LARS Contract Report 120279

Final Report

Annual Technical Summary

by **D. A. Landgrebe and Staff**

December 1, 1978 - November 30, 1979

Prepared for
National Aeronautics and Space Administration
Johnson Space Center
Earth Observation Division
Houston, Texas 77058
Contract No. NAS9-15466
Technical Monitor: J. D. Erickson/SF3

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

80-10117

AgRISTARS

SR-P9-00414
NAS9-15466 NASA CR-160551

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

Supporting Research

November 1979

Final Report

Annual Technical Summary

by D. A. Landgrebe and Staff

(E80-10117) [REMOTE SENSING RESEARCH STUDIES] Final Report, 1 Dec. 1978 - 30 Nov. 1979 (Purdue Univ.) 126 p HC A07/MF A01 CSCL 02C G3/43 N80-30823 Unclas 00117

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

1 Report No	2 Government Accession No	3 Recipient's Catalog No
SR-P9-00414		

4 Title and Subtitle	5 Report Date
Annual Technical Summary Report	November 30, 1979
	6 Performing Organization Code
7 Author(s)	8 Performing Organization Report No
D.A. Landgrebe and Staff	120279
9 Performing Organization Name and Address	10 Work Unit No
Laboratory for Applications of Remote Sensing Purdue University 1220 Potter Drive, West Lafayette, Ind. 47906	11 Contract or Grant No NAS9-15466
	13 Type of Report and Period Covered Final 12/1/78 to 11/30/79
12 Sponsoring Agency Name and Address NASA/Johnson Space Center Earth Observations Division Houston, Texas 77058	14 Sponsoring Agency Code

15 Supplementary Notes

Principal Investigator: D.A. Landgrebe
Technical Monitor: J.D. Erickson

16 Abstract

This report summarizes the work and results obtained in a broad set of remote sensing research studies. These include: field studies of the spectral characteristics of crops and soils; development of multispectral data acquisition system for field research; status of a field research data base and a data handling and analysis system for it; application and evaluation of Landsat training, classification and area estimation procedures for crop inventory; initial development of spectral inputs for corn yield models; research of multistage and contextual classifiers; ambiguity reduction for training sample labeling; construction and evaluation of multitype data sets involving Landsat MSS, synthetic aperture radar, ancillary data; and status and accomplishments of a multilocation, multiterminal computer processing system for supporting remote sensing research. More complete, detailed reports of the studies are contained in the four volume final report of the contract.

17 Key Words (Suggested by Author(s))	18 Distribution Statement
Field research, spectral characteristics, crop identification, multitype data, multistage and contextual classifiers.	

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

NASA - JSC

SR-P9-00414
NAS9-15466
LARS 120279

ANNUAL TECHNICAL SUMMARY REPORT

BY

D.A. Landgrebe and Staff

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

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

November 1979

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Page

Page

Preface

This report summarizes the work and results obtained under contract NAS9-15466 during the period from December 1, 1978 to November 30, 1979. The complete report of this work is contained in the contract Final Report published in four volumes as follows:

Vol. I	Agricultural Scene Understanding and Supporting Field Research
Vol. II	Processing Techniques Development Part 1: Crop Inventory Techniques
Vol. III	Processing Techniques Development Part 2: Data Preprocessing and Information Extraction Techniques
Vol. IV	Computer Processing Support
Vol. V	Technical Summary

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

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

Technical reports and papers published in 1979 by Contract NAS9-15466

Ahlrichs, J.S. and M.E. Bauer. 1979. Relation of Crop Canopy Variables to the Multispectral Reflectance of Spring Wheat. (LARS Tech. Rpt. 072479)

Hixson, M.M., B.J. Davis and M.E. Bauer. 1979. Stratification and Sample Selection for Multicrop Experiments. (LARS Contract Rpt. 112278).

Hixson, M.M., B.J. Davis, and M.E. Bauer. 1979. Sampling for area estimation: an evaluation of the bias and precision of several sampling schemes. In Proc. Fifth Symp. on Machine Processing of Remotely Sensed Data, Purdue Univ. p. 97-104. (LARS Tech. Rpt. 062779).

Landgrebe, D.A., and M.J. Muasher.. 1979. Training samples for classification of multispectral earth observation data. In Proc. Fifth Symp. on Machine Processing of Remotely Sensed Data, Purdue Univ. p. 359.

Mobasseri, B.G., C.D. McGillem and P.E. Anuta. 1979. A parameter model for multispectral scanners. In Proc. Fifth Symp. on Machine Processing of Remote Sensed Data, Purdue Univ. pp. 213-221.

Peterson, J.B., B.F. Robinson and R.H. Beck. 1979. Predictability of change in soil reflectance on wetting. In Proc. Fifth Symp. of Machine Processing of Remotely Sensed Data, Purdue Univ. p. 264-274.

Robinson, B.F., M.E. Bauer, L.L. Biehl and L.F. Silva. 1978. The Design and Implementation of a Multiple Instrument Field Experiment to Relate the Physical Properties of Crops and Soils to Their Multispectral Reflectance. In Proc. International Symp. on Remote Sensing for Observation and Inventory of Earth Resources and Environment, Freiberg, West Germany.

Robinson, B.F., and L.L. Biehl. 1979. Calibration Procedures for Measurement of Reflectance Factor for Remote Sensing Field Research. In SPIE Vol. 196, Measurements of Optical Radiations, pp. 16-26.

Robinson, B.F., M.E. Bauer, D.P. DeWitt, L.F. Silva and V.C. Vanderbilt. 1979. Multiband Radiometer for Field Research. In SPIE Vol. 196, Measurements of Optical Radiations, pp. 8-15.

Scholz, D., N. Fuhs, and M.M. Hixson. 1979. An Evaluation of Several Different Classification Schemes: Their Parameters and Performance. In Proc. 13th Int'l. Symp. on Remote Sensing of Environment, Ann Arbor, Michigan.

Stoner, E.R., M.F. Baumgardner, R.A. Weismiller, L.L. Biehl and B.F. Robinson. 1979. Extension of laboratory-measured soil spectra to field conditions. In Proc. Fifth Symp. on Machine Processing of Remotely Sensed Data, Purdue Univ. p. 253-263.

Stoner, E.R., M.F. Baumgardner, L.L. Biehl and B.F. Robinson 1979. Atlas of Soil Reflectance Properties. (LARS Tech. Rpt. 111579).

Swain, P.H., H.J. Siegel and B. Smith. 1979. A method for classifying multispectral remote sensing data using context. In Proc. Fifth Symp. on Machine Processing of Remotely Sensed Data. Purdue Univ. p. 343-353. (LARS Tech. Rpt. 062379).

Vanderbilt, V.C., M.E. Bauer, and L.F. Silva. 1979. Prediction of solar irradiance distribution in a wheat canopy using a laser technique. Ag. Meteorology 20:147-160. (LARS Tech. Rpt. 042079).

Wiersma, D.J., and D.A. Landgrebe. 1979. An analytical approach to the design of spectral measurements in the design of multispectral sensor. In Proc. Fifth Symp. on Machine Processing of Remotely Sensed Data. Purdue Univ. p. 331-342. (LARS Tech. Rpt. 062279).

Wiersma, D.J. and D.A. Landgrebe. 1979. The analytical design of spectral measurements for multispectral remote sensor systems. (LARS Tech. Rpt. 12267

1A. FIELD RESEARCH EXPERIMENT DESIGN AND DATA ANALYSIS

Marvin E. Bauer

Major advancements have been made in recent years in the capability to acquire, process, and interpret remotely sensed multispectral measurements of the energy reflected and emitted from crops, soils, and other earth surface features. As a result of experiments such as the Large Area Crop Inventory Experiment (LACIE), the technology is moving rapidly toward operational applications. There is, however, a continuing need for more quantitative knowledge of the multispectral characteristics of crops and soils if further advancements in technology development and application are to be made.

Understanding of the relationships between the spectral characteristics and important biological-physical parameters of earth surface features can best be obtained by carefully controlled studies over fields and plots where complete data describing the condition of targets are attainable and where frequent, timely spectral measurements can be obtained. It is these attributes which distinguish field research from other remote sensing research activities.

In 1975, field research activities in support of LACIE were initiated. Spectral, agronomic, and meteorological measurements were made at LACIE test sites in Kansas, South Dakota and North Dakota for three years. The LACIE Field Measurements data form one of the most complete and best documented data sets acquired for agricultural remote sensing research. Analyses of the field data have provided insight into the spectral properties and spectral identification and assessment of crops. The analyses have included investigations of the spectral separability of barley and spring wheat; determination of the effects of cultural and environmental factors on the spectral reflectance of wheat; development of predictive relationships between spectral and agronomic variables related to wheat growth and yield; and comparisons of Landsat MSS and thematic mapper spectral bands for crop identification and assessment.

The objective of this task has been to: (1) design a multiyear technical program of Supporting Field Research which will result in broadly applicable quantitative and predictive relations between crop reflectance and emittance spectra and crop attributes of interest and (2) perform initial data analyses to support objective 1.

1. Field Research Experiment Design

The overall objectives of the multicrop supporting field research are to design, acquire and make available well-documented spectral data sets with detailed agronomic and physical observations. The data will be used for analysis and modeling to obtain a quantitative understanding of the radiation characteristics of crops and their soil background and to assess the capability of current, planned and future satellite systems to capture available useful spectral information (Figure 1A-1). Specific research objectives for which research was initiated in 1979 are:

- Determine the relationship of crop growth and development stages to the reflectance and radiant temperature characteristics of corn and soybeans.

- Evaluate the potential for estimating from spectral measurements important agronomic attributes of crop canopies such as leaf area index, biomass and percent soil cover which may be related to grain yield.

- Determine the effects of stressed, normal, and above average growing conditions on the raidation patterns and spectral identification of corn, soybeans, and small grains. Stresses of particular interest are moisture deficits, nutrient deficiencies and diseases.

- Identify and quantify the agronomic, environmental and measurement conditions which are the dominant factors determining observed spectral characteristics.

- Determine the effects of soil background conditions such as color, texture, roughness and moisture on the spectral response of crop canopies.

- Compare and evaluate the capability of present, planned and possible future sensors to capture available spectrally derived information describing crops and soils.

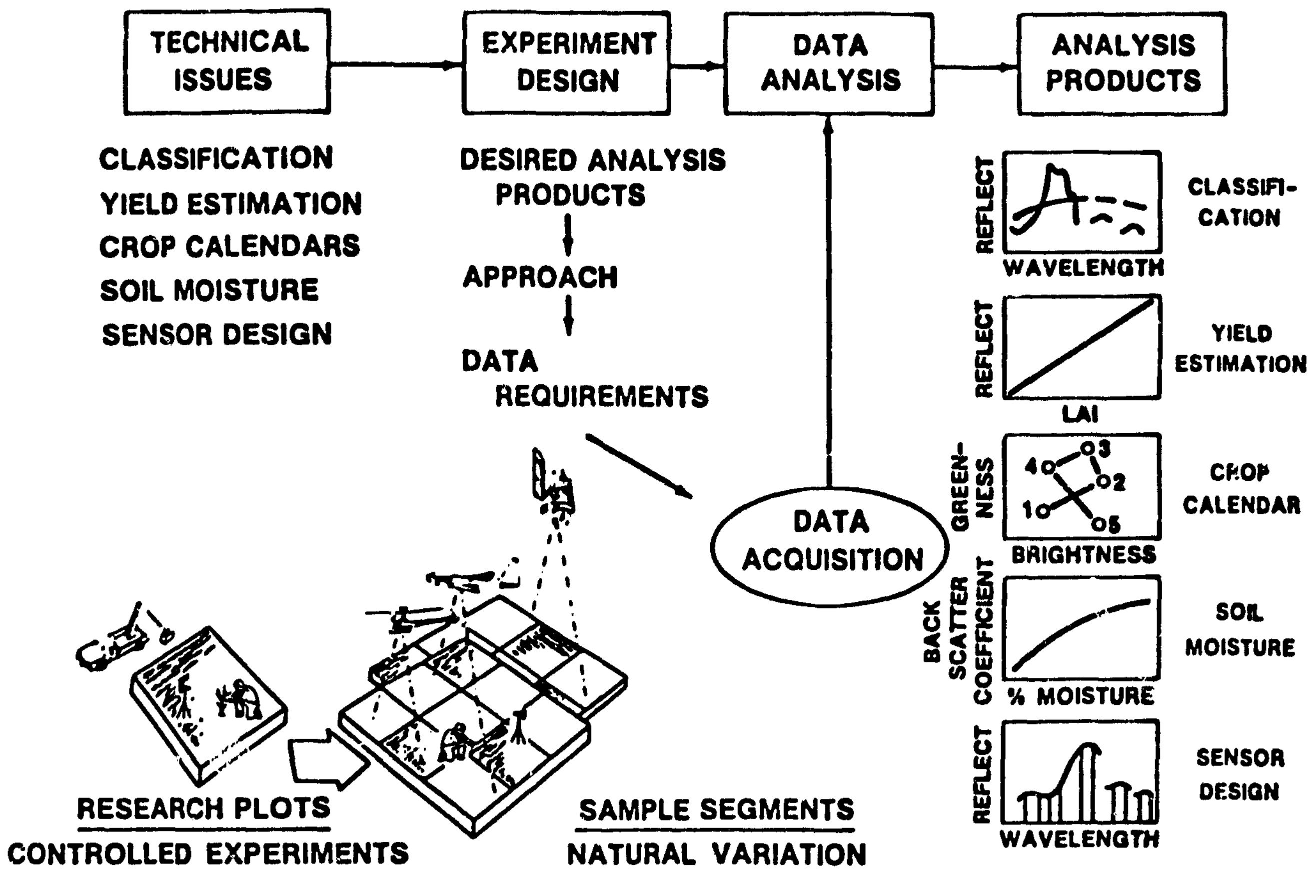


Figure 1A-1. The role of field research in the development and application of remote sensing technology.

At the beginning of the project the technical issues and specific objectives to be addressed with field research data were defined. This led to the definition of the experimental design for data acquisition and processing and the initial definition of data analysis plans and products.

A multistage approach to data acquisition with areal, vertical, and temporal staging is used. Areal sampling is accomplished with test sites in different parts of the Great Plains and Corn Belt. Vertical staging, or collection of data by different sensor systems at different altitudes, ranges from mobile towers to Landsat. Temporally, data is collected at seven to 21-day intervals to sample important crop growth stages, and during several years to obtain a measure of the year-to-year variations in growing conditions and their influence on spectral response.

The test locations are two-five by six mile segments in Hand County, South Dakota, and Webster County, Iowa, and two agricultural experiment stations: the Purdue University Agronomy Farm at West Lafayette, Indiana, and the University of Nebraska, Sandhills Agricultural Laboratory near North Platte, Nebraska(Figure 1A-3). Winter wheat and spring wheat are both major crops in Hand County, while the other sites were selected for study of corn and soybeans.

The staging of data acquisition is summarized in Figure 1A-2. Helicopter-spectrometer and aircraft-scanner data are collected over commercial fields in a series of flightlines over the sites in Hand and Webster counties. Landsat MSS data is also acquired and processed for each Landsat overpass during the seasons over the entire test site. These data provide a measure of the natural variation in the temporal-spectral characteristics of wheat, corn, and soybeans and surrounding cover types.

The truck-mounted spectrometer and radiometer collect spectra of crops in controlled experimental plots of corn, soybeans, and winter wheat at the Purdue Agronomy Farm. These data combined with the more detailed and quantitative measurements of crop and soil conditions which can be made on the plots, enable more complete understanding and interpretation of the spectra collected from commercial fields. Past experience

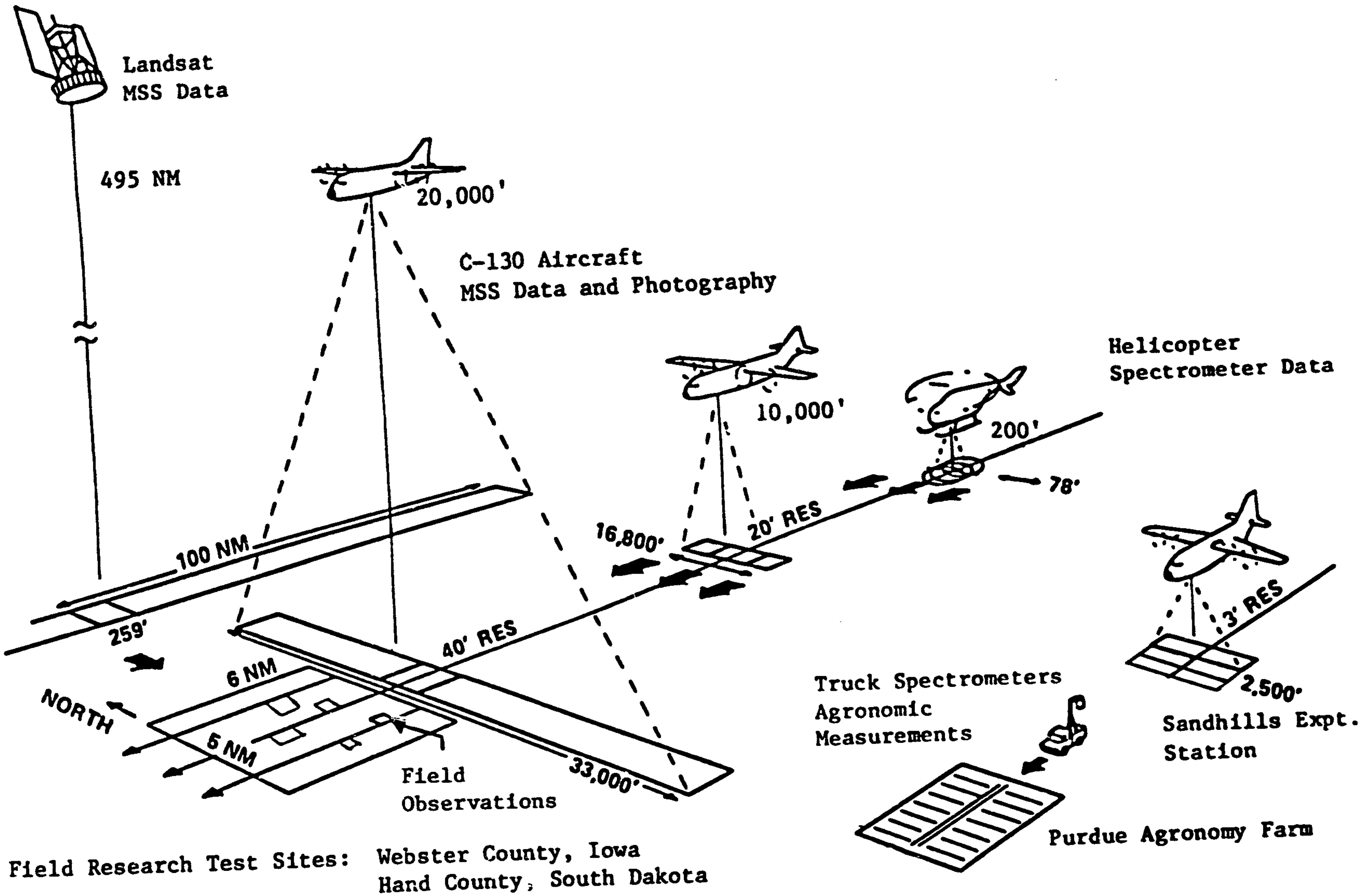


Figure 1A-2. Schematic illustration of multistage approach to multicrop field research data acquisition.

1. Webster County, Iowa
2. Hand County, South Dakota
3. Purdue Agronomy Farm, West Lafayette, Indiana
4. Sandhills Experiment Station, North Platte, Nebraska

Figure 1A-3. Locations of 1979 Supporting Field Research test sites.

has shown that there are generally too many interacting variables in commercial fields to determine exact causes of observed differences in spectral response. With data from plots where only two to four factors are varied undercontrolled conditions, it is possible to determine more exactly and understand more fully the energy-matter interactions occurring in crops.

The spectral measurements are supported by descriptions of the targets and their condition. The observations, counts, and measurements of the crop canopy include: maturity stage, plant height, biomass, leaf area index, percent soil cover, and grain yield. Also, included are measurement conditions such as sensor altitude and view angle, as well as measurements of the atmospheric and meteorological conditions. The data are supplemented by aerial photography and ground-level vertical and oblique photographs of the fields and test plots.

A data library of all spectral, agronomic, meteorological and photographic data collected is maintained at Purdue/LARS. The data are processed in standard data formats and measurements units and are made available to NASA/JSC-supported investigators and other interested researchers upon request.

2. Data Analysis

During the past year two studies of wheat described in last year's final report have been completed and will be submitted to technical journals. In completing the work, our previous analyses of the relationships of crop canopy variables such as leaf area index were extended to evaluate prediction equations on independent data sets. The results provide additional evidence that measures of vegetative growth, including leaf area index, biomass, and percent cover can be predicted from multispectral measurements. In the second investigation, additional analyses of the source of variation in the spectral response of spring wheat canopies were completed. In this experiment, planting date and soil moisture were the primary factors which affected the reflectance of canopies from tillering to maturity. Differences in reflectance were primarily caused by differences in leaf area index, biomass, percent soil cover, and maturity stage. Additional analyses and modeling of the spectral response of spring wheat canopies as a function of maturity stage, solar zenith and azimuth angles, sensor view angle and direction, and wavelength are being conducted and will be reported at a later date.

Preliminary results from two investigations with corn and soybean data acquired in 1978 at the Purdue Agronomy Farm have been obtained and are briefly summarized in the following sections. Additional analyses of these data are currently being performed and collection of a second year's data on the same experiments has recently been completed.

2.1 Soybean Canopy Reflectance Characteristics

The objectives of the first experiment were to determine the relationship of reflectance and agronomic characteristics of soybean canopies throughout the growing season, including (1) examination of the effects of cultural practices on reflectance and (2) prediction of canopy characteristics from reflectance measurements.

Reflectance spectra and agronomic measurements were acquired for 81 soybean plots throughout the 1978 growing season. The factors investigated were three row spacings (15, 46 and 91 cm), three plant populations (185, 259, and 334 thousand plants per hectare) and three cultivars (Amsoy, Wells, and Elf). At intervals of 5 to 10 days, spectral reflectance was measured in four Landsat MSS wavelength bands using an Exotech 100 radiometer. Agronomic measurements included plant height, biomass, leaf area index, percent soil cover, maturity stage and surface soil moisture.

In analyzing the data it was found that plant population had no significant effect on reflectance, height, biomass, leaf area index or percent soil cover.

There were no significant differences among the three cultivars except near the end of the growing season when differences in maturity stage caused significant changes in reflectance. Row width caused differences in both red and infrared reflectance until the wide rows reached approximately 100 percent soil cover (Figure 1A-4).

High correlations were found between reflectance in both red and near infrared wavelength bands (Figure 1A-5) to percent soil cover. These relationships were used to develop a simple regression model to predict percent soil cover (Figure 1A-6). Similar analyses are being conducted for other canopy variables.

In summary, (1) no significant spectral or agronomic differences due to population were found, (2) the cultivars were spectrally and agronomically similar until near maturity, (3) row width which caused different proporations of soil and vegetation in the scene, was a dominant influence on reflectance throughout most of the season, (4) spectral reflectance was highly correlated with soil cover, and (5) prior to 100 percent soil cover, sun angle and row width interact to influence spectral reflectance.

2.2 Corn Canopy Reflectance Characteristics

In a second experiment, the relationship of the agronomic characteristics, particularly the effects of nitrogen fertilization, growth stage and leaf area and biomass to multispectral reflectance were investigated. Nitrogen deficiency has been shown to alter the single leaf reflectance characteristics of corn and other species, but few studies have been conducted with field-grown canopies.

In this experiment reflectance spectra and agronomic measurements were acquired throughout the growing season of 12 plots of corn. The experimental treatments were 0, 67, 134, and 202 kg per hectare of nitrogen fertilizer in a randomized complete block design with three replications. Bidirectional reflectance factor of the canopies was measured at approximately weekly intervals over the 0.4 to 2.4 μm wavelength range using the Exotech 20C spectroradiometer. Agronomic

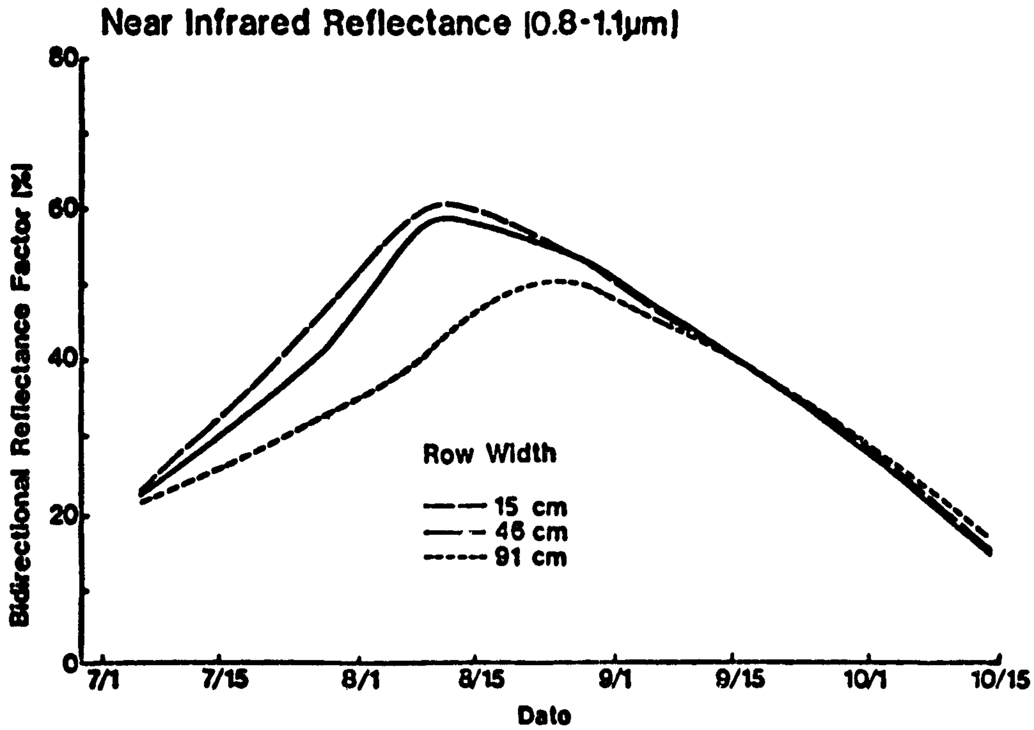


Figure 1A-4. Effect of row width on soybean canopy reflectance as a function of measurement date.

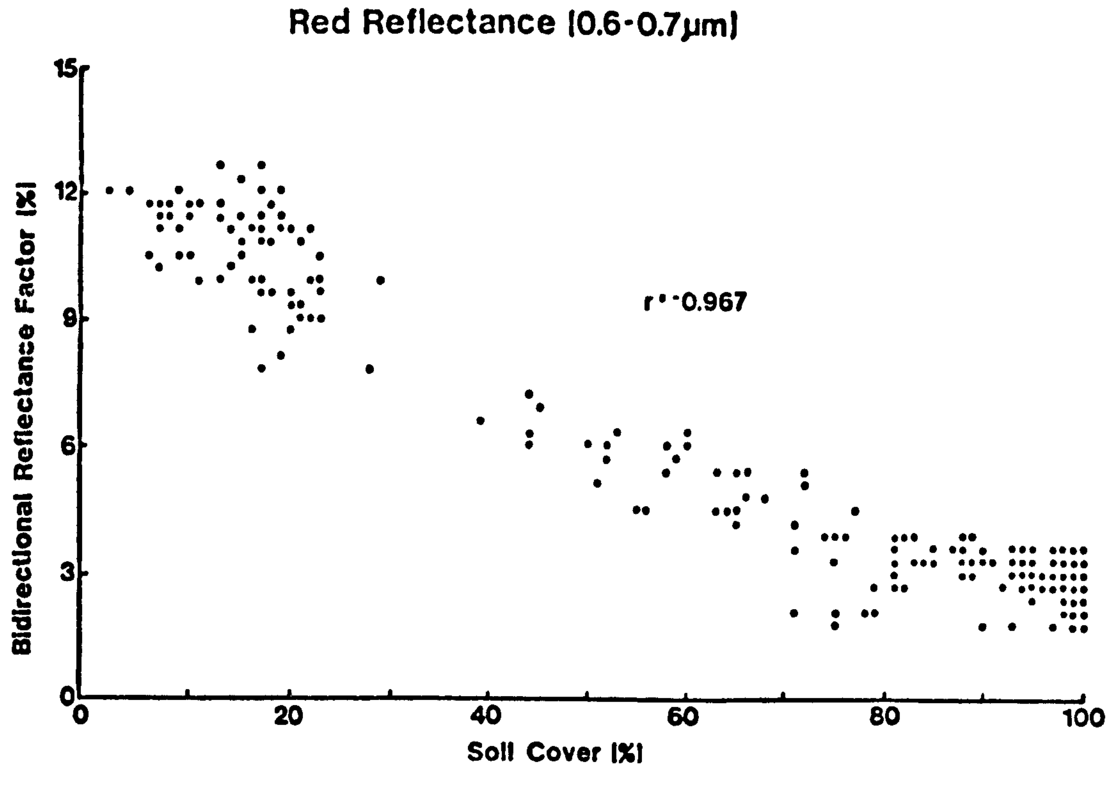


Figure 1A-5. Correlation of red reflectance with percent soil cover of soybean canopies.

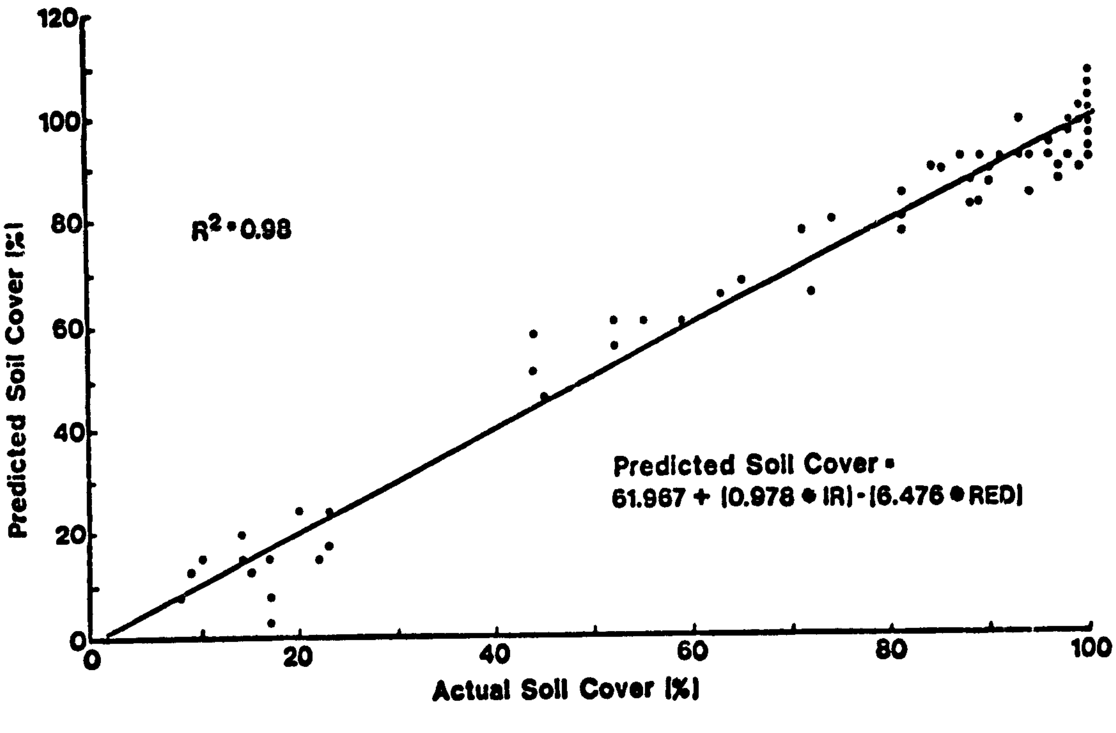


Figure 1A-6. Comparison of actual measurements of percent soil cover of soybean canopies with that predicted from reflectance measurements.

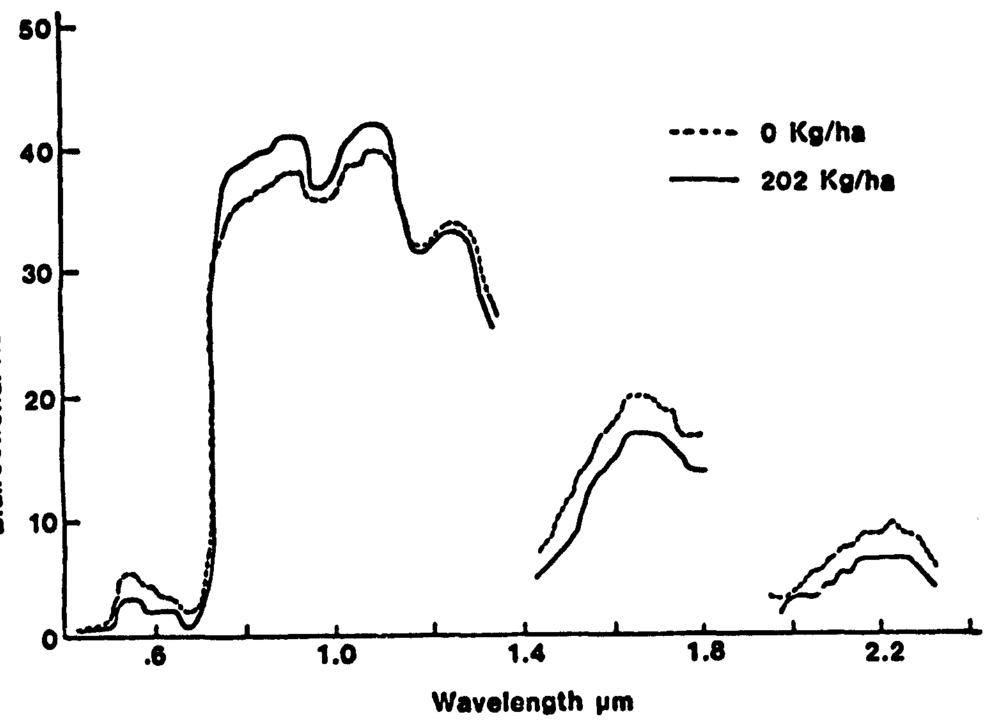


Figure 1A-7. Effect of nitrogen fertilization level on the spectral reflectance of corn canopies.

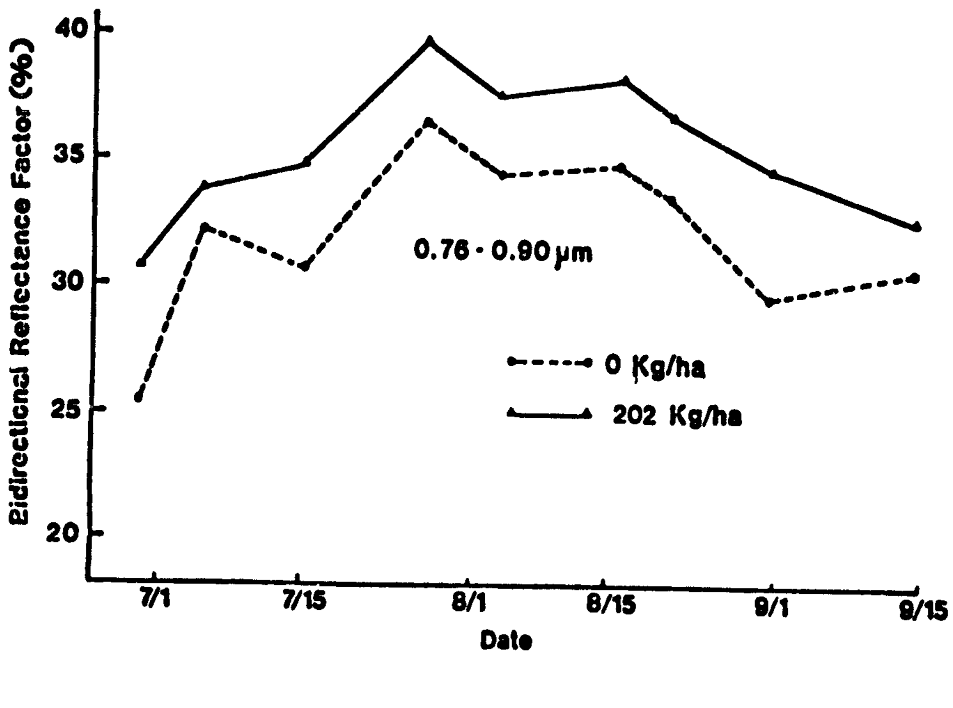


Figure 1A-8. Effect of nitrogen fertilization on the near infrared reflectance of corn canopies as a function of measurement date.

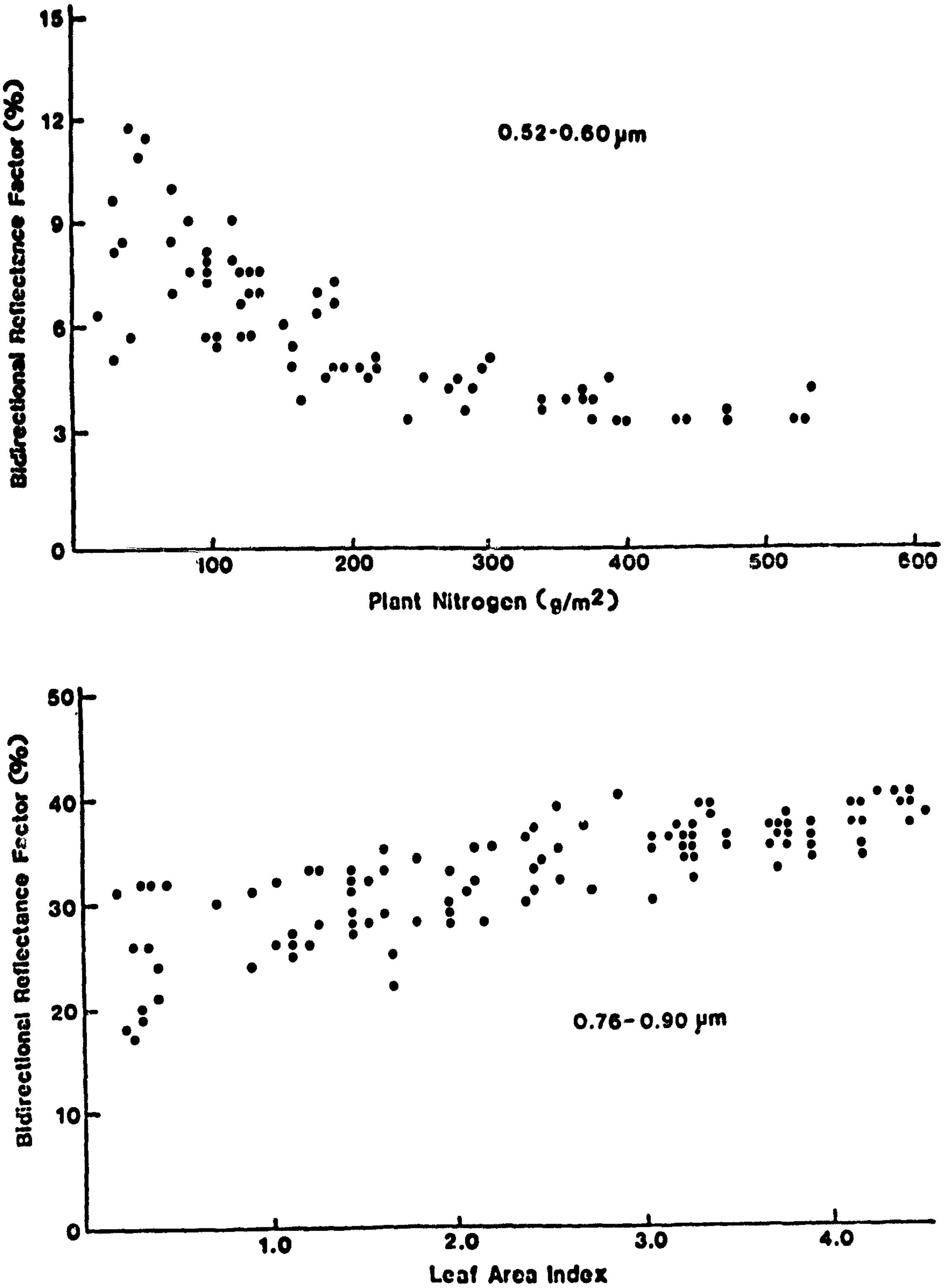


Figure lA-9. Relationship of plant nitrogen content and leaf area index to reflectance of corn canopies.

measurements of the canopies included growth stage, leaf area index, percent soil cover, biomass, and nitrogen concentration.

Nitrogen fertilization caused significant changes in spectral response (Figures 1A-7 and 1A-8) due to changes in several measures of the amount of vegetation, i.e., less nitrogen, less biomass, leaf area, and ground cover. The agronomic characteristics of the canopies were in turn manifested in the spectral response of the canopies (Figure 1A-9). Reflectances in the green and red wavelength regions were mostly strongly related to leaf and plant nitrogen concentrations. Leaf area index, biomass and percent soil cover were most strongly related to near infrared reflectance. The results indicate the potential of using multispectral remote sensing to monitor crop condition, but considerable more research is required to develop a complete and quantitative understanding of the effects of crop stresses, such as nitrogen deficiency, or the spectral response of crops.

1B. FIELD RESEARCH DATA ACQUISITION AND PREPROCESSING

Larry L. Biehl and Craig S.T. Daughtry

The objectives of the task were to acquire and preprocess the data required to accomplish the objectives of the Supporting Field Research project as described in section A. The overall objectives of the task are: (1) develop data acquisition and preprocessing plans, (2) acquire the 1979 field research data as defined in the plans, and (3) preprocess the field research data for use by researchers. The data were acquired at the Purdue University Agronomy Farm and at test sites in Webster County, Iowa, Hand County, South Dakota, and MacPherson County, Nebraska.

1. Purdue Agronomy Farm Experiments

During 1978 experiments to investigate the spectral characteristics of corn and soybean crops were initiated at the Purdue Agronomy Farm. These experiments emphasized determination of the effects on the spectral responses of maturity stage, canopy variables such as leaf area index and biomass, stresses such as moisture and nutrient deficiencies, and cultural practices such as row spacing. These experiments, with some modification, and additions were continued to obtain additional years of data (1978 was an unusually good year with respect to weather although planting was late).

The following overall objectives were selected for the experiments to be conducted at the Purdue Agronomy Farm:

- To determine the reflectance and radiant temperature characteristics of corn and soybeans as a function of maturity stage and amount of vegetation present.
- To determine the effects of stresses including moisture deficits, nutrient deficiencies and disease on the reflectance and radiant temperature properties of corn, soybeans, and winter wheat.
- To determine the effect of important agronomic practices (e.g., planting date, plant population, fertilization) and environmental factors on the spectral characteristics of corn and soybeans.
- To support the development of corn and soybean yield models which use as an input spectral response as a function of crop development stage under stressed and normal growing conditions.

- To assess, using present and future Landsat spectral bands, the spectral separability of corn, soybeans, and other typical Corn Belt crops and cover types as a function of date and maturity stage and soil background conditions (color, texture, moisture, tillage).

- To determine the effects of measurements conditions such as sensor altitude and solar elevation and azimuth angles on crop reflectance.

Ten experiments were developed for the Purdue Agronomy Farm to accomplish the objectives of Supporting Field Research. The experiments included studies of crop stress, cultural practices, instrument observation parameters and canopy geometric characteristics. A summary of the experiments, treatments, and spectral instrument systems are given in Table 1B-1.

The spectral measurements of the experiments were made by either the Exotech 20C spectroradiometer system or the Exotech 100 radiometer system. Both systems also include Barnes PRT-5 sensors and 35 mm cameras, sighted to view the same area as the spectrometers. Spectral measurements also included vertical and oblique radiant temperatures with a Barnes PRT-5 instrument. The spectral measurements for the experiments are summarized in Table 1B-2. A special laboratory spectrometer was used in a pilot test to collect leaf reflectance measurements from the corn nitrogen experiment.

To obtain data which can be readily compared, the two instruments systems are operated following similar procedures. The instruments are operated from aerial towers at three to six meters above the target at heights which minimizes the shadowing of skylight and yet ensures that the field of view of the instrument includes only the desired subject. Care is taken to avoid scene shadowing and minimize the reflective interaction due to personnel or vehicles. The routine data taking mode of the instruments is straight down, for determination of bidirectional reflectance factor. Measurements of the $BaSO_4$ painted reference panel are made at 15 minute intervals. Two measurements of each plot are typically made by moving the sensor so that a new scene within the plot fills the field of view.

Table 1B-1. Summary of the 1979 Supporting Field Research Experiments at the Purdue Agronomy Farm.

Experiment, Treatments, and Primary Sensor System

Winter Wheat: Nitrogen Fertilization and Disease (Exotech 20C and Exotech 100)

3 Cultivars
3 Nitrogen Fertilizer Rates (0, 60, and 120 kg/ha)

Corn: Cultural Practices (Exotech 100)

3 Planting Dates (May 2, 16, and 30)
3 Plant Populations (25, 50, and 75 thousand plants/ha)
2 Soil Types (Chalmers-dark and Fincastle-light)
2 Replications

Soybeans: Cultural Practices (Exotech 100)

3 Planting Dates (May 10, 24, and June 7)
2 Cultivars (Amsoy-narrow, group II maturity and Williams-bushy, group III maturity)
2 Soil Types (Chalmers-dark and Fincastle-light)
2 Replications

Corn: Nitrogen Fertilization (Exotech 20C)

4 Nitrogen Fertilizer Rates (0, 67, 134, 202 kg/ha)
3 Replications

Corn: Disease Southern Corn Leaf Blight(Exotech 20C)

3 Leaf Blight Treatments (None-resistant, early and late infection)
2 Hybrids (Pioneer 3545 and DeKalb XL43)
2 Replications

Corn and Soybeans: Moisture Stress (Exotech 20C)

3 Moisture Levels

Corn: Soil Background (Exotech 100)

2 Surface Moisture Levels (moist and dry)
2 Surface Tillage Conditions (rough and smooth)
2 Replications

Soybeans: Row Direction and Solar Azimuth and Zenith Angles (Exotech 100)

9 Row Directions(90,105,120,135,150,165,180,210,240 degrees planted in 76 cm rows, plus narrow rows (25 cm) and bare soil)

Table 1B-1. Con't

Experiments, Treatments and Primary Sensor System
Soybeans and Corn: Sensor Altitude and Field of View (Exotech 100)
10 Sensor Heights from 0.5 to 10 m above canopy
Wheat, Corn and Soybeans: Geometric Characterization
2 Crops (corn and soybeans) 3 Methods (laser, point quadrat and photographic)

Table 1B-2. Summary of spectral measurements collected at the Purdue Agronomy Farm for the 1979 field research experiments.

Experiment, Instrument System, and Spectral Measurements
Stress Experiments (nutrition, moisture, disease)
Exotech 20C Field Spectrometer System
Bidirectional reflectance factor (0.4-2.4 μm)
Radiant temperature
Color photographs
Cultural Practices Experiments (Corn, Soybeans)
Exotech 100 Field Radiometer System
Bidirectional reflectance factor (Landsat MSS spectral bands)
Radiant temperature
Color photographs
Instrument Parameter Experiments
Exotech 100 Field Radiometer System
Bidirectional reflectance factor (Landsat MSS spectral bands).
Color photographs

Table 1B-3. Summary of agronomic measurements collected at the Purdue Agronomy Farm for the 1979 field research experiments.

Agronomic Measurements
Crop Development Stage
Amount of Vegetation
Plant Height
Percent soil cover
Number of plants per square meter
Number of leaves per plant
Leaf area index
Total fresh and dry biomass (g/m^2)
Dry biomass of leaves, stems, and heads, ears or pods (g/m^2)
Crop Condition
Percent leaves green, yellow, and brown
Plant water content (g/m^2)
Presence and severity of stress
Soil Background Condition
Percent moisture
Munsell color
Roughness
Additional Data for Specific Experiments
Leaf nitrogen and chlorophyll concentrations (wheat and corn nitrogeu fertilizer experiments)
Leaf water potential (moisture stress experiments)
Leaf blight infection levels (corn blight experiment)
Grain Yield

Data recorded at the time of each measurement describing the measurement parameters include: date, time, reference illumination, air temperature, barometric pressure, relative humidity, wind speed and direction, percent cloud cover and type, field of view, latitude, longitude, and zenith and azimuth view angles.

Detailed agronomic measurements of the crop canopies included crop development stage, vegetation measurements, crop condition, soil background condition, grain yield, and additional measurements for specific experiments. The agronomic measurements are summarized in Table 1B-3. The agronomic and spectral data are supplemented by vertical and horizontal color photographic of each plot.

Augmenting the spectral and agronomic measurements were meteorological data. The meteorological data included air temperature, barometric pressure, relative humidity, wind speed, and wind direction. A record of the irradiance was collected by a total incidence pyranometer on strip charts. Additional environmental data including precipitation, pan evaporation, dew point, solar radiation, and net radiation were acquired hourly by a computerized agricultural weather station located on the Agronomy Farm.

Spectral measurements, along with agronomic and meteorological data, were acquired on each day that weather conditions permitted. During 1979 over 5000 spectra of winter wheat, corn, and soybeans were acquired. Crop maturity stages from seedling to senescence for 1979 are represented on these data.

Preprocessing of the 1978 Exotech 20C spectrometer data and the Exotech 100 Landsat band radiometer data collected at the Purdue Agronomy farm were completed during this year. Preprocessing of the 1978 FSS data collected at the Hand County, South Dakota, intensive test site were also completed.

A major portion of the data collected at the Purdue Agronomy Farm during 1979 has been completed and is available for analysis. All the Exotech 100 data from May through July have been processed. Additional agronomic data will be added to the identification records as it becomes available.

2. Data Acquisition and Preprocessing at Other Test Sites

Field research test sites utilized in 1979 in addition to the Purdue Agronomy Farm included Hand County, South Dakota, Webster County, Iowa and the University of Nebraska Agriculture Research Station in MacPherson County, Nebraska. The Iowa and Nebraska test sites were added during 1979 to expand the corn and soybean test sites. The major crops at the Hand County test site are small grains (spring and winter wheat).

The test sites in South Dakota and Iowa represented commercial fields. The major spectral systems were the NASA/JSC helicopter-mounted spectrometer (FSS) and aircraft multispectral scanner systems. The Nebraska test site included controlled plots of corn moisture stress and irrigation experiments. The NASA/JSC aircraft multispectral scanner was the major spectral system used in Nebraska.

1C. DEVELOPMENT OF MULTIBAND RADIOMETER SYSTEM

Barrett F. Robinson

1. Introduction

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

The objective of this task is to develop a multiband radiometer system for agricultural remote sensing field research. The radiometric instrument will be a multiband radiometer with 8 bands between 0.4 and 12.5 micrometers; the data acquisition system will record data from the multiband radiometer, a precision radiation thermometer, and ancillary sources. The radiometer and data handling systems will be adaptable to helicopter, truck, or tripod platforms. The system will be suitable for portable hand-held operation. The general characteristics of the system are that it will be: (i) comparatively inexpensive to acquire, maintain, and operate; (ii) simple to operate and calibrate; (iii) complete with the data handling hardware and software and (iv) well-documented for use by researchers.

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

This report describes the design specifications of the multiband radiometer and data recording modules, preparation of system and user's manuals, construction of a truck-mounted boom, and development of data handling software.

2. Description of the Multiband Radiometer

The prototype unit will be equipped with a standard set of spectral bands which match, as nearly as is practical, the seven bands of the Thematic Mapper multispectral scanner. Filters will be durable and suitable for use under field conditions of temperature and humidity. A summary of the spectral bands is shown in Table 1C-1 and Figure 1C-2. Examination of Figure 1C-3 will show that, while the four Landsat MSS bands sample the vegetation spectrum coarsely and over a limited range, the eight bands provide complete and rather detailed coverage of the spectrum.

The multiband radiometer will simultaneously produce analog voltages which are proportional to scene radiance in each of eight spectral bands. The radiometer will be a stand-alone device capable of operation with a variety of data recording systems. The prototype radiometer will be capable of operation from 0° to 60°C, when mounted on a tripod, truck, boom, helicopter, or small plane.

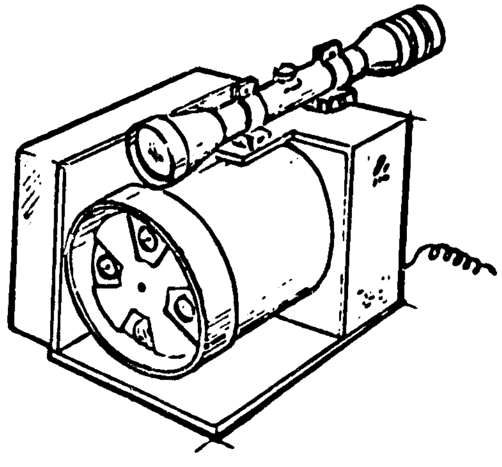


Figure 1C-1. Sketch of multiband radiometer.

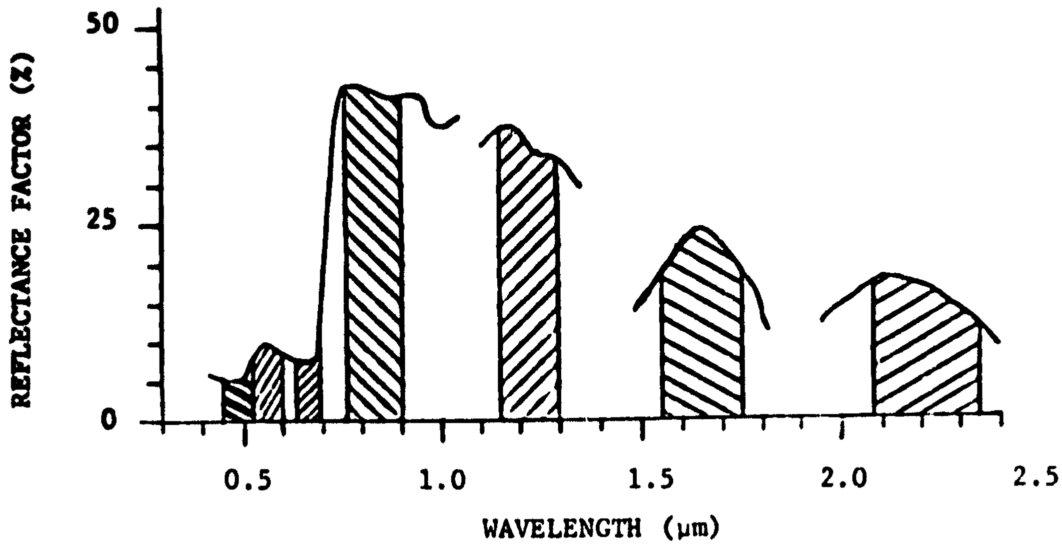


Figure 1C-2. Spectral distribution of passbands superimposed on a typical vegetation spectrum.

Table 1C-1. Spectral band specifications.

Band	50% Response Wavelengths (μm)	Detector	L* $W \cdot m^{-2} \cdot sr^{-1}$
1	0.45 - 0.52	Silicon	31
2	0.52 - 0.60	Silicon	27
3	0.63 - 0.69	Silicon	25
4	0.76 - 0.90	Silicon	45
5	1.55 - 1.75	PbS	16
6	2.08 - 2.35	PbS	6
7	10.40 - 12.50	$LiTaO_3$	8-32
8	1.15 - 1.30	PbS	21

Other features of the radiometer include: variable field of view, adjustable dynamic range, modular configuration, control panel, gain status signal, camera boresight and control, low battery warning signal, system temperature signals, electronic filtering, and 12 volt battery operation.

The multiband radiometer is under construction by Barnes Engineering, Stamford, CT. Delivery is anticipated by July 15, 1980.

3. Description of the Data Recording Module

The Data Recording Module is presently under construction by Purdue/LARS and completion is expected by July 1981. The Data Recording Module (DRM) will digitize, format, and store analog data in a solid state memory. The DRM will accept analog signals from the multiband radiometer and other sources as appropriate to the measurement situation. It will operate under the same environmental conditions as the multiband radiometer.

The main function of the DRM is to record the data from the multiband radiometer and other data channels within 2.5 milliseconds - corresponding to 15 cm at 61 meters per second. Additionally, the unit will provide a suitable interface for a printing calculator to allow on-site evaluation of system performance and to provide a means for analysis of limited quantities of data (typically, a H.P. 97S will require about 30 seconds to process a single observation (all channels) to the desired final form and print the results. The principal transfer mechanism will be 16 bit parallel with handshake which is well suited for entry to many micro-processors and computers. A parallel to serial conversion may be required for some systems but can be easily accomplished external to the DRM.

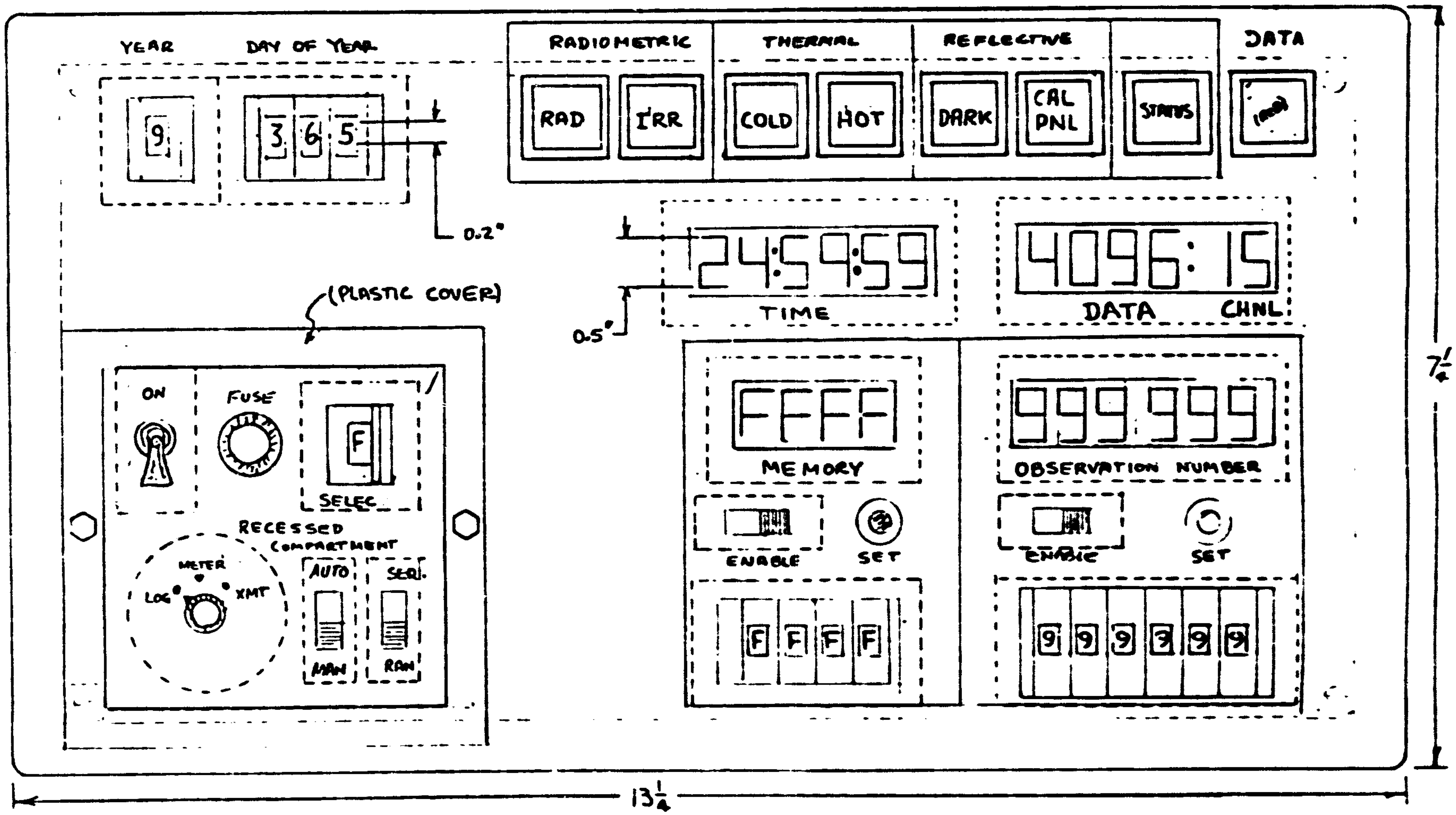


Figure 1C-3. Sketch of front panel of data recording module.

4. Construction of Truck-Mounted Boom

As part of the task, a pick-up truck-mounted boom was constructed. A discussion of the trade-offs in the mechanical design, appropriate stress computations, and a description of the boom are documented in LARS Technical Report 090879, Design and Evaluation of a Pick-up Truck-Mounted Boom for Elevation of a Multiband Radiometer System, by Roy Tsuchida. The specifications for the boom design incorporate several years of field experience by LARS with similar structures. Except for discussion of performance tests, which will be conducted when construction of the boom is completed, this document is complete and has been supplied to the Technical Monitor. The boom assembly and its operation are shown in Figures 1C-4 and 1C-5.

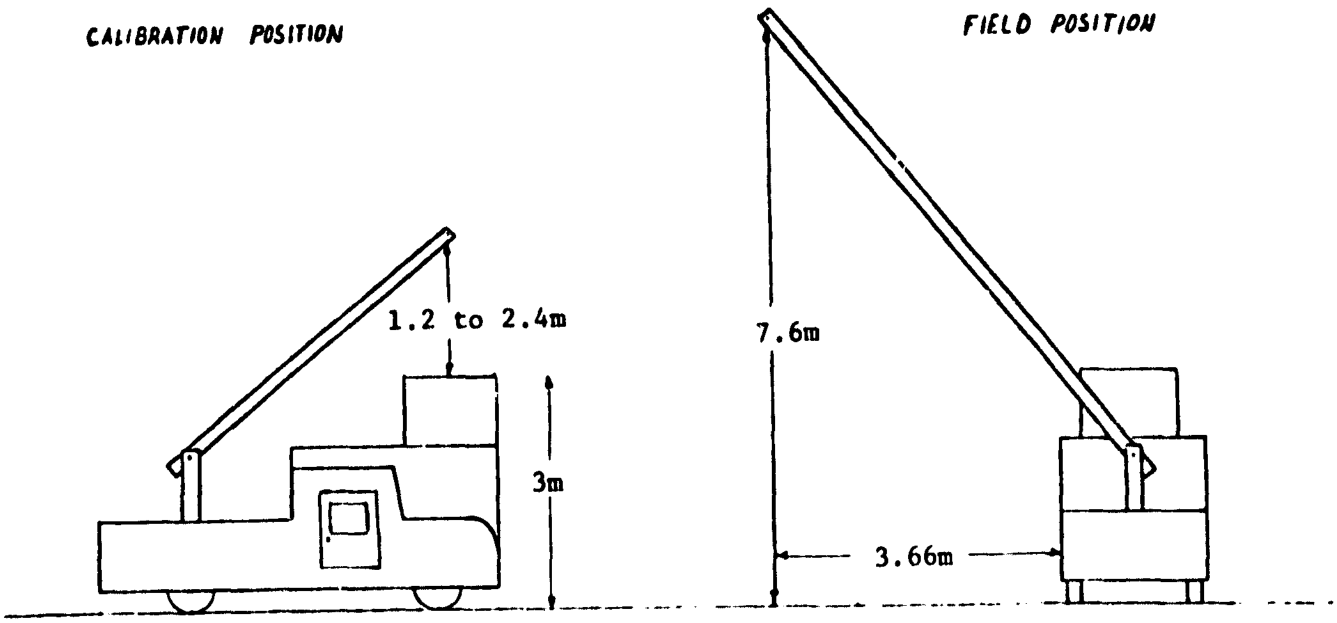


Figure 1C-4. Positioning of the pick-up truck-mounted boom.

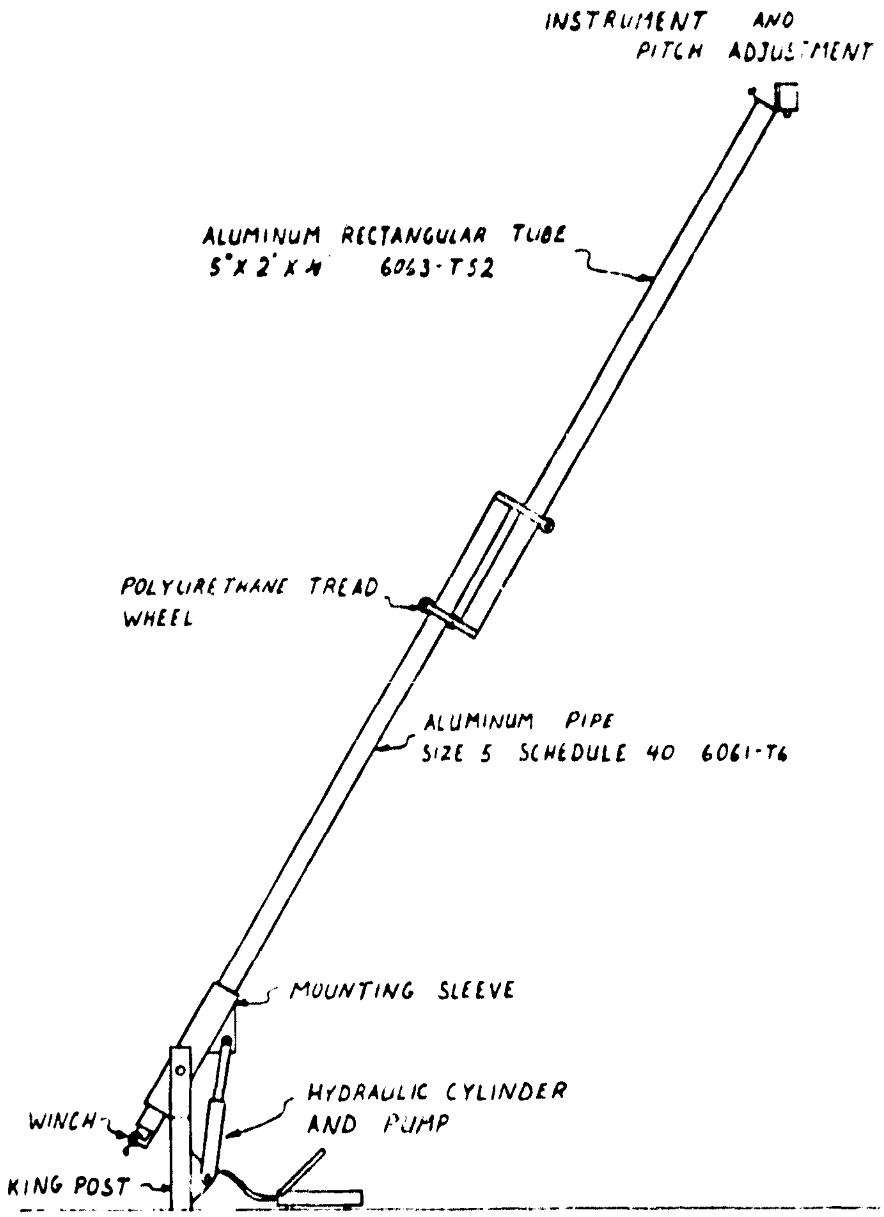


Figure 1C-5. Sketch of pick-up truck-mounted boom.

5. System and User Manuals

System manuals are being prepared for the multiband radiometer and data recording modules. Each manual includes the following topics:

- General Information (description, specifications, etc.)
- Installation
- Operation
- Theory of Operation
- Maintenance and Repair
- Testing Procedures

In addition, a users manual is being prepared to provide information to researchers and system users on the following topics:

- Fundamentals of Measurement and Calibration
- Field Measurement and Calibration Procedures
- Performance Evaluation Tests
- Experimental Design (including a sample experiment)

6. Development of Data Handling Software

A software system was designed and implemented to aid in the handling and preprocessing of multiband radiometer data along with associated agronomic, meteorological, and other ancillary data. The software system was initially built around an Exotech 100 Landsat band radiometer system and later made more general for any multiband radiometer system having one to twelve channels.

The software system presently uses 80 column computer cards for radiometer data input. Presently for the Exotech 100 system, the data is recorded on printed paper tape, copied to 80 column record sheets, and keypunched. The program is designed, however, to accept data input directly from the Data Recording Module. The software system calibrates the data according to one of several algorithms selected by the researcher.

7. Interface Hardware and Software

In preparation for entry of data from the Data Recording Module, a DR-11C Bit Parallel I/O Interface and a DD11-DF Back Plane were installed on the Purdue/LARS PDP 11/34A minicomputer. The installation is being brought on line and should be completed early in December 1979. A program for testing the interface has been prepared and additional interface software is on order.

Entered data will be quickly stored in core memory until the entry is complete. Data will then be stored on a disk awaiting transfer from the PDP 11/34A to the IBM 3031 for calibration and re-formatting to LARSPEC compatible format.

1D. ESTABLISHMENT OF SOILS DATA BASE

Eric R. Stoner and Marion F. Baumgardner

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

Modern soil classification systems emphasize the importance of information about the quantitative composition of soils. In order to differentiate among soil groups, it is necessary to rely on laboratory measurements of selected soil properties. Physical and chemical determinations of most soil properties follow well established procedures of laboratory analyses. Certain of these soil properties are selected as diagnostic criteria in the soil classification process, based on their importance in understanding the genesis of the soil. By a procedure of empirical correlation, critical limits between sets of soils are established, designed to reflect the influence of the soil forming factors of climate, parent material, relief, biological activity, and time.

Quantitative measurements of soil spectral properties have become available as a diagnostic tool for the soil scientist with the advent of such instruments as the Exotech Model 20C spectroradiometer. However, the climatic and genetic effects on the relationships between measured spectral properties and specific chemical, physical, and biological properties of the soil are not well understood. Whereas soil color is used as diagnostic criterion in the U.S. Soil Taxonomy, the determination of soil color by comparison with a color chart continues to be a rather nonquantitative and subjective procedure. Spectral characterization of soil "color" by means of quantitative spectroradiometric measurements may add to the precision with which soils can be differentiated. With this increased precision of soil spectral characterization, the relationships with the more important diagnostic soil characteristics or qualities that are not so easily and accurately observed may be better understood.

A study was, therefore, initiated in 1977 to develop a data base of soil reflectance and physical-chemical properties and investigate the relationships of reflectance and important physical-chemical properties of soils. During the current year the data base and initial statistical analysis of the data have been completed and are summarized in this report.

1. Objectives

The objective was to define quantitatively the relationships between soil reflectance and physicochemical properties of soils of significance to agriculture. Selection of soil samples with a wide range of important soil characteristics by statistical stratification of continental United States climatic zones permits the evaluation of climatic and genetic effects on the relationships between multispectral reflectance and these soil properties. A further objective was to define the relationships sufficiently to design further research to quantify the contributions which different soil components make to the multispectral characteristics of specific soils. The ultimate objective of this research approach is to provide a body of knowledge in the form of a soils data base which will render remote multispectral sensing a valuable tool for mapping soils, determining land use capabilities and soil productivity ratings, identifying crops and predicting crop yields.

2. Information Approach

Approximately 250 soils, representing a statistical sampling of the more than 10,000 soil series in the United States were selected for this investigation. Stratification of soil sampling was based on series type location within climatic zones. Climatic strata included the frigid, mesic, thermic, and hyperthermic soil temperature regimes as defined by the U.S. Soil Taxonomy as well as the perhumid, humid, subhumid, semiarid, and arid moisture regions as identified by Thornthwaite's 1948 Moisture Index. The 500 soil samples from 39 states were acquired by the USDA/Soil Conservation Service.

Soil reflectance was measured using an Exotech Model 20C spectroradiometer adapted for indoor use with a reflectometer equipped with an artificial illumination source, transfer optics, and sample stage. Spectral readings were taken in 0.01 μm increments over the 0.52-2.32 μm wavelength range.

An identification record containing 100 items of information including complete soil taxonomic classification along with site characteristics and laboratory analyses in computer tape format was prepared for all of the soils in this study. This information together with digitized soil reflectance data is accessible for editing and rapid retrieval of all soils information by means of the LARSPEC software package.

3. Soil Atlas

An abbreviated format was chosen for presentation of selected soil properties in an atlas of soil reflectance properties to be published as a LARS Technical Report and as a Purdue University Agriculture Experiment Station research bulletin. Soils are arranged in this atlas by alphabetical order of the 39 states from which they were sampled. Spectra of four soils are displayed on each page, along with 16 items of information describing the location, soil classification, and physical-chemical characteristics of each sample (Figure 1D-1. Examples of various soils with widely-differing chemical, physical and spectral properties are illustrated in Figure 1D-2.

Wavelength, expressed in micrometer (μm) units, denotes the portion of the electromagnetic spectrum under consideration. Wavelength regions frequently referred to are the visible (0.38-0.72 μm), near infrared (0.72-1.3 μm), and middle infrared (1.3-3.0 μm).

4. Numerical Analysis

For the purpose of statistical correlation all soil spectral curves were represented by ten spectral bands.

1) Sample location
2) Soil subgroup name
3) Soil family modifiers
4) Moisture zone
5) Parent material
6) County
7) Horizon
8) Slope class
9) Internal drainage class
10) Texture class
11) Percent sand, silt and clay
12) Munsell color
13) Organic matter content
14) Cation exchange capacity
15) Iron oxide content
16) Moisture content

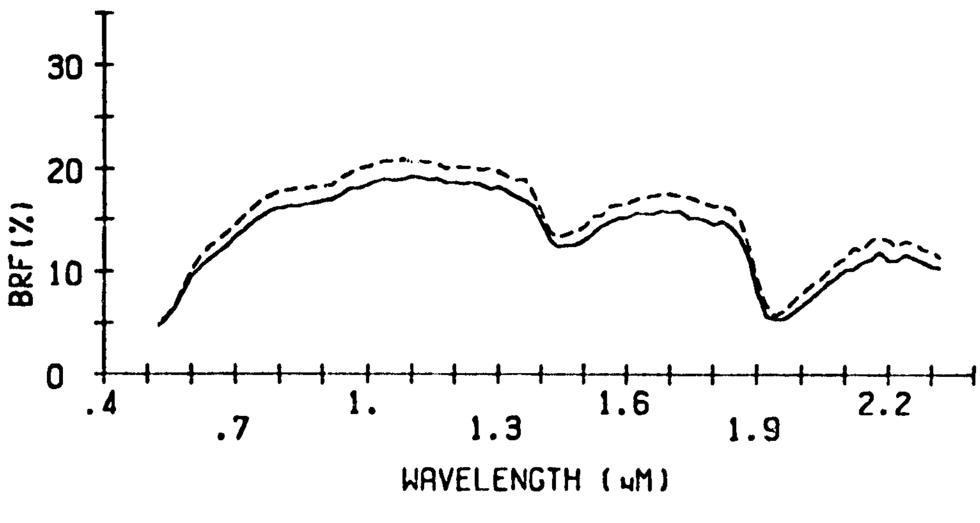
17) 

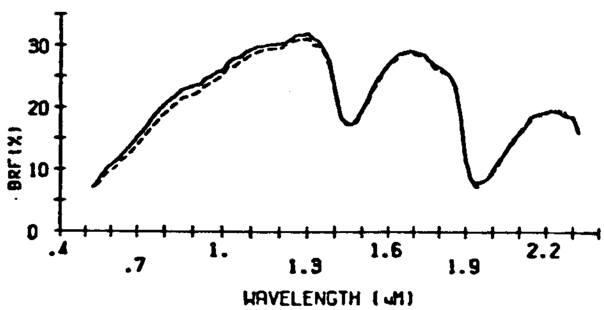
Figure 1D-1. Numbered guide corresponding to narrative key to soil information

PRATT(KS)

Psammentic Haplustalf
sandy, mixed, thermic
subhumid zone
sandy eolian deposits
Pratt Co.

——	----
Ap horizon	Ap horizon
B slope	B slope
well drained	well drained
fine sandy loam	fine sandy loam
73%S 24%Si 3%C	61%S 37%Si 2%C
10YR 3/3 (moist)	10YR 4/3 (moist)
7.5YR 6/4 (dry)	7.5YR 6/2 (dry)
0.55% O.M.	0.44% O.M.
2.8 meq/100g CEC	1.9 meq/100g CEC
0.31% Fe_2O_3	0.25% Fe_2O_3

11.0 MW%• —— 13.4 MW%• ----

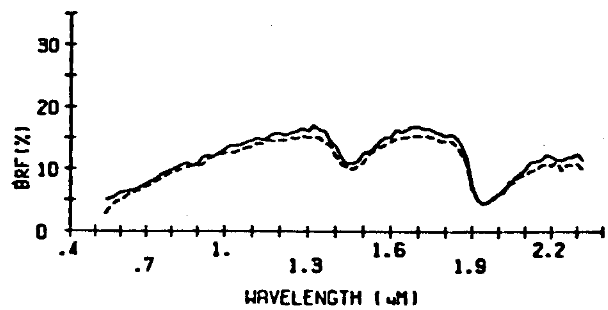


RICHFIELD(KS)

Aridic Argiustoll
fine, montmorillonitic mesic
semiarid zone
silty eolian sediments
Grant Co.

——	----
Ap horizon	Ap horizon
A slope	A slope
well drained	well drained
silt loam	silt loam
8%S 72%Si 20%C	12%S 70%Si 18%C
10YR 3/2 (moist)	10YR 3/2 (moist)
10YR 5/3 (dry)	10YR 5/2 (dry)
2.14% O.M.	1.78% O.M.
21.4 meq/100g CEC	21.3 meq/100g CEC
0.79% Fe_2O_3	0.86% Fe_2O_3

37.3 MW%• —— 35.6 MW%• ----

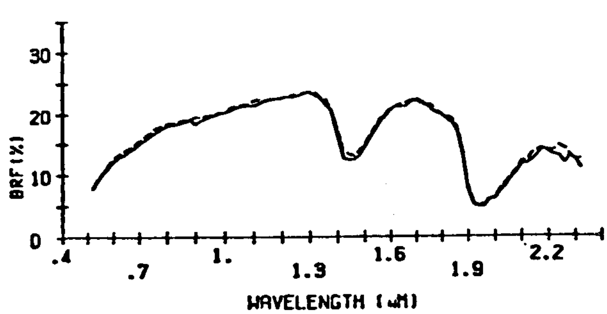


COLBY(KS)

Ustic Torriorthent
fine-silty, mixed, calcareous, mesic
semiarid zone
calcareous silty material
Hamilton Co.

——	----
Ap horizon	Ap horizon
A slope	A slope
well drained	well drained
silt loam	silt loam
22%S 54%Si 24%C	15%S 62%Si 24%C
10YR 5/3 (moist)	10YR 5/3 (moist)
10YR 6/4 (dry)	10YR 6/4 (dry)
1.24% O.M.	0.85% O.M.
30.3 meq/100g CEC	30.2 meq/100g CEC
0.69% Fe_2O_3	0.68% Fe_2O_3

37.3 MW%• —— 36.6 MW%• ----



NEWARK(KY)

Aeric Fluventic Haplaquept
fine-silty, mixed, nonacid, mesic
humid zone
mixed alluvium
Daviess Co.

——	----
Ap horizon	Ap horizon
A slope	A slope
s. poorly drained	s. poorly drained
silt loam	silt loam
25%S 57%Si 18%C	4%S 79%Si 18%C
10YR 4/2 (moist)	10YR 4/3 (moist)
10YR 6/3 (dry)	10YR 6/4 (dry)
1.83% O.M.	2.84% O.M.
15.7 meq/100g CEC	17.0 meq/100g CEC
1.05% Fe_2O_3	1.93% Fe_2O_3

29.0 MW%• —— 34.1 MW%• ----

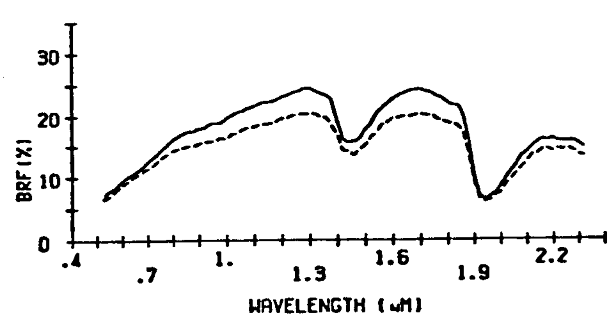


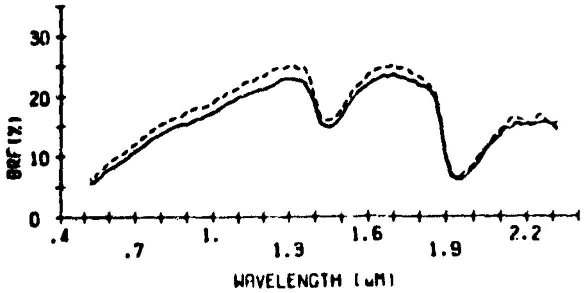
Figure 1D-2. Reflectance curves, soil test results, and site characteristics in the soil data base.

WHITLEY(KY)

Typic Hapludult
fine-silty, mixed, mesic
humid zone
part alluvium, part acid residuum
Laurel Co.

____	----
Ap horizon	Ap horizon
B slope	B slope
well drained	well drained
silt loam	silt loam
23%S 57%Si 20%C	16%S 65%Si 19%C
10YR 4/3 (moist)	10YR 4/3 (moist)
10YR 6/4 (dry)	10YR 6/4 (dry)
3.50% O.M.	2.57% O.M.
13.7 meq/100g CEC	14.2 meq/100g CEC
1.55% Fe_2O_3	2.11% Fe_2O_3

18.5 MW%• ____   35.9 MW%• ----

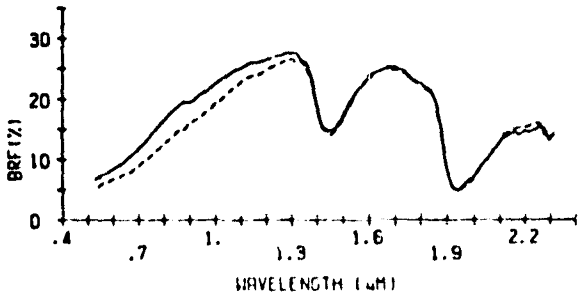


MIDLAND(LA)

Typic Ochraqualf
fine, montmorillonitic, thermic
humid zone
clayey sediments
Acadia Parish

____	----
Ap horizon	Ap horizon
A slope	A slope
poorly drained	poorly drained
silty clay loam	silty clay loam
5%S 57%Si 38%C	3%S 65%Si 32%C
10YR 4/2 (moist)	10YR 3/1 (moist)
10YR 6/3 (dry)	10YR 6/3 (dry)
2.42% O.M.	2.32% O.M.
25.1% meq/100g CEC	27.3 meq/100g CEC
0.88% Fe_2O_3	0.62% Fe_2O_3

37.7 MW%• ____   41.2 MW%• ----

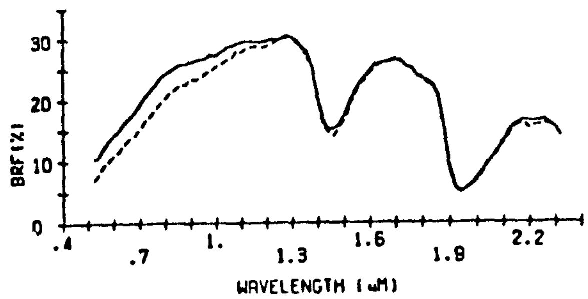


CALHOUN(LA)

Typic Glossaqualf
fine-silty, mixed, thermic
humid zone
loess
East Baton Rouge Parish

____	----
A1 horizon	A1 horizon
A slope	A slope
poorly drained	poorly drained
silt loam	silt loam
15%S 71%Si 14%C	20%S 69%Si 10%C
10YR 5/3 (moist)	10YR 5/3 (moist)
10YR 7/3 (dry)	10YR 6/4 (dry)
1.74% O.M.	2.40% O.M.
7.1 meq/100g CEC	11.4 meq/100g CEC
0.60% Fe_2O_3	0.72 Fe_2O_3

34.6 MW%• ____   33.7 MW%• ----



KENNER(LA)

Fluvaquentic Medisaprist
euic, thermic
humid zone
herbaceous plant remains with clayey
alluvium
Jefferson Parish

____	----
Oe1 horizon	Oe1 horizon
A slope	A slope
v. poorly drained	v. poorly drained
muck	muck
4%S 40%Si 56%C	3%S 31%Si 66%C
7.5YR 2/0 (moist)	10YR 2/1 (moist)
10YR 2/1 (dry)	10YR 2/1 (dry)
55.14% O.M.	54.38% O.M.
73.6 meq/100g CEC	82.1 meq/100g CEC
0.00% Fe_2O_3	0.00% Fe_2O_3

77.2 MW%• ____   73.1 MW%• ----

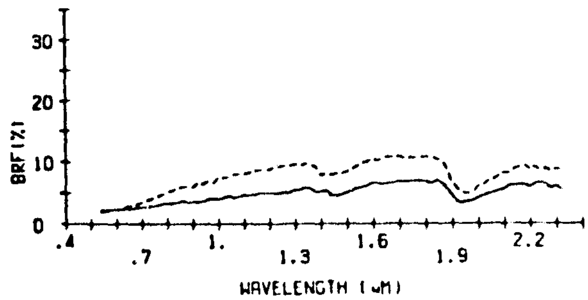


Figure 1D-2. Continued.

Soil Spectral Bands for Correlation Analysis

Band	Wavelength (μm)	Spectral Region	Band	Wavelength (μm)	Spectral Region
1	0.52-0.62	visible	6	1.02-1.12	near IR
2	0.62-0.72	visible	7	1.12-1.22	near IR
3	0.72-0.82	near IR	8	1.22-1.32	near IR
4	0.82-0.92	near IR	9	1.55-1.75	middle IR
5	0.92-1.02	near IR	10	2.08-2.32	middle IR

Reflectance in all of the spectral bands was found to be negatively correlated with the natural logarithmic transformation of organic matter content. Reflectance in the important 2.08-2.32 μm band is negatively correlated, in addition, to moisture content, cation exchange capacity, iron oxide content and clay content, while it is positively correlated with medium and fine sand contents.

A forward stepwise inclusion procedure was used to estimate the reflectance in individual wavelength bands using certain agronomic and site characteristics as predictors. The importance of climatic zone, parent material, and drainage characteristics in explaining soil reflectance was brought out in this analysis. Other soil parameters had an expected influence on reflectance, with organic matter being the single most important variable in all but band 10, where moisture content was the first variable to be included in the regression equation. Although the R^2 values were not high for this analysis of 481 soils, the inclusion of site characteristics pointed to the importance of these soil-forming factors as contributors to soil reflectance.

Prediction models for certain soil parameters using only soil reflectance data as inputs reveal high R^2 values when soils are grouped by specific climatic zone. The importance of visible, near infrared, and middle infrared reflectance data is seen repeatedly. Iron oxide has high predictive values only in the humid frigid and some arid regions. Results are highly climate specific. As in previous studies of this type, cation exchange capacity often shows higher R^2 values than other soil parameters which are known to exhibit inherent spectral behavior. Cation exchange capacity seems to be acting as a natural integrating factor for several other soil parameters such as organic matter and particle size.

Extending these results to the level of airborne remote sensors, it is likely that reflectance data from carefully selected wavelength bands could be used to extract information from bare soil areas that could be related to levels of organic matter, soil moisture, iron oxide content, particle size content, or even an indicator of potential productivity such as cation exchange capacity for certain specified climatic areas. Where prior information is available about soil drainage and parent material classes, even better correlations can be expected within more homogeneous areas of soil inference.

Based on results of statistical analyses as well as on qualitative evaluation of soil reflectance/absorption characteristics, the following wavelengths are critical for identification of soil reflectance characteristics: 0.52 to 0.62 μm (green wavelength region highly correlated with organic matter content), 0.7 μm and 0.9 μm (ferric iron absorption wavelengths), 1.0 μm (ferrous iron and hydroxyl gibbsite absorption wavelength), 1.22 to 1.32 μm and 1.55 to 1.75 μm (regions of highest reflectance for many soils, correlated with many soil properties), 2.08 to 2.32 μm (region of highest correlations with soil moisture). Although spectral bands for the Thematic Mapper sensor include 0.52 to 0.60 μm, 1.55 to 1.75 μm and 2.08 to 2.35 μm, the 0.76 to 0.90 μm near infrared wavelength band is too broad for specific iron oxide studies in soils, a fact that could limit its usefulness in erosion studies as well as soil productivity surveys.

It can be seen that soils with widely-differing physicochemical and site characteristics are no more similar in their reflectance properties than are different species of plants throughout their growth cycles. To treat soil reflectance as a constant, unchanging characteristic from location and from date to date is to ignore the well-ordered physical and chemical relationships that impart diverse spectral reflective character to soils. The soils data base developed in this study should help further an understanding of the diverse nature of soils as they would be viewed by remote sensors.

1E. FIELD RESEARCH DATA BASE MANAGEMENT AND SOFTWARE DEVELOPMENT

Larry L. Biehl

The development of the field research data library at Purdue/LARS was initiated in the fall of 1974 by NASA/Johnson Space Center (JSC) with the cooperation of the United States Department of Agriculture (USDA) as a part of the Large Area Crop Inventory Experiment (LACIE). The purpose of the data base is to provide fully annotated and calibrated multitemporal sets of spectral, agronomic, and meteorological data for agricultural remote sensing research.

Milestones achieved during the past year have been:

- Data library and distribution
    - Inclusion of 1978 crop year data
    - Inclusion of part of 1979 crop year data
    - Inclusion of 1978 soils data
    - Distribution of data to researchers
    - Update of field research data library catalog
    - Preparation of report on the evaluation of spectral data calibration procedures.
- Development and documentation of analysis software
- Development and documentation of processing software

1. Field Research Data Library and Distribution

The general organization of the field research data library is illustrated in Figure 1E-1. The data have been collected over several test sites and for different crops as illustrated in Table 1E-1. The test sites are of two types, controlled experimental plots and commercial fields. The spectrometer data are processed into comparable units, bidirectional reflectance factor, in order to make meaningful comparisons of the data acquired by the different sensors at different times and locations. The multispectral scanner data are approximately linearly related to scene radiance. The information is available for the researcher to calibrate the scanner data to in-band bidirectional reflectance factor if desired.

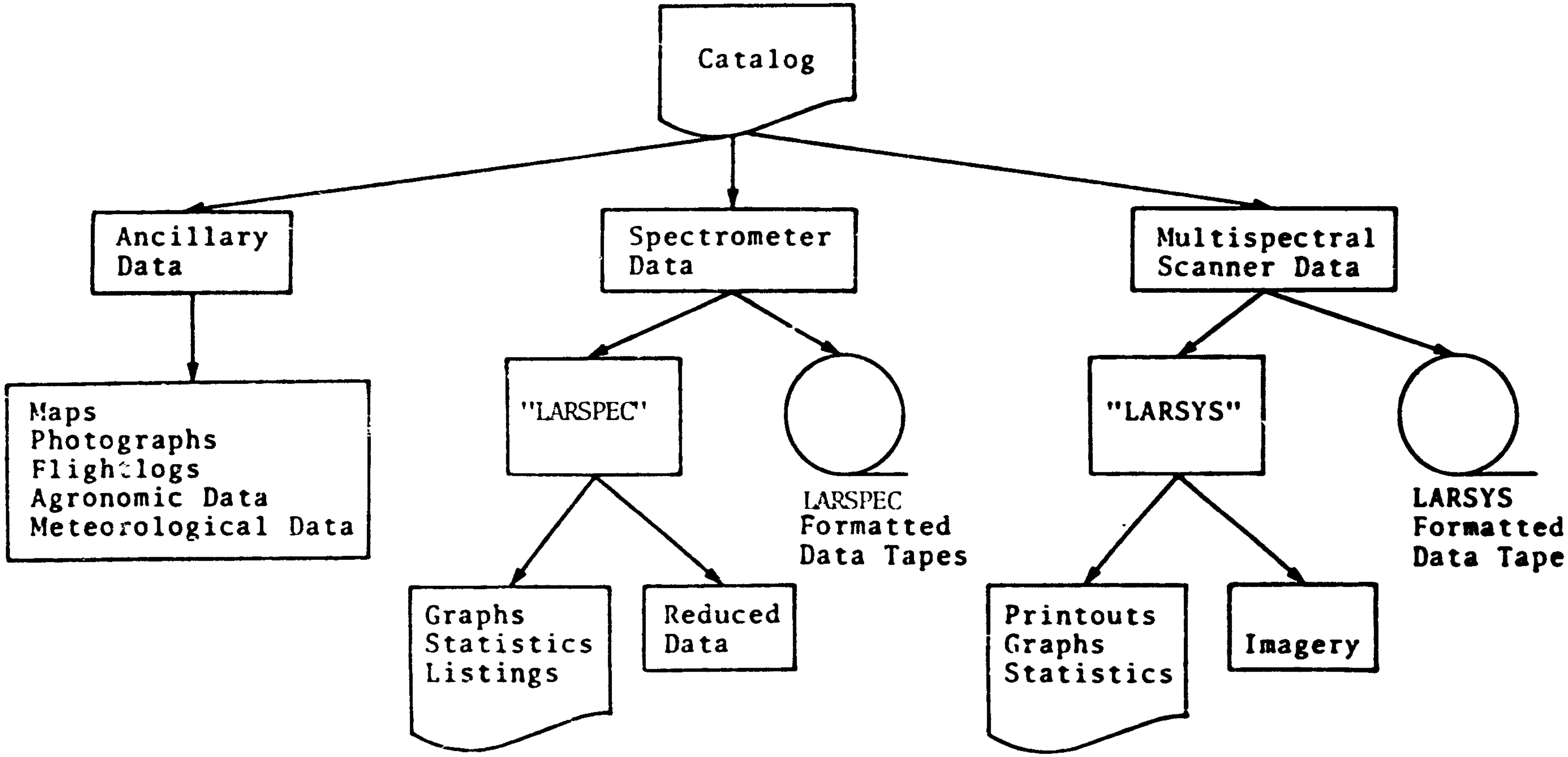


Figure IE-1. Organization of field research data library. LARSPEC and LARSYS are Purdue/LARS software systems to analyze spectrometer/radiometer and multispectral scanner data.

The Field Research Data Library Catalog summarizes the data available. The catalog includes a separate volume for each crop year during which data were collected. In the past twelve months, five aircraft scanner runs and 18,000 spectrometer/radiometer observations for the 1978 crop year have been made available to researchers. The 1978 data includes spectral observations of over 500 soil samples from 39 states of the United States as well as Brazil, Costa Rica, Sudan, Spain and Jordon along with their physical and chemical properties. Also, 4000 spectrometer/radiometer observations for the 1979 crop year have been processed and made available to researchers. A summary of the spectral data is given in Table 1E-2.

Six institutions, ERIM, NASA/JSC, NASA/GSFC, GISS, Ecosystems, Inc., and Purdue/LARS, have received or accessed field research data during the past year. Investigators at NASA/JSC, ERIM, and Purdue have direct access to the digital data via remote terminals to the LARS' computer. Table 1E-3 summarizes the data that have been distributed during the past three years and indicates which institutions received data during this past year.

2. Data Analysis Software Development and Documentation

During the field research project over the past five years, over 140,000 observations of calibrated spectrometer data and 300 flight lines of aircraft scanner data have been collected. Researchers require more powerful research tools to analyze the data efficiently and effectively. During this past year several analysis software tools were developed and/or documented. The achievements include:

Expansion and documentation of LARSPEC
Implementation of Statistical Analysis System (SAS)
Implementation of 3-dimensional graphics software

2.1 Expansion and Documentation of LARSPEC

Several additions were made to LARSPEC, the software system on the Purdue/LARS computer which accesses the spectrometer/radiometer data and associated agronomic and meteorological measurements: (1) ability to access multiband radiometer data, such as data from the Exotech 100 Landsat band radiometer, (2) increased ability to transfer agronomic data

Table 1E-1. Summary of field research test site locations and major crops.

Crop Year	Finney Co. Kansas	Williams Co. N. Dakota	Hand Co. S. Dakota	Tippecanoe Co. Indiana	Webster Co. Iowa	McPhearson Co. Nebraska
1975	Winter Wheat	Spring Wheat	-	-	-	-
1976	Winter Wheat	Spring Wheat	Spring & Winter Wheat	-	-	-
1977	Winter Wheat	Spring Wheat	Spring & Winter Wheat	-	-	-
1978	-	-	Spring & Winter Wheat	Corn Soybeans	-	-
1979	-	-	Spring & Winter Wheat	Corn Sovbeans Winter Wheat	Corn Soybeans	Corn

Table 1E-2. Summary of spectral data in the field research data library by instrument and data type.

Instrument/Data Type	Crop Years 1975-1979
Landsat MSS Frames	124
Aircraft Multispectral Scanner Dates/Flightlines	47/306
Helicopter Mounted Spectrometer Dates/Observations	74/120,000
Truck Mounted Spectrometer Dates/Observations	207/9000
Truck Mounted Multiband Radiometer Dates/Observations	48/9000

Table 1E-3. Summary of recipients of field research data.

Organization	Number of Requests
NASA/Goddard Space Flight Center Greenbelt, Maryland	>5√
General Electric Corporation Philadelphia, Pennsylvania	2
University of South Florida Tampa, Florida	1
USDA, Agriculture Research Service Weslaco, Texas	1
Goddard Institute for Space Studies New York, New York	3√
Ecosystems, Inc. Gambrills, Maryland	*√
Environmental Research Institute of Michigan Ann Arbor, Michigan	+√
NASA, Johnson Space Center Houston, Texas	+
Purdue University, Laboratory for Applications of Remote Sensing West Lafayette, Indiana	+

* Received from NASA/Goddard Space Center
+ SR&T users of field research data.
√ Recipients of field research data during 1979.

to disk for use by the SAS or SPSS statistical routines, (3) addition of crop productivity parameters to identification records. With the addition of new graphics and clustering capabilities in 1978, the decision was made to develop a new LARSPEC User's Manual containing a control card dictionary, examples of inputs and outputs, and error messages. The manual should enhance the usefullness of the LARSPEC software system for field research data analysis.

2.2 Implementation of Graphics and Statistics Software

During this past year, a set of 3-dimensional graphics software routines were implemented on the LARS' computer. The software is the 3-D Graphics Compatibility System (GCS) acquired from the U. S. Army Corps of Engineers at the Waterways Experiment Station, Vicksburg, Mississippi. The software routines, in Fortran, can be used in the development of researcher tools for more complex displays and analyses of spectral-agronomic data. The software has the capability for plotting 3-D graphics in perspective or orthographic projections for 3-D rectangular, spherical, or cylindrical coordinate systems. An example is given in Figure 1E-2. In addition, the Statistical Analysis System (SAS) was implemented to augment other statistics routines.

3. Data Processing Software Development

Two major processing software capabilities were developed and documented this year. The first is the software to process the Exotech 100 Landsat band radiometer data and the second is software to update spectrometer identification records. The Exotech 100 processing software, or more correctly, the multiband radiometer data processing software is described in section 1C.

The software to update LARSPEC identification records provides the capability to process the Exotech 20C and Exotech 100 spectral data soon after it is collected prior to having all the agronomic measurements. This capability is the reason that a large part of the 1979 Purdue Agronomy farm data is available for analysis before the end of the year.

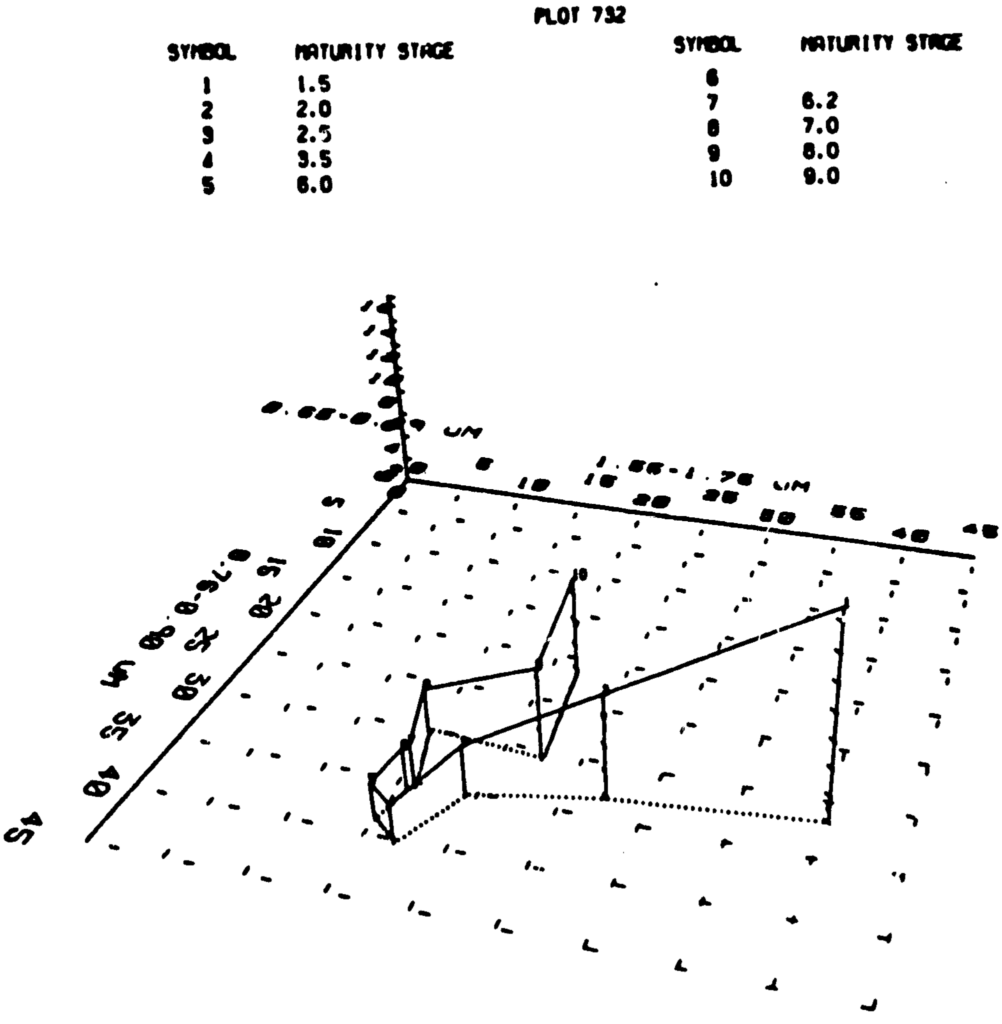


Figure 1E-2. Example of 3-dimensional line type plot using 3-D Graphics Compatibility System software. This plot is a temporal (growth stage) plot of three Thematic Mapper bands for corn.

1F. EVALUATION OF CALIBRATION PROCEDURES FOR MEASUREMENT OF BIDIRECTIONAL REFLECTANCE FACTOR IN REMOTE SENSING FIELD RESEARCH

Barrett F. Robinson and Larry L. Biehl

1. Introduction

The use of measurements of optical radiation for identification and area estimation of agricultural crops from earth satellites has reached near operational status for some crops. As the new technology develops, there is growing awareness of the need for quantitatively understanding the optical properties of crops and soils.

Researchers are heading to the fields in increasing numbers with a variety of instruments to measure radiation reflected and emitted by soils and crop canopies. While many types of measurements will prove to be useful, there is a need for data which may be compared from site to site and instrument to instrument, independent of atmospheric conditions.

This section describes and evaluates a reflectance calibration procedure which has been used since 1974 by Purdue/LARS and NASA/JSC in the field research project sponsored by NASA/JSC to obtain field measurements which are being analyzed across sites, instruments, and dates in ongoing agricultural experiments.

2. Bidirectional Reflectance Factor (BRF)

A reflectance factor is defined as the ratio of the radiant flux actually reflected by a sample surface to that which would be reflected into the same reflected beam geometry by an ideal (lossless) perfectly diffuse (Lambertian) standard surface irradiated in exactly the same way as the sample.

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, Figure 1F-1. For small fields of view (less than 20 degrees full angle) the term bidirectional reflectance factor has been used to describe the measurement: one direction being associated with the viewing angle (usually 0 degrees from normal) and the other direction being the solar zenith and azimuth angles.

3. Discussion of the Approach

The objective of the approach is to obtain a property of the scene which is nearly independent of the incident irradiation and atmospheric conditions at the time of the measurement. The majority of measurements are made by viewing along the normal to the subject and the level reference surface with solar angles similar to those for satellite overpasses. Under these conditions, the properties of the subject are measured which are related to the response of a satellite-borne sensor.

An understanding of the factors which affect the bidirectional reflectance factor of agricultural scenes will provide a basis for improved identification and characterization of agricultural crops. A deeper understanding will be required in the future with the advent of improvements in characterizing the optical properties of the atmosphere over frames of satellite-borne sensor data and future satellite sensors which may view the surface at angles significantly different from normal.

4. Reference Surfaces

An assumption for measuring BRF is that the reflectance properties of the reference surface are known. Three kinds of reference surfaces have been used by Purdue/LARS and NASA/JSC for field research, Figure 1F-2. They are pressed barium sulfate powder, painted barium sulfate, and canvas. All three reference surfaces have medium to high reflective properties, Figure 1F-3, and are highly diffuse for the solar angles encountered, Figure 1F-4.

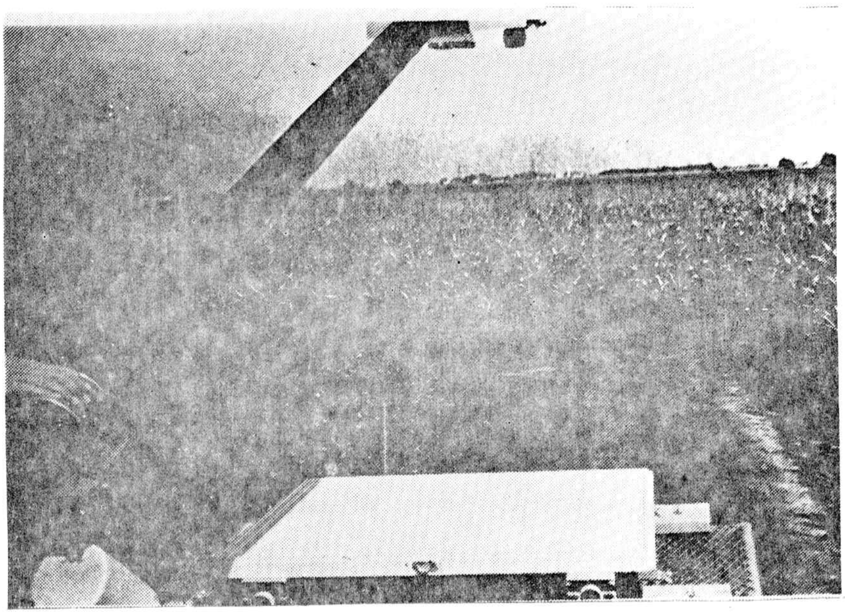


Figure 1F-1. Multiband radiometer positioned over level painted barium sulfate reference panel for field reflectance calibration.

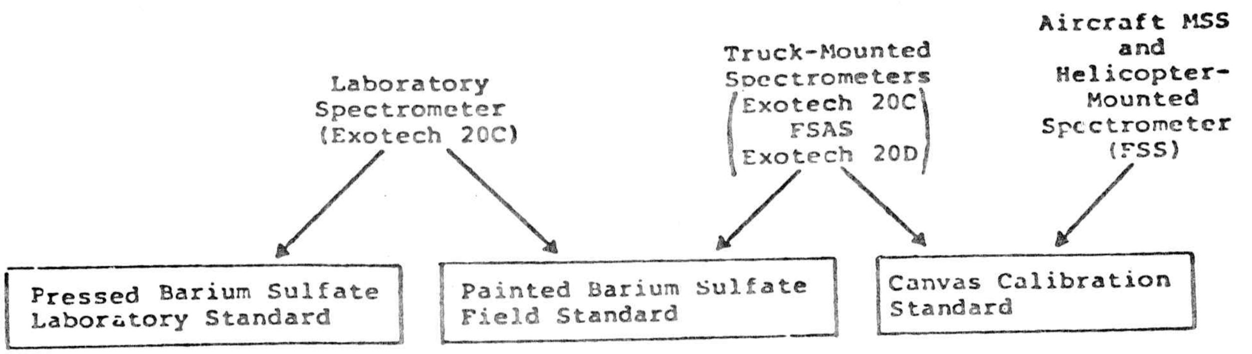


Figure 1F-2. Bidirectional reflectance factor calibration transfer of reference surfaces used for agricultural field research.

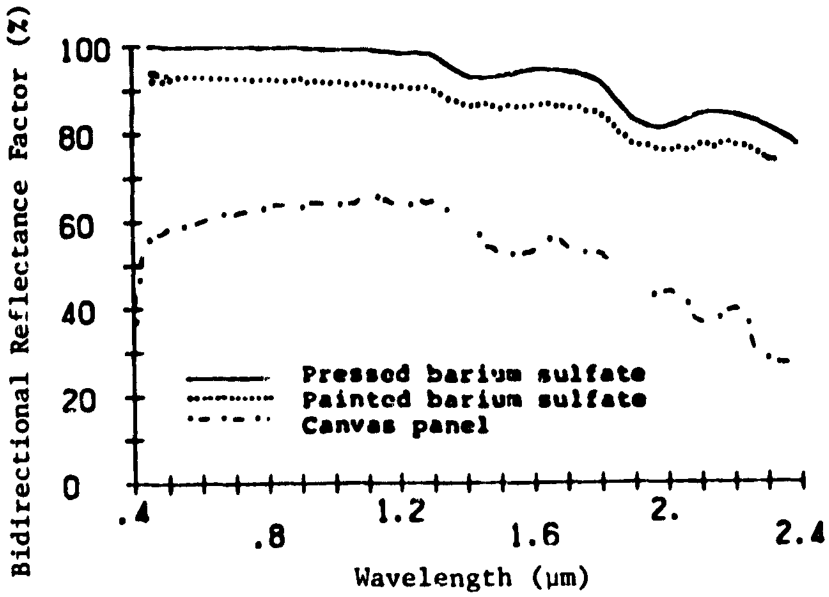


Figure 1F-3. Spectral bidirectional reflectance factor of three reference surfaces used for field research.

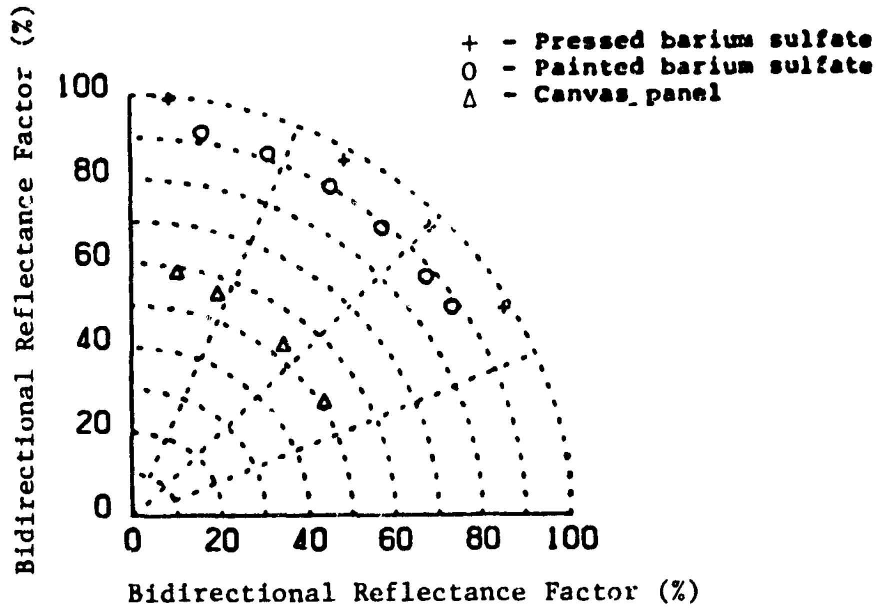


Figure 1F-4. Bidirectional reflectance factor at 0.6 µm, for several incident angles, θ, and reference surfaces. Polar coordinates.

5. Effects of Nondirectional Illumination

An assumption for the measurement of BRF is that the incident radiation is directional. Clouds and skylight cause the irradiance of the subject and reference surface to be other than directional. Clouds can be avoided by taking data on clear days. Skylight, however, can not be avoided.

The relationship of measured bidirectional reflectance factor, R_F, which includes skylight and the true bidirectional reflectance factor R_t, may be described as:

$$R_F = R_t \left[1 + K_1 \cdot K_2 \right] \qquad (1)$$

where K_1 represents the fractional amount that the bidirectional reflectance distribution function of the scene differs from a Lambertian surface, and K_2 represents the ratio of the diffuse component of the irradiance to the total irradiance.

The maximum value of $K_1 \cdot K_2$ is 0.03 for a representative bidirectional reflectance distribution for vegetation and previously measured skylight to total irradiance ratios. Therefore, the reflectance factor measured including the skylight will differ from the true reflectance factor by a systematic 3% of value on a hazy day (visibility = 8 km) for the spectral band 0.5 to 0.6 μm.

6. Comparison of Spectral Bidirectional Reflectance Factor Measurements made by Spectrometer Systems

During the Large Area Crop Inventory Experiment (LACIE), four different spectroradiometer systems were used to collect agricultural field spectral BRF measurements. The spectrometer systems were used in five different test sites in North Dakota, South Dakota, and Kansas. In July of 1977, three of the spectrometer systems were brought together to the Williams county, North Dakota, test site for a formal comparison study.

The spectral BRF measurements by the three field systems of the canvas panels (5 to 60 percent), compared quite favorably, as illustrated in Figure 1F-5. The BRF measurements from the separate field systems agree to within 4 percent of value for BRF ranges from 5 to 60 percent, the normal range for most normally viewed agricultural crop canopies. The results of the instrument comparison study, indicate that the procedure developed for the measurement of BRF is sound. Moreover, quantitative information about the comparison of the measurements from different spectroradiometer systems is available. The use of common scenes such as canvas panels is a valuable aid in accessing the comparability of several spectrometer systems.

7. Conclusions

Bidirectional reflectance factor (BRF) is an appropriate and useful optical property for remote sensing field research because it is a fundamental property of the subject. The described procedures for use of reflectance surfaces provide a good approximation to the true BRF of the subject because the irradiance is dominated by its directional component, the reference surface is nearly Lambertian and the BRF of the subject is not radically different from Lambertian.

The described procedure is an effective means to acquire data which may be meaningfully compared from time to time, site to site, and instrument to instrument because:

- It is relatively easy to train instrument operators to obtain repeatable results
- The reference surfaces can be prepared and tested at central locations; therefore, most researchers do not need sophisticated calibration apparatus.
- The performance of different instruments can be easily compared under field conditions.

Acquisition of meaningful data requires that measurements be made at the appropriate scale. This entails positioning the sensor at a proper distance above the subject, and careful consideration of the field of view. Without careful planning, these factors and other procedural errors can seriously limit the usefulness of well calibrated data.

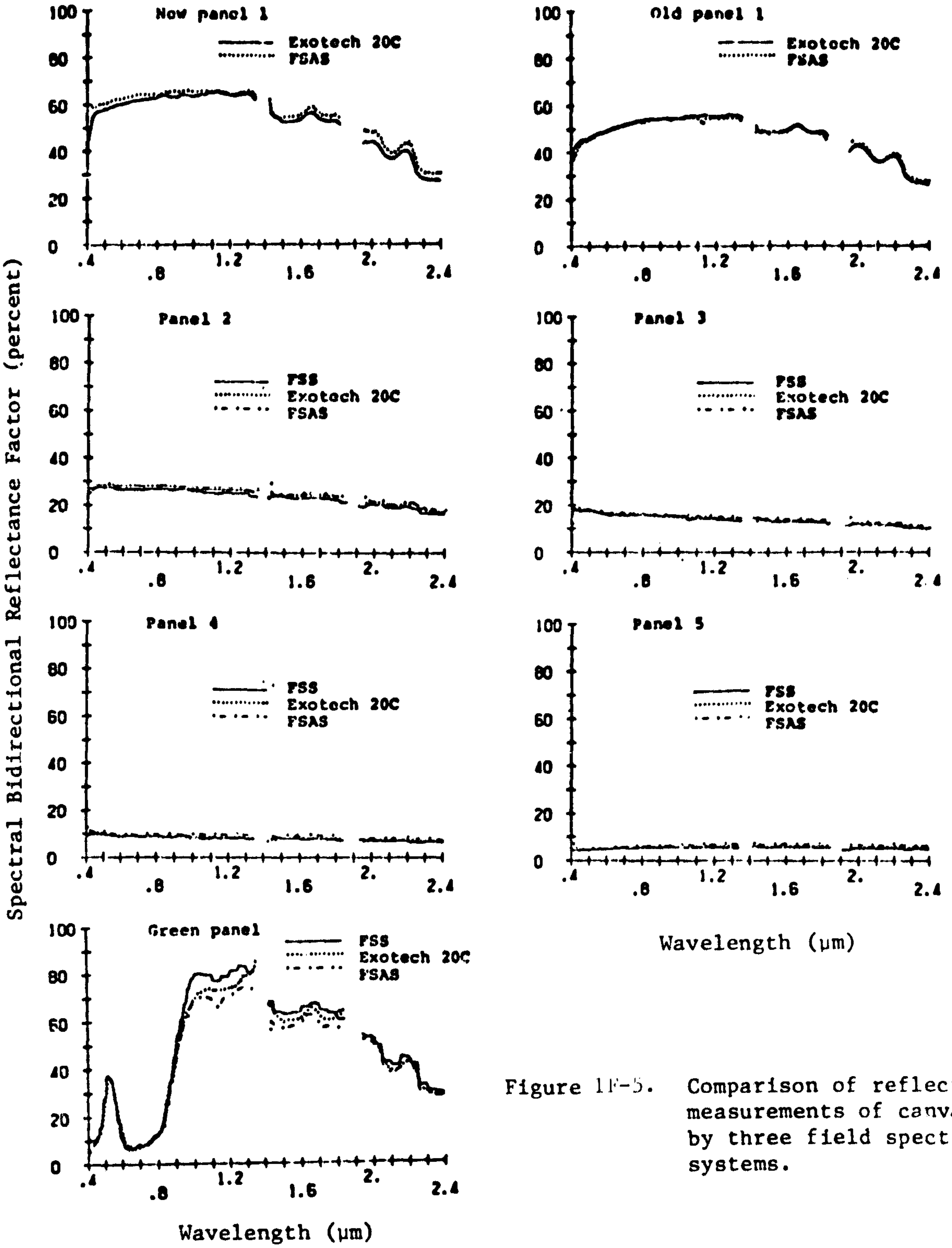


Figure 1F-5. Comparison of reflectance measurements of canvas panels by three field spectrometer systems.

2A. APPLICATION AND EVALUATION OF LANDSAT TRAINING, CLASSIFICATION, AND AREA ESTIMATION PROCEDURES FOR CROP INVENTORY

Marilyn Hixson

1. Introduction

This task is the second year of a specific LARS task which resulted from a proposal in response to the Applications Notice It is also part of the second year of effort in a larger, multiyear, multi-organizational effort to extend LACIE-like technology to crops other than the small grains. The accuracy and precision of area estimates obtained from Landsat data are affected by a combination of training, classification, and area estimation procedures used.

2. Objectives

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

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

3. Experimental Approach

The data set used in this study was drawn from the 1978 multicrop data over four test areas in Iowa, Illinois, and Indiana (Figure 2A-1). LACIE-type sample segments (5 x 6 nm in size) were selected, generally two per county. Landsat data acquired included multitemporally registered MSS data tapes and film writer imagery (PFC Product 1). Color infrared prints of aerial photography with ground inventory overlays were obtained for 69 segments.

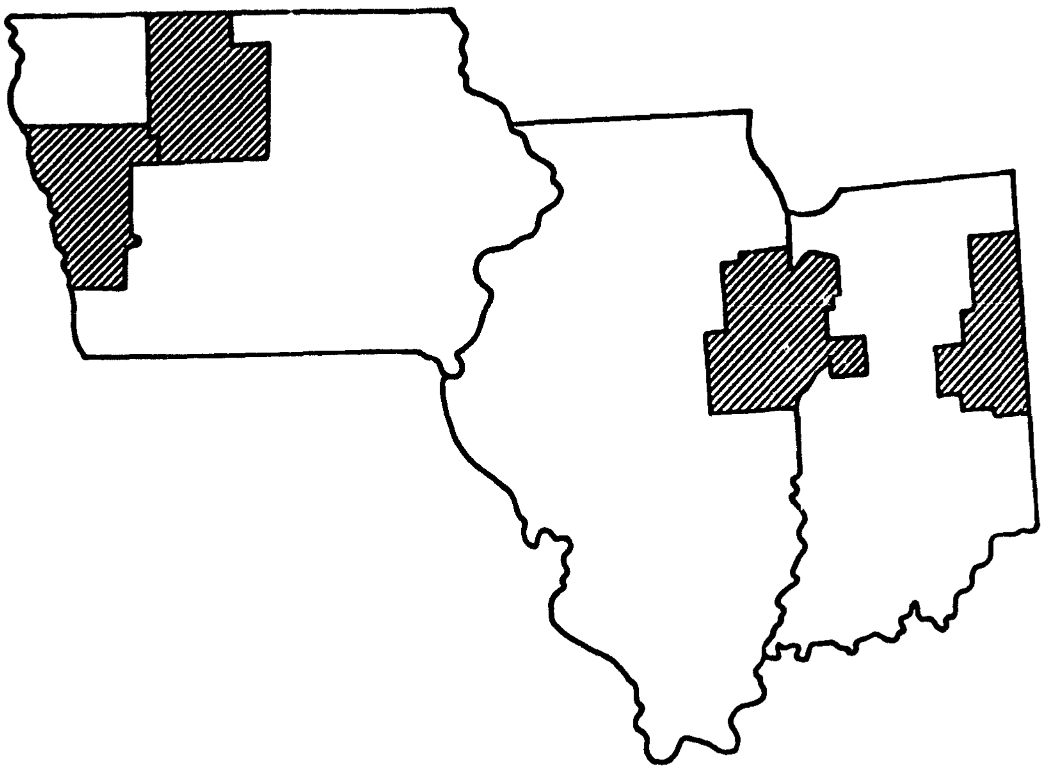


Figure 2A-1. Locations of the four test areas containing sample segments used in this investigation.

4. Sampling Effects in Clustering and Classification

A study was conducted to investigate the "best" subset of bands for crop classification. Multitemporal data from four Landsat acquisitions were analyzed to select the best combination of four from 16 available channels based on average transformed divergence. The first channel on each date was very rarely selected; the two near infrared bands were both selected with high frequency on all dates. It was discovered that of the 30 best channel combinations in four segments, neither two visible, nor two infrared channels from the same date were ever selected. Channels two and four from each date appear to give a good subset to classify with or select another set from. Other analyses confirmed that there was no significant difference between use of eight vs. 16 bands in a four date analysis. But, use of less than half of the wavelength bands caused significant decreases in accuracy.

5. Evaluation of Alternative Training Methods

Several aspects of Procedure 1 were investigated and their effects on estimates were assessed. No significant differences in estimates of corn or soybeans were found using L1 or L2 distance in the LABEL processor. No significant differences in estimates of corn or soybeans were found when clusters smaller than 100 points were deleted, but slightly higher classification accuracies were obtained. An evaluation of the number of iterations used in ISOCLS showed that the distribution of cluster standard deviations improved very slightly with more iterations. The space covered by cluster means did not change substantially with additional iterations.

A second study examined procedures used in a modified supervised training approach. Four acquisitions were analyzed, one from each of four time periods: preplant to eight leaves of corn, ten leaves to tasseling, tasseling to beginning dent, and dent to maturity. Weighted and unweighted separability measures were used to select the best four of six or eight channels for use in classification. In the majority of cases, the same subset was selected. If a different subset was selected, the weighted method produced classification results of higher accuracy. Deleting small clusters (15-20 points) consistently increased classifier performance.

6. Comparison of Performance for Five Classification Algorithms

The overall objective of this study was to evaluate the performance of five classification schemes on agricultural data sets which sampled winter wheat in Kansas, spring wheat in North Dakota, and corn and soybeans in Indiana, Illinois, and Iowa. Classifications were performed using four bands selected from four acquisition dates which temporally sampled the crop calendar. Five classifiers were selected for study; Classifypoints, Classify, Minimum Distance, Layered, and ECHO. In order to insure that differences in classification accuracies were the result of classifier differences and not training methods, the same set of training statistics was used for all classifiers. Training fields were clustered within cover type to define spectral subclasses for each of the classes of interest. Also used with CLASSIFY was a training method using a random selection of pixels to define initial cluster centers for clustering the entire area.

Segment-to-segment variability was highly significant (Table 2A-1). Segment variability was attributed to factors other than the classifier selected, including spectral data quality and scene characteristics such as field sizes and number of confusion crops present.

There was no significant difference among classifiers in percent correct classification of corn, soybeans, or other in the five Corn Belt segments. In addition, there was no significant difference in overall classification accuracy among classifiers for all seven segments. The sum-of-normal-densities classifier using LARSYS statistics, however, had significantly higher small grain classification accuracy (about 2% improvement). Although differences were nonsignificant overall, the LARSYS training method provided a consistent improvement over the ISOCLS training method in six of the seven segments.

The classification schemes varied considerably in ease of use. ECHO requires setting the parameters for cell homogeneity testing and cell size. The LAYERED classifier requires the additional step of designing a decision tree, which can be very complex if many spectral classes and features are needed to characterize the scene of interest.

Table 2A-1. Comparison of classifier performance (percent correct) classification by test site.

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

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

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

7. Landsat Data Acquisition History

The objectives of this study were:

- Assess the accuracy of early season estimates.
- Determine a minimum number and distribution of acquisitions necessary for accurate estimation of corn and soybean areas.
- Compare minimum distance, maximum likelihood, and sum-of-densities classifiers in other band/date combinations than previously assessed.

The accuracy of early season estimates is illustrated in Figure 2A-2. During the first defined time period, corn and soybeans were not spectrally separable. In the Corn Belt, however, relatively accurate identification can be made of corn and soybeans together at that time. It is not until after the corn has tasseled that consistently high classification accuracies are obtained. The classification accuracy does not improve by using later season information when the crops of interest have reached maturity.

It seems that the availability of acquisitions from time periods one (about emergence) and three (after tasseling of the corn) provides a minimal set for accurate identification of corn and soybeans (Figure 2A-3). No combination of acquisitions which does not include growth stage three gives high classification performance; a stage one acquisition appears to be less critical if acquisitions from other time periods are available. The overall accuracy of the third period alone was only 85%, illustrating that classification using the single best acquisition period is not as accurate as can be obtained using multitemporal information.

A comparison of the minimum distance, maximum likelihood, and sum-of-densities classifiers showed that the difference in overall classification accuracies was significant ($\alpha = .01$), with the sum-of-densities classifier having the highest accuracy and the minimum distance classifier having the lowest accuracy. This pattern held for individual combinations of acquisition periods and segments in general. Most of the performances

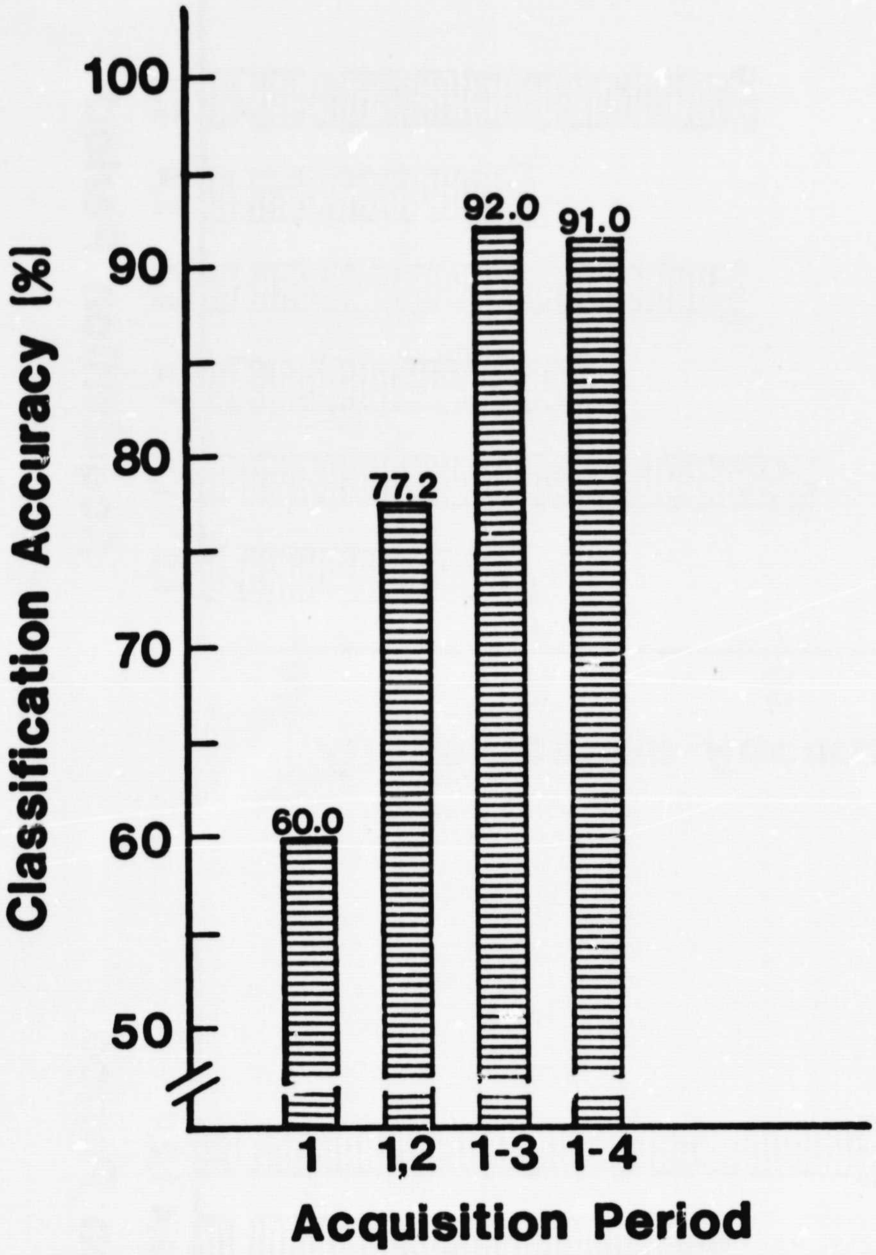


Figure 2A-2. Overall classification performance using cumulative spectral information with a minimun distance classifier and subsets of two, four, six, and eight channels.

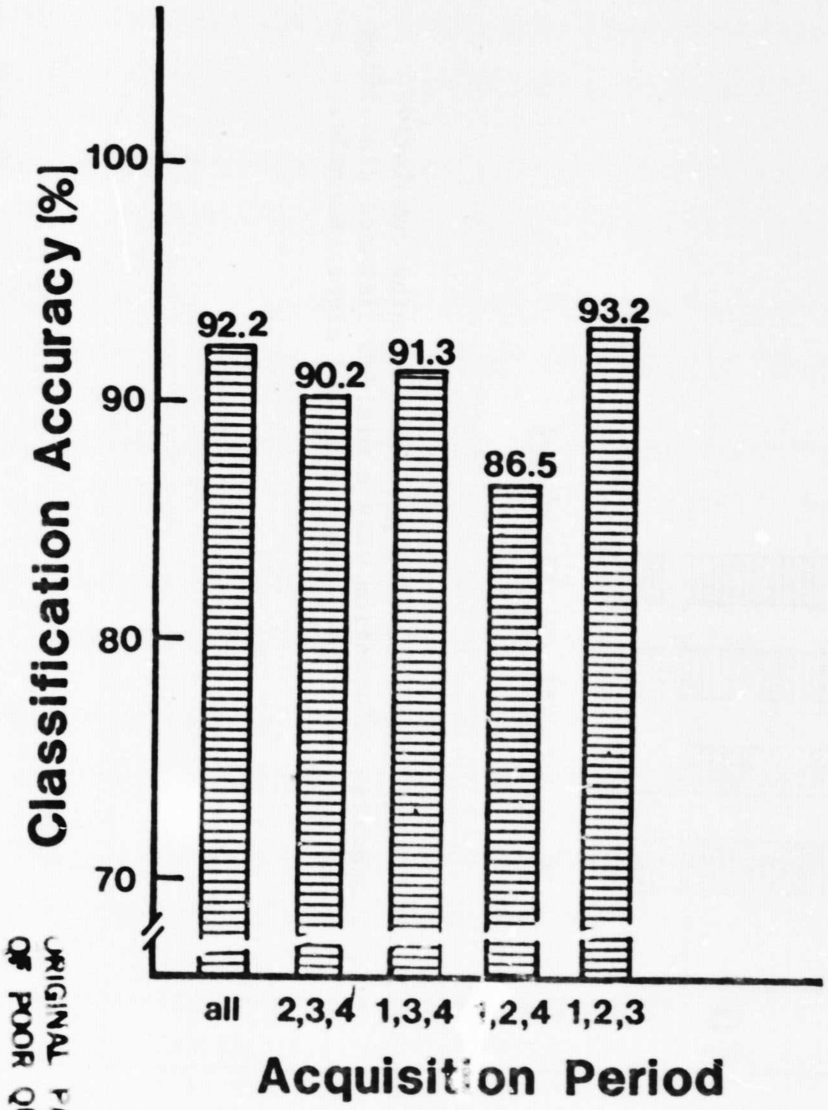


Figure 2A-3a. Overall classification accuracies of three and four date classifications.

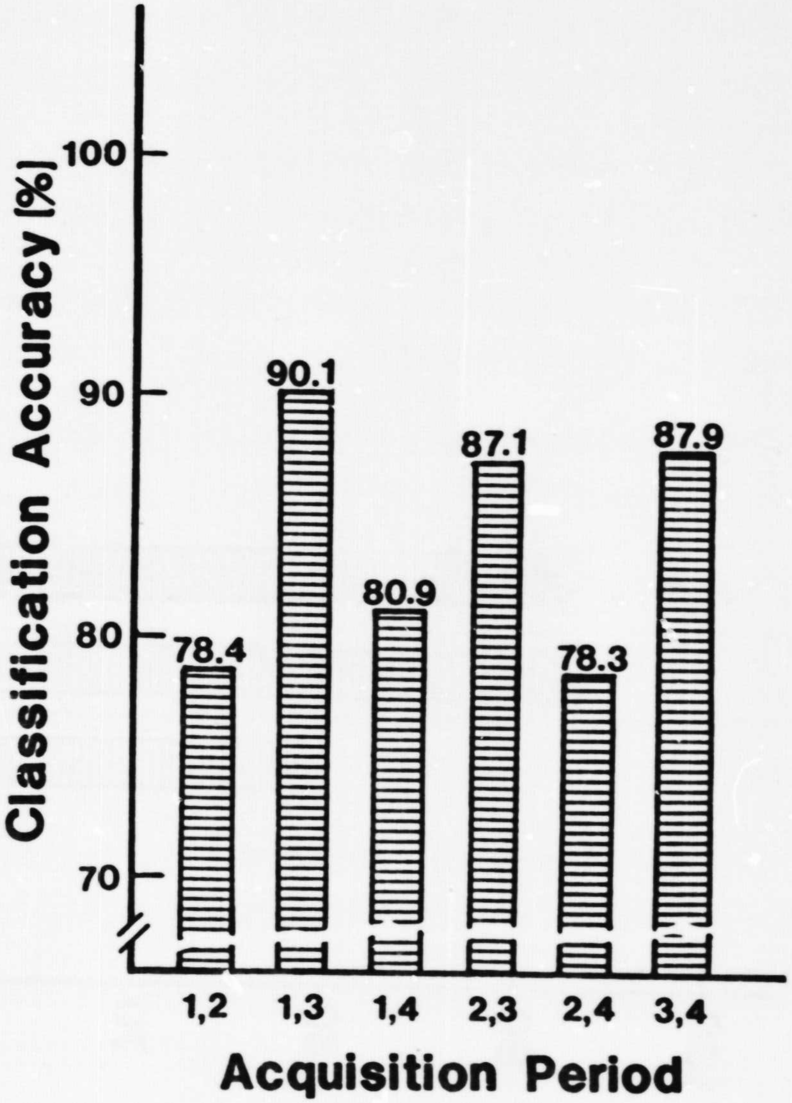


Figure 2A-3b. Overall classification accuracies of two date classifications.

were within about 2% for all classifiers, so classification costs (which increase in the same order as performance) should probably be considered in the choice of a classifier.

8. Future Plans

This investigation will be continued during the next contract year. Further studies on training unit size and training data selection will be conducted. A wider variety of segments across the U.S. Corn Belt will be classified and performance related to scene characteristics. A study investigating separation of the functions of sampling for training and sampling for area estimation is also planned.

2B. INITIAL DEVELOPMENT OF SPECTROMET YIELD MODELS FOR CORN

C.S.T. Daughtry

1. Introduction

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

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

2. Objectives

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

- Identify important factors in determining and predicting corn yields.
- Determine how these factors can be observed or estimated from alternate sources of data.
- Define long-term data requirements for continued model development.
- Select and further develop several candidate approaches for corn yield modeling.
- Identify and obtain data required for these yield models.
- Conduct initial calibrations and tests of models using spectrometer and Landsat MSS data.

3. Description of Data

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

The other major source of spectral data was Landsat MSS data acquired over commerical corn fields in nine 5 x 6 mile segments located in six states. Within each of these segments up to 10 corn fields were identified and periodically observed throughout the growing season by USDA personnel. These observations consisted of notes on plant height, percent soil cover, maturity stage, and recent field operations. Grain yield in each field was either estimated by the USDA representative or acquired during an interview with the farmer. Grain yields ranged from 50 bushels per acre in Ballard, KY to 158 bushels per acre in Iroquois, IL.

4. Results and Discussion

4.1 Factors Influencing Crop Yields and Prediction of Crop Yields

The economic end product of crop production is often the seed which comprises about [illegible]5 percent of the above ground dry weight of corn. This accumulation of dry matter requires not only the availability of the proper substrates (CO_2, H_2O, NH4+and/or NO_3^-, and other nutrients) in the environment but also a great deal of energy which the plant derives from sunlight.

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

1) crop factors - e.g., photosynthetic rate, stress tolerance, leaf area index, leaf area duration, growth rate

2) soil factors - e.g., drainage, water-holding capacity, fertility

3) management factors - e.g., planting date, weed, disease, and insect controls, cultivar selection.

4) weather factors - e.g., solar radiation, air temperature, precipitation, evaporation.

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

If weather is truly the most important factor controlling crop yields, how can the effects of weather on crop response (yield) be quantified? Reviews of research on environmental and physiological aspects of crop yield have identified and generally attempted to quantify optimum conditions for assimulation processes, growth, development, and ultimately yields for various crops.

Of the various physical measurements of the environment, temperature, moisture and solar radiation are most frequently used to estimate crop yields. Researchers have used various experimental techniques to relate hourly, daily, weekly or monthly means of temperature, moisture (precipitation or soil moisture) and/or solar radiation to yields. Some have used selected weather variables from the entire growing season while others have preferred to identify physiologically important periods during which they felt crops were most sensitive to the effects of weather. While these fitted parameters may be associated with reasonable proportions of the variance in fitted crop yield series, the predictive equations generally explain disappointingly little of the crop yield variance in independent tests.

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

4.2 Data Requirements and Sources of Data

Data requirements for crop model development vary greatly depending on the specific type of model employed. The most commonly recorded physical measures of the environment are daily maximum and minimum air temperatures and daily total precipitation. Less common measurements include solar radiation, evaporation, wind travel, soil temperatures and soil moisture on daily and in some cases hourly basis. These data are frequently used in crop models either by design or necessity since other data are available only in special instances. Variability of precipitation patterns in time and space makes precipitation both the most important and most error-prone in any water budget or weather and crop yield study.

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

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

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

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

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

Since crop response to weather may differ from year to year, a homogeneous series of crop and weather factors are required for continued model development. Changing segment locations annually does not contribute to a homogeneous series of crop and weather factors and generally serves to confound an already complex situation.

4.3 Approaches for Crop Yield Modeling

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

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

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

where,

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

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

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

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

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

Weather Factor

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

1) Simulation or causal models which describe crop performance as a series of functions with daily solar radiation, air temperature and moisture. Simulation models are broadly applicable, require short historical data bases for development, and can provide local detail.

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

3) Hybrid models which seek to combine some of the best features of both simulation and statistical models by condensing the effects of weather on crops into a single weather index which can be related to yield.

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

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

Spectrally-derived information about crop condition can be used to estimate intercepted solar radiation. An estimate of intercepted solar radiation based on spectrally derived LAI should more accurately depict conditions in the field than the averaged values of LAI. An alternative, which has yet to be investigated, is to estimate intercepted solar radiation directly as a function of reflectance and incident solar radiation.

Regardless of which crop model is employed, its spatial resolution is limited by the distribution of weather stations. The best estimate of yield that can be expected from any of these models is the mean of a region. If there exists considerable variation in yields within a region due to, for example, soil fertility then these models are not likely to estimate yields very precisely or accurately at the local level. Spectral data, on the other hand, is limited by the spatial resolution of the sensor which is 0.45 ha for Landsat MSS. On a temporal basis meteorological data is available much more frequently than most spectral data. Thus combining the better temporal resolution of meteorological data with the better spatial resolution of spectral data should produce yield estimates with local detail unsurpassed by either method alone.

Information about crop yields at the local level that is contained in spectral data is illustrated in the Corn Nitrogen experiment. The weather, soil type, and soil moisture conditions were as nearly alike as possible in all plots of this experiment. Only the amount of nitrogen fertilizer differed from plot to plot. The best estimate of corn yields for this area by a meteorologically-based yield model would be the mean yield. Figure 2B-2 illustrates the departures of individual plot yields from mean yield due to nitrogen fertility and how some of this variation

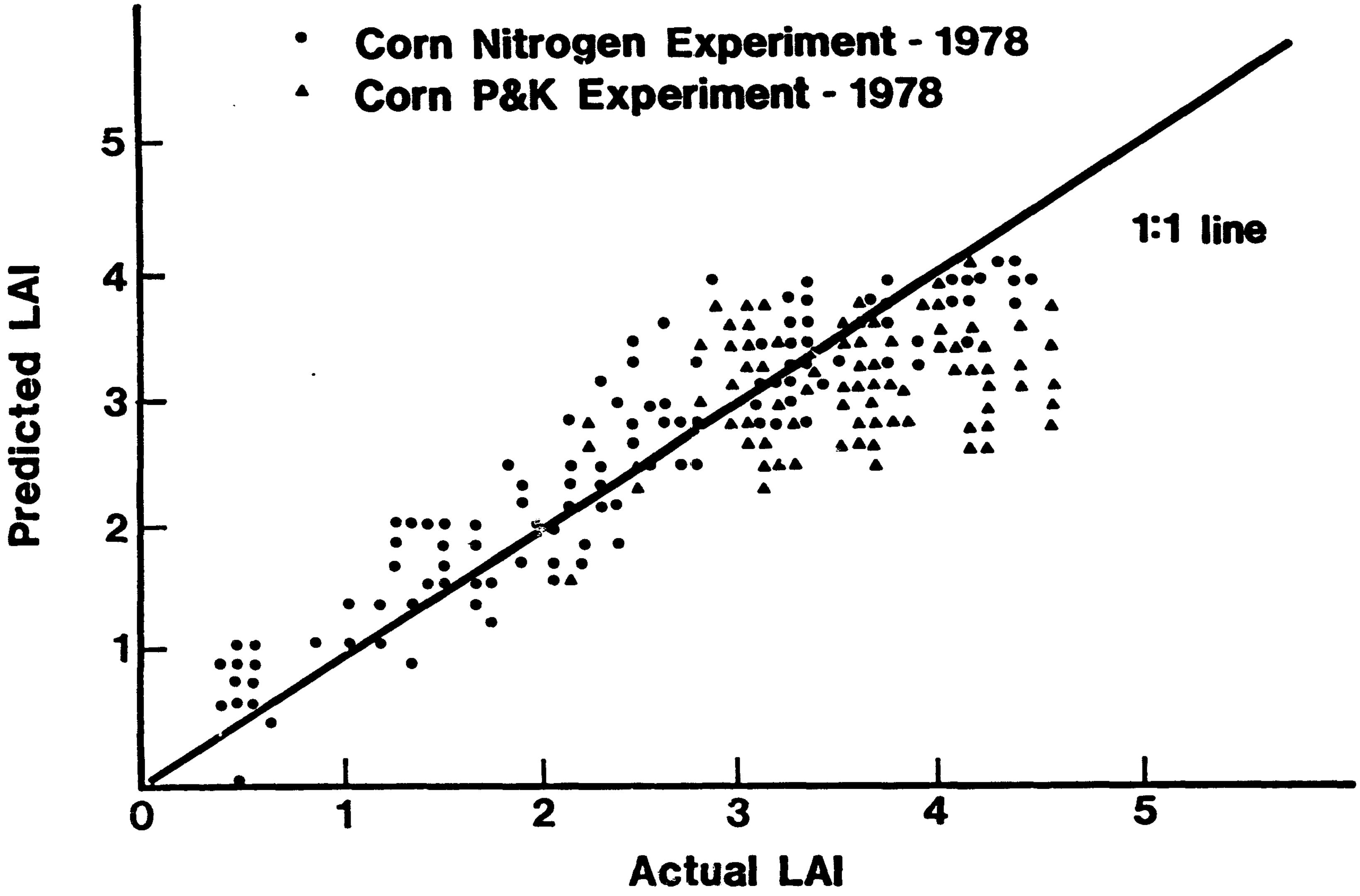


Figure 2B-1. A comparison of measured leaf area index (LAI) and LAI predicted from spectral data in the four Landsat MSS bands for two experiments at Purdue Agronomy Farm in 1978. The coefficients of the regression equation were derived with data from the Corn Nitrogen Experiment and were plotted with data from both experiments. $\widehat{LAI} = 0.523 - 0.953 * B50 + 0.399 * B60 + 0.154 * B70 + 0.380 * B80$.

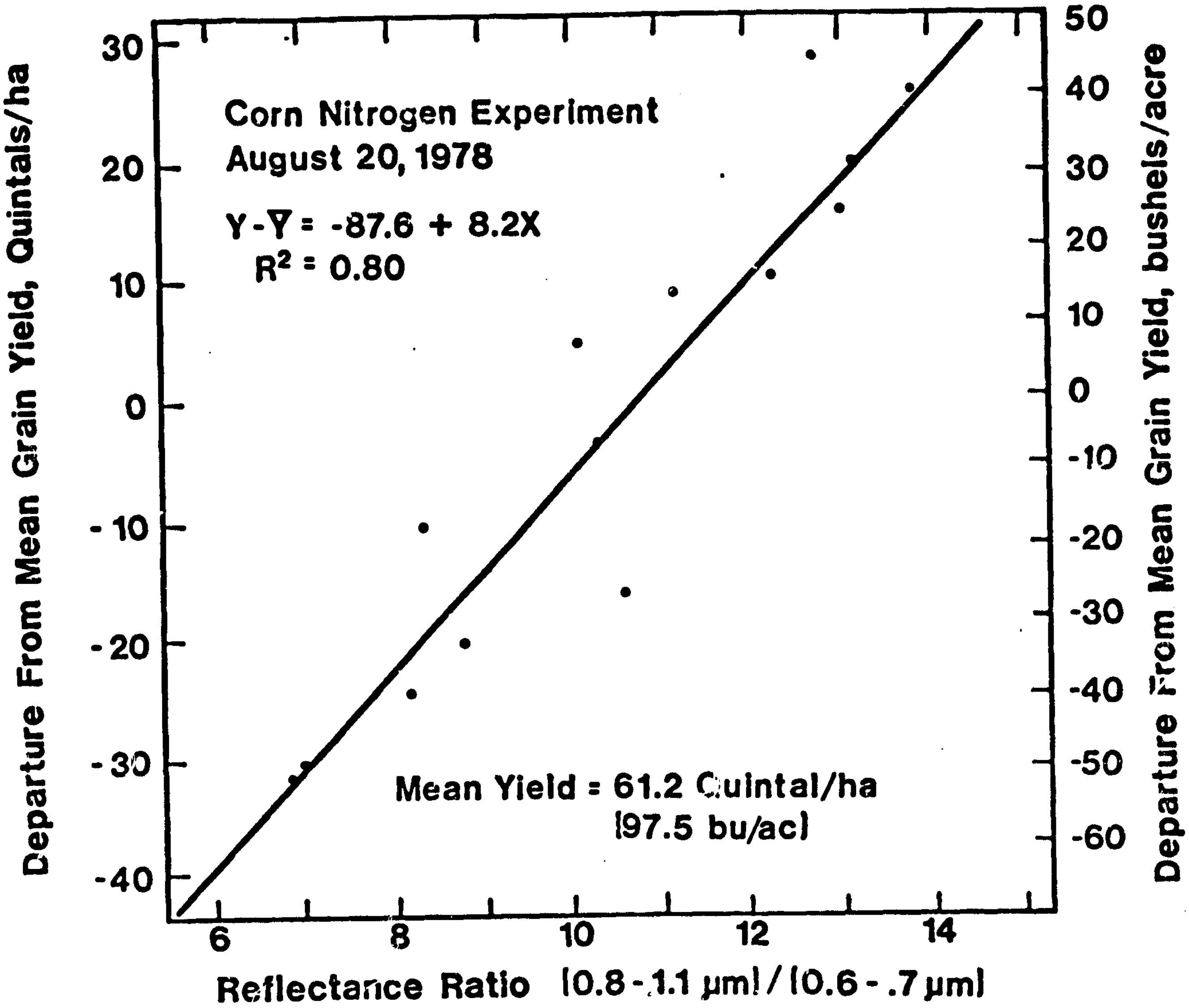


Figure 2B-2. Association of the ratio of reflectances in the near infrared (0.8 - 1.1 µm) band and the red (0.6 - 0.7 µm) band with departures from mean grain yield for the Corn Nitrogen Experiment in 1978.

about the mean is associated with a spectral variable such as the ratio of reflectances in 0.8-1.1 and 0.6-0.7 μm bands. This relationship appears to be rather stable for 4 to 6 weeks during the tasseling and grain filling periods of corn. From this limited data set it appears that this period occurs at or shortly after the time when the maximum IR/red ration of corn is reached (Figure 2B-3). Together Figures 2B-2 and 2B-3 represent a potential method, not only to adjust yield predictions from meteorological models, but also to identify the time interval when remotely-sensed data are most highly correlated with corn yields.

Extension of these simple concepts developed from spectrometer data gathered at an agricultural experiment station to Landsat MSS data acquired over commercial fields represented quantum leaps in scene complexity and potential sources of unaccounted for variability. Initial examinations of the Landsat MSS data from selected corn fields indicated that maximum Kauth Greenness occurred at or shortly after tasseling (Figure 2B-4) as expected from spectrometer data.

Figure 2B-4 represents typical fields of corn in Pottawattomie County, Iowa and Tippecanoe County, Indiana and have basically similar shapes. The abrupt changes in greenness over a two day period are caused by from consecutive-day passes with Landsat MSS. The influence of the atmosphere on spectral response was not considered and may account for some of the abrupt changes in greenness over 9 to 18 day periods.

Correlations of Greenness and IR/Red ratio with yields are greatest near tasseling and are consistent with results from agricultural experiment station data. Preliminary indications are that simple correlations of Landsat MSS data with departures from mean yield for each segment will not be sufficient to explain the variation in yields observed in individual fields. Additional research is in progress to examine these relationships fully. Alternative approaches which will use spectral data indirectly to estimate yields are also being pursued.

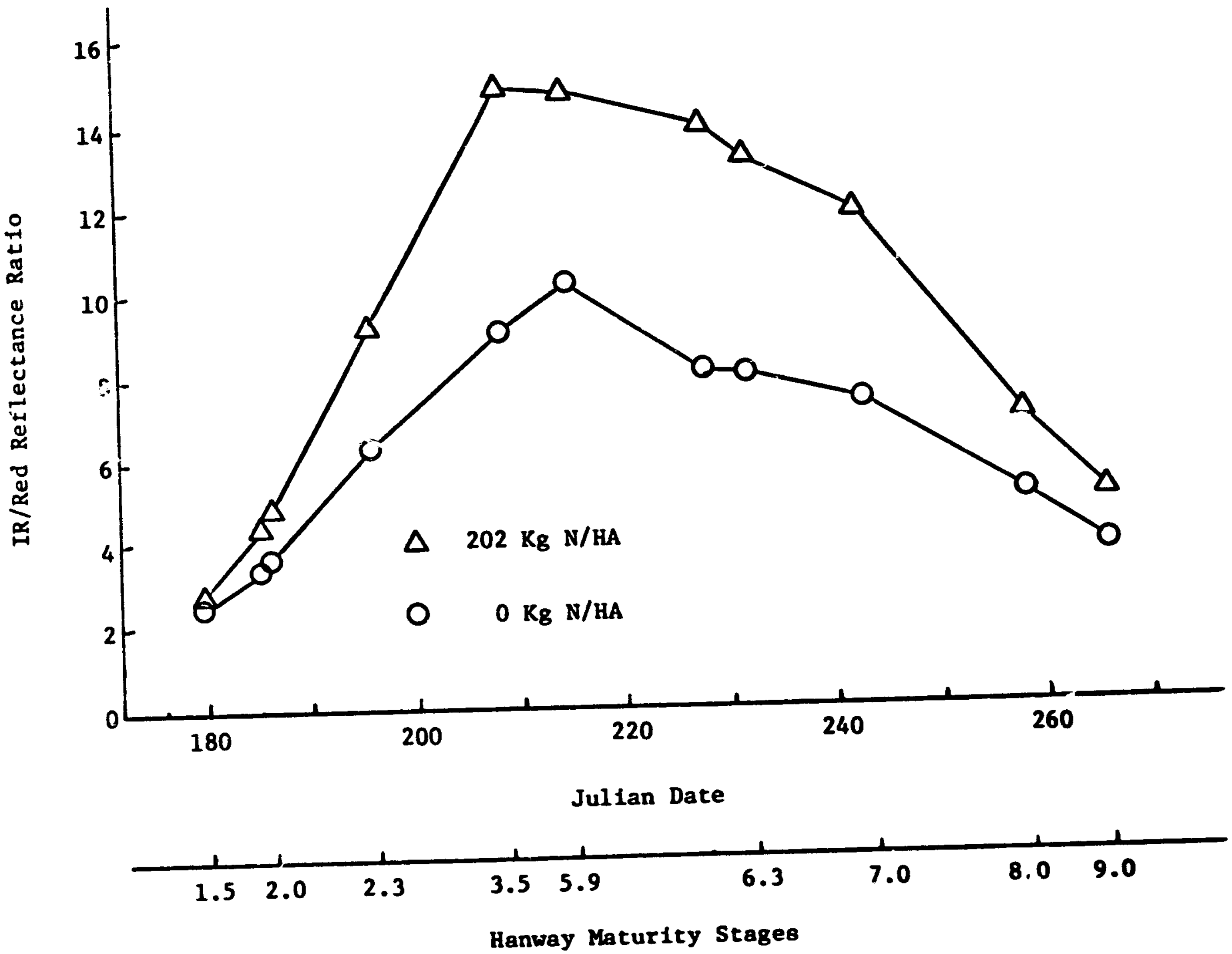


Figure 2B-3. Seasonal changes in ratio of reflectances in a near infrared (0.8 - 1.1 μm) band and a red (0.6 - 0.7 μm) band for the Corn Nitrogen Experiment in 1978. Note that the maximum reflectance ratio occurs near time of tasseling (Maturity Stage 5). Only high and low N treatments are shown for clarity.

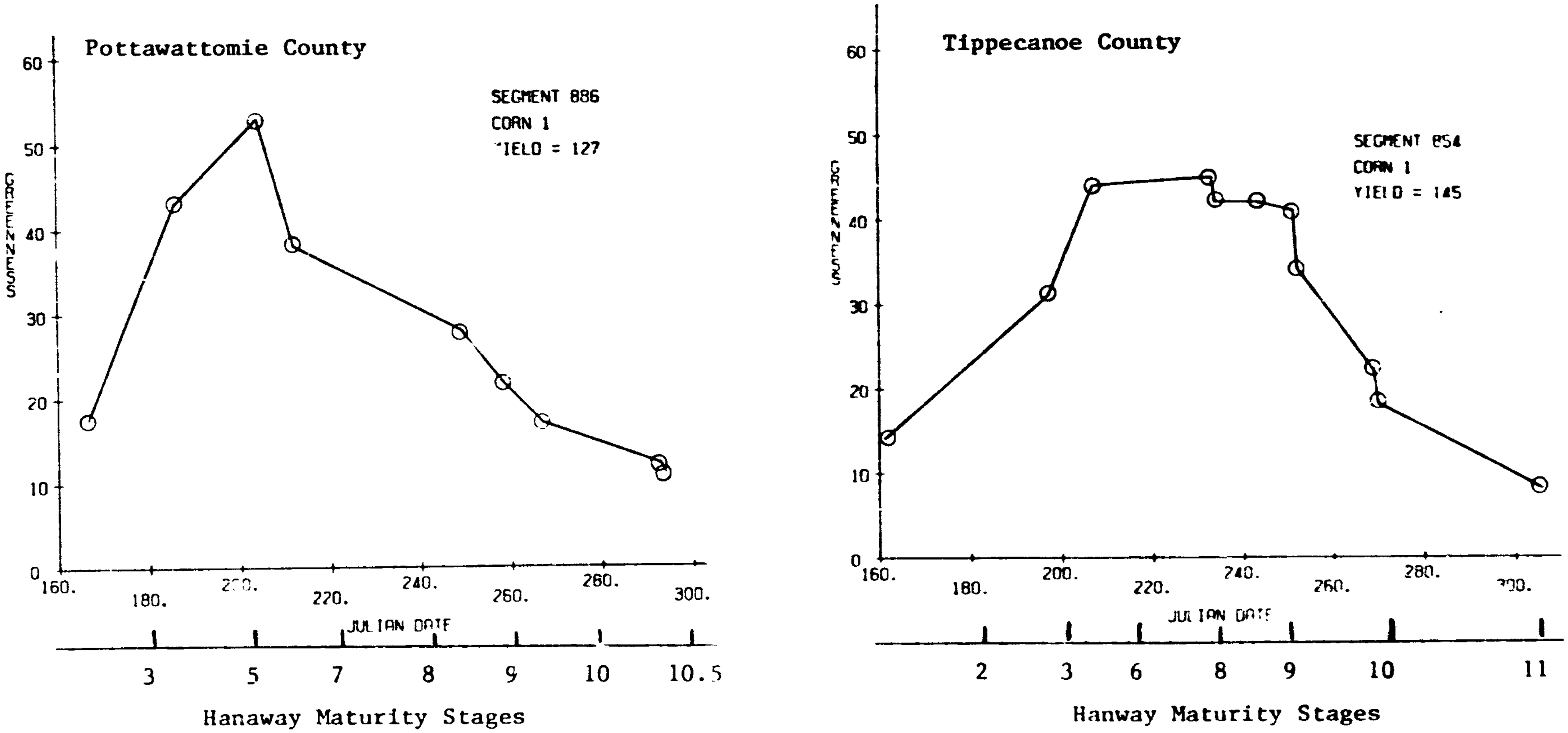


Figure 2B-4. Seasonal changes in Kauth Greenness transformation of Landsat MSS data acquired over corn fields in Iowa and Indiana in 1978. Yields are in bushels per acre. Maximum greenness occurs near tasseling/silking (Maturity Stage 5).

3A. MULTISPECTRAL DATA ANALYSIS RESEARCH

Philip H. Swain and David A. Landgrebe

Task 2C consists of three subtasks involving research into advanced methods for classifying multispectral remote sensing data. The first two are multiyear investigations resulting from proposals submitted to NASA in response to the 1978 Applications Notice, OSTA-78-A (April 19, 1978).* The first year of work on both of these subtasks is reported here.

The third subtask resulted from a proposal submitted to NASA Johnson Space Center during the contract year.$ The work was funded quite late in the year and will be continued. A background discussion is contained in this report.

1. Multistage Classification

A number of different types of classifiers are now in routine use in remote sensing. The more advanced applications of the future, however, will require even more detailed classes than are now feasible. To serve this need, more sophisticated sensors, such as Thematic Mapper, are scheduled and there is appearing in the user community an increased use of quantitative geographically distributed ancillary data. These developments point to the need for a concomitant advance in classifier technology.

A task has been initiated to develop classifiers which can address this need. Specifically the development of multistage decision logic is being studied. Previous work has indicated that such techniques have the potential for (1) enabling the optimal classification of multitype (combined multispectral and ancillary) data sets, (2) improving the accuracy

---

* Proposals entitled "Design and Applications of Multistage Classifiers for Earth Resources Data Analysis" and "Analysis of Multispectral Earth Resources Data Using Context." Principal Investigator on both proposals was Philip H. Swain.

$ Proposal entitled "An Addendum to Research in Remote Sensing of Agriculture, Earth Resources and the Environment." Principal Investigator: D. A. Landgrebe.

of classification, and (3) improving the computation efficiency. The concept of decision tree logic may be illustrated with the following hypothetical (but not unrealistic) example. The standard means for obtaining a crop species classification from multispectral data is illustrated in Fig. 3A-1. In order to achieve adequate accuracy for this multiclass classification, the dimensionality of the multispectral data must be relatively large (e.g., five to seven) and a final classification of each pixel is achieved in one step. A decision tree approach might lead to the classifier logic illustrated in Fig. 3A-2. Since the discrimination between vegetation and nonvegetation is a simple one, it might be accomplished without loss of accuracy, using low dimensionality (e.g., one or two spectral bands). Thus many pixels, those which are "nonvegetation," would be processed only at low dimensionality, providing a savings in computer processing.

At the second stage, subdividing vegetation, a feature set especially suited to discriminating between vegetation subcategories might be used. Because of this specialization and because additional features needed for subdividing soils and urban classes need not be used, an improvement in both accuracy and efficiency might be achieved.

It sometimes happens that the separability in terms of spectral features alone is simply not adequate for desired classes. Suppose this is the case with the classes corn and urban 2. However, it would be possible to register onto the multispectral data manually determined boundaries of urban areas. Thus this ancillary data channel could be used to discriminate between the otherwise inseparable classes using a different (and simpler) algorithm.

All of these types of improvements, efficiency, accuracy, and use of multitype data have been demonstrated empirically with actual data in individual cases. The remaining problem is that a technique for systematically designing the decision tree must be devised.

This work is still in its early stages and no final results are ready to be reported. Progress to date has included the assembly of a survey of relevant papers in the scientific literature, initiation of a

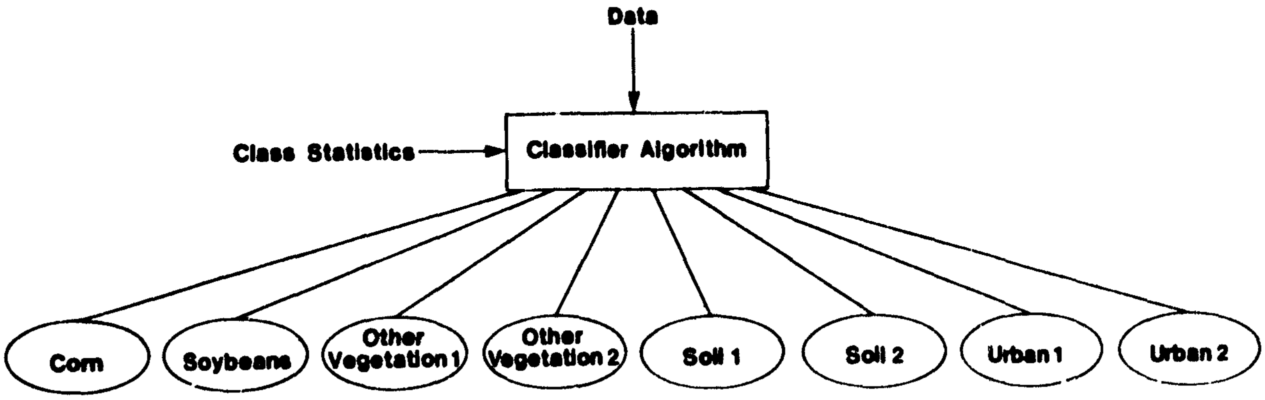


Figure 3A-1. Standard approach for crop species classification.

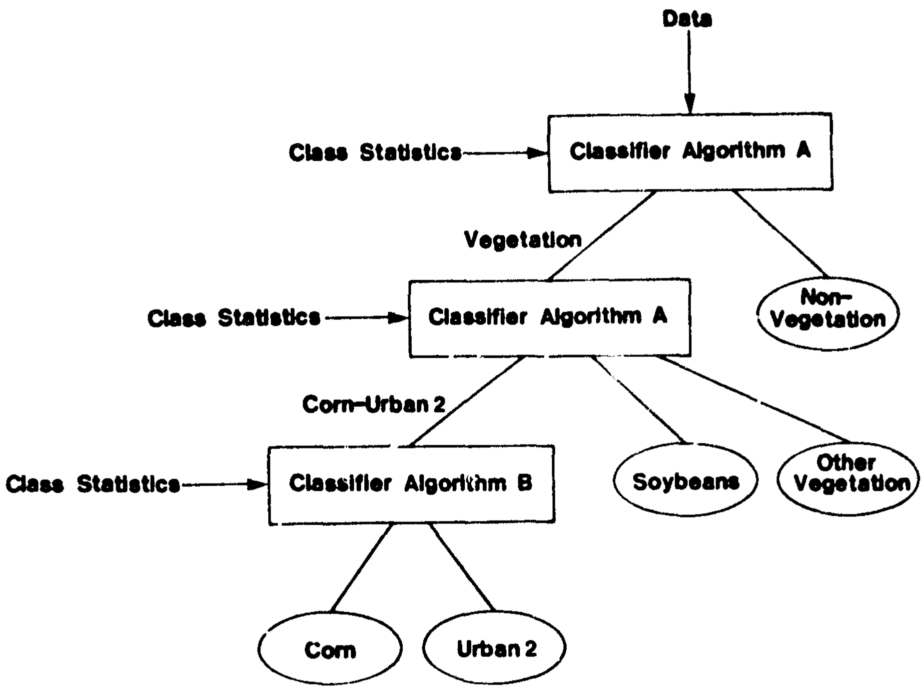


Figure 3A-2. Classifier logic for a decision tree approach to crop species classification.

preliminary study intended to lead to a list of alternative approaches, and assembly of preliminary software and test data sets with which to do the research.

The plan is to complete preliminary studies by mid-1980. Based on what is learned from these studies, a final approach will be selected and developed.

2. Contextual Classification

2.1 Introduction

Multispectral image data collected by remote sensing devices aboard aircraft and spacecraft are relatively complex data entities. Both the spatial attributes and spectral attributes of these data are known to be information-bearing; but to reduce the magnitude of the computations involved, most analysis efforts have focused on one or the other. Only within the last few years have serious efforts been made to utilize them jointly. For example, one approach uses the spectral homogeneity of "objects," such as agricultural fields, to segment the scene and then uses sample classification to assign each object as a whole, rather than its individual pixels (picture elements), to an appropriate ground cover class. Another approach involves extraction of features based on gray-tone spatial-dependency matrices from which texture-like characteristics are developed.

In this project, we are developing a more general way to exploit the spatial/spectral context of a pixel to achieve accurate classification. Just as in written English one can expect to find certain letters occurring regularly in particular arrangements with other letters (qu, ee, est, tion), so certain classes of ground cover are likely to occur in the "context" of others. The former phenomenon has been used to improve character recognition accuracy in text-reading machines. We have demonstrated that the latter can be used to improve accuracy in classifying remote sensing data. Intuitively this should not be surprising since one can easily think of ground cover classes more likely to occur in some contextsthan in others. One does not expect to find wheat growing in the midst of a housing subdivision, for example. A close-grown, lush vegetative cover in such a location is more likely the turf of a lawn.

This report contains the theoretical foundations of a contextual classifier, experimental results from applying the contextual classifier to a variety of very different sets of data, and an extensive discussion of multiprocessor implementation of the classifier algorithm.

2.2 The Contextual Classifier Model

Consistent with the general characteristics of imaging systems for remote sensing, we assume a two-dimensional array of $N = N_1 x N_2$ pixels of fixed but unknown classification, as shown in Figure 3A-3.

Associated with the pixel having image coordinates (i,j) is its true state or true classification $\theta_{ij} \varepsilon \Omega = \{\omega_1, \omega_2, \ldots, \omega_m\}$, and a random measurement vector (observation) $X_{ij} \varepsilon R^n$ having class-conditional density $p(X_{ij}|\theta_{ij})$. Further, $\{p(X|\omega_i)$, $i=1,2,\ldots m$ is the set of class-conditional probability density functions associating the multispectral measurement vector X with the classes.

We specify some arrangement of p pixel locations including a pixel to be classified. Call this arrangement the p-context array, several choices of which are shown in Figure 3A-4. Let $\underline{\theta}^p \varepsilon \Omega^p$ and $\underline{X}^p \varepsilon (R^n)^p$ stand respectively for p-vectors of classes and n-dimensional measurements. Each component of $\underline{\theta}^p$ is a variable which can take on values in Ω. Each component of $\underline{X}^p$ is a random n-dimensional vector which can take on values in the observation space. Correspondence of the components of $\underline{\theta}^p$ and $\underline{X}^p$ to the positions in the p-context array is fixed but arbitrary except that the pixel to be classified in the array will always correspond to the pth component. The notation $\underline{\theta}_{ij}$ and $\underline{X}_{ij}$ will refer to the particular instance of $\underline{\theta}^p$ and $\underline{X}^p$ associated with pixed (i,j).

A straightforward decision=theoretical development leads to the following rule:

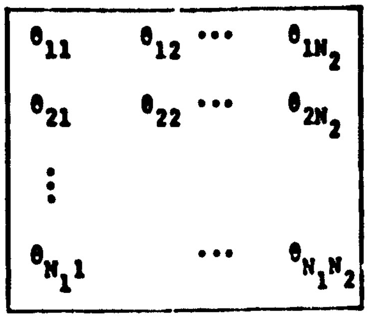
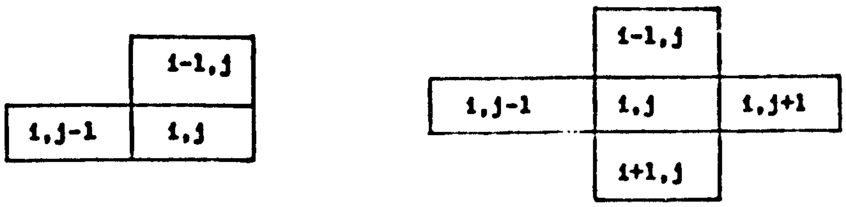


Figure 3A-3. A two-dimensional array of $N = N_1 \times N_2$ pixels.



a p=3 choice a p=5 choice

Figure 3A-4. Examples of p-context arrays.

$d(\underline{X}_{ij})$ = the action a which maximizes

$$g_a(\underline{X}_{ij}) = \left[\sum_{\substack{\underline{\theta}^p \in \Omega^p \\ \theta_p = a}} G(\underline{\theta}^p) \prod_{i=1}^{p-1} p(X_i | \theta_i) \right] p(X_p | a)$$

where $X_1, X_2, \ldots, X_p$ are the multispectral measurement vectors comprising $\underline{X}_{ij}$.

The expression shown above provides a set of discriminant functions for the classification problem.

The optimal choice of $d(\cdot)$ cannot actually be determined because it depends on $G(\underline{\theta}^p)$ which is unknown. We can, however, expect that, at least for large $N = N_1 x N_2$, a decision rule in which $G(\underline{\theta}^p)$ is replaced by an estimate of $G(\underline{\theta}^p)$ based on the $\underline{X}_{ij}$ will have risk approximating that of the optimal rule. In the interest of a practical solution to the problem, estimates of $G(\underline{\theta}^p)$ were derived by simply counting the occurrences of each $\underline{\theta}^p$ obtained in a classification of the scene without the use of context. Our experimental results have demonstrated the effectiveness of this "bootstrap" procedure.

Given this approach to classification using contextual information, it remained to show that a contextual classifier could indeed produce significant improvement in classification performance. Furthermore, since it was clear that a great deal of computation is involved in the method, investigation of a multiprocessor implementation of the classifier, using advanced computer technology, was deemed essential. Our work during the contract year has dealt with both of these issues.

2.3 Experimental Results

Experiments were performed to explore the effectiveness of contextual classification as applied to the analysis of multispectral remote sensing data. First, simulated data were used to determine the degree to which contextual classification might improve the analysis results (as compared to no-context classification), given that the class-conditional densities and the context distribution for the scene were known. The simulated data were used again to investigate candidate methods for estimating the context distribution since, as noted above, it usually cannot be assumed that the context distribution is known a priori. Finally, contextual classification was applied to real data to determine the extent to which the conclusions drawn from the simulated-data experiments could be extended to the more realistic case.

Details of all experimental results are presented in the report. Briefly, the simulated data experiments have demonstrated that contextual information can produce a dramatic improvement in classification accuracy. See Figure 3A-5. The results are, however, sensitive to the accuracy with which the p-context distribution is estimated, and although we have developed a "bootstrap" method for performing this estimation (Figure 3A-6) experiment with real data have shown that further investigation of this problem is necessary. Initial results with a new distribution estimation procedure appear promising (Figure 3A-7).

2.4 CDC Flexible Processor System

Classification algorithms such as the contextual classifier (and even much simpler algorithms used for remote sensing data analysis) typically require large amounts of computation time. One way to reduce the execution time of these tasks is through the use of parallelism. Various parallel processing systems that can be used for remote sensing have been built or proposed. The Control Data Corporation Flexible Processor System is a commercially available multiprocessor system which has been recommended for use in remote sensing.

The basic components of a Flexible Processor (FP) are shown in Figure 3A-8. Each FP is microprogrammed, allowing parallelism at the instruction level. An example of the way in which N FPs may be configured into a system is shown in Figure 3A-9. There can be up to 16 FPs linked together, providing much parallelism at the processor level. The FPs can communicate

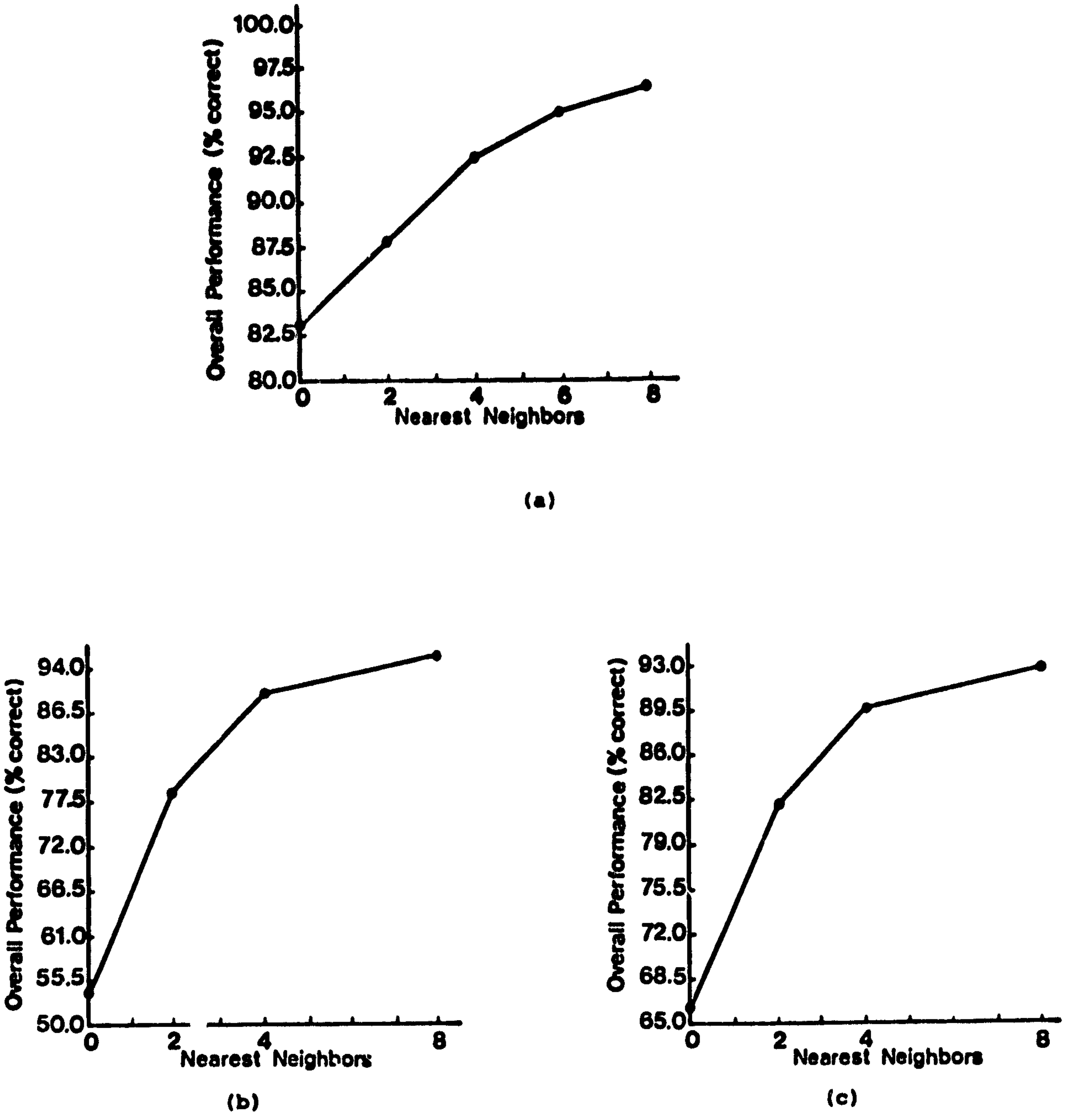


Figure 3A-5. Contextual classification of simulated data.
(a) Data set 1. (b) Data set 2a. (c) Data set 2b.

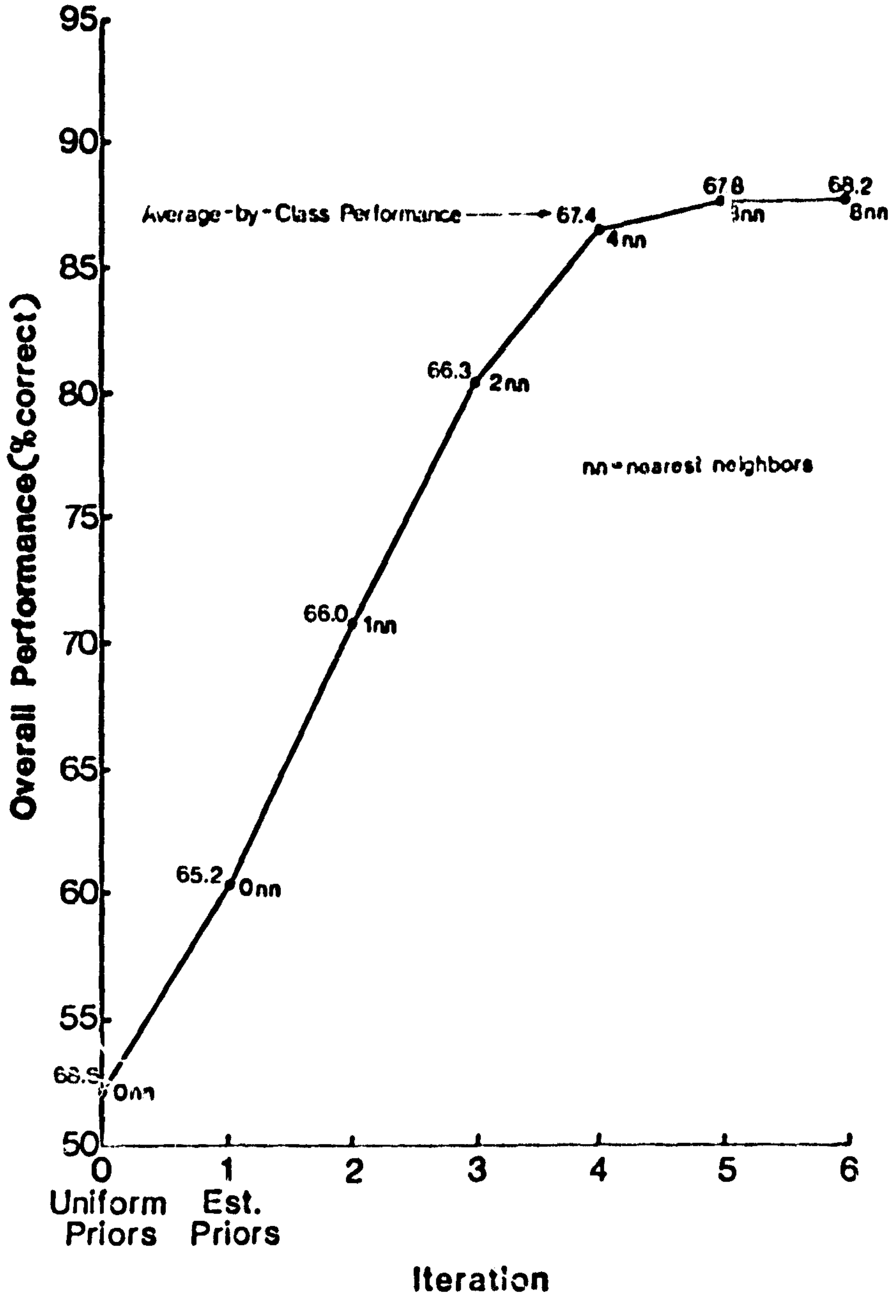


Figure 3A-6. Contextual classification of simulated data based on simplified iterative technique (simulated data set 2a).

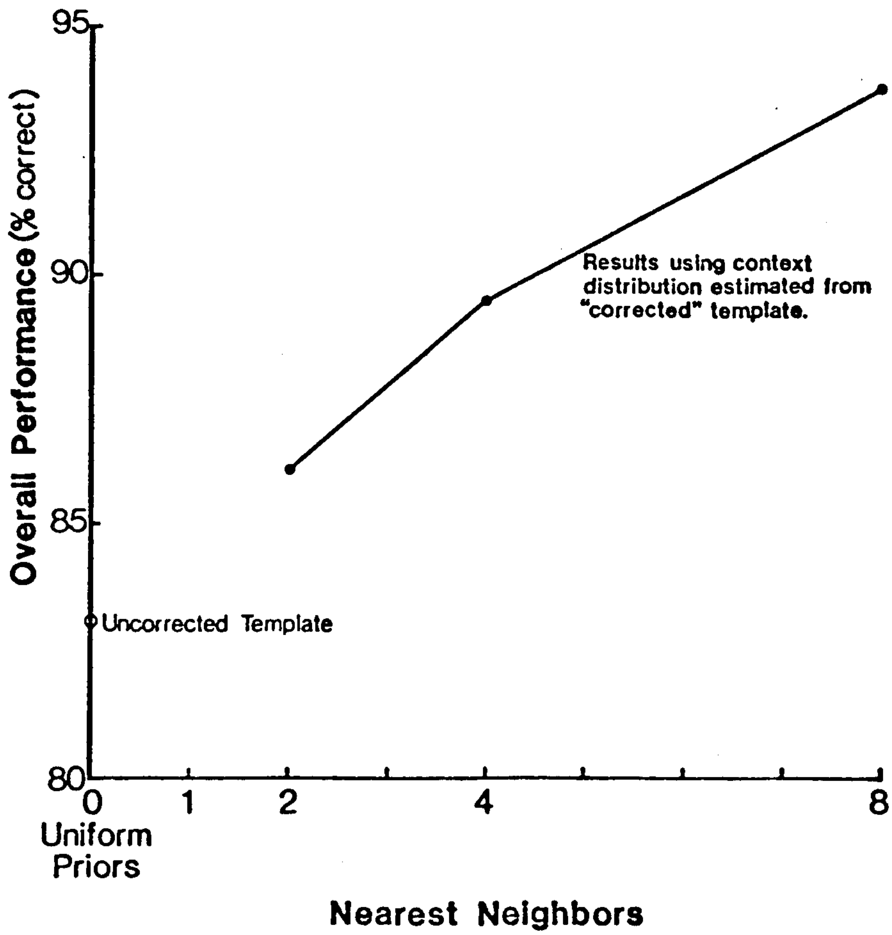


Figure 3A-7. Performance on real data using manual template correction for estimating the context distribution.

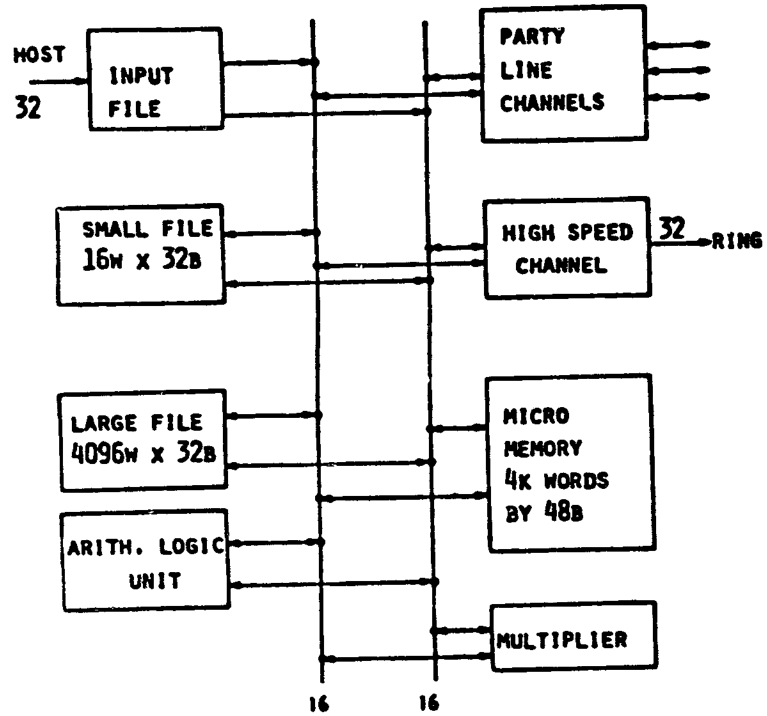


Figure 3A-8. **Data path organization in the CDC Flexible Processor.**

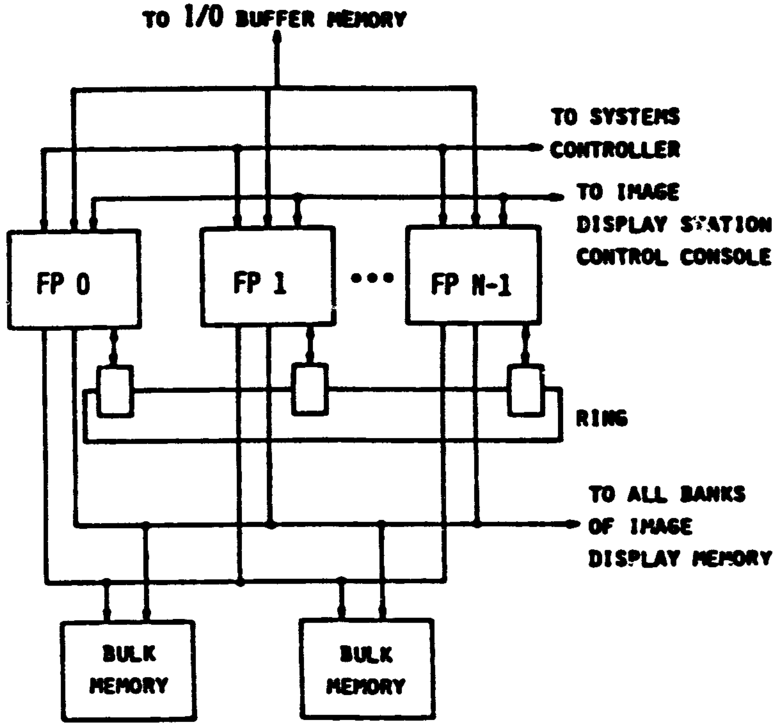


Figure 3A-9. **Block diagram of typical Flexible Processor array.**

among themselves through the high-speed ring or shared bulk memory. The clock cycle time of each FP is 125 nsec (nanoseconds). Since 16 FPs can be connected in a parallel and/or pipelined fashion, the effective throughput can be drastically increased, resulting in a potential effective cycle time of less than 10 nsec.

FP is programmed in "micro-assembly language," allowing parallelism at the instruction level. For example, it is possible to conditionally increment an index register, do a program jump, multiply two 8-bit integers, and add two 32-bit integers, all simultaneously. This type of operational overlap, in conjunction with the multiprocessing capability of the FPs, greatly increases the speed of the FP array.

The following list summarizes the important architectural features of an FP:

User microprogrammable;
Dual 16-bit internal bus system;
Able to operate with either 16- or 32-bit words;
125 nsec clock cycle;
125 nsec time to add two 32-bit integers;
250 nsec time to multiply two 8-bit integers;
Register file (with 60 nsec access time) of over 8,000 16-bit words.

2.5 Maximum Likelihood Classification on a Flexible Processor System

The pointwise maximum likelihood classification (MLC) of pixels using an FP array is discussed below. The contextual classifier performs computations similar to those used by the maximum likelihood classifier, but is complicated by the involvement of "neighboring" pixels. The analysis approach that has been taken is to first investigate the implementation of the maximum likelihood classifier and then extend the results to the contextual classifier.

Assume there are N FPs in the array. Then, one way to implement the MLC algorithm is to have each FP process 1/N of the classes. For example, consider the case where there are N classes. The host initially sends the inverse covariance matrix, the mean vector, and constants for class i to the ith FP, $0 \leq i < N$. The host then sends the current data vector X to FP 0, then to FP 1, FP 2, etc. As soon as the FP receives the data vector, it begins the calculation of the value of the discriminant function. After the host gives all FPs the data for pixel (i,j), it waits until FP 0 has calculated the value for its discriminant function. The host then retrieves the value of the discriminant function and loads FP 0 with the data vector for the next pixel. The host executes this process for all the FPs. When the last FP has transmitted the result, the host does a compare and stores the class index corresponding to the maximum of the discriminant values computed for this pixel. Thus, the compares are done by the host while the FPs are computing the discriminant functions for the pixel, minimizing delay. Nearly a factor of N improvement over using a single FP can be obtained.

Another way to approach the problem is to divide the image into N subimages and have each FP perform the complete maximum likelihood classification for all pixels in its subimage. This is shown in Figure 3A-10 for an A-by-B image. Again, nearly a factor of N improvement is possible. The choice between these two implementation methods will depend on such factors as the processing capabilities of the host, the number of classes, and the size of the image.

The maximum likelihood classifier has been programmed on a simulator for a FP array. The simulator displays the contents of the main registers and provides a variety of tools for debugging microcode. Preliminary tests indicate that a <u>single</u> FP will perform a maximum likelihood classification faster than a PDP-11/70.

The experience gained through the use of the simulator has made evident the following advantages and disadvantages of the system.

Advantages:

Multiple processors (up to 16)

User microprogrammable -- parallelism at the instruction level

Connection ring for inter-FP communications

Shared bulk memory units

Separate arithmetic logic unit and hardware multiply

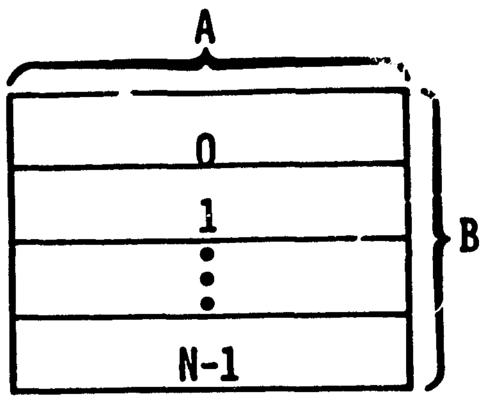


Fig. 3A-10. An A by B image divided among N Flexible Processors.

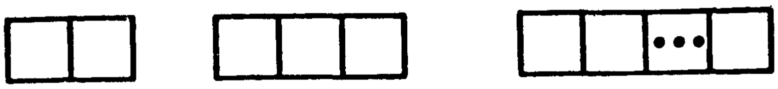


Fig. 3A-11. Horizontally linear neighborhoods. Each box is one pixel.

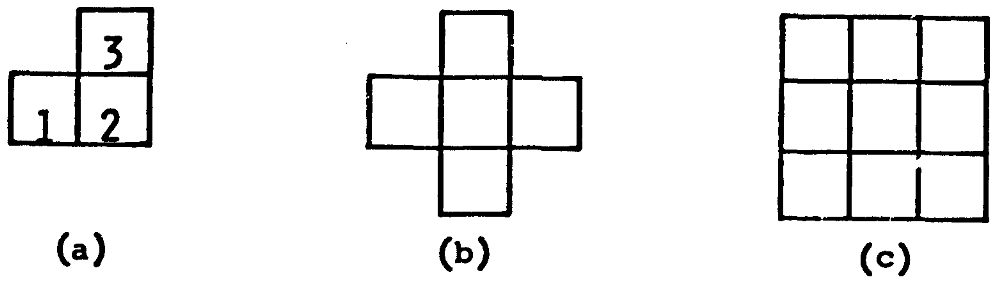


Fig. 3A-12. Nonlinear neighborhoods. Each box is one pixel.

Disadvantages:

No floating point hardware
Micro-assembly language -- difficult to program
Program memory limited to 4k microinstructions

2.6 Contextual Classification on a Flexible Processor System

Consider the implementation of a contextual classifier on an array of FPs. Assume the neighborhood is horizontally linear, as shown in Figure 3A-11. Divide the A-by-B image into subimages of B/N rows A pixels long, as shown in Figure 3A-10. The entire neighborhood of each pixel is included in its subimage. Therefore, nearly a factor of N improvement is attained. Vertically linear and diagonally linear neighborhoods can be processed similarly.

Consider nonlinear neighborhoods, that is, neighborhoods which do not fit into one of the linear classes. For example, all of the neighborhoods in Figure 3A-12 are nonlinear. Figure 3A-12(a) and its rotations represent the simplest nonlinear neighborhood. It is included in all other nonlinear neighborhoods. Thus, that neighborhood is called the nonlinear kernel neighborhood.

It can be shown that there is no way to partition an A-by-B image into N (not necessarily equal) sections such that a contextual classifier using a nonlinear neighborhood can be performed without data transfers among FPs. This will be demonstrated for the nonlinear kernel and will thus be true for all nonlinear neighborhoods. There are three cases to consider. If there is a vertical border between two subimages stored in different FPs, then pixels 1 and 2 in Figure 3A-12(a) will be in different FPs. If there is a horizontal border, pixels 2 and 3 will be in different FPs. If there is a diagonal border, pixels 1 and 2 will be in different FPs. The way in which to assign pixels to FPs in order to minimize computation time will depend upon the particular image size, number of FPs used, the time required for inter-FP communications, and the shape and size of the neighborhood. These factors will also determine the effectiveness of the use of the FP array for performing context classifications based on a given neighborhood.

2.7 Summary and Concluding Remarks

During this contract year, notable progress has been achieved with respect to the research objectives set out for this task. Specifically:

1. Procedures have been investigated for determining and representing the contextual information in a given scene. The performance of the contextual classifier is found to be sensitive to the accuracy with which the p-context distribution is estimated. Although good results have been achieved, both with real and simulated data, further work is needed on methods for determining the context distribution.

2. The contextual classifier algorithm has been analyzed with respect to achieving efficient implementation on a multiprocessor system. It has been shown that under rather severe restrictions on the shape of the contextual neighborhood, an "ideal" speedup by a factor of N, for an N-processor system, can be achieved. Easing of these restrictions definitely incurs a cost in terms of computation time, the details of which are the subject of ongoing analysis.

3. Actual implementation of the contextual classifier on multiprocessor systems has been limited to development of a simulator for the CDC Flexible Processor Array System and implementation, on the simulator, of a maximum likelihood classifier. Computations performed by the maximum likelihood classifier are identical to many of the computations required for the contextual classifer, but the overall algorithm is considerably simpler. Thus, implementing the maximum likelihood classifier provided a useful means for beginning to learn how to program a Flexible Processor Array System.

In support of the above achievements, the mathematical formulation of the contextual classifier has been put on firmer ground and some insights gained into the nature of the spatial context. A significant amount of effort has gone into understanding the architectural details of the Flexible Processor, in order to use its facilities effectively.

At this point in the study, we may conclude that the contextual classifier does indeed lead to improved classification accuracy by utilizing spatial context information in multispectral earth resources data. Although the computational demands of the proposed contextual classifier are substantial, multiprocessor systems such as the CDC Flexible Processor Array System can be used to achieve efficient implementation of this and other image processing algorithms.

Ongoing research in connection with this project will be directed toward better understanding the nature of contextual information in multispectral image data and exploiting the computational efficiencies to be gained through parallelism and other special features of advanced data processing system architectures.

3. Ambiguity Reduction for Training Sample Labeling

The proper training of the classifier in a remote sensing data analysis system is one of the pivotal steps to good system performance. The original method used for training classifiers was to define a set of classes based on user need, then to choose an adequate set of prelabeled sample pixels of these classes by which to compute class statistics. Because it was assumed that the labels would be established by ground observation, they were always assumed perfectly accurate.

However, in some application situations, ground observations (or at least observations from the ground) are not always possible. Thus, cases arise where the labels associated with training pixels are not entirely accurate.

In any application situation, there nearly always exists a wide assortment of ancillary information, some of which is subjective in nature, other objective, which should be able to materially contribute to the accuracy of such a pixel labeling process. Examples are data about the terrain, weather and climate, seasonal characteristics and the spatial context of pixels. The question is what mechanism should be used to incorporate such information into the labeling process. Thus, the objective for the current work is:

> To devise and evaluate quantitative and objective means for optimally arriving at classifier training sets using remotely

sensed spectral observations together with any other types of ancillary data and knowledge which may become available.

In selecting an approach to pursue this objective, it is important to note that the ancillary data to be used are varied in type and not well defined. The information content of such data may be known only somewhat vaguely a priori, and in some cases it will certainly be difficult to quantify. Such a situation suggests defining an approach which, instead of being based on a direct deterministic calculation, might be iterative in nature so as to provide a "convergence of evidence."

A problem in the field of picture processing with somewhat similar characteristics is being studied using a method, known as relaxation, which is iterative in character. It was therefore decided to study this approach to see if it might be adaptable to the problem at hand. Basically, the idea would be to use the ancillary information to reduce any ambiguity which might be associated with a given label on a given pixel. At the outset there would exist an exhaustive list of labels and a set of pixels with a (preliminary) label association for each. There would be a measure of certainty of the accuracy of this association in quantitative form. The process would then be one of utilizing the ancillary information in an iterative fashion to cause reinforcement of the degree of certainty for the correct label of the pixel at the expense of all of the incorrect ones.

This task is a new one, begun only two months before the end of this contract period. The work so far has consisted of a survey of available literature (compiled in the annual report) and the formulation of a more detailed technical approach.

3B. MULTISENSOR, MULTIDATE SPATIAL FEATURE MATCHING, CORRELATION REGISTRATION, RESAMPLING AND INFORMATION EXTRACTION

Paul E. Anuta

The research completed under this task constitutes the second and final year of a research investigation into the problems of combining and utilizing multiple data types for remote sensing surveys. The study considered the merging of different remote sensing data types, information extraction from the combined data and digitization and merging of ancillary data. The multidata-merging problem was explored and results reported in the final report of the first year [1]. Information extraction of merged Landsat and SAR data is discussed in this, the second-year, report as well as ancillary data digitization. A new concept for a multidata merging system emerged from the study.

Registration of the Landsat MSS and SAR data types was studied in detail in the first year of the project and results were reported in the final report issued in November 1978. The merged SAR/MSS data set formed the basis of research done in the current year.

The Landsat data are from frame 5-792-16152 imaged on June 19, 1977 over Phoenix, Arizona. A 512 by 512-point grid was defined over the crop area between Sun City and Phoenix, AZ, at the 25 meter resolution. The Landsat and SAR data were registered to this grid, using the LARS registration system and results of the previous year's registration study [1]. The agricultural "scene" analyzed consisted primarily of cotton fields with urban encroachment by Phoenix on the east and Sun City on the west. In order to simulate a segment size area, a 3 by 5-mile block was selected from the agricultural area. Within the 3 by 5-mile segment containing 15 sections, there were 76 ground-truthed fields. The contents of these fields and the number of pixels in each are indicated in Table 3B-1.

The cluster block approach was taken in training the classifier. Blocks of pixels containing samples from each class were clustered, using the LARSYS * CLUSTER routine.

Table 3B-1. Classes Analyzed in SAR/Landsat Data

Class	No. of Fields	No. of Pixels
Cotton	40	9377
Alfalfa	12	3345
Barley	2	364
Urban	22	3968
TOTAL	76	17,054

The Landsat/SAR data set was classified using both pixel and field classifiers and using Landsat only and Landsat plus SAR bands. The results of these tests are presented in Table 3B-2. The best overall results were obtained using the field classifier and spectral data only. Addition of the SAR channel reduced test classification accuracy in most cases, except alfalfa and barley. In general, the SAR seems to reduce separability of the spectral classes and it would appear that direct addition of this particular SAR data to the spectral data is undesirable.

The availability of high resolution imagery of the earth scene from the satellite platform provides the opportunity to employ scene structure as an input to the classification process. The basic concept is to use a single channel high resolution image as a mapping band to define scene structure and to then use spectral samples from within the objects in the scene for classification of those objects. Using this approach, the results in Table 3B-2 can be reinterpreted in terms of knowing all field boundaries in the scene. These would be obtained from high resolution SAR, RBV, MLA or any source of current imagery of the scene to be analyzed.

In the experiments carried out here, both field and pixel classification was carried out for the SAR Landsat data set. The field classifier results using spectral data were seen to be better than the pixel results but neither was very good. Knowing the field boundaries allows the results of pixel classifixation to be analyzed according to majority or plurality rules. In this approach, the class having a majority or a

Table 3B-2. Classification Results for Phoenix Site Test Fields, % Correct.

Class	Spectral Pixel Classifier	Spectral Field Classifier	Spectral + SAR Pixel	Spectral + SAR Field	Spectral Majority Classifier	Spectral Plurality Classifier
Cotton	70.2	87.6	48.1	43.3	95.2	100.0
Alfalfa	59.8	35.5	67.8	80.7	60.1	93.3
Barley	63.5	42.3	20.3	55.2	100.0	100.0
Urban	46.9	55.0	28.5	49.3	25.9	84.7
Overall	62.6	68.8	46.8	52.3	72.3	95.1

plurality is assumed to be the correct class for all points in the field. These results are significantly better than for the individual pixel or field classifier results. Thus the plurality rule applied to pixel classifier results for the case of known homogeneous fields appears to be an attractive approach.

The basic problem motivating this project was the combination of dissimilar data types to enable coordinated digital analysis for various applications. The concept of a self-defining data set (SDDS) was set forth as an approach to solving the problem of merging arbitrary areas from diverse data types.

To test the SDDS idea, a basic software system was developed in this task which carries out the basic functions needed. The problem has two parts: (1) Generating an SDDS, and (2) combining SDDS's to form a merged data set in a user-defined coordinate reference.

The data-merging system developed was tested on one data set for the study. The Landsat frame from the Picayune, Miss., area was selected since it was the first frame for which both fully geometrically corrected and uncorrected data were available. The control point finding and regression modeling portions were tested on the uncorrected frame. Both data forms, after conversion to SDDS format, were processed to produce a north-oriented data set for one 1:24000 USGS quadrangle.

The multidata merging and information evaluation task investigated several aspects of utilization of different data types for remote sensing surveys. The primary topics studies were: (1) Merging of different types of remote sensing data, (2) digitization and merging of ancillary data, and (3) information extraction from the combined data sets. Due to difficulties in obtaining data, only Landsat and synthetic aperture radar data types were studied. Digitization and merging of color map ancillary data sources were studied and a color classification method was validated. A self-defining digital data set structure was defined and implemented to facilitate merging of different data types.

1. Anuta, P. E., Hixson, M. M. and Swain, P. H., "Vol. III Processing Techniques Development," Final Report of NASA Contract NAS9-15466, Contract Report 112778, Laboratory for Applications of Remote Sensing (LARS), Purdue University, Nov. 1978.

4. COMPUTER PROCESSING AND DATA BASE SERVICES

James L. Kast

1. Background

The Computer Processing Support Task enables members of Johnson Space Center's Research Community to access a data processing environment appropriate for researching remote sensing of agriculture. The ingredients of a shared data processing facility provided directly by the Computer Processing Support Task have included:

*Access to a computer facility designed and implemented to support remote sensing research needs;
*Training in the use of hardware and software facilities of the computer system;
*Consulting support to remote users of the facility;
*Software development supporting and enhancing the utility of the shared facility;
*Identification and investigation of the potential benefits which may be derived when geographically dispersed research centers working on the same problem (researching remote sensing of agriculture) share a computational environment.

The full implementation of a shared data processing environment (either in the form of a network or a shared central facility) will provide the following potential benefits;

*The opportunity to better mold geographically dispersed research groups into a more informed and coordinated research team;
*A mechanism for efficient transfer of information between research centers, NASA, and other participating government agencies;
*Faster, less redundant software development;
*Faster transfer of newly developed analysis techniques and research results to and from participating research groups;

*Concentration of systems programming, data acquisition, data base, and certain computer services at a small number of locations (frequently one).

To build a suitable environment for the implementation, evaluation and exchange for remote sensing data processing techniques, Purdue has concentrated its efforts in providing access to suitable hardware, software, utilities and data. Purdue is also providing consulting and training support to foster user knowledge of the facilities available on the shared system.

2. 1979 Activities

2.1 Hardware Components

The most notable hardware change during 1979 was the replacement of the IBM 370/148 at Purdue with an IBM 3031 machine. Since June, 1978 IBM System 370, Model 148's work load was well above its capacity during the day and evening shifts, leading to extremely slow response times and people waiting for the computer. LARS attempted to more evenly distribute the computer load throughout the 24 hour day, encourage more efficient use of the computer, investigate system bottlenecks and investigate hardware alternatives. Several actions were taken which were successful in providing more even distribution of the computer load throughout the day, in more efficient use of the computer and in the identification of system bottlenecks; however, they failed to impact the day and evening shifts saturation problems. Therefore, the decision was reached to pursue the acquisition of an IBM 3031.

Installation of a 3031 was expected to provide:

*More efficient computer users after installation due to greatly improved response time;
*No major applications software conversions;
*Capacity for additional computer projects;
*Minimal down time during switch over;
*A need to secure continued funding for LARS computational facility at FY79 levels.

Benchmark tests run on the 3031 indicated at a 3 to 1 improvement in performance could be expected over the 148. The 3031 was up and ready for general use on Monday, September 10 at 8 am, one day ahead of schedule.

During the second and third quarters of 1979 the IBM 3330 capatible disk units were installed. These units provided storage space needed by the SRT data base, CSMP and the expanded needs of the JSC user community. These drives also serve as replacements for the 2314 drives which are being phased out. Figure 4-1 and 4-2 present the configuration of the Purdue/LARS computer as of 12/1/78 and 12/1/79, respectively.

2.2 Systems Access

During the second quarter of 1979, the IBM 2780 printer/reader/punch terminal at JSC was replaced by a Data 100 having a tape transfer capability. The Data 100 functions as a HASP work station using HASP protocol which is more efficient than the protocol used by the 2780. In addition, the Data 100 is a more reliable piece of equipment.

Two IBM 2741 keyboard terminals which work at 13.8 characters per second, were replaced during the first quarter by Trendata terminals functioning at 30 characters per second. During the fourth quarter the Trendata terminals were upgraded to 120 character per second operation. The elimination of the 2741's made it possible to replace the clock dedicated to them with a 1200 baud clock. This in turn, made possible the installation of the Statistical Multiplexor and more efficient use of the bandwidth dedicated to keyboard terminal communication. During the first quarter, the Environment Research Institute of Michigan (ERIM) was provided printer and card reader as well as dial-up keyboard access to the LARS facility.

During the fourth quarter, a statistical multiplexor was installed and the two Trendata terminals at JSC were upgraded to 120 character per second operation. The statistical multiplexor allows the device utilizing the modem to make maximal use of the available bandwidth upon demand, rather than dedicating the specific amount of bandwidth to each device.

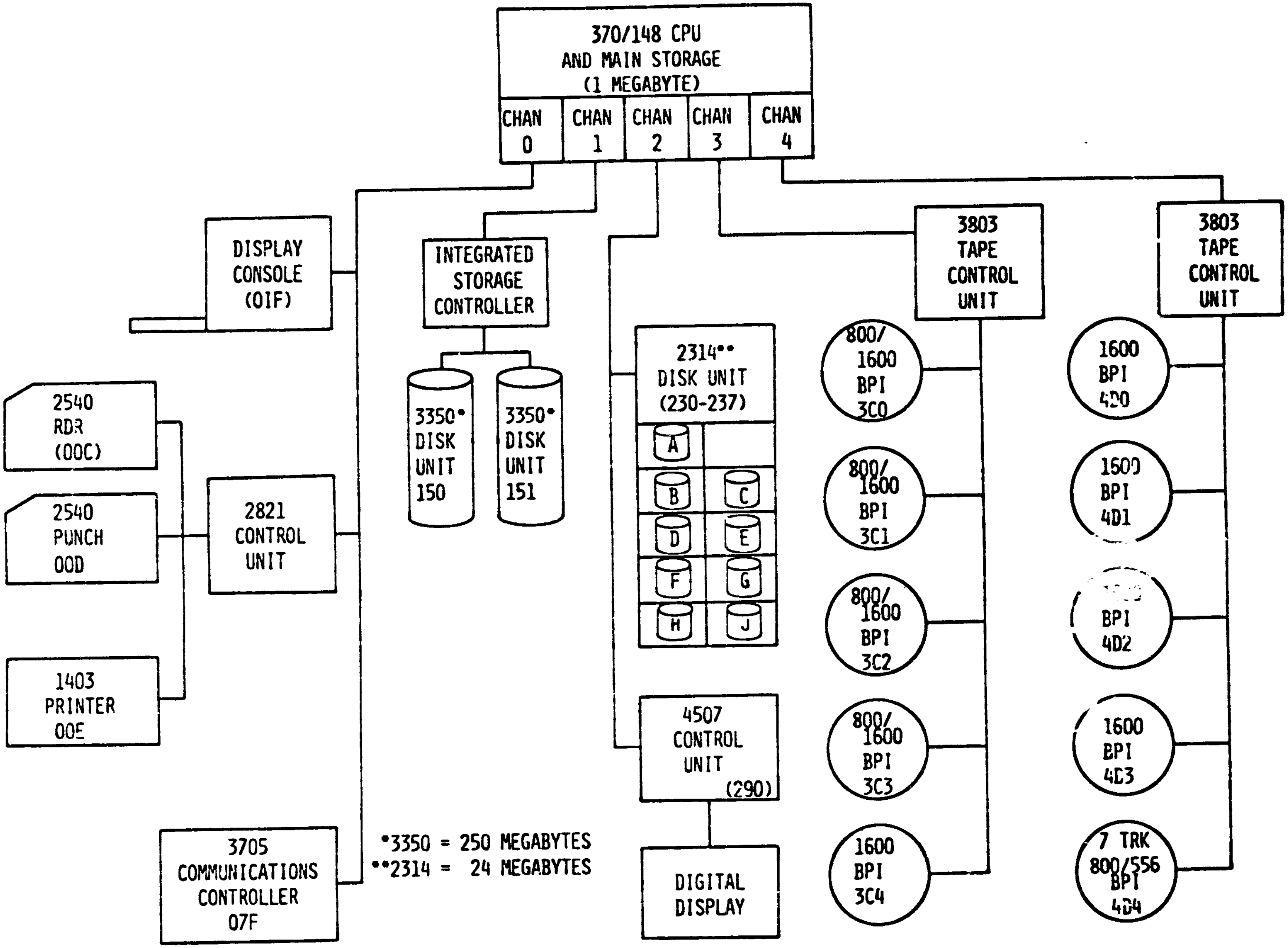
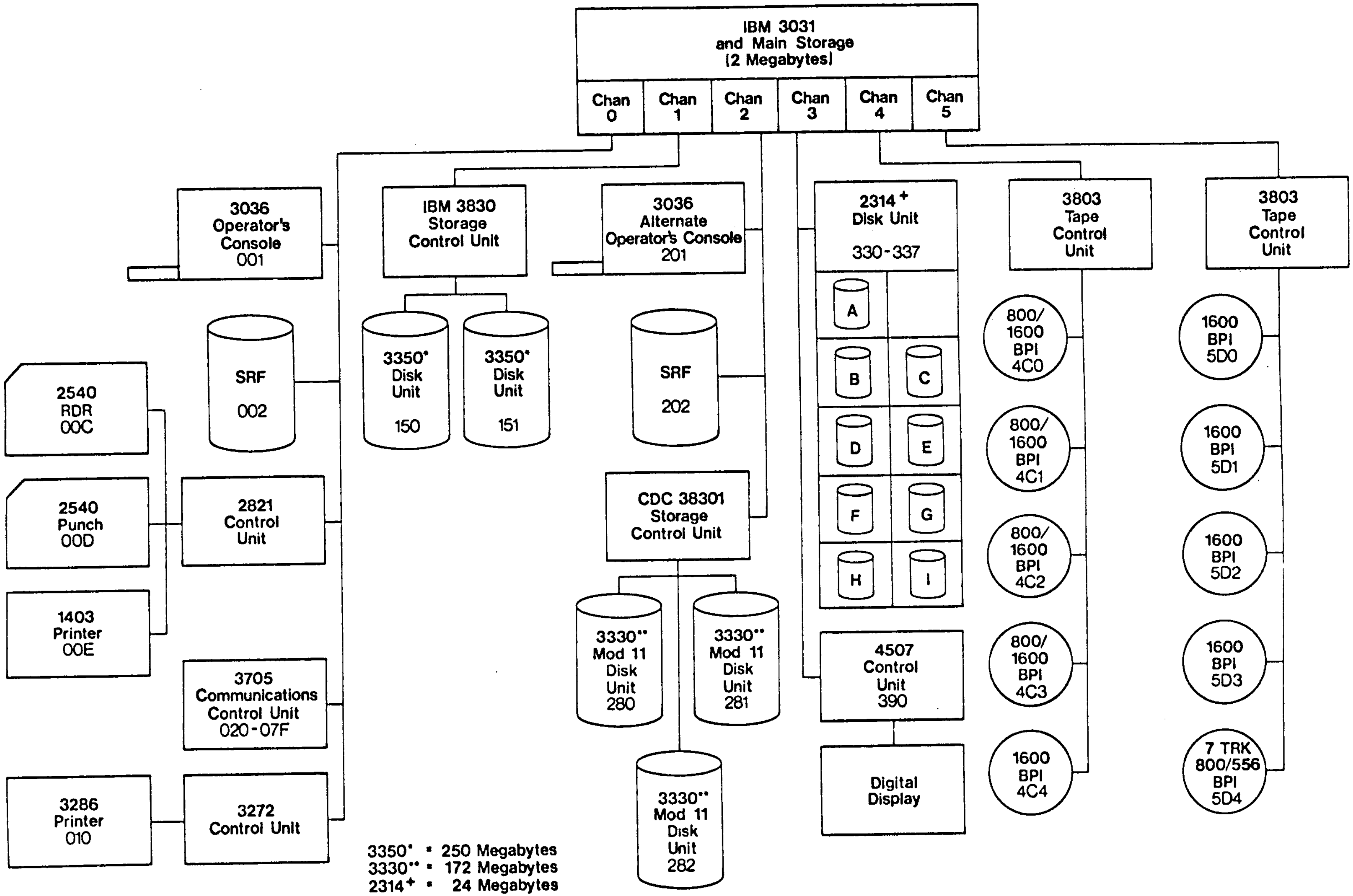


Figure 4-1. Purdue/LARS computer configuration as of 12/1/78.

3031 Hardware Configuration



3350* = 250 Megabytes
3330** = 172 Megabytes
2314+ = 24 Megabytes

Figure 4-2. Purdue/LARS computer configuration as of 12/1/79.

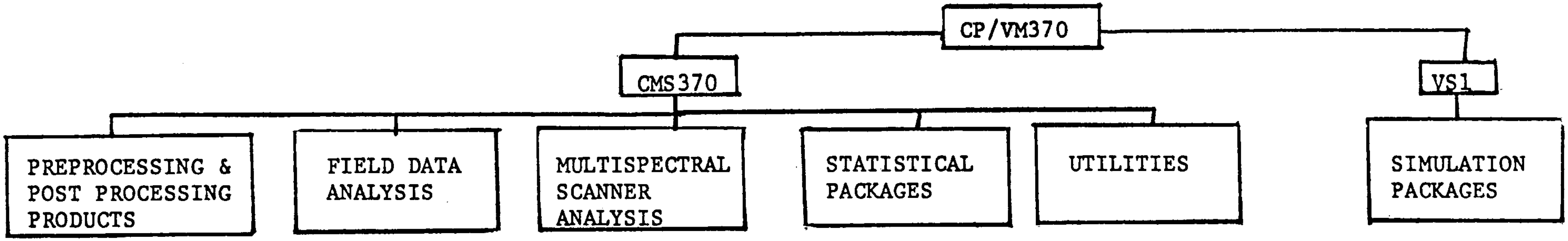
One substantial cost of a shared data processing system is the charge for the communications line (long distance or dedicated line charge). During FY79, LARS investigated the relative cost of the private network now operating to those of telecommunications facility vendors, hoping to find a way to reduce these line costs. Results of this investigation indicate that at the current time no significant cost reductions to LARS and its user community are possible. Packet switching networks will be re-evaluated when telecommunication vendors begin offering synchronous communications as part of their service.

2.3 Software Development

Figure 4-3 lists the software which has been implemented on a LARS computer. Software development supporting users of the shared system during 1979 included:

*An on-line problem reporting system (TROUBLE);
*A system allowing users to mail memos to each other (MAIL);
*An upgraded SR&T news facility to inform users of major changes to the shared data processing environment;
*The installation of SPSS Release 8;
*The installation of Addition 7 of IMSL;
*The installation of SAS;
*An upgrade to the tape transfer software (TAPTRAN);
*Development of four computer programs allowing communication between Purdue's LARSYS and JSC's EODLARSYS;
*Major upgrade to the LARS Spectral Analysis system (LARSPEC);
*Development of software for requesting computer resources (ID's, tapes, etc.);
*An upgrade to the Accounting-By-User-Group software;
*Modifications to the BATCH system and the addition of a new Batch machine (BATHOUST);
*The implementation of the 3-dimensional graphic system (3-D graphics compatibility system - GCS).

SOFTWARE AVAILABLE ON THE LARS COMPUTER



REFORMATTING	LARSPEC	EOD LARSYSP1	SPSS	CMS370 BATCH	CSMP
GEOMETRIC CORRECTIONS		LARSYS	BMD	EXECUTIVE CONTROL	
A/D CONVERSIONS		LARSYSDV	IMSL	EDITOR	
PHOTO PROCESSING		ECHO	SAS	FORTRAN G	
DATA DIGITIZATION		LAYER		FORTRAN H	
REGISTRATIONS		CLASSY		ASSEMBLER	
-IMAGE/IMAGE		AMOEBA		DEBUG PACKAGE	
-IMAGE/MAP		UNIFORM CHROMATICITY TRANSFORM		TAPE, DISK, CORE DUMPS	
-IMAGE/ANALLARY		LIST		TAPE TRANSFERS	
		GLMAOV		NEWS FILES	
		PROCEDURE M		GCS	
				DATA BASE ACCESS	
				ACCOUNTING	
				USER MAIL	
				TROUBLE REPORTING	
		PLANNED:		PLANNED:	
		EODLS DEVELOPMENT		NEW BATCH SYSTEM	

Figure 4-3. Software available on the LARS computer.

2.4 Data Base Management

To support the research needs for the SR&T community, the following data bases were acquired for the LARS computing system:

*LACIE PHASE I
*LACIE PHASE 3 Blind site ground Truth
*Multicrop
*Corrected Phase 3 Ground Truth
*Transition Year Foreign Data

These Data Bases were transmitted from JSC to Purdue. Upon receipt at Purdue, the data were inspected for proper blocking factor, file organization and readability. When a data base became complete and verified, it was then entered into the RT&E Segment Catalog. Users were notified of new data base installations through announcements in SRTNEWS and SCANLINES.

During 1979 the entire data base was checked for data integrity. During 1978 and early 1979 quality control checks of the tapes generated in Building 30 were insufficient. Discussion with JSC and LFC personnel nas resulted in a marked improvement in the reliability of the delivered tapes.

An unexpected problem arose with the discovery that geographical location associated with LACIE segment numbers were not always unique. Initially the segment catalog was organized with the segment number being the only master key. Since this arrangement could lead to geographic ambiguity, a new method of organization was devised including latitude and longitude with the segment number as master keys.

Prior to the installation of the IBM 3330 compatible disk drives, there was insufficient disk storage space to house the segment catalog. During March the full segment catalog was loaded onto the new 3330 drives and became generally accessible to all users of the shared system.

Data management software implemented during 1979 includes the programs to build the ground observation index for Ground Truth Data, the user routine GTINFO for queuing the Ground Truth Data Base, and the segment catalog editor. The segment catalog editor is a significant piece of software which aids in the maintenance of the data base.

2.5 Communications Consulting and Training

This section deals with the personnel services portion of the Computer Processing Support Task. A service-oriented support group is essential to the creation and maintenance of an efficient, effective and coordinated Data Processing environment.

Five visiting consultant trips took place during 1979. During these trips, experts from LARS computer facility helped users at JSC overcome problems they had encountered using the LARS system, evaluate approaches to software design and implementation. Visiting consultants also gain insight into remote user operating environments and resource needs, and review new system capabilities with JSC users during their visits.

A CMS Short Course was presented at JSC during February 1979 and a new modular CMS Short Course was designed and ready for presentation during December for users at JSC. A Hands-on demonstration of LARSPEC was presented to an ERIM representative during July 1979.

In April a communications study was conducted to evaluate the responsiveness of LARS personnel to user problems, identify the utility of means of communication with LARS System Services, and identify what users view is a major communication problem they have with LARS. The study included a survey of users at Purdue, JSC and other remote sights. 74% of those participating in the survey rated the overall responsiveness of the LARS staff to user problems as very good or good. The rating by users at JSC was even higher. In terms of effectiveness, JSC users rated personal contact (usually achieved by visiting consultant trips) highest, followed by phone calls, communication through the terminals, SCANLINES, memos, correspondence, SRTNEWS, and documentation. Problem areas were identified on a scale from three to zero, going from very significant to none. The results of the survey indicated that very few JSC users felt there was a significant communication problem with the LARS support staff. The problem with the highest weighted average was the inability to contact a person due to his absence from his desk. This problem had a weighted average of .87 among JSC respondents. The on-line TROUBLE reporting system and the user MAIL system were developed to help alleviate this problem. Tables 4-1 and 4-2 show the results of the survey on the overall responsiveness of System Services personnel to user problems and major communication problem areas.

3. Results

User statistics for the Computer Processing Support Project indicate the success of the shared system concept. Computer resources consumed for the year are nearly double those consumed during FY78 (see Table 4-3). Figure 4-4 is a graph of the 370/148 equivalent CPU hour usage for the Computer Processing Support Task from December 1977 to November 1979. Not only was a large increase experienced during 1979, but 690 148-equivalent CPU hours were consumed during August through November -- that's a rate of 2070 148-equivalent CPU hours per year (roughly 700 3031 CPU hours per year). Thus, coupled with the increase in usage at LARS, the decision to acquire 3031 appears to be justified.

The cost of a 370/148-based facility to LARS would have been 90% of the cost of the 3031 facility and would have been more than adequate for LARS' computational needs. The 3031 facility increased computer power available to the JSC-Purdue user community threefold while increasing costs only 10%. As a result of 3031 acquisition, the rate charged for computer time in November 1979 is approximately half the rate charged for computer time in December 1978.

Table 4-1. Survey of Computer Users Estimation of Overall Responsiveness to LARS Computer Support

	NUMBER OF RESPONDENTS	VERY GOOD	GOOD	ADEQUATE	NEEDS WORK	TOTALLY INADEQUATE
3A Users	16	7	6	2	0	1
LARS Users	15	4	7	2	2	0
System Services Staff	16	4	7	4	1	0
TOTALS	47	15	20	8	3	1

Table 4-2. Communication Problem Areas

PROBLEM AREA	MAGNITUDE OF PROBLEM/WEIGHT				
	Very 3	Some 2	Little 1	None 0	Weighted Average
1. Inability to contact person due to absence from desk.	2	0	7	6	.87
2. Information given is not adequate.	0	4	2	8	.71
3. Promises for services are made, then not kept.	1	1	4	8	.64
4. Misinformation is given.	1	1	2	9	.54

Write-in problems included:

- Don't know whom to contact should something go wrong.
- Long delays in placing long-distance calls.
- Too few people to interface with at JSC on a technical level.
- Inadequate IBM and LARS documentation available.

Table 4-3. Number of Users and CPU Usage by Computer Processing Support Task.

	Dec '77	Nov '78	Nov '79
JSC Users	26	71	96
LARS Support	3	12	19
Deleted ID's		8	3
Batch Machines		3	4
Library ID (JSC Disk)		1	1
Total ID's	29	95	123
370/148 CPU Hours*, year ending	70	694	1374

*Equivalent 370/148 CPU hours. 3031 CPU hours are multiplied by three to approximate 370/148 hours.

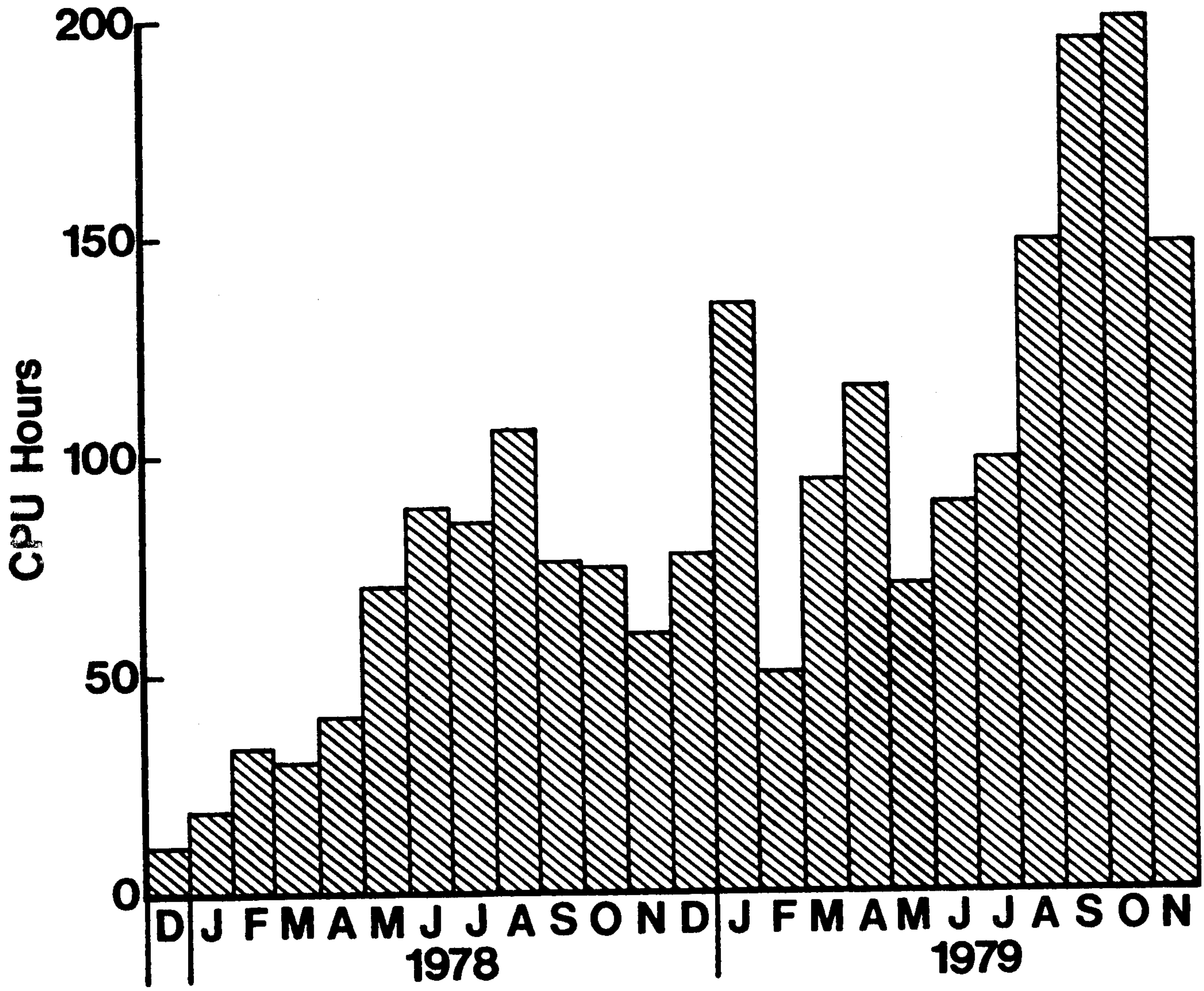


Figure 4-4. CPU usage by computer processing support task during 1978 and 1979 (370/148 equivalent CPU consumption).

4. Evaluation

The Computer Processing Support Task has been beneficial to the JSC research community. This is evidenced by the near doubling of computer use at JSC; the reduction of cost of computation for all users; the sharing of software and data by Purdue, NASA, IBM, Lockheed, and, to a much more limited extent, ERIM personnel; and the creation and use of certain computer user communication facilities.

The Computer Processing Task at LARS has serviced as a "pilot" for the concept of a shared SR&T computational environment. Such an environment could be supplied by centralized computer system or through a network of computers. In one way or another JSC will:

*Pay the bills for computer, personnel, and all other expenses incurred by all the members of the JSC-sponsored research community.
*Benefit from those fruitful new techniques which can successfully be integrated into pilot and LSAT analysis systems.

The Computer Processing Support Task has demonstrated that a shared computational environment can provide:

*User access at all user locations to data, software, and documentation contained in the shared environment;
*Sharing of expense of portion to the processing hardware at a cost advantage;
*Sharing of software allowing flexibility in software maintenance, addition, and updating at a cost advantage over independent, non-compatible systems; and
*Ease of training users, and sharing and comparing new techniques through standard data formats, terminology, and communication channels.

For example, users at JSC are making use of the field measurements data library and the LARSPEC software at Purdue without having to:

*copy, transport, and verify the copy of the data base for JSC users;
*convert the LARSPEC software to run on a different computer with a different operating system;
*wait or pay for the two items above to be accomplished;
*support the updating of JSC's local LARSPEC or field measurements data base each time it is updated at LARS; or
*constantly running verification test to make sure the two implementations remain functionally equivalent.

Similar statements could be made of LARS use of EODLARSYS and LARS' and JSC's use of SAS, SPSS, IMSL, LARSYS, the RT&E Data Base, CSMP, GCS, etc.

5. Recommendation

5.1 EOD Computer System Development

To meet the expanded computational needs placed on NASA's Earth Observation Division (EOD) by the AgRISTAR's project, EOD is pursuing the acquisition of computer. This computer is not anticipated to be large enough to supply the computational needs of JSC's entire research community including the universities and ERIM. The following recommendations are made for this system:

1. It should be as compatible as possible with the system at Purdue/LARS.

2. It should be used primarily for the development of configuration controlled pilot systems (system integration) and for the execution of pilot tests.

3. The LARS machine should be used for much of the research and techniques development work which will exceed the capacity of the EOD system.

4. The EOD and LARS systems should be networked so that the user at any site may elect to utilize either system.

5. This network should be available to users at all major research sites supporting EOD.

5.2 Data Processing Task Force

A research-community-wide data processing task force should be established to handle such problems as the development of programming conventions, documentation, and software delivery standards; mechanisms for new technique communication; and the definition of a baseline analysis system. This task force would be responsible for tackling community-wide data processing problems and identifying approaches to solutions to those problems. Topics such as local site support and computer user training courses would also be within this committee's perview.