-77.

3. DeWitt, D.P. and B.F. Robinson. 1976. Description and Evaluation of a Bidirectional Reflectance Factor Reflectometer. Information Note 091576, Laboratory for Applications of Remote Sensing: Purdue University.

4. LACIE Field Measurements Data Library Catalogs, Volume 1: 1974-75 Crop Year; Volume 2: 1975-76 Crop Year; Volume 3: 1976-77 Crop Year (Latest update - August 1977).

5. Bauer, M.E., W.R. Simmons, and B.F. Robinson. 1976. LACIE Field Measurements Project Status Report, Laboratory for Applications of Remote Sensing, Purdue University.

6. Biehl, L.L., B.F. Robinson, W.R. Simmons, and M.E. Bauer. 1976. LACIE Field Measurements Data Qualtiy Evaluation and Verification Recommendations. Report prepared for NASA/ Johnson Space Center by Laboratory for Applications of Remote Sensing, Purdue University.

7. Landgrebe, D.A., L.L. Biehl, and W.R. Simmons. 1977. An Empirical Study of Scanner System Parameters. IEEE Transactions on Geoscience Electronics, Vol. GE-15. No. 3.

8. Spectrometer Comparison Experiments Plan for LACIE Field Measurements. 1977. Report prepared for NASA/Johnson Space Center by Laboratory for Applications of Remote Sensing, Purdue University.

9. Egbert, D.D. 1977. A Practical Method for Correcting Bidirectional Reflectance Variations. Proceedings Machine Processing of Remotely Sensed Data Symposium, Purdue University, June 21-23. IEEE Catalog Number 77 CH 1218-7 MPRSD.

VII. Acknowledgements

The leader of the field measurements project at LARS is Marvin Bauer; Larry Biehl and Marilyn Hixson are the project managers for data acquisition, processing, and verification and data analysis, respectively. Barrett Robinson has been responsible for directing the acquisition of data and William Simmons, Paul Spencer, and Jeanne Etheridge have been responsible for data processing. The following personnel are acknowledged for their efforts in various aspects of the project:

Data Acquisition:

John Ahlrichs, Chris Parker, Don Crecelius, Charles Curtis, Joe Smith, Joe Wojda, Craig Daughtry, Jeff Kollenkark, Kevin Healy, Vern Vanderbilt, and Vic Fletcher.

Data Processing and Software Development:

David Freeman, Nancy Fuhs, Keith Phillip, Bill Shelley, Jim Kast, Chuck Smith, Jeff McMeekin, Keith Steva, Don McLaughlin, Jerry Majkowski, Jack West, Ken Dickman, and Andy Teetzel.

Data Analysis:

John Ahlrichs, Vern Vanderbilt, Craig Daughtry, Chris Stellon, and Jeff Kollenkark.

Data Library:

Kathi Freeman, Cathy Axtell

Faculty consultants, Virgil Anderson, LeRoy Silva, and Byron Blair, are acknowledged for their contributions to the project. Special appreciation is due to Lisa Afanador, Beverly Carpenter, and Julie Hanover for secretarial support during the project. Finally, the assistance and cooperation provided by the North Dakota Agricultural Experiment Station at Williston is appreciated.

B. Thermal Band Canopy Modeling

I. Background

The final grain yield of wheat crops has a strong dependence on the moisture stress during the growing season. It has been demonstrated that moisture stress can readily be evaluated by comparing the wheat biomass temperature to the surrounding air temperature.[1,2] Plants with adequate moisture available are cooler than the surrounding air due to transpiration. But, transpiration is inhibited in water stressed plants and consequently the temperature of water stressed plants increases. An advantageous technique for determining wheat canopy temperatures of large areas is by remote sensed thermal radiation measurements.

On days without significant cloud cover there exists a temperature gradient in the wheat canopy. Also, the soil temperature is greater than the wheat biomass temperature. This feature of crop canopies has two significant influences on determining the moisture stress of the crops:

(1) The radiance temperature of a wheat canopy measured from an overhead position is an integration of the soil temperature and the various temperatures of the wheat. A knowledge of the various canopy temperatures that exist must be attained in order to assess the overhead radiance temperature.

(2) Many models used to predict moisture stress assume one biomass temperature. Therefore, an understanding of the temperature gradients in wheat canopies could aid in the development of more realistic and general models to predict water status and final grain yield.

II. Introduction

The goal of the thermal modeling task is to gain an understanding of the relationship between the canopy temperatures (biomass and soil) and the canopy geometry. To achieve this goal the following working objectives were set:

1. To relate canopy parameters including: radiance temperature measured from an overhead position, vertical temperature profile, soil temperature, and canopy geometry.

2. To determine wind velocity effects on the wheat canopy.

3. To observe diurnal thermal phenomena.

4. To determine the quality of the laser technique in providing geometric characterization of the canopy.

The measurements required to accomplish the above goals are: vertical temperature profile, overhead temperature, soil temperature, air temperature profile, and canopy geometry. The thermal measurements were obtained on two thermally different canopies; one canopy shielded from the wind, the other canopy was under prevailing wind conditions. The geometric characterization was obtained from measurements performed on the shielded canopy.

III. Theory of Method

The method used to relate the vertical temperature profile and the overhead radiance temperature is:

1. Determine the canopy geometry.
2. Divide the canopy into 10 horizontal layers.
3. Determine the average temperature of each horizontal layer.
4. Sum the product of the average temperature of each layer and the fraction that the layer contributes to the total view, for all 10 layers, that is,

$$T = \sum_{i=1}^{10} \left[x_i T_i^4 \right]^{1/4}$$

where x_i = fraction the i^{th} layer contributes to the total view, T_i = average radiance temperature of the i^{th} layer, and T = predicted overhead radiance temperature.

IV. Test Site

Spectral-thermal measurements were made on June 28, and July 1, 1977 on wheat canopies under different environmental conditions. All measurements were made on the Larry O'Brian farm adjacent to the University of North Dakota Agriculture Research Station at Williston, North Dakota. The wheat had the following agronomic description: hard red spring wheat, fully headed, approximately 82 cm in height, row spacing 17-18 cm.

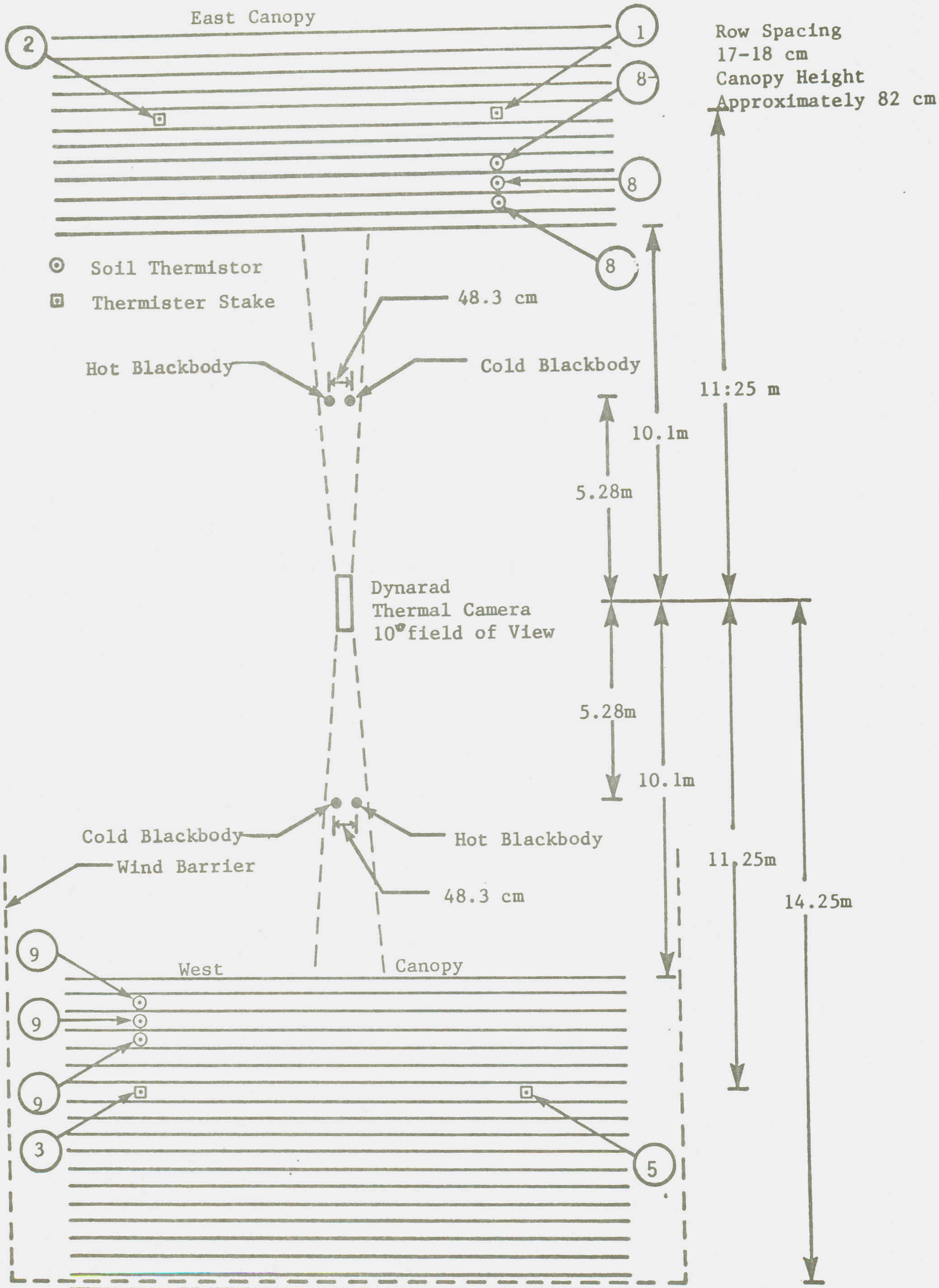
Figures 1 and 2 illustrate the procedures for the data acquisition and the location of the instruments used for the thermal measurements.

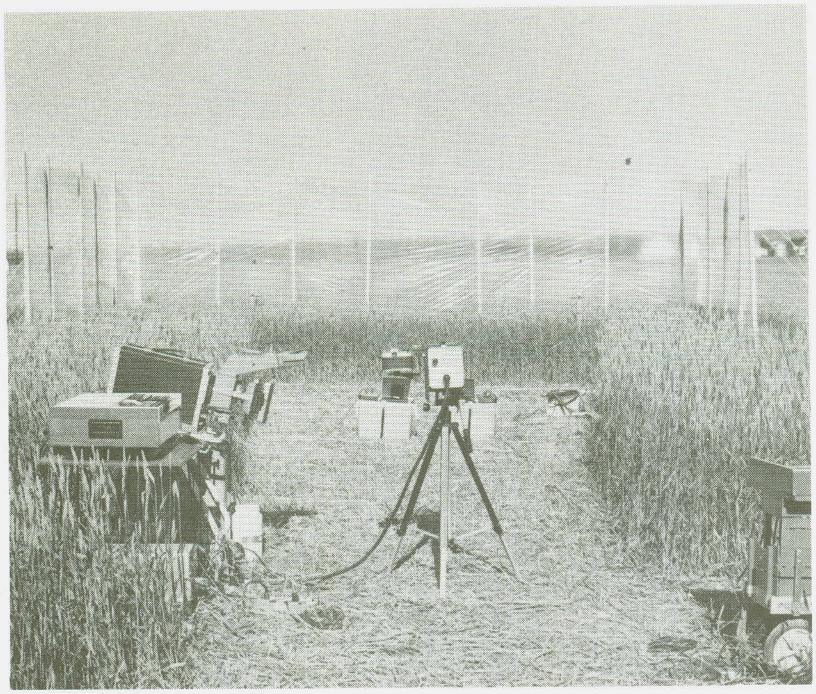
V. Experimental Procedure

The Dynarad 209A thermal scanner, an optical-mechanical scanner and display unit, was used to determine the vertical temperature profiles (Figures 2A,B,C). The scanner utilizes reflective optics and a dichroic mirror to focus the image with an optical resolution of 1.74 milliradians and a total field of view of 10°. The detector is mercury-cadmium-telluride, operating in the 8 to 14 micrometer spectral region and is mounted in a liquid nitrogen cooled dewar. All thermal image data was recorded on a Sony video tape recorder (VTR), model 2850.

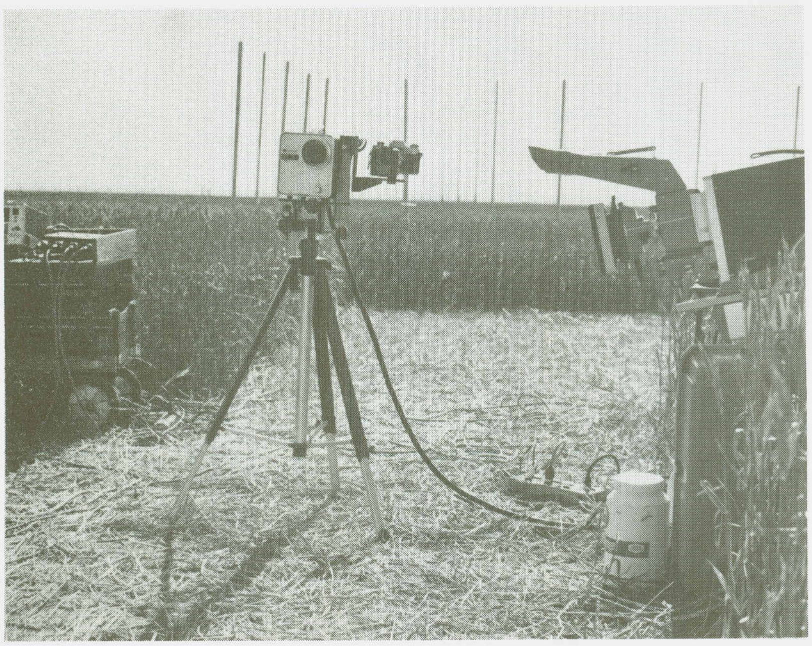
The wheat canopy, planted in north-south rows, was observed from the east and west directions. The west canopy was shielded from the wind by a 2.4 meter high polyvinyl barrier to provide a virtually motionless canopy

Figure B-1. Equipment setup for thermal measurements taken on Larry O'Brian Farm adjacent to University of North Dakota Agriculture Research Farm in Williston, North Dakota.





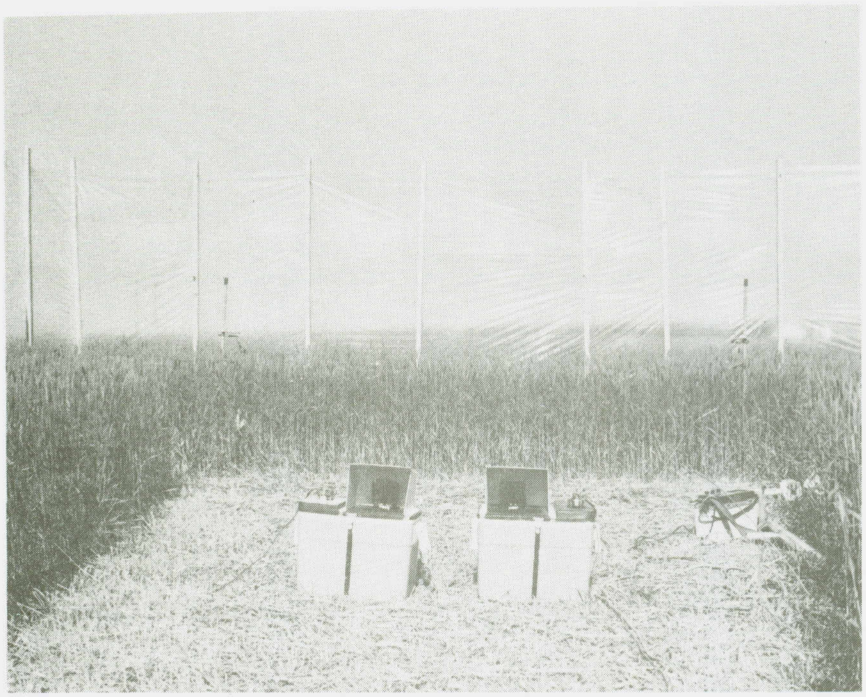
(A) West canopy as viewed by Dynarad Thermal Scanner.



(B) Dynarad Thermal Camera setup east canopy in background.

Figure B-2. Test site and experimental procedures at Larry O'Brian farm adjacent to University of North Dakota Agriculture Research Farm in Williston, North Dakota, July 1977.

Figure B-2 (cont.)



(C) West canopy: Blackbodies, thermistor stakes, and wind barrier.

(D) Overhead Radiance Temperature measured with PRT-5 on east canopy.

(Figure 2A,B,C). The east canopy was not shielded from the wind and provided a canopy under prevailing conditions. The purpose of the two types of canopies (with and without wind velocity) was to determine the effect of wind velocity on a wheat canopy and to provide a motionless canopy for determining geometric characterization using the laser technique.[3]

The Barnes Precision Radiation Thermometer, model PRT-5, was used to measure the overhead radiance temperature at a zenith angle of 0° (Figure 2D). The PRT-5 is an optical-electronic unit that is designed to give equivalent blackbody temperatures from -20°C to 75°C. The radiation detector is a hyperimmersed thermistor bolometer mounted in a controlled reference temperature cavity. The optical elements are an objective lens which defines the instrument field of view (nominal 2°) and a spectral filter which limits the measurements to the 8 to 14 micrometer spectral region.

The PRT-5 was mounted on a support, which could be rotated, 2.75 meters above the soil. An average overhead radiance temperature was obtained by taking measurements at six positions across the front of the canopy behind the fourth row. At each position the PRT-5 was rotated at 30° increments in a semicircle; a total of 42 measurements were obtained for each set of measurements. This procedure provided an average of the soil and biomass temperatures.

The laser technique yielded the necessary data to determine the geometric characterization of the wheat canopy. This technique consisted of pointing a laser at the canopy at a zenith angle of 0° and measuring the height of the intersection between the laser beam and wheat canopy. From this information it is possible to determine the fraction that each horizontal layer contributes to the total normal view.

Thermistors were used to measure the soil temperature and air temperature profile within the wheat canopy; see Figure 1 for thermistor probe placement.

VI. Summary of Measurements

A total of 24 sets of thermal data were collected: 8 on June 28, and 16 on July 1, 1977. Half of the data sets were collected on the east canopy and half on the west (motionless) canopy. All thermal measurements on each canopy were collected in a 10 minute period in the following order: thermistor, PRT-5, thermal scanner. The data sets were collected in pairs; an east data set and a west data set were acquired within 10-15 minutes of each other. In addition to the thermal data five sets of laser data were collected for a total of approximately 2000 data points.

The radiance temperature profiles were obtained using the Dynarad and the VTR. The thermal scanner display unit was operated in the line scan (A-scan) mode and had an illuminated grid on front of the display screen. The temperature reduction procedure consisted of the following:

1. The top and bottom of the canopy were located with respect to line scan position.

2. The tape was viewed on the display unit to locate a 10-15 second time span where the instrument sensitivity did not change appreciably.

3. The line scan was set for the blackbody location. A polaroid photograph was taken showing the hot blackbody as a high peak on the grid and the cold blackbody as a low peak on the grid.

4. The line scan was set at the bottom of the canopy then moved up at regular intervals (7.5 cm) while each position was photographed.

5. The canopy scale readings were related to the blackbody scale readings.

6. The scale readings were converted into temperatures in a linear calibration.

7. The temperatures are then plotted on a temperature versus height graph. See Figure 3 and Table 2 for results.

The overhead radiance temperatures measured with the PRT-5 are given in Table 1. Each value is an average of 42 measurements and is an integration of all canopy temperatures as observed from an overhead position.

The geometric characterization of the wheat canopy is determined by the following method. The canopy is divided into 10 horizontal layers, one soil layer and 9-10 cm horizontal layers within the wheat canopy. The fraction that each layer contributes to the total normal view is calculated by dividing the number of hits in each layer by the total number of hits in all of the layers. A hit is the intersection between the wheat or soil surface and the laser beam.

$$x_i = \frac{H_i}{H}$$

where, x_i = fraction the i^{th} layer contributes to the total view, H_i = number of hits in the i^{th} layer, and H = total number of hits in all layers. Figure 4 shows the results of the determination of vertical biomass distribution.

The soil temperatures are given in Table 1. The air temperature profiles are given in Table 3.

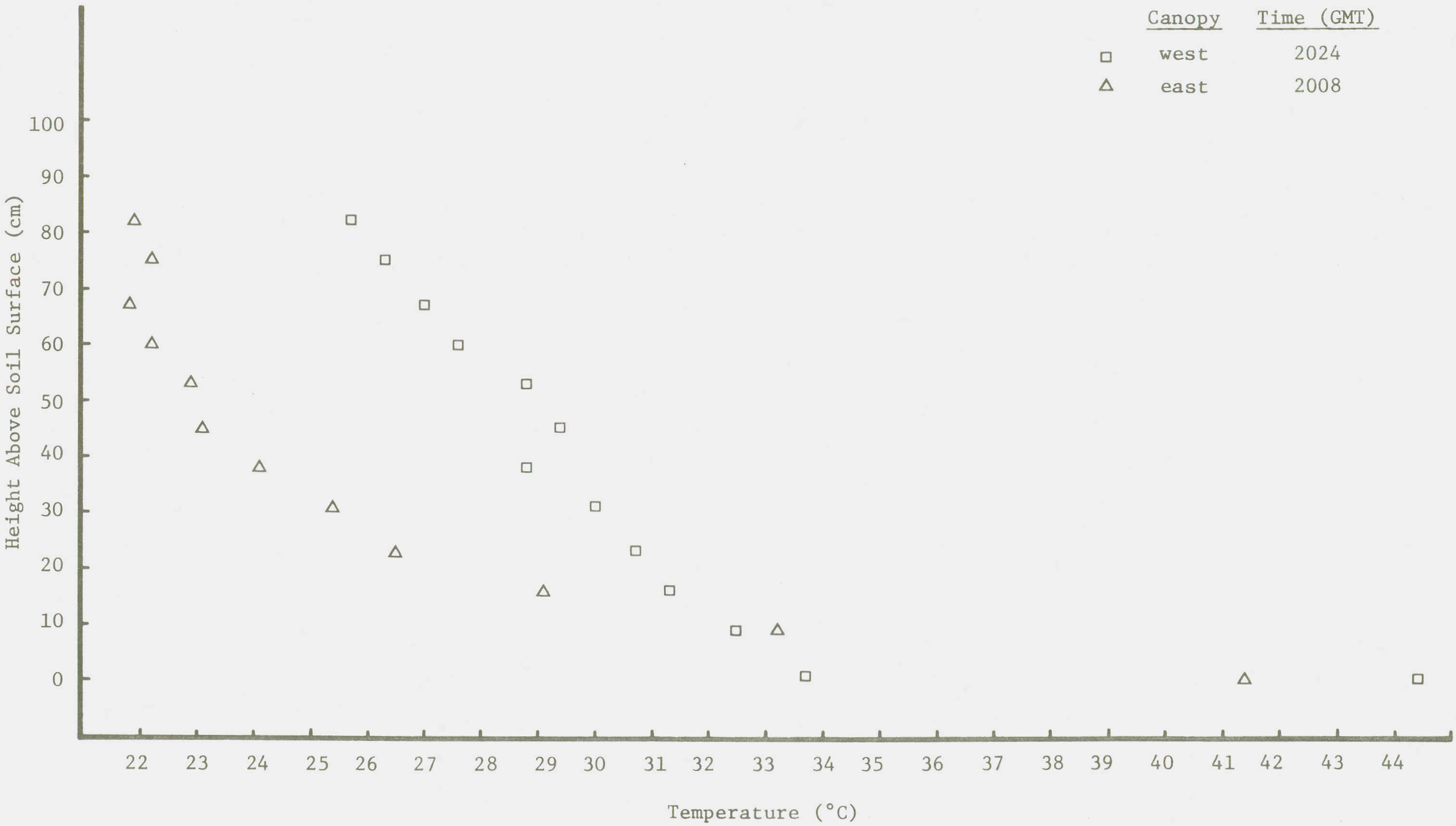


Figure B-3. Vertical temperature profiles of the east and west canopies. The east canopy under prevailing wind conditions, the west canopy was shielded from the wind. Soil temperatures measured with thermistors, 1 July 1977.

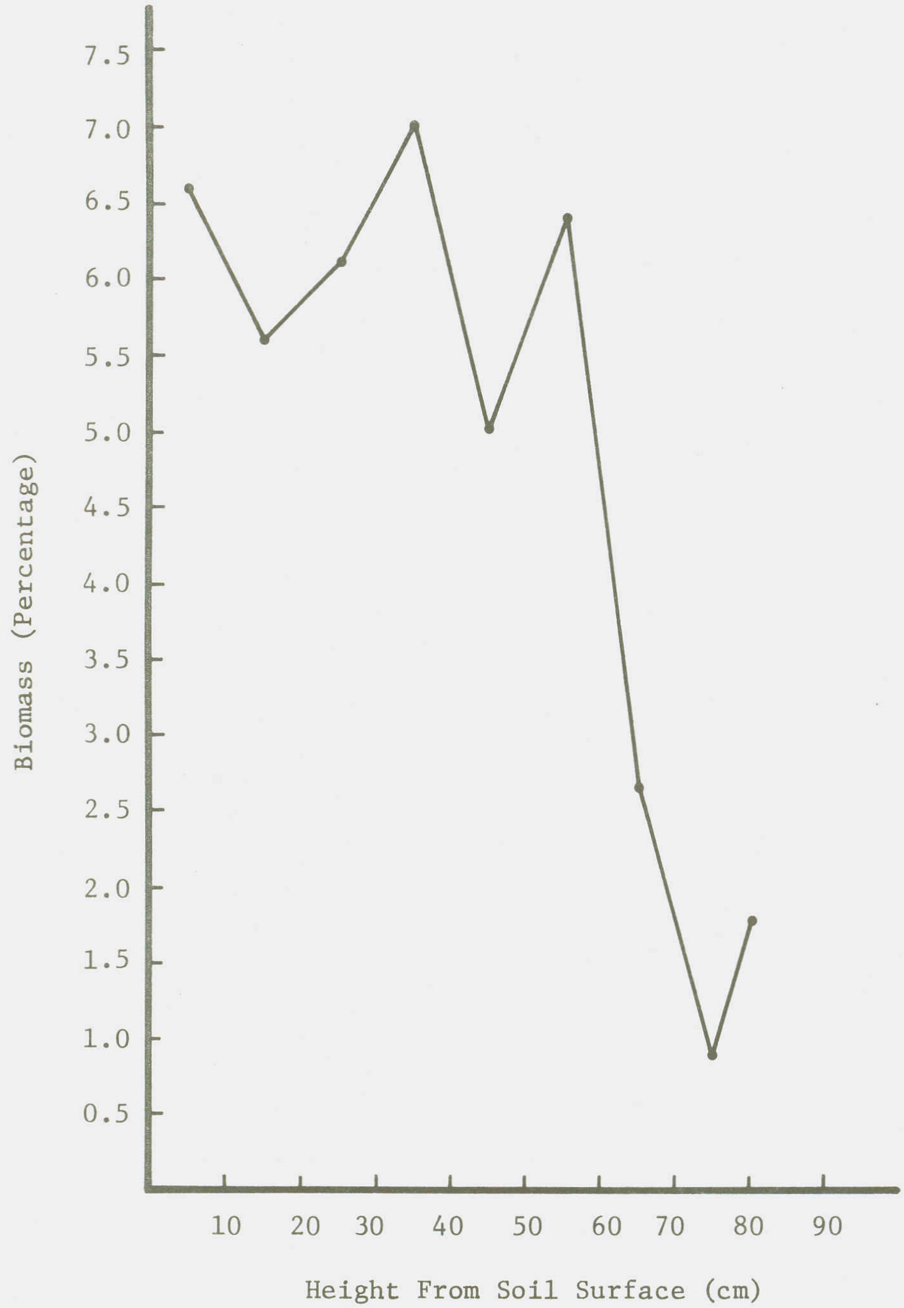


Figure B-4. Biomass Vertical Distribution as viewed from normal overhead position (Laser Technique, 1 July 1977) soil surface 57.7% of normal view.

Table B-1. Canopy temperatures: Soil temperature, measured and predicted overhead radiance temperature, and temperature gradient.

Set	Time GMT	Canopy	Soil Temperature (°C)	Overhead Measured (°C)	Overhead Predicted (°C)	Temperature Gradient (°C/cm)
1	1408	East	17.0	17.4	-	-
1	1426	West	22.2	21.4	-	-
2	1610	West	29.7	29.1	27.4	-0.108
2	1620	East	26.4	27.0	24.4	-0.055
3	1810	East	37.3	32.7	31.8	-0.146
3	1828	West	41.8	40.4	36.1	-0.105
4	1908	West	47.8	43.2	42.1	-0.078
4	1922	East	43.8	36.7	37.4	-0.154
5	2008	East	42.4	36.0	36.5	-0.132
5	2024	West	45.4	42.3	39.6	-0.092
6	2109	West	43.1	40.4	37.9	-0.096
6	2122	East	41.1	34.4	34.9	-0.140
7	2205	East	38.3	33.7	34.6	-0.138
7	2221	West	38.2	36.3	33.9	-0.076
8	2334	West	32.7	30.9	29.3	-0.031
8	2349	East	31.0	28.5	29.4	-0.082

Table B-2. Radiance temperature profiles July 1, 1977.

Temperature Profile (°C)

East Canopy

Height Above Soil (cm)	1620	1810	1922	2008	2122	2205	2349	Time (GMT)
0*	26.4	37.3	43.8	42.4	41.1	38.3	31.0	
5	22.9	30.6	35.4	35.4	31.2	35.0	30.9	
15	22.3	26.4	31.2	30.8	29.9	33.2	30.7	
25	21.4	24.0	27.8	27.6	26.6	30.5	27.6	
35	20.9	22.8	26.2	25.8	24.4	26.4	25.8	
45	20.9	21.4	24.8	24.4	22.8	26.4	24.6	
55	20.0	18.8	23.6	23.8	21.8	25.6	23.9	
65	19.0	20.0	22.9	23.0	21.2	25.0	23.8	
75	19.0	18.8	22.6	23.2	20.8	25.0	24.3	
82	19.0	18.8	22.0	22.8	20.8	25.0	24.5	

Temperature Profile (°C)

West Canopy

Height Above Soil (cm)	1610	1828	1908	2024	2109	2221	2334	Time (GMT)
0*	29.7	41.8	47.8	45.4	43.1	38.2	32.3	
5	29.2	33.4	35.4	33.8	33.8	31.5	26.5	
15	26.1	30.0	31.2	32.8	32.2	29.7	25.5	
25	23.9	27.4	27.8	31.4	30.6	27.5	24.8	
35	22.7	26.2	26.2	30.4	29.4	26.2	24.2	
45	2.15	25.0	24.8	30.0	28.4	25.6	23.8	
55	21.0	24.2	23.6	29.4	27.7	25.8	23.5	
65	21.0	24.6	22.9	28.2	27.2	25.5	23.9	
75	21.0	24.4	22.6	27.6	26.8	24.7	24.1	
82	21.0	24.0	22.0	26.8	26.9	24.7	24.1	

* All soil temperatures were measured with thermistor probes.

Table B-3. Air temperature profiles within wheat canopy 1 July 1977.

Temperature Profiles (°C)

East Canopy

Height Above Soil (cm)	1408	1620	1810	1922	2008	2122	2205	2349	Time (GMT)
96	18.5	24.1	27.3	28.4	29.4	30.9	30.5	27.9	
76	19.1	24.7	27.8	29.2	29.1	30.6	30.7	27.7	
31	19.6	28.3	31.9	34.3	33.1	33.8	34.4	28.8	
6	19.1	29.5	34.5	38.3	37.1	36.3	36.2	29.8	
0	17.0	26.4	37.3	43.8	42.4	41.1	38.3	31.0	

Temperature Profiles (°C)

West Canopy

Height Above Soil (cm)	1426	1610	1828	1908	2024	2109	2221	2324	Time (GMT)
96	21.6	26.4	29.6	29.1	31.3	32.1	30.1	27.8	
76	21.5	26.7	31.4	29.3	31.4	32.1	30.2	27.6	
31	23.2	30.1	35.8	31.9	35.0	33.3	32.3]	28.7	
6	22.2	32.0	36.3	36.7	36.6	36.3	34.6	30.2	
0	22.2	29.7	41.8	47.8	45.4	43.1	38.2	32.3	

Table B-4. Regression analysis of soil and overhead radiance temperature, 1 July 1977.

Regression Analysis:

	Equation	r^2
East:	$T = 6.971 + 1.21\ T_{prt}$	0.9482
	$T = 6.38 + 0.708\ T_{soil}$	0.9670
	$T_{prt} = 7.10 + 0.68\ T_{soil}$	0.9785
West:	$T = 1.058 + 0.905\ T_{prt}$	0.9803
	$T = 4.80 + 0.760\ T_{soil}$	0.9969
	$T_{prt} = 2.26 + 0.884\ T_{soil}$	0.9898

T = Predicted Overhead Radiance Temperature

T_{prt} = Measured Overhead Radiance Temperature

T_{soil} = Soil Temperature

VII. Results of Model

The data collected on July 1, 1977 provided all of the canopy parameters used in the model to predict the overhead radiance temperature. Temperature profiles were obtained for most of the data sets and average temperatures for each horizontal layer of the canopy were determined from the temperature profiles and soil thermistor temperature measurements. The results of the predicted overhead radiance temperature are given in Table 1. Figure 5 shows the predicted vs measured overhead radiance temperatures. The east and west points are located on different curves therefore, linear regression was performed on each set of points with regression coefficients of 0.948 and 0.980 for the east and west curves, respectively.

VIII. Discussion of Results

The model used to predict the overhead radiance temperature gave results that provide insight into the thermal characteristics of the wheat canopies. Figure 5, the graph of the measured versus predicted overhead radiance temperature, reveals a measurable effect of several environmental parameters on the model and on the measurements used in the model including: solar power and position, wind velocity, air temperature, and canopy geometry. An analysis of the above parameters and their effect on the model provides an explanation for the difference between the east and west data points of Figure 5.

The most noticeable characteristic of Figure 5 is the difference between the east and west data points. Most of the east predicted overhead temperatures were greater than the measured temperatures but, the west predicted overhead temperatures were all less than the measured temperatues.

The parameter with the largest variation between the east and west canopies is wind velocity. The wind velocity affected the following: canopy temperatures, temperature gradient, and canopy geometry.

The soil temperature was less in the east canopy (with wind) than the west canopy due to an increase in convection in the east canopy. The wheat biomass temperature was reduced in the east canopy because the evapotranpiration was enhanced by the wind velocity. Table 1 shows the temperature gradients that exist within the wheat canopies. A larger temperature gradient existed in the east canopy (with wind velocity) than in the west canopy.

To determine the effect of wind velocity on canopy geometry, linear regression was performed on the following temperatures: measured overhead and soil, predicted overhead and soil; for both east and west canopies. The results are given in Table 4 and show that the overhead radiance temperature exhibits a stronger dependence on the soil temperature in the west canopy than in the east canopy. Therefore, the movement in the east canopy, due to wind, resulted in a reduction in the fraction the soil layer contributed to the normal view. The soil is the warmest layer during the day. Consequently, the model using the geometric characterization obtained from a motionless canopy predicted high temperatures in the east canopy (with wind). Early

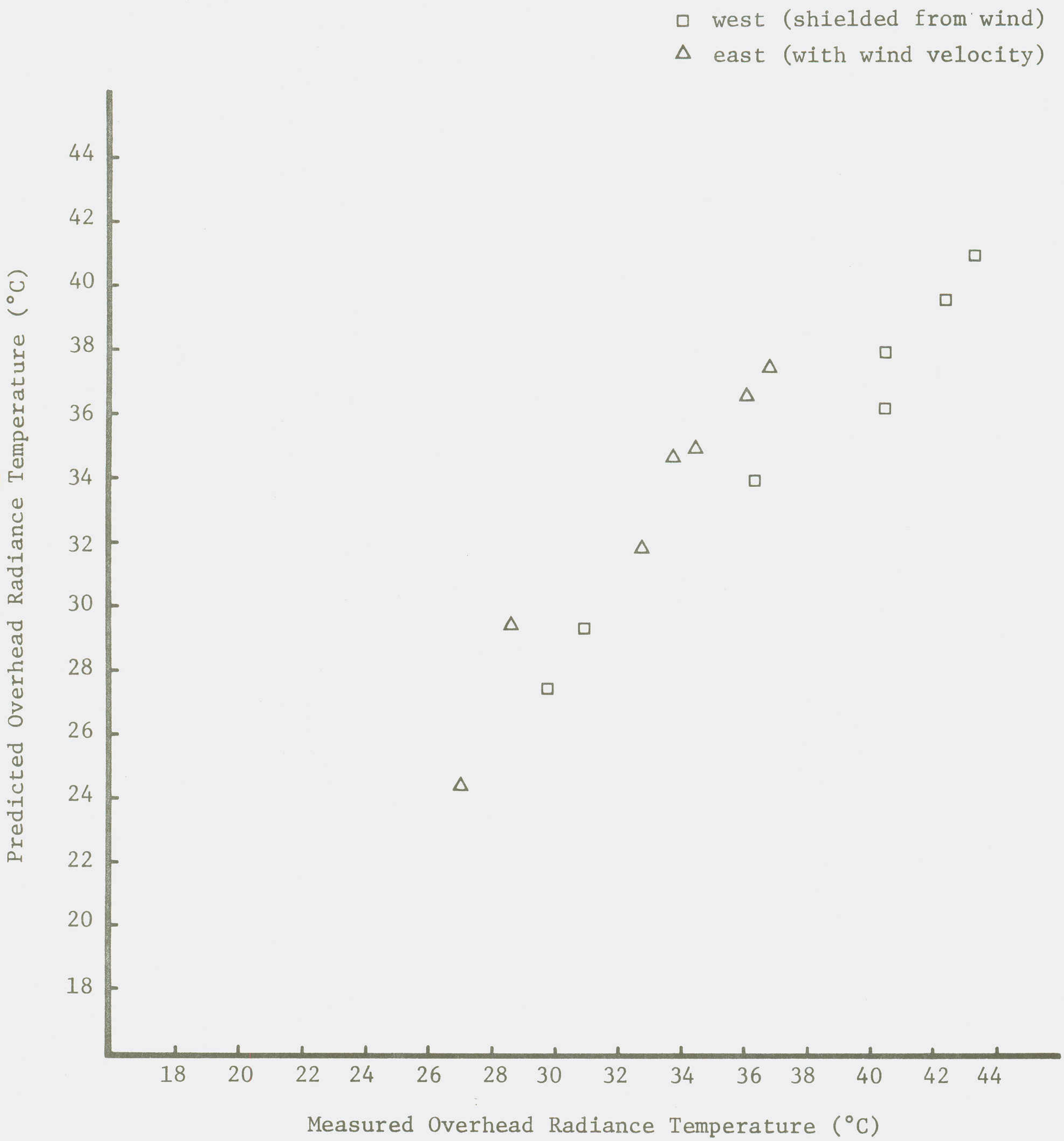


Figure B-5. Measured versus Predicted Overhead Radiance Temperature for two different wheat canopies. East under prevailing wind, west shielded from wind, July 1, 1977.

in the morning before the solar irradiation penetrates the canopy to the soil, the soil may be the coolest part of the canopy. Then, the effect of using the geometric characterization obtained from a motionless canopy on a canopy with wind velocity would be the opposite. The model would predict overhead radiance temperatures lower than the measured temperatures.

The solar azimuth angle had an effect on the east canopy thermal measurements during the late afternoon. The radiance temperature profile of the east canopy increased instead of a continual temperature decrease that usually occurs after solar noon. This increase in the temperature profile was attributed to the direct solar irradiation on the front vertical plane of the canopy viewed by the Dynarad thermal scanner, (and not an increase in moisture stress) because the overhead radiance temperature was declining when the temperature profile was increasing.

It has been shown that wind velocity has a significant influence on the thermal characteristics of wheat canopies. Figure 6 shows soil temperature minus overhead temperature versus overhead temperature for both the east and west canopies. The east and west data points are located on two different curves. The east canopy (with wind) has a larger soil minus overhead temperature difference than the shielded west canopy. This shows that a unique relationship between soil temperature and overhead radiance temperature may exist for every wind velocity.

The maximum soil temperature and overhead radiance temperature measured occurred at approximately solar noon (1900-1930 GMT) for both the east and west canopies. The vertical biomass temperature gradient was relatively constant from 1600 to 2200 GMT with a decrease in both canopies at 1510 and 1410 GMT for the east and west canopies respectively.

The laser technique provided good geometric characterization of the west motionless canopy but, the application of the motionless geometric characterization to a canopy with motion will result in the following error. The laser technique yields a high value for the soil layer's contribution to the overhead view. Therefore, using the motionless geometric characterization in the model to predict overhead radiance temperature will give a high predicted temperature. An automated system to perform the laser technique would be required to correctly characterize wheat canopies with wind velocity. This is due to the following: the measurements (height of intersection between the laser beam and canopy) must be performed with much greater speed, due to the canopy movement, than can be achieved with the manual technique. Furthermore, the required number of data points to achieve a desired statistical confidence level would be greater due to greater variability of the canopy.

IX. Applications of Results

Generally, the models used to predict the moisture status and final yield of wheat crops require the leaf (or biomass) temperature and the surrounding air temperature. Idso, Jackson, and Reginato [4,5] have developed the "Stress Degree Day" concept to predict final yield. They hypothesized that final yield is linearly related to the stress degree day accumulated over some critical period.

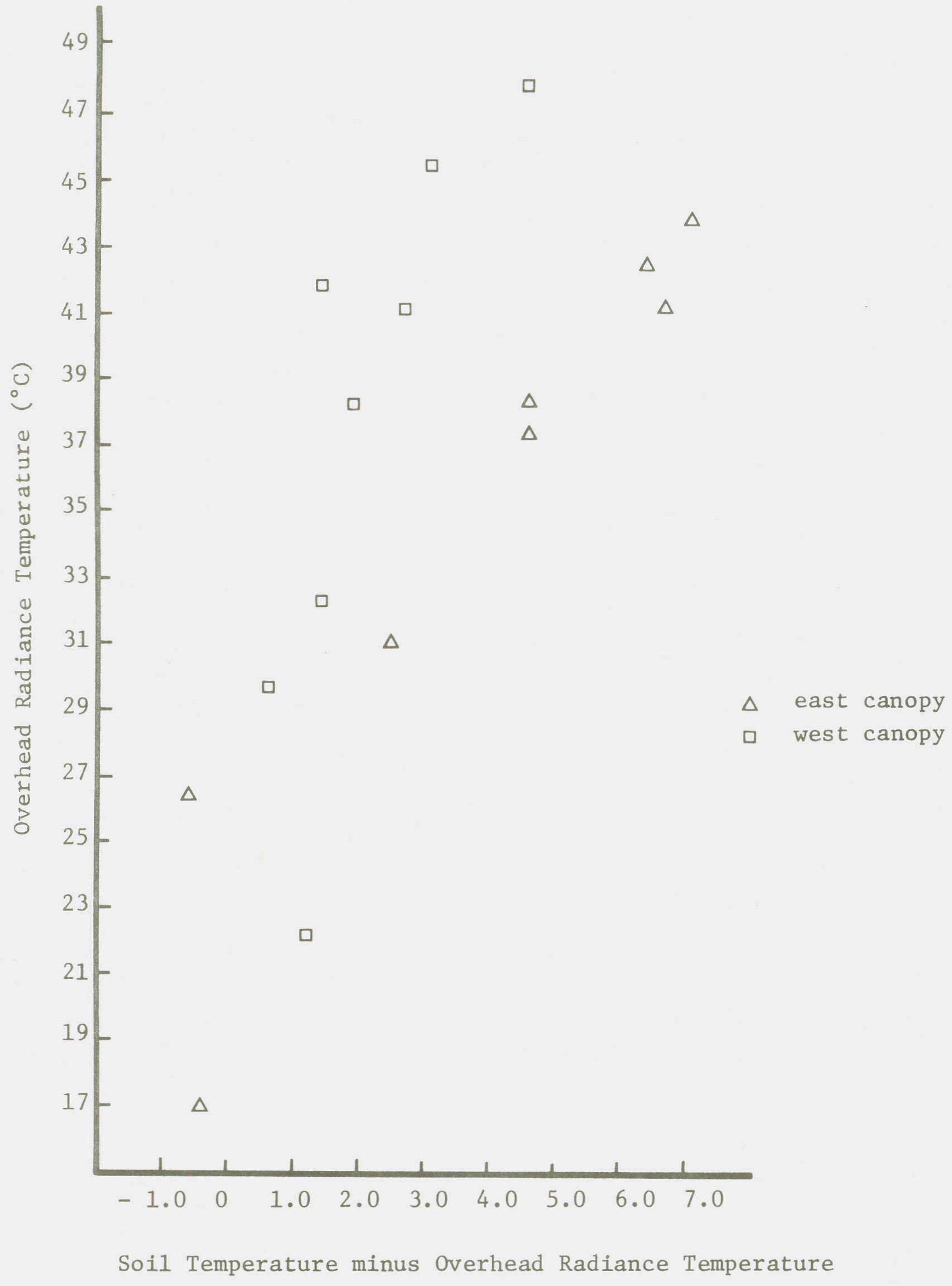


Figure B-6. Relationship between Soil Temperature and Overhead Radiance Temperature (measured with PRT-5) 1 July 1977.

$$Y = \alpha - \beta\left(\sum_{i=b}^{e} SDD_i\right)$$

where Y = final yield, SDD_i is the midafternoon leaf temperature minus air temperature on day i, and b and e are the beginning and ending days of the summation procedure.

Idso, Jackson, and Reginato applied the "Stress Degree Day" concept to thermal measurements they obtained on a 72x90-m field of winter wheat (Triticum durum Desf. Var. Produra) in Arizona. The field was divided into six north-south rectangular plots. The six plots were irrigated differently in order to observe plots with different moisture status, ranging from very dry to very wet.

The thermal measurements were made with 2 and/or 20 degree field of view Barnes PRT-5 radiation thermometers. The 20 degree F.O.V. PRT-5 viewed the canopy from overhead positions at zenith angles of 0° and 45°. The 2 degree F.O.V. PRT-5 viewed the canopy at an angle such that only plant parts were viewed.

They reported results for all angles were equally good with regression coefficients of around 0.976.

The above experiment was done on wheat canopies under different moisture stress but, other environmental conditions, such as wind, were close to being constant for all plots due to the close proximity of all the plots. Therefore, the "Stress Degree Day" concept as presented by Idso, Jackson, and Reginato shows the relationship between final yield and the summation of leaf temperature minus air temperature but, it does not contain the leaf temperature dependence on wind velocity. Consequently, the "Stress Degree Day" concept would predict different final yields for two canopies with the same moisture stress if the wind velocity significantly varied (between the two canopies) for the summation period. This is due to the following two effects: (1) the wind changes the thermal characteristics of the canopy, (2) the wind changes the canopy geometry.

The summation procedure of the "Stress Degree Day" concept should reduce the effect of variation in environmental condition. But, it is apparent that the application of the concept to canopies under different environmental conditions will require an understanding of the canopy thermal and geometric responses to environmental conditions, including wind velocity. Once the effect of wind velocity on canopy parameters is determined for various wind velocities then, the success of predicting water stress and crop yield will be improved.

X. Summary

The thermal measurements and geometric characterization provided the necessary data required to determine the relationship between canopy and environmental parameters including: solar power and position, wind velocity canopy temperature, and canopy geometry. The model used to predict overhead radiance temperatures is a powerful tool to assess the relative dependence of the overhead radiance temperature on the different parts of the canopy (heads, leafs, stem, and soil).

The significant results can be summarized as the following:

1. Wind velocity has a significant influence on the overhead radiance temperature and the effect has been quantized. It alters biomass and soil temperatures, temperature gradient, and canopy geometry.

2. The temperature gradient of the wheat canopy is relatively constant during the day.

3. The temperature gradient is a function of wind velocity.

4. The laser technique provides good quality geometric characterization, and an automated system would be required to characterize canopies with wind velocity.

The results of the measurements and the analysis indicate the strong and weak aspects of the experiment. The major contribution of the experiment was to show the effect of environment conditions on wheat canopies and to demonstrate the relationship between canopy parameters and radiance temperature. A limitation of the experiment was the quantity of data collected, a result of budgetary restrictions. As a consequence, it is difficult to provide thorough quantitative results, but the qualitative analysis has provided an understanding of the thermal phenomena that exists in wheat canopies. Also, the experimental results provide insight into the relative importance of various canopy and environmental parameters.

To gain a comprehensive understanding of the various environmental and canopy parameters that have an influence on remotely sensed thermal measurements a large quantity of measurements on canopies under different environmental conditions would be required. Also, more data will be required to determine the discrepancy between the measured overhead radiance temperature and the predicted temperature.

Most of the important parameters needed to assess the impact of the various environmental conditions on crop canopies could be obtained from the thermal and geometric measurements used in the experiment to relate canopy parameters. In addition, an automated system that could perform the laser technique on canopies with motion would be required. Measurements that would aid in future analysis would include wind velocity profile within the wheat canopy, soil temperature distribution, and solar power distribution within the wheat canopy (using laser technique).[3] With these additional measurements, the heat transfer processes in wheat and other canopies could be analyzed more comprehensively. Also, with an increase in the amount of data acquired with an automated system to perform the laser technique it might be possible to assess the canopy geometric response to moisture stress.

By gaining an understanding of the important parameters affecting crop canopies, models such as the "Stress Degree Day" could be applied to a larger range of conditions and crops.

REFERENCES

1. Blad, B. L., Heilman, J. L., Kanemasu, E. T., and Rosenberg, N. J. "Thermal Scanner Measurement of Canopy Temperatures to Estimate Evapotranspiration", Remote Sensing of Environ., 5, 137-145 (1976).

2. Bartholic, J. F., L. N. Namken, and C. L. Weigand, "Aerial Thermal Scanner to Determine Temperatures of Soil and Crop Canopies Differing in Water Stress", Agron. J., 64, 603-608, 1972.

3. Vanderbilt, V. C., L. F. Silva, and M. E. Bauer, "A Laser Technique for Characterizing the Geometry of Plant Canopies", LARS Information Note 120776 (1977).

4. Idso, S. B., R. D. Jackson, and R. J. Reginato, "Remote-Sensing of Crop Yield", Science, Vol. 196, 1 April 1977.

5. Idso, S. B., R. D. Jackson, and R. J. Reginato, "Wheat Canopy Temperature: A Practical Tool For Evaluating Water Requirements", Water Resources Research, Vol. 13, No. 3, June 1977.

C-1. Forestry Applications Project.

The overall thrust of the FAP activity has been the continued development of key elements in the design of forest inventory systems utilizing Landsat data and computer-aided analysis techniques. During the current contract period, FAP activities have involved four specific tasks, as follows:

1. To complete work initiated in 1976 for defining the acceptability of Landsat acreage estimates as input to a forest inventory design.

2. To complete the work on defining and documenting an efficient and cost effective method of developing an optimal set of training statistics for mapping forest cover.

3. To make a comparison of five different classification techniques in terms of cost, accuracy, and the characteristics of the output products.

4. To define the objectives and scope of the forestry research activities to be pursued by LARS within the FAP year program.

The first three of these objectives have each resulted in significant findings, as described in the following three sections of this report. As reported in the last Quarterly Progress Report, activity on the fourth task was considered to be finished at that time, since discussion with FAP personnel indicated that there would be no forestry applications research as part of the LARS SR&T program after the current contract period.

It should be pointed out that this report contains only a synoposis of the complete final report involving Task 2 (i.e. definition of an efficient and cost effective method for developing an optimal set of training statistics for mapping forest cover). The synoposis contained herein briefly documents some of the key aspects of the complete study. LARS Technical Report No. 112177 contains the complete findings of this study. However, because of the size of Technical Report 112177, it is not being appended to this final contract report, but copies have been forwarded to the FAP Program Office at J.S.C., and for those interested, copies can be obtained upon request from LARS or from the FAP Program Office.

C. FORESTRY APPLICATIONS PROJECT

C. I. Introduction

Background. Personnel at LARS have been working with the Forestry Applications Program at NASA/JSC since 1974. Our major thrust during this involvement has been in assessing the applicability of computer-assisted analysis techniques to forest mapping and inventory. Specifically, over the last year-and-a-half our focus has been in the application of Landsat collected and computer-analyzed data to national forestry inventory.

This report will discuss our recommended approach for using remote sensing techniques to meet a portion of the forest information needs imposed by the 1974 Resources Planning Act. Examples of how the existing Landsat data and computer-assisted techniques can benefit on-going Forest Survey will be drawn up on. Suggested inventory procedures together with recommendations for additional work in forest inventory will be outlined.

Objective. The results which follow are the culmination of work under the objective to: Develop inventory methods using remote sensing technology with application forest and range land information needs.

II. Approach

Problem Definition. The objective we were given was nebulus in that it did not contain a well defined problem statement. Therefore, our first task was to identify the portion of the problem we could most affect with the limited resources we had available. We accomplished this by collecting and reviewing pertinent forest inventory literature. The bibliography included with this report identifies the material reviewed. This literature can be grouped into the following categories: Forest Service publications, forest mensuration and statistics texts, symposia proceedings and UN-FAO inventory manuals.

The 1974 Resources Planning Act (RPA) and the 1975 Resource Assessment were emphasized heavily during our initial review. These documents best defined needs and requirements for information on a national scale whereas the other literature dealt with specific needs and situations. Table 1 is a distillation of the RPA needs and their relationship to remote sensing inputs. Based on analysis of Table 1 and previous experience with Landsat data, we determined that the satellite data could be most useful for estimating the areal extent of various resource cover types. Furthermore, area estimation was not discussed extensively in any of the literature we reviewed. Therefore, we identified our problem in terms of determining if computer-assisted analysis of Landsat data could provide accurate areal estimates of gross forest cover.

III. Current National Survey Situation

Intuitively, the estimation of the total area in forest type should be an important variable to any sample design. The only discussion of the importance of area is related to on-going Forest Survey, Therefore we

Table C-1. Summary Review of Resource Planning Act Assessment (1).

System	Unit of Measure (2)	Measurable by R. S. (3)	Beneficient to Decision (4)
Timber			
Type	Area	+	D
Ownership	Area	0	I
Productivity	Volume	0	I
Range			
Type	Area	+	D
Ownership	Area	0	I
Productivity	Volume	0	I
Water			
Type-Impounded	Area	+	D
Flowing	Area	0	I
Use	Volume	-	I
Fish and Wildlife			
Habitat Type	Area	+	I
Population Estimates	Census	-	I
Recreation and Wilderness			
Participation	Census	-	I
Human and Community Development	Economic	-	N

(1) ref: The Nation's Renewable Resources-An Assessment, 1975, USDA-Forest Service.

(2) General classification of units of measure reported in (1).

(3) Measurable by remote sensing techniques.
Includes all measures possible from satellite and/or aircraft systems.
+ Directly measurable from satellite or aircraft data alone.
0 Measures can be inferred with the support of ancillary type data.
- Measure which is not even indirectly affected by satellite or aircraft inputs.

(4) Benefit of Remote Sensing inputs to Decision Process.
Inferred based on inputs.
D - Direct not utilizing other inputs.
I - Indirect utilizing other inputs.
N - Not Obvious.

undertook an evaluation of current Forest Survey requirements to determine if we had identified a valid problem.

The initial activity of any survey involves the measurements of the areas present in each class of interest. For discussion purposes these classes will include: forest, water, and other. Once acres have been estimated for an area, sample points will be allocated within the area and identified by photo-interpretation. The allocation of points may be proportional to areas or they can be optimized for some specific variable if enough information is available for the resource being studied.

The sample points thusly selected are further categorized as being:

a) commercial forest by timber type, commercial size class (i.e. poletimber or sawtimber) and stand density or stocking
b) non-commercial forest
c) non-forest with or without trees
d) water

Once points have been selected and interpreted a subset or sample of the points identified as forest are ground checked. Extensive information is collected about the merchantable as well as the growth characteristics of only the commercial forest plots. This ground information is then expanded back over the entire sample area to give an estimate of the total forest resource of the site. The above example assumes that the initial estimate of acreage from the aerial photo-interpretation was correct. Because survey rarely enjoys the luxury of new aerial photos (new being dated within a year of the survey), acreage estimates must be adjusted to reflect changing land use patterns. Since Landsat data is theoretically available on a timely basis, changes in land use trends would be identified on this data and allocation of the photo and field samples should therefore be improved although not optimized. The basic requirement would then be that the acreage estimate provided from the Landsat data analysis must be within the current accuracy standards for the survey.

Hypothesis. The salient question which therefore remains to be answered can be stated in two parts:

1. Are the acreage-estimates from machine-assisted Landsat analysis compatible to existing survey standards?
2. If acreage estimates from machine-assisted analysis of Landsat data is not compatible to existing survey standards, are the results repeatable and of sufficient quality to input into the inventory design?

Obviously, if the response to both questions is negative, machine-assisted analysis of Landsat data would not be a suitable input to the inventory design. The probability of not utilizing Landsat inputs appear to be low based on results reported by various investigators using Landsat data. Determining which of the two parts of the question is relevant will be the problem addressed by the end of the contract period.

Design Definition. Given that either of the above questions can be positively answered an approach can be defined which addresses specific inventory needs; e.g. gross fiber volume, species composition, etc. Figure 1 is a schematic which identifies how an inventory may flow from general Landsat derived information to more specific inventory needs. Table 2 lists the accuracy requirements which the Landsat estimates must meet if they are to be considered for Forest Survey use. Also the table identifies boundary conditions which should be met if remote sensing and satellite data can be expected to input into Forest Survey or RPA needs. The remainder of this report will address the feasibility of utilizing Landsat results for the first level of inventory intensity.

IV. Data

The hypothesis we defined required that we have access to acreage statistics for gross forest land for comparison to Landsat classification results. Furthermore, we needed information for forest land that was recorded in a form similar to our Landsat results which were classified on a county-by-county basis. Forest Survey, collects and reports forest acreage statistics, in addition to other information, on a county basis as a part of their repetitive State Survey cycle. Forest Service resources bulletins report this information by county as acres in commercial and noncommercial forest cover, species mix, gross volume, and volume distribution by timber size class. For this study we were only concerned with obtaining information on a county-by-county basis for gross forest acres.

Landsat classification results for a number of counties were available at LARS from two previous research projects. Data were available for parts of States comprising the boundary of the Great Lakes Watershed and for a block of counties in north-central Missouri. Statistics from these results included a forest class as a major component of the Landsat classification. Furthermore, these data were analyzed on a county-by-county basis. Therefore, acreage or at least the percent of the area for each county in a forest class could be directly compared to Forest Survey published estimates for these same areas.

Table 3 indicates the states and years of the most recent published Forest Survey data. Table 4 gives similar information relating to the type and data of the Landsat classification results used for the comparison. Figure 2 indicated the geographic area for which our comparison would be made.

Given the data available, we would be able to compare acreage statistics and make inferences regarding our hypothesis. The following comparisons were possible:

1. The comparison between Forest Survey and Landsat estimates for gross forest acreage by county.
2. The comparison between Forest Survey and Landsat estimates for gross forest acreage by Survey Unit.
3. The comparison between Forest Survey and Landsat estimates for gross forest acreage for the State of Michigan.

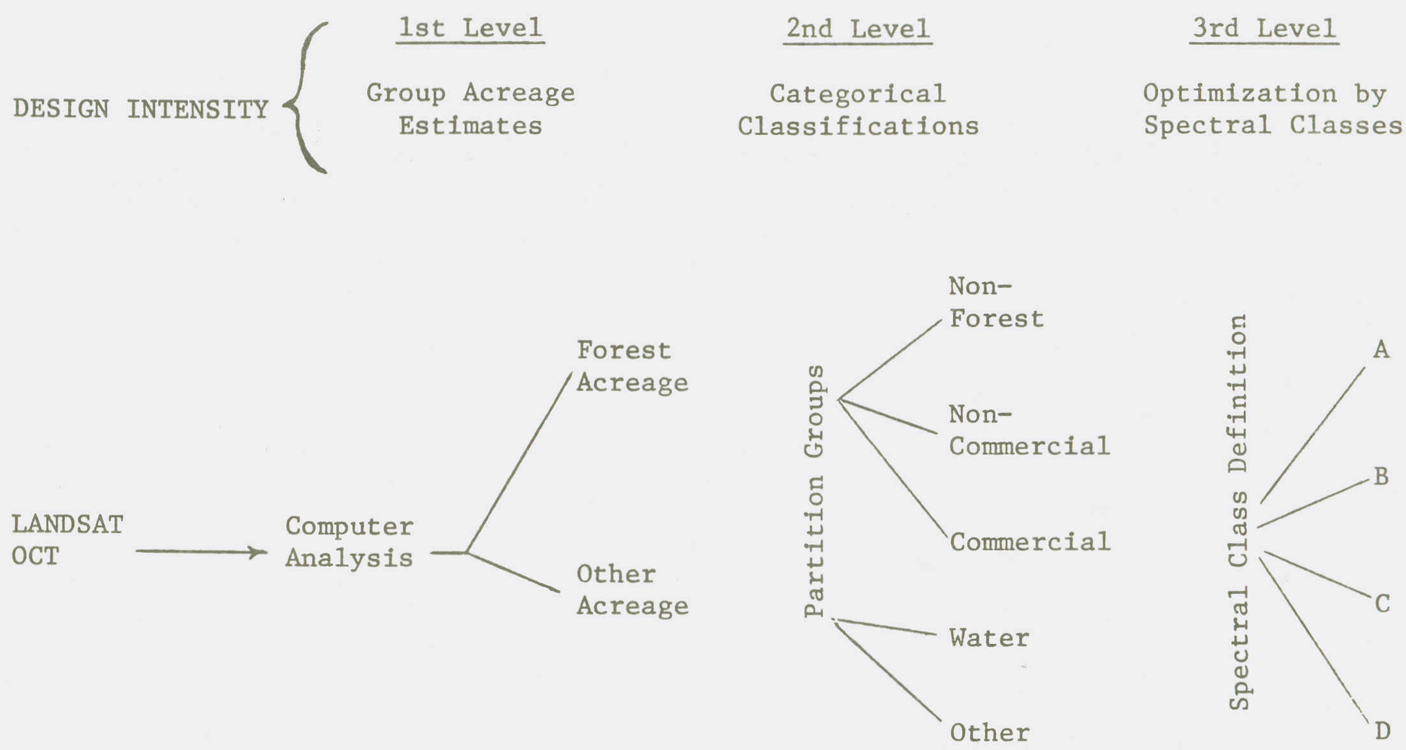


Figure C-1.Schematic flow anticipated in design definition.

Table C-2. Evaluation criteria and boundary conditions necessary for Landsat Survey design.

LANDSAT FOREST SAMPLE DESIGN GUIDELINES

A. ALLOWABLE ERROR TO MEET FOREST SURVEY REQUIREMENTS
C.V OF 3% PER MILLION ACRES FOR COMMERCIAL FOREST
C.V OF 10% PER MILLION FOR NON-COMMERCIAL FOREST

B. DESIRABLE BOUNDARY CONDITIONS FOR LANDSAT SURVEY TO MEET FOREST SURVEY REQUIREMENTS.

1. OVERALL SURVEY EFFICIENCY MUST BE EQUAL TO OR BETTER THAN EXISTING SURVEYS WITH REGARD TO:
   A. TIMELINESS OF DATA
   B. TIMELINESS OF ACREAGE ESTIMATES
2. PHYSICAL LIMITATIONS OF SURVEY
   A. CLASSIFICATION RESULTS MUST BE REPEATABLE FOR LARGE AREAS
   B. DATA THROUGH PUT SHOULD NOT BE A LIMITING FACTOR
3. A PRIORI INPUTS BASED ON ANCILLARY INFORMATION SHOULD BE USED TO:
   A. IMPROVE SPEED OF CLASSIFICATION
   B. IMPROVE CLASSIFICATION ACCURACY

Table C-3. Forest Survey Reports containing gross forest acreage data which was used for comparison with Landsat classification of forest area.

Forest Survey Region	Survey Date	State	Survey Units	No. of Counties	Survey Report
North Central	1966	MI	4	83	Res Bull NC-9
	1970	MO	5	114	Res Bull NC-30
	1968	WI	8	71	Res Bull NC-15
Northeastern	1968	NY	8	55	Res Bull NE-20

Table C-4 States within Forest Survey Regions which had Landsat classification results for forest acreage in a format comparable to published Forest Survey Statistics

Forest Survey Region	Year of Data	State	No. of Survey Units Analyzed	No. of Counties Analyzed	No. of Analyst
North Central	72-73	MI	4	83	10+
	74	MO	4	21	1
	72-73	WI	3	30	7+
Northeastern	72-73	NY	5	25	4

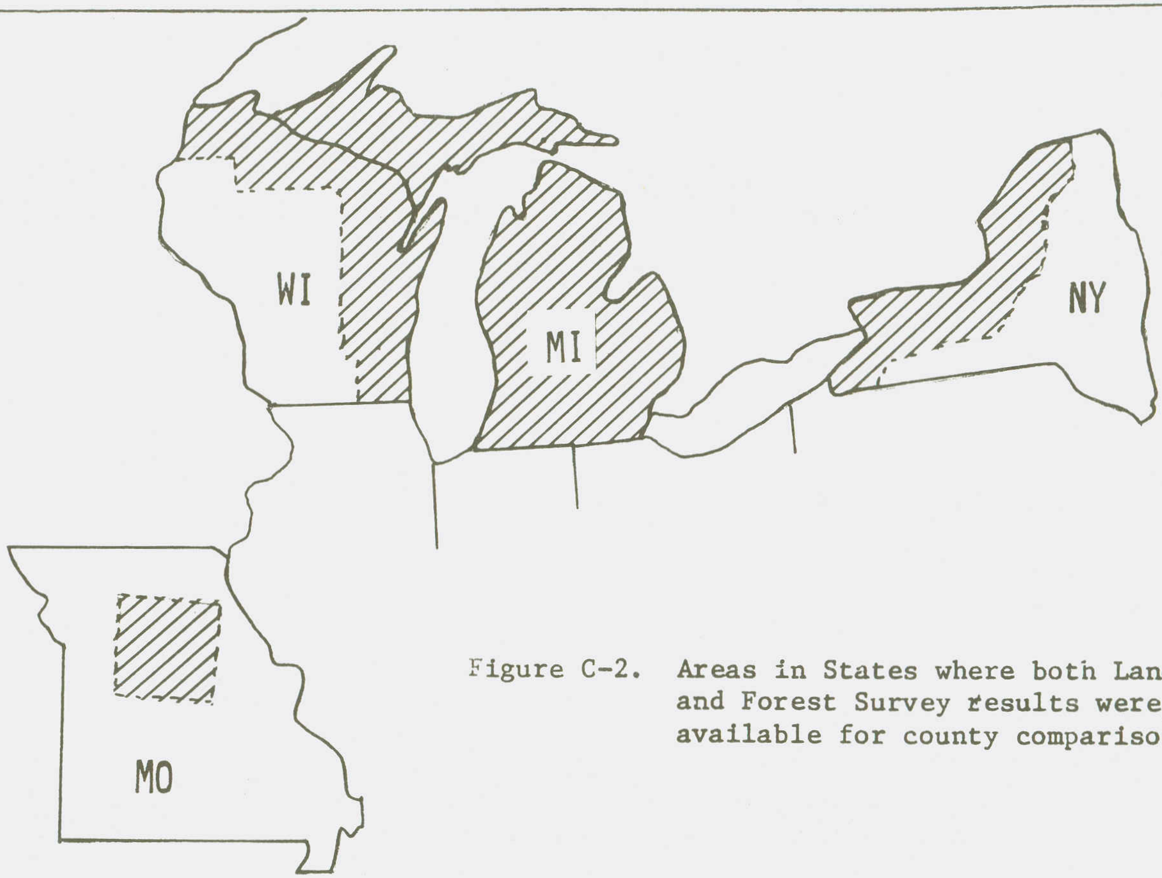


Figure C-2. Areas in States where both Landsat and Forest Survey results were available for county comparisons.

V. Results

Correlations. Forest Survey data and Landsat data were compiled in such a way so that the percent of gross forest area was comparable on a county-by-county basis. Since the minimum reporting unit used by Forest Survey is the county, we felt that areal Landsat estimates on a county basis would be a valid test. Figures 3,4,5 and 6 are scattergrams indicating the correlation between the Landsat estimate and the reported Forest Survey estimates. For each comparison, except in the case of Figure 4, the State of Missouri, the high value for the correlation coefficient indicated that the Landsat estimate of gross forest acreage is highly correlated with the published Forest Survey estimate.

Statistically, one would expect that for the number of counties studied the under-and over-estimates of forest acreage would average so that the high correlation coefficient could be expected. Indeed, if one studies the scattergrams, the amount of under-and over-estimation by the Landsat results becomes apparent. However, one may expect to see the same type of scatter if one were to compare Forest Survey estimates to actual forest acreage on a county-by-county basis. Survey in effect reports estimates derived from a coarser sample than Landsat. A more indicative measure of survey accuracy can be derived by comparing state wide and survey unit statistics.

Therefore, we realized that it would be necessary for us to draw comparative statistics on a forest survey unit basis and if possible on a state wide basis. Fortunately, we were able to combine enough county estimates to make a comparison for all or parts of 16-survey units in the four states that we had data available for study. Figure 7 shows a scattergram of this analysis. Again, the high correlation coefficient of r = .94 strongly suggest that Landsat estimates for gross forest acreage are highly correlated with Forest Survey estimates of gross forest acreage on a Survey Unit basis. Generally, one could conclude from this graph that Landsat tends to overestimate gross forest acreage except for two Survey Units where that acreage is grossly underestimated. These two cases occurr in the State of Missouri which will be discussed later.

Unfortunately, only state-wide data were available for one state. A complete Landsat analysis was available for the State of Michigan. These data were used to make a comparison of estimates by county, survey unit, and state-wide from Landsat and Forest Survey. Table 5 shows the results of that comparison broken down on an acreage basis. The last column indicates the difference between the Landsat and Forest Survey estimates. A negative number indicates a Landsat over-estimate compared to Forest Survey whereas a positive number would indicate a Landsat underestimate. Table 6 summarizes the comparison of Survey Unit statistics and state wide statistics for Michigan. Generally, we interpret these results in the table to indicate that the counties in the Upper Peninsula are under estimated by Landsat classifications in the magnitude of 6-9%. Whereas in the Lower Peninsula the Northern Survey Unit is almost equal to the Forest Survey estimate, whereas the Southern Unit is grossly over-estimated. Comparing the state-wide totals, Landsat compares favorably with Forest Survey estimates the differences being

Figure C-3. Scattergram comparison of Landsat and Forest Survey estimates of forest area for the State of Michigan.

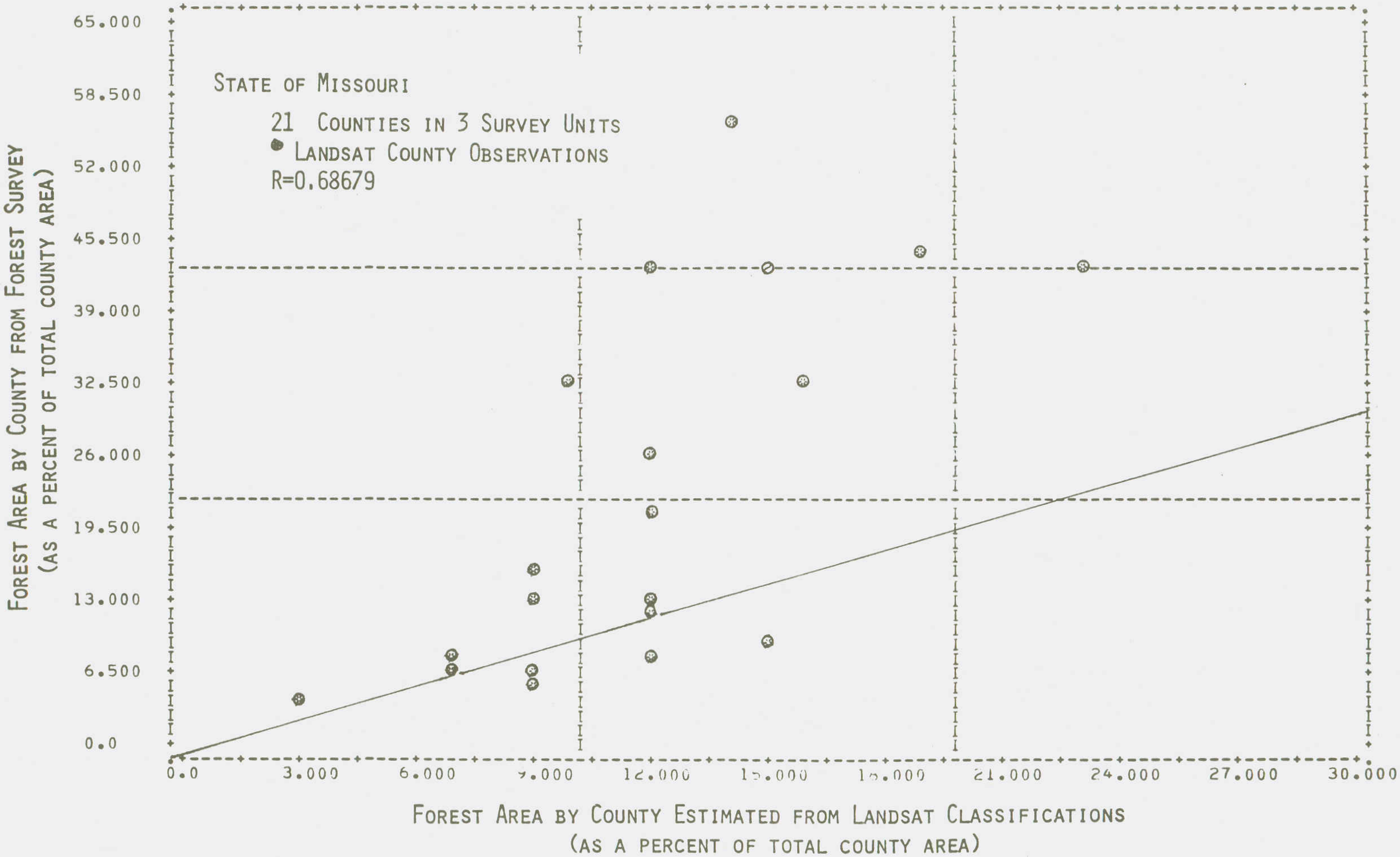


Figure C-4. Scattergram comparison of Landsat and Forest Survey estimates of forest area for the State of Missouri.

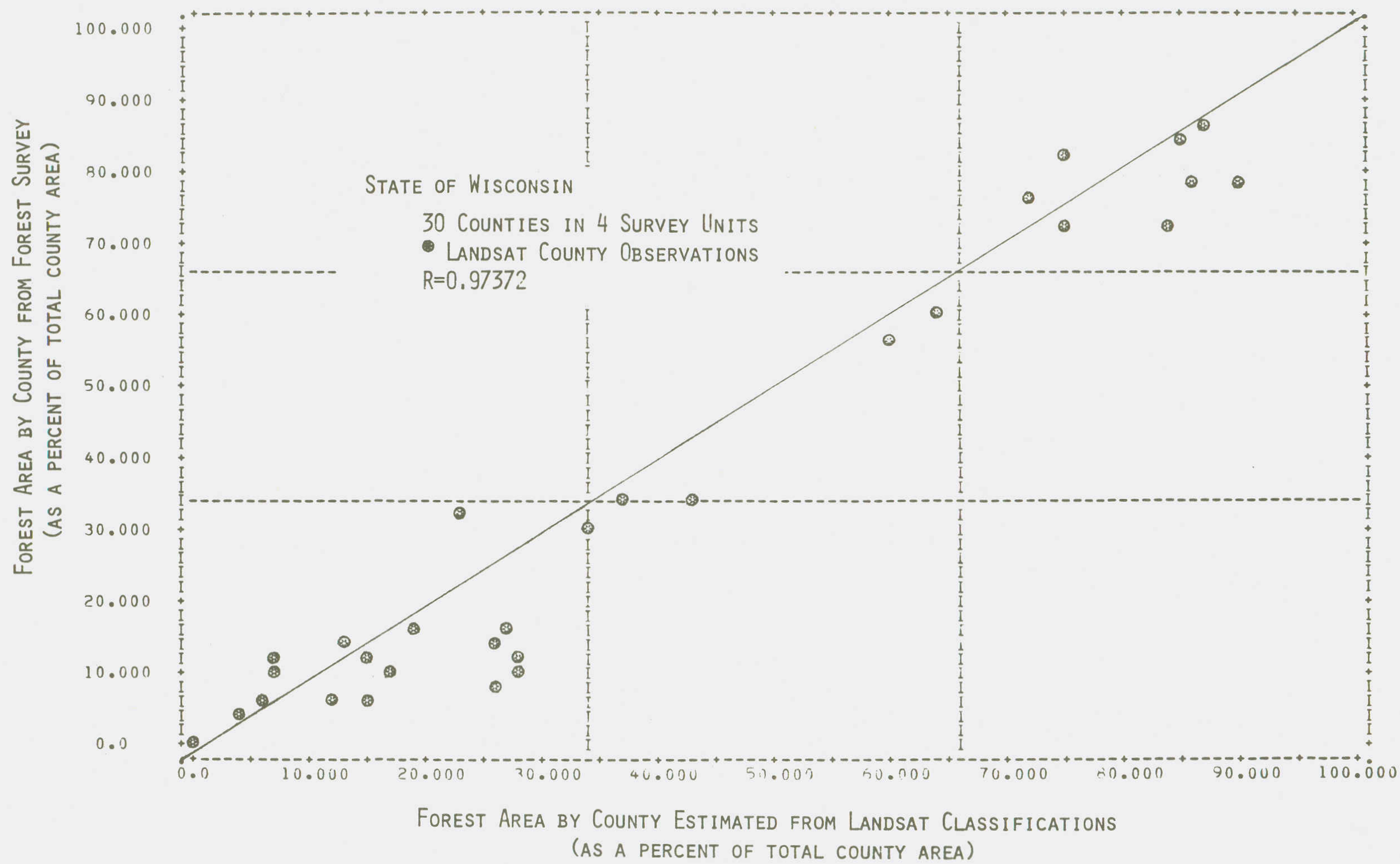


Figure C-5. **Scattergram comparison of Landsat and Forest Survey estimates of forest area for the State of Wisconsin.**

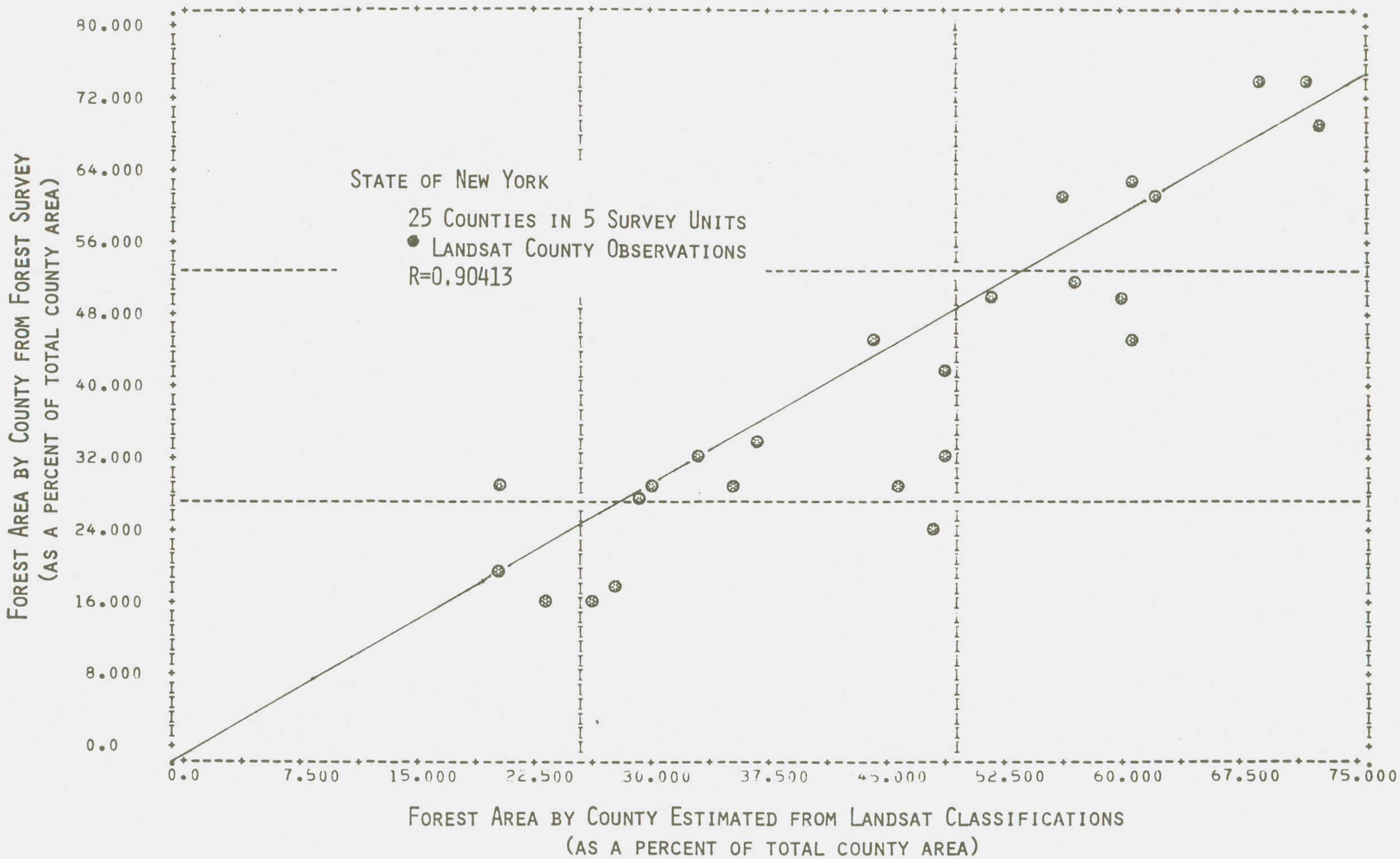


Figure C-6. Scattergram comparison of Landsat and Forest Survey estimates of forest area for the State of New York.

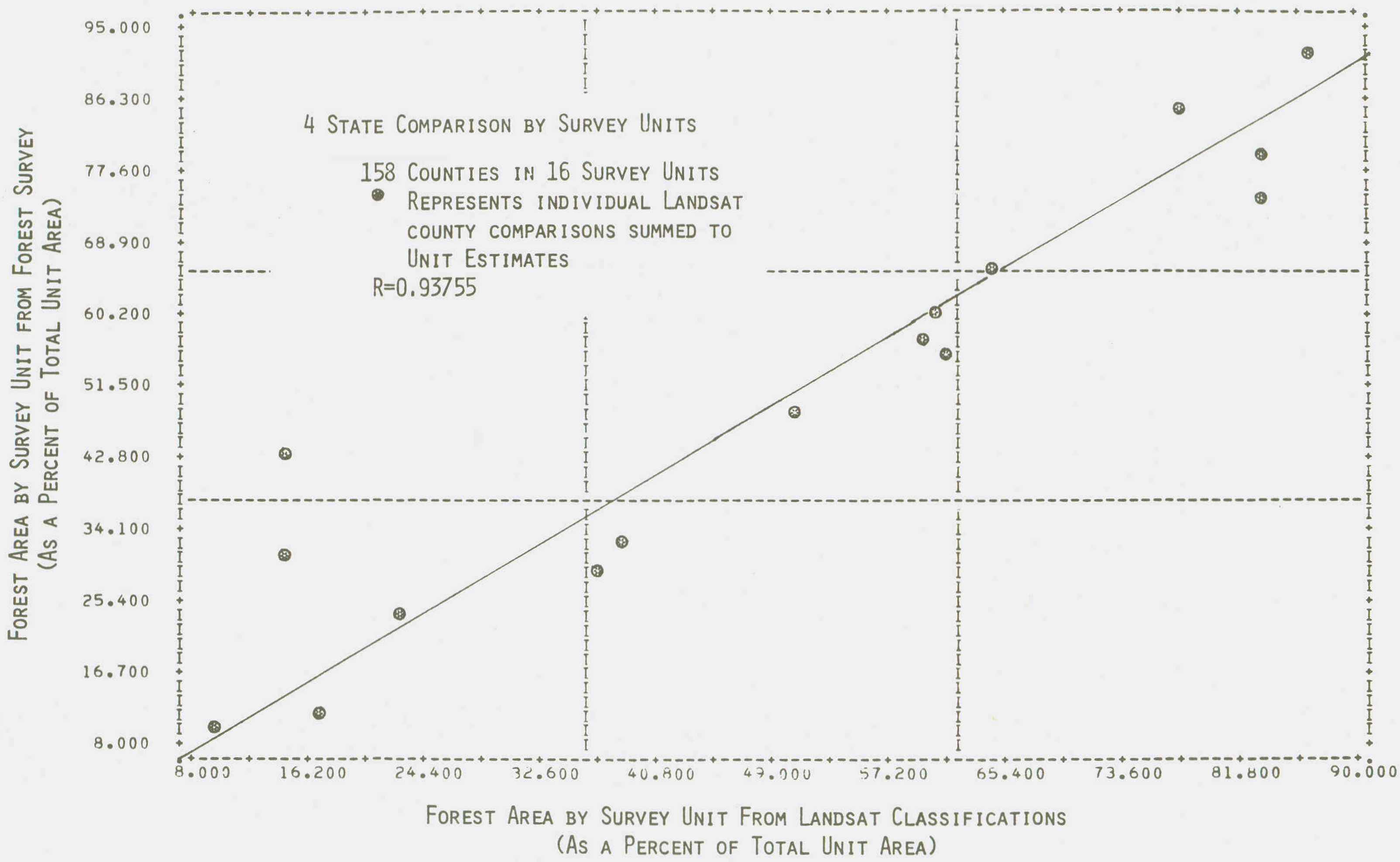


Figure C-7. Scattergram comparison of Landsat and Forest Survey estimates for Forest Survey Units comprising 158 counties in four States.

Table C-5. Comparison of Forest Survey[1] and Landsat[2] Estimates of Forest Acreage for the State of Michigan by Forest Survey Unit and County.

WESTERN UPPER PENINSULA UNIT

County:	Gross Forest[3] Land (In thousands of acres) Forest Survey	Landsat[4]	Difference[5]
Baraga	543.3	551.9	-8.6
Dickinson	454.0	437.5	16.5
Gogebic	666.4	611.4	55.0
Houghton	576.9	583.4	99.6
Iron	708.7	609.1	
Keweenaw	341.2	280.7	60.5
Marquette	1,108.6	1,072.2	36.4
Ontonagon	784.2	713.4	70.8
TOTALS	5,183.3	4,859.6	323.7

EASTERN UPPER PENINSULA UNIT

County:	Gross Forest[3] Land (in thousands of acres) Forest Survey	Landsat[4]	Difference[5]
Alger	540.6	454.6	86.0
Chippewa	805.1	675.5	129.6
Delta	638.3	654.8	-16.5
Luce	543.3	508.4	34.9
Mackinac	572.0	474.4	97.6
Menominee	527.5	411.5	116.0
Schoolcraft	654.5	679.2	-24.7
TOTALS	4,281.3	3,858.4	422.9

Table C-5 Cont.

NORTHERN LOWER PENINSULA UNIT

County:	Gross Forest[3] Land (In thousands of Areas) Forest Survey	Landsat[4]	Difference[5]
Alcona	324.1	319.3	4.8
Alpena	212.2	255.2	-43.0
Antrim	186.4	156.6	29.8
Arenac	106.4	97.7	8.7
Bay	48.3	28.3	20.0
Benzie	124.7	130.6	-5.9
Charlevoix	162.6	223.7	-61.1
Cheboygan	340.3	370.7	-30.4
Clare	230.3	239.4	-9.1
Crawford	313.6	320.3	-6.7
Emmet	189.4	246.6	-57.2
Gladwin	189.7	140.3	49.4
Grand Traverse	161.7	143.2	18.5
Iosco	243.9	208.7	35.2
Isabella	97.9	85.3	12.6
Kalkaska	271.1	322.0	-50.9
Lake	297.1	217.1	80.0
Leelanau	114.2	100.3	13.9
Manistee	221.3	204.6	16.7
Mason	160.1	118.9	41.2
Mecosta	149.6	113.5	36.1
Midland	171.7	180.7	-9.0
Missaukee	227.4	245.8	-18.4
Montmorency	287.5	306.9	-19.4
Newaygo	311.7	311.6	0.1
Oceana	152.3	115.9	36.4
Ogeman	241.4	275.9	1.9
Osceola	176.1	163.2	12.9
Oscoda	302.3	325.1	-22.8
Otsego	251.8	261.5	-9.7
Presque Isle	290.8	310.5	-29.7
Roscommon	271.9	260.7	11.2
Wexford	221.9	197.8	24.1
TOTALS	7,051.7	6,997.9	53.8

Table C-5. Cont.

SOUTHERN LOWER PENINSULA UNIT

County:	Gross Forest Land[3] (In thousands of acres) Forest Survey	Landsat[4]	Difference[5]
Allegan	150.7	168.7	-18.0
Barry	94.4	141.9	-47.5
Berrien	71.0	75.7	-4.7
Branch	56.6	71.9	-15.3
Calltoun	99.7	91.2	8.5
Cass	69.7	98.7	-29.0
Clinton	45.0	49.2	-4.2
Eaton	55.9	50.1	5.8
Genessee	56.2	160.5	-104.3
Gratiot	49.4	45.6	3.8
Hillsdale	70.3	74.6	-4.3
Huron	62.7	55.8	6.9
Ingham	58.6	67.6	-9.0
Fonia	63.5	72.9	-9.4
Jackson	96.2	117.3	-21.1
Kalamazoo	79.9	146.9	-67.0
Kent	129.3	127.9	1.4
Lapeer	83.8	92.5	-8.7
Lanawee	64.8	41.9	22.9
Luingston	92.5	117.7	-25.2
Macomb	46.5	40.0	6.5
Montcalm	134.4	110.7	23.7
Monroe	35.1	24.1	11.0
Muskegon	170.0	138.4	31.6
Oakland	154.6	164.5	-9.9
Ottawa	89.1	105.1	-16.0
Saginaw	99.0	104.5	-5.5
St. Clair	79.4	67.3	12.1
St. Joseph	57.7	77.7	-20.0
Sanilac	70.4	76.9	-6.5
Shiawassee	42.5	83.6	-41.1
Tuscola	105.7	156.1	-50.4
Van Buren	100.1	76.9	23.2
Washtenaw	84.2	137.9	-53.7
Wayne	38.2	47.0	-8.8
TOTALS	2,857.1	3,279.3	-422.2

TABLE C-5. FOOTNOTES
(Cont.)

1. Michigan Forest Survey results reported in USDA-Forest Service Research Bulletin NC-9

2. Land use Acreage results reported for Michigan in LARS Information Note 011077

3. Gross Forest includes figures for commercial, non-commercial and productive reserved acres

4. Landsat forest acreage estimates were corrected for U.S. Bureau of Census county acreages

5. Difference = (Forest Survey reported county acreage)-(Landsat estimated acreage)
   + = Landsat under estimate compared to Forest Survey
   - = Landsat over estimate compared to Forest Survey

Table C-6. Summary. Comparison of Landsat estimates for gross forest acreage to Survey Unit acreage figures published for the State of Michigan in USFS Resource Bulletin NC-9, 1970.

Survey Unit	Gross Forest Acres (in thousand acres)		Differences (in percent)
	1966 Forest Survey	1973 Landsat Analysis	
Western Upper Peninsula	5,183.3	4,859.6	-6.3
Eastern Upper Peninsula	4,281.3	3,858.4	-9.8
Northern Lower Peninsula	7,051.7	6,997.9	-0.8
Southern Lower Peninsula	2,857.1	3,279.3	+14.8
State Totals	19,373.4	18,995.2	-1.95

a negative 1.95%. This result is well within the 3% maximum imposed by current survey restrictions.

Scattergrams for both New York and Missouri do not follow the same pattern as do the scattergrams for Wisconsin, Michigan and the Survey Unit totals. In New York Landsat estimates generally overestimated those reported in the 1968 Forest Survey Resource Bulletin. This was concerning since the analysis for New York state had the same objectives as did the analysis for Michigan and Wisconsin. Therefore, we would assume a similar pattern in the scattergrams. Upon investigation of the Forest Service publications, NE-20, we learned that the State of New York, had experienced an increase in gross forest acreage between the 1950 and 1968 survey periods. If we could assume that forest acreage was still on the increase, and we have no reason to doubt otherwise, we could rationalize that the Landsat estimates of forest acreage would be over those reported in the 1968 survey. To support this, we decided to make a projection of one survey unit. The Lake Plain Survey Unit area statistics were expanded to determine a projected rate of growth between 1968 and 1974 in gross forest acreage. Our assumption, was that the forest acreage would continue growing at the same rate as it had grown between the 1950 and 1968 survey periods. Table 7 shows the results of our projection for the Lake Plain Survey Unit of New York. The change column indicates that the gross forest land area would have increased by 7% between the 1968 and the projected 1974 survey period. The Landsat estimate of forest acreage increase between 1968 and the time the Landsat data was classified indicated a 6.5% increase. Unfortunately, a more complete data set was not available for comparisons. Similar results were not available for survey units in the North Central Region for the States of Michigan, or Wisconsin or Missouri so similar comparisons could not be made.

The results for the State of Missouri show in general that the Landsat data grossly underestimates forest acreage compared to the published Forest Survey results. The counties in Missouri were not analyzed with the same objective as those in Michigan, Wisconsin and New York. However, Missouri was unique in that more control was imposed on the analysis than in the other States, since primarily only one analyst was responsible for developing training statistics. This result for Missouri was counter to our belief that analysts should have a greater effect on the classification accuracy, and, therefore, the acreage comparisons.

The correlation study indicated that a more detailed evaluation of the classifications would be in order. We reviewed our knowledge of the material that we were using for the comparative evaluations. For the data from the Great Lakes Watershed classifications we estimated that we had sufficient information to evaluate:

1. The effect that a number of different analysts had on the acreage comparisons given that they were all following a uniform classification objective.
2. The effect on acreage comparisons of extending training data to adjacent counties.

Table C-7. Comparison of a Landsat estimate of forest acreage with a projected estimate of forest acreage from one survey unit in New York.

	Acres (in Thousand acres)		Projected Acres	Change (in percent)	
Survey Unit	1950 Survey	1968 Survey	1974	1974 projected	1974 Landsat
Lake Plain	1440	1889	2038	+7	+6.5

Projection based on regression estimate assuming constant rate of forest growth through time.

3. The effect of commercial forest acreage on training set selection and if this variable can affect the predicted outcome of Landsat classifications.

Analysis-of-Variance. An analysis-of-variances test was run using the difference between the Landsat estimate of percent gross forest by county and the Forest Survey published estimate for each of the variables listed above. The objective of this analysis was to determine how the three variables being tested affected classification performance, or our ability to used Landsat data to estimate forest cover. The test, we felt, would highlight areas needing more intensive study.

In general, for the purpose of this study the analysis-of-variance raises more questions than it is able to provide answers. Furthermore, the inferences that one can make from these statistics follow intuitive lines of logic to those familiar with Forest Survey procedures and remote sensing technology.

A summary of our results indicate:

- For the three variables tested (analysts, data set extension, and training set) for the three States it was not obvious which variable was most significant with regard to classifying forest cover.
- The ranking of the variables significance within States varied for each state.
- The effect of between variable interaction was inconsistent from State to State.
- Generally it appears that training set may be the most significant variable.
- Possibly the effect of analyst, given a well defined analysis is the least significant variable.

VI. Conclusions

Based on our study to determine if Landsat data can be used to help implement a large area forest resource inventory such as required by the RPA we can conclude:

- Computer-assisted analysis of Landsat data can be used as an estimator of gross forest acreage.
- Landsat acreage estimates can be derived on county-by-county bases, given sufficient resources to adequately define training statistics.
- Landsat data and computer-assisted analysis can provide State wide estimates of forest acreage within the current Forest Survey requirements.
- Training statistics cannot be extended across Survey Unit Boundaries.

Given the foregoing we conclude that we have positively answered our hypothesis. We can therefore state that Landsat data can be used as a tool to derive forest acreage.

VII. Discussion

One must remember, that for this comparison the Landsat classification results were not specifically generated for the purpose of separating forest from non-forest types. Therefore, we have attempted, where feasible, to present comparisons for gross forest acreage. By so doing we have knowingly lumped Forest Survey classes of commercial, non-commercial and productive reserved lands together. Our primary objective was to determine if Landsat data could be used in a coarse sense to provide up-to-date acreage statistics We have accomplished this objective.

Logically, our next step would be to progress to the next level of survey detail. However, we feel that such a progression would be premature. We feel that these results indicate that our ability to predict the outcome of Landsat classifications are below par. That is to say we have not been able to identify the sources of variability that affect classifications accuracy in terms of accurately predicting surface area of forest area by county.

The study indicates that there are a number of variables that need to be studied and understood before a Landsat-based inventory system can be defined. A partial list of these variables follow:

- The effect of data date on classifications performance.
- The effect of size and locations of training set data.
- The location and distribution of training data based on a-priori inputs from:
    1. previous forest surveys
    2. potential natural vegetation
- The predictability of spectral class structure based on previous classification experience.
- The utility of Forest Survey data for allocation of training data. This is now possible since Forest Survey now records a land class for a variable area plot around the field point. Survey data we had access to only recorded information for the one-acre field point.
- The distance that training statistics may be extended within Survey Units.

Once these areas of concern have been addressed we may proceed to work on providing more detailed information from the Landsat classifications. Likewise, the design for using Landsat classifications in a sample model could be approached and design test could be performed that would determine the data's ultimate utility.

VIII. Recommendations

Based on the foregoing we recommended the following:

- Work is needed to determine the extent to which Landsat and computer-assisted analysis techniques are capable of estimating forest acreage.
- Work to determine the allocation of training samples.
- Work to determine the efficiencies involved in classifying difference between the forest and non-forest categories using different classification techniques.
- Work in classifying spectrally, the difference between commerical and non-commercial forest classes.

For preliminary design considerations:

- We recommend the use of survey units in defining the final Landsat sampling procedure.
- We strongly recommend the use of a-priory knowledge contained in potential vegetation maps for defining the sample intensity.
- We strongly recommend the evaluation of training sample size and distribution based on the predominate land use in a survey unit.

COMPARISON OF CLASSIFICATION TECHNIQUES

Introduction. Within the past several years, many advances have been made in the use of computer-aided analysis techniques for classifying and mapping forest cover. As the previous section of this final report has indicated, there are several approaches that can be used in developing an optimal set of training statistics. However, once a satisfactory set of training statistics has been defined, there are several different approaches that can be used to classify the data. During the past several years, several different classification techniques have been developed, each of which has its own characteristics, advantages, and disadvantages. This phase of the FAP project has therefore involved the testing and comparison of several of these different classification techniques.

Objectives. The objective of this study is to compare five different computer classification techniques in terms of their cost, accuracy, and characteristics of the output products. Classification techniques to be compared include the (a) Perpoint classifier, (b) ECHO classifier, (c) Layered classifier, (d) Minimum Distance to the Means classifier, and (e) Levels classifier.

Materials and Methods. Previous studies have shown that where large geographic areas are involved a variety of techniques can be used to develop the training statistics. However, once the training statistics are defined, one of several different supervised techniques can be utilized to classify the data. Many of the supervised classification techniques that have been developed utilize the powerful maximum likelihood approach based upon normal distribution of the data, although there are other criteria that can be used for classification algorithms to select the identification of a particular pixel of data. Even when using the maximum likelihood approach, we find that several different types of maximum likelihood classifiers have been developed and these vary in the number of decisions that must be made for each pixel or group of pixels. The following paragraphs will briefly describe the characteristics of each of the classifiers tested in this study.

(A) The Maximum Likelihood Perpoint Classifier. This is often referred to simply as the Perpoint classifier and is the simplest and most straightforward of the classifiers tested. This classifier calculates the probability of a particular pixel belonging to each of the training classes specified using the means, variances, and correlation matrices for a specified set of spectral channels, and assuming Gaussian distribution of the training data. The training class having the maximum likelihood (highest probability) is selected and the pixel being classified is assigned to that particular class by the algorithm. A Bassian classification approach using the maximum likelihood perpoint classifier can be achieved by using a priori probability weightings. Use of the maximum likelihood perpoint classifier involves only one decision for each pixel, which is to determine the most probable class for that pixel. This classification algorithm does assume that the distribution of training data is Gaussian (i.e. normal) and that the training sample is of sufficient size to accurately estimate the statistical parameters. Each pixel of data is systematically classified. The result is stored on a tape containing the classification decision for each pixel and also the probability of that pixel being in the class to which it was assigned. The probability data allows later use of the thresholding option in evaluating the classification results.

(B) A Maximum Likelihood Boundary Finding Per-field, or "ECHO" Classifier. "Simultaneous analysis and classification of a group of pixels, all assumed to be drawn from the same class, has been shown to be a powerful method for incorporating simple spatial information (adjucency) into the analysis process." (Kettig and Landgrebe, LARS Information Note 050975). By automatically defining a group of pixels (sample location) and then classifying each group as a unit (sample classification), both the accuracy and efficiency of the data analysis tasks can be increased. An approach combining both sample location and sample classification has been developed at LARS and has been referred to as the "ECHO" classifier which is an acronym, for "Extraction and Classification of Homogeneous Objects" (Swain, LARS Information Note 111276).

The ECHO classifier is a two step classification algorithm. First homogeneous areas or cells are defined by the computer and then each field or cell is classified as a unit. For the first step, three parameters must be defined by the analyst. First, the cell size must be selected by the analyst (i.e. 2 x 2, 3 x 3, etc.). The second parameter determines if the cell is significantly heterogeneous (i.e. whether the cell must be split or not); the second parameter determines if adjacent cells are similar and can be combined to form a larger field. The classification algorithm then makes a single decision on the identification of each field, using the maximum likelihood decision rule. If there is more than one pixel in a cell (which is usually the case), in addition to the means, variances, and correlation metrices of the training classes, this classifier uses the means, variances, and correlation matrices for the field in the classification step (a per-field approach). As with the maximum likelihood perpoint classifier, the ECHO classifier only uses one set of channels (specified by the analyst) for use throughout the analysis sequence.

(C) A Maximum Likelihood Multi-layer Perpoint, or "Layered" Classifier. The Layered classification algorithm differs from the normal maximum likelihood perpoint classifier in terms of the number of decisions that are made for each pixel. Instead of a single decision as to the identification of that pixel, several decisions are made, on a pixel by pixel basis. The multi-decision nodes allow different subsets of channels to be used at the various nodes, and fewer choices among spectral classes are involved in each decision. Although a maximum likelihood decision rule is utilized at each of the nodes, when the classification is completed no single probability value can be assigned to a particular pixel. However, it should be noted that a priori probability weighting can be utilized with the layered classifier. The part of the classifier that is most difficult to develop and is least understood is the "tree structure" which is required to define the relationships between the various decision points (or nodes) and the informational and spectral classes involved. It should be noted that the layered classifier and the tree type of structure lend themselves very well to analysis of multitemporal and multi-scource data sets.

(D). A Minimum Distance to the Means Perpoint Classifier. One of the two classifiers compared in the study that does not utilize the maximum likelihood decision rule is the so-called "Minimum Distance Classifier". In this classifier, a minimum Euclidean distance is found between the data vector being classified and the mean of each class of training statistics classified. Therefore, this classifier does not use variances and correlation matrices (this is equivalent to assuming that the variances are all equal and there is no correlation between the channels). Since the maximum likelihood rule is not utilized, probability values cannot be obtained for each pixel nor can a priori probability weighting be utilized. The algorithm systematically makes a pixel by pixel classification of the entire data set. As was the case in the first two classification algorithm techniques discussed, a single decision is made for each pixel, using a single set of channels specified by the analyst.

(E). A Parallel Piped Perpoint, or "Levels" Classifier. The final classification technique tested in this study was the so-called levels classifier. The major difference between the levels classifier and the other classification algorithms discussed thus far is that the levels classifier utilizes the ranges of spectral values in all channels to define each spectral-informational class. These ranges of spectral values for each channel to be utilized must be specified to the computer by the analyst. By not utilizing the means, variances, and correlations matrices, probabilities cannot be obtained and a priori probability weighting cannot be utilized. Any pixel that has a data vector which is not included in any of the ranges of the classes specified is put into an "other" class. For this reason, the levels classifier can produce a result in which many pixels are not included in any of the spectral classes defined. This is quite different than the other classification techniques discussed since they will all classify every pixel into one of the classes defined, even though the probability of actually belonging to that class is very low. The levels classifier utilizes a single decision to classify data on a pixel by pixel basis, using a single set of channels specified the analyst.

Results. The U.S.G.S. Platoro quadrangle area was classified with each of the five classification algorithms, at both the Level 2 and Level 3 degrees of detail. Because of the nature of the test site, some of the spectral classes defined as Level 3 included some deciduous species in with the premarily coniferous cover types. Because the spruce-fir cover type could be subdivided as a function of crown closure, these classifications actually involved a higher level of detail than is normally considered to be included in a Level 3 classification. The informational classes defined for the Level 2 and Level 3 classifications are shown in Table 8.

Table C-8. Informational classes utilized in the computer-aided analysis of the Platoro quadrangle for both level 2 and level 3+ degrees of detail.

Level 2	Level 3+
Barren	Barren (Exposed rock and soil)
Water	Water
Grassland	Range land Agricultural
Deciduous	Aspen Oak
Coniferous	Ponderosa Pine/Oak Spruce-fir/Aspen Spruce-fir, 60-80% Crown Closure Spruce-fir, 80-100% Crown Closure

To improve the comparison of the results obtained, as many variables as possible were maintained constant. The same Landsat MSS data set was used for all classifications (Frame 1424 - 17132, collected on September 20, 1973). A single statistics deck was developed using a multi-cluster blocks approach (described in the previous section of this report) for the Platoro quadrangle. All of the classification algorithms use basically the same statistical parameters contained in the statistics deck (the means and variances), with the exception of the Levels Classifier, for which the spectral ranges had to be determined. Therefore, for the Levels classifier, the ranges for each class were defined as plus and minus 2½ standard deviations around the mean. A priori weights were not utilized with any of the classifiers, although this was an option available on several. Except for the layered classifier, all four channels were used for each decision in the classification process. The layered classifier utilizes a subset of channels for several nodes (decision points), so all four channels were eventually utilized; but not in a single decision step. All classifications were obtained on the IBM 370/138 computer using the software configuration currently available.

The results for each classification algorithm were evaluated relative to the amount of CPU time required, the extent of analyst involvement (in terms of man-hours), and estimates of the accuracy of the results. Classification of the entire 15,303 hectare (37,812 acres) Platoro quadrangle was used as the basis for comparing relative speed of the different algorithms. Even though the area is not particularly large, it does provide a valid relative index of the computer CPU time required for the various algorithms. The amount of analyst time required in this study was minimal in all cases except the Levels classifier, since it only involved setting up the classification decks using an already existing set of training statistics. The accuracy estimates were based upon a detailed set of test fields that had been defined and thoroughly field checked by personnel of the Institute of Artic and Alpine Research (INSTAAR) of the University of Colorado.

Table 9 contains the results summary for the various classifications. In each case, the CPU time given is based upon the classifications for the Level 3 degree of detail. Table 10 contains the individual preformance tables for the Level 2 test field results for each of the classification algorithms. The Level 3+ test field results for each of the classification algorithms are shown in Table 11. The following paragraphs will discuss the classification results for each algorithm in sequence.

(a). Maximum Likelihood Perpoint. This classification was the most expensive to obtain, primarily because of the large amount of CPU time required for the classification. The total cost was 0.218 cents per hectare. A relatively large amount of CPU time was required by the algorithm since the probability for each individual pixel belonging to each of the spectral classes had to be calculated. However, this procedure gives the algorithm considerable power and makes it relatively accurate. For most work involving computer-aided analysis techniques, this algorithm is considered the standard against which other techniques are compared because of its high level of accuracy, no matter what cover types are involved. The analyst involvement in classifying the data with this algorithm was minimal.

Table C-9. Summary of the costs and accuracies obtained for the five classification algorithms evaluated.

		Classification Algorithm Used: Maximum Likelihood perpoint	ECHO	Minimum Distance to the Means	Layered	Levels
TIME	CPU Time (sec.)[1]	449.4	267.4	168.5	258.6	144.7
	Man Hours (min.)[2]	10	15	10	30	120
COST[3]	Total Cost ($)	32.87	21.06	13.37	27.73	30.05
	Cost per hectare (¢)	0.218	0.140	0.089	0.184	0.200
OVERALL ACCURACY[4]	Level 2	93.5	92.8	93.5	92.8	68.0
	Level 3	75.0	73.0	75.9	74.5	53.6

[1] Based upon classification of the entire Platoro quadrangle (15,303 hectares or 37,190 acres).

[2] It should be noted that these times do not include the time (cost) for developing the training statistics, except in the case of the Levels Classifier.

[3] Calculated on the basis of @$250/CPU hour & $10/man-hour.

[4] Based upon evaluation of tested areas involving 3,704 acres and which have been carefully field checked.

Table C-10 Classification performance tables for each of the algorithms tested, at the Level 2 degree of detail.

(A) Maximum Likelihood Perpoint Classifiers

GROUP	NO OF SAMPS	PCT. CORCT	NUMBER OF SAMPLES CLASSIFIED INTO EXPOSED	GRASS	DECID	CONIFER	WATER	BADDATA
1 EXPOSED	41	63.4	26	1	0	14	0	0
2 GRASS	1068	88.2	95	942	15	16	0	0
3 DECID	127	47.2	1	59	60	6	0	1
4 CONIFER	1897	99.4	0	3	2	1885	6	1
5 WATER	234	100.0	0	0	0	0	234	0
TOTAL	3367		122	1005	77	1921	240	2

OVERALL PERFORMANCE(3147/ 3367) = 93.5

AVERAGE PERFORMANCE BY CLASS(398.2/ 5) = 79.6

(B) ECHO Classifier

GROUP	NO OF SAMPS	PCT. CORCT	NUMBER OF SAMPLES CLASSIFIED INTO EXPOSED	GRASS	DECID	CONIFER	WATER	BADDATA
1 EXPOSED	41	68.3	28	0	0	13	0	0
2 GRASS	1068	86.5	122	924	10	12	0	0
3 DECID	127	44.9	2	58	57	9	0	1
4 CONIFER	1897	99.2	1	4	3	1882	6	1
5 WATER	234	100.0	0	0	0	0	234	0
TOTAL	3367		153	986	70	1916	240	2

OVERALL PERFORMANCE(3125/ 3367) = 92.8

AVERAGE PERFORMANCE BY CLASS(398.9/ 5) = 79.8

(C) Minimum Distance Classifier

GROUP	NO OF SAMPS	PCT. CORCT	NUMBER OF SAMPLES CLASSIFIED INTO EXPOSED	GRASS	DECID	CONIFER	WATER	BADDATA
1 EXPOSED	41	58.5	24	2	0	15	0	0
2 GRASS	1068	89.6	70	957	22	19	0	0
3 DECID	127	48.0	0	57	61	8	0	1
4 CONIFER	1897	98.7	0	3	3	1873	18	0
5 WATER	234	100.0	0	0	0	0	234	0
TOTAL	3367		94	1019	86	1915	252	1

OVERALL PERFORMANCE(3149/ 3367) = 93.5

AVERAGE PERFORMANCE BY CLASS(394.9/ 5) = 79.0

Table C-10 Cont.

(D) Layered Classifier

GROUP	NO OF SAMPS	PCT. CORCT	EXPOSED	GRASS	DECID	CONIFER	WATER	BADDATA
			NUMBER OF SAMPLES CLASSIFIED INTO					
1 EXPOSED	41	56.1	23	1	0	17	0	0
2 GRASS	1068	87.7	100	937	15	16	0	0
3 DECID	127	47.2	0	59	60	6	1	1
4 CONIFER	1897	98.6	0	3	3	1871	20	0
5 WATER	234	100.0	0	0	0	0	234	0
TOTAL	3367		123	1000	78	1910	255	1

OVERALL PERFORMANCE(3125/ 3367) = 92.8

AVERAGE PERFORMANCE BY CLASS(389.7/ 5) = 77.9

(E) Levels Classifier

GROUP	NO OF SAMPS	PCT. CORCT	EXPOSED	GRASS	DECID	CONIFER	WATER	BADDATA	OTHER
			NUMBER OF SAMPLES CLASSIFIED INTO						
1 EXPOSED	41	51.2	21	1	0	9	0	0	10
2 GRASS	1068	56.6	53	605	5	15	0	0	390
3 DECID	127	27.6	0	40	35	5	0	0	47
4 CONIFER	1897	81.1	1	2	1	1538	11	0	344
5 WATER	234	38.0	0	0	0	0	89	0	145
TOTAL	3367		75	648	41	1567	100	0	936

OVERALL PERFORMANCE(2288/ 3367) = 68.0

AVERAGE PERFORMANCE BY CLASS(254.5/ 5) = 50.9

Table C-11. Classification performance tables for each of the algorithms tested, at the Level 3+ degrees by detail.

(A) Maximum Likelihood Perpoint Classifier

GROUP	NO OF SAMPS	PCT. CORCT	NUMBER OF SAMPLES CLASSIFIED INTO BARREN	RANGE	MOIST	ASPEN	SF/A	SF 70	SF 90	WATER	BADDATA
1 BARREN	41	63.4	26	1	0	0	1	13	0	0	0
2 RANGE	1047	64.2	95	672	250	14	13	3	0	0	0
3 MOIST	21	85.7	0	2	18	1	0	0	0	0	0
4 ASPEN	127	47.2	1	51	8	60	6	0	0	0	1
5 SF/A	104	70.2	0	1	0	2	73	28	0	0	0
6 SF 70	476	88.9	0	1	0	0	33	423	18	0	1
7 SF 90	1317	77.4	0	1	0	0	9	282	1019	6	0
8 WATER	234	100.0	0	0	0	0	0	0	0	234	0
TOTAL	3367		122	729	276	77	135	749	1037	240	2

OVERALL PERFORMANCE(2525/ 3367) = 75.0
AVERAGE PERFORMANCE BY CLASS(597.0/ 8) = 74.6

(B) ECHO Classifier

GROUP	NO OF SAMPS	PCT. CORCT	NUMBER OF SAMPLES CLASSIFIED INTO BARREN	RANGE	MOIST	ASPEN	SF/A	SF 70	SF 90	WATER	BADDATA
1 BARREN	41	68.3	28	0	0	0	0	13	0	0	0
2 RANGE	1047	59.2	122	620	284	9	12	0	0	0	0
3 MOIST	21	90.5	0	1	19	1	0	0	0	0	0
4 ASPEN	127	44.9	2	47	11	57	9	0	0	0	1
5 SF/A	104	73.1	0	3	1	3	76	21	0	0	0
6 SF 70	476	91.4	1	0	0	0	20	435	19	0	1
7 SF 90	1317	75.2	0	0	0	0	3	318	990	6	0
8 WATER	234	100.0	0	0	0	0	0	0	0	234	0
TOTAL	3367		153	671	315	70	120	787	1009	240	2

OVERALL PERFORMANCE(2459/ 3367) = 73.0
AVERAGE PERFORMANCE BY CLASS(602.5/ 8) = 75.3

Table C-11 Cont.

(C) Minimum Distance Classifier

GROUP	NO OF SAMPS	PCT. CORCT	NUMBER OF SAMPLES CLASSIFIED INTO BARREN	RANGE	MOIST	ASPEN	SF/A	SF 70	SF 90	WATER	BADDATA
1 BARREN	41	58.5	24	2	0	0	1	11	3	0	0
2 RANGE	1047	68.6	70	718	220	20	15	4	0	0	0
3 MOIST	21	85.7	0	1	18	2	0	0	0	0	0
4 ASPEN	127	48.0	0	47	10	61	7	1	0	0	1
5 SF/A	104	73.1	0	1	0	2	76	24	1	0	0
6 SF 70	476	86.6	0	1	0	1	44	412	18	0	0
7 SF 90	1317	76.8	0	1	0	0	11	275	1012	18	0
8 WATER	234	100.0	0	0	0	0	0	0	0	234	0
TOTAL	3367		94	771	248	86	154	727	1034	252	1

OVERALL PERFORMANCE(2555/ 3367) = 75.9
AVERAGE PERFORMANCE BY CLASS(597.3/ 8) = 74.7

(D) Layered Classifier

GROUP	NO OF SAMPS	PCT. CORCT	NUMBER OF SAMPLES CLASSIFIED INTO BARREN	RANGE	MOIST	ASPEN	SF/A	SF 70	SF 90	WATER	BADDATA
1 BARREN	41	56.1	23	1	0	0	1	16	0	0	0
2 RANGE	1047	64.2	100	672	245	14	13	3	0	0	0
3 MOIST	21	85.7	0	2	18	1	0	0	0	0	0
4 ASPEN	127	47.2	0	53	6	60	6	0	0	1	1
5 SF/A	104	70.2	0	1	0	2	73	28	0	0	0
6 SF 70	476	88.9	0	1	0	1	33	423	18	0	0
7 SF 90	1317	76.3	0	1	0	0	9	282	1005	20	0
8 WATER	234	100.0	0	0	0	0	0	0	0	234	0
TOTAL	3367		123	731	269	78	135	752	1023	255	1

OVERALL PERFORMANCE(2508/ 3367) = 74.5
AVERAGE PERFORMANCE BY CLASS(588.6/ 8) = 73.6

Table C-11 Cont.

(E) Levels Classifier

GROUP	NO OF SAMPS	PCT. CORCT	NUMBER OF SAMPLES CLASSIFIED INTO BARREN	RANGE	MOIST	ASPEN	SF/A	SF 70	SF 90	WATER	BADDATA	OTHER
1 BARREN	41	51.2	21	1	0	0	1	8	0	0	0	10
2 RANGE	1047	39.4	53	412	182	5	13	2	0	0	0	380
3 MOIST	21	47.6	0	1	10	0	0	0	0	0	0	10
4 ASPEN	127	27.6	0	32	8	35	5	0	0	0	0	47
5 SF/A	104	62.5	0	1	0	1	65	9	0	0	0	28
6 SF 70	476	71.6	1	1	0	0	32	341	13	0	0	88
7 SF 90	1317	63.2	0	0	0	0	9	237	832	11	0	228
8 WATER	234	38.0	0	0	0	0	0	0	0	89	0	145
TOTAL	3367		75	448	200	41	125	597	845	100	0	936

OVERALL PERFORMANCE(1805/ 3367) = 53.6

AVERAGE PERFORMANCE BY CLASS(401.1/ 8) = 50.1

(b) ECHO Classifier. This classification algorithm allows more than one data point to be at a time, thereby reducing the number of individual decisions that must be made and thus reducing the amount of CPU time required. The involvement by the analyst is relatively minimal and simply involves selecting two classification parameters. However, it must be pointed out that the optimum values for these two classification parameters are not thoroughly understood. For this reason, in this test the classification parameters that were utilized were those recommended for use with the ECHO classifier by Swain (LARS Information Note 111276). These recommended parameters were cell size = 2 x 2; cell splitting 60; annexiation - 1.0.

Classification of groups of pixels rather than individual pixels has a definite effect on the characteristics of the classification results, producing an output map that has much less of the "salt & pepper" effect. Since the algorithm uses spatial as well as spectral information, it would appear that the identification of each cell defined by the algorithm should be more accurate. However because some informational classes are often in small isolated units (e.g. The barren category) the effect of the ECHO classifier in smoothing the data or removing the salt & pepper effect can cause a slight decrease in the classification accuracy of certain informational classes. In this study, the overall classification accuracy was not significantly influenced by the use of the ECHO classifier as compared to the maximum likelihood perpoint classifier. The overall cost per hectare, however, was reduced to 0.14 cents per hectare.

(c) Minimum Distance to the Means Classifier. Through use of the Euclidean distance measure, the CPU time required by the minimum distance classifier was much less than for any of the maximum likelihood classifiers. Analyst involvement was also minimal, since only a single job deck defining the area to be classified and wavelength bands to be used was required. The overall cost was therefore the lowest of all classifications algorithms compared, being only 0.089 cents per hectare. Surprisingly, the classification accuracy with the minimum distance classifier was relatively high, equalling the classification accuracy of the maximum likelihood perpoint classifier at Level 2 and being slightly higher than the maximum likelihood perpoint classifier for the Level 3 classification results. This relatively high accuracy of the minimum distance classifier may have been due in part to the relatively heavy reliance on the clustering algorithms (which is a minimum distance algorithm) in developing the spectral/informational classes defined by the training statistics deck. Therefore, the clustering algorithm developed classes which were characterized by the minimum distance assumption. However, as the previous study indicated, use of the clustering algorithm to develop the training statistics provides an optimal method for developing training statistics in such a spectrally complex area. These results may therefore indicate a need for further study of the inter-relationship between developing training statistics using the clustering algorithm to aid the analyst and in the use of a minimum distance algorithm for the actual classification of the data set.

(d) The mulitple nodes and tree structure of the layered classifier allowed various subsets of channels to be used in the classification. This resulted in a considerable reduction in CPU time required as compared to the single-layer approach involved in the maximum likelihood classifier. However, the CPU time for the layered classifier was still much higher than that involved in the minimal distance classification. Although obtaining the truly "optimum" tree structure is difficult, a procedure has been developed to "automatically" generate the tree, thereby reducing the analyst involvement to that of selecting two parameters and preparing three job decks (which could be reduced to only one job deck with with further programming). The accuracy of the resulting classification was slightly less than the maximum likelihood classifier or the minimum distance algorithm, but not by an amount that is considered significant. The overall classification cost was 0.184 cents per hectare, which is significantly lower than that involved in the maximum likelihood perpoint classifier but higher than either the ECHO or the minimum distance classifiers.

(e) Levels Classifier. In the use of this classifier, the ranges of the values used to define each class are difficult to obtain accurately. These values are essentially the boundaries between the classes, and must be defined by the analyst for all channels and for each spectral class. In this study, we found that defining the ranges required considerable analyst time to simply convert the means and standard deviations to the ranges to be used. In this study, each range was defined as $\pm$ 2½ standard deviations of the mean. Therefore, further programming could certainly reduce the amount of analyst time involved in this step of the analysis. The biggest limitation in the use of the levels classifier involves the fact that the analyst does not know the optimal set of ranges for each spectral class and therefore a large number of pixels are often classified as "other", thereby reducing the classification accuracy by a significant amount. In addition, high numbers of data points which are unclassified for all practical purposes makes usage of acreages from this data very difficult. The classification accuracy probably could have been increased by further adjustments of the levels, but this would have increased both man-hours and CPU time considerably. In this study, only a single attempt was made to classify the data using each algorithm. Therefore, the CPU time for the levels classifier was the lowest of any of the algorithms used but the analyst involvement increased the total cost considerably. The final result showed a 0.2 cents per hectare total cost, which is second only to that of the maximum likelihood perpoint classifier, but the classification accuracy at both Level 2 and Level 3 were significantly less than any of the other four classification algorithms used.

To compare the characteristics of the output products obtained, two approaches were used. First, Varian printer maps were obtained for Platoro quadrangle using each of the five classification algorithms tested. These were compared and showed that the maximum likelihood perpoint classification map (shown in Figure 8) was very similar in appearance to the minimum distance to the means and the layered classification results. The ECHO classifier produced an output map that was "blockier" in appearance (see Figure 9), and without the "salt and pepper" effect that the other algorithms produced because they classified each pixel independently. The levels classifier had many data points that were classified as "other", thereby producing a classification map that was not very statisfactory because of the large amount of data that was not included in any of the informational classes that had been defined (see Figure 10).

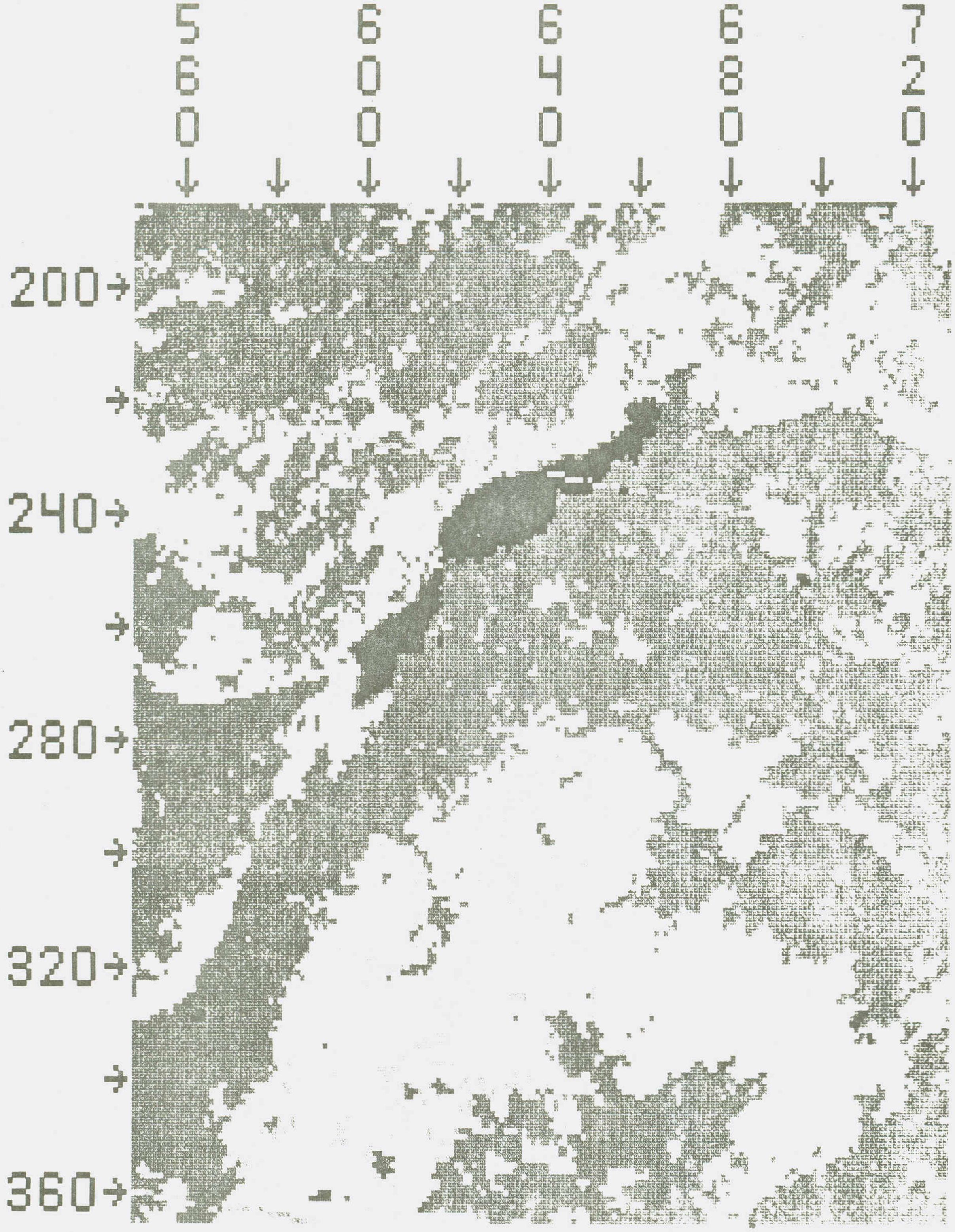


Figure C-8. Varian Display of Maximum Likelihood Perpoint Classification.

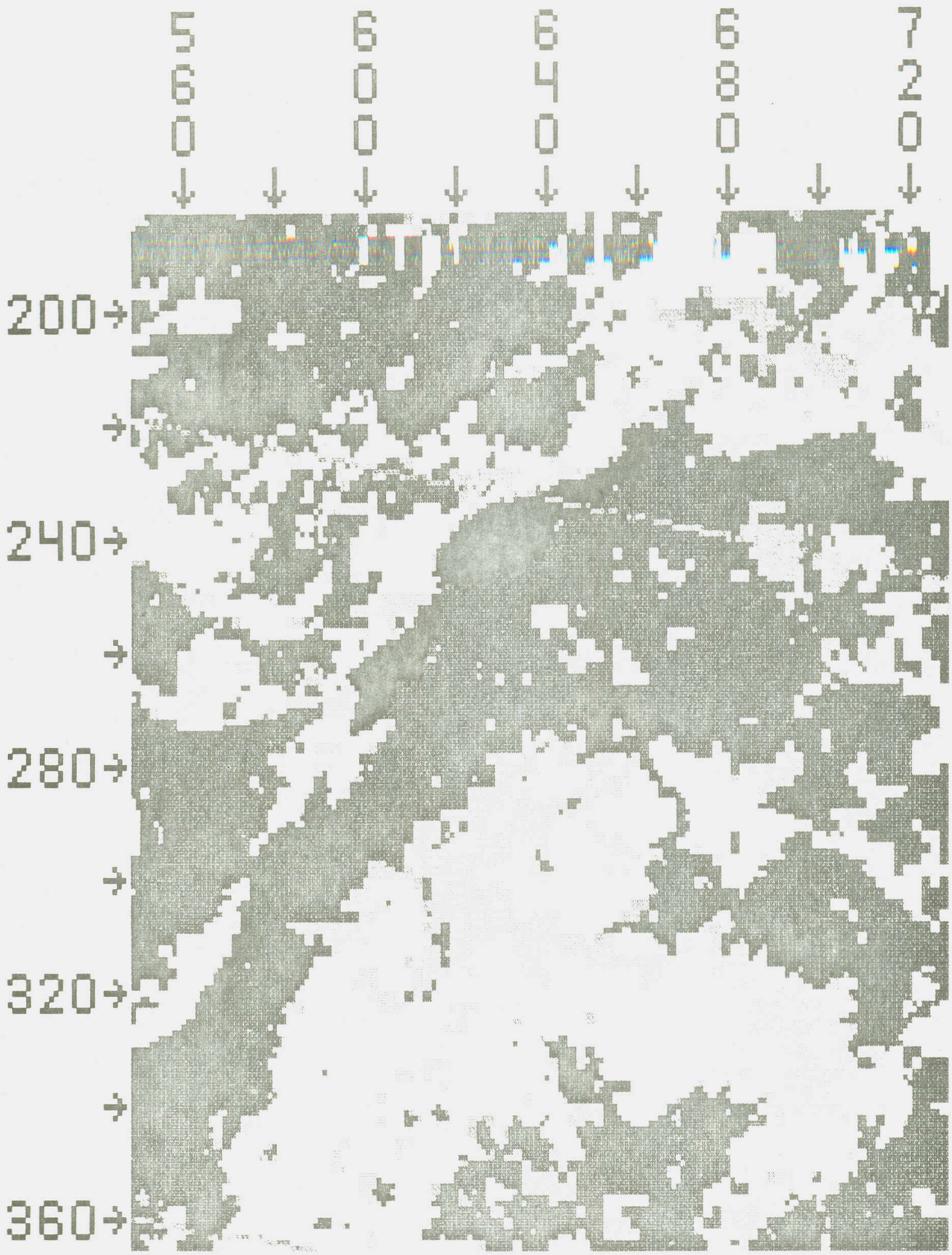


Figure C-9. Varian Display of ECHO (Extraction and Classification of Homogeneous Objects) Classification.

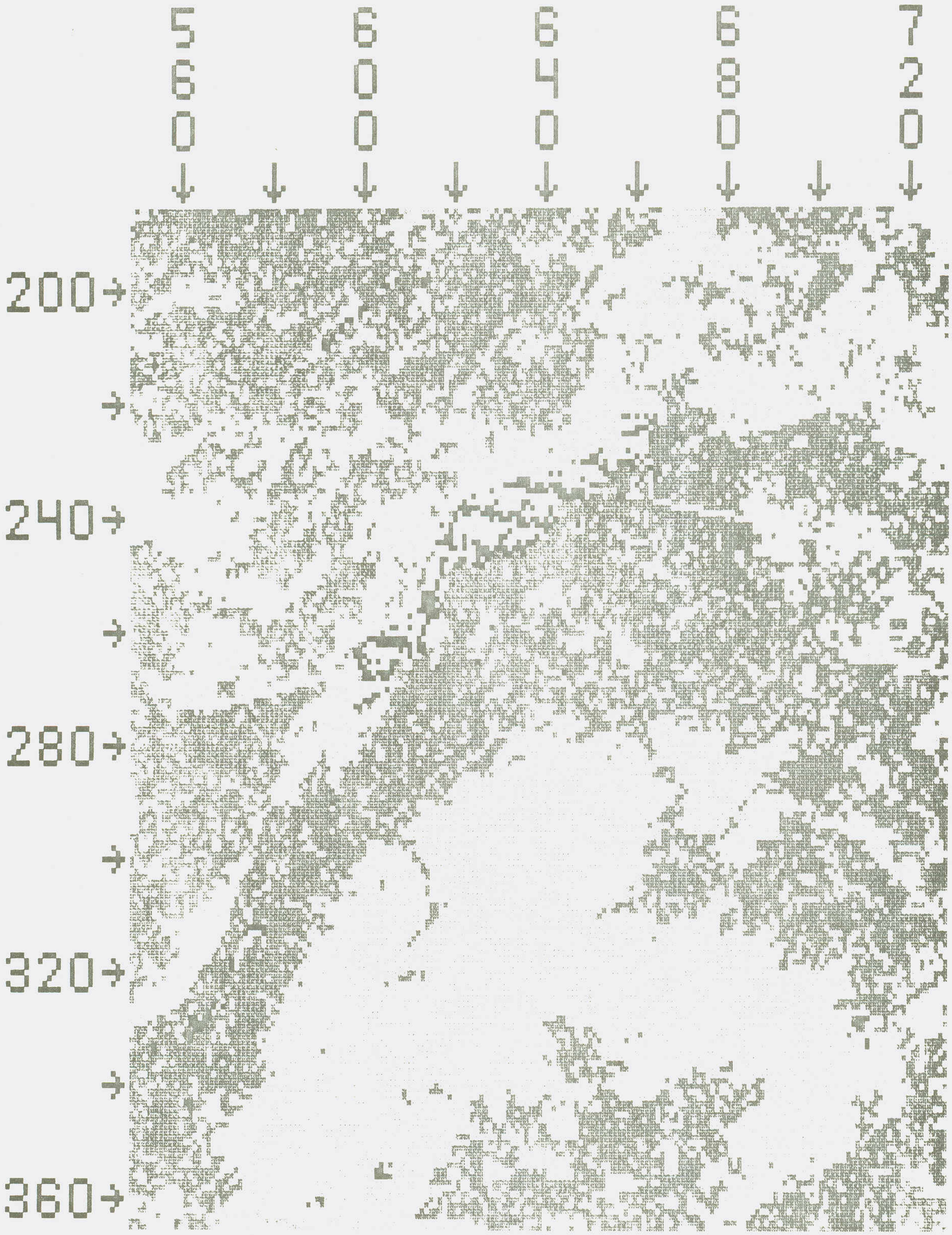


Figure C-10. Varian Display of Levels Classification.

The second evaluation of the classification maps involved producing a map product in each of several different output devices each having a different format, including a line printer (Figure 11), Calcomp plotter (Figure 12), Varian (Figure 9), and digital display (Figure 13) output devices. These figures were each obtained using the ECHO classification results, and indicate the wide variety of map output formats that can be obtained from the same classification result.

Evaluation of the different formats of map products indicated that the levels classification result was not accurate enough to produce a satisfactory map product. The other algorithms produced very similar classification performances: Map outputs having similar appearances (except for the ECHO classifier) could be obtained with any of the output devices. The particular output devices to be used, and the choice of a perpoint or an area classifier (i.e. ECHO) is a function of the user needs and should be specified by the user. The line printer is usually a 1:24,000 scale, and changes in this scale can best be obtained photographically when reproducing the original 1:24,000 scale printout. The Calcomp printer does not work very well when many adjacent pixels are classified differently, so is most appropriate for use with the ECHO classifier. The Varian output is inexpensive to obtain and produces a map of a large geographic area in a compact format, with line and cloumn numbers designated, and at various scales. The digital display produces high quality photographic outputs at various scales and can obtain color-coded outputs, but is more time consuming and expensive to use. In summary, it appears that the scale, format, and characteristics of the map output are quite flexible, and it is the user's choice as to which algorithm and output device will produce the output map that best meets his information needs.

Summary and Conclusions. Five different classification algorithms were utilized in this study. Several of the classification algorithms utilized as much of the statistical description of the spectral classes as possible to increase the accuracy of the classifications. The maximum likelihood classifiers (i.e. maximum likelihood perpoint, ECHO, and the layered classifier) all utilize the mean, variances, and covariance matrices with the ECHO classifier utilizing additional information in terms of the spatial variation. Some algorithms attempt to reduce the CPU time by reducing the number of decisions required (as with the ECHO classifier), or by maximizing the use of channel combinations by using more numerous but simplier decisions involving fewer channels (as with the layered classifier), or by reducing the amount of information utilized in the decision process (as with the minimum distance classifier). Several of these algorithms have minimized the analyst involvement in the classification step of the analysis to essentially zero. This is particularly true of the maximum likelihood perpoint and the minimum distance classifier. The layered classifier and ECHO classifier require analyst involvement only to select several parameters. However, it should be pointed out that some of the parameters which must be defined by the analyst are not well understood at present, thereby decreasing the current value of these algorithms.

The optimal classification algorithm is one which is designed to reduce the analyst time required and CPU time to a minimum, while at the same time maximizing the classification accuracy. Each of the algorithms tested have advantages and disadvantages. Various types of output map products can be obtained using any of the classification algorithms tested. A key point that must be made is that the development of an effective set of training statistics

NUMBER OF POINTS DISPLAYED IS 19980

NUMBER OF POINTS DISPLAYED IS 12420

Figure C-11. Line-Printer Output Showing ECHO Classification Results.

Figure C-12. Calcomp Output Showing ECHO Classification Results.

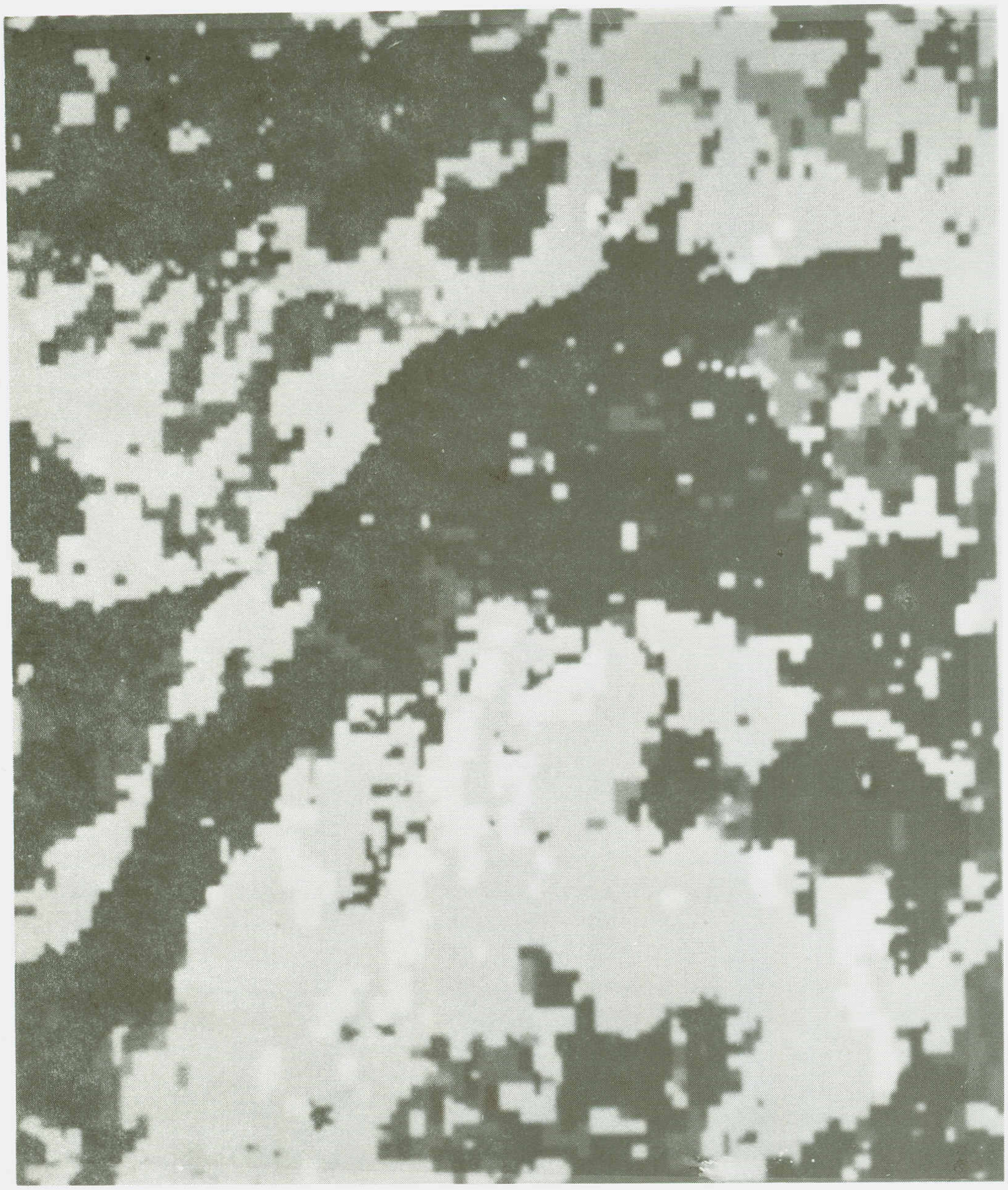


Figure C-13. Digital Display Output Showing ECHO Classification Results.

is a critical part in obtaining a high level of accuracy in the final classification results. Results of this study indicate that use of different classification algorithms can cause a significant difference in the cost of classifying the data without adversely effecting the classification accuracy. Overall, the results obtained in this study indicated that the minimum distance to the mean classification algorithm was the most efficient since it minimized the analyst time and computer time while the accuracy was as high or higher than was obtained using any other classification algorithms.

Comparison of Six Approaches for Developing Training Statistics

Introduction. Over the past decade, tremendous progress has been made in the development of computer-aided analysis techniques (CAAT) involving the implementation of pattern recognition theory on computer systems to analyze remotely sensed data obtained in the form of multispectral scanner data and digitized aerial photography. The two basic approaches by which the computer is taught to recognize and identify cover types are the "supervised", involving a training sample approach, and the "non-supervised", a clustering approach, both of which have been used by many researchers with considerable success. Much of the previous work involving the application of computer-aided analysis techniques in a variety of disciplines has shown the capabilities and limitations of these two approaches.

Preliminary attempts to map forest cover in areas of complex vegetation types and rugged terrain using both the supervised and non-supervised techniques indicated several difficulties with each approach in relating the cover type categories present on the ground to the spectral classes present in the data. For example, in the mountainous terrain of the San Juan Mountains of southwestern Colorado, selection of a training data set was extremely difficult due to spectral differences caused by variations in slope and aspect, as well as to the many spectral differences in the cover types themselves. It was therefore essential that a more effective procedure be defined to accurately map forest and other cover types when utilizing the computer-aided analysis techniques and Landsat MSS data obtained over spectrally complex areas.

If computer-aided analysis techniques are to be used operationally as a reliable source of information for a land manager, the technique must meet three basic requirements. First, the process must become more effective by optimizing the analyst/data interactions, thus reducing human bias and increasing the reliability of results (i.e., making the process more of a science than an art). Secondly, the analysis sequence must be defined as a series of identifiable and repeatable steps which are not data or analyst dependent. Thirdly, the entire process must be more efficient, i.e., faster, while reducing both computer cost and analyst time. Although there may be no single or "best" way to perform the analysis for all applications and situations, a better understanding of the capabilities and limitations offered by computer-aided analysis techniques and Landsat MSS data must be developed and documented.

Objective. The major objective of this study was to define an effective and efficient computer-aided analysis technique that can be utilized to map forest lands in areas of rugged topography using digital multispectral scanner data collected from satellite altitudes.

In order to accomplish this objective, a series of experiments were set up to develop and evaluate the basic approaches to computer-aided analysis techniques. The following five steps were pursued:

1) Define several alternate analysis procedures for training a pattern recognition classifier.

2) Develop and illustrate the steps involved in each analysis approach to maximize the classification accuracy and optimize the interactive and computational efficiency of each.

3) Evaluate these analysis techniques to better understand the capabilities and limitations of each.

4) Compare the amount of support data and human involvement required, as well as the computer cost and accuracy of the various analysis techniques.

5) Describe and illustrate, in detail, the recommended computer-aided analysis technique for analyzing Landsat MSS data in an operational system.

Test Sites. The primary study areas involved in this investigation were the Platoro quadrangle (15,303 hectares or 37,812 acres) and the Southern San Juan Mountain Planning Unit (S.S.J.M.P.U.) (540,580 hectares or 1,335,755 acres). The relatively small Platoro quadrangle test site was utilized as an intensive study site for development and preliminary testing of the various analysis techniques. The S.S.J.M.P.U. study area was specified by personnel of the U. S. Forest Service, and includes parts of three National Forests -- the San Juan and the Rio Grande in Colorado, and the Carson in New Mexico. This larger test site was used for a more complete testing and evaluation of the effectiveness and efficiency of the different techniques. Both the Platoro quadrangle and S.S.J.M.P.U. were classified at a Level 2 and at a Level 3 degree of detail. Table 1 gives the cover types or informational classes included in both the Level 2 and Level 3 classifications. In addition, six other test sites in the Colorado Rocky Mountains, ranging in size from 14,600 to 1,012,000 hectares (36,000 to 2,500,000 acres), were also used to evaluate the recommended technique over a variety of cover types and data sets.

Analysis Procedures. The initial phases of the study involved defining several alternate analysis procedures for developing a set of training statistics. Evaluation of the two basic approaches (supervised and non-supervised or clustering) indicates that these two approaches are in opposition, both in terms of the method of defining the training sample (data points) and the method of grouping the training sample into unimodal training classes (training statistics). In the supervised approach, a scattered sample of small fields are carefully selected by the analyst, who then groups the fields into unimodal informational training classes. The non-supervised approach is just the opposite in that the training sample is selected and grouped into spectral training classes entirely by the clustering algorithm. Due to this contrast, the possibility exists for intermediate steps between the two extremes in both parts of the training procedure (i.e., selecting the training sample and grouping the training data into spectral training classes).

Table C-12. Informational classes utilized in the computer-aided analysis of both the Platoro quadrangle and the S.S.J.M.P.U. for Level 2 and Level 3.

Level 2	Level 3
Barren	Barren (Exposed rock and soil)
Water	Water
Grassland	Rangeland Agricultural
Deciduous	Aspen Oak
Coniferous	Ponderosa Pine/Oak Spruce-fir/Aspen Spruce-fir, 60-80% crown closure Spruce-fir, 80-100% crown closure

Six different approaches for developing training statistics were therefore defined, based upon three methods for selecting the training sample and three methods for grouping the training sample into training classes. The three methods for selecting the training sample represent three levels of supervision by the analyst, whereas the three methods for grouping the training sample represent three levels of supervision using the clustering algorithm. The three methods of selecting the training sample are:

1) HOMOGENEOUS FIELDS - This requires a high degree of involvement by the analyst to select informationally pure fields which represent the spectral variation for each cover type of interest. The relatively small areas of homogeneous cover type in the Landsat data makes it necessary for the analyst to select numerous small fields, usually less than 10 lines by 10 columns of data in size.

2) HETEROGENEOUS BLOCKS - This requires the analyst to select a series of areas that each contain a mixture of several cover types. The variation in spectral response for each cover type of interest is represented in at least one of the blocks. The size of the blocks varies from less than 30 lines by 30 columns to over 100 lines by 100 columns. Since the blocks are not informationally pure, the spectral classes must be obtained using the clustering algorithm.

3) STATISTICAL SAMPLE - This requires very little analyst supervision. The entire area of interest is systematically sampled to represent the variation in spectral response for all cover types in the area. The training data points are 'automatically' selected by the clustering algorithm.

The three methods for converting the training sample data into training classes are:

1) NO CLUSTERING - This requires a high degree of involvement by the analyst to group the training fields into unimodal spectral classes without the aid of the clustering algorithm.

2) MULTIPLE CLUSTERING - This requires the analyst to divide the training sample into groups which are clustered separately. The numerous cluster classes must then be combined into meaningful training classes.

3) SINGLE CLUSTERING - This requires very little analyst supervision. The entire training sample is grouped into spectral classes by the clustering algorithm, essentially unsupervised.

As the above comments indicate, the different levels of supervision required represent a balance between involvement by the analyst and dependence on the computer system, primarily the clustering algorithm. Figure 14 graphically shows the differences between these six approaches, and indicates that the different approaches defined represent a combination of the method used to select the training sample and the method used to group the training sample into training classes. Figure 15 shows the names that were given to the various analysis techniques. The major aspects of each of the different techniques are discussed in the following paragraphs. (A more detailed description is given in the full report of this study, LARS Information Note 112177.)

The supervised approach for developing the training statistics includes: selection and delineation of supervised training fields by the analyst, grouping the fields into unimodal training classes, and calculation of statistical parameters for each training class. The training statistics are then evaluated to determine if they are acceptable. In this approach, the clustering algorithm is not used to group the training sample into unimodal classes.

The multi-cluster fields approach for developing training statistics involves: selection and delineation of supervised training fields by the analyst, individual clustering of each informational class (i.e., a series of clusterings, each clustering involving all training fields of one cover type or information class), and combining the spectral-informational classes into a single set of training classes. The training statistics are then evaluated to determine their acceptability. The key difference between the multi-cluster fields analysis procedure and the supervised approach is the use of the clustering algorithm to divide each informational class into a number of spectrally distinct classes. The clustering algorithm provides a mechanism for dividing each informational class into approximately Gaussian spectral classes in a faster, non-supervised manner.

The mono-cluster fields approach is the least supervised technique for separating supervised training fields into spectral classes, since all training fields are clustered at one time. The major steps in this approach are: selection and delineation of supervised training fields, a single clustering of all fields, and identification of the cluster classes. The training statistics are then evaluated to determine if they are acceptable. As compared to the multi-cluster fields approach, it was anticipated that the mono-cluster fields approach would allow the analyst's time to be considerably reduced but might cause a considerable increase in computer time.

The non-supervised approach involves clustering a statistical sample of data points from the entire site. It therefore differs from the mono-cluster fields approach primarily in the method used to select the training sample. The mono-cluster fields approach relied on supervised training fields to represent each information class, while the non-supervised approach uses a statistical sample of data points from the active study area to represent the spectral characteristics of the area. The non-supervised approach for developing the training statistics therefore involves the following major steps: clustering a statistical sample of points from the entire site into a specified number of spectral classes; and identification of the cluster classes, which requires a classification map. Note that the training statistics can not be identified before a classification is performed. Also note that the clustering of data from the entire site can involve every data point in the study site, but this requires so much computer

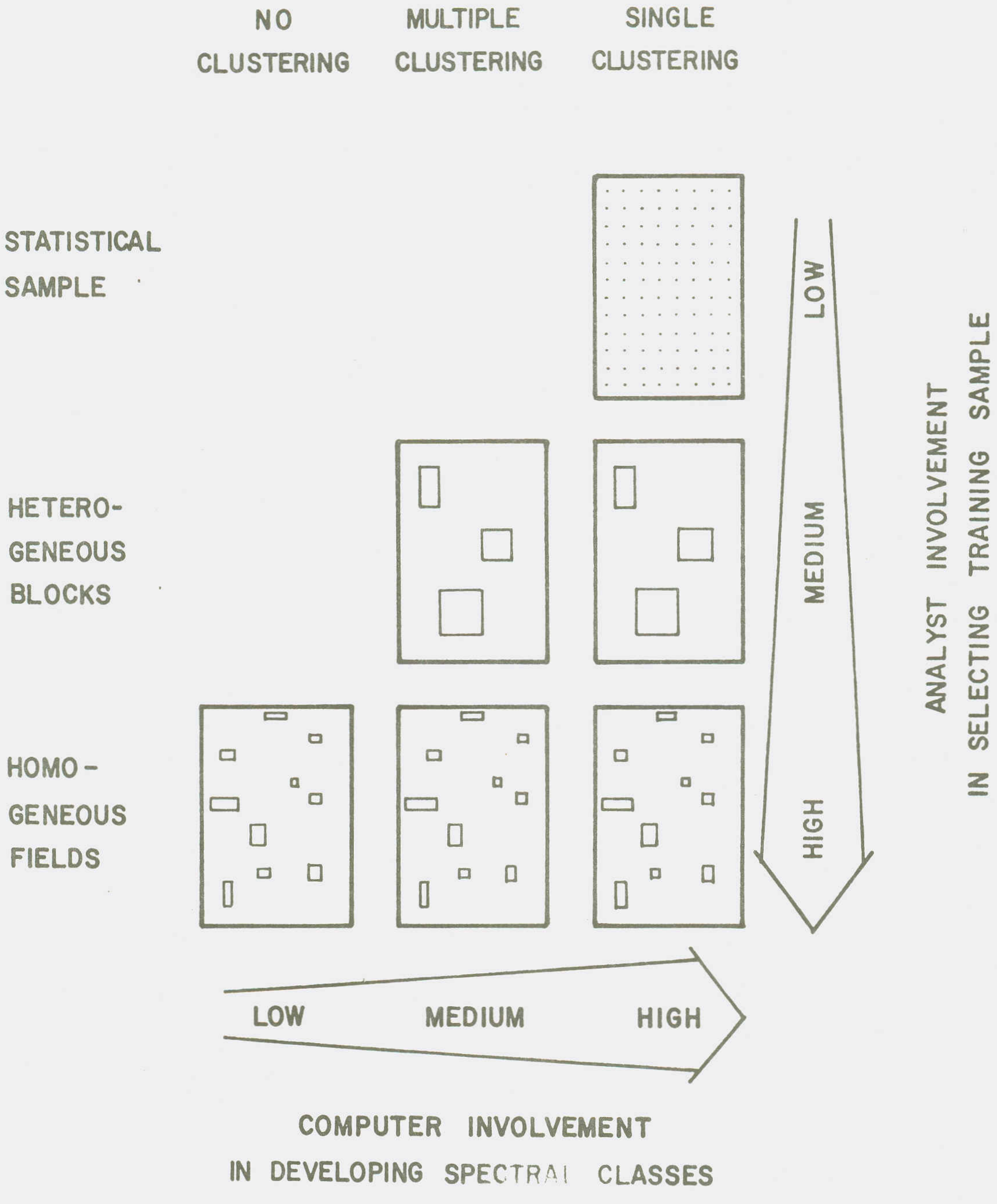


Figure C-14. The six techniques defined for developing training statistics, emphasizing the different levels of analyst and computer involvement.

		UN- SUPER- VISED
	MULTI- CLUSTER BLOCKS	MONO- CLUSTER BLOCKS
SUPER- VISED	MULTI- CLUSTER FIELDS	MONO- CLUSTER FIELDS

Figure C-15. The names given to each of the six techniques defined for developing training statistics.

storage if the study site is very large that a sample of every m^{th} line and n^{th} column must be utilized. In areas where some of the features of interest are relatively small in size and sparce in occurrance, a sampling procedure results in an insufficient number of data points to adequately define the training blocks statistics for these cover types.

In the mono-cluster blocks approach informationally heterogeneous blocks of data scattered throughout the site are selected by the analyst. All blocks are then clustered as a single group to obtain the training statistics. In the mono-clustering blocks approach, the major steps are: selection and delineation of heterogeneous blocks of data, clustering of all blocks as a single group, and indentification of spectral classes. The training statistics are then evaluated to determine if they are acceptable.

In the multi-cluster blocks approach (also refered to as the modified cluster technique (Fleming, et al, 1975) heterogeneous blocks of data are selected by the analyst, but are then clustered individually (rather than together, as it done in the mono-cluster blocks approach), as compared to the mono-cluster blocks approach ed the individual clustering of blocks increases the number of cluster classes and the sample size considerably. It was anticipated that the computer time required would be much less for the multi-cluster blocks approach as compared to the mono-cluster blocks approach, but that the analyst time would be increased. This approach for developing training statistics includes the following major steps: selection and delineation of informationally heterogeneous blocks, individual clustering of each block, identification of the spectral classes for each block, and pooling of similar classes to obtain spectrally separable classes. The training statistics are then evaluated to determine if they are acceptable.

Evaluation Procedures. Once an adequate training set has been defined, it is not difficult to classify a large geographic area using computer analysis techniques. However, unless one can verify the accuracy of such computer classification results, little has been accomplished by simply classifying data over various areas of interest. The real question is whether or not the resultant classification maps and tables are reasonably accurate and have a reasonable level of reliability. In this study, a combination of two techniques proved most satisfactory to achieve the best possible indication of the classification performace. The two techniques utilized to evaluate the classification results included:

1) Qualitative comparison of the computer classification maps and the maps obtained from interpretation of aerial photos, as well as direct comparison to the aerial photos.

2) Quantitative evaluation of classification performance, utilizing a system of statistically defined test areas.

In the qualitative evaluation, the maps resulting from interpretation of aerial photos obtained by NASA were utilized as reference data. The computer classification map and the photo interpretation map were compared, and a general assessment of the level of agreement between the computer classification map and the photo interpretation map was made. This approach provides a general indication of the computer classification performance, but it is difficult to make in-depth or detailed comparisons of classification results.

A quantitative evaluation technique utilizing test areas provides a more effective method for evaluating computer classification results. Two different approaches to quantitative evaluation were attempted. The first approach estimated the extent to which the informational classes match the spectral classes, using as series of individually selected test fields which represent the information classes. This enabled "standard" classification results tables to be obtained for the classifications. The second approach to a quantitative evaluation involved thresholding the classification results to determine the number of points that did not "fit" (low probability) any of the training spectral classes. Thresholding estimates the extent to which the spectral classes used for training match those actually present in the data.

The costs -- both for personnel and computer -- involved in developing the training statistics were monitored. These data were used in evaluating the efficiency and effectiveness of the various approaches to defining a recommended training procedure.

Results and Discussion. After defining the six alternative approaches for training a pattern recognition classifier, the steps in each analysis approach were developed to maximize their classification accuracy while optimizing their interactive and computational efficiency. The different techniques were then tested on both the relatively small Platoro quadrangle (15,303 hectares) and the large S.S.J.M.P.U. area (540,580 hectares) to determine the limitations, capabilities and requirements of each. These results were compared to determine the "optimum" approach in terms of CPU time, man-hours, and support data required, as well as accuracy achieved.

We attempted to maximize the validity of the study by reducing the sources of variation not directly associated with the analysis techniques. Throughout the entire study, several factors were held as constant as possible. The most important ones were: the number of training classes, the test fields, the classification algorithm, the convergence value of the clustering algorithm, and the training areas used for each method of selecting the training sample. An effort was made to develop 16 and 25 training classes for each analysis approach on the Platoro quadrangle and S.S.J.M.P.U. areas, respectively.

The classification accuracy was evaluated using two sets of test fields defined by personnel from the Institute of Artic and Alpine Research (INSTAAR), University of Colorado independent of the computer-aided analysis. In the Platoro quadrangle, a set of "true" test fields were defined by a photointerpreter and checked and rechecked several times, whereas the S.S.J.M.P.U., a "systematic" set of test fields were statistically defined over six quadrangle areas and not checked. (This difference in the method used to define the test fields and the amount of checking have caused some of the difference in accuracy levels that were found in the Platoro quadrangle results as compared to the S.S.J.M.P.U.)

The "systematic" fields in the S.S.J.M.P.U. were classified using a maximum likelihood perfield algorithm. The "true" set of fields were classified by a maximum likelihood perpoint algorithm to evaluate all analyses of the Platoro quadrangle. A convergence value of 99 percent was used for each application of the clustering algorithm. The same areas were used for each method of selecting the training sample. For example, the same set of supervised training fields were utilized for the three analysis techniques that required them. By minimizing these influences, the differences between the analysis techniques were due to the approach rather than bias by the analyst or extraneous factors. The results for the Platoro quadrangle are shown in Figure 16 and the results for the S.S.J.M.P.U. are shown in Figure 17.

In evaluating these results, it must be pointed out that the cost in terms of CPU time and man-hours should not be interpreted as absolute, and would not accurately represent the expected cost in an operational situation. The amount of CPU time varies considerably depending upon the computer hardware on which the algorithms are implemented and the efficiency of the programs themselves. The number of man-hours also varies, depending upon the analyst's knowledge and experience with the MSS data, the analysis techniques, and the complexity of the area. However, in this test, all the analysis techniques were evaluated on the S.S.J.M.P.U. and Platoro quadrangle by one analyst using the LARS computer system. Therefore, the relative quantities can be used to accurately compare the various approaches to developing training statistics for computer-aided analysis of Landsat MSS data.

An evaluation of the results in Figures 3 and 4 indicates that the multi-cluster techniques generally were more accurate than the mono-cluster techniques. The main influence was the increase in sample size and number of candidate spectral classes. It was also found that when the analyst supervision in selecting the training sample was reduced (from training fields to training blocks to systematic sample), a better statistical sample of the spectral variation in the area increased the classification accuracy. The difference between the techniques having the lowest and highest accuracy was considerable -- 8.8 percent for the Platoro quadrangle and 14.1 percent for the S.S.J.M.P.U. This indicates that the type of analysis approach does have a definite impact on the classification accuracy. The difference in range of accuracy for the two sizes of areas was probably due to the increased spectral variation in both the number of cover types and within each cover type.

Comparison of the six approaches indicated that the supervised approach was the least accurate method of developing training statistics, and the multi-cluster blocks approach was the most accurate, for both areas considered. The S.S.J.M.P.U. analysis by the multi-cluster blocks approach was estimated to be 78.8 percent correct, according to the test field results. However, a detailed comparison between the classification results and aerial photography indicated that the classification was closer to 85 or 90 percent correct. This again indicates the difficulty in estimating the accuracy of computer-aided analysis techniques, due to the lack of information with which to compare results.

KEY		
analysis approach		UNSUPERVISED
man-hours		6 hours
CPU time		17.0 min.
accuracy		90.0 %

	MULTI-CLUSTER BLOCKS	MONO-CLUSTER BLOCKS
	5 hours	4 hours
	6.0 min.	6.1 min.
	90.1 %	89.0 %

SUPERVISED	MULTI-CLUSTER FIELDS	MONO-CLUSTER FIELDS
13 hours	11 hours	9.5 hours
12.8 min.	3.2 min.	3.9 min.
83.3 %	87.1 %	87.4 %

Figure C-16. Results from the Platoro quadrangle areas(15,303 hectares) of the evaluation of the six techniques for developing training statistics.

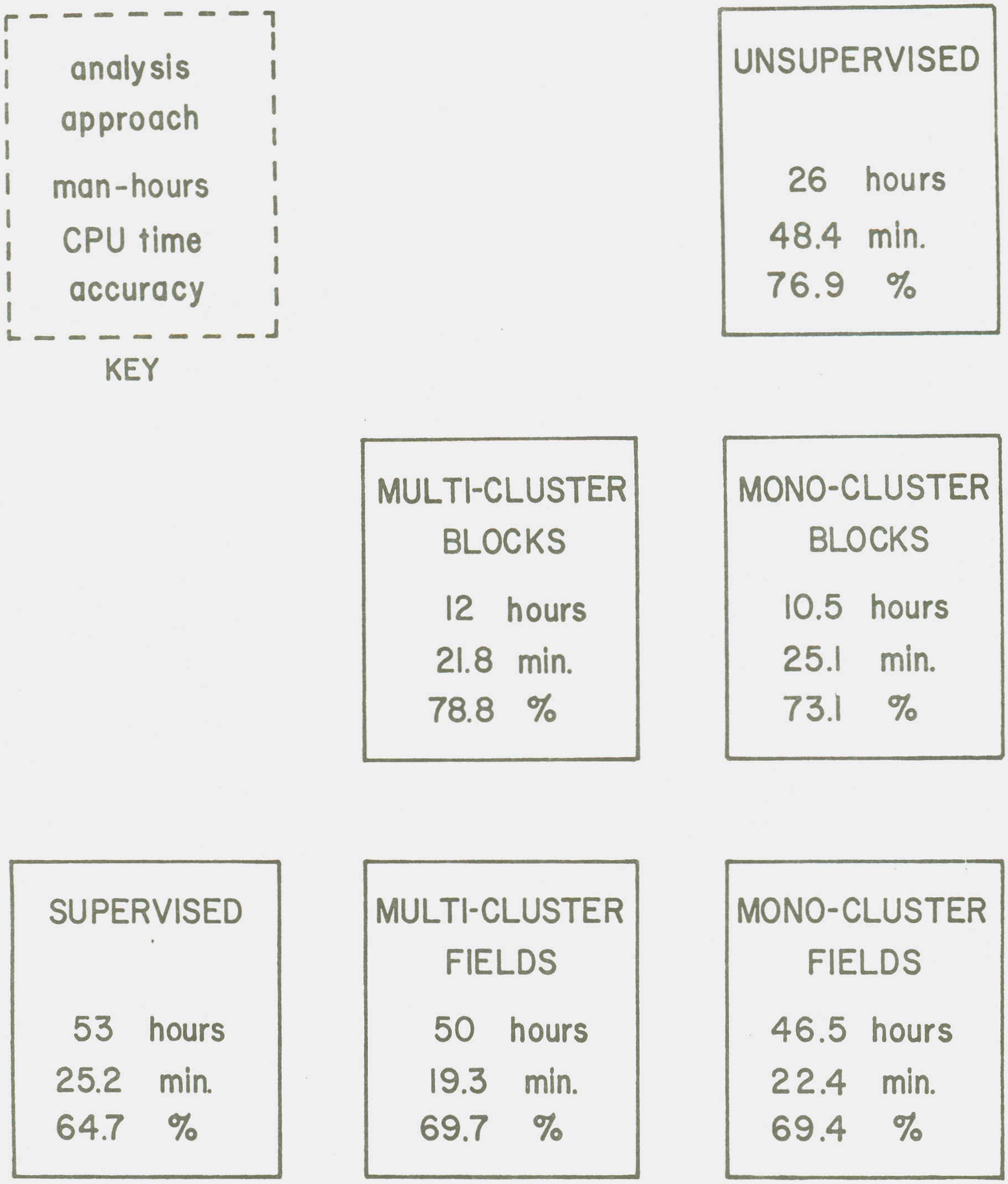


Figure C-17. Results from the Southern San Juan Mountain Planning Unit (540,580 hectares) of the evaluation of the six techniques for developing training statistics.

Most of the computer time used was for defining the spectral classes, either using the *STATISTICS or *CLUSTER processor. Different methods of calculating the statistical parameters caused the *STATISTICS processor to be less efficient in terms of CPU time, but allowed essentially an unlimited sample size. The *CLUSTER processor was more efficient, but the maximum sample size was 10,000 to 11,000 data points with four channel Landsat MSS data.

The cost in terms of CPU time was influenced by the degree of supervision in selecting the training sample and the method of developing the spectral classes. As the extent of supervision increased, the sample size and the utilization of clustering techniques decreased, thereby reducing the CPU time required. The inefficiency of the *STATISTICS processor caused the CPU time to be greater than that necessary to develop the spectral classes using the clustering algorithm. With proper utilization of the clustering algorithm (i.e. multi-clustering) the sample size limitation could be avoided.

The main factors influencing the CPU time required by the clustering algorithm were the convergence value, sample size and number of classes. The CPU time was significantly reduced upon lowering the convergence value from 100 to 99 percent. While not affecting the overall classification accuracy, the CPU time was reduced by 50 to 70 percent.

Most efficient use the clustering algorithm could be achieved by multiple rather than single clustering. The CPU time for all of the multi-clustering techniques was less than that for mono-clustering comparable training samples. The effective use of multi-clustering techniques resulted in several advantages, including a dramatic increase in the number of spectral classes, an almost unlimited sample size, and a reduction in CPU time.

In evaluating the various approaches, we found that, in terms of the support data required, the type, scale, age and amount of aerial photography all influenced the accuracy of the interpretation and the amount of information that could be obtained. The amount of aerial photography necessary varied with the type of analysis approach. The heterogeneous block methods for selecting the training sample areas required the least amount of photography. For example, only isolated frames scattered throughout the test site were required in this study -- three (1:15,840) for the Platoro and six (1:100,000) for the S.S.J.M.P.U. analysis.

The number of man-hours required for the various techniques is basically a measure of the supervision by an analyst. The majority of the analyst involvement is necessary for input of the support data either before, during, or after formation of the spectral classes. The three methods of adding the support data represent respectively the three levels of supervision in selecting the training sample; homogeneous fields, heterogeneous blocks, and statistical sample.

The three techniques requiring the selection of homogeneous fields used a considerable number of man-hours. A large number of fields had to be selected to adequately represent all cover types. The supervised techniques would be efficient if the objective was to map a relatively few or a single cover type (i.e. snow, water). However, the supervised method of selecting homogeneous fields is not suitable for large areas due to the requirements, in part, for support data. In this study, we found that the spectral variation due to the topographic relief made collecting sufficient training samples to characterize every cover type impractical.

Other than the three supervised foelds approaches, the technique that required the next greatest number of man-hours was the non-supervised. This approach required very little time to select the training sample, but a considerable amount to identify the spectral classes. The limited number of data points per spectral class in a single field of view on the Zoom Transfer Scope made identification of the spectral classes difficult and time consuming.

The fastest and easiest method to add the ancillary information to the analysis procedure was by selecting and identifying cluster classes for heterogeneous blocks. This approach required very little time to select the training sample and less time to identify the spectral classes that the non-supervise ed approach. By interacting with the analysis procedure in several places, adjustments can be made to closely control the progress. The analyst interacts with the data at the beginning of the procedure by selecting training fields, in the middle of identifying spectral classes and at the end by pooling the spectral-informational classes. With this apporach, the amount of supervision was reduced at each step.

Summary and Conclusions. The various approaches for developing the training statistics were basically a function of (1) the method used for selecting and identifying the training sample and (2) the technique used for generating unimodal spectral classes. The differences between the various techniques were evaluated in terms of the amount of CPU time and man-hours required to define the training statistics and the classification accuracy that was achieved.

The multi-cluster blocks approach was found to be the most effective method of developing training statistics for computer-aided analysis of the Landsat data since it: (a) reduced the CPU time, (b) required relatively few man-hours of time, (c) utilized the lowest amount of support data, and (d) resulted in the highest overall classification accuracy.

A detailed description of the multi-cluster blocks approach and recommended procedures for the implimentation of this technique is contained in the complete report on this project (LARS Technical Report 112277) which is entitled: "Computer-aided Analysis Techniques for an Operational System to Map Forest Lands Utilizing Landsat MSS Data", by M.D. Fleming and R.M. Hoffer.

Bibliography

A list of material available for review in preparation for developing a Landsat forest sample design.

Aldrich, R.C. 1976. Evaluation of Skylab (EREP) Data for Forest and Rangeland Surveys. USDA-Forest Service Research Paper PSE-113. 74 pp.

Department of Natural Resources. 1976. Forest Inventory Phase II. Washington Forest Productivity Study. Prepared for the Pacific Northwest Regional Commission, Operations Research Section, Division of Technical Services, Department of Natural Resources, Olympia, Washington. p. 231-279.

Eisgruber, L.M. 1972. A Three-Stage Sampling Model for Remote Sensing Application. Laboratory for Applications of Remote Sensing Information Note 021072, 27 pp.

Ek, A.R. and J.N. Issos. 1976. Bayesian Estimator, Methodology for Forest Inventory. Presented at Midwest Mensurationists Meeting, Glen Arbor, MI. 16 pp.

Frayer, W.E., G.B. Hartman and D.R. Bowen. 1974. INVENTORY DESIGN AND ANALYSIS. Proceedings of Workshop sponsored by the Soc. of American Foresters. 368 pp.

Freese, F. 1967. Elementary Statistical Methods for Foresters. USDA-Forest Service. Agricultural Handbook 517, 87 pp.

Heller, R.C., R.C. Aldrich, R.S. Driscoll, R.E. Francis and F.P. Weber. 1974. Evaluation of ERTS-1 Data for Inventory of Forest and Rangeland and Detection of Forest Stress. USDA-Forest Service Pacific Southwest and Rocky Mountain Forest and Range Experiment Station. 255 pp.

Hildebrandt, G. 1973. Symposium Proceedings: Remote Sensing in Forestry. Proceedings of Symposium of Subject Group S6.OR of the International Union of Forestry Research Organizations. 535 pp.

Hirsch, B. 1963. FOREST MENSURATION AND STATISTICS. The Ronald Press Company. 474 pp.

Hirsch, B. 1971. PLANNING A FOREST INVENTORY. FAO Forestry and Forest Products Studies, No. 17, 120 pp.

Johnson, J.R. 1975. Small Scale Photo Probability Sampling and Vegetation Classification in Southeast Arizona as an Ecological Base for Resource Inventory. Ph.D. Thesis submitted to Oregon State University. 179 pp.

Kalensky, Z. 1976. Automation of Thematic Mapping Based on Remote Sensing and Computerized Image Interpretation. Forest Management Institute. Canadian Forestry Service, Department of Environment, Ottawa, Canada. 7 pp.

Keays, J.L. 1975. Projection of World Demand and Supply for Wood Fiber to the Year 2000. Journal of the Technical Association of the Pulp and Paper Industry. Vol. 58, No. 11. p. 90-95.

Krebs, P.V. and Staff. 1976. Multiple Resource Evaluation of Region 2 U.S. Forest Service Lands Utilizing Landsat MSS Data, Type III. Final Report prepared for Goddard Space Flight Center. 298 pp.

Langley, P.G. 1976. Sampling Methods Useful to Forest Inventory when using Data from Remote Sensing. In proceedings International Union of Forestry Research Organizations. Remote Sensing in Forestry, p. 313-322.

Mroczynski, R.P. 1976. Application of Satellite Collected and Computer Analyzed Data to the Management of the Central Hardwood Forest. In proceedings International Union of Forestry Research Organization. Remote Sensing in Forestry. p. 171-182.

Mroczynski, R.P. and T.M. Lyons. 1977. A Remote Sensing Evaluation of Forest Resources in Indiana's Coastal Zone. In proceedings First Conference on the Economics of Remote Sensing Information Systems. p. 325-339.

Mouat, D.A. 1974. Relationship between Vegetation and Terrain Variables in Southeastern Arizona, Ph.D. thesis submitted to Oregon State University, 241 pp.

Nichols, J.D., R.A. Harding, R.B. Scott and J.R. Edwards. 1975. Forest Inventory of Western Washington by Satellite Multi-Stage Sampling. Operations Research Section, Division of Technical Services, Department of Natural Resources, Olympia, Washington, 37 pp.

Probine, M.C. 1976. Second Quarterly Report Landsat II Investigation Program 28230. Report No. 531. Physics and Engineering Laboratory, Department of Scientific and Industrial Research, New Zealand. 80 pp.

Rafsnider, G.T. and R.H. Rogers. 1977. Forestland Type Identification and Analysis in Western Massachusetts: A Linkage of Landsat Forest Inventory to an Optimization Study. In Eleventh International Symposium on Remote Sensing of Environment, 9 pp.

Schrumpf, B.J. 1975. Multiseasonal-Multispectral Remote Sensing of Phenological Change for Natural Vegetation Inventory. Ph.D. thesis submitted to Oregon State University. 213 pp.

Titus, S.J. and R.N. Colwell. 1976. Forestry Applications Project/Timber Resource. Sam Houston National Forest Inventory and Development of a Survey Planning Model. Univ. of California, Space Sciences Laboratory, Series 17, Issue 55. 75 pp.

United Nations. 1973. Manual of Forest Inventory with Special Reference to Mixed Tropical Forests. Food and Agriculture Organization of the United Nations. 200 pp.

United Nations. 1975. Second FAO/SIDA Training Course on Forest Inventory. FAO Rome. 291 pp.

United Nations. 1976. Forest Management - A Bibliography with Abstracts, U.S. Dept. of Commerce, NTIS/PS-76/050/157 pp.

Welch, R.L. and R.A. Cathey. 1976. Sampling and Analytical Techniques for an Interim Survey in the South Carolina Lowcountry. USDA-Forest Service Research Paper SE-154. 8 pp.

Williams. D.L. and G.F. Haver. 1976. Forest Land Management by Satellite: Landsat-Desorted Information as input into a Forest Inventory System. NASA/Goddard Space Flight Center, Intralab Project No. 75-1. 36 pp.

References

Chase, Clarence D., Ray E. Pteifer and John S. Spencer, Jr. 1970. The Growing Timber Resource of Michigan 1966. USDA Forest Service Resource Bulletin NC-9, 62pp.

Ferguson, Roland H. and Carl E. Mayer 1970. The Timber Resources of New York. USDA Forest Service Resource Bulletin NE-20, 193 pp.

Fleming, M.D., J. Berkebile and R. Hoffer. 1975. "Computer-Aided Analysis of Landsat-1 MSS Data: A Comparison of Three Approaches Including the Modified Clustering Approach." Proceedings of the Symposium on Machine Processing of Remotely Sensed Data. Purdue University, West Lafayette, Indiana. Pp. 1B-54 to 1B-61. Also available at LARS Information Note 072475.

Fleming, M.D. and R.M. Hoffer. 1977. "Computer-Aided Analysis Techniques for an Operational System to Map Forest Lands Utilizing Landsat MSS Data." LARS Technical Report 112177.Purdue Univ., W. Lafayette, Ind.

Kuchler, A.W. 1964. Potential Natural Vegetation of the Carterminous United States. American Geographical Society Special Publication No. 36, 116 pp.

Spencer, John S. Jr. and Barton L. Esses. 1972. Timber in Missouri, USDA Forest Service Resource Bulletin. NC-30, pp. 108.

Spencer, John S. Jr. and Harry W. Thome. 1972. Wisconsin's 1968 Timber Resource-A Perspective. USDA Forest Service Resource Bulletin. NC-15, 80 pp.

Swain, Philip H. 1976. "Land use Classification and Mapping by Machine-Assisted Analysis of Landsat Multispectral Scanner Data. LARS Information Note 111276. Purdue University, West Lafayette, Indiana.

U.S. Forest Service. 1965. Forest Survey Handbook USDA-Forest Service Manual and Handbook. ref. 4810, 113 pp.

U.S. Forest Service. 1975. The Nation's Renewable Resources-An Assessment 1978. USDA-Forest 3 Vol's. 1100pp.

U.S. Geological Survey. 1974. Earth Resources Survey-Benefit-Cost Study. Appendix 4. Earth Satellite Corp prepared by Boox-Allen Applied Res. Inc. 250 pp.

Weismiller, R.A. 1977. Land Use Inventory of the Great Lakes Basin Using Computer Analysis of Satellite Data. LARS Information Note 011077 in publication.

D. Soil Classification and Survey

For the past ten years soil scientists at Purdue University have been studying the reflectance and emittance from soils in order to establish the relationships between the spectral properties and physicochemical properties of soils. The rationale is that the greater the understanding of these relationships the more useful will be multispectral measurements for mapping and characterizing soils. These studies at the Laboratory for Applications of Remote Sensing have been supported by data from aircraft multispectral scanners, the Landsat 1 and 2 scanners, and an Exotech Model 20 field spectroradiometer.

Since the late 1960's several studies have been conducted using digital analysis of aircraft and satellite scanner data to delineate and map soils with different physical and chemical properties. In support of these studies a limited amount of research has been conducted with the field spectroradiometer in an attempt to define the effects of different soil constituents on the spectral properties of soils.

The research reported under this contract includes an extension of investigations to (a) define quantitatively the soil variables which affect soil reflectance and (b) examine the differences in soil patterns of the same area which are spectrally separable with Landsat data obtained on three different dates.

Spectral Relationships-Outdoor Exotech Experiment. The Exotech Model 20-C spectroradiometer was used to measure the spectral reflectance from surface soils at the Purdue University Agronomy Farm during the week of May 11-12, 1977. An experiment was set up to measure the effects of organic surface residue and soil moisture content on the reflectance of two soils differing greatly in surface soil color, organic matter content, and natural drainage.

Treatment combinations consisted of two levels of moisture content along with two surface soil conditions, i.e., with and without organic residue. Soil moisture differences were obtained by saturating the soil with water several hours before reflectance measurements were taken on half of the plots. The other half of the plots remained at the field moisture content. Organic residue of corn stover was applied at a rate of 2.2 metric tons per hectare to half of the plots. This amount of corn residue represents about four times the amount necessary to reduce erosion (1,2) while not obscuring the soil background to a large extent.

The soils investigated were Chalmers silty clay loam, a dark colored soil developed under tall prairie grass, and Fincastle silt loam, a light colored soil developed under forest vegetation. Chalmers soils occur in depressions on broad, gently undulating till plains on loam to light clay loam glacial till. Fincastle soils occur on broad, gently undulating divides on the border of the prairie uplands from a shallow covering of silt overlying glacial till. According to the USDA/SCS system of Soil Taxonomy(3), Chalmers silty clay loam is classified as a fine loamy, mixed mesic, Typic Argiaquoll. Fincastle silt loam is classified as a fine loamy, mixed, mesic, Aeric Ochraqualf. The observed differences in soil properties for this experiment are summarized in Table D-1.

Two plot sites were selected representing the two soils under investigation. On each site twelve plots measuring three meters by three meters each were marked off, providing three replications of the treatment combinations.

Spectral measurements were made with the Exotech Model 20-C spectroradiometer over the spectral range from 0.4 to 2.4μm. Using the 15° FOV mode, the soil surface area viewed at a height of 6m was approximately 1.6m in diameter. Measurements were taken on two consecutive days, May 11 and 12, 1977. The May 11 readings were performed late in the day at a low sun angle and resulted in a definite specular reflectance component being present in measurements taken from the moistened plots and plots with corn stover. Thus, the May 11 readings were not comparable to the May 12 measurements in quality and were not utilized as repeated observations to verify the repeatability of the spectral responses.

The May 12 measurements were of excellent data quality. Spectral response from the three replications for each treatment showed that the data were consistent and did not deviate from mean response by more than approximately ten percent for any wavelength band studied.

The general shapes of the reflectance curves for the two soils followed the findings of previous studies (4,5) even when corn stover residue was present. The Fincastle soil displayed a smooth convex reflectance curve in the 0.4-1.3μm wavelength region, while the Chalmers soil displayed a smooth concave reflectance curve in the same wavelength region. The fact that the general shape of these curves was not affected by the presence of surface corn stover residue is quite significant for soil survey work because it indicates that spectral response from soils with organic residue does not result in confusion between different soil types.

Another observation was made that in all cases of soil with corn stover residue present on the surface, there was a slight increase in reflectance at 1.3μm. This alteration in the usual soil reflectance curve at the 1.3μm wavelength may indicate that this would be a valuable wavelength region for determining the presence or absence of organic residue on soil.

For the purpose of analysis, the spectral response curves were broken up into the following ten wavelength bands: 0.47-0.52μm, 0.52-0.60μm, 0.63-0.69μm, 0.74-0.80μm, 0.80-0.91μm, 0.98-1.08μm, 1.09-1.19μm, 1.20-1.30μm, 1.55-1.75μm, and 2.10-2.35μm. These represent bands proposed for the thematic mapper satellite in addition to middle infrared bands used on the Skylab S192 multispectral scanner.

Results show that the 1.20-1.30μm band provides the greatest contrast among the eight soil/treatment combinations. Summarized results of spectral response are given in Table D-2 for the 1.20-1.30μm band. The

Table D-1. Summarized soil properties for soil reflectance experiment.

Soil properties	Soil	
	Chalmers silty clay loam	Fincastle silt loam
Natural drainage class	Very poorly drained	Moderately well drained
Dry Munsell color	10YR4/1	10YR6/2
Organic matter content	4.74%	1.39%
Cation exchange capacity (CEC)	38.1 meq/100g	14.6 meq/100g
Base saturation	68.7%	58.4%
Moisture content (upper 1 cm)		
Field moisture level	3.05%	3.64%
Saturated treatment	24.57%	24.40%

Table D-2. Spectral response of soil/treatment combinations in the 1.2-1.3μm wavelength band.

Soil/Treatment Combination	Mean Response Bi-Directional Reflectance Factor	Standard Deviation	Percent Deviation
Fincastle sil*, bare, dry	49.86	4.12	8.27
Fincastle sil, bare, moist	31.39	1.18	3.75
Fincastle sil, corn stover, dry	54.27	5.38	9.91
Fincastle sil, corn stover, moist	40.96	2.97	7.26
Chalmers sicl**, bare, dry	26.71	0.70	2.62
Chalmers sicl, bare, moist	19.85	1.74	8.78
Chalmers sicl, corn stover, dry	37.93	2.02	5.33
Chalmers sicl, corn stover, moist	29.32	1.52	5.17

* silt loam

** silty clay loam

rather large variation in response of the dry Fincastle soil both with and without surface residue caused the greatest problem in selection of an ideal reflectance band. The two middle infrared bands (1.55-1.75μm and 2.10-2.35μm) were not considered optimal because the variance in the response of the dry, light colored soil was too large in this region. Also the ability to distinguish between the dry and moist soil treatments increased with increasing wavelength. Both organic residue and moisture content appeared to affect the magnitude of soil response without altering the characteristic curves of soil reflectance.

There does not appear to be any "masking" effect of organic residue or moisture content on the general response curve of the two soils studied. Observed spectral differences of similar soils using Landsat data may very well be explained by differences in organic residue and moisture content as well as tillage effect, which were not studied here. As indicated by the results of this experiment, Landsat band 7 (0.8-1.1μm) of the current Landsat bands provides the greatest contrast for discrimination of soil moisture differences as well as detection of organic residue.

Recommendations for Further Study. Future field experimentation designed to measure soil spectral response should include the factor of tillage in the experimental design along with surface residue and moisture treatments. Different amounts and kinds of organic surface residue should be studied. Several moisture treatments should also be included in the design. Such a design would be useful for determining whether the observed increase in reflectance in the 1.3μm region can actually be attributed to the effect of surface organic residue.

Conclusions. The ability to identify soil conservation tillage practices from a remote position would permit soil managers to evaluate and monitor the extent of these practices. Results of this soil reflectance experiment show that the effect of surface corn residue can be readily distinguished from bare soil, regardless of soil moisture content.

Cooperative efforts with soil survey personnel serve to gain from a better understanding of the factors affecting soil reflectance in the field situation. The ability to characterize the spectral response of Fincastle and Chalmers soils is not adversely affected by surface organic residue or moisture differences.

Although this study strongly suggests that reflectance data can be used to distinguish between some soil conditions, it must be noted that the data examined represent a very limited number of soils and conditions. It is important that future studies include the effects of (a) kind and quantity of green vegetative cover, (b) surface roughness, (c) a wide range in condition and quantity of plant residue on the surface, and (d) a wide range in soil moisture content on the reflectance from surface soils.

Analysis of Landsat Data. The objective of this phase of the soils study was to examine the spectral patterns produced from Landsat data obtained on different dates over Tippecanoe County, Indiana and to relate those patterns to meaningful soils differences, such as texture, surface color, internal drainage, and organic matter content.

Description of the Study Area. Tippecanoe County is located in west central Indiana within the Tipton Till Plain which is a part of the Central Lowland Province. The underlying bedrock consists of flint, shale, sandstone and limestone of the Mississippian period and is exposed as rock terraces in the Wabash Valley and on the upland in the western part of the county.

The Wabash Valley is the most striking physiographic feature of the county. The bottomlands of the river are approximately 500 m wide. The meandering, rapidly flowing streams of the Wabash and its tributaries have very narrow, discontinuous bottomlands, which are cut off in many places. This is shown very clearly on images and classifications produced from Landsat data. The bottomlands are subject to frequent flooding, and the soils in these floodplains developed from glacial drift washed from the uplands and terraces. Floodplains soils include the well drained Genesee series, yellowish to brown in color; the moderately well drained Eel series, and the imperfectly drained Shoals series.

Terraces exist along the major streams separated from the bottomlands by steep slopes. Soils of the terraces include the excessively drained, dark Elston series associated with the dark Wea series on the outwash plain. The brown, well drained Ockley soils are associated with the poorly drained Westland and Abingdon series. The well drained Fox series developed under a mixed forest cover and occur mostly on terraces above the larger streams.

The upland soils are on the glacial till plain which is about 200 meters above sea level. Several different upland series are found in the county. The Hennepin series occurs in the valleys of the Wabash and Tippecanoe Rivers and Wildcat Creek and is associated with the well drained Miami series. The very dark grayish Odell is an imperfectly drained soil which occurs on the nearly level prairie land. It is associated intricately with the very poorly drained Chalmers and Romney series. The Sidell series has a relatively high percent of organic matter and good natural drainage. Very poorly drained soils, developed in swales and depressions belong to the Brookston series which is associated with the imperfectly drained Crosby and Fincastle series. On slightly elevated areas the somewhat poorly drained Raub and very poorly drained Ragsdale series occur.

Procedure. The multispectral scanner data used in this study were obtained during Landsat passes on 9 June 1973, 6 April 1975 and 29 June 1976. More than two centimeters of precipitation were recorded several hours prior to the Landsat pass on 6 April 1975. Dry weather conditions and unusually high temperatures prevailed for several days prior to the 9 June 1973 pass. Rainfall was recorded on each of the three days preceding the 29 June 1976 pass.

The Landsat data for three dates for Tippecanoe were geometrically corrected (6), overlaid, and spatially registered to ground control points selected from a set of U.S. Geological Survey 7½-minute quadrangle sheets. The resulting multiband, multidate data set when printed on a computer line printer in pictorial form had a scale of 1:24,000.

The 6 April 1975 data for Tippecanoe County were subjected to nonsupervised clustering procedures to obtain 28 spectrally separable cluster classes. This number was selected because it represents two times the 13 soil associations of the county plus two classes for man-made features. Limited ground observation data and black and white aerial photography were used as aids in identifying and describing the 28 cluster classes. Total MSS relative reflectance values (magnitude) and visible/infrared reflectance ratio values for these 28 classes from the three dates (Table D-3) were used to explain probable soil patterns and other surface features.

Three broad categories--land areas without vegetative cover, land areas with vegetative cover, and water--were separated spectrally. Terrestrial ecosystems having green vegetative cover are highly reflective in the near and middle infrared and lowly reflective in the visible spectrum. In general, the areas having ratio values less than 1 were classified as green vegetation, ratios between 1.1 and 1.6 as non-vegetated areas, and ratios above 1.7 as water.

In the merging of the 28 spectral classes into the three broad categories for the three dates it is immediately apparent that April data have definite advantages in identifying and mapping soils differences (Table D-3). For example, on 6 April 1975 only 18.9% of the county was classified as having green vegetative cover. On 29 June 1976 98% of the county was vegetated. It should be noted that even within the urbanized area of Greater Lafayette, a significant percentage of the June data were classified as green vegetation because of parks and residences with lawns and trees. Water bodies, residential areas and roads showed some interesting spatial and areal shifts from April to June. The separability and measurability with digital analysis of Landsat data of each of these features seem to be affected significantly by the appearance of green vegetation. In many instances in the late June data green vegetation has come to dominate the spectral response in the areas along roads, streams, and among residences where it did not dominate a few weeks earlier. For this reason the classes of small or narrow bodies of water and roads may have been more easily identified and separated spectrally early in the season than those same features encircled by green trees or grass later in the season.

When the spectral classification objective includes the identification and mapping of different classes of green vegetation, bare soils and water both total relative reflectance for all four Landsat bands and the visible/infrared ratio values are useful statistics. However, in this study to delineate soils boundaries the total relative reflectance or magnitude was the principal data feature used.

For a general overview of the spectral patterns of Tippecanoe County on the three different dates, gray scale images for each MSS waveband at a scale of 1:60,000 were produced on an electronic printer/plotter. An area in the northeast corner of the county where the Tippecanoe River

Table D-3. Spectral statistics from three dates of Landsat MSS data for twenty-eight cluster classes in Tippecanoe County, Indiana.

Spectral Class	Dates of Landsat Pass					
	6 April 1975		9 June 1973		29 June 1976	
	Mag*	Ratio**	Mag*	Ratio**	Mag*	Ratio**
1	143	0.85	155	0.69	145	0.63
2	126	0.91	155	0.90	131	0.55
3	88	0.98	128	1.08	116	0.50
4	104	0.85	152	1.01	122	0.72
5	154	0.73	163	0.69	153	0.51
6	142	0.61	165	0.62	162	0.42
7	104	1.09	155	0.80	125	0.43
8	165	0.99	158	0.68	158	0.56
9	154	1.21	180	1.16	163	0.76
10	127	1.32	174	1.21	128	0.71
11	118	1.14	166	1.07	138	0.71
12	115	1.38	198	1.29	133	0.65
13	106	1.36	173	1.18	135	0.73
14	114	0.93	172	1.07	139	0.68
15	98	1.39	181	1.26	125	0.74
16	93	1.18	143	0.56	133	0.32
17	91	1.47	156	1.26	112	0.69
18	82	1.48	151	1.11	107	0.66
19	70	1.67	128	1.26	100	0.58
20	92	1.74	139	1.26	104	0.90
21	106	1.48	150	1.15	129	0.76
22	109	1.67	143	1.56	133	1.21
23	204	1.73	167	1.73	187	1.40
24	99	1.25	175	0.86	125	0.49
25	87	2.11	124	1.39	98	0.86
26	90	2.30	114	1.79	103	0.81
27	75	2.49	120	1.07	97	0.71
28	60	3.58	92	1.92	63	1.18

* Mean total relative reflectance for 4 Landsat MSS bands.

**Visible/infrared reflectance ratio values.

Table D-4. Summary of three major categories of surface features on three dates in Tippecanoe County, Indiana.

Classification	6 April 1975		9 June 1973		29 June 1976	
	Area (Ha)	% of County	Area (Ha)	% of County	Area (Ha)	% of County
Vegetated	23,848	18.9	69,763	55.4	123,490	98.0
Non-vegetated	101,268	80.4	55,145	43.8	635	0.5
Water	875	0.7	1,083	0.8	1,866	1.5
Total	125,991	100	125,991	100	125,991	100

joins the Wabash River is reproduced in this report, showing gray scale images of the 0.8-1.1μm waveband for the three dates (Figures D-1, D-2, D-3). The same general soils patterns are visually distinguishable in these three images.

The soils in the northwest corner of the county were formed under tall prairie grass and are quite dark throughout the surface or A horizon. A classification of that area was produced by the merging of seven cluster classes from the 6 April 1975 Landsat data. The resulting 1:24,000 scale map gave a good delineation between the Odell-Chalmers association (M) and the Raub-Ragsdale (F,/) association (Figure D-4). A comparison of the spectral classification with the existing soil map revealed a close similarity, showing that spectral data could be used to separate different dark soils having low relative reflectance. The same area classified with 29 June 1976 Landsat data in which 10 cluster classes were merged produced similar soils patterns (Figure D-5).

The outwash and bottomland soils along the Wabash River generally have a higher reflectance than do the upland prairie soils. A spectral classification of an area along the Wabash River was produced using the four bands from the available 12 bands (4 each from three dates) which gave the best separability of the features in the scene (Figure D-6). Prominent bottomland soils are the well drained Genesee silt loam (L) on the upper or left side of the Wabash River and the Eel and Genesee silty clay loam (I) on the right of the river.

Conclusions. From this study the following conclusions were drawn:

1. Digital analysis of Landsat MSS data can be used effectively in identifying and mapping meaningful soils patterns at a scale as large as 1:24,000.

2. Mean total relative reflectance for cluster classes is a valuable data feature for determining important differences between soils and for determining which spectral classes can best be merged.

3. Digital analysis of Landsat data for delineating soils differences can greatly reduce the field work required for detailed soil surveys.

Figure D-1. Gray scale image of area around the junction of the Tippecanoe and Wabash Rivers, Tippecanoe County, Indiana. Landsat MSS band 7 (0.8-1.1μm), 6 April 1975. Scale 1:60,000.

Figure D-2. Gray scale image of area around the junction of the Tippecanoe and Wabash Rivers, Tippecanoe County, Indiana. Landsat MSS band 7 (0.8-1.1μm), 9 June 1973. Scale 1:60,000.

Figure D-3. Gray scale image of area around the junction of the Tippecanoe and Wabash Rivers, Tippecanoe County, Indiana. Landsat MSS band 7 (0.8-1.1μm), 29 June 1976. Scale 1:60,000.

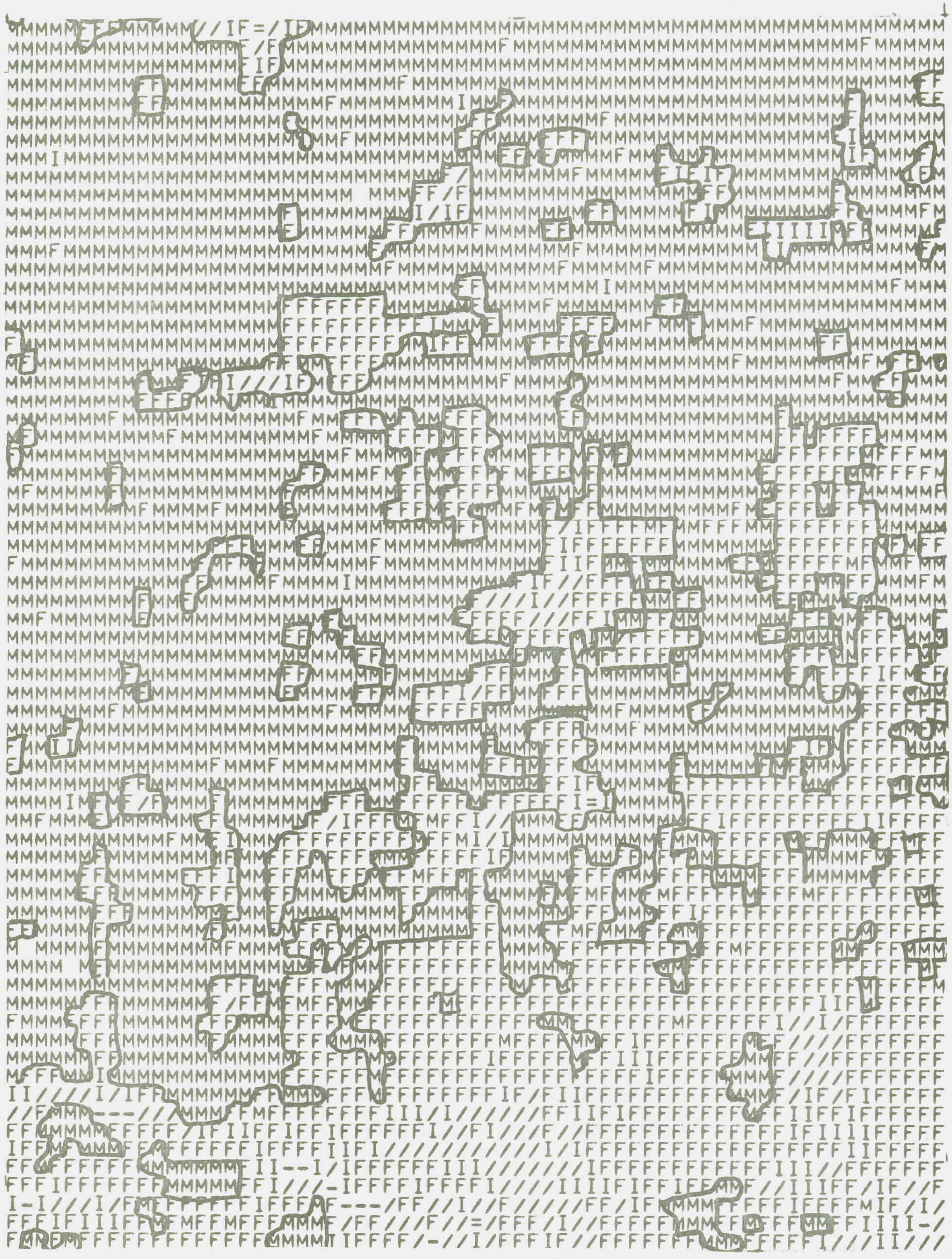


Figure D-4. Spectral classification of 6 April 1973 MSS data for area in northwest Tippecanoe County, Indiana delineating two associations of upland soils: Odell-Chalmers (M) and Raub-Ragsdale (F,/). Scale 1:24,000.

Figure D-5. Spectral classification of 29 June 1976 MSS data for northwest Tippecanoe County, Indiana delineating two associations of upland soils: Odell-Chalmers (M) and Raub-Ragsdale (F,0,4,I,/). Scale 1:24,000.

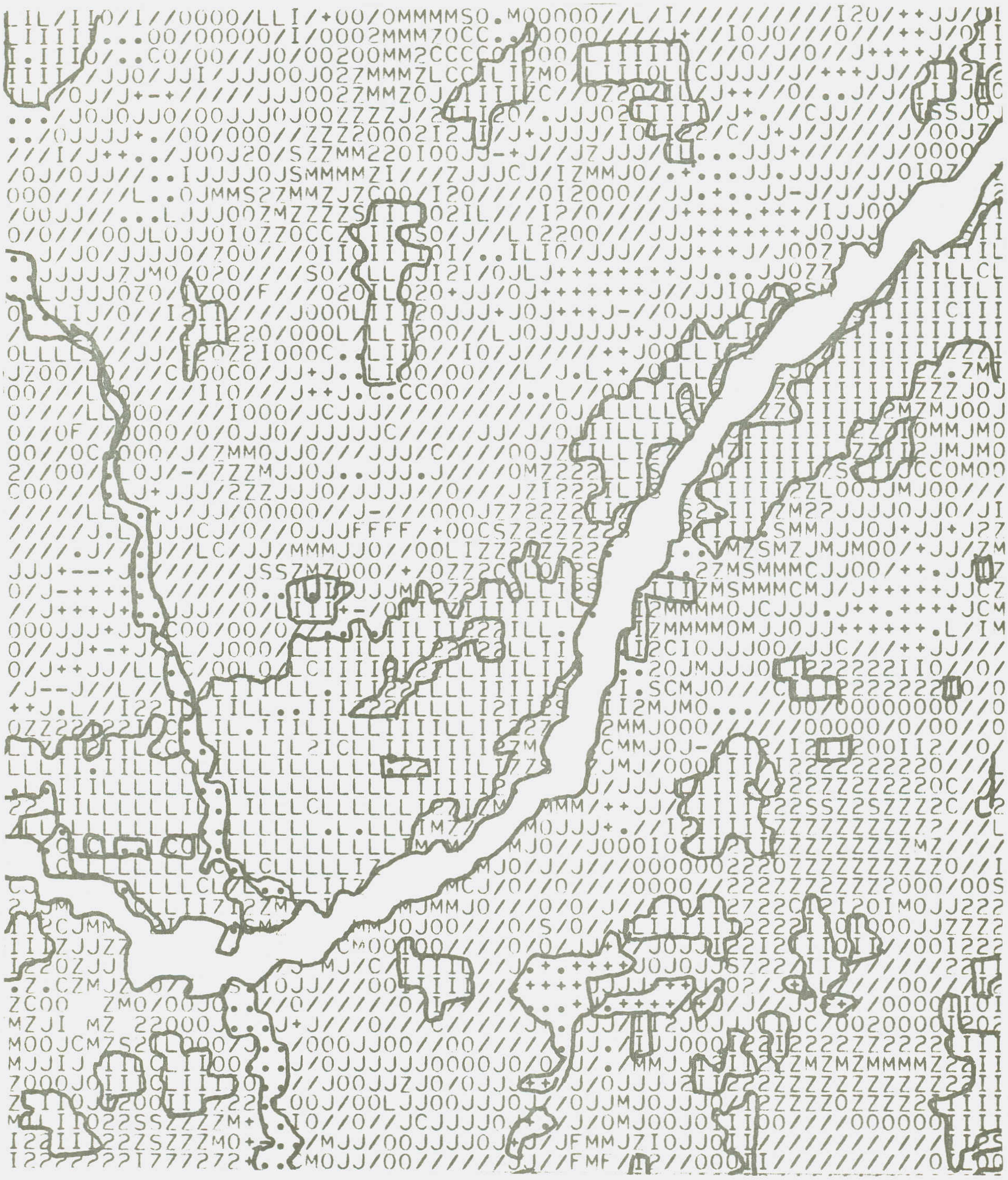


Figure D-6. Spectral classification using best four of twelve bands (3 dates of Landsat MSS data) to delineate soils on terraces and bottomlands along Wabash River, Tippecanoe County, Indiana. Scale 1:24,000.

L - Genesee silt loam

I - Genesee and Eel silty clay loam

REFERENCES

1. Mannering, J. V., J. D. Meyer, and L. D. Meyer. 1963. The effect of various rates of surface mulch on infiltration and erosion. Soil Sci. Soc. Am. Proc. 27:84-86.

2. Griffith, D. R., J. V. Mannering, and W. C. Moldenhauer. 1977. Conservation tillage in the eastern corn belt. J. Soil and Water Conservation 32:20-28.

3. Soil Survey Staff. 1975. Soil taxonomy--a basic system of soil classification for making and interpreting soil surveys. Soil Conservation Service, U.S. Dept. of Ag., Agriculture Handbook No. 436, Washington, D.C.

4. Condit, H. R. 1970. The spectral reflectance of American soils. Photogrammetric Eng. 36:955-966.

5. Cipra, J. E., M. F. Baumgardner, E. R. Stoner, and R. B. MacDonald. 1971. Measuring radiance characteristics of soil with a field spectroradiometer. Soil Sci. Soc. Amer. Proc. 35:1014-1017.

6. Anuta, P. A. 1973. Geometric correction of ERTS-1 digital MSS data. Laboratory for Applications of Remote Sensing. Information Note 103073. Purdue University, West Lafayette, Indiana.

Final Report Distribution List

NAS9-14970

NAME	NUMBER OF COPIES
NASA/Johnson Space Center Houston, Texas 77058	
ATTN: J. D. Erickson/SF3	(1)
ATTN: M. C. Trichel/SF3	(1)
ATTN: L. F. Childs/SF	(1)
ATTN: K. J. Demel/SF5	(1)
ATTN: F. Weber/SF5	(1)
ATTN: G. O. Boatwright/SF3	(1)
ATTN: K. Baker/SF4	(1)
ATTN: H. G. DeVezin, Jr./FM8	(1)
ATTN: R. P. Heydorn/SF3	(1)
ATTN: M. C. McEwen/SF3	(1)
ATTN: D. H. Hay/SF12	(1)
ATTN: D. L. Amsbury/SF5	(1)
ATTN: J. G. Garcia/SF3	(1)
ATTN: F. G. Hall/SF2	(1)
ATTN: B. L. Carroll/C09	(1)
ATTN: E. Laity/SF121	(2)
ATTN: R. Shirkey/JM6	(4)
ATTN: J. T. Wheeler/AT3	(1)
ATTN: G. E. Graybeal/SF4	(2)
ATTN: I. D. Browne/SF3	(5)
IBM Corporation FSD Mail Code 56 1322 State Park Drive Houston, Texas 77058	
ATTN: R. E. Oliver	(1)
Department of Mathematics Texas A&M University College Station, Texas 77843	
ATTN: L. F. Guseman, Jr.	(1)
ERIM P. O. Box 8618 Ann Arbor, Michigan 48107	
ATTN: R. F. Nalepka	(1)
ATTN: W. A. Malila	(1)
ATTN: R. C. Cicone	(1)
Kansas State University Department of Statistics, Calvin 19 Statistical Lab Manhattan, Kansas 66506	
ATTN: A. M. Feyerherm	(1)

NAME	NUMBER OF COPIES
U. S. Department of Interior Geological Survey GSA Building, Room 5213 Washington, D. C. 20242	
ATTN: Mr. W. A. Fischer	(1)
NASA Wallops Wallops Station, Virginia 23337	
ATTN: Mr. James Bettle	(1)
U. S. Department of Interior EROS Office Washington, D. C. 20242	
ATTN: Dr. Raymond W. Fary	(1)
U. S. Department of Interior EROS Office Washington, D. C. 20242	
ATTN: Mr. William Hemphill	(1)
NAVOCEANO, Code 7001 Bay St. Louis, MS 39520	
ATTN: Mr. J. W. Sherman, III	(1)
University of Texas at Dallas Box 688 Richardson, Texas 75080	
ATTN: Dr. Patrick L. Odell	(1)
Department of Mathematics University of Houston Houston, Texas 77004	
ATTN: Dr. Henry Decell	(1)
U. S. Department of Agriculture Statistical Reporting Service Room 4833, South Bldg. Washington, D. C. 20250	
ATTN: W. H. Wigton	(1)
Goddard Space Flight Center National Aeronautics & Space Administration Greenbelt, Maryland 20771	
ATTN: Mr. W. Alford, 563	(1)
ATTN: Dr. J. Barker, 923	(1)

NAME	NUMBER OF COPIES
EROS Data Center U. S. Department of Interior Sioux Falls, South Dakota 57198 ATTN: Mr. G. Thorley	(1)
U. S. Department of Agriculture Soil & Water Conservation Research Division P. O. Box 267 Weslaco, Texas 78596 ATTN: Dr. Craig Wiegand	(1)
U. S. Department of Interior USGA National Center Mail Stop 115 Geography Program Reston, Virginia 22092 ATTN: Dr. James R. Anderson	(1)
Director, Remote Sensing Institute South Dakota State University Agriculture Engineering Building Brookings, South Dakota 57006 ATTN: Mr. Victor I. Myers	(1)
U. S. Department of Agriculture Forest Service 240 W. Prospect Street Fort Collins, Colorado 80521 ATTN: Dr. Richard Driscoll	(1)
University of California School of Forestry Berkeley, California 94720 ATTN: Dr. Robert Colwell	(1)
Environmental Remote Sensing Applications Laboratory Oregon State University Corvallis, Oregon 97331 ATTN: Dr. Barry J. Schrumpf	(1)
U. S. Department of Interior Director, EROS Program Washington, D. C. 20242 ATTN: Mr. J. M. Denoyer	(1)

NAME	NUMBER OF COPIES
Remote Sensing Laboratory 129 Mulford Hall University of California Berkeley, California 94720 ATTN: C. M. Hay	(1)
Department of Mathematics University of Houston Houston, Texas 77004 ATTN: H. P. Decell, Jr.	(1)
NASA Lyndon B. Johnson Space Center Public Affairs Office, Code AP Houston, Texas 77058	(1)
National Aeronautics and Space Administration Scientific and Technical Information Facility Code KS Washington, D. C. 20546	(1)
Department of Watershed Sciences Colorado State University Fort Collins, Colorado 80521 ATTN: Dr. James A. Smith	(1)
NASA/Johnson Space Center Earth Resources Program Office Office of the Program Manager Houston, Texas 77058	(1)
NASA/Johnson Space Center Earth Resources Program Office Program Analysis & Planning Office Houston, Texas 77058 ATTN: Dr. O. Glenn Smith/HD	(1)
NASA/Johnson Space Center Earth Resources Program Office Systems Analysis and Integration Office Houston, Texas 77058 ATTN: Mr. Richard A. Moke/HC ATTN: Mr. M. Jay Harnage, Jr./HC	(1) (1)
Earth Resources Laboratory, GS Mississippi Test Facility Bay St. Louis, Mississippi 39520 ATTN: Mr. D. W. Mooneyhan	(1)

NAME	NUMBER OF COPIES
Lewis Research Center National Aeronautics & Space Administration 21000 Brookpark Road Cleveland, Ohio 44135	
ATTN: Dr. Herman Mark	(1)
John F. Kennedy Space Center National Aeronautics & Space Administration Kennedy Space Center, Florida 32899	
ATTN: Mr. J. P. Claybourne/AA-STA	(1)
Texas A&M University Institute of Statistics College Station, Texas 77843	
ATTN: Dr. H. O. Hartley	(1)
Code 168-427 Jet Propulsion Laboratory 4800 Oak Grove Drive Pasadena, California 91103	
ATTN: Mr. Fred Billingsley	(1)
NASA Headquarters Washington, D. C. 20546	
ATTN: Mr. W. Stoney/ER	(1)
ATTN: Mr. Leonard Jaffee/ER	(1)
ATTN: Mr. M. Molloy/ERR	(1)
ATTN: Mr. James R. Morrison/ERR	(1)
ATTN: Ms. Ruth Whitman/ERR	(1)
Texas A&M University Remote Sensing Center College Station, Texas 77843	
ATTN: Mr. J. C. Harlan	(1)
USGS National Center Mail Stop 115 Geography Program Reston, Virginia 22092	
ATTN: James Wray	(1)
Canada Centre For Remote Sensing 2464 Sheffield Road Ottawa, Canada K1A 0Y7	
ATTN: Dr. David Goodenough	(1)