123168

Volume 1 Number 4

Remote Multispectral Sensing in Agriculture

Laboratory for Agricultural Remote Sensing

Purdue University

CHAPTER 1

INTRODUCTION

Believe Report Represents Status
as of end of 1968
YJB 5/2021

Mentions research done in 1968 and
plans for 1969.

GENERAL

The research activities reported herein are a continuation of the programs initiated in 1966 at the Laboratory for Agricultural Remote Sensing (LARS) under the sponsorship of the National Aeronautics and Space Administration in cooperation with the United States Department of Agriculture and Purdue University. A principal objective of these programs continues to be the development of data acquisition, processing and distribution techniques for incorporation into earth resources information systems of the future. Timely, accurate, and comprehensive data are essential to prudent resource development planning and resource use management.

Planet Earth and its populations have entered into a new era--one marked by a growing awareness that the resources of the planet are limited; this in face of an ever-increasing demand for their use. Man is beginning to appreciate, upon consideration of predictable demands, the need for fixing limits on the utilization of atmosphere, water, soil, minerals, and chemicals. Man himself is one of the few resources which continue to increase in number. Agriculture constantly makes primary demand for the use of earth's resources. Improved resource development planning and day-by-day resource use management can be realized only if adequate information systems are developed through applications of modern technology to provide accurate, comprehensive, and timely information on resource uses, availability, productivity, and potential. The lack of such information is a major obstacle in the economic development of the undeveloped regions of the world and a significant obstacle in the formulation of important policies and programs in the more fully developed regions.

Consideration of the magnitude of the data loads which must be dealt with in such information systems has led to the institution of research programs to develop automatic recognition techniques for agricultural and forestry remote sensing applications. Information systems incorporating these techniques will be able to take advantage of the sublime data acquisition capabilities of aerospace platforms--aircraft and spacecraft.

Future information systems will depend on observations and measurements collected from the ground, air, and space. The observations collected by future earth-orbiting satellites will be supplemented by measurements and observations by aircraft systems. Ground observations will continue to be of value. Regardless of how data are required, there is a tremendous need to be able to process such data automatically for their information content. The major thrust of the research programs at Purdue University has been oriented toward meeting this requirement.

It is the opinion of responsible researchers at the Laboratory for Agricultural Remote Sensing that research funding continues to be somewhat inadequate to permit appropriate rate of progress. A more intensive research program is required to prepare a firm foundation for future anticipated needs.

PROGRAM OBJECTIVES AND PLAN OF WORK

The Laboratory for Agricultural Remote Sensing at Purdue University has been designated by the U. S. Department of Agriculture, in cooperation with the National Aeronautics and Space Administration, as a principal focal point in the scientific community for coordinating and conducting research to develop observation and measurements systems for the benefit of agriculture. LARS also has the specific responsibility to develop automatic recognition techniques for agricultural and forestry remote sensing applications. LARS is to diligently plan and carry out research on data collection, processing, and distribution techniques designed to provide a capability for the remote sensing of agricultural resources.

Investigations shall be made of spectral, temporal, and spatial reflectance and emmittance radiance characteristics of plants, soils, water, animals, and other agricultural features.

LARS shall conduct investigations in the following program areas:

- Biogeophysical remote sensing
- Measurements
- Data processing
- Agricultural requirements and applications
- Aerospace systems

BIOGEOPHYSICAL REMOTE SENSING PROGRAMS

Investigations directed toward establishing a capability for the detection, identification, and mapping of various vegetation and soil features of significant importance through the use of multispectral scanner systems and automatic data handling and processing techniques shall be continued. Relationships between spectral response patterns and properties of various agricultural materials and atmospheric conditions influencing these materials shall be investigated. The research effort shall involve the collection and investigation of a variety of data, including that obtained in the laboratory and field and from aircraft, and shall include scanner and spectrometer measurements, photographs, meteorological measurements, and field descriptions and measurements as follows:

Spectral Response Investigations--investigations conducted and designed toward the following objectives:

(1) Determination of various agricultural features which can be differentiated and identified on the basis of spectral, spatial, and temporal data

(2) Definition of wavebands in the electromagnetic spectrum which yield characteristic and consistent spatial, spectral, and temporal information to be used in identifying and characterizing important agricultural situations

(3) Evaluation of seasonal and geographic variations in multispectral response patterns

(4) Determination of the agronomic causes of variations in multispectral response.

Ground Truth Investigations--investigations conducted and designed toward the following objectives:

(1) Determination and identification of significant items of ground truth data to be obtained

(2) Determination of techniques

and instrumentation to be used in obtaining ground truth data of significant importance

(3) Determination of relationships between plant and soil spectral measurements and micrometeorological data.

MEASUREMENTS PROGRAMS

Investigations shall be continued toward the conception, design, calibration and operation of instruments for the measurement of observable physical agricultural features.

The scope shall include laboratory, field, aircraft and projected satellite experiments designed to define, measure, and utilize electromagnetic radiation phenomena in the timely identification of agricultural and vegetative conditions. The objectives are:

(1) Investigation of design proposals for measurement apparatus, particularly for multichannel single-aperture airborne scanners and field spectrometers of compatible data format.

(2) Determination of agricultural satellite requirements with regard to bandwidth and resolution choices in photographic and electro-optic sensing instruments.

(3) Provision for the development of a mobile ground truth data collection system complete with field spectroscopy instrumentation

(4) Construction of laboratory apparatus capable of determining reflectance and transmittance variations in different plant parts due to various induced plant conditions such as moisture and nutrient stress

(5) Investigations of photoemulsion calibration techniques and requirements sufficient to assure consistent analysis of panchromatic, color, and false color imagery.

DATA PROCESSING PROGRAMS

Research in data processing methods for multispectral remote sensing in agriculture will be established and carried out.

Specifically, LARS will use, maintain, and improve the current data handling computation system in agriculture and such other disciplines as USDA/NASA specify.

LARS will improve the operational program for the Phase I data handling system so that:

(1) Data from the Michigan scanner system can be calibrated, using the most recent calibration procedures

(2) Data from the interferometers and other field-based instruments can be calibrated and reformated according to the needs of the user

(3) Ground truth data can be stored in a near optimum fashion

(4) All data can be made available for research according to desirable and efficient procedures.

LARS will study proposed remote sensing measurement techniques and participate in recommending specifications for systems capable of supplying quality data adequate and convenient for data processing.

LARS will remain cognizant of the data handling hardware/software system components technology to improve the man-data communication capability necessary for both research and operational systems. Recommendations will be made to the USDA to implement such technology as appears appropriate.

LARS will develop and study new pattern recognition techniques which show promise for agricultural applications. Areas of attention should include but are not limited to:

(1) Feature selection

(2) Training sample selection

(3) Delineation of categories

(4) Pattern classification by by algorithms applicable to high data volume.

LARS will use, maintain, and improve an operational pattern recognition research program to:

(1) Study the geographic variability of signatures

(2) Study operational promise of current techniques in agriculture and other disciplines

(3) Analyze problems of current techniques.

LARS will reduce and analyze data collected over agricultural areas in the Weslaco, Texas, Davis, California and other agricultural test site regions in cooperation with scientists who planned the data collection over the particular site.

AGRICULTURAL REQUIREMENTS AND APPLICATIONS PROGRAMS

Attempts shall be made to achieve the definition of information requirements in agriculture and the potential applications of remote sensing and automatic data processing in obtaining agricultural data to:

(1) Study information needs of the farmer-stockman-producers, government agencies, industry, research scientists, and international agricultural development agencies

(2) Compile and categorize a list of the kinds of information needed, when it is required and how it is used

(3) Organize information of this study in anticipation of the publication of a monograph on "Applications of Remote Sensing in Agriculture."

Attempts will also be made to achieve the extension of LARS capabilities for obtaining and analyzing data into other areas of agricultural research such as:

(1) Soil classification and survey

(2) Land-use planning and classification.

(3) Water resources research

(4) Plant ecology studies.

AEROSPACE SYSTEMS PROGRAMS

LARS will assist USDA and NASA to plan and conduct agricultural application experiments in the NASA/USDA earth resources aerospace programs. To do this LARS will:

(1) Assist USDA and NASA in developing mission plans of future aerospace systems in earth resources applications

(2) Integrate research results into aerospace application experiment plans which involve determination of equipment requirements for aircraft and spacecraft missions and equipment integration.

ORGANIZATION

The Laboratory for Agricultural Remote Sensing is interdisciplinary in its operation and in its organization. Co-directors of the Laboratory represent the School of Agriculture and the Schools of Engineering. The LARS staff is comprised of members of various departments of the schools of Purdue University.

The Laboratory for Agricultural Remote Sensing is organized as shown in Figures 1 through 6.

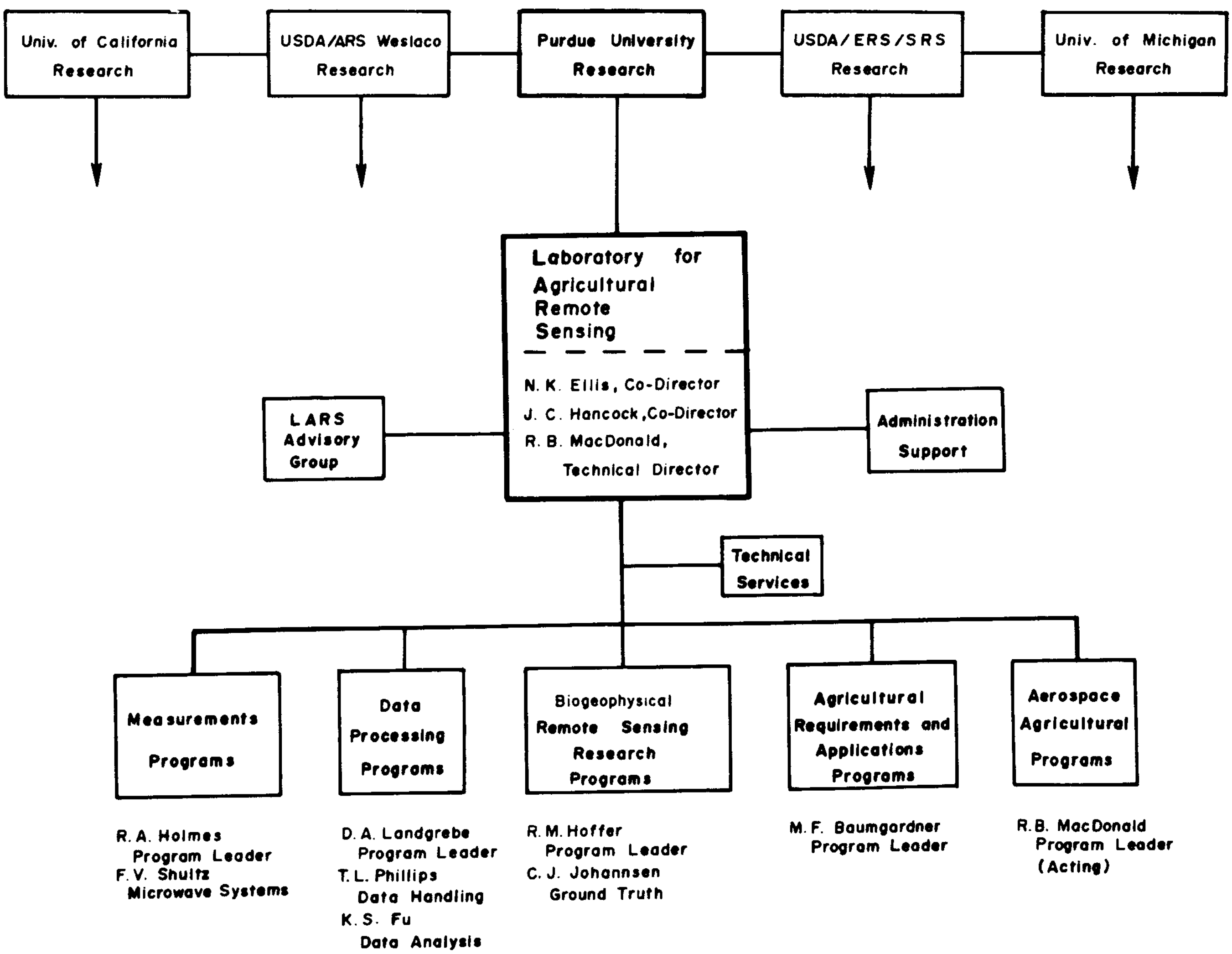


Figure 1. Organization of Laboratory for Agricultural Remote Sensing

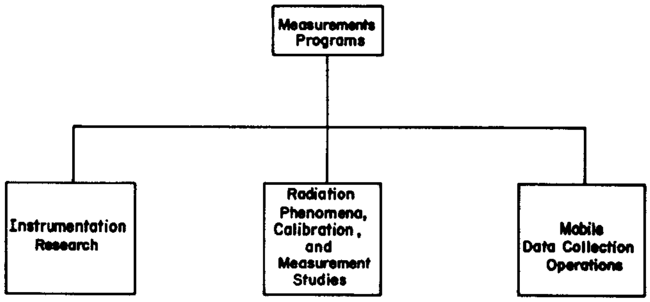


Figure 2. Measurements Program

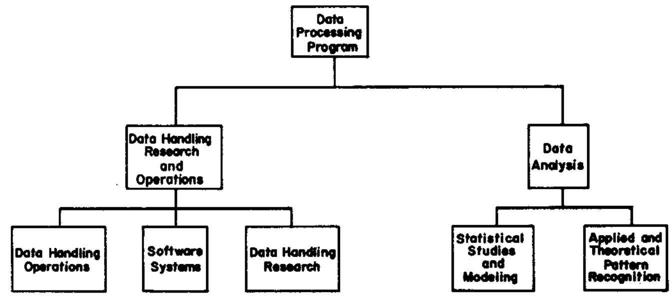


Figure 3. Data Processing Program

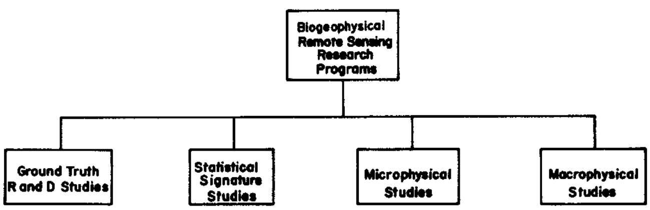


Figure 4. Agricultural Remote Sensing Research Programs

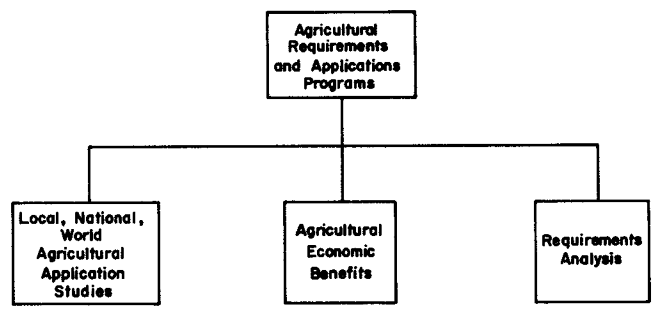


Figure 5. Agricultural Remote Sensing Requirements Program

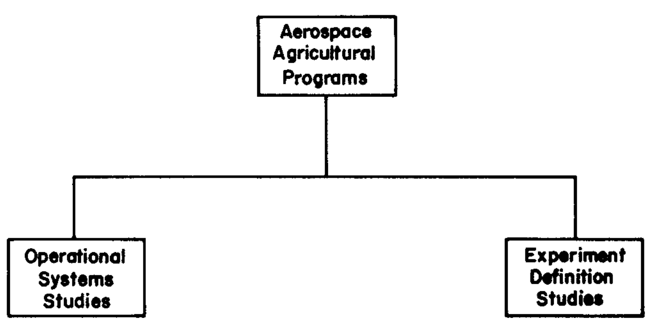


Figure 6. Aerospace Agricultural Programs

A scientific advisory committee serves to critically review current and proposed LARS research programs. Additionally, the committee reviews research reports generated by the Laboratory staff. Members of the advisory committee are:

. Virgil Anderson, Professor of Statistics

. Ludwig M. Eisgruber, Associate Professor of Agricultural Economics

. William H. Hayt, Jr., Professor of Electrical Engineering and Assistant (for Television) to the Vice President for Academic Affairs

. Herbert H. Kramer, Director of the Agricultural Experiment Station

. Robert D. Miles, Professor of Civil Engineering and Head of the Air Photo Laboratory

. Robert W. Stanley, Associate Professor of Physics

Dr. Hayt is chairman of the committee and R. B. MacDonald serves as secretary.

LEVEL OF EFFORT

Table 1 indicates the approximate level in man months maintained in designated program areas during the period from January 1968 through September 1968. This includes the efforts of individuals regularly assigned to and working on the LARS project. Not included are the efforts of such persons as Dr. N. K. Ellis, Dr. J. C. Hancock, and Dr. H. H. Kramer. Also excluded are the man hours contributed by other University staff members through service on several committees working with LARS.

RESEARCH PROGRAMS TO DATE

DATA PROCESSING PROGRAMS

LARS had previously developed a limited capability for the automatic classification of crops from airborne multispectral scanner data. During 1968, these limited capabilities were combined into an effective computer software package known as LARSYSAA which greatly increased the efficiency with which LARS data analysis research could be carried forward. The basic intention of the package was and is to greatly facilitate man-machine conversation so that the analysis scientist may more quickly pose questions of data sets and obtain the results along with LARSYSAA. Data preprocessing programs have been developed which also aid in bringing about a short turnaround time in the preprocessing of data for research use.

The first new data mission flown for LARS since the 1966 growing season was carried out during July 1968. LARS used this opportunity to test out, in a simulated operational mode, the designing of a classifier and the achieving of the classification of data in a period small enough such that conditions in the field had not changed appreciably between the time the data was gathered and the analysis results were available. Specifically, classification results became available within 48 hours after the data became available to Purdue.

A continuing research project in

Table 1. Name, Title, and Amount of Participation by LARS Staff Members

NAME	TITLE	MAN MONTHS[a/]
H. H. Kramer	Director, Agricultural Experiment Station	b/
N. K. Ellis	Associate Director, Agricultural Experiment Station	b/
J. C. Hancock	Head, Department of Electrical Engineering	b/
R. B. MacDonald	Senior Research Engineer	10
R. M. Hoffer	Research Associate in Department of Forestry and Conservation	9.3
R. A. Holmes	Associate Professor of Electrical Engineering	6
D. A. Landgrebe	Associate Professor of Electrical Engineering	8.5
M. F. Baumgardner	Associate Professor of Agronomy	6
K. S. Fu	Professor of Electrical Engineering	1.6
F. V. Schultz	Professor of Electrical Engineering	1.5
T. L. Phillips	Research Engineer	10
C. J. Johannsen	Research Agronomist	10
P. E. Anuta	Research Engineer	10
P. H. Swain	Graduate Student	7.3
S. J. Kristof	Research Agronomist	10
J. E. Halsema	LARS Photographer	10
H. T. Breece, III	Graduate Student	6
J. W. Clevenger	Graduate Student	6
R. E. Becker	Graduate Student	5
W. R. Simmons	Engineer	4
	Student	1.5
D. A. Germann	Graduate Student	3.5
P. J. Min	Graduate Student	3.5
J. I. Ash	Graduate Student	1.2
T. Huang	Graduate Student	.5
D. Bernstein	Statistician	3
	Graduate Student	2.5
F. Phillips	Student	6.5
J. Cooley	Secretary	2
C. Roe	Secretary	7
G. Holt	Secretary	4
P. Small	Secretary	7
T. Renne	Secretary	2
T. Martin	Student	3
J. Holmes	Student	3
D. Remsburg	File Clerk	10
T. Sinclair	Graduate Student	1.5

a/ This is the approximate number of man-months of participation by staff members for the contract period.

b/ Time devoted to LARS is part of administrative duties and jobs are not funded by LARS.

the overlay of data gathered through different multispectral scanner apertures reached a significant milestone during 1968. More specifically it became possible, on a relatively routine basis, to overlay multispectral images gathered through different scanner apertures on a given aircraft flight. The purpose of this research at the present time is to make available from a multiple-aperture scanner the type of data which could be collected with a single-aperture scanner, thus in carrying out crop classifications making available additional parts of the electrical magnetic spectrum. This research program will be continued after single-aperture scanners become available since a data overlay capability will make possible the use of data gathered at different times of the day and from other data gathering instruments (e.g., radar) in classification routines.

During 1968 a limited amount of data gathered over a geologically significant site for the U. S. Geological Survey (USGS) was analyzed and classified in concert with USGS scientists. In this work an attempt was made to identify, by use of multispectral pattern classification techniques, different types of soil and rock outcroppings. This study indicated that techniques developed by LARS for agriculture do have applicability to other disciplines.

During 1968 the visual digital image display was designed to enable analysis of larger quantities of data. When this device becomes available, it will greatly facilitate man-data communications in situations where the data is in image form and of exceedingly large volume. The system will be instrumental in processing data collected by future satellite systems.

In summary, the general progress of LARS during 1968 moved the state-of-the-art of data handling and analysis to the point from being able to handle data from a five-square-mile area to a 500-square-mile area.

MEASUREMENTS PROGRAMS

Infrared calibrated spectra were taken over more than 150 targets in conjunction with NASA aircraft flights. These spectra provided calibration checks on aircraft scanners and were used in LARS recommendations on the decision of wavelength band choices for the NASA scanner proposal request.

Leaf-scattering spectral and spatial data were obtained from normal corn and soybean leaves. These data will serve as base line data for spectral determination of plant stress due to moisture and nutrient deficiencies and disease in the coming year.

Close liaison was maintained with NASA-Houston on scanner parameter choices prior to the issuance of the request for proposal to industry to achieve an instrument useful for agricultural purposes.

AGRICULTURAL REQUIREMENTS AND APPLICATIONS PROGRAMS

Significant research results in the application of remote sensing and automatic data processing techniques to problems of agricultural production can be obtained only after such techniques have reached a certain stage of development. Research efforts are now beginning to make available techniques which can be applied to research in the areas of landscape surveys, soils mapping, land use planning, surveys of crop areas and conditions, nutrient deficiencies, erosion control studies, and watershed management. With these potential applications in mind, significant progress has been made in recent months in outlining a long-range research program. Initially, this research will involve the design of an information system for Tippecanoe County, Indiana, an area of 501 square miles. This experiment will include data gathering and analysis techniques and a study of information utilization for planning and management of agricultural resources in a specific region of importance. This expanded test site is to be a fundamental agricultural test site in early test satellite programs.

BIOGEOPHYSICAL RESEARCH PROGRAMS

Using data from two different geographical areas, automatic classification of fields of bare soil was carried out. Classifications were based upon spectral characteristics of surface soils. Various soil groupings were made. In some cases only light and dark soils were identified and mapped, whereas in others six, eight, and twelve soil groups were determined. It is believed that many more soil groupings can be obtained in a quantitative manner using spectral scanner data. A reasonable degree of success has been demonstrated in this soils classification and mapping. Further research will be needed to develop this into a reliable technique to aid the various soils mapping programs of federal, state, and local agencies. This research represents a major breakthrough in methods used for soils mapping work in that it (1) allows quantitative determination of soil class boundaries, (2) allows quantitative determination of areas having similar surface characteristics, and (3) is an automatic technique which makes possible the analysis of data and its reduction to useful information in a timely manner.

The capability to successfully classify data obtained over large geographic areas was demonstrated. A flight line 70 miles in length between Indianapolis and Bedford, Indiana, was classified into basic cover types including bare soil, green vegetation, and water. Mature vegetation and man-made objects such as roads and houses served as thresholds. More detailed classification groupings were also studied. These involved coniferous and deciduous forest cover, winter wheat and pasture categories. In one instance, it was observed that polluted water had a different spectral signature than unpolluted river or quarry water. Different soils groupings were studied in detail.

The major objective in this study was to determine the capability to work with data obtained over very large geographic areas and automatically identify and map primary cover types in this region in a timely manner and with high reliability.

Analysis of spectral data collected over widely separated geographic areas (Weslaco, Texas; Davis, California; and, Lafayette, Indiana) provided evidence that automatic processing techniques derived with LARS research programs are generally applicable in processing data collected on a global basis.

While limited ground truth was available for the Texas and California data, tentative identification (later shown to be correct) could be made on such things as rice or bare soil, based upon analysis of the spectral signature of the data. Fields of similar crop materials could be identified even though the specific crop type was not known.

AEROSPACE PROGRAMS

Some aerospace programs have been designed but not implemented. The objectives include the applications of research results to the design of experimental programs leading to the development of operational earth resource aircraft and satellite systems.

FUTURE OBJECTIVES

Even though automatic recognition techniques for applications in earth resources information systems of the future have been established, much remains to be done before such techniques are reduced to practice. Relatively few data have been analyzed in but a limited number of situations. The proper time for collecting data has not yet been established for the majority of situations in agriculture. Certainly, that information which can be collected regularly throughout the growing season is still to be defined in future studies.

LARS believes, on the basis of experience to date, that automatic recognition techniques will require periodic training utilizing measurements from known training sites. In 1969, investigations are planned to better define how often and under what conditions retraining is required. This analysis is to be performed with data collected over a relatively large area of 500 square miles.

A series of flights have been scheduled over the 500-square-mile Tippecanoe County area in 1969 as part of a year-long experimental program at LARS to further develop automatic recognition capabilities with remote sensing data. Five data acquisition flight missions have been scheduled at cardinal points in an agricultural season in the Corn Belt region around LARS. These are scheduled from tilling and planting in April-May through the growing season to harvest in September-October. An additional flight is scheduled for November-December to investigate fall-plowed areas and areas planted to winter wheat. Some 280 square miles of data are to be collected per flight.

Primary data are multispectral measurements stored on magnetic tapes. Black and white and multiband photographic data are to be collected as support data. Additional ground truth data are to be collected by ground crews. Only limited rate-of-change detection analysis can be conducted with these data. An investigation of the information contained in temporal variations of radiated energy requires data collected with clusters of flights at critical times in the year. Such studies will be continued in a following year.

LARS researchers feel that more attention must be given to the radiation characteristics of vegetation under different stress conditions in order to develop reliable remote sensing techniques for identifying and mapping of such conditions. Experiments are to be conducted under carefully controlled conditions during 1969 with certain vegetative species. The primary objectives are to determine spectral intervals in which different levels of stress are detectable. The experiments will involve several different sources of stress.

The data to be acquired in 1969 over Tippecanoe County will be utilized in further analyzing automatic recognition capabilities for surface soil type delineations. Major objectives include the formulation of relationships between spectral characteristics and surface soil conditions of interest and the effect of conditions such as different soil moisture content on surface soil type classification processes.

It is also planned that the staff will further investigate the feasibility of automatically classifying surface water bodies into different categories on the basis of radiated energy characteristics.

CHAPTER 2

AGRICULTURAL REQUIREMENTS

INTRODUCTION

Defining the applications and requirements of an Agricultural Industry, a Government Agency and an individual is a continuing process. Their needs vary at different times of the year due to growing seasons and according to the crop production of specific areas. The emphasis of management or outside interest groups can further influence the information needs of these groups.

Efforts in this program area concentrate on determining the requirements of agricultural industries and in assisting industry in the preparation of cooperative proposals to plan for better development and utilization of remote sensing systems.

A proposal is being written with International Minerals & Chemical Corporation to cover research for a five year period. Some preparatory research must be done and work completed before the start of this five year period. This time period has been termed Phase "O" and will be reported here.

REQUIREMENTS OF AGRICULTURAL CORPORATIONS

Seven Agricultural Corporations were chosen for survey to determine what their requirements might be from remote sensing systems. These industries were selected because of their wide range of background and their interests in different phases of agriculture. Many visits with representatives from other companies are not included in this writing but results also reflect these visitations.

Companies which were visited are as follows:

Deere and Company, Moline, Illinois
DeKalb Agricultural Association, DeKalb, Illinois
International Harvester Company, Chicago, Illinois
Swift Agricultural Chemicals, Chicago, Illinois
Sinclair Petrochemicals, Inc., Chicago, Illinois
American Oil Company, Chicago, Illinois
Allis-Chalmers, Milwaukee, Wisconsin

Presentations were made to the different companies according to the level of interest expressed by contacts within the company. In most companies, research personnel were visited, but in a few cases presentations were made to management personnel. After a formal or informal presentation, the company's personnel were asked how they would visualize they could use such a system. Questions were also asked as to how the company now obtains agricultural information and how quickly it needs this type of information.

Kinds of information which the different companies generally require would fall into two main categories. They need information which is used for short range planning and a different set of information for long range planning and decision making. Some of the companies were reluctant to discuss in detail what type of information they used, how it was used, and how it was obtained. They indicated that most of these procedures were company secrets and that they did not want their competitors to find out how they obtained this information.

Some of the companies gave the impression that they did not want to be bothered with a new technique such as remote sensing until it had been proven that it would actually work. When specific information could be supplied, such as the total acreage of a certain crop that is planted by a certain date, they would be interested in visiting further about the usefulness of this information.

PHASE ZERO OF IMC/PURDUE PROPOSAL

The title of the proposal will be "Aerospace Remote Sensing System Project for Agricultural Resource Development Planning and Management." The research work of LARS **in** data generation by remote sensing and in data manipulation plus IMC's broad experience in computerized data analysis for business management and distribution systems provide the foundation for this five year proposal.

Some preliminary foundation work is a prerequisite to the implementation of this proposal. This is because not all the necessary knowledge, methods, and techniques required to implement exist. There is also a need for the development of an efficient and effective working relationship between Purdue and IMC in order that the research and supporting services may interface satisfactorily. Phase Zero described herein is designed to fill these needs and requirements and to provide an adequate and more complete foundation from which the aerospace remote sensing systems project may be launched.

The following is a description of the areas in which major effort will be expended during phase zero. More detailed plans and programmed schedules must be developed for the research and supporting services, the pilot observation and utilization systems, and the educational program. This would include a seminar workshop utilizing the leading authorities in remote sensing. Its purpose would be to establish a well documented foundation for the take off of the program and to obtain the best judgement available concerning the optimum course for the program. The cost and benefits of alternative sources of aircraft and sensing equipment for each part of the program would be evaluated. A schedule of aircraft services to be used and flight missions to be flown would be developed. The output of different sensors would be considered and how to interface with rapid data processing would be determined.

An initial program would be planned and conducted with field cooperators in Tippecanoe County. This would involve interviewing farmers to obtain basic information such as soil type, tillage, and cropping program. Other information users would be interviewed to obtain basic business information and to determine potential uses of data from remote sensing. Orientation meetings would be conducted for selected program cooperators. Working with the above users on test cases would be necessary to determine if the proposed program (variables to be measured and sensed) will satisfy the needs of major users. A planned program of instruction and on-the-job training for new personnel for data handling, photo and data extraction as well as interpretation and ground truth determination would be implemented.

The number and size of sample observation areas required to give statistically reliable data for both Tippecanoe County, Indiana and the Corn Belt would be evaluated. This would be done in order to be sure that the sampling procedures developed are usable in larger geographic areas and are compatible with satellite operations. A statistically reliable system to extrapolate from sample data to population data by counties for both satellite and aircraft data would be developed.

Necessary design characteristics of data analysis systems would be determined. This would include further development of pattern recognition and multispectral signatures for remote multispectral sensing. Also, tests would be carried out on significant quantities of data using alternative recognition techniques. Hopefully this would help to better understand the capabilities of multispectral sensing of major Corn Belt crops and environmental variables. Data extraction and computer training techniques usable for significant quantities of data on a scheduled basis would be evaluated.

The output of Tippecanoe County flights that measure acreage of major crops would be utilized to test information generation and distribution systems that are applicable to broader geographic areas. This would include the development and evaluation of preliminary schedules and programs for the selection and reproduction of information output from the data extraction and analysis systems. The variables that need to be quantified as the basis for yield and production projections throughout the growing season would be evaluated. Also, the characteristics of decision assistance models that will be of most value to information users would be determined.

Computer programs must be prepared to handle proposed volumes of data. These programs would need to be written with the capability to be adapted to revised methods of data collection and output needs. Collation of field data with data collected from scanners and photographs through the use of computer programs would be implemented. Preliminary field and greenhouse fertility trials would be conducted. Spectral measurements would be made to assist in determining experimental design and analysis procedures. Ground truth parameters and improved systems of gathering these parameters would be studied so that data from many locations can be quickly obtained and use for interpretation of scanner data.

The National Commission of Space Activities of Brazil would be assisted in the development and implementation of an effective program to provide agricultural information collected by remote sensing to fulfill major needs in Brazilian agriculture. Work would be carried out with the seven test sites at the South Central Research and Agricultural Experiment Station (Kilometer 47) and the Rural University of Minas Gerais. This would involve the selection of aircraft and sensors for data acquisition; the planning of agronomic research and ground truth systems; the determining of spectral signatures and computer training methods; the carrying out of program planning and personnel training and the planning of data analysis, distribution and use systems.

Consideration would also be given to assisting in the formulation and implementation of information systems using remote sensing in Mexico and other countries where such programs could be initiated.

Studies would be made to aid in

determining the change of wavelength response due to altitude. These response changes are generally due to moisture and dust particles in the atmosphere and add "noise" to the data. The effect of altitude could therefore be simulated by adding noise to data taken at low altitudes. After classifying an area according to its different agricultural features, the original analog tape would be re-digitized with some noise added to simulate a higher altitude. The second digitized tape would then be classified; the accuracy of classification would be compared to the original classification. The original tape would be re-digitized with increasing noise levels for increasing simulated altitude levels. Effect of altitude on different wavelength bands would then become apparent by the classification results obtained.

CHAPTER 3

DATA PROCESSING

INTRODUCTION

During a previous year, LARS developed a limited capability to utilize airborne multispectral scanner data obtained to automatically classify crops. During 1968 these capabilities were combined into an effective computer software package known as LARSYS, which greatly increased the efficiency of LARS data analysis. The basic intention of the data analysis part of the package (LARSYSAA) was and is to greatly facilitate man-machine conversation so that the analysis scientist may more quickly pose questions of data sets and obtain the results. Data preprocessing programs (LARSYSAH) have been developed which aid in bringing about a short turnaround time in the preprocessing of data for research use.

A discussion of LARSYS is given in the following two sections of this report. The following section briefly discusses LARSYS, LARSYSAH[1/] and presents in detail the capabilities of LARSYSAA. In another section, the technical and programming considerations used in designing LARSYSAA are presented.

In the past year, a continuing research project in the overlay of data gathered through different multispectral scanner apertures reached a significant milestone. More specifically, it became possible on a relatively routine basis to overlay multispectral images gathered through different scanner apertures on a given aircraft flight. The purpose of this research at the present time is to make available from a multiple-aperture scanner the type of data which could be collected with a single-aperture scanner, thus making available for crop classifications additional parts of the electromagnetic spectrum. This research program will be continued after single-aperture scanners become available because a data overlay capability will make possible the use of data gathered at different times of the day and year and from other data gathering instruments (e.g., radar).

A scanner data overlay section presents a summary of the overlay work done to date. A complete report of this work was presented as LARS Information Note 103068.

The LARS data handling and analysis system has been used for data other than LARS and agricultural data. A report of these efforts is included in the section entitled, Analysis of Non-LARS Data.

Analysis of data collected over Purdue test sites in 1966 continues. Methods of training sample selection and class selection are being researched.

During July, 1968 the first new data mission flown for LARS since the 1966 growing season was carried out. LARS used this opportunity to test in a simulated operational mode the designing of a classifier and the resulting classification of data in a short enough period that conditions in the field would not be changed appreciably between the time the data were gathered and the analysis results

---

1/The rationale for LARSYSAH has been more fully described in "Remote Multispectral Sensing in Agriculture," Volume No. 2 and Volume No. 3, Annual Reports, Agricultural Experiment Station Research Bulletin No. 832 and No. 844, Purdue University, Lafayette, Indiana, 1967 and 1968.

were available. Specifically, classification results became available within 48 hours after data became available to Purdue. The results of these data analysis efforts are reported in Chapter 4, Biogeophysical Research.

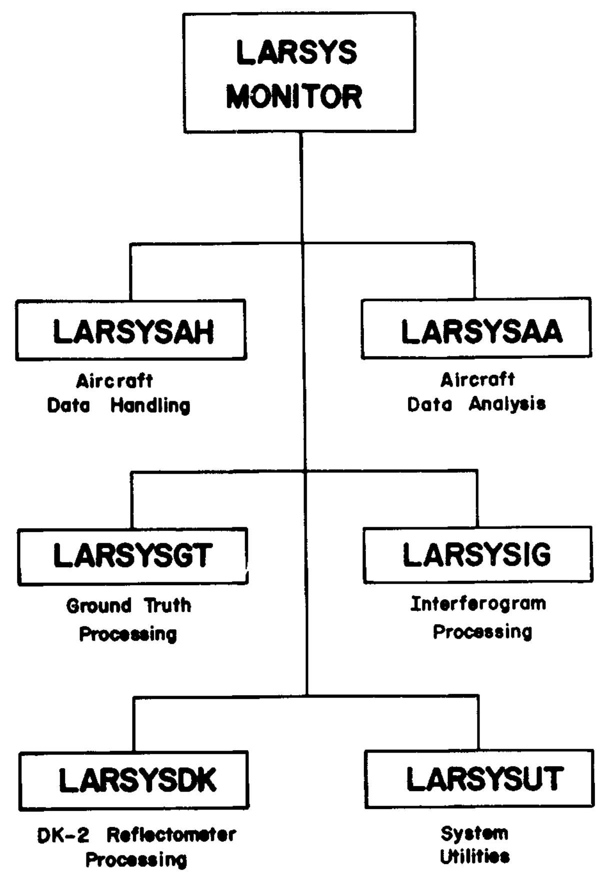
The development of the digital display concept has progressed so that the hardware design can be finalized. This device was designed to facilitate the analysis of larger quantities of data. When it becomes available, the digital display will greatly facilitate man-data communications in situations of an exceedingly large volume of data in image form. The system will be instrumental in processing data collected by future satellite systems.

In summary, the general progress of LARS in the past year has seen the application and extension of state-of-the-art data handling and data analysis techniques to the processing of data from areas several tens of square miles in size on an accurate and timely basis.

LARSYS, A PROCESSING SYSTEM FOR AIRBORNE EARTH RESOURCES DATA

The system used by LARS for research in agriculture and other earth resources systems is embodied in a set of computer programs known as LARSYS (See Figure 7). The purpose of this section is to give examples of the types of output which can be produced for the aircraft data portions of this system.

The chief purpose of LARSYSAH (Figure 7) is to produce aircraft data storage tapes (see next sub-section), alphanumeric pictorial printouts and, in the future, digital display data images. LARSYSAH output is still essentially data in an unreduced form. Once the data storage tapes and pictorial printouts have been generated, LARSYSAA may then be used by researchers to reduce the data to useful information. As illustrated in Figure 8 LARSYSAA contains four processors, each controlled by its own supervisor. These examples were all generated using data from the University of Michigan 12-channel scanner although data from other scanners could also have been used. Correspondence between the channel numbers and spectral bands for this particular scanner are given in Table 2.



LARS PROGRAMMING SYSTEM (LARSYS)

Figure 7. Diagram of LARSYS

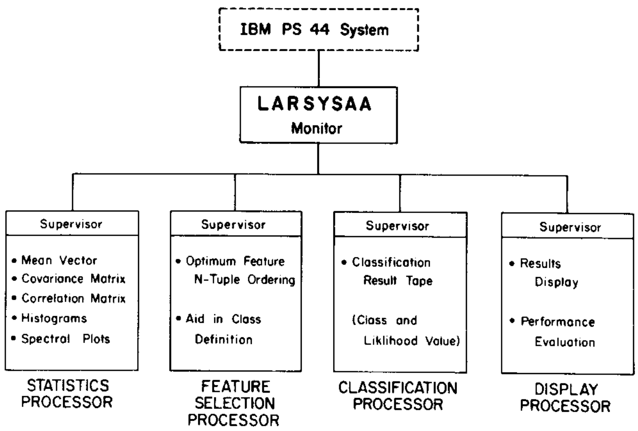


Figure 8. Organization of LARSYSAA

Table 2. Channel Numbers and Spectral Bands for University of Michigan 12-Channel Scanner

Channel Number	Spectral Band (Microns)
1	.40- .44
2	.44- .46
3	.46- .48
4	.48- .50
5	.50- .52
6	.52- .55
7	.55- .58
8	.58- .62
9	.62- .66
10	.66- .72
11	.72- .80
12	.80-1.00

DATA STORAGE TAPE

The data storage tape is a digital magnetic tape produced by LARSYSAH from the aircraft analog tape. It contains the data for each resolution element stored in a packed format and has a specific address for each point in the form of a scan line number and sample number. Certain other information, such as run number and date, necessary for a machine controlled storage and retrieval system and data for calibration purposes derived from the aircraft analog tape are also stored in a convenient format. For further details see Volume 3 Annual Report.1/

ALPHANUMERIC PICTORIAL PRINTOUT

Two channels of a portion of run number 26600061 are shown in pictorial printout form in Figures 9 and 10. Parameters which may be varied in this type of presentation include the spatial resolution, the radiance-level-to-symbol correspondence and the symbols used.

Spatial Resolution

This particular run was digitized so that on the average there is neither underlap nor overlap of adjacent samples on the data storage tape. In this case, based on the aircraft altitude and the scanner resolution, this required that every seventh scan line on the analog tape be digitized and 220 samples be made in each scan line. It is shown on the two printouts (Figures 9 and 10) that every other sample point of every other line on the data storage tape has been printed. Experience has shown that this choice of spatial parameters is convenient and very adequate for most purposes. Printouts of greater resolution are easily obtained, however, and an illustration of one in

1/ "Remote Multispectral Sensing in Agriculture," Volume No. 3, Annual Report, Agricultural Experiment Station Research Bulletin No. 844, Purdue University, Lafayette, Indiana, 1968.

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

RUN NUMBER ... 26600061	DATE 6/28/66
FLIGHT LINE .. C1	TIME 1229
TAPE NUMBER .. 102	ALTITUDE ..... 2600 FEET
CHANNEL NUMBER 9	SPECTRAL RANGE 0.62 TO 0.66 MICRONS

THE CHARACTER SET USED FOR DISPLAY IS

FROM 0 TO 166 DISPLAYED AS
FROM 167 TO 172 DISPLAYED AS -
FROM 173 TO 177 DISPLAYED AS =
FROM 178 TO 181 DISPLAYED AS /
FROM 182 TO 186 DISPLAYED AS I
FROM 187 TO 192 DISPLAYED AS *
FROM 193 TO 195 DISPLAYED AS X
FROM 196 TO 198 DISPLAYED AS O
FROM 199 TO 200 DISPLAYED AS 8
FROM 201 TO 211 DISPLAYED AS M
FROM 212 TO 218 DISPLAYED AS W
FROM 219 TO 225 DISPLAYED AS W
FROM 226 TO 232 DISPLAYED AS W
FROM 233 TO 239 DISPLAYED AS W
FROM 240 TO 246 DISPLAYED AS W
FROM 247 TO 255 DISPLAYED AS W
ILLEGAL DATA VALUE DISPLAYED AS $

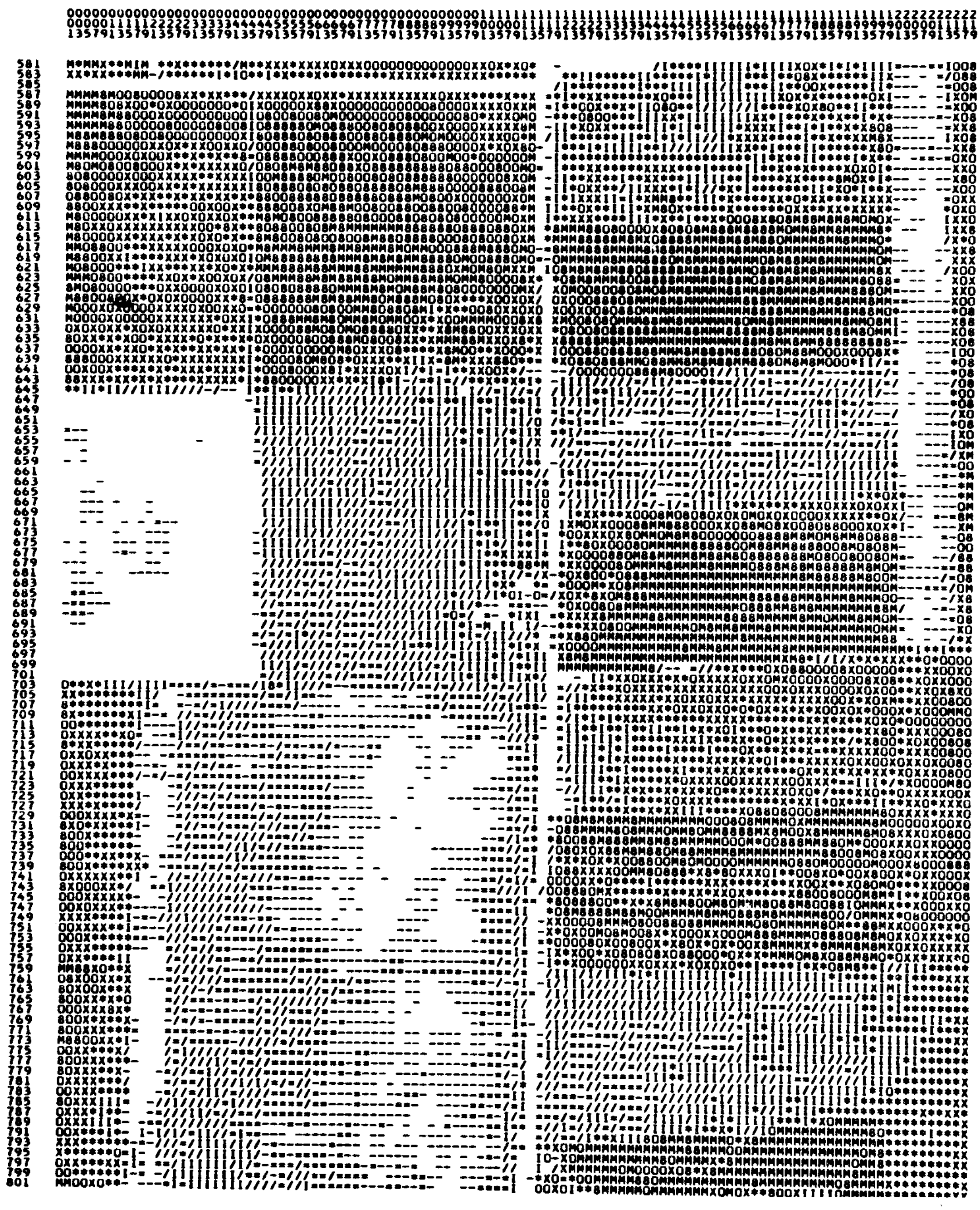


Figure 9. Alphanumeric Printout of .62-.66 Micrometer Band of Run Number 26600061

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

RUN NUMBER ... 26600061	DATE 6/28/66
FLIGHT LINE .. C1	TIME 1229
TAPE NUMBER .. 102	ALTITUDE ..... 2600 FEET
CHANNEL NUMBER 12	SPECTRAL RANGE 0.80 TO 1.00 MICRONS

THE CHARACTER SET USED FOR DISPLAY IS

```
FROM   0 TO 148 DISPLAYED AS
FROM 149 TO 155 DISPLAYED AS -
FROM 156 TO 164 DISPLAYED AS =
FROM 165 TO 169 DISPLAYED AS /
FROM 170 TO 173 DISPLAYED AS I
FROM 174 TO 177 DISPLAYED AS *
FROM 178 TO 180 DISPLAYED AS X
FROM 181 TO 181 DISPLAYED AS O
FROM 182 TO 185 DISPLAYED AS 8
FROM 186 TO 200 DISPLAYED AS M
FROM 201 TO 209 DISPLAYED AS W
FROM 210 TO 218 DISPLAYED AS W
FROM 219 TO 227 DISPLAYED AS W
FROM 228 TO 236 DISPLAYED AS W
FROM 237 TO 245 DISPLAYED AS W
FROM 246 TO 255 DISPLAYED AS W
ILLEGAL DATA VALUE DISPLAYED AS $
```

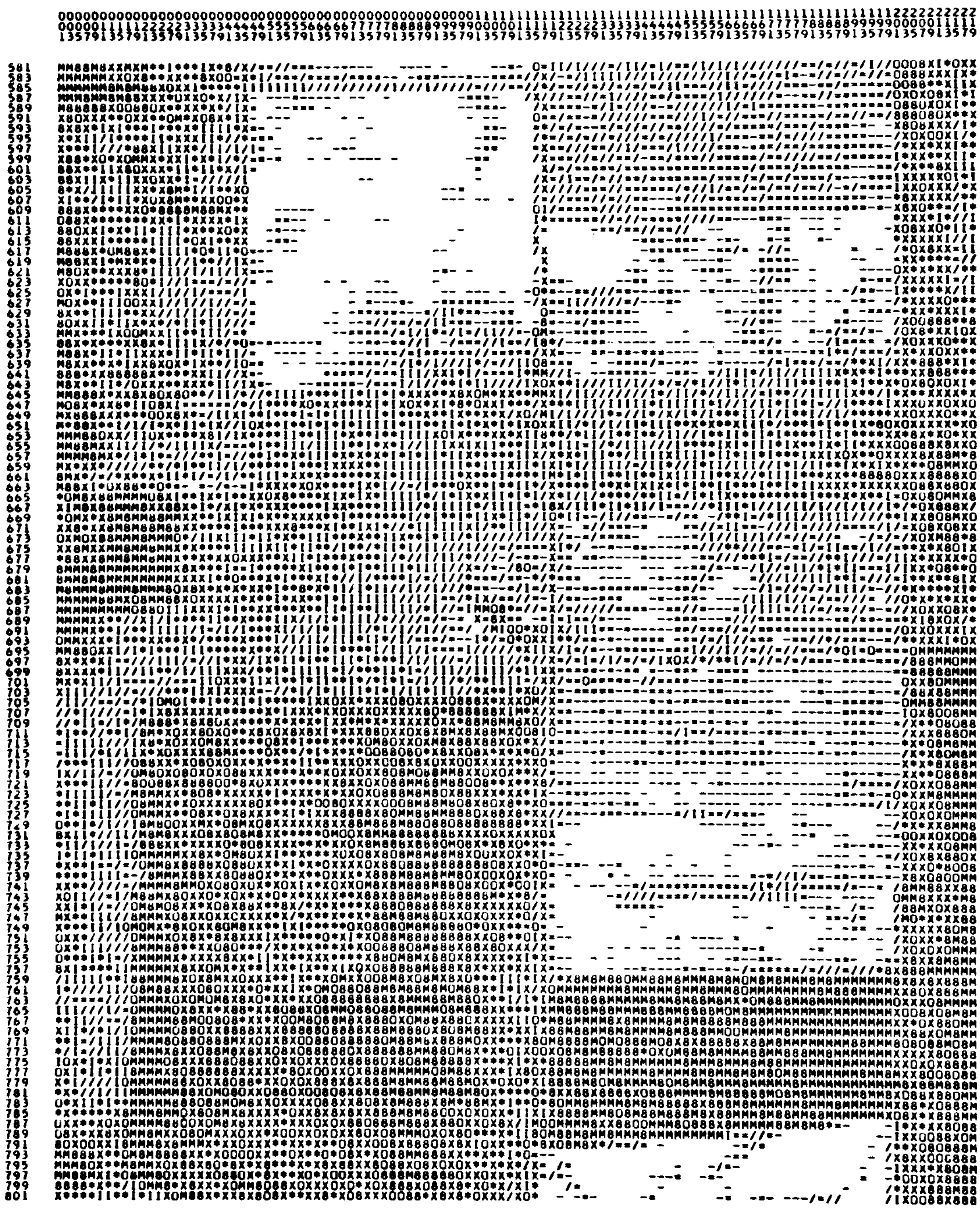


Figure 10. Alphanumeric Printout of .8-1.0 Micrometer Band of Run Number 26600061

photographically reduced form is given in Figure 11.

It should also be noted that when the data is to be analyzed in image form, as opposed to analyzing the spectrum of each point in quantitative form, overlap of samples is usually desirable. Underlap on the other hand may be desirable to reduce the data load when a detailed, high resolution study is not required.

Radiance-Level-to-Symbol Correspondence

The analog data is quantized to 8-bit accuracy. Each resolution element of each spectral band will have one of 256 possible values, according to the radiance of that element in that band. Experience has shown that from 10 to 16 symbols should be used to simulate the gray scale tones in the printout. Therefore, each of the 256 levels must be assigned to one of the 16 symbols in some fashion. Each symbol is assigned to its own group of gray levels.

One way to do this might be to use equal-sized groups, so as to assign levels 0 to 15 to the first symbol, 16 to 31 to the second, and so on. However, this does not usually give satisfactory results since the data is rarely if ever uniformly distributed over the entire dynamic range of radiances.

One of the unique advantages of the digital approach to image data handling is the simplicity with which the radiance level to gray scale may be varied. Photographically this would be the film density versus radiance curve. By using the above procedure, a curve of arbitrary shape, whether linear, nonlinear, or even multivalued, can be achieved.

LARSYSAH provides for three means of specifying the group edges for each symbol: (1) use of a preassigned set, (2) assignment of an arbitrary (ad hoc) set for the current job, (3) computation of a set which will result in equal activity for each of the 16 symbols. The first of these three is the most economical insofar as both the researcher's and computer's time are concerned. Based on experience, a set of group edges has been picked which gives satisfactory but suboptimum results for most data.

The second method may be used when the researcher knows of a good set of symbols from previous computation or if he has some specific use of the printout in mind.

The third method takes longer computationally, but it automatically provides the maximum contrast over the entire range of radiance. In this method, the data is first histogrammed, then group edges are set so that all groups have equal area under their portion of the histogram. It is also possible to accumulate the histogram from one area or a group of areas and use the resulting group edges to print another area. For the printouts shown in Figure 9 and 10, the group edges were determined from histograms of the first 950 lines of run 26600061.

The radiance-level-to-symbol correspondence for a specific printout is always given in tabular form as part of the heading (See Figures 9 and 10).

Symbols Used

The set of symbols used in Figures 9 and 10 has been picked as a standard

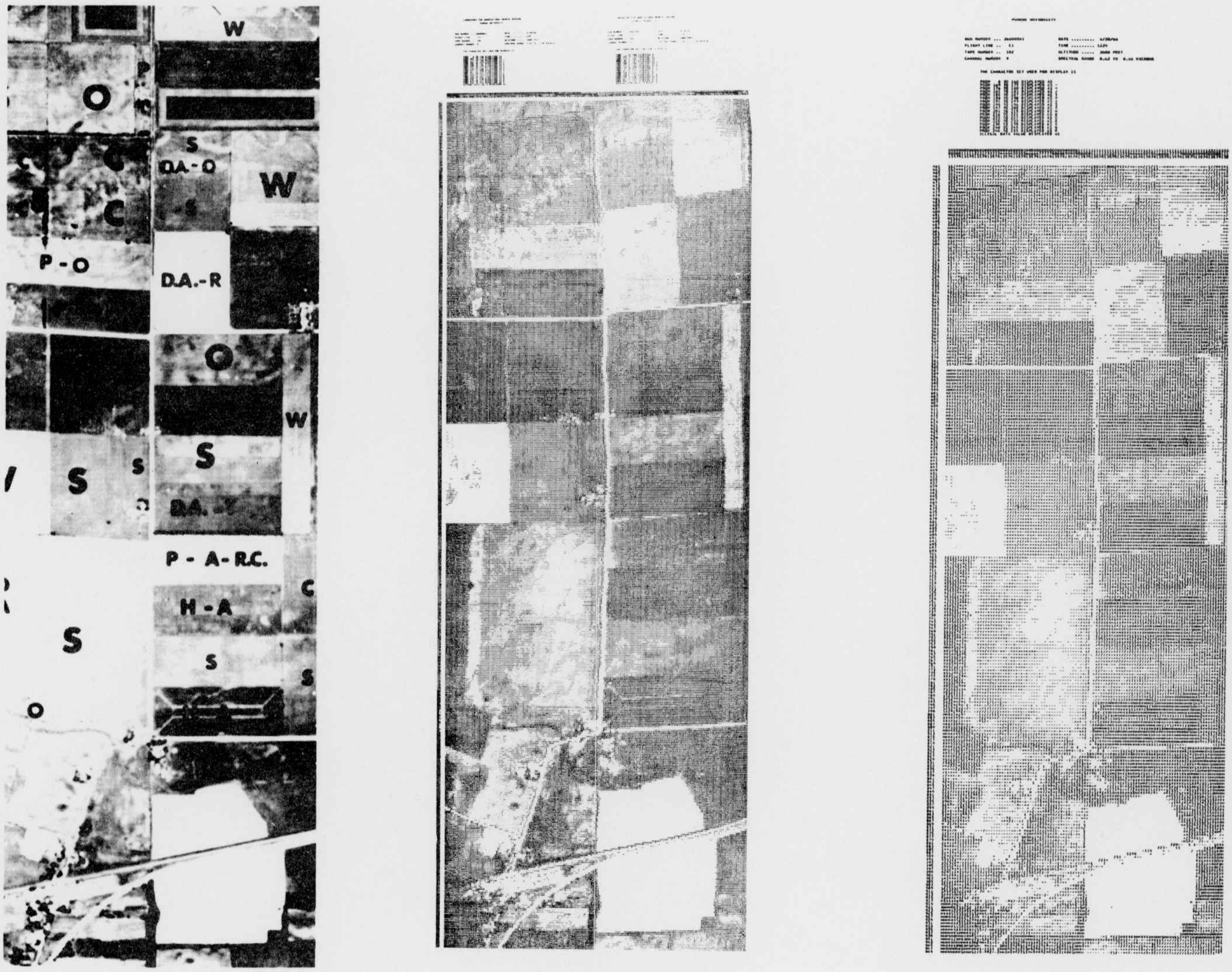


Figure 11. Photographically Reduced Alphanumeric Pictorial Printouts

set based on experience and study. However, any other set may be specified as needed for specific purposes. For example, if a contour map is desired which shows areas having a certain temperature/emissivity from a channel in the thermal region, it could be obtained by appropriately assigned group edges. This would be option (2) with blanks being used for all groups except the desired one.

Photographically Reduced Alphanumeric Pictorial Printouts

The illustration in Figure 11 shows two photographically reduced pictorial printouts of run 26600061, lines 461 to 949. The area displayed includes that shown in Figures 9 and 10. The left printout in Figure 11 again has every other sample of every other line presented, while in the right printout every sample of every line is given. In both examples there are 10 active gray level groups (symbols). A panchromatic aerial photograph is given in Figure 11 for comparison.

Output in this format is between viewing the data as a large number of quantitative spectra and viewing it as an image. For example, one may view the printout in Figure 11 as a low resolution image and also (perhaps with the aid of a hand-held magnifier) see that there is a point in Channel 9 at line number 609 and column 213 near the right edge of the wheat field which is displayed as an asterisk (*) and therefore, from the tabulation in the heading has a relative radiance of 188-190 on a scale of 0 to 255.

PHOTOGRAPHS OF DIGITAL IMAGE DISPLAY PRESENTATIONS

It is not possible to show an example of this type of data format since procurement of the digital display has not yet been possible. However, when it becomes available it is expected to provide images of at least studio TV quality. While the original primary purpose for the digital display was to speed the data editing function, it will in addition provide a considerable number of capabilities never before possible in man/data communication. This is because it will match good quality image production with the flexibility provided by digital control.

Note in particular the comments in the previous section about group selection and gray scale control also apply here. Additionally, images created from linear combinations of bands, reconstituted color and false color images are expected to be quickly possible.

EDITED DATA

The chief reason for devising the pictorial printout and digital display was to enable the editing of data. That is, one must be able to extract from the data storage tape the data belonging to a specific area on the ground. Refer to Figure 9 or 11 showing pictorial printouts and aerial photo. Note the corn field at the left edge just above the wheat field. By referring to the printout, the addresses on the data storage tape of the data field (i.e., set of contiguous resolution elements rectangular in shape) in the center of this agricultural field can be determined as lines 603 to 625 and columns 13 to 33. These addresses are valid for all channels of data gathered by this 12 channel detector set.

Once located in this fashion and

extracted from the data storage tape, this data can be stored in the computer for further processing. It can also be printed, punched out, or written on magnetic tape for external use by the researcher. Figure 12 shows a printout of the above data.

HISTOGRAMS

The LARSYAA statistics processor can provide as one type of output a histogram of a field or a group of fields. The histograms in three arbitrarily selected bands for the data from the previously mentioned corn field are given in Figure 13. The abscissa is relative radiance (brightness) and increases from left to right. (The numerical abscissa values decrease for increasing radiance because the scanner output signal is inverted.) The ordinate gives the number of resolution elements with a given relative radiance.

In Figure 14 the histograms for another corn field from another flight line several miles away are shown. It is seen that by overlaying the first set of histograms over the second, perhaps on a light table, a quick but useful comparison between data from the two fields can be obtained. In this case they are similar in Channel 1 but different in Channels 9 and 12.

Figure 15 shows histograms of the data from these same two fields put

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

LARSYSAA ILLUSTRATION

CORN I

RUN NO. 26600061, FIELD 12-9

NO. OF SAMPLES- 132, FROM LINES 603 TO 625 (EVERY 2 LINES), SAMPLES 13 TO 33 (EVERY 2 SAMPLES)

EDITED FROM CHANNEL 9

LINE	SAMPLE VALUES										
	13	15	17	19	21	23	25	27	29	31	33
603	197	194	197	196	196	195	194	194	195	195	192
605	195	193	194	198	196	195	194	190	194	192	195
607	196	194	191	193	194	191	192	193	188	194	194
609	195	191	192	193	191	192	191	191	191	196	197
611	196	195	193	191	195	164	195	194	197	193	197
613	194	195	193	194	195	193	194	193	194	195	198
615	195	194	191	194	194	193	191	194	192	195	197
617	198	187	190	190	195	195	194	194	193	196	197
619	194	184	190	189	189	192	193	193	193	192	193
621	192	183	187	184	193	193	192	191	192	195	194
623	196	192	191	191	190	194	194	196	193	192	195
625	196	197	191	188	169	197	195	194	196	196	196

Figure 12. Data Edited From .62-.66 Micrometer Band of Run Number 26600061

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

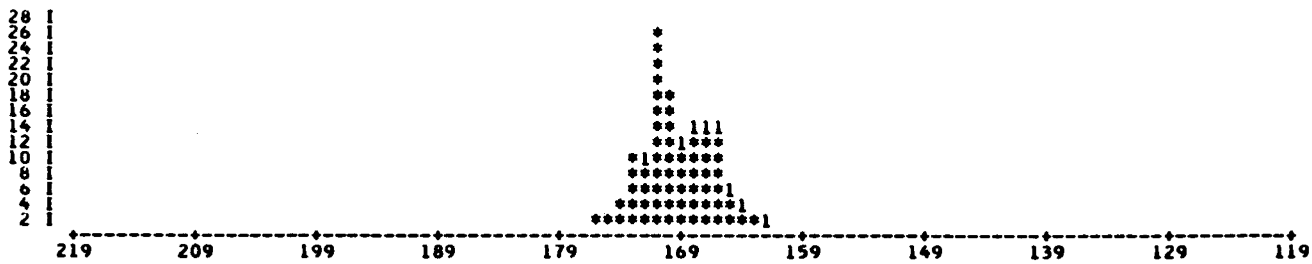
CORN I

RUN NO. 26600061, FIELD 12-9
NO OF SAMPLES = 132, FROM LINES 603 TO 625 (EVERY 2 LINE(S)), SAMPLES 13 TO 33 (EVERY 2 SAMPLE(S))

HISTOGRAM FOR ... 12-9

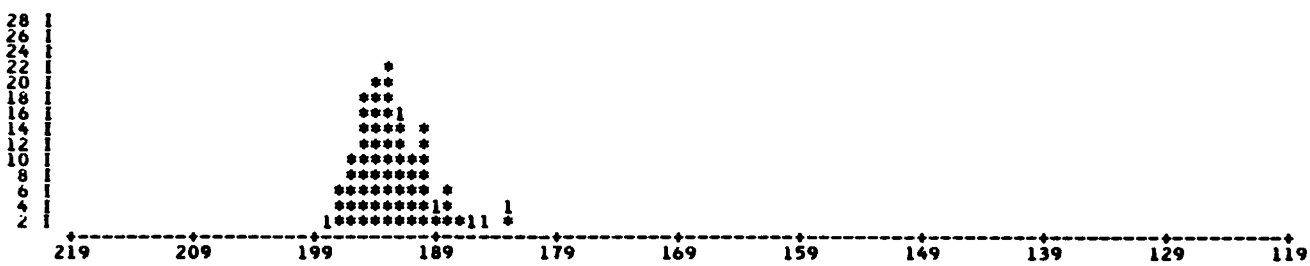
CHANNEL 1 0.40 - 0.44 MICRONS

EACH * REPRESENTS 2 POINT(S).



CHANNEL 9 0.62 - 0.66 MICRONS

EACH * REPRESENTS 2 POINT(S).



CHANNEL 12 0.80 - 1.00 MICRONS

EACH * REPRESENTS 2 POINT(S).

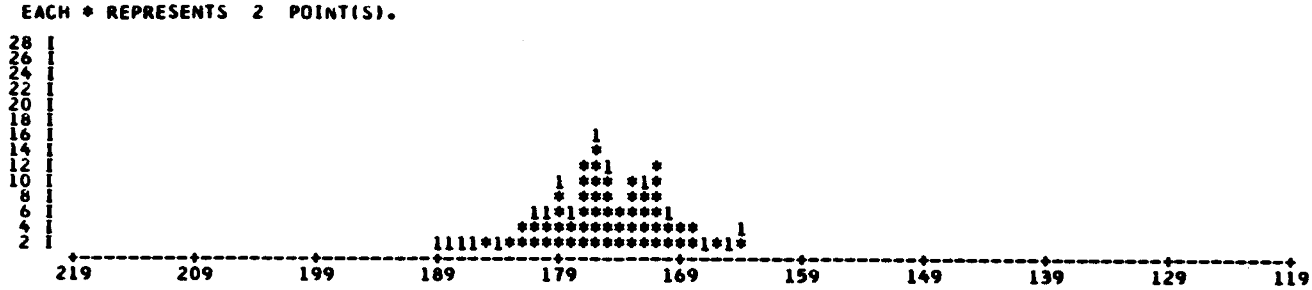


Figure 13. Histograms of Data From a Corn Field in Run 26600061

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

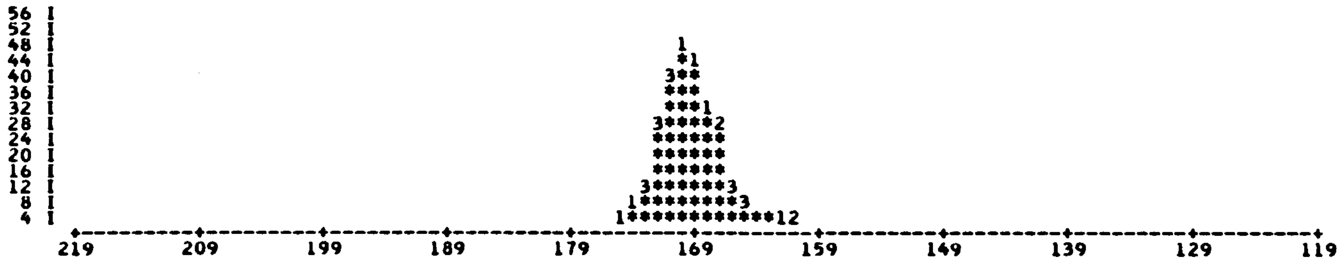
CORN I

RUN NO. 26600081, FIELD 8-20
NO OF SAMPLES = 253, FROM LINES 669 TO 713 (EVERY 2 LINE(S)), SAMPLES 171 TO 191 (EVERY 2 SAMPLE(S))

HISTOGRAM FOR ... 8-20

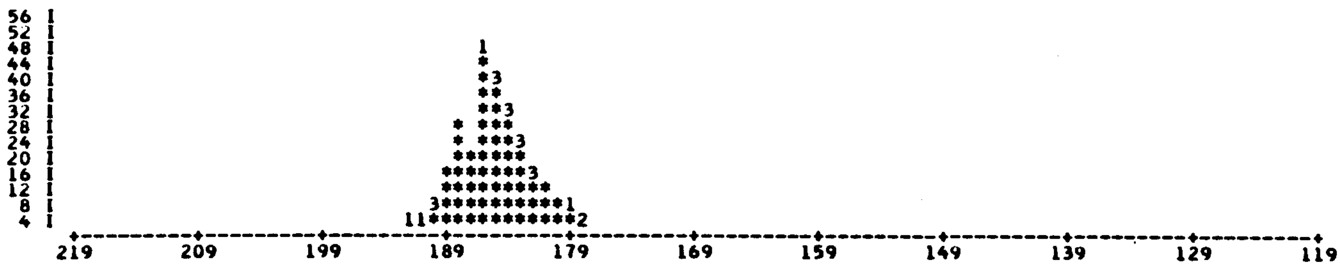
CHANNEL 1 0.40 - 0.44 MICRONS

EACH * REPRESENTS 4 POINT(S).



CHANNEL 9 0.62 - 0.66 MICRONS

EACH * REPRESENTS 4 POINT(S).



CHANNEL 12 0.80 - 1.00 MICRONS

EACH * REPRESENTS 3 POINT(S).

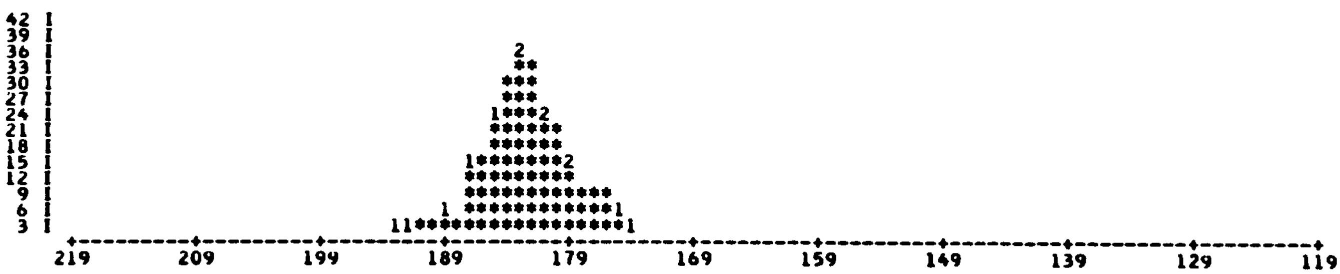


Figure 14. Histograms of Data From Another Corn Field in Run 26600061

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

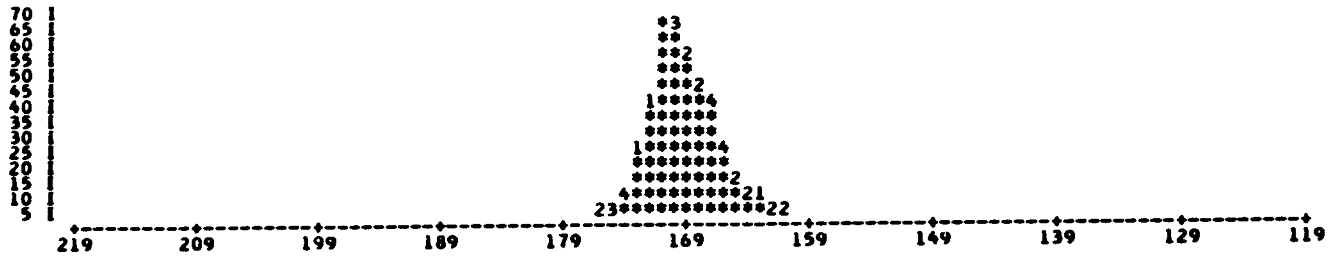
*** LARSYSAA ILLUSTRATION ***

CORN I

HISTOGRAM FOR ... CORN I

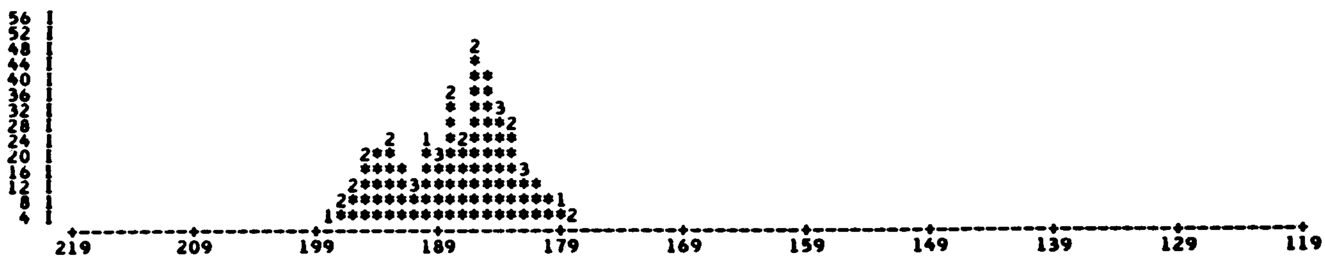
CHANNEL 1 0.40 - 0.44 MICRONS

EACH * REPRESENTS 5 POINT(S).



CHANNEL 9 0.62 - 0.66 MICRONS

EACH * REPRESENTS 4 POINT(S).



CHANNEL 12 0.80 - 1.00 MICRONS

EACH * REPRESENTS 3 POINT(S).

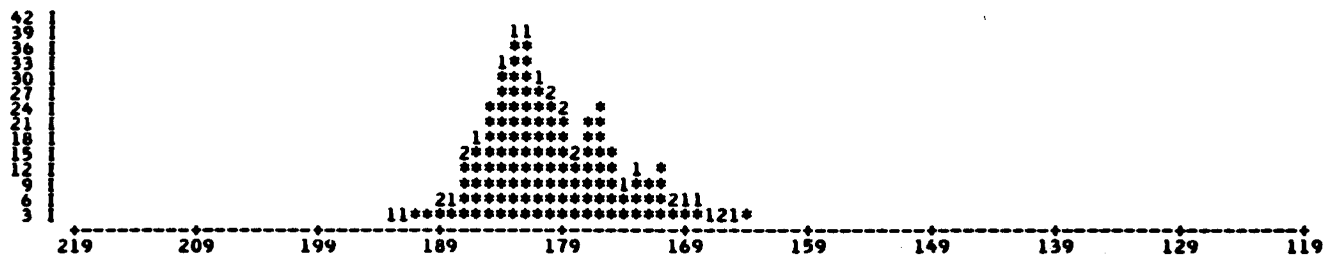


Figure 15. Combined Histograms of Data From Two Corn Fields in Run 26600061

together as a class (a group of fields not necessarily contiguous nor even from the same flight line). The double lobes in Channels 9 and 12 predicted from the two separate histograms are apparent.

In summary, note that a data field is any rectangular region in the data. If desired, the whole flight line could be defined to be a data field.

SPECTRAL PLOTS

Another type of output provided by the LARSYSAA program is the spectral plot as illustrated in Figures 16, 17, 18, and 19. In this type of

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CORN I

RUN NO. 26600061, FIELD 12-9
NO OF SAMPLES = 132, FROM LINES 603 TO 625 (EVERY 2 LINE(S)), SAMPLES 13 TO 33 (EVERY 2 SAMPLE(S))

SPECTRAL PLOT (MEAN PLUS AND MINUS ONE STD. DEV.) FOR TRAINING FIELD 12-9

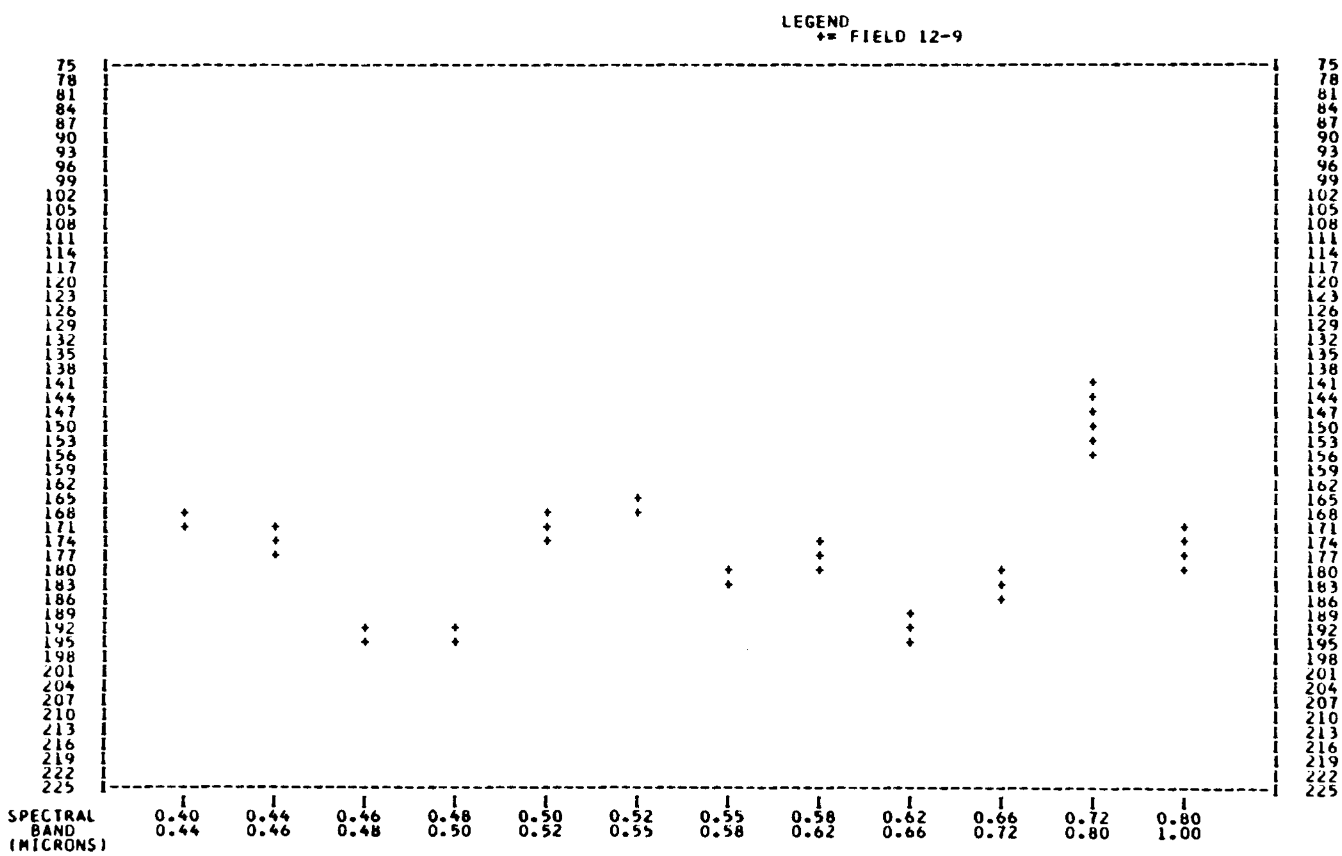


Figure 16. Spectral Plot of a Corn Field in Run Number 26600061

presentation, the 12 channels are indicated along the abscissa while the ordinate is relative radiance (brightness). A vertical line two standard deviations in length centered opposite the mean radiance is drawn using alphanumeric symbols.

The first two spectral plots given in Figures 16 and 17 are from data

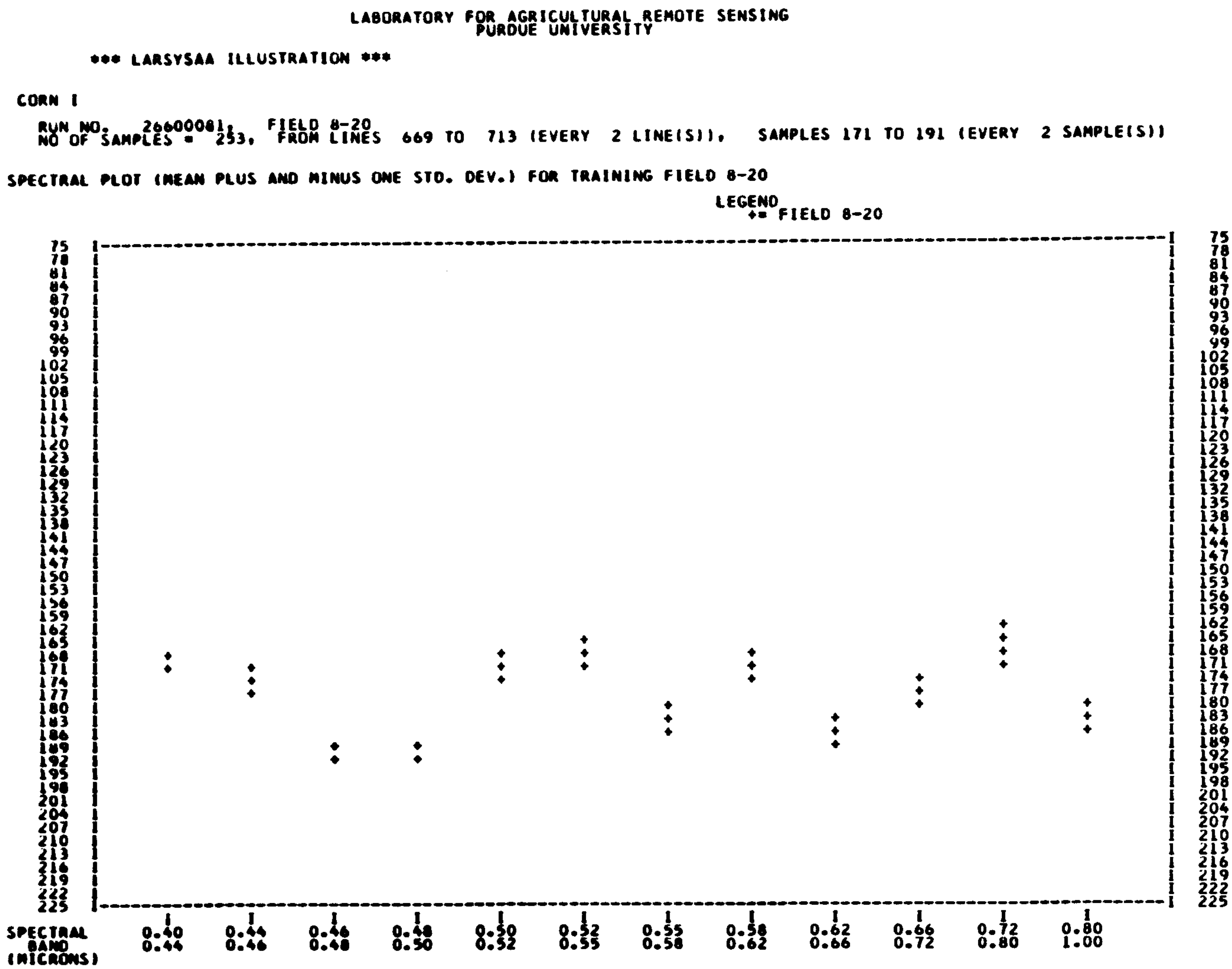


Figure 17. Spectral Plot of Another Corn Field in Run Number 26600061.

obtained from the two corn fields used in the histogram example. Again, overlaying one upon another provides a simple but useful means for comparing two fields. The third spectral plot (Figure 18) is that of these two fields combined into a class.

The fourth spectral plot (Figure 19), referred to as a coincident spectral plot,

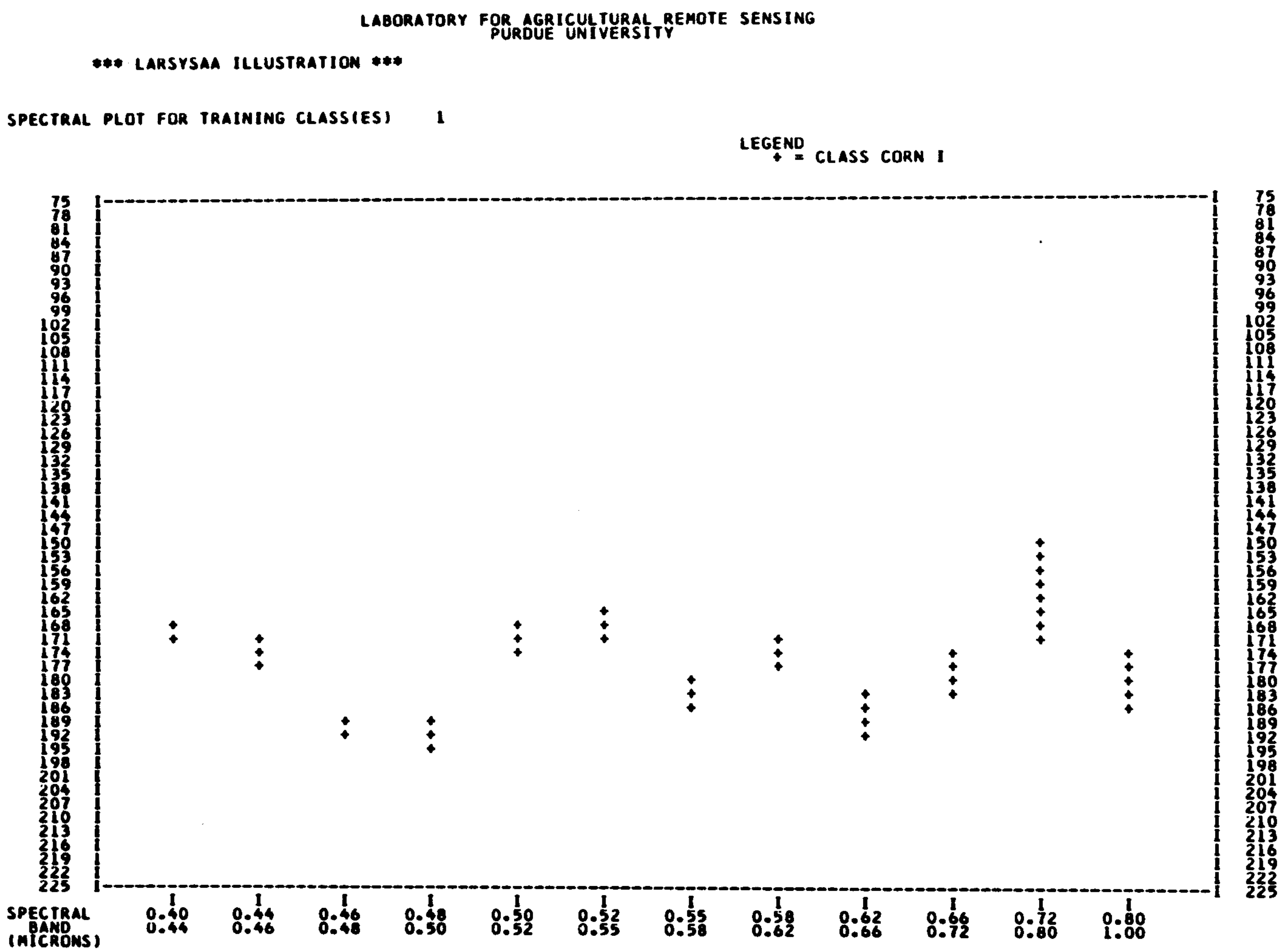


Figure 18. Combined Spectral Plot of Two Corn Fields in Run Number 26600061

provides another means for comparing two data sets. In this case four classes have been displayed on the same coordinates. By overlaying coincident spectral plots upon one another a considerable number of fields or classes can be simultaneously compared.

In summary, it must be pointed out that the scanner system is not capable of an accurate absolute calibration between bands.

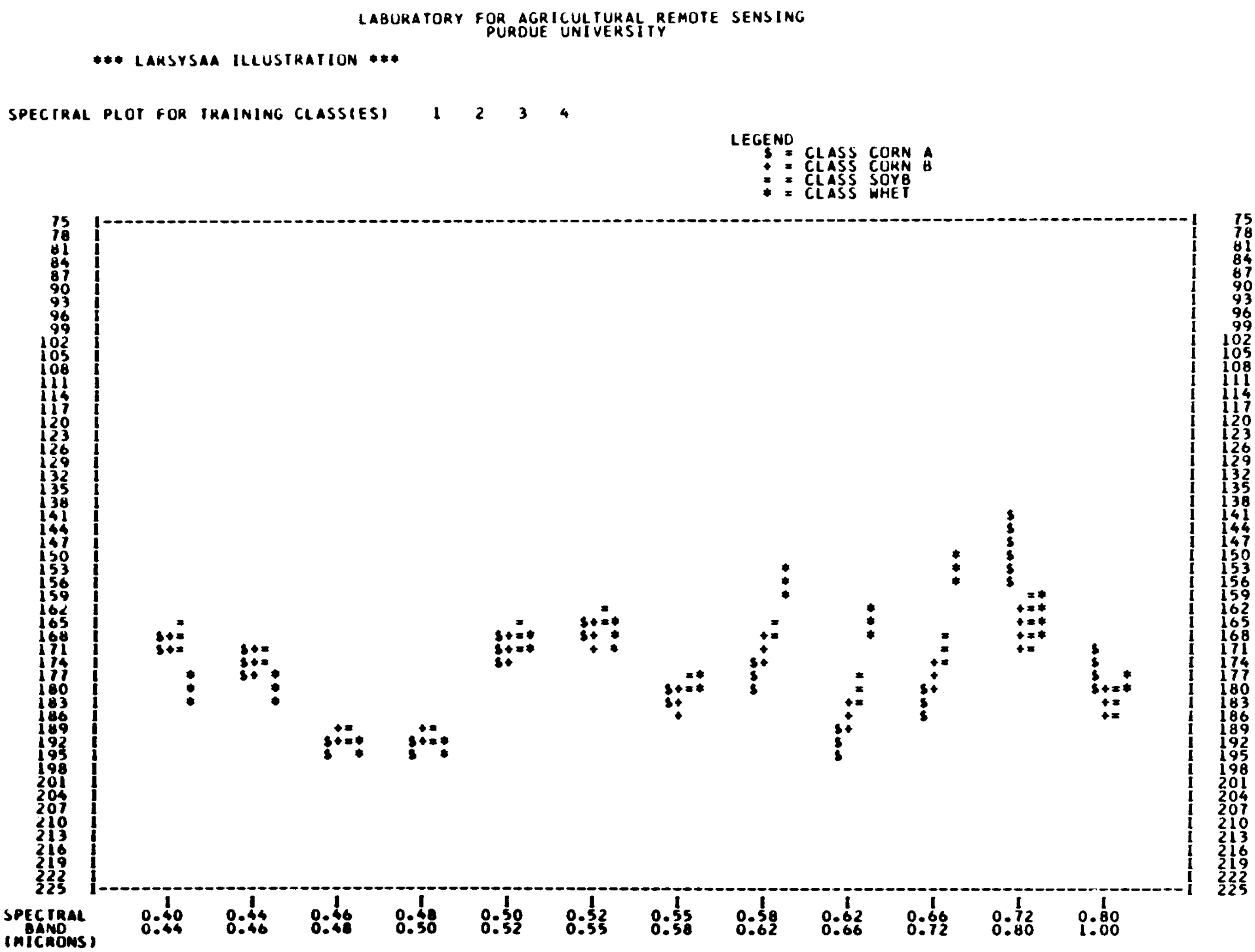


Figure 19. Spectral Plots of Four Fields in Run Number 26600061

STATISTICS

The LARSYSAA program will also provide as output a one page display of statistics as illustrated in Figures 20, 21, or 22. The first two (Figures 20 and 21) are from the two individual corn fields used previously and the third is from these two taken together as a class.

The format is largely self-explanatory. Near the top the means and standard deviations for the data of the designated region in each spectral band are given. Below this the covariance matrix is given. For example, the entry in the third row of the first column, which is 2.92 for the first field, is the covariance between Channels 1 and 3. Since this matrix is symmetrical by definition, only that part below the major diagonal is presented.

The correlation matrix, which is a normalized version of the covariance matrix is presented in the lower portion of the figure. Options in the program provide for the presentation of these statistics for only a subset of these bands as desired by the researcher.

By use of another LARSYS option, the means and covariance matrix may also be obtained in punched card form. A line printer listing from a card deck obtained in this fashion is given in Figure 23.

FEATURE SELECTION/SEPARABILITY

The LARSYSAA program provides the capability for measuring the degree of separability of Gaussianly distributed classes and determining the optimum set of channels for doing so. This is done by calculating the statistical distance in N-dimensional space between the classes. The particular distance measure implemented is known as divergence.[1,2] It has been shown that under appropriate conditions a relationship exists between numerical divergence values and the degree of separability of the classes as measured by percent classification error. Other separability measures have been tested and could be implemented if it becomes desirable.

The program is written so that the researchers may make judgments and supervise the calculations as they are proceeding. It also has many other options, not all of which will be discussed here. Figure 24 is an example of the type of output provided. In this example, five classes designated for study are given in Table 3.

Table 3. Five Classes and Symbols Used In the Example Output

Class Name	Symbol
Soybeans	S
Corn	C
Oats	O
Wheat I	W
Wheat II	M

1/ Marill, T. and D. M. Green, "On the Effectiveness of Receptors in Recognition Systems," IEEE Transactions on Information Theory, January, 1963.

2/ Min, P. J., D. A. Landgrebe, and K. S. Fu, "On Feature Selection in Multiclass Pattern Recognition," Proc. Second Annual Princeton Conference on Information Sciences and Systems, March 25-26, 1968, Princeton University, Princeton, N. J.

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CORN I

RUN NO. 26600061, FIELD 12-9
NO OF SAMPLES = 132, FROM LINES 603 TO 625 (EVERY 2 LINE(S)), SAMPLES 13 TO 33 (EVERY 2 SAMPLE(S))

THE COVARIANCE AND MEAN FOR TRAINING FIELD 12-9

	0.40-0.44	0.44-0.46	0.46-0.48	0.48-0.50	0.50-0.52	0.52-0.55	0.55-0.58	0.58-0.62	0.62-0.66	0.66-0.72	0.72-0.80	0.80-1.00
MEAN	169.39	174.33	192.80	193.45	171.08	165.83	182.29	177.59	192.54	182.98	148.94	175.16
ST DEV	2.86	3.09	1.62	1.71	2.89	2.32	2.04	2.96	2.95	3.06	8.49	4.99

COVARIANCE MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
8.18											
4.82	9.54										
2.92	2.70	2.62									
2.40	3.48	1.70	2.92								
4.77	5.67	2.99	2.85	8.37							
3.31	4.25	1.99	2.40	4.95	5.40						
2.06	3.43	1.62	2.08	3.52	3.48	4.16					
4.76	6.50	2.78	3.43	5.70	4.75	4.07	8.76				
4.81	5.90	2.67	3.60	5.26	4.12	4.09	7.27	8.69			
4.56	5.11	2.60	3.04	5.66	5.40	4.65	7.35	6.77	9.34		
-5.59	-6.29	-3.57	-5.71	-1.52	2.68	-0.43	-7.64	-10.02	-1.19	72.16	
-3.50	-3.01	-2.69	-2.91	-0.94	2.15	0.01	-3.44	-5.24	0.44	33.64	24.90

CORRELATION MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
1.00											
0.55	1.00										
0.63	0.54	1.00									
0.49	0.66	0.61	1.00								
0.58	0.64	0.64	0.58	1.00							
0.50	0.59	0.53	0.60	0.74	1.00						
0.35	0.54	0.49	0.60	0.60	0.73	1.00					
0.56	0.71	0.58	0.68	0.67	0.69	0.67	1.00				
0.57	0.65	0.56	0.71	0.62	0.60	0.68	0.83	1.00			
0.52	0.54	0.53	0.58	0.64	0.76	0.75	0.81	0.75	1.00		
-0.23	-0.24	-0.26	-0.39	-0.06	0.14	-0.02	-0.30	-0.40	-0.05	1.00	
-0.25	-0.20	-0.33	-0.34	-0.07	0.19	0.00	-0.23	-0.36	0.03	0.79	1.00

Figure 20. Statistics of a Corn Field in Run Number 26600061

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CORN I

RUN NO. 26600081, FIELD 8-20
NO OF SAMPLES = 253, FROM LINES 669 TO 713 (EVERY 2 LINE(S)), SAMPLES 171 TO 191 (EVERY 2 SAMPLE(S))

THE COVARIANCE AND MEAN FOR TRAINING FIELD 8-20

	0.40-0.44	0.44-0.46	0.46-0.48	0.48-0.50	0.50-0.52	0.52-0.55	0.55-0.58	0.58-0.62	0.62-0.66	0.66-0.72	0.72-0.80	0.80-1.00
MEAN	169.28	173.22	191.10	191.51	170.11	168.64	182.66	172.11	185.11	176.97	166.38	182.50
ST DEV	2.46	2.61	1.59	1.77	2.90	2.39	2.12	3.10	2.69	3.33	4.69	3.45

COVARIANCE MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
6.06											
2.89	6.79										
2.42	2.28	2.53									
1.56	2.81	1.42	3.12								
3.26	3.64	2.47	2.46	8.41							
2.84	4.10	2.14	2.86	4.46	5.69						
1.88	3.48	1.75	2.48	3.32	3.86	4.49					
2.85	5.18	2.77	3.60	6.27	5.88	4.92	9.60				
1.76	4.21	2.01	3.20	4.40	5.03	4.46	6.73	7.21			
1.85	4.38	2.42	3.66	5.45	5.75	5.06	8.44	7.15	11.08		
1.99	2.13	1.76	1.70	4.05	4.11	3.14	5.71	3.48	6.42	22.04	
1.93	2.44	1.69	1.35	4.27	3.79	2.34	4.81	3.05	5.23	9.10	11.91

CORRELATION MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
1.00											
0.45	1.00										
0.62	0.55	1.00									
0.36	0.61	0.51	1.00								
0.46	0.48	0.54	0.48	1.00							
0.48	0.66	0.57	0.68	0.64	1.00						
0.36	0.63	0.52	0.66	0.54	0.76	1.00					
0.37	0.64	0.56	0.66	0.70	0.80	0.75	1.00				
0.27	0.60	0.47	0.67	0.56	0.79	0.78	0.81	1.00			
0.23	0.51	0.46	0.62	0.56	0.72	0.72	0.82	0.80	1.00		
0.17	0.17	0.24	0.20	0.30	0.37	0.32	0.39	0.28	0.41	1.00	
0.23	0.27	0.31	0.22	0.43	0.46	0.32	0.45	0.33	0.45	0.56	1.00

Figure 21. Statistics of Another Corn Field in Run Number 26600061

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CORN I

THE COVARIANCE AND MEAN FOR TRAINING CLASS CORN I

	0.40-0.44	0.44-0.46	0.46-0.48	0.48-0.50	0.50-0.52	0.52-0.55	0.55-0.58	0.58-0.62	0.62-0.66	0.66-0.72	0.72-0.80	0.80-1.00
MEAN	169.31	173.60	191.68	192.17	170.44	167.67	182.54	173.99	187.65	179.03	160.40	179.98
ST DEV	2.60	2.83	1.79	1.97	2.93	2.71	2.10	4.01	4.49	4.31	10.38	5.34

COVARIANCE MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
6.77											
3.57	7.99										
2.62	2.85	3.20									
1.89	3.52	2.26	3.90								
3.79	4.57	3.01	3.01	8.59							
2.92	3.43	1.01	1.46	4.00	7.36						
1.93	3.36	1.56	2.17	3.29	3.96	4.40					
3.63	6.99	4.86	5.94	7.26	2.00	4.15	16.07				
2.98	6.64	5.08	6.60	6.31	-0.01	3.69	16.09	20.17			
2.92	6.13	4.78	6.08	6.83	1.80	4.39	15.48	17.09	18.61		
-1.03	-5.13	-6.74	-8.51	-1.69	14.68	3.39	-20.44	-30.41	-19.86	107.79	
-0.11	-1.27	-2.62	-3.34	0.87	7.88	2.16	-7.10	-12.11	-6.38	46.36	28.48

CORRELATION MATRIX

0.40 - 0.44	0.44 - 0.46	0.46 - 0.48	0.48 - 0.50	0.50 - 0.52	0.52 - 0.55	0.55 - 0.58	0.58 - 0.62	0.62 - 0.66	0.66 - 0.72	0.72 - 0.80	0.80 - 1.00
1.00											
0.49	1.00										
0.56	0.56	1.00									
0.37	0.63	0.64	1.00								
0.50	0.55	0.57	0.52	1.00							
0.41	0.45	0.21	0.27	0.50	1.00						
0.35	0.57	0.41	0.52	0.54	0.70	1.00					
0.35	0.62	0.68	0.75	0.62	0.18	0.49	1.00				
0.25	0.52	0.63	0.74	0.48	-0.00	0.39	0.89	1.00			
0.26	0.50	0.62	0.71	0.54	0.15	0.49	0.90	0.88	1.00		
-0.04	-0.17	-0.36	-0.42	-0.06	0.52	0.16	-0.49	-0.65	-0.44	1.00	
-0.01	-0.08	-0.27	-0.32	0.06	0.54	0.19	-0.33	-0.51	-0.28	0.84	1.00

Figure 22. Combined Statistics of Two Corn Fields in Run Number 26600061

The program was instructed to determine the 30 best sets of four channels (features) for separating or correctly classifying these five classes. In addition, it was specified that the separability of the 4-tuple of Channels 1, 5, 10, and 12 should be displayed even if it was not in the top 30. The results are shown in Figure 24. The best 4-tuple of channels is seen to be 1, 6, 10, and 12. To the right of the column labeled D(TOT) are listed the interclass divergences. For example, interclass divergence for soybeans and corn classes (S and C) for Channels 1, 6, 10, and 12 is 36, while that for soybeans and oats (S and O) is 84. This means that a classifier will have less difficulty separating soybeans from oats than soybeans from corn. It is inappropriate here for a complete discussion of the size of these numbers beyond the following general statement. Divergence above 400 indicates very good separability while one below about 20 would be marginal.

The entries in the column labeled D(TOT) are the sum of the interclass divergences, and it is on the basis of these numbers that the feature sets are ordered. The column labeled DIJ(MIN) gives the minimum interclass divergence for each feature set for easy reference. Also, a maximum divergence input may be considered in ordering the feature sets. This is desirable in order to prevent one highly separable pair of classes from unduly influencing the ordering of feature sets. The maximum set in this example is 350, and the entries indicated by three dots in Figure 24 are those exceeding this maximum.

A minimum divergence may also be set. If any interclass divergence does not exceed the minimum for a given feature set, that feature set will be

```
MODULE TRAINING FIELD DECK FOR LARSYSAA
CLASS    1FIELD     2FEAT     12VECTR 1 2 3 4 5 6 7 8 9101112131415161718CAL 1
FLDINF        26600061  603   625    13    33     1  26600081   669   713   171   191      1
CLSDESC    CORN 1
FREQ   0.40  0.44   0.46   0.48   0.50   0.52   0.55   0.58   0.62   0.66   0.72   0.80
FREQ   0.44  0.46   0.48   0.50   0.52   0.55   0.58   0.62   0.66   0.72   0.80   1.00
NUPTS        385
MEAN   0.16931429E 03 0.17360258E 03 0.19168051E 03 0.19217401E 03 0.17044415E 03
MEAN   0.16767271E 03 0.18253506E 03 0.17399220E 03 0.18765454E 03 0.17902856E 03
MEAN   0.16039999E 03 0.17998181E 03
COVAR 0.67681541E 01 0.35705357E 01 0.79900970E 01 0.26241064E 01 0.28466644E 01
COVAR 0.32023535E 01 0.18930798E 01 0.35224628E 01 0.22588673E 01 0.38993235E 01
COVAR 0.37897320E 01 0.45675926E 01 0.30146637E 01 0.30136490E 01 0.85860653E 01
COVAR 0.29208326E 01 0.34321022E 01 0.10097532E 01 0.14555397E 01 0.39973011E 01
COVAR 0.73613634E 01 0.19251480E 01 0.33616266E 01 0.15568037E 01 0.21722670E 01
COVAR 0.32929783E 01 0.39594221E 01 0.44004593E 01 0.36274548E 01 0.69690823E 01
COVAR 0.48646908E 01 0.59414635E 01 0.72636865E 01 0.20026512E 01 0.41474085E 01
COVAR 0.16065033E 02 0.29786453E 01 0.66410036E 01 0.50768461E 01 0.65393366E 01
COVAR 0.63126888E 01-0.91856048E-02 0.36931343E 01 0.16091049E 02 0.20174606E 02
COVAR 0.29180803E 01 0.61285706E 01 0.47773809E 01 0.60835562E 01 0.68284225E 01
COVAR 0.18010416E 01 0.43935261E 01 0.15479389E 02 0.17088013E 02 0.18611160E 02
COVAR-0.10348949E 01-0.51296873E 01-0.67442703E 01-0.85098953E 01-0.16911449E 01
COVAR 0.14683332E 02 0.33921871E 01-0.20436966E 02-0.30410934E 02-0.19863007E 02
COVAR 0.10778749E 03-0.11406249E 00-0.12702646E 01-0.26230106E 01-0.33405771E 01
COVAR 0.87007570E 00 0.78820543E 01 0.21607952E 01-0.70964956E 01-0.12105255E 02
COVAR-0.63822908E 01 0.46361450E 02 0.28476227E 02
```

Figure 23. Line Printer Listing of Punched Statistics by LARSYS

deemed unacceptable and not be considered further. This is indicated by the blank lines in the interclass divergence table in Figure 24.

Further, it is possible to apply

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

RESTORED DATA
RETENTION LEVEL .. 325
MINIMUM 25
MAXIMUM 350

	FEATURES	DIJ(MIN)	D(TOT)	INTERCLASS DIVERGENCE									
				SC (10)	SO (10)	SW (10)	SM (10)	CO (10)	CW (10)	CM (10)	OW (10)	OM (10)	WM (0)
1.	1 6 10 12	36.	1538.	36	84	194	179	120	...	347	95	133	0
2.	1 6 10 11	33.	1534.	33	81	196	184	120	...	...	90	130	0
3.	1 6 8 11												
4.	1 6 9 11	27.	1509.	27	69	201	189	83	...	...	104	136	0
5.	1 6 9 12	29.	1501.	29	73	200	175	82	...	343	111	138	0
6.	1 6 8 12	26.	1495.	26	82	170	206	93	311	...	98	159	0
7.	6 9 10 12	35.	1484.	35	66	175	153	108	...	...	116	131	0
8.	6 8 10 12	35.	1483.	35	67	160	164	109	...	...	108	140	0
9.	6 8 10 11	33.	1477.	33	60	166	183	105	...	...	104	126	0
10.	6 9 10 11	33.	1474.	33	58	179	168	105	...	...	108	123	0
11.	1 6 9 10	31.	1469.	31	54	220	192	112	...	...	71	89	0
12.	1 8 10 11	28.	1461.	28	71	166	212	95	295	...	93	151	0
13.	2 6 10 12	35.	1448.	35	76	166	154	114	...	325	95	133	0
14.	1 9 10 11												
15.	1 7 10 11	29.	1447.	29	74	174	183	101	310	...	87	139	0
16.	1 8 9 11												
17.	2 6 10 11	32.	1438.	32	69	169	158	112	...	332	90	126	0
18.	1 6 8 10	29.	1437.	29	47	206	209	106	...	...	52	88	0
19.	1 8 10 12	30.	1436.	30	70	160	208	92	274	...	95	157	0
20.	1 7 9 11												
21.	6 8 9 11	26.	1433.	26	51	164	176	71	...	...	113	132	0
22.	1 8 9 12	26.	1432.	26	70	162	205	79	279	...	102	159	0
23.	6 8 9 12	27.	1429.	27	59	158	151	76	...	344	119	145	0
24.	1 9 10 12	27.	1428.	31	67	171	193	92	299	...	100	125	0
25.	6 10 11 12	35.	1427.	35	72	152	147	106	...	326	111	128	0
26.	2 6 9 11	26.	1426.	26	55	175	159	72	...	...	105	134	0
27.	2 6 9 12	28.	1419.	28	63	172	144	74	...	336	112	140	0
28.	1 7 10 12	32.	1413.	32	75	169	179	99	290	343	89	137	0
29.	2 6 8 11												
30.	1 10 11 12	28.	1407.	32	74	166	174	94	300	343	107	117	0
*** 34.	1 5 10 12	31.	1403.	31	72	170	161	91	279	...	89	140	0

Figure 24. Sample LARSYS Feature Selection/Separability Output

varying weights to the interclass divergences as indicated by the numbers in parentheses under the symbols in this Figure. In this example, even though it is necessary to define two wheat categories, it is not desired to separate them from each other. Therefore, a weight of zero has been specified for the WM interclass divergence.

Note also that the feature set specifically called for, 1, 5, 10, and 12, is displayed and marked by asterisks (***). Its rank turned out to be 34.

CLASSIFICATION AND DISPLAY

Proof of separability of classes may be obtained by using the classification and display processor of LARSYSAA. The classification processor carries out the actual classifications using a Gaussian maximum likelihood scheme and stores the results on magnetic tape. The display processor is then used to read the results from these tapes, apply various options, analyze and display the results. Currently the display is presented in line printer form but the digital display device will also be used after its procurement.

A classification display example of the area previously shown in the pictorial printouts is given in Figure 25. For this classification, nine classes were defined; training samples from run 26600061 were selected for each class; and their statistics were computed and punched using the statistics processor. Four features (Channels 1, 6, 10, and 12) were used in the classification as indicated in the printout heading.

After classification, the display processor is used to evaluate the results. Symbols appropriate to the classes were assigned by the user. The choice of symbols is entirely arbitrary. For example, instead of viewing the results as in Figure 25, it may be desirable to display the same classification only using S for soybeans and blanks for all others in order to make more obvious the location of all soybeans.

In addition to assigning each resolution element to one of the nine categories, a threshold capability is also provided in the display processor. The classification of each element is compared to a user-assigned threshold. If the element classification does not exceed the threshold because the element does not look enough like a member of the class to which it has been tentatively assigned even though this is the most likely class, then final classification of the element is declined. The element is assigned to a null category and displayed as a blank. Different thresholds may be assigned individually to each of the classes if desired.

A satisfactory qualitative determination of the classifier performance may be obtained by studying such a display, but a more quantitative analysis of results is usually desirable. To facilitate this analysis, the display processor is able to receive the addresses of additional data fields called test fields and the designation of the class to which each field belongs. The display processor can then tabulate the performance of the classifier in these fields. The areas so designated as test fields in this classification are outlined with + on

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CLASSIFICATION STUDY .. SERIAL NO. 705807300
CLASSIFICATION DATE .. JULY 5, 1968

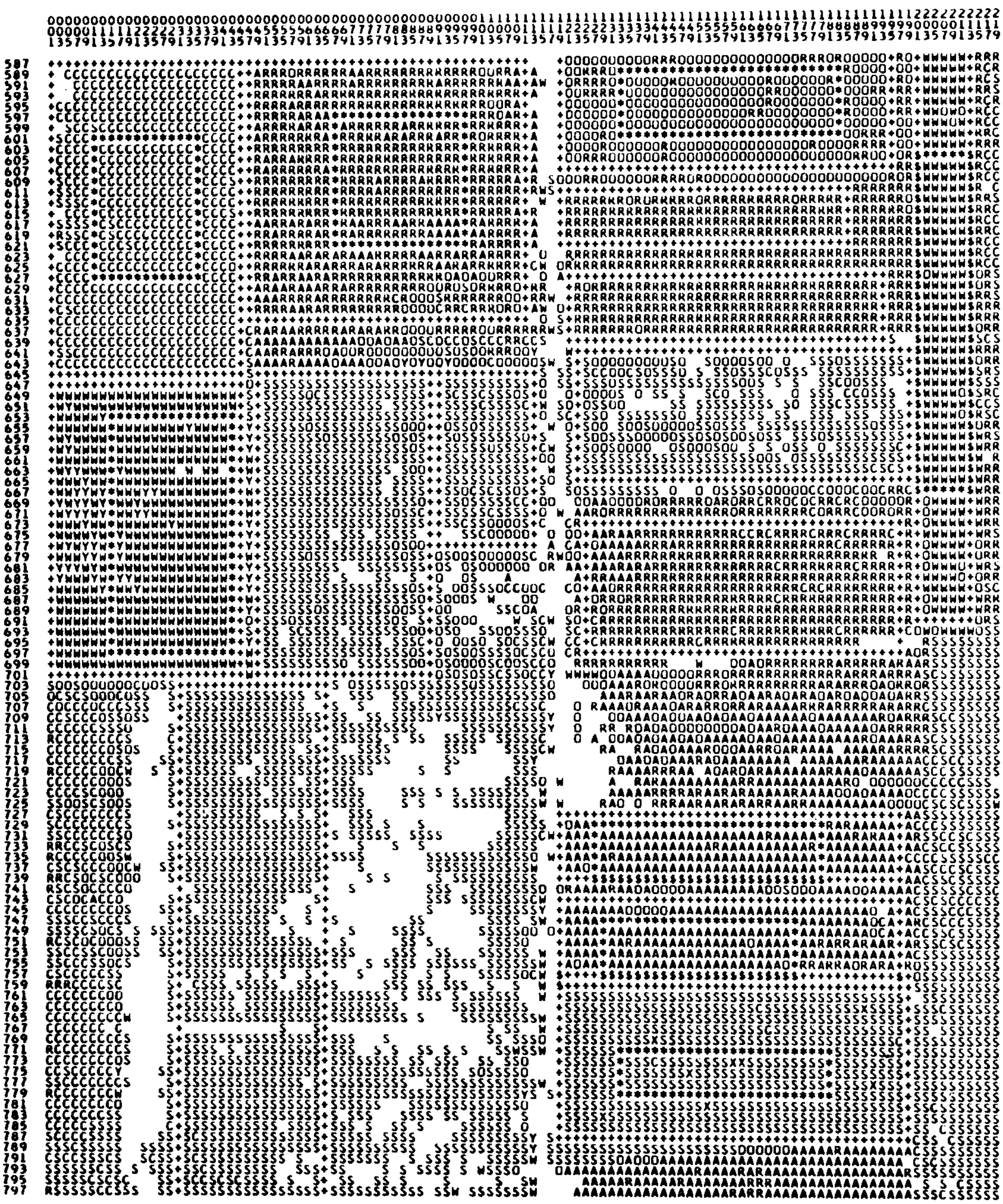
RUN NUMBER-----26600061 DATE------ 6/28/66
FLIGHT LINE---- C1 TIME------1229
TAPE NUMBER---- 102 ALTITUDE-- 2600 FEET

CLASSES CONSIDERED

SYMBOL	CLASS	THRESHOLDS
S	SOYBN I	14.900
C	CORN I	14.900
O	OATS	14.900
W	WHEAT I	14.900
R	RD CL I	14.900
A	ALFALFA	14.900
Y	RYE	14.900
X	BR SOIL	14.900
W	WHEAT II	14.900

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
6	0.52	0.55
10	0.66	0.72
12	0.80	1.00



TOTAL NUMBER OF SAMPLED POINTS = 11660

Figure 25. Classification Display

the display and the tabulation is given in Figure 26. The same type of tabulation but giving the performance on a per class basis is given in Figure 27. In addition to the test data fields, the program can produce the same type of outlining and tabulation for the training samples. Training fields are outlined with astericks (*) as indicated in Figure 25.

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CLASSIFICATION STUDY .. SERIAL NO. 705807300
CLASSIFICATION DATE .. JULY 5, 1968

RUN NUMBER-----26600061 DATE------ 6/28/66
FLIGHT LINE---- C1 TIME------1229
TAPE NUMBER---- 102 ALTITUDE-- 2600 FEET

CLASSES CONSIDERED

SYMBOL	CLASS	THRESHOLDS
S	SOYBN I	14.900
C	CORN I	14.900
O	OATS	14.900
W	WHEAT I	14.900
R	RD CL I	14.900
A	ALFALFA	14.900
Y	RYE	14.900
X	BR SOIL	14.900
W	WHEAT II	14.900

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
6	0.52	0.55
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TEST FIELDS

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOYB	CORN	OATS	WHEA	RED	ALFA	RYE	SOIL	THRS
7-27	SOYB	407	63.4	258	12	84	0	0	0	0	0	53
12-7	SOYB	513	88.9	456	4	31	0	0	0	0	0	22
12-2	SOYB	150	79.3	119	11	18	0	0	0	0	0	2
12-3	SOYB	752	89.2	671	8	0	0	0	0	0	0	73
7-23	SOYB	546	97.3	531	4	0	1	0	0	0	10	0
12-9	CORN	588	94.0	25	553	1	0	1	0	0	0	8
7-1	OATS	370	84.9	0	0	314	0	56	0	0	0	0
7-2	WHEA	260	93.5	0	0	17	243	0	0	0	0	0
12-10	WHEA	546	90.1	0	0	0	492	0	0	48	0	6
12-8	RED	713	80.2	2	3	27	0	572	109	0	0	0
7-29	RED	128	96.9	0	0	4	0	124	0	0	0	0
7-28	RED	175	98.9	0	0	2	0	173	0	0	0	0
	RED	385	86.8	0	17	5	0	334	24	0	0	5
7-24	ALFA	190	93.2	0	0	2	0	11	177	0	0	0
7-24	ALFA	266	83.8	0	4	19	0	18	223	0	0	2
	TOTAL	5989		2062	616	524	736	1289	533	48	10	171

OVERALL PERFORMANCE = 87.5

Figure 26. Tabulation of Classification Results of Test Fields

LABORATORY FOR AGRICULTURAL REMOTE SENSING
PURDUE UNIVERSITY

*** LARSYSAA ILLUSTRATION ***

CLASSIFICATION STUDY .. SERIAL NO. 705807300
CLASSIFICATION DATE .. JULY 5, 1968

RUN NUMBER-----26600061 DATE------ 6/28/66
FLIGHT LINE---- C1 TIME------1229
TAPE NUMBER---- 102 ALTITUDE-- 2600 FEET

CLASSES CONSIDERED

SYMBOL	CLASS	THRESHOLDS
S	SOYBN I	14.900
C	CORN I	14.900
O	OATS	14.900
W	WHEAT I	14.900
R	RD CL I	14.900
A	ALFALFA	14.900
Y	RYE	14.900
X	BR SOIL	14.900
W	WHEAT II	14.900

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
6	0.52	0.55
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TEST CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO								
				SOYB	CORN	OATS	WHEA	RED	ALFA	RYE	SOIL	THRS
1	SOYB	2368	85.9	2035	39	133	1	0	0	0	10	150
2	CORN	588	94.0	25	553	1	0	1	0	0	0	8
3	OATS	370	84.9	0	0	314	0	56	0	0	0	0
4	WHEA	806	91.2	0	0	17	735	0	0	48	0	6
5	RED	1401	85.9	2	20	38	0	1203	133	0	0	5
6	ALFA	456	87.7	0	4	21	0	29	400	0	0	2
	TOTAL	5989		2062	616	524	736	1289	533	48	10	171

OVERALL PERFORMANCE = 87.5
AVERAGE PERFORMANCE BY CLASS = 88.3

Figure 27. Tabulation of Classification Results on a Class Basis

LARSYS AA

A large number of applications of LARS research is possible and not all have been clearly defined. Generally, each application requires a pattern recognition system with significantly distinctive characteristics. In such a situation, it is necessary to have an efficient and flexible method of designing pattern classifiers, for performing statistical analysis and feature selecting and calculating other parameters.

This section describes how LARSYS

may be used to perform these functions. An important feature of this system is the considerable degree of user-system interaction which achieves the flexibility required by this type of research.

DATA HANDLING AND DATA ANALYSIS

Figure 28 is a block diagram of the overall data flow for the Laboratory for Agricultural Remote Sensing Data Processing System (LARSYS). The principal data input is multispectral data collected on analog tape by a multichannel optical-mechanical scanner and tape recorder mounted aboard an aircraft.1/ The LARS Aircraft Data Handling Processor (LARSYSAH) prepares the data for use by the researcher. The data are digitized, calibrated, and recorded on digital tape in a packed format to reduce the physical volume. To make the data readily accessible to the user, line-sample coordinates (much like x-y coordinates) are added during the digitization process. A special computer subroutine is available which will read any desired area of data as specified by a set of line-sample coordinates into core memory and pass it to the user's program in unpacked form.

Also available as part of LARSYSAH is a program which prints gray-level displays of selected data on a computer line printer. As noted previously, these are useful in coordinating the ground truth with the multispectral scanner data.

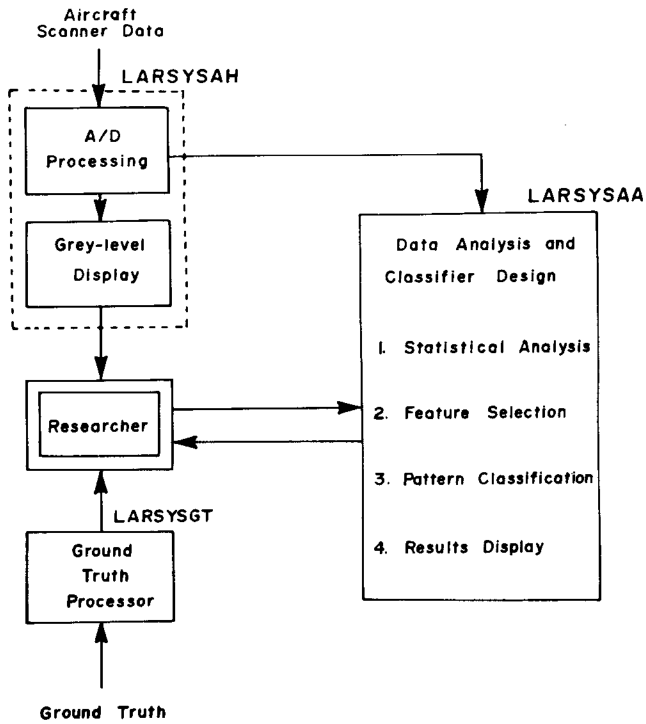


Figure 28. Data Flow in LARSYS for Aircraft Data Analysis

The other form of data utilized is "ground truth," which is collected on film and in the form of detailed written field reports. Ground truth, including such information as crop species, crop varieties, soil types, and percent ground cover is cataloged and made available to the researcher in convenient form by the LARS Ground Truth Processor (LARSYSGT).

1/ The airborne optical/infrared scanning equipment used in this research was made available by the U. S. Army Electronics Command (USAECOM), Fort Monmouth, N. J. on a no-cost basis to the University of Michigan (who collected the data) for use on contracts administered by USAECOM.

Mainly, the LARS Aircraft Data Analysis and Classifier Design System (LARSYSAA) which performs the function of pattern classifier design based on the data from the aircraft mounted multispectral scanner will be discussed.

THE LARS AIRCRAFT DATA ANALYSIS AND CLASSIFIER DESIGN SYSTEM (LARSYSAA)

The User - System Interface

There are at least three important reasons why the user-system interface has received considerable attention in the development of the analysis and classifier design system. These are:

(1) Optimal classifier design requires a substantial amount of interaction between the various phases of the design system.1/ At present, this interaction is best coordinated by the researcher.

(2) Remote sensing applications involve huge masses of data. As a result, the quantity of data input and the required results output for a classifier design task consumes a considerable amount of computer time. It is essential, therefore, that the analysis and classifier design system be largely immune from user errors (such as control card errors) so that errors in the later stages of processing will not result in loss of all the previous work.

(3) In the face of the previous two requirements, the experimental status of the remote sensing problem makes it desirable that most or all of the processing system be written in a high level compiler language so that modifications to the system may be made quickly and easily by the researcher.

The third requirement has been satisfied through use of FORTRAN IV except for a few minor utility functions which can be accomplished most efficiently through use of assembly language. The ways in which the other two requirements have been met are discussed in the next two sections.

LARSYSAA System Monitor - Free-form Card Format

Figure 8 illustrates the control structure of the LARSYSAA system. The figure indicates that the system is composed of a Monitor and four distinct processing phases, each processing phase directed by its own supervisor. The multiphase structure results largely from the need to minimize the amount of core memory occupied at any one time by program instructions, in order to maximize the amount of memory available for data. For the same reason, the individual processors are also decomposed into multiple phases which are only called into core memory as needed by the respective supervisors.

Processing under LARSYSAA commences when the computer operating system recognizes a job control card calling for the LARSYSAA system and loads the LARSYSAA Monitor into core memory from

1/ Three phases of classifier design are distinguished: statistical analysis, in which general statistical properties of the data are measured; feature selection, in which an optimal set of features are selected for use in the recognition process; and classifier synthesis, in which the classifier is designed (or "trained") and tested using the results of the preceding phases. The precise nature of the computations carried out in each phase depends on the particular feature selection and pattern recognition algorithms used by the researcher.

the program library.

The principal responsibility of the LARSYSAA Monitor is to recognize and interpret Monitor Control Cards which request loading of the processor phases from the program library. These control cards and all other control cards and data cards read by LARSYSAA may have an almost arbitrary format. A special key word must be punched first on control cards, followed by any other key words associated with the control card, separated by commas. On data cards, parameters are punched in a specific order, separated by at least one blank. Should the user inadvertently punch an unrecognizable keyword or inconsistent parameter, the card in error is printed on the console typewriter along with instructions as to the action necessary to resume processing. Diagnosis and correction of errors in this manner before they cause abnormal termination of processing prevents the loss of intermediate results stored in core memory by preceding stages of processing. The researcher will generally operate the computer and punch control cards as they are needed, If desired, he may type them in at the console. By this system the researcher is thus freed from the burden of remembering and adhering to rigid input format requirements and may concentrate more fully on the analysis and design problem at hand. The experienced user is able to operate the computer program with ease, the novice finds that the key word, free-form card input and attendant error checking speeds the process of effectively learning to use the system.

The Processor Supervisors

The responsibilities of the processor supervisor are threefold: interpretation of processor control cards, dynamic memory allocation, and processor control. Once all of the processor control cards have been read and any necessary error recovery performed, the processor scans the list of operations (processing options, see Table 4) that have been requested. The amount of core memory required to perform those operations is then calculated. If the needed core memory exceeds that available, the supervisor reports this fact to the user who may then reenter the processor control cards with appropriate changes. In pattern recognition terms, these changes generally involve trading off the number of processing options requested against the number of pattern features and/or pattern classes.

After a suitable set of processor control cards has been read, the supervisor calculates base addresses and sizes of all variable arrays so that core memory will be efficiently utilized. The size and address information is then available to the processor subroutines as needed.

The flexibility gained through the use of dynamic memory allocation significantly broadens the range of problems that can be handled, even in the face of fairly severe core memory limitations.

Once the memory allocation procedure is complete, the analysis and design operations requested are carried out under the direction of the supervisor.

The Processors

To discuss the processors in general terms, a possible analysis and design procedure will be described.

Table 4. LARSYSAA Processing Facilities

<u>Statistical Analysis Facilities</u>

Compute mean vector & covariance matrix for each class.
Compute mean vector & covariance matrix for each field.
Punch data deck containing statistics and other pertinent information for future use with Classification Processor.
Histogram selected features for each class.
Histogram selected features for each field.
Print spectral plots for each class.
Print spectral plots for each field.
Print as many spectral plots as desired, each displaying results for up to four different classes.

<u>Feature Selection Facilties</u>

Determine optimal sets of 1, 2, 3, ... features.

<u>Classification Facilties</u>

Perform pattern recognition using any subset of classes and features made available by the Statistical Processor.

<u>Display Facilities</u>

Print information as to source of training data.
Outline training sets if they appear in results display map.
Print results of training operations.
Use a specified symbol set for results display map.
Compute and print classifier performance evaluation for training set
(1) on per class basis
(2) on per field basis
List areas used as test samples for performance evaluation.
Outline on results map the areas used as test samples.
Compute and print classifier performance evaluation for test set
(1) on per class basis
(2) on per field basis
Apply likelihood thresholding to establish a rejection class.
Recompute and print performance evaluations on the basis of any specified grouping of classes.

Given a new set of digitized scanner data and the associated ground truth, the researcher's first task is to select a set of training patterns. To accomplish this, he obtains a set of gray-level printouts which aids in locating the boundaries of the agricultural fields, roads, bodies of water and so forth. Data from fields of known classification are then used by the Statistical Processor to calculate various statistical quantities for the pattern classes and produce the graphical data displays listed in the first part of Table 4 and discussed in the section on LARSYSAA System Monitor - Free-Form Card Format.

Because of uncertainty of ground truth details such as the precise locations of field boundaries, the researcher may at this point have limited confidence in the training set he has selected. However, by using the graphical output of the Statistical Processor, he may select a set of pattern features which appear to be useful for differentiating between the classes and, temporarily bypassing the Feature Selection Processor, use the Classification Processor and Display Processor to produce a classification based on the tentative training set. It has been found that even such "crude" classification can yield a printout or map of the data which is considerably more detailed than a gray-level printout. This is probably because the information contained in several spectral channels is condensed into a single display. The researcher can use this result to refine the training set, perhaps then performing one or more reiterations of the same procedure to achieve additional refinement.

Once a reasonable amount of confidence in the training set has been attained, the Feature Selection Processor is brought into the processing loop. The Feature Selection Processor provides information as to the best one, two, three, so forth, features to be used to obtain optimal classification for the specific problem at hand. Using this information and additional passes through the Classification and Display Processor, the researcher's principal remaining task is to decide, on the basis of reasonable computation time and the desired level of classification reliability, which optimal feature set (i.e., how many features) should be used.

PROGRAM MODIFICATION FACILITIES

As noted in the previous section, the research environment requires that the analysis and classifier design program be easily modified - for example, to test new classification algorithms. This flexibility is achieved by (1) the dynamic storage allocation approach discussed previously; (2) inter- and intra-program communication via common storage areas; (3) residence of the source language program on a tape which is easily modified by an editing program; and (4) a self-directed System Construction Program which, once initiated, performs all of the steps necessary to go from source language to operational program. Use of common storage areas for inter- and intra-program communication increases considerably the importance of the System Construction Program which has the responsibility of inserting all COMMON cards into the source language "deck" during system construction. This relieves the user of the chore of modifying the COMMON cards in every program and subroutine each time a change is made involving the common variables.

CONCLUSION

This section has presented some aspects of a system of computer programs which handles a fairly complex processing situation with demanding core memory requirements. The system allows the user to solve a wide range of problems without actual modification of the program. When modifications are unavoidable, the system structure allows the changes to be implemented easily with the aid of a System Construction Program. A "conversational mode" of operation which is of particular value in the research environment has been achieved. This has been done through the development of techniques which optimize man-machine communication and minimize the inefficiencies which usually result from a high level of on-line user-system interaction.

Results of pattern recognition studies in which this system was used are given in Chapter 7 of this volume.

SCANNER DATA OVERLAY

The aircraft scanner data being used for LARS multispectral pattern recognition research contains an inherent interchannel spatial alignment error which prevents simultaneous utilization of all channels of data. Only 10 of the 17 channels sensed are usable as a set without further preprocessing. A study was carried out by the data handling staff to determine the basic nature of the problem and to design a data handling system which would make all scanner channels available to researchers. A system of programs has been completed which can produce 17 channel data files with a moderate degree of spatial accuracy. Work is continuing on the problem to improve the speed and accuracy of the system. New aspects of the general image alignment problem are being studied. A full report of current work is reported in LARS Information Note 103068.

There are four separate sensor systems combined into the current (University of Michigan) multispectral scanner system. These four sub-systems detect energy in the following channels: (1) 80° field of view, .4 to 1.0 micrometers - 10 channels; (2) 80° field of view reflective infrared - 3 channels; (3) 40° field of view reflective infrared - 3 channels; and (4) thermal infrared, 1 channel, 40° field of view. The data from these four sub-systems, referred to here as apertures, is from the same general ground area but is recorded on separate tapes. The data on the tapes is not aligned so that a particular point on the ground is represented in coincidence in all 17 channels of data. The alignment must be very exact if pattern recognition research is to be carried out based on a 17 element measurement vector for each resolution element on the target. Thus the data from the four tapes must be combined on one tape in such a manner that the data vectors have coincident information in all channels.

The alignment problem is due specifically to the fact that the data is sensed by different units, and the channels are recorded separately, digitized separately, and reformatted separately. In this process, the starting points in scanner lines from different tapes may be different; sampling frequency may be different; and lines may be lost on some tapes due to bad data or digitizing problems. Furthermore, the formats of the data lines are different and the angular fields of view from the apertures are different. These factors make it impossible to use the four apertures for research by addressing the tapes separately.

Image matching is the general problem solved by the data alignment system. A human analyst would align two negatives of the same scene by placing one on top of the other and moving one with respect to the other, while viewing them against an illuminated background until a "match" was found. This is essentially what is required for LARS multichannel imagery. The term "overlay" has come into use to describe this process as a symbolic extension of the previous example. Mathematically the process is called auto-correlation for identical data sets and cross-correlation for non-identical sets. The scanner imagery is available after the analog to digital conversion process of a rectangular array as binary numbers, each number related to the reflectance or emittance levels in a small area surrounding the geometrical position of the point on the target surface. The array of points forms a digital image, and exact overlay of two digital images is the goal of the research effort.

The initial work on the problem took the obvious approach of cross-correlating the mis-aligned imagery and searching for a point of match in the two dimensional correlation function. The function which was used is:

$$\phi_{xy}(K,L) = \frac{1}{M \cdot N} \sum_{i=1}^{M} \sum_{j=1}^{N} X_{i+K,j+L} y_{ij}$$

$$-M+1 \leq K \leq M-1$$

$$-N+1 \leq L \leq N-1$$

where: $x_{i,j}$ = Data sample from the ith row and jth column of the reference aperture

y_{ij} = Data sample from the ith row and jth column of the aperture to be aligned with the x_{ij}.

M = Number of rows in the area being correlated.

N = Number of columns in the area being correlated.

$\phi_{xy}(K,L)$ = Cross correlation function for data from x and y apertures.

Two factors exist which influence the use of this function. One is that LARS computer memory size is limited. Therefore, processing of two dimensional arrays is constrained by the total memory available minus the program core requirement which presently leaves little space for two dimensional arrays. The other factor is that the misalignment in the imagery can be assumed to be small since initial alignment can be achieved visually to a few resolution elements and subsequent errors are small throughout a flight line. This characteristic allows the image array problem to be separated into two one dimensional problems. The image array is thus processed in the overlay system in terms of lines and columns rather than as a matrix. The above function is used with i or j, a set of a constant for correlating lines or columns respectively. Results obtained with cross-correlating lines and columns of data were inconsistent. Postive correlation was obtained for very few lines and columns, and it became evident early in the project that some means of enhancing some feature of the imagery would be necessary if automatic cross-correlation was to be achieved.

BOUNDARY ENHANCEMENT

Research was carried out to find a means of enhancing features in the imagery from the two apertures so that lock on could be achieved using the cross-correlation process. The similarities in the visible and IR data are geometrical and not spectral. The visual data values represent reflectance of energy whereas the thermal IR data represent emissive energy due to thermal activity in the target. A visible band area which has a high reflectance at a certain color is not necessarily at a high temperature or does not necessarily have a high emissivity. It was concluded that some means of transforming the data would be necessary in order to enhance the geometrical features of the target so that correlation could be achieved on the characteristic common to both types of data.

It is important to note that the ground areas being observed are highly cultured; man has altered the features of the land to suit his needs. In so doing man has partitioned the surface so that borders are seen which did not originally exist. These borders are edges of agricultural plots, road edges, buildings, city areas and the like. These borders exist in addition to natural borders such as riverbanks and edges of forests. It became apparent that these borders were characteristics of the target that could be used to enable cross-correlation and scanner data overlay. Borders in an area will be common to data from different sensors whereas the data characteristics from areas, say within a field, may be markedly different. This notion was used as a basis of the overlay research.

Several methods of performing border enhancement were studied. All of the methods fall into two categories: (1) those which utilize the data values and (2) those which use the derivative of the data. The nature of the changes in reflectance and emittance of the data strongly suggests that differentiation of the image should cause borders to stand out while suppressing the data within bordered areas. Differentiation experiments constitute the major part of the work done to date; however, step function curve fitting of the unaltered data is a process which was also studied. The methods of enhancement which were studied were:

(1) Processes using differentiated data
   a. Horizontal and vertical 1st difference
   b. Prefiltered 1st differences

(2) Processes using data values
   a. Step function curve fitting
   b. Filtering data before curve fitting

DIFFERENTIATION OF IMAGERY

Two factors exist which suggest that differentiating imagery will enhance borders. One is the previously indicated fact that scanner data tends to remain constant across agricultural fields and features of other large areas and "jumps" across boundaries to new values. The second is that in some cases, data are available from more than one coincident channel as in the case of the 10 channel .4 to 1.0 micrometer scanner. The derivatives from the multiple channels can be combined to form a composite derivative which will accumulate the effect of borders in each channel and provide a high "contrast" border enhancement process.

The most work in the overlay requirement project was the combination of 10 channel visible and reflective IR data with one channel thermal IR data. The visible band data enable boundary derivative accumulation by a factor of 10 whereas the thermal IR has only one channel. Thus a "sharp" and a "noisy" border enhancement result was expected for the two types of data, respectively. The first difference for discrete data corresponds to the first derivative and is taken of adjacent points in the horizontal and vertical dimensions in the program which was developed. The magnitude of the difference is summed since negative and positive differences both indicate a boundary irrespective of whether the data is increasing or decreasing in value. Also, the data value for the same border may be decreasing in one channel and increasing in another and to achieve an accumulation of border indication summing of positive values is required. The operations performed are:

$$\Delta LN_{ij} = \sum_{k=1}^{N} \left| x^k_{i,j} - x^k_{i,j-1} \right| \quad j=2,NPTS$$

ΔLN_{ij} = 1st line difference for ith line, jth sample (column)

k = channel index

N = Number of channels available

i = line

j = column

x_{ij} = sample value for ith line and jth column

NPTS = number of columns in line

The difference between lines for column j is:

$$\Delta CO_{ij} = \sum_{k=1}^{N} \left| x^k_{ij} - x^k_{i-1,j} \right|$$

ΔCO_{ij} = 1st column difference for ith line, jth column

The above expressions are implemented as part of the border enhancement phase and several flight lines of data have been processed. The horizontal (line) and vertical (column) differences are written on a data storage tape along with the sum of the two so that a complete border enhancement pictorial printout can be obtained for display purposes. The cross-correlation process, however, uses the differences separately.

Results of scanner imagery differentiation are shown in Figure 29. The center portion of Figure 29 is a gray scale pictorial printout of the Purdue C-1 test area south of the Wabash River (at top of strip) in Tippecanoe County, Indiana. On the left in Figure 29 is a thresholded printout of the differentiated data accumulated for the 10 coincident visible band channels. On the right is a thresholded printout of differentiated data accumulated from the three coincident channels from the reflective IR aperture. Most of the borders in the agricultural areas are clearly defined although "noisy" in that

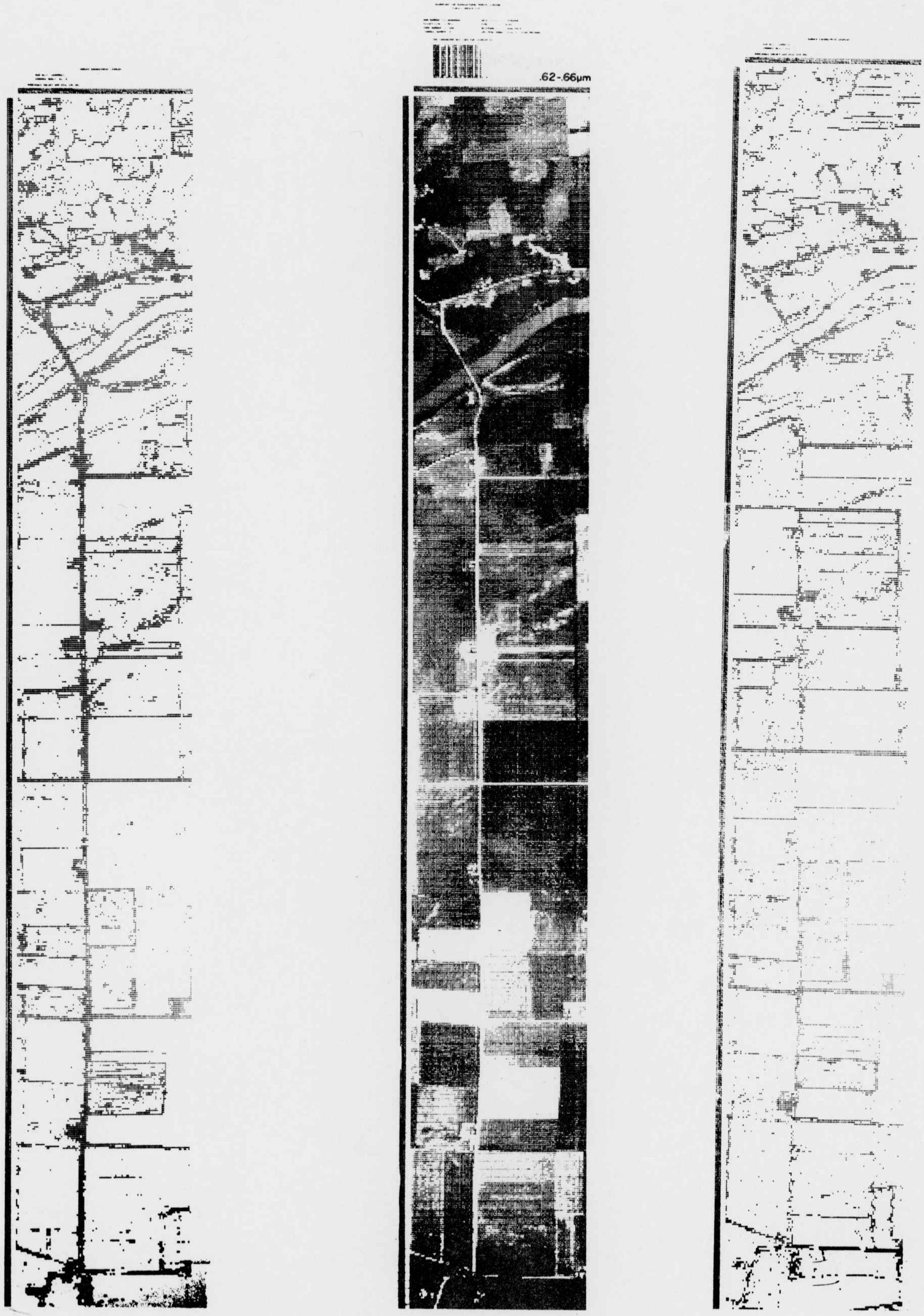


Figure 29. Border Enhancement Display - Flight Line PF 21

they are intermittent and too wide in many places. This case represents the best results obtained for border enhancement. The one channel thermal IR enhancement gave relatively poor results. When this method was used on other areas containing a large number of trees and irregular shaped borders, enhancement results were poor. However, these judgements are visible ones and the cross-correlation function can respond to border indications which the human eye would tend to reject as unacceptable.

The step function approach will not be discussed because work to date has not progressed to the point where feasibility can be determined. Low pass filtering was determined to significantly improve the performance of the experimental step function border finding algorithm. It did not sufficiently improve borders found by differentiation to warrant filtering for this method. Any useful step function method will most likely use some sort of filtering in an operational processor.

DATA OVERLAY

The image differentiation process was used to preprocess the data before cross-correlation. The overlay system is organized as shown in Figure 30. The data

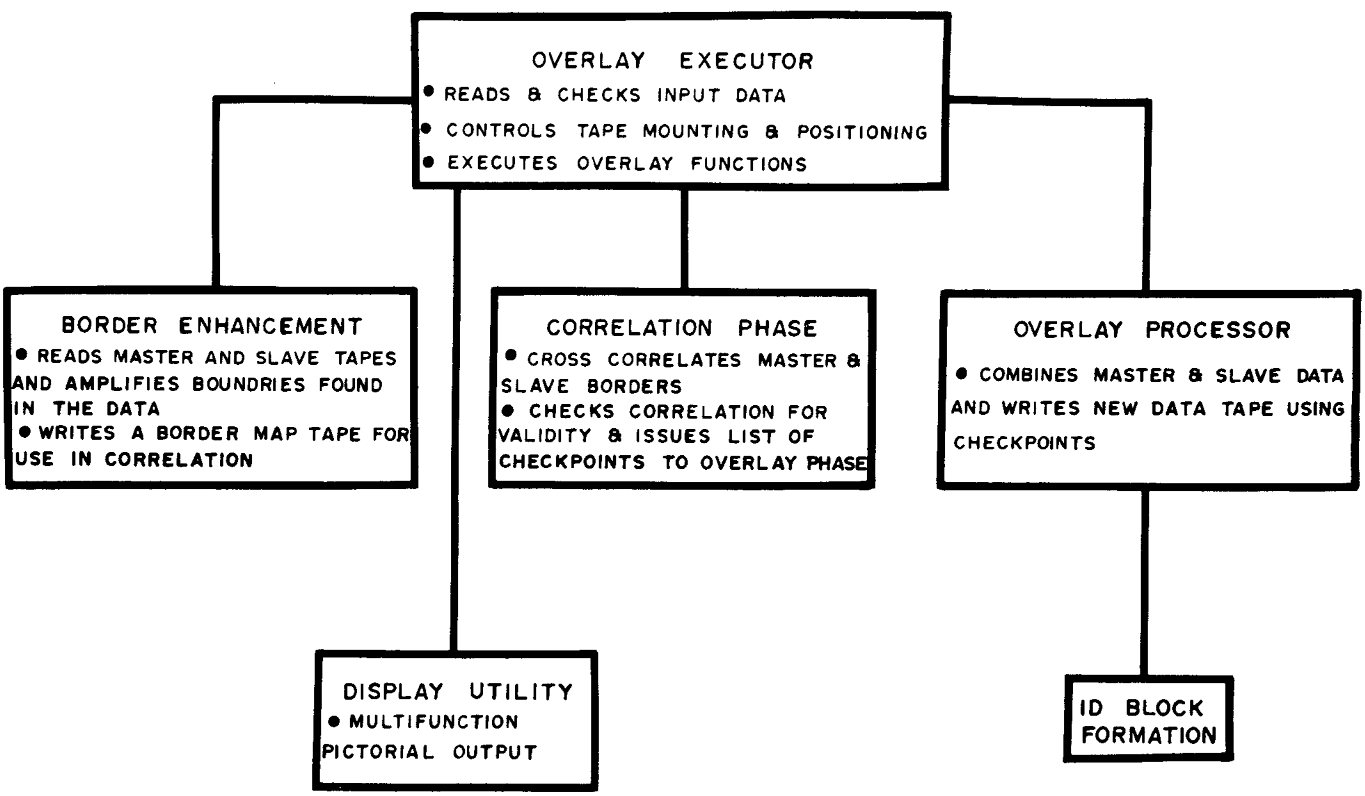


Figure 30. LARS Overlay System

is first enhanced by differentiation and sent to the second phase which cross-correlates the imagery to be combined. Spatial correction data are sent to the third phase where the imagery data are stored in coincidence on output tape. Three passes through the overlay processor are necessary to obtain a final tape containing data from the four apertures. The .4 to 1.0 micrometer data are first combined with that from the reflective IR apertures. Next the second reflective IR aperture data are added and finally the 8-14 micrometer thermal IR data are added forming a final 17 channel tape.

The final output data storage tape can be used by researchers to carry out data analysis and pattern recognition work with a full set of wavelength band measurements rather than with the previously available restricted set.

The average accuracy of the automatic overlay process is two to three resolution elements. Higher accuracy is achieved for areas containing many sharp straight line boundaries such as the C-1 agricultural area. For areas containing many irregular regions generally covered with trees or brush, the image correlation becomes less accurate. Since the data are dependent on the type of ground surface in the area, further work is required to insure that overlay can be carried out on a wide variety of imagery data.

FURTHER WORK

The boundary enhancement and image correlation work in the overlay studies have possible value for other problems of remote sensing. When a single aperture multispectral scanner is developed and in use, the need for overlay to solve the multiaperture problem will be eliminated. General imagery correlation, however, will continue to be a function of interest to researchers at LARS and elsewhere. Some applications of overlay which would be of interest are:

(1) Automatic combination of aircraft scanner data from flight lines over the same area taken at different times would be a very valuable capability. For example, this facility would enable research to be carried out on June, July, and September data from the same ground resolution elements. Samples would be callable in coincident data vectors from each ground element as the time coincident channels are presently callable. This feature would then add an additional, as well as a new, form of measurement dimension to the scanner data. The time dimension could replace a set of wavelength dimensions or the total NT* NC[1/] dimensions could be supplied to the researcher.

(2) Boundary detection in ground areas where pictorial output display does not produce sharp border representation would be possible. The overlay system performs a border enhancement process which produces a map of boundaries. This intermediate output form would be useful to persons doing identification of data and selection of training samples. Features which may not be recognizable at all in one or two channels can be enhanced.

(3) The step function border detection method which was studied produces the average value for each channel in the areas outlined by the

1/ NT = Number of times samples available for, NC = Number of channels at each time.

detected borders. Thus, only one sample from each "field" would have to be classified by PR techniques, and the results assigned to the total area enclosed by the border. This approach would be advantageous if the total time for border enhancement plus pattern recognition were less than for a total PR process for each resolution element in the field.

The data overlay project thus has possible benefits beyond the solution of the present scanner alignment problem. The project is exploring the region of the two dimensions (spatial and temporal) not yet studied at LARS. Study of the nature and detectability of geometric shapes could be stimulated by the boundary enhancement investigation, and the availability of time as a data dimension may be of value in multispectral pattern recognition work.

THE ANALYSIS OF NON-LARS DATA

Shortly after the procedures for digitizing, reformatting, printing, and analyzing data were used on LARS data, the question of their usefulness for data from other parts of the country was asked. During the year, data from several parts of the country were sent to LARS and partially or completely processed.

In August and September 1967, ten runs of data were digitized, reformatted and printed out for the University of California School of Forestry. Areas of agricultural ground cover and forest ground cover were processed. One area known as the Rice Field Test Site was classified into categories of water, soil, and vegetation without the aid of ground truth. The results of all processing were sent to the University of California for analysis or request for further work. A limited amount of ground truth and analysis by the California personnel indicated that the classification results were encouraging.

In November 1967, 91 runs of data were digitized, reformatted, and printed out for USDA-ARS at Weslaco, Texas. All area covered were primarily agricultural ground cover. Again, one area with an oxbow lake was classified into categories of water, soil, and vegetation without aid of ground truth except for an out-dated aerial photograph. Using this, the vegetation category was divided into trees and other green vegetation. Later, ground truth photos taken near the time of data collection showed these early results to be encouraging. All processing results were sent to Weslaco for analysis or request for further work. Further work in the form of choosing different gray level bins for some of the printouts was completed and sent to Weslaco.

Data gathered over a geologically significant test site, Mono Lake, California, were processed in December 1967. Four runs over the site were digitized, reformatted, and printed out before a visit from William Hemphill of USGS. During the visit, LARS and USGS scientists analyzed and classified the data into categories of soil types, rock outcroppings, and water. The purpose of the experiment was to test with a limited amount of data the use of multispectral pattern classification techniques developed by LARS on non-agricultural data. The study indicated that the techniques do have applicability to other disciplines.

In April 1968 a proposed highway

site in southern Indiana was flown for a project sponsored by the Indiana Department of Highways. The data, two runs approximately 70 miles long, were digitized, reformatted, and printed out. This was the first data collected in connection with a LARS project since 1966. The data were first analyzed from an agricultural point of view. Using only a small segment of the data for ground truth and training, the entire 70 miles was successfully classified into water, soil, and vegetation categories. The data were also used to classify different soil covers for LARS and significant geologic classes for the highway project. This work is reported in other sections of this report.

Data collected over an area of range land were processed in July 1968 for Oregon State University. Six runs were sent for analysis or request for further work. Some analysis of the data is completed, but classification has not yet started. In October 1968, four runs of data of geological significance were printed out for Lee Miller at Colorado State University. Thermally active areas of Yellowstone National Park were flown and processed.

A test of the LARS system on data collected by a scanner other than the University of Michigan scanner was conducted in November 1968. A flight line of data was flown by the Bendix Corporation and processed at Purdue to test the compatibility of the two systems. This was successful except for the analog tape specifications. The Bendix analog tape unit was required during the digitization process. The data were collected during questionable weather and only a small amount of the data could be successfully classified.

These experiments indicate the flexibility of the system to accept different kinds of data and the usefulness of the techniques over areas having other than agricultural significance. In the next year greater amounts of data will be utilized in research, and techniques will be developed to extend the state-of-the-art of data handling and analysis to larger areas and other disciplines.

CHAPTER 4

BIOGEOPHYSICAL RESEARCH PROGRAMS

INTRODUCTION

The purpose of the Biogeophysical Research Program is to determine the capability for mapping various vegetation and soil features, through the use of multispectral scanner systems and automatic data handling and processing techniques. Research is therefore conducted to study the relationships between the spectral response patterns and the physical and chemical properties of the plant and soil materials. The specific objectives are as follows:

(1) To determine the various agricultural features that can be differentiated and identified on the basis of the spectral, spatial, and temporal data.

(2) To define the optimum portions of the electromagnetic spectrum for such identifications.

(3) To determine the amount of seasonal and geographic variation in such multispectral response patterns and to define the optimum time of year to obtain remote multispectral data on crop species and ground covers of interest.

(4) To determine the agronomic causes of variations in multispectral response and to define the importance of these variations in identifying the crop or its condition.

In pursuit of these objectives, data must be collected and studied in the laboratory, field and aircraft. This includes scanner, spectrometer, and meteorological measurements, photographs, crop identification and field descriptions. Thus, the collection and analysis of such data must be in conjunction with the Measurements, Data Processing, and Support Programs personnel. The information reported in this chapter is therefore the result of efforts by many members of the LARS research team.

Other aspects of the analysis of the 1966 and 1968 multispectral scanner data and the interpretation of the color and color IR film are discussed in Chapters 6 and 7.

VEGETATION, BARE SOIL, AND WATER MAPPING OF A 70-MILE FLIGHT LINE

INTRODUCTION

One of the analysis tasks undertaken by LARS has been the classification of multispectral scanner data obtained over a significantly large geographical area. This was to answer a major question posed to researchers concerning the capability to classify large geographic areas and to extrapolate a limited number of training samples from a large test area. Much work done with laboratory spectra and aerial photography, scanner imagery, and digital multispectral data has indicated marked spectral differences between many types and conditions of soil, water, and vegetation. There are statistically significant spectral variations with each of these very basic categories. The objectives of this research were to: (1) determine the capability to accurately identify all green vegetation, bare soil, and water in the flight line regardless of natural variability and (2) determine the capability to accurately classify data over a large geographic area and utilizing the same very small set of training samples taken from one segment of the flight line.

TASK

For this task a set of scanner data over a 70-mile flight line in central Indiana was used.[1/] The scanner data were collected from an altitude of 3200 feet in 17 discrete wavelength bands. For this analysis task, only the first 12 wavelength bands, in the 0.4-1.0 micrometer range, were considered. At 3200 feet altitude, the path width is approximately one mile, thereby producing a seventy square mile area to be classified.

Figure 31 shows the location of this flight line area which followed the path of a proposed highway location. The flight line extended south from Indianapolis along the White River, through a predominately agricultural area, then through forest and pasture land with more rolling terrain.

ANALYSIS

For the initial analysis, approximately ten training areas for each class were selected. These included light and dark soils areas, river and pond water, and a variety of green vegetation conditions including pasture winter wheat, and trees. The color photographs in Figures 32 and 33 give an indication of the variability of these three cover types. All training samples were selected from one relatively small segment of the data. Frequency histograms[2/] and multispectral response graphs were obtained for each training class of cover type and analyzed

Figure 31. Location of 70-Mile Flight Line

1/ The aircraft scanner data were collected by the Institute of Science and Technology, University of Michigan, and the photographic data were collected by the Indiana State Highway Commission. These data were made available to LARS by the Airphoto Interpretation and Photogrammetry Laboratory, School of Civil Engineering, Purdue University. These data were obtained for research sponsored by the Indiana State Highway Commission, and Bureau of Public Roads, U. S. Department of Transportation.

2/ See Chapter 3 Data Processing for more information on histograms.

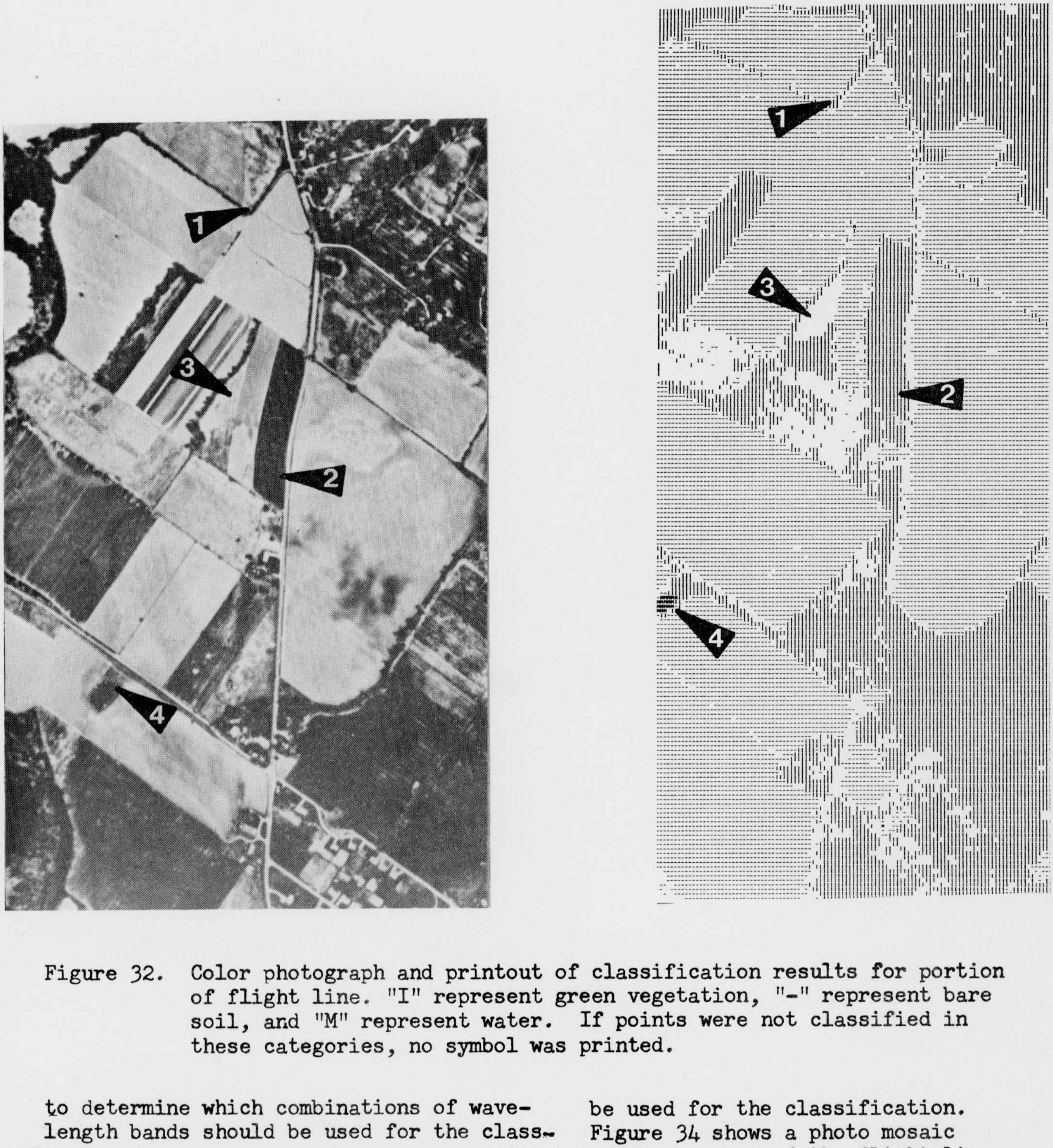


Figure 32. Color photograph and printout of classification results for portion of flight line. "I" represent green vegetation, "-" represent bare soil, and "M" represent water. If points were not classified in these categories, no symbol was printed.

to determine which combinations of wavelength bands should be used for the classification. It was decided that the 0.40-0.44, 0.58-0.62, 0.66-0.72, and 0.80-1.0 micrometer wavelength bands should be used for the classification. Figure 34 shows a photo mosaic of one segment of the flight line along with gray scale printouts of two of the wavelength bands used in the classification.

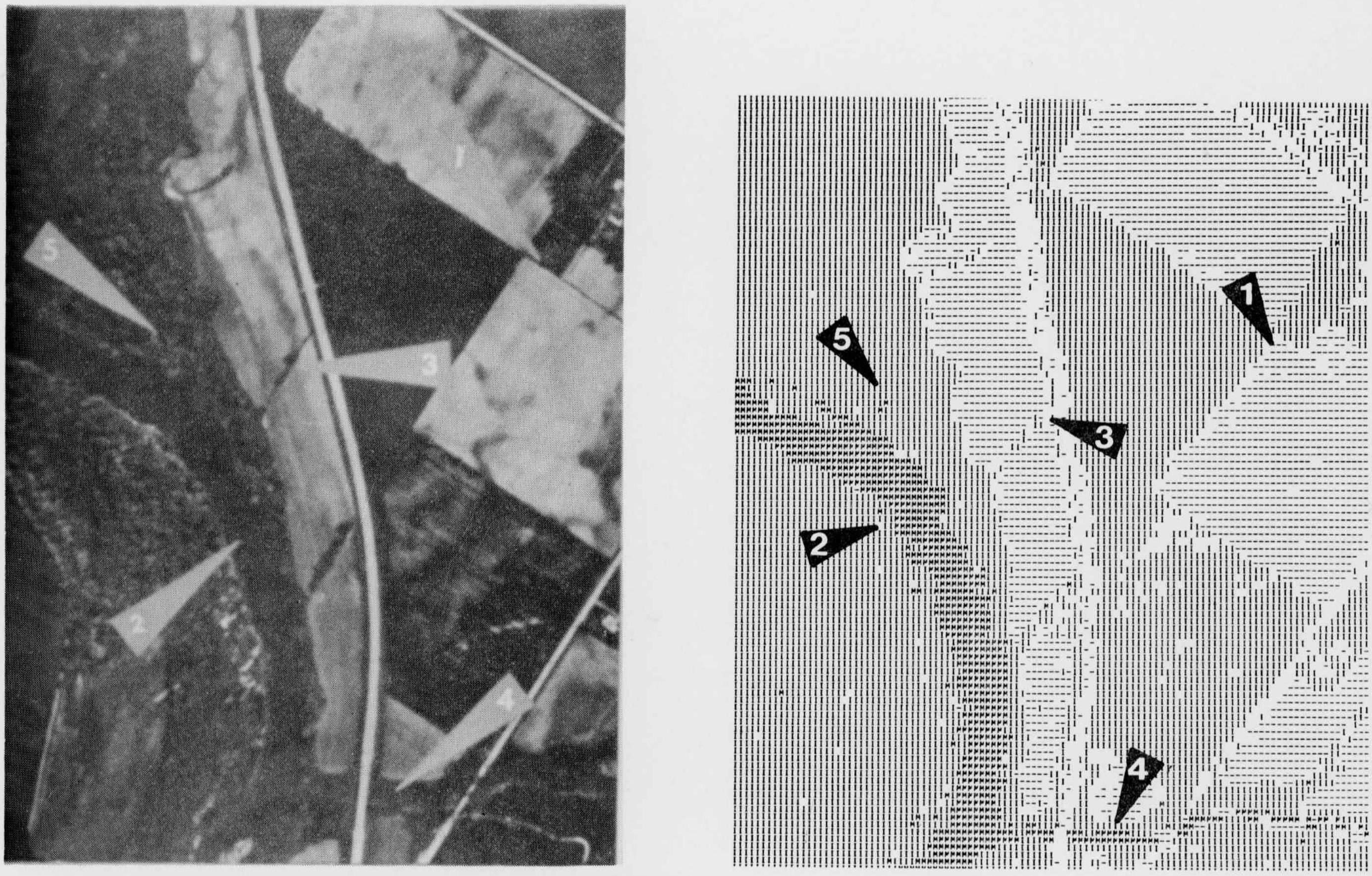


Figure 33. Color Photograph and Printout of Classification Results for Portion of Flight line

Figure 35 is the multispectral response graph of the selected training samples for these three categories and shows the marked differences in spectral response that may be observed in some of the wavelength bands. Note especially the 0.80-1.0 micrometer reflective infrared band, where water is very absorbtive and therefore has a relatively low reflectance and low response on the graph. Green vegetation is highly reflective and has a relatively high response on the graph. Figure 35 also indicates that throughout the visible portion of the spectrum (0.40-0.72 micrometers) the green vegetation and water have very similar spectral responses. This helps to explain why some of the water ponds surrounded by green vegetation are not more obvious on the color photography, but do appear with striking contrast on the color infrared photography.

The entire 70 miles of flight line data were then automatically classified into green vegetation, bare soil and water categories. Figure 36 shows the classification results for the same area seen in Figure 34. It should be pointed out that the photo mosaic on the left in Figure 36 was compiled from data

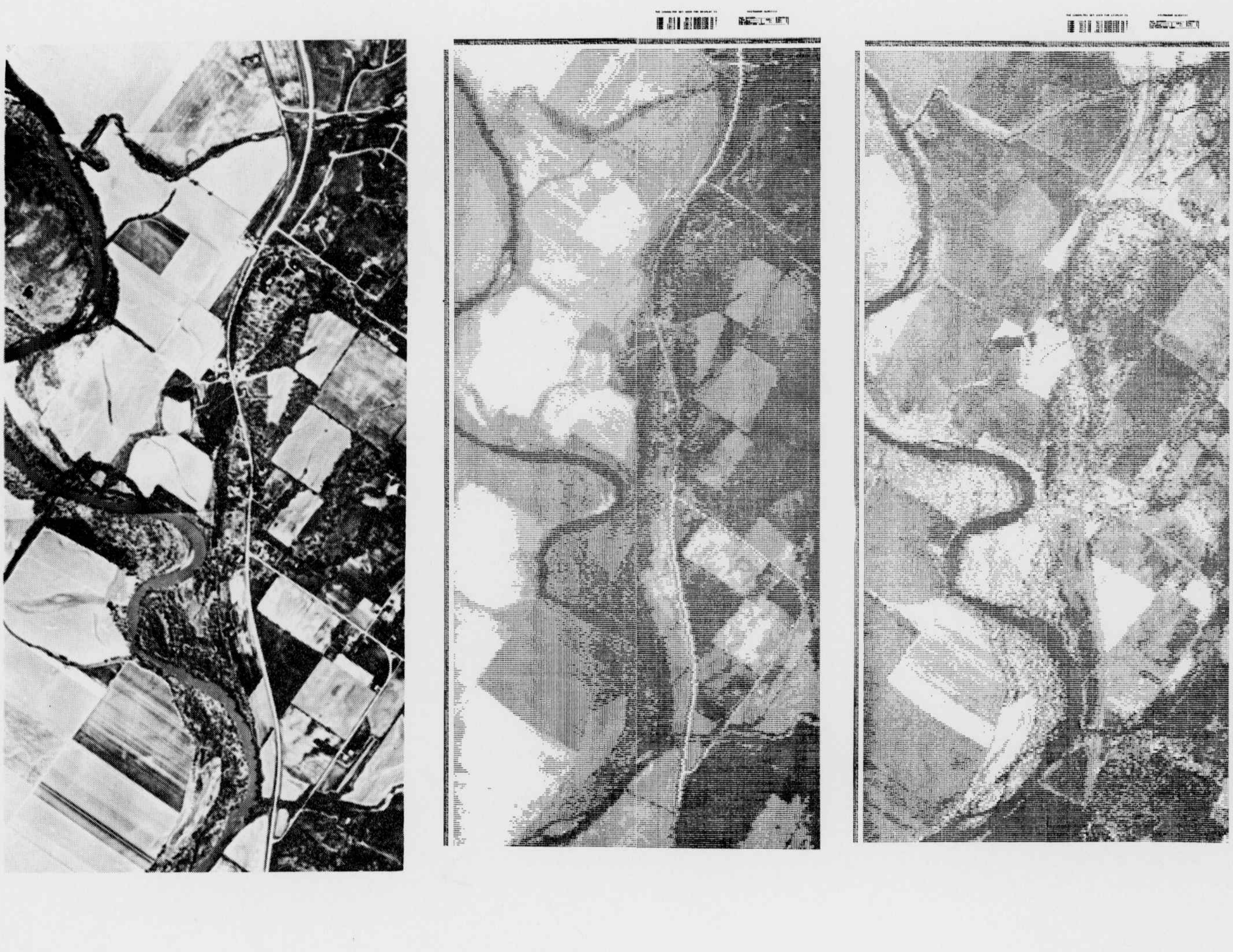


Figure 34. Aerial Photo Mosaic of Northern Portion of the Flight Line and Two Gray Scale Printouts

There is no page 62

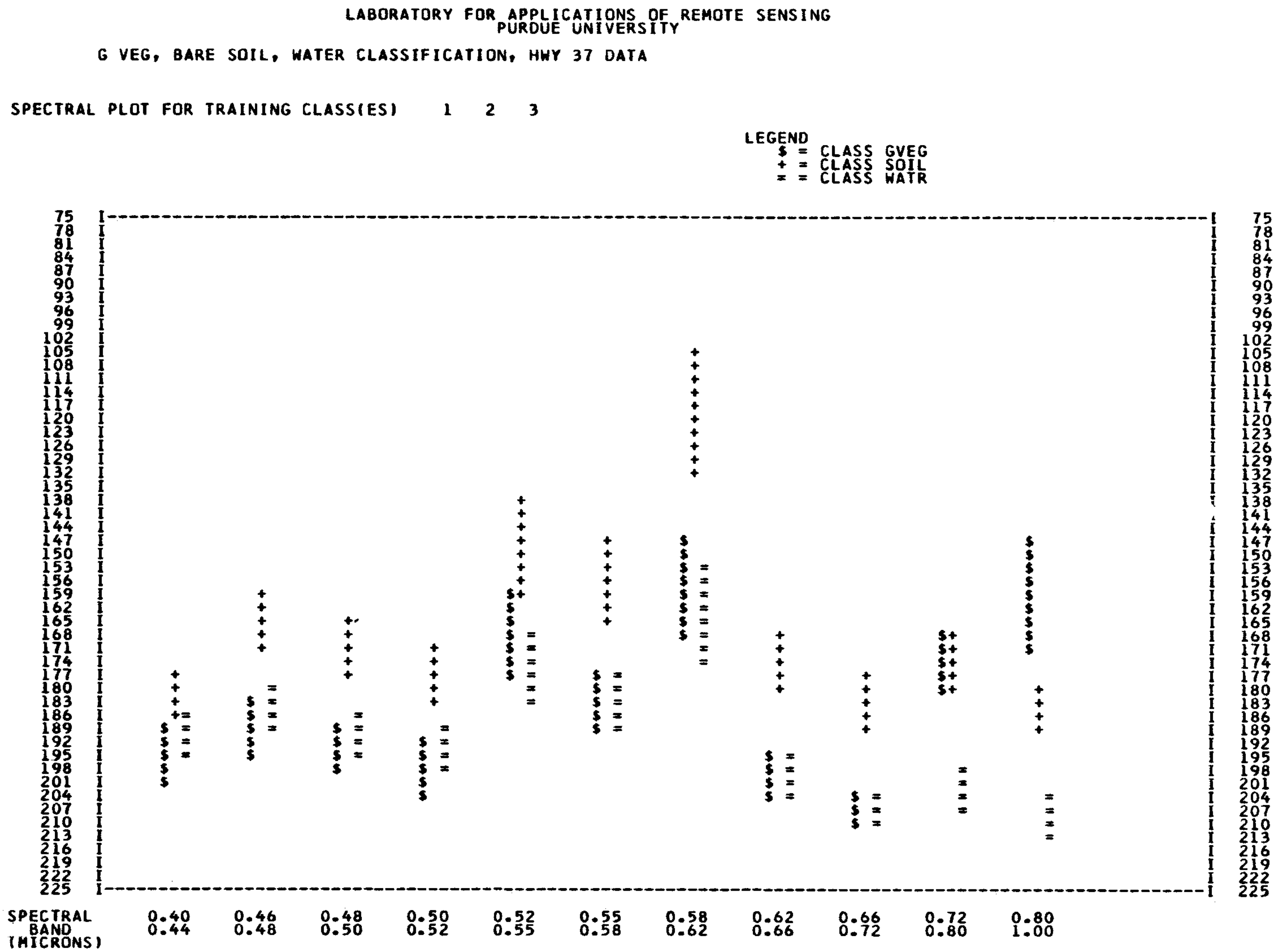


Figure 35. Multispectral Response Graph of Selected Training Samples for 3 Categories of Green Vegetation, Bare Soil, and Water

flown for planning purposes about ten days before the scanner data were flown. There were several areas which were green vegetation when photographed but were plowed (bare soil) before the scanner data was flown. Therefore, these particular areas, which appear to be green vegetation on the photos and bare soil on the printout were correctly classified.

The image labeled "B" on Figure 36 is the computer printout showing the automatic classification results for this area. The light tones were created using the computer symbol "-"

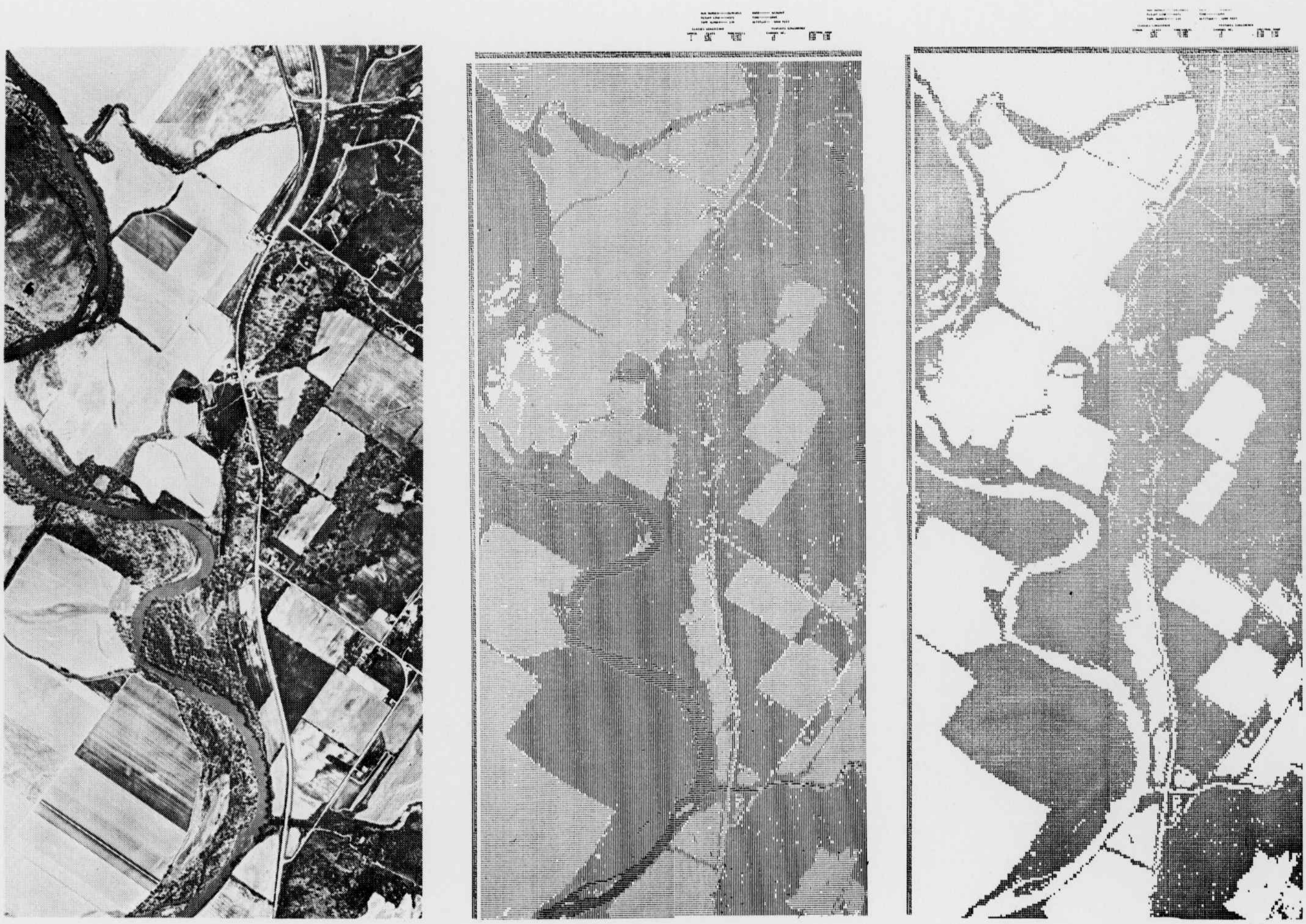


Figure 36. Aerial photograph and Classification Results of Area Shown in Figure 34

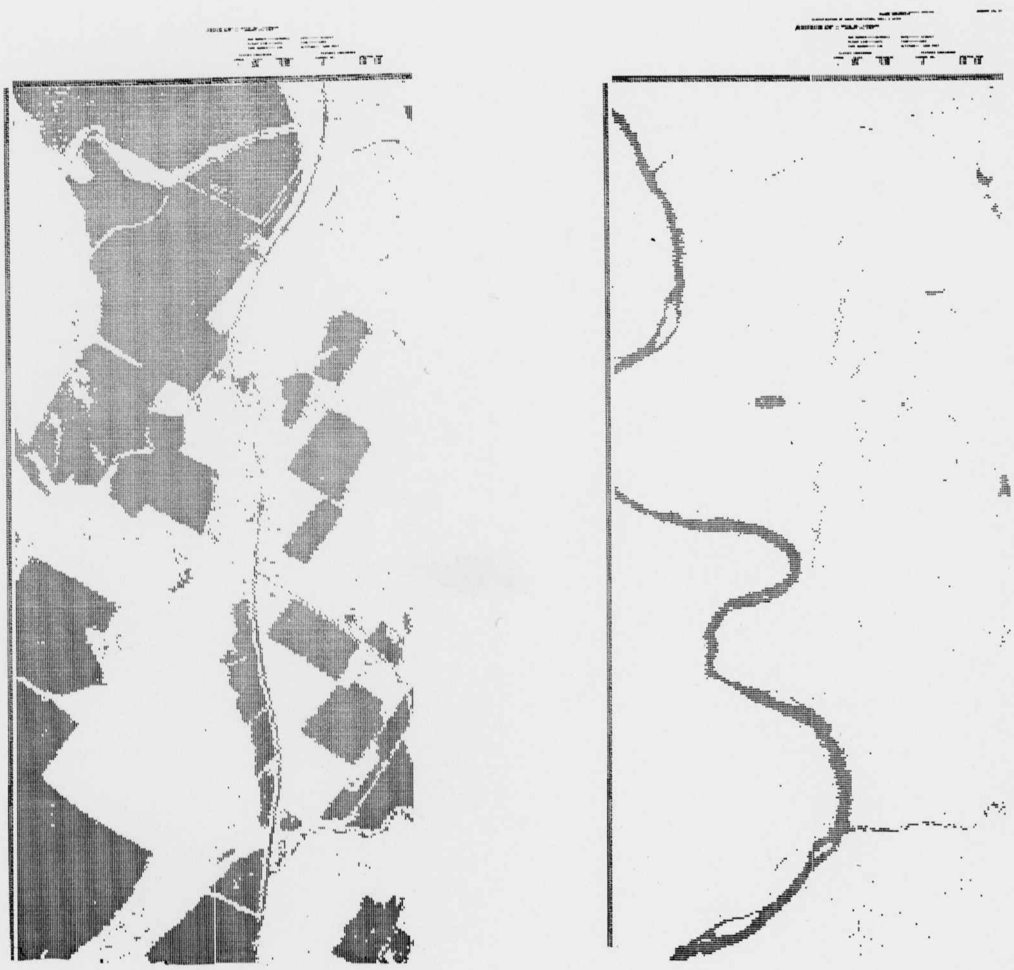


Figure 36. Continued From Previous Page

and indicate the areas classified as bare soil. The medium tones (symbol "/") indicate the areas classified as green vegetation, and the dark tones (symbol "M") indicate water. All other ground cover conditions (such as dry, dead vegetation, roads, rooftops, etc.) have been thresholded and are shown as "blank" (no symbol) areas on the computer printout.

The printouts labeled "C", "D", and "E" in Figure 36 show only the areas classified into bare soil, green vegetation and water, respectively. These are displayed in this manner to emphasize particular cover types which have been classified and to evaluate the accuracy of the automatic classification.

Figure 32 and 33 are representative examples of the classification results which show more detail than could be seen in Figure 36.
In Figure 32 a hedgerow (Arrow 1) shown on the photograph was correctly identified on the computer printout on the right. Arrow 2 indicates a field of winter wheat, which is at a very green dense stage of maturation. The symbol "/" used to identify green vegetation on the computer printout indicates that this area was accurately classified. The dry brown vegetation shown at Arrow 3 has been "thresholded" in the automatic classification process, which means that the multispectral response in this area was unlike the response for green vegetation, bare soil, or water.1/ As may beseen in this figure, the roads and houses were also "thresholded" and are shown as "no symbol" areas. Arrow 4 is a pond of water which was correctly classified and indicated by the symbol "M" on the printout.

In Figure 33, Arrow 1 is at the corner of a field of green vegetation and two fields of bare soil and Arrow 2 is the river. Arrow 3 shows a small hedgerow between two areas of bare soil, which was accurately identified on the printout. A tributary which was correctly identified is indicated by Arrow 4. Arrow 5 is a very small pond of water that is almost indistinguishable on the color photograph.

Many of the points identified as water appeared as scattered individual RSU's (Remote Sensing Units or individual computer symbols on the printouts) or as very small groups of RSU's. These were at first thought to be misclassifications. Additional checking revealed that there was ponded water or water in drainage ditches at most of these points. It is of interest to note that there were several such instances in which water was correctly identified automatically, but was previously overlooked on the aerial photographs. Several of these water areas were difficult to see on the aerial photos because of overhanging tree branches or a lack of distinctive color differences between the water and other materials in the vicinity.

Some problem areas were found to exist in the water classification, primarily because of shadows. For instance, heavy, distinct cloud shadows tended to be classified as water. This is probably because both water and shadows reflect relatively small amounts of energy, particularly in the infrared wavelengths. A similar

1/ See Chapter 3, Data Processing, for more detail about thresholding.

problem was encountered in some of the forested areas where deep shadow areas between the tree crowns were misclassified as water.

By placing a moderately heavy threshold on the water category, the shadow areas could be thresholded and shown as a blank because no symbol was printed by the computer on the printout. This is shown in Figure 37 where a heavy cloud shadow near the southern end of the flight line was classified as water, as seen on the printout labeled "A". On printout "B", a threshold on the water category has caused most of the cloud shadow to be thresholded. However, by following this procedure, we also found that some RSU's that were correctly classified as water were thresholded. This would indicate that a larger, more completely representative set of training samples for water should be obtained.

To indicate the quality of these classification results over the entire 70 mile flight line, Figure 38 shows the classification results for an area south of Bedford, near the extreme southern end of the flight line. The previous illustrations (Figure 36) of classification results were from an area about ten miles south of Indianapolis, near the northern end of the flight line. The single set of training samples from a relatively small area did allow for an accurate classification to be obtained over the entire 70 mile flight line.

QUANTITATIVE RESULTS

The above results indicated a rather good capability for automatic identification of these spectrally simply cover types. To obtain a more quantitative check on the classification results of the entire area, 89 test sample areas were randomly selected from the computer data. Based upon study of the photos and the ground truth data collected at the time of the scanner flight mission, 24 test areas were determined as belonging to the green vegetation category, 29 as soil, and 36 as water. The automatic classification of sample points showed that 83 of the 89 test areas had a 92 percent or higher classification accuracy. The average classification accuracy was 99.2 percent for green vegetation, 97.0 percent for soil and 99.7 percent for water. These classification results are shown in Tables 5 and 6. Table 5 contains the summary classification results for the training areas used in classifying the entire 70 mile flight line. Table 6 contains the classification results for all of the individual RSU's contained in the 89 test sample areas.

In examining the printout results from areas approximately 50 miles south of where the training areas were selected, it was observed that a larger proportion of the RSU's were being thresholded. This could indicate slight changes in spectral response of materials along the length of the flight line.

Four possible reasons for spectral changes are: (1) change in atmospheric conditions between the northern and southern portions of the flight line; (2) slight changes in the spectral characteristics of the materials due to geographic variability (such as trees being leafed out more in the southern part of the flight line, or the water quality of the East Fork White River perhaps being somewhat

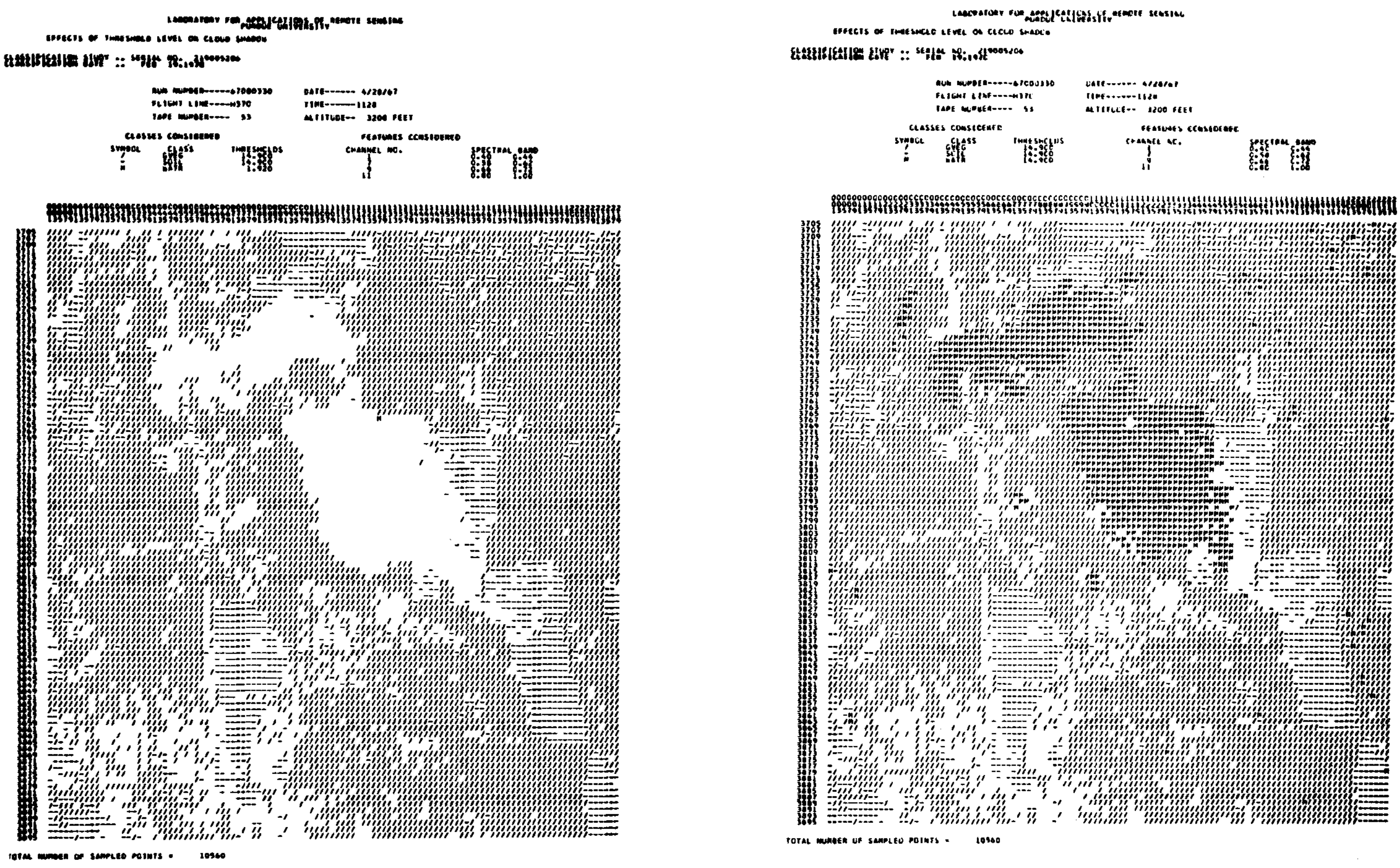


Figure 37. Heavy Cloud Shadow Near Southern End of Flight Line (A). Results When Threshold Put on Water Category (B).

Figure 38. Classification Results on Southern End of the 70-Mile Flight Line

Table 5. Summary of Classification Results for the Draining Areas Used to Classify the Flight Line

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

TRAINING FIELD CLASSIFICATION RESULTS

CLASSIFICATION STUDY .. SERIAL NO. 309005214
CLASSIFICATION DATE .. MAR 9,1970

RUN NUMBER-----67000100 DATE------ 4/28/67
FLIGHT LINE---- 3RD TIME------1055
TAPE NUMBER---- 202 ALTITUDE-- 3200 FEET

CLASSES CONSIDERED

SYMBOL	CLASS	THRESHOLDS
/	GVEG	14.900
-	SOIL	14.900
M	WATR	99.900

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
7	0.58	0.62
9	0.66	0.72
11	0.80	1.00

CLASSIFICATION SUMMARY BY TRAINING CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO GVEG	SOIL	WATR	THRS
1	GVEG	731	100.0	731	0	0	0
2	SOIL	846	100.0	0	846	0	0
3	WATR	168	100.0	0	0	168	0
	TOTAL	1745		731	846	168	0

OVERALL PERFORMANCE =100.0
AVERAGE PERFORMANCE BY CLASS =100.0

different than the West Fork White River) thereby producing differences in spectral response; (3) adjustments in instrumentation setting, although this will normally produce dramatically different results; and (4) electronic drift in some of the data gathering, recording, or processing instrumentation. In this case, the third possibility was not believed to be a factor, but any of the other possibilities alone or in concert could have caused the observed slight shift in spectral response.

In this classification, the shift in response was slight, and occurred over a relatively long flight line. The results were not greatly affected, because the materials being classified were spectrally very different at least in some wavelength bands. However, these observed changes indicate a need for further study of this type of phenomena, so that proper instrument calibration can be planned as well as proper techniques developed for delineating an optimum set of training samples. If the problem

Table 6. Classification Results of Individual RSU's in the 89 Test Sample Areas

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY
MAR 9,1970

3 CLASS RESULTS, ENTIRE 70 MILE FLIGHTLINE, USING SEGMENTS

CLASSIFICATION STUDY .. SERIAL NO. 302005213
CLASSIFICATION DATE .. MAR 4,1970

RUN NUMBER-----67000330 DATE------ 4/28/67
FLIGHT LINE----H370 TIME------1128
TAPE NUMBER---- 53 ALTITUDE-- 3200 FEET

CLASSES CONSIDERED

SYMBOL	CLASS	THRESHOLDS
/	GVEG	14.900
-	SOIL	14.900
M	WATR	99.900

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
7	0.58	0.62
9	0.66	0.72
11	0.80	1.00

CLASSIFICATION SUMMARY BY TEST CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO GVEG	SOIL	WATR	THRS
1	GVEG	1095	99.2	1086	1	5	3
2	SOIL	3105	97.0	2	3012	0	91
3	WATR	887	99.7	1	1	884	1
	TOTAL	5087		1089	3014	889	95

OVERALL PERFORMANCE = 97.9
AVERAGE PERFORMANCE BY CLASS = 98.6

cannot be corrected by instrument calibration techniques procedures for updating the training samples may have to be developed. This might involve such procedures as using a new set of training samples every twenty miles of the flight line and weighting the new training samples more than the previously existing training samples.

SUMMARY

Seventy square miles of data were automatically classified into three basic cover types--green vegetation, bare soil, and water. This was the first time that such a large geographic area has been automatically mapped. The training samples were selected from an area near one end of the flight line. In general, the classification results were highly accurate for the entire flight line. There were some problems with shadow areas being classified as water, but careful thresholding generally alleviated most of these errors. It was also observed that there was a slight degradation of the results as the classification got farther and farther from the end where the training samples were obtained. This was indicated by a larger percentage of

RSU's being thresholded near the southern end of the flight line. There are several possible reasons for this change in accuracy suggesting that additional research into this problem is necessary. The principal result from this work, however, was that a significantly large geographic area was automatically mapped into the basic cover types with a relatively high degree of accuracy, thereby demonstrating the capability to successfully extrapolate from training samples taken in one small segment of the flight line to the entire 70 square mile area.

AGRICULTURAL SPECIES IDENTIFICATION

Figure 39 shows a flight line area about 4 miles in length and one mile in width. An aerial photo mosaic of the flight line area is shown on the left. Individual crop species and cover types were identified by personnel gathering ground truth at the time of the flight mission. These are indicated on the photo mosaic by letters, using C for corn, S for soybeans, W for wheat, O for oats, H for hay, A for alfalfa, P for pasture, RC for red clover, T for timothy, and DA for diverted acres. To the right of the photo mosaic in Figure 39 are four gray-scale computer printouts, showing the relative amplitude levels of reflected radiation in the flight line for each of the indicated wavelength bands. These computer printouts are similar to cathode-ray tube scanner imagery. The data have been digitized and the individual computer symbols shown are referred to as RSU's, or remote sensing units. These RSU's represent measured radiance values for small resolution elements on the ground. For each RSU, the radiance level in each of these four wavelength bands can be combined into a multispectral response pattern. This pattern is then categorized by a pattern recognition algorithm, and the individual RSU's are classified as to crop species or cover type.

A closeup of the classification results that were obtained by the LARS-Purdue computer for a portion of this four mile flight line is shown in Figure 40. On the lower left, the letter symbols annotated on the aerial photo indicate the crop that was actually present. In the computer printout is the identification by the computer of the individual RSU's, using W for wheat, O for oats, etc. From these results, the field can be quite readily distinguished on the computer printout. Comparing the series of "W"'s in the computer printout, you see that actually it was a field of wheat as designated on the annotated aerial photo, or a group of "S"'s correspond to a field of soybeans. Notice in particular that the tones of gray between the field of soybeans and the field of corn on the left side of the aerial photo are very much alike. Likewise, the wheat and oat fields on the aerial photo have about the same tone of gray. This would make identification of these crop species from small scale black and white aerial photos extremely difficult. However, by using essentially the same type of "tonal measurements," but using multispectral scanner data obtained in many relatively narrow bands of the spectrum, it appears to be possible to differentiate and identify crop species rather accurately using automated pattern recognition techniques.

In Figure 40, the area identified as DA (diverted acres) was not classified into any particular crop type

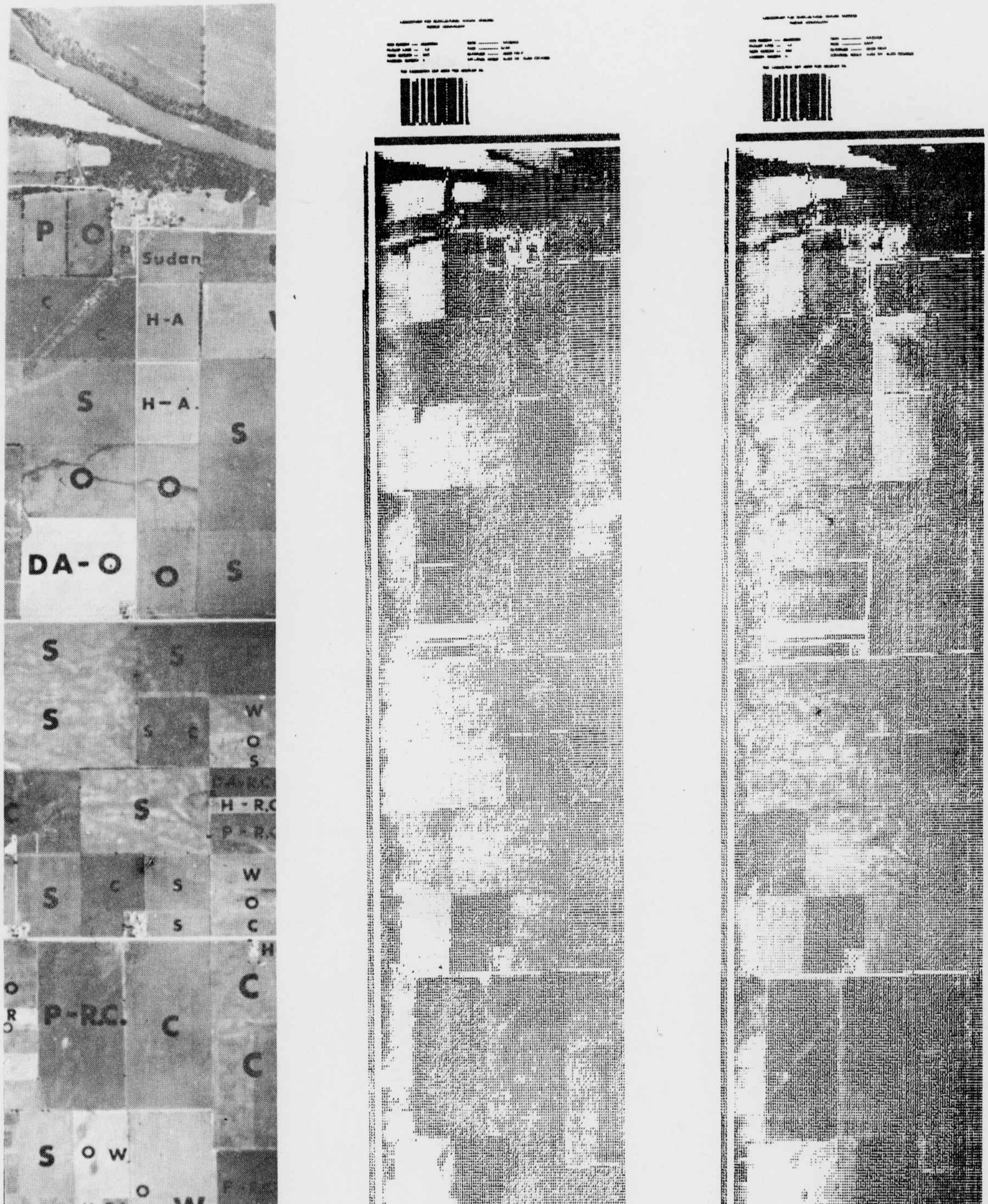


Figure 39. Photograph with Ground Truth Symbols and Four Gray Scale Printouts of Area

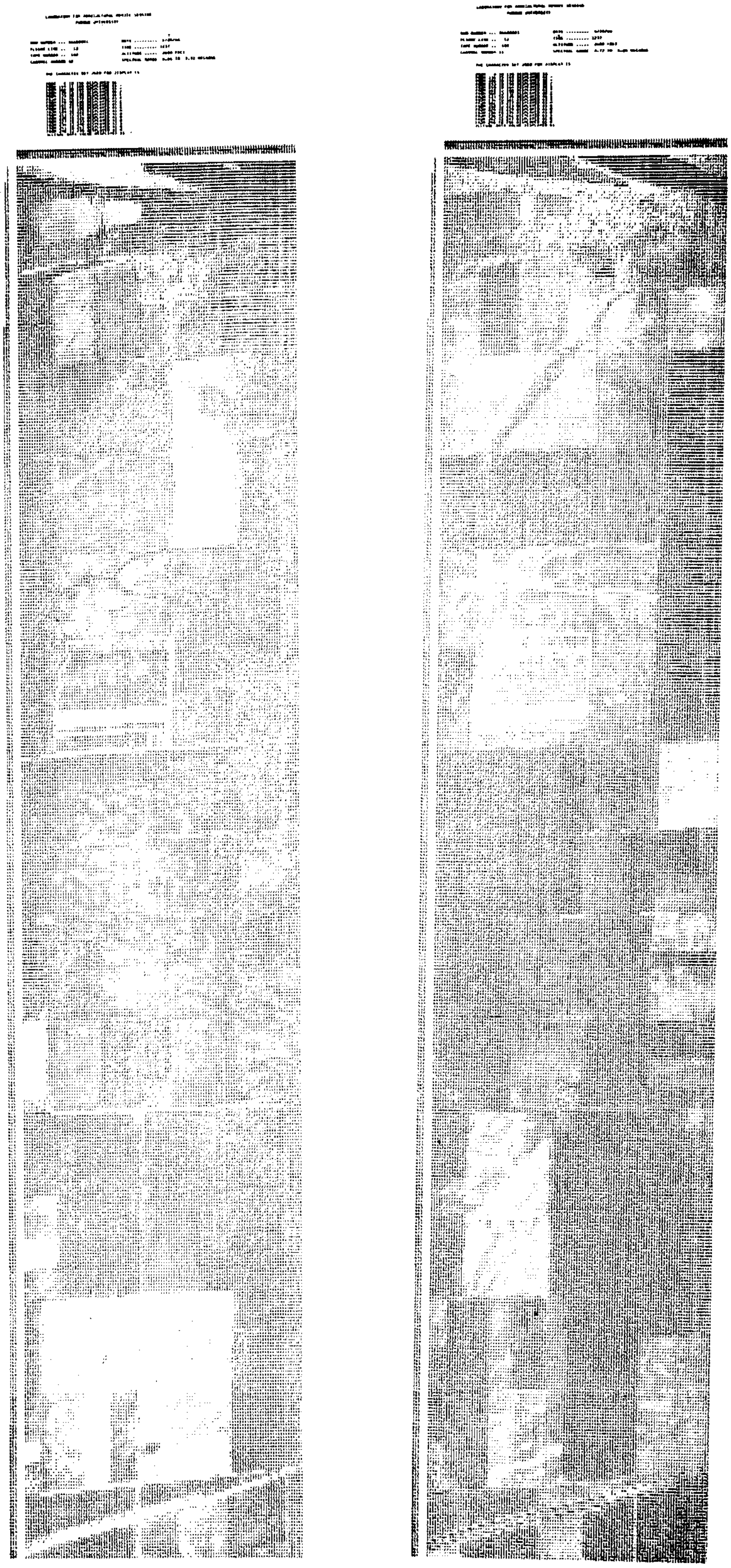


Figure 39. Continued From Previous Page

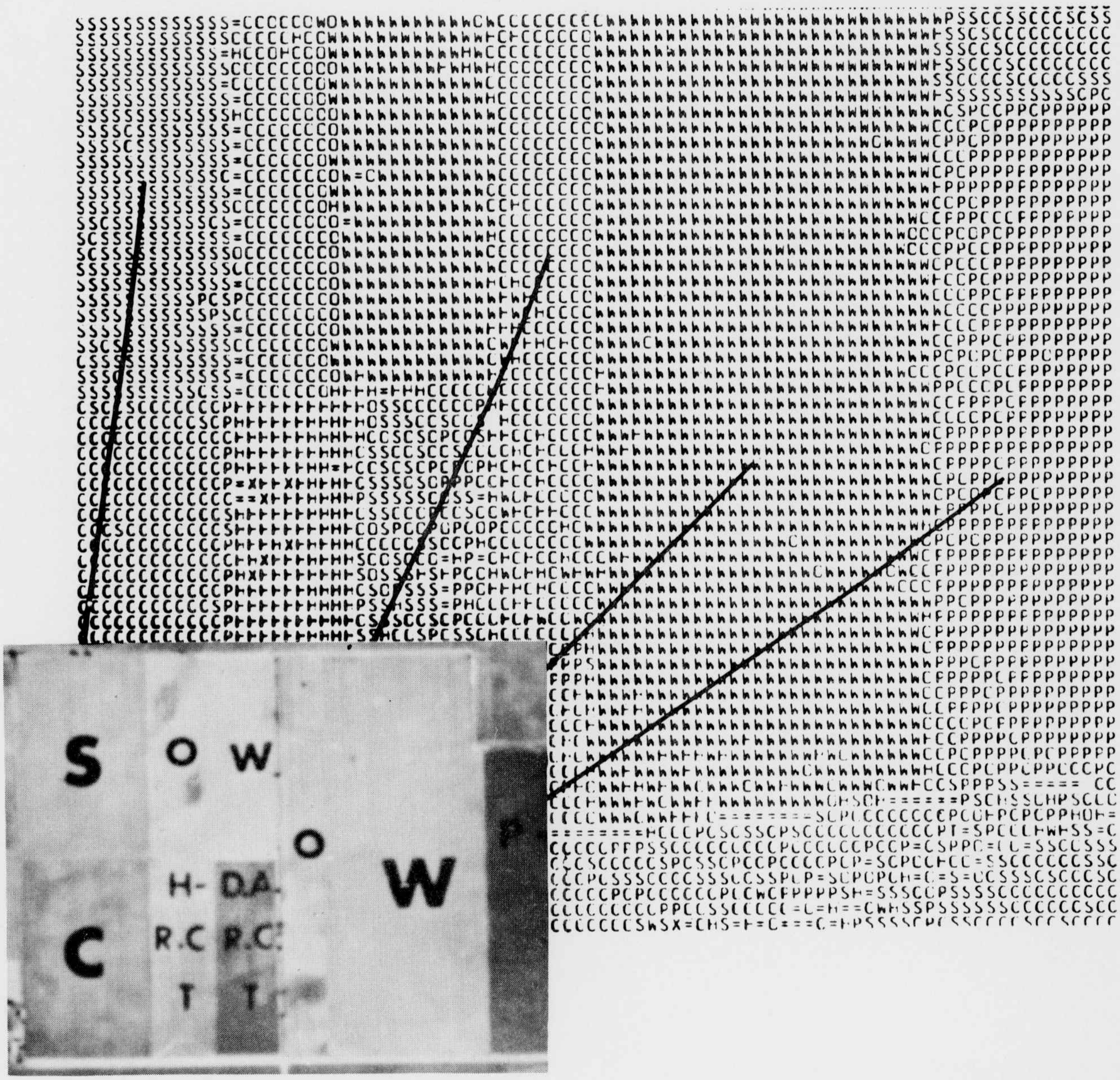


Figure 40. Close Up of Computer Printout of Crop Species Identification Results and Corresponding Aerial Photograph

because the computer was not trained to recognize diverted acres as a distinct cover type. Therefore, the individual RSU's in this field were classified according to the cover type which they looked most like. As will be described later, this field would have thereby fallen into a group called "indeterminate," in which case, it would not have been identified as to any one crop cover or species group.

The particular classification scheme used in obtaining these results was designed to identify ten different types of ground cover as follows: corn, soybeans, wheat, oats, pasture, hay, bare soil, roads, trees, and water. Because of the natural vari[illegible]ty in agricultural crops, it is o[illegible] impossible to identify all fields of wheat, for example, using a single training class for wheat. In this classification task, three subclasses were used to describe various conditions of wheat, at their different stages of maturity. Likewise, there were three subclasses of oats, five of corn, and four of soybeans. Hay was divided into hay-red clover and hay-alfalfa classes. Therefore, although there were only 10 agricultural categories being classified, a total of 24 training classes or subclasses were utilized.

Computer classification results for the entire flight line are shown in Figure 41. The annotated aerial photo mosaic is on the left, while in the center is a computer printout showing the automatic identification of areas of row crops. The fields in the flight line area for which ground truth was obtained and which were positively identified by ground truth personnel as corn or soybeans are outlined by hand on an overlay. They are designated with a "C" or "S". A couple of areas can be seen which were not corn and soybeans but were largely identified as a row crop. A field on the left side about two thirds down the page actually was soybeans, but few RSU's in the field were identified as such. In the printout on the right, the computer was instructed to print out an L for each RSU which was identified as a small grain, either wheat or oats. All fields in the flight line area which were in fact either wheat or oats were largely identified as such. There were a few RSU's in each of the fields that were apparently not correctly classified. Likewise, there are a few points scattered throughout the flight line which were erroneously classified into the small grain category. These results indicate accurate classification for either row crops or small grains. Such an evaluation, however, is rather qualitative and creates a desire to obtain a more quantitative classification result.

Figure 42 shows the aerial photomosaic for this flight line and a computer printout with rectangles designating the areas in the flight line for which quantitative classification results were summarized. The coordinates for the corners of these fields were designated, and the number of RSU's classified into each class within the designated field areas were tabulated into the crop categories. A large portion of the entire flight line was included in the quantitative classification results. There are a few areas which were not included because they consisted of diverted acres, barnyards, houses, and others. Because of the natural variabilities of diverted acres it is extremely difficult to classify this land use category. Therefore, diverted acres is not included in the quantitative analysis results. (A photo illustrating this problem is shown later.)

Figure 43 is a summary of the quantitative analysis. The percentage of correct classifications for each cover type is given. The accuracy range is from 100 percent for water to 80.7 percent for hay.

Figure 44 illustrates how these figures are obtained. (Because this table is a computer printout and the program allowed only four letter headings, wheat was spelled as "whet.") On the left of Figure 44 the ground cover class is designated. Next column contains the number of samples or RSU's which were known to belong to that class according to the test areas within

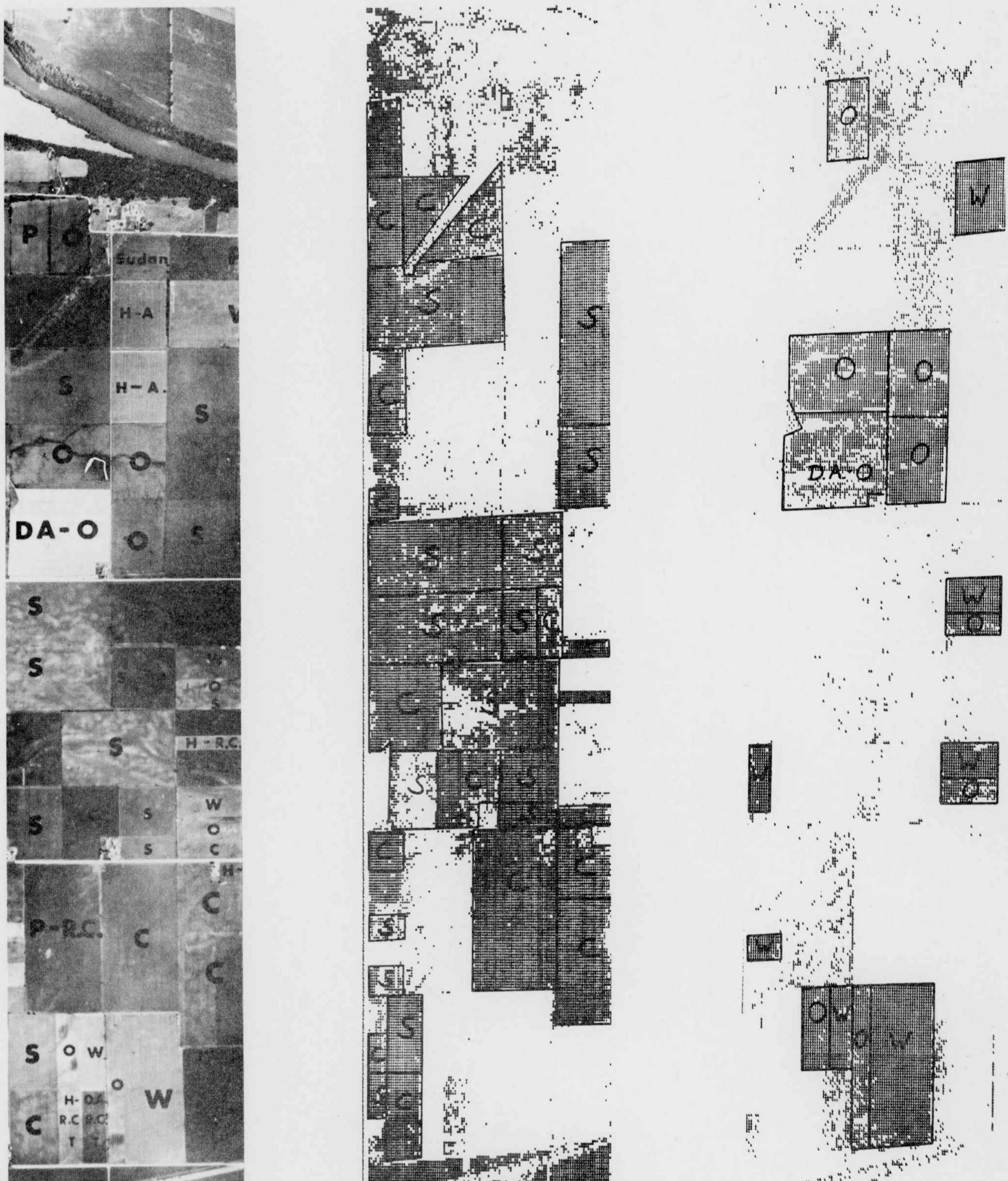
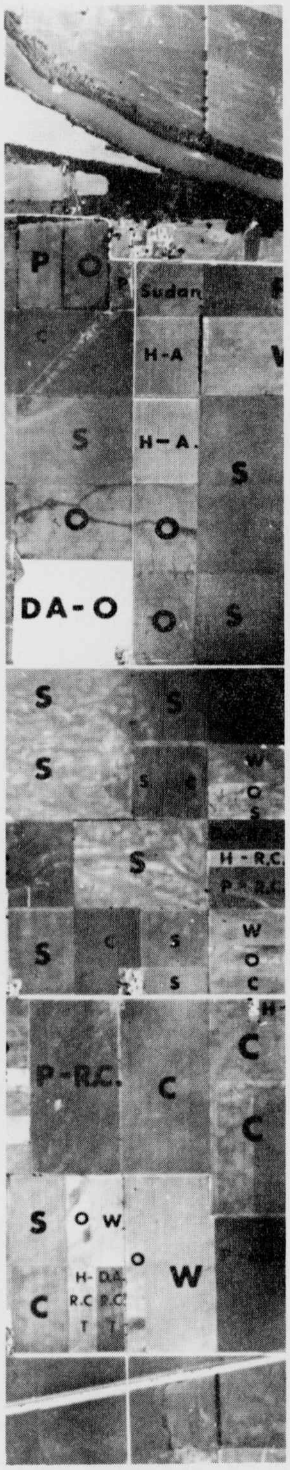


Figure 41. Automatic classification results for row crops and cereal grains. The symbols represent: S - soybeans; C - corn; O - oats; W - wheat; A - alfalfa; T - timothy; RC - red clover; R - rye; Sudan - sudan grass; P - pasture; DA - diverted acres; and, H - hay.



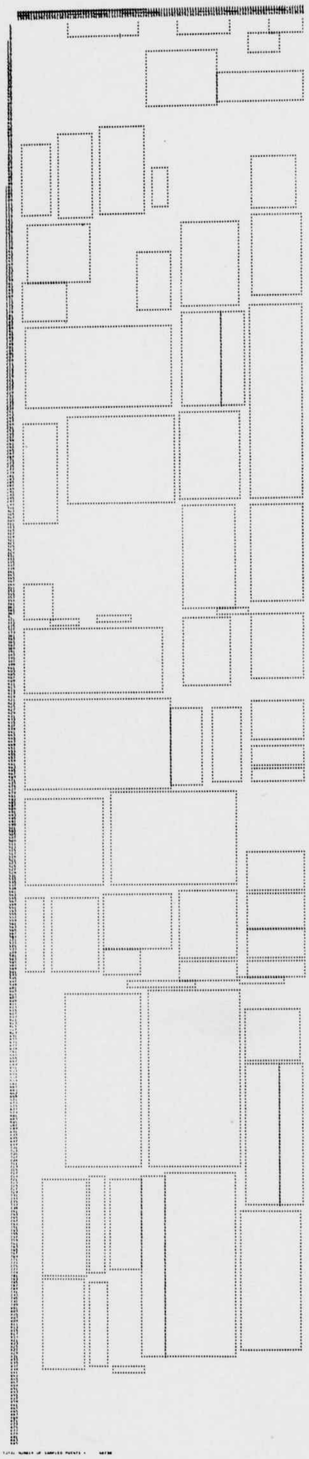


Figure 42. Photo Mosaic and Computer Printout with the Test Areas Outlined from which Quantitative Results were Summarized

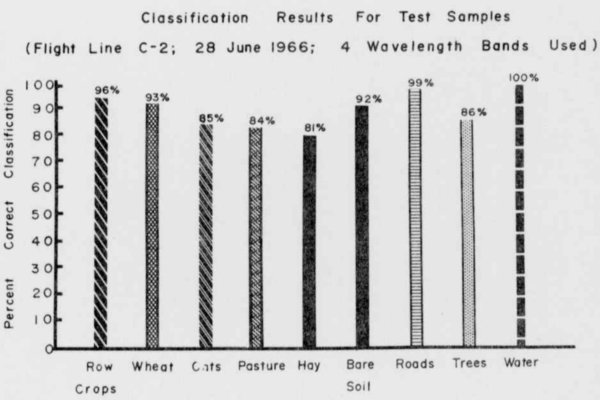


Figure 43. Quantitative Classification Results for Areas Outlined in Figure 42

the fields outlined in the previous figure. The remainder of the table gives the number of samples or RSU's which were actually classified into the various cover types. In other words, out of the 15,185 samples designated as belonging to the row crop category, 14,035 were actually identified as a row crop (either corn or soybeans). Therefore, the correct percentage recognition was 92.4 percent. These percentage figures correspond to those shown on the graph in Figure 43.

The overall performance of this classification is 89.7 percent. If each category is weighed equally, an average performance is 90.2 percent. The overall performance is thought to be more meaningful because it weighs each cover type according to the portion of the total area which is covered by that type. For this example, of the total 27,000 samples classified, more than half of these (15,000) belong to the row crop category. This gives

CLASSIFICATICN SLMMARY BY TEST CLASSES

NC CF SAMPLES CLASSIFIED INTC

	CLASS	NC CF SAMPS	PCT. CCRCT	RCW	WHET	CATS	PAST	HAY	SCIL	RCAC	TREE	WATE	THRS
1	RCW	15185	92.4	14035	3	55	956	119	2	1C	4	1	C
2	WHET	2736	92.8	2	2540	168	C	24	C	1	1	0	0
3	CATS	3465	84.7	175	21	2934	184	128	C	1	21	C	1
4	PAST	3595	83.7	159	1	260	3010	2	C	C	161	0	2
5	HAY	1291	8C.7	17	2	226	C	1C42	4	C	C	0	C
6	SCIL	1C4	92.3	6	0	C	C	1	96	1	C	C	0
7	RCAC	85	98.8	C	0	C	C	C	1	84	C	0	0
8	TREE	72C	86.C	8	0	18	75	C	C	C	619	C	0
9	WATE	131	1CC.C	C	0	0	C	C	C	C	0	131	0
	TCTAL	27312		144C2	2567	3661	4225	1316	103	97	806	132	3

CVERALL PERFCRMANCE = 89.7

AVERAGE PERFCRMANCE BY CLASS = 9C.2

Figure 44. Computer Output Showing the Quantitative Classification Results and the Number of Samples Used to Obtain the Percentage Correct Recognition Figures

row crops over 50 percent of the weight in the overall performance figure of 89.7 percent.

These figures are based upon entire fields and are, therefore, perhaps somewhat lower than is actually correct. For example, in some fields of wheat there might be a patch of weeds covering perhaps 10 percent of the total area in the field. However, the entire field was assumed to be wheat. If the computer classifies all the field as wheat, there is 100 percent accuracy. If the computer classifies that patch of weeds as anything other than wheat, the classification results are only 90 percent accurate. Actually, it was quite correct not to classify the patch of weeds as wheat. Situations such as this would cause these results to be somewhat lower than may actually be the case.

These classification results may be examined in another manner. The table in Figure 44 represents a total of 66 test areas as outlined in Figure 42. These were quantitatively analyzed on an individual field-by-field basis and then summarized. Sixty out of the 66 test areas had a correct percentage recognition within the field of 75 percent or more. (Thirty-six test areas had 95 percent or better correct recognition.) Therefore, if 75 percent or better classification of the RSU's in any field would identify a field as being a certain crop species or cover type, then 60 out of the 66 test areas in the flight line were correctly identified according to their known ground cover. Of the six remaining fields, one field of soybeans would have been incorrectly identified as a pasture area. The other five fields did not have 75 percent of the RSU's classified into any particular cover type and, therefore, could not be identified by this method. These fields would be designated as "indeterminate."

In an operational system, such areas which could not be definitely classified could be designated accordingly. A photo interpreter could then be asked to identify the land use or cover type in such areas. However, by first putting all the data through an automatic classification procedure capable of making acreage determinations and maps of cover types, a great amount of the data load might be taken off the photo interpreter and put onto the other remote sensing systems. Thus, automatic pattern recognition capabilities could be used in this manner as a supplementary tool, not a replacement, for the photo interpreter.

PRELIMINARY SOILS STUDIES

INTRODUCTION

As the demand for soil maps increases, the need for determining new techniques to assist the soil mapper in improving his output becomes greater. The initial results obtained in mapping small soil areas with the use of a computer appear promising. Even though these results are preliminary, it is visualized that they should be a great aid in helping to map and understand soil patterns.

DESCRIPTION OF RESEARCH LOCATION

The areas studied in this research are near U.S. Highway 37 in Morgan County, in South Central Indiana (Figure 45). Additional data collected on this flight line were discussed in the section, "Vegetation, Bare Soil, and Water Mapping of a 70-Mile Flight Line."

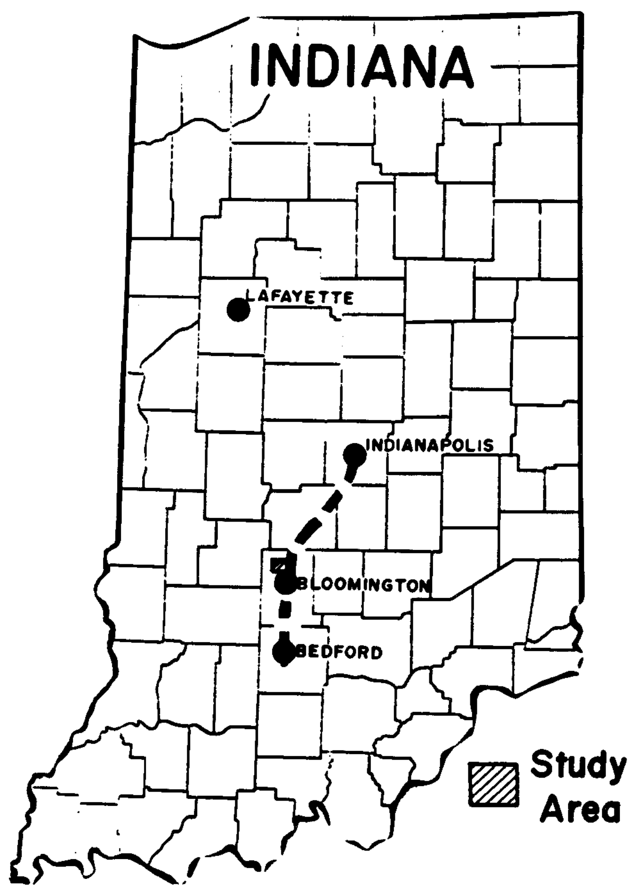


Figure 45. Location of Research Area in South Central Indiana

The study area was covered by the early stage of Wisconsin glacier and is referred to as the Tipton Till plain. The soils are gray-brown podzolic soils developed mostly under a dense forest cover. The surface colors are quite varied and are generally light in color and low in organic matter. The topography is nearly level to very rolling due to the geologic features such as river bottoms, outwash terraces, and glacial drift deposits. Farming in this area is diversified with grain crops and livestock enterprises.

DATA COLLECTION AND AUTOMATIC PROCESSING

Data were collected in wavelengths from approximately 0.3-15 micrometers, and were handled as other LARS data.1/ Bare soil fields showing various color patterns were chosen as training fields. The spectral response pattern or signature of the soils were found to vary in a similar manner (Figure 46). The differences in three soil categories shown in Figure 46 are not as great as the spectral responses of green vegetation, bare soil, and water (Figure 47).

RESULTS AND DISCUSSION

On the aerial photograph in Figure 48 dark and light patterns of the soils can be seen. The computer was trained with samples of dark and light colored areas from the fields shown on the photograph. Green vegetation and water categories were given to the categorizer but no symbol was assigned. Therefore these categories were left blank. The computer was further instructed to print an "I" when it recognized dark soil and a "-" when it recognized light colored soil, and these results are shown in Figure 48. A comparison of the computer printout with the aerial photograph shows that the dark and light soil patterns are rather accurately displayed.

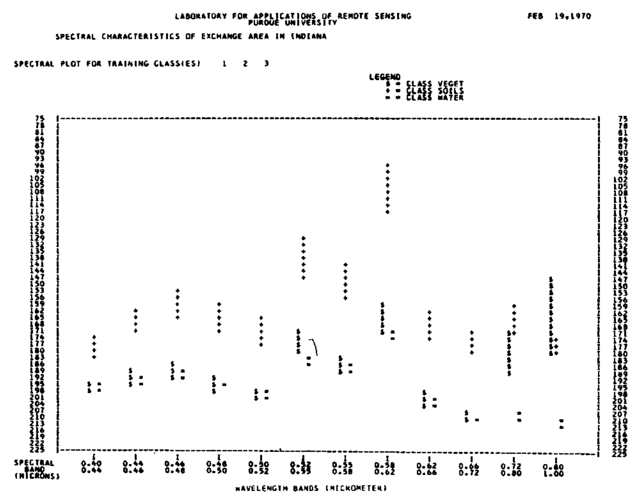


Figure 46. Spectral Signature of Soils

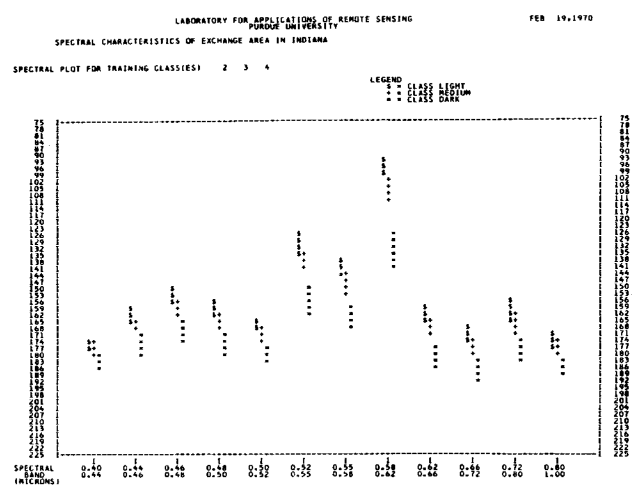


Figure 47. Spectral Signatures of Green Vegetation, Bare Soil, and Water

1/ See Chapter 3, Data Processing

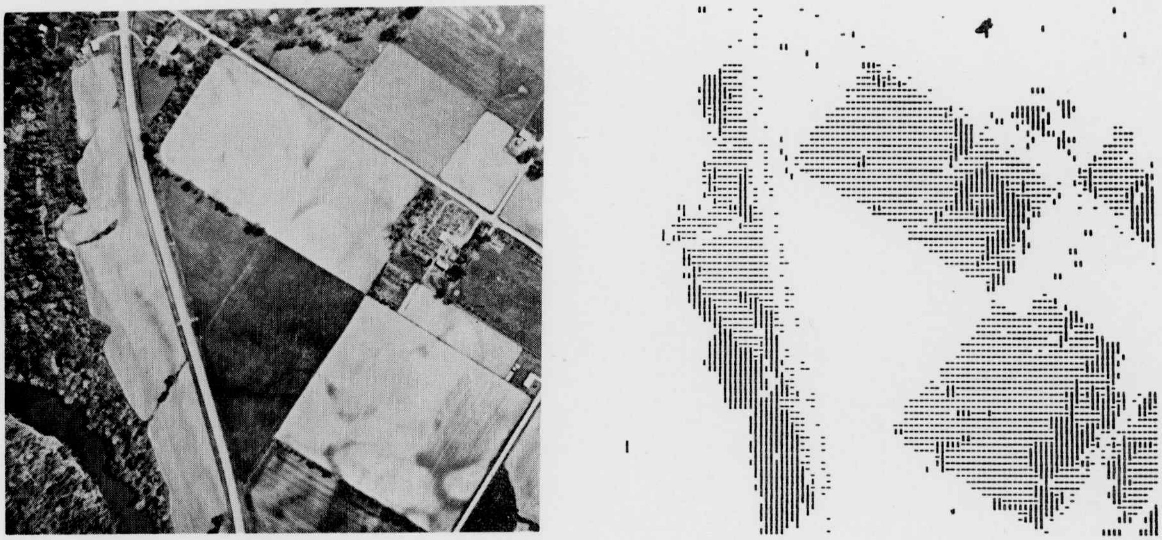


Figure 48. Left is an Aerial Photograph of Study Area. Right is a Computer Printout Showing Dark ("I") and Light ("-") Soil.

In a closer examination of the aerial photograph, the variations of the dark and light colored areas should be noted. Samples from selected vector points were obtained for the darkest areas, the intermediate tone areas, and the light areas. Spectral plots of these categories were shown previously in Figure 46. Using similar techniques as described for the two categories of soil, the computer was instructed to print out three categories. These are shown in comparison with an aerial photograph in Figure 49, and appear quite accurate.

It was noted that there was a large standard deviation of the mean responses for these categories of soil. This indicated there was a possibility of distinguishing more distinct categories based on the spectral properties. It was found that by selecting individual training points and comparing the spectral response with other points, many more soil categories could be defined. Figure 50 shows an attempt at mapping six distinct soil categories.

Examination of these categories in the field showed that the overall tone of the soil may not have been entirely responsible for the categories as they are shown. Tone variations were not evident in every case. Variations in surface structure such as roughness or crusting factors may have played an important role in the spectral responses of soils. The effects of moisture content has been shown to vary spectral responses of soil, also.

An area containing several large fields of bare soil and located a short distance from the

Figure 49. Aerial Photograph and Computer Printout Showing Three Soil Categories

Figure 50. Aerial Photograph and Computer Printout Showing Six Soil Categories

first study area was chosen to see if the classification technique could be repeated. An attempt was made to classify six soil categories of the bare soil area and these are compared to an aerial photograph of the same area in Figure 51. Again, the results are quite striking. There are some variations within one of the bare soil fields which show that perhaps tillage practices may have an effect on soil categories mapped by this technique. These variations and other factors are being further studied.

Soil patterns within vegetation are sometimes quite evident. Attempts are now being made to determine if soil categories can be mapped using the spectral variations of the vegetation for training purposes. Preliminary results show this to be quite promising and indicating that one would not necessarily need to have a bare soil area to make a spectral map of that region.

Soil could be classified into many different categories using this technique. The limits seem to be determined by the amount of detail desired by the soil mapper and the ultimate user of the survey map, the surface color variations caused by temporal and spatial variations in soil moisture, and the spectral variations due to differences in cultural practices. Attempts are now being made to determine how accurately these soil categories compare with soil surveys made by soil mappers in the field. Cultural practices and moisture variations cer--tainly affect the overall computer mapping accuracy, but it is believed that the computer printout with the mapped soil categories would be a great aid to professional personnel in soil mapping.

CONCLUSIONS

The use of multispectral sensing techniques and automatic data recognition techniques developed at the Laboratory for Applications of Remote Sensing at Purdue University appear to have potential in the area of soil mapping.

The possibility of flying over an area and obtaining a map of different soil patterns within hours after the data were collected seems to be of great value to the ultimate users--city planners, soil conservation personnel and others. The presented study illustrates that 6 different categories of soil can be mapped with reasonable accuracy by computer techniques. Observations of actual surface moisture and surface roughness factors would have greatly aided in the interpretation of this data. This indicates how important it is to have accurate information about the study area at the time the scanner data are obtained. Figure 52 illustrates that the same data collected to map soils can be utilized to map green vegetation, bare soil, and water. To do this, the computer was trained with selected areas of green vegetation, bare soil, and water.

CLASSIFICATION STUDIES IN DIFFERENT GEOGRAPHIC AREAS

One of the more difficult problems involved in automatic classification of agricultural materials is the proper selection of an adequate training set of data. The difficulty lies in the spectral variability of agricultural materials. To examine the problems which might be encountered for data obtained from different geographical areas, preliminary analysis work has

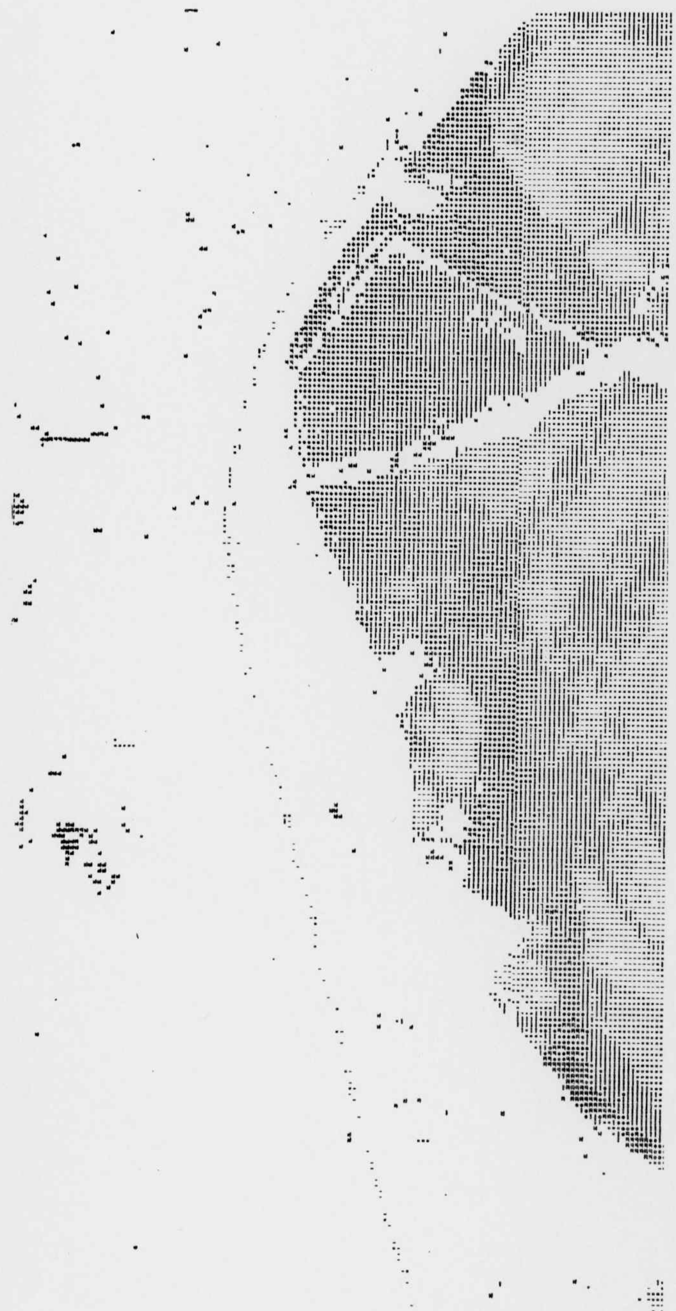


Figure 51. Aerial Photograph and Computer Printout Classification of Six Soil Categories in an Area North of the Study Area. The Training Samples were from the Study Area.

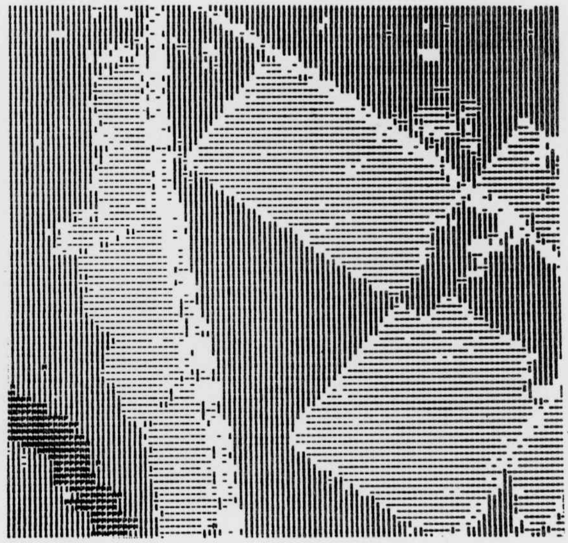


Figure 52. Aerial Photograph and Computer Printout of Green Vegetation (I), Soil (-), and Water (M).

been done with data collected at the Weslaco, Texas, and California Test Sites. The problems of agricultural variability and training sample selection were found to be very similar to those encountered for data obtained in Indiana.

Scanner data from an area designated as Moon Lake, near Weslaco, Texas, was utilized for the first analysis. The data had originally been digitized for the researchers at Weslaco. The only ground truth information for this area was the aerial photograph shown in Figure 53. As can be seen in the photograph, there are bare soil areas; fields with different amounts, types and conditions of vegetation; as well as lake and river water. Since the identify of the various species of vegetation present in this area was not known five groups of training samples were used to represent all cover types for the automatic classification. These were: water, bare soil, trees, and two spectrally different groups of green agricultural vegetation. The data were classified and the results indicated a 97 percent accuracy for the training samples. Figure 54 shows a printout of these classification results.

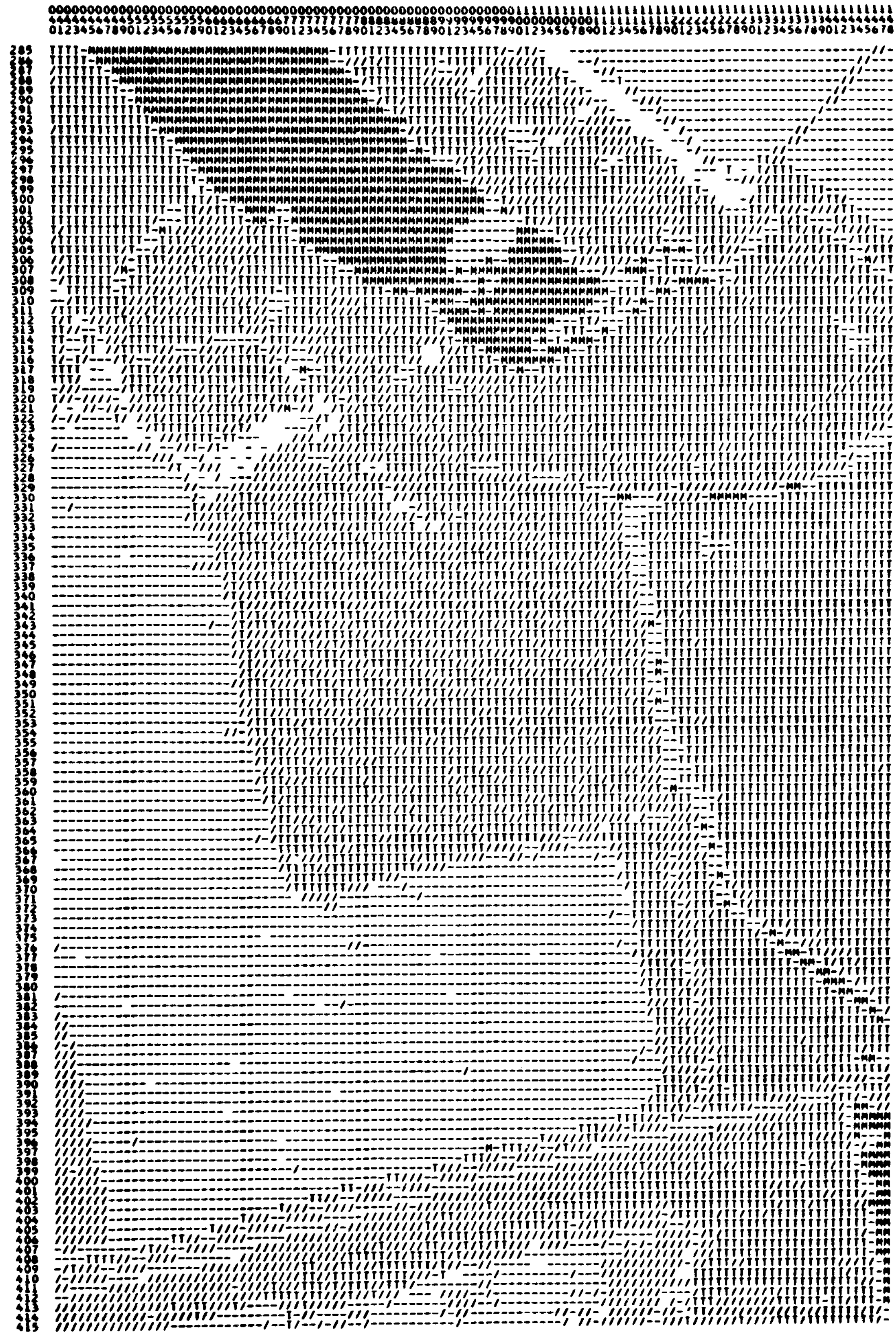
As seen in Figure 53, the classification results for the entire area (of which the training samples are only a small portion) appear qualitatively to be highly accurate. Notice the narrow strips of vegetation in the field of bare soil indicated with arrows labeled "A." Also note the field labeled "B" where vegetation is partially covering the soil, but a great deal of soil is still visible from above. The classification of this area is a mixture of RSU's classified as bare soil and as green

Figure 53. Aerial Photograph of Moon Lake Area, Weslaco, Texas

agricultural vegetation.

The orchard (labeled "C") just below Moon Lake also offers an interesting classification result. A higher spatial resolution of the classification of this area is shown in Figure 55, and illustrates the combination of "T" and "/" symbols representing the "tree" and "green agricultural vegetation," respectively. The individual RSU's do not indicate individual trees, as was expected due to the resolution of the scanner system. It is significant that the orchard trees were identified as such, since the "tree" training samples were obtained from a forested area beside Moon Lake. These trees had similar physiognomic characteristics but were probably not the same species. Also, the number of RSU's identified as "trees" and as

Figure 54. Double Width Computer Printout of Classification Results of Moon Lake Area, Weslaco, Texas



TOTAL NUMBER OF SAMPLED POINTS = 14279

Figure 55. Classification of Orchard Just Below Moon Lake Area, Weslaco, Texas
M = Water, - = Soil, / = Vegetation, T = Trees

"green vegetation" appeared to be the proportion expected after viewing the aerial photograph. It is of interest to note that the pattern of RSU's identified as trees and as "green vegetation" is a distinct moire pattern.1/

Scanner data from California was utilized as a second example of classification of data from a completely different geographical location. This area had quite different vegetative species and conditions. This data were collected over the University of California, Davis test site and had been digitized previously for the University of California at their request. Initially, the species of vegetation present was not known. The multispectral gray scale printouts were examined and based upon past experience in interpreting spectral response of basic cover types, the following categories for classification were selected: bare soil, water, green vegetation, combination of green vegetation and water (believed to be rice, it was not known that rice paddies were present in the flight line). The automatic classification into these four categories appeared to be good, but the evaluation of these results was limited since no positive ground truth information was available. Information concerning the cover type or crop species present was obtained later. The areas believed to be soil were indeed bare soil, the combination of green vegetation and water was immature rice, and the green vegetation was either safflower or mature rice. Using these crop species identifications, the data were again classified, using the following classes: soil (both light and dark), immature rice, mature rice, safflower, and water. The tabular classification results showed a 98 percent accuracy for the training fields. Figure 56 shows imagery of this area in the .46 to .48 micrometer wavelength bands. Figure 57 is the same area in the 0.8 to 1.0 micrometer wavelength bands. These were two of the four bands used in the classification; the other two bands were .55 to .58 and 0.72 to 0.80 micrometers. The crop species identification is shown in Figure 58. Figure 58 shows the computer printouts of classification results for the entire area, on a crop by crop basis.

The accuracy appears to be very high. The bright spots seen in Figures 56 and 57 in the paddies of immature rice were thresholded on the computer printouts and not identified as rice, even though surrounded by rice.

1/ Laboratory for Agricultural Remote Sensing, Remote Multispectral Sensing in Agriculture, Purdue University, Agricultural Experiment Station, Research Bulletin No. 831, 1967.

Figure 56. Imagery of the California Area in the .46 to .48 Micrometer Wavelength Bands

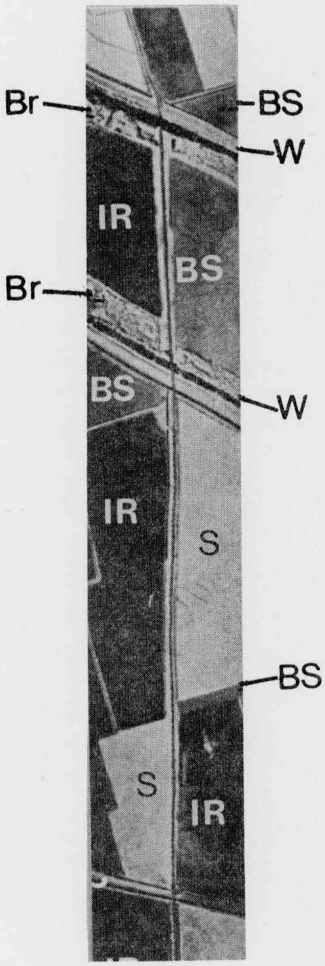


Figure 57. Imagery of the California Area in the 0.8 to 1.0 Micrometer Wavelength Bands. BS=bare soil, W=water, Br=brush, IR=immature rice, S=safflower, MR= mature rice

Figure 58. Computer Printout Classification Results of California Area Showing (a) Rice (M), Immature Rice (-); (b) Safflower (M); (c) Soil (-,=), Water (M)

The potential of this technique to measure acreage of the crop actually present, rather than the total acreage of fields planted to that crop, as determined by the field boundaries is demonstrated.

Through the use of these classification results from California and Texas data it was concluded that this multispectral data processing technique could be satisfactorily used for data from different geographic locations. The problems of agricultural variability and training sample selection were similar to Indiana for the Texas and California data. The results show a reasonably high degree of accuracy for both locations.

LABORATORY STUDIES OF LEAF AND SOIL REFLECTANCE

COLOR AND COLOR INFRARED PHOTOGRAPHY

In the past few years increased use of color and color infrared photography has placed greater emphasis upon the spectral differences among plant species and the spectral changes that occur because of maturity or stress conditions. Correct interpretation of regular color film is not exceedingly difficult since our eyes are also sensitive to this portion of the spectrum. However, because our eyes are not sensitive to the infrared wavelengths, the spectral characteristics of plant and soil materials in these wavelengths are relatively unfamiliar. Therefore, interpretation of photography in the infrared wavelengths is more difficult. Since color infrared film is sensitive to both the visible and infrared wavelengths, it can be erroneously interpreted.

SPECTRAL BEHAVIOR OF PLANTS AND SOILS

As the potentials for the use of color infrared film and multispectral scanners become better known, the factors affecting reflectance and emittance from various plant and soil materials will need to be more completely understood. This will be necessary so that (1) we can interpret the data more completely and precisely, and (2) we can reliably predict those conditions under which we would expect to observe changes in the spectral characteristics of these materials.

The manner and portion of the spectrum in which the change will occur must be known to evaluate the feasibility for using remote sensing systems to located and map areas where special situations exist. To gain this type of understanding of the spectral behavior of plant and soil materials requires an integrated research program of laboratory, field, and aircraft experiments. Without such an integrated program, we will find ourselves continually dealing with new problems by emperical methods.

In an attempt to gain additional insight to the reflectance properties of plants and soils, DK-2 spectral data were obtained for over 2300 samples of various plant and soil materials. A detailed account of the techniques used, the plant species, soil type, and the condition of the samples used in obtaining these spectra is given in a previous LARS Annual Report.[1/]

1/ Laboratory for Agricultural Remote Sensing. 1968. Remote Multispectral Sensing in Agriculture, Volume No. 3 (Annual Report). Agricultural Experiment Station Research Bulletin No. 844. pp. 18-25.

Some of the analyses conducted with these spectral data were reported previously.1/ Additional analyses were conducted during this contract period. In these analyses, particular attention was given to the (1) effects of pigmentation on leaf reflectance, (2) effects of moisture content on leaf reflectance, and (3) effects of moisture content on soil reflectance.

Figure 59 is a brief review of the characteristic spectral reflectance of a green leaf and indicates that the 0.4 to 2.6 micrometer portion of the spectrum can be roughly divided into three areas. First is the visible wavelength region, in which plant pigments (especially the chlorophylls) dominate the spectral response of plants. Second is the region from approximately 0.72 to 1.3 micrometers where there is very little absorption by a leaf, and therefore most of the energy reaching the leaf must be either transmitted or reflected. The third region is from 1.3 to 2.6 micrometers where water is absorbed.

Note in Figure 59 that a normal green, healthy leaf will have four primary absorption bands. Two of these are in the visible wavelenghts and are caused by chlorophyll absorption, one at approximately 0.45 micrometers in the blue region and one at approximately 0.65 micrometers in the red region. Water absorption accounts for the strong decrease in reflectance at wavelength bands located at approximately 1.45 and 1.95 micrometers in the infrared.

EFFECT OF PIGMENTATION ON LEAF REFLECTANCE

As previously stated, leaf pigmentation can cause marked differences in spectral response in the visible wavelengths, as illustrated in Figure 60.

The white Coleus leaf without any apparent pigmentation has a very high level of reflectance throughout the 0.5 to 0.9 micrometer region. The leaf dominated by chlorophyll pigmentation shows the characteristic curve for a green leaf, with relatively low reflectance at 0.5 micrometer, a peak in the green at approximately 0.55 micrometer, low reflectance again in the red at about 0.65 micrometers, and then the usual sharp increase at about 0.7 micrometers to the reflective infrared wavelengths. A red leaf has a low reflectance throughout the blue and green portions of the spectrum, then a marked increase and very high level of reflectance throughout the red wavelengths. A deep reddish-purple leaf has a relatively low level of reflectance throughout the visible region, and then a sharp rise which coincides with that of the green leaf.

Figure 61 extends the wavelength band of these four leaf pigment conditions out to 2.6 micrometers. There is very little difference in reflectance throughout the reflective infrared wavelengths, in spite of the marked differences in the visible wavelengths caused by the pigmentation.

Figure 62 shows distinct differences in reflectance of two silver maple

1/ The results of this DK-2 study were presented in a paper by R. M. Hoffer and C. J. Johannsen, entitled "Ecological Potentials in Spectral Signature Analysis" since published by the University of Georgia Press, Athens, Georgia in a book edited by P. L. Johnson, entitled Remote Sensing in Ecology. (1969)

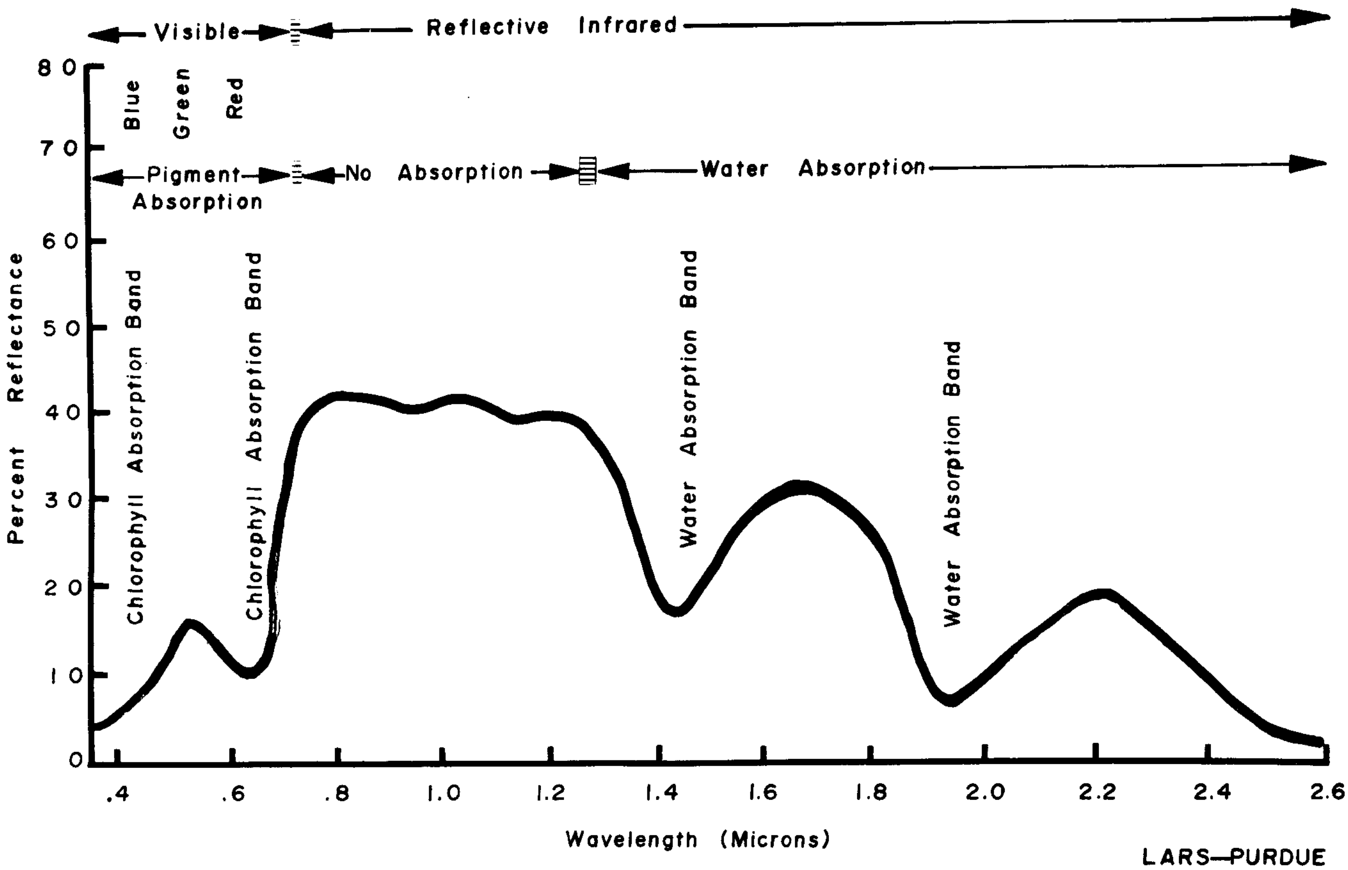


Figure 59. Characteristic Spectral Reflectance Curve of a Green Leaf

(Acer saccharinum L) leaves. These spectra were obtained in the fall and are typical of several obtained for leaves at this stage of maturity. One of these leaves was a normal green, the other a brilliant red. The red coloration was caused by the presence of anthocyanins, which are often produced in maple trees in the fall after chlorophyll production has ceased. As was the situation with the red Coleus leaf, the red maple leaf has a relatively low reflectance at wavelengths below about 0.60 micrometers, then has a sharp rise in reflectance. In this case, however, there is a decrease in reflectance at approximately 0.66 micrometers in the red chlorophyll absorption band, thus indicating the presence of some small amounts of chlorophyll. However, the generally high reflectance throughout the 0.62 to 0.72 micrometer portion of the spectrum dominates, thus giving the leaf its brilliant red coloration.

Spectra for leaves of tuliptree (Liriodendron tulipifera L) are shown

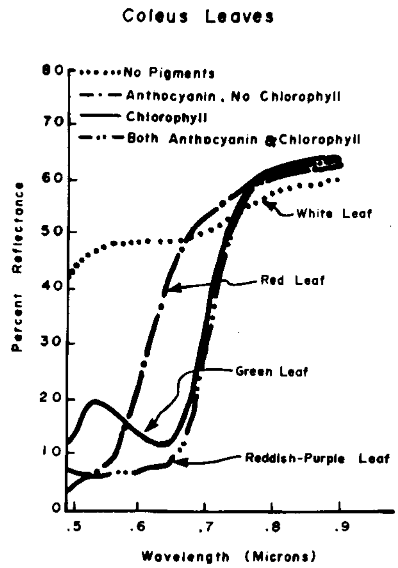


Figure 60. Reflectance Curves for Coleus Leaves with Different Pigmentation

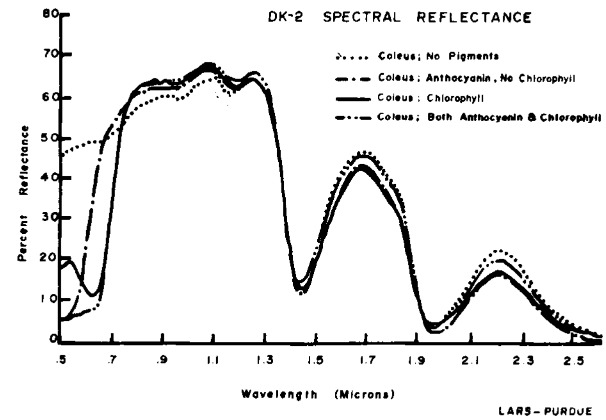


Figure 61. Reflectance Curves for Coleus Leaves Extended To 2.6 Micrometers

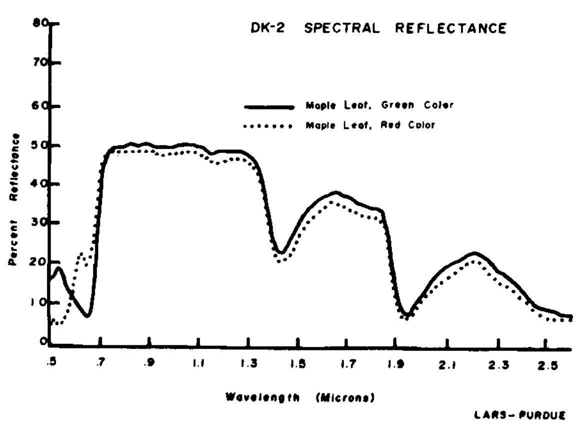


Figure 62. Reflectance Curves for Two Silver Maple Leaves

in Figure 63. Both leaves were quite succulent and at normal high moisture contents. However, one leaf was a normal, deep green color while the other was a bright yellow. The yellow coloration was due to the normal autumn breakdown of the chlorophylls, which were not reformed, thus allowing the presence of the carotenes and xanthophylls to become evident. These carotenoid pigments were present before the chlorophyll breakdown, but were masked by the chlorophylls. The green tuliptree leaf has the usual green spectral curve, but the yellow leaf has a very sharp increase in reflectance starting at 0.50 micrometer and continued high reflectance throughout the green and red portions of the visible spectrum. In the infrared wavelengths, the yellow leaf has 2 to 3 percent lower reflectance than the green leaf. This could cause a somewhat darker tone on infrared film.

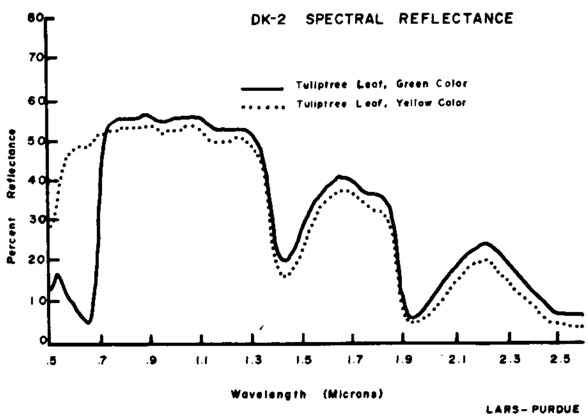


Figure 63. Reflectance Curves for Two Tuliptree Leaves

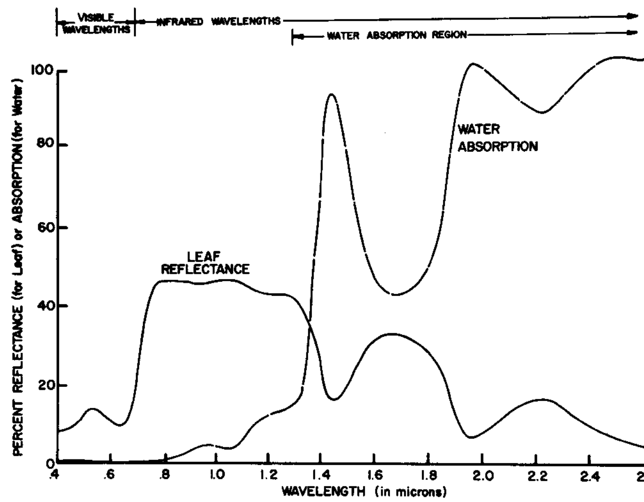


Figure 64. Relationship Between Reflectance Curve of Green Leaf and Water Absorption

EFFECT OF MOISTURE CONTENT ON LEAF REFLECTANCE

Reflectance of leaves in the 1.3 to 2.6 micrometer portion of the spectrum should be examined with particular attention to the water content of the leaves. Forsythe and Christison 1/ showed the absorption for one millimeter thickness of water in these wavelengths. Figure 64 indicates the close relationship between water absorption and reflectance for a healthy, turgid green leaf. In wavelengths where water absorption is high, leaf reflectance is low. This is most apparent in the primary water absorption bands centered at 1.45 and 1.95 micrometers. There are also slight increases in water absorption at approximately 0.96 and 1.2 micrometers. These minor water absorption bands cause slight decreases in reflectance of the leaf. However, because water absorption is generally very low in the 0.7 and 1.3 micrometer band, its influence upon leaf reflectance is minor.

The influence of the drying of leaf tissue and the marked spectral changes that take place in the water absorption region are shown in Figure 65. These curves are plotted from averages for groups of corn leaves at four different moisture content levels. Marked increases in reflectance with decreasing moisture content are observed throughout the 0.5 to 2.6 micrometer region. The curve for corn leaf samples in the 0 to 40 percent 2/ moisture

1/ Forsythe, W.E. and F.L. Christison. 1930. The Absorption of Radiation from Different Sources by Water and by Body Tissue. Journal of Optical Society of America 20:693-700.

2/ There were no leaves below 4 percent moisture content within this group.

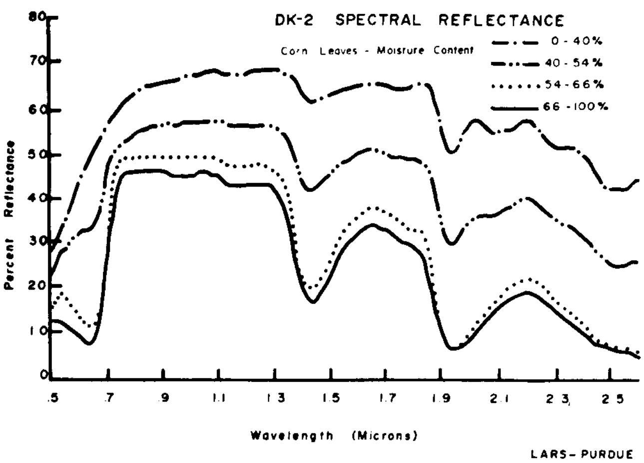


Figure 65. Reflectance Curves for Corn Leaves with Different Moisture Content

content range approaches 70 percent reflectance in the unabsorbed wavelengths, decreasing to about 42 percent at 2.50 micrometers. Only slight decreases are seen in the primary water absorption bands. These leaves lack chlorophyll and therefore have no absorption in the red visible wavelengths. The curve for the 40 to 54 percent moisture content range shows a 10 to 18 percent decrease in reflectance throughout the reflective infrared wavelengths. The primary water absorption bands are quite evident, and the general shape of the curve in the water absorption region from 1.3 to 2.6 micrometers resembles that of normal succulent leaves more than does the curve for 0 to 40 percent moisture content. There is a slight decrease in reflectance at 0.64 to 0.66 micrometers, indicating that there was a small amount of chlorophyll present.

The two remaining curves in Figure 65 (54-66 percent and 66-100 percent1/ moisture content) are generally similar. Reflectance differences in the primary water absorption bands are from 0 to 3 percent, although in the regions between these water absorption bands there is a 3 to 4 percent difference in reflectance. In the unabsorbed region, there is a fairly consistent difference of 3 to 4 percent. Examination of leaf cross-sections indicated that this difference is not due to water absorption as much as to the structural changes taking place in the leaf that are accompanied by the loss in moisture content. The 54 to 66 percent moisture content curve has a higher response in the green visible wavelenghts due to a lighter color.

These curves indicate that reflectance measurements are strongly influenced by the moisture content of the leaves, particularly but not exclusively, in the 1.3 to 2.6 micrometer region. Changes in the leaf structure and pigmentation accompanying the changes in moisture content also have strong effects on reflectance. The observed differences in reflectance as related to moisture content of leaves indicate a possible potential for remotely detecting moisture stress in plants using bands in the 1.3 to 2.6 micrometer portion of the spectrum.

Reflectance of leaves from different species but in the same moisture content range may be significantly different. Curves of averaged reflectance values of leaf samples with 66 to 80 percent moisture content are shown in Figure 66. These average

1/ There were no leaves above 76 percent moisture content within this group.

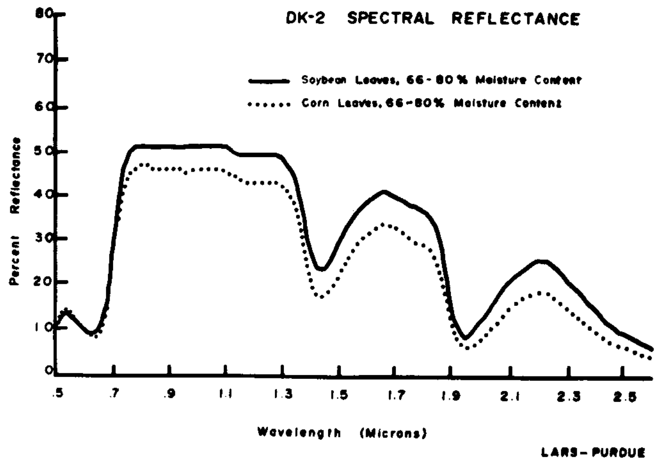


Figure 66. Reflectance Curves for Corn and Soybean Leaves Having the Same Moisture Content

values are from 308 and 382 samples of soybeans and corn leaves, respectively. The curves are very similar throughout the visible wavelengths. However, in the entire reflective infrared region from 0.72 to 2.6 micrometers, the soybean leaves have a higher reflectance than the corn leaves. This is thought to be caused primarily by the structural differences between monocotyledonous and dicotyledonous leaves. The dorsi-ventral structure of the dicotyledonous leaves usually results in higher reflectance. This was found true for the leaves of several plant species examined in detail.[1/] Reflectance and transmission spectra for monocotyledonous and dicotyledonous leaf structures indicate that very little energy is absorbed in 0.72 to 1.3 micrometer wavelengths band.

Spectra for the leaves of three tree species are shown in Figure 67.

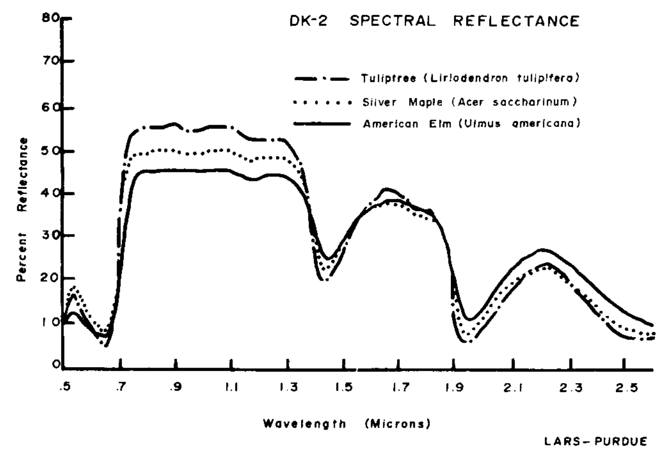


Figure 67. Reflectance Curves of Leaves of Three Tree Species.

Although all three spectra have approximately the same general shape, characteristic of green vegetation, there are significant differences in amplitude of reflectance in certain wavelength bands. These relative differences in reflectance within the various wavelength bands are significant in pattern recognition by helping to identify unknown samples through correlation. Most evident in these spectra are the marked differences in reflectances throughout the unabsorbed wavelengths. The tuliptree had about a 5 percent higher reflectance than the silver maple, which in turn had about 5 percent higher reflectance than the American elm. In the visible wavelengths, the peak in the green shows highest reflectance from the maple

1/ Sinclair, T. R. 1968. Pathway of Solar Radiation Through Leaves. Unpublished M.S. Thesis. Purdue University, Lafayette, Indiana.

leaf and lowest reflectance from the elm. A crossover takes place in the red chlorophyll absorption band, however, and the tuliptree reflects the least in this wavelength band. The maple still has highest reflectance. In the 1.9 to 2.6 micrometer region, the elm has a somewhat higher reflectance than either the maple or tuliptree, which are quite similar in reflectance.

EFFECT OF MOISTURE CONTENT ON SOIL REFLECTANCE

Spectral measurements were obtained on 250 soil samples. Ten different soil textures, four drainage profiles and three major soil horizons were represented in these samples. Spectral samples of sandy and clay soils were selected to show the effect of moisture on soil reflectance. The mean spectral curves for the clay soils at two different moisture levels and sandy soils at three different moisture levels are shown in Figures 68 and 69, respectively. The curves for both soil categories show a very large decrease in reflectance with an increase in moisture. Curves for the clay soils maintain the same spectral shape, but the sandy soils exhibit a marked change in the shape of the spectral curve with a change in moisture. The water absorption bands at approximately 1.45 and 1.95 micrometers become pronounced for sandy soils with a moisture content of over 4 percent. Pure samples of Bentonite, Muscovite, and Kaolinite clays at only 0.1 percent moisture still had strong water absorption bands. This evidence and the lack of change in the characteristic curves at low moisture contents seem to indicate that bound water may be exerting an influence on the reflectance. The bound water in sandy soils is very low compared to clay soils. Thus, the sandy soils at low moisture contents produce a rather flat, uniform reflectance curve throughout the reflective infrared wavelengths. This curve is without the decrease in reflectance in the strong water absorption bands seen for the clay soils at low moisture contents.

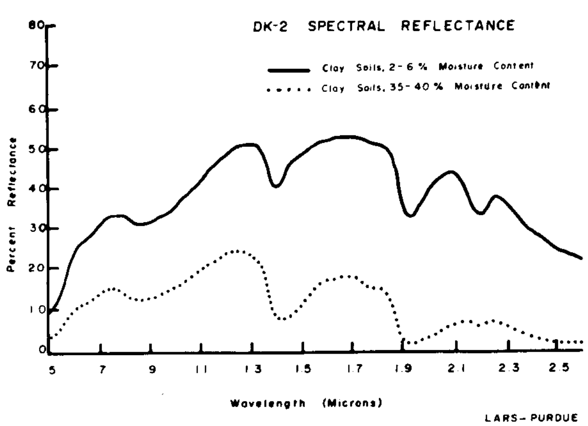


Figure 68. Reflectance Curves for Clay Soil at Two Moisture Levels

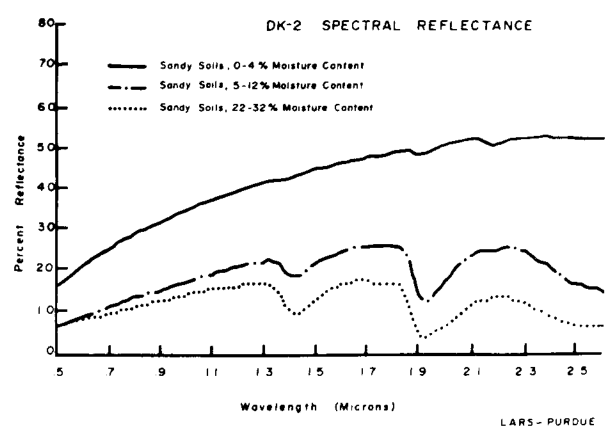


Figure 69. Reflectance Curves for Sandy Soil at Three Moisture Levels

The ability to delineate soil types and to survey the soil moisture content would be of great interest to ecologists, soil scientists, and others. The data just presented represent reflectancy from the soil surface. As noted, the reflectance is greatly reduced with an increase in moisture content of the soil surface. It is well known that soil tends to dry on the surface, forming a thin dry crust. This crust can develop within a few hours after a rain.[1/] Therefore, reflective measurements using remote sensing devices could show that the soil appears to be dry, whereas the soil profile might be very wet. This indicates a possible limitation in the utility of reflectance data alone. However, the use of the emissive or microwave wavelengths, perhaps in combination with the reflective wavelengths, may allow soil moisture information to be obtained by remote sensing techniques.

Comparisons of corn leaf reflectance at low moisture (0 to 20 percent) with reflectance of clay soils at low moisture (2 to 6 percent) and reflectance of corn leaves at high moisture (80 to 90 percent) with clay soils at high moisture (22 to 40 percent) are shown in Figures 70 and 71, respectively. The overall shapes of the low moisture curves are somewhat similar. The corn samples generally have a 10 to 15 percent higher reflectance in the water absorption regions. A characteristic decrease in reflectance for these clay soils at about 0.80 to 0.90 micrometers result in a difference

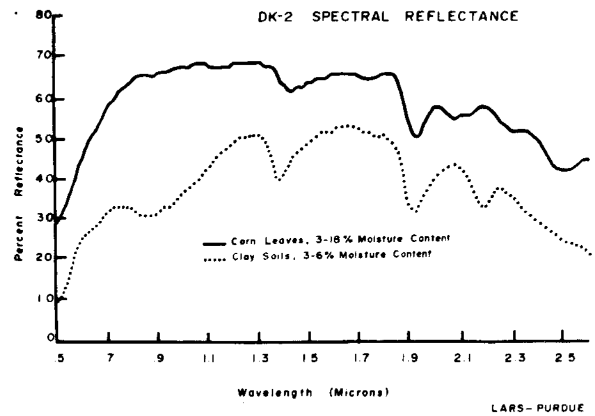


Figure 70. Reflectance Curves for Corn Leaves and Clay Soils with Low Moisture Content

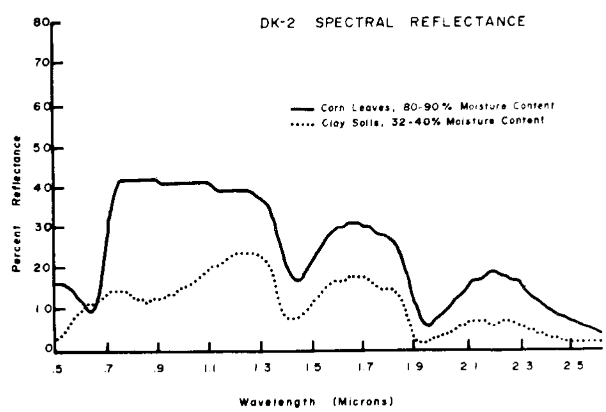


Figure 71. Reflectance Curves for Corn Leaves and Clay Soils with High Moisture Content

in reflectance of over 30 percent. Also, the characteristic decrease in reflectance for these clay soils at 2.20

1/ Mannering, J. V. 1967. The Relationship of Some Physical and Chemical Properties of Soils to Surface Sealing. Unpublished Ph.D. Thesis. Purdue University, Lafayette, Indiana.

micrometers is not observed in the curve for corn at low moisture.

In the comparison of high moisture clay and leaf samples, the clay soils were completely saturated at 32 to 40 percent moisture. The corn samples were also at a maximum moisture level at 80 to 90 percent moisture content. Because of differences in corn and clay materials and in the total amount of water present, the clay has a lower reflectance in the water absorption region even though the corn has a higher percentage moisture content than the clay. Again, the similarity in shape of the curves in the water absorption region (1.3 to 2.6 micrometers) is striking. Of course, the clay does not have the characteristic "green vegetation" reflectance curve in the visible wavelength bands, or the sharp rise at about 0.7 micrometers. In the 0.72 to 1.0 micrometer region, the clay soils have about 30 percent lower reflectance than the corn leaves. This difference would explain the high response of green vegetation on infrared film compared to the low response generally exhibited by soils.

With the use of these leaf and soil spectra, an attempt has been made to show the utility of laboratory spectral reflectance data for interpretation purposes. The portion of the spectrum where greatest differences in reflectance among various plant and soil materials occur can be determined; effects of differences in moisture content, pigmentation, and internal leaf structure can be studied. Laboratory studies are a very necessary part in developing an understanding of energy interactions with plant and soil materials. However, it must be stressed that a capability for differentiating plant or soil materials on the basis of laboratory spectra does not mean that the same capability will be obtained in the field. Laboratory spectra are obtained on very small plant or soil areas normally oriented at an angle perpendicular or nearly perpendicular to the incident beam of energy. Such spectral data are not necessarily comparable to the spectral response measured remotely. For example, aircraft spectral data of a particular plant species may represent a mixture of leaves oriented at many different angles; some leaves are green and succulent, some brown and dry; shadow areas are present; and some soil may be showing (with accompanying differences in response due to soil type, surface texture and moisture conditions).1/ The number of layers of leaf material through which radiation will pass also varies and this may cause marked differences in spectral response.2/

Therefore, these analyses of individual leaf and soil spectral data obtained with laboratory instrumentation must be regarded with caution. Nevertheless, proper analysis of

---

1/ Laboratory for Agricultural Remote Sensing. 1968. "Remote Multispectral Sensing in Agriculture." Volume No. 3 (Annual Report). Agriculture Experiment Station Research Bulletin 844. pp. 45-58.

2/ Meyer, V. I., C. L. Wiegand, M. D. Heilman, and J. R. Thomas. Remote Sensing in Soil and Water Conservation Research. Proc. Fourth Symposium on Remote Sensing of Environment, I.S.T., University of Michigan, Ann Arbor, Michigan. pp. 801-813.

spectral data are a valuable and necessary phase of developing a more complete understanding of the interrelationships of spectral response in plant and soil materials. Also, such studies will significantly aid in developing the necessary capabilities for proper interpretation of remotely sensed multispectral scanner data.

CHAPTER 5

MEASUREMENTS

LEAF SCATTERING STUDY

The scattering of light from corn and soybean leaves in vivo was measured with a leaf scattering apparatus. The measurements were made in a plane normal to the leaf plane at nineteen wavelengths from 375 nanometers to 1000 nanometers, at angles of incidence of 0°, 15°, 30°, 45°, and 60°.1/

The scattering measurement chamber is shown in Figure 72 with the white arrow indicating the incident light beam. Optical interrupt tabs are

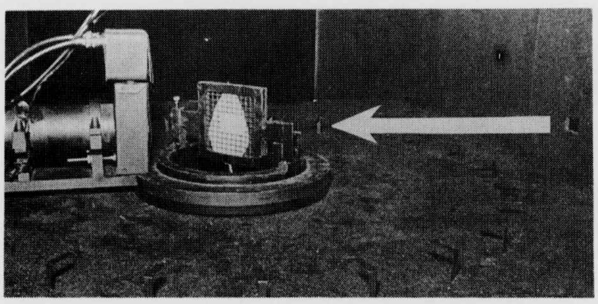


Figure 72. The Scattering Measurement Chamber Showing a Sample Leaf Mounting and the Incident Beam Axis

mounted on the floor of the chamber to cause a sampling trigger every 15° of rotation of the photomultiplier mount. The leaf mounting is shown in more detail in Figure 73, and the soybean plant beneath the scattering chamber is shown in Figure 74. The monochromator and chopper detail are shown in Figure 75. The electronic portion of the system provided amplification and synchronous demodulation of the sensor output in the customary manner. In addition this portion of the system contained a sine and cosine generating circuit keyed to the optical interrupt tabs to trigger signals so that direct polar plotting of the sensor output was possible on a storage oscilloscope. A diagram of the electronic system is shown in Figure 76 and a typical storage oscilloscope presentation of the polar data is shown in Figure 77. These data were photographed on 35 millimeter film strips, one frame per wavelength.

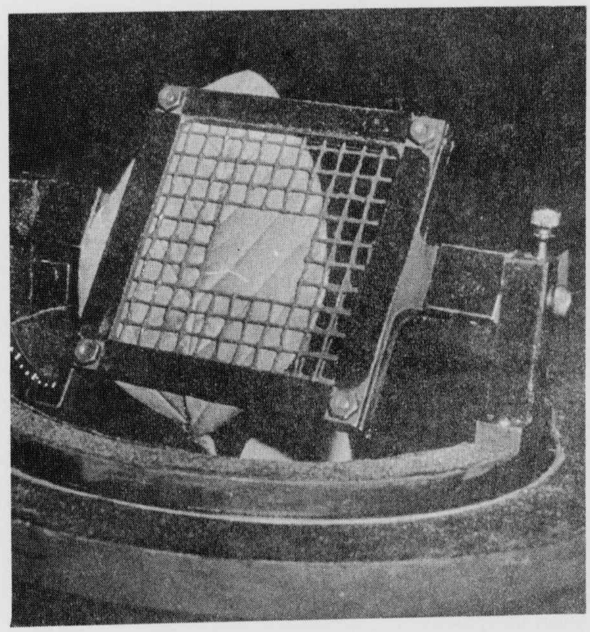


Figure 73. The Leaf Mount in the Scattering Chamber

Prior to the main data collection, two control experiments were run.

1/ Breece, Harry T. III, 1969, Bidirectional Scattering Characteristics of Healthy, Green Soybean and Corn Leaves in Vivo. Ph.D. Thesis, Purdue University.

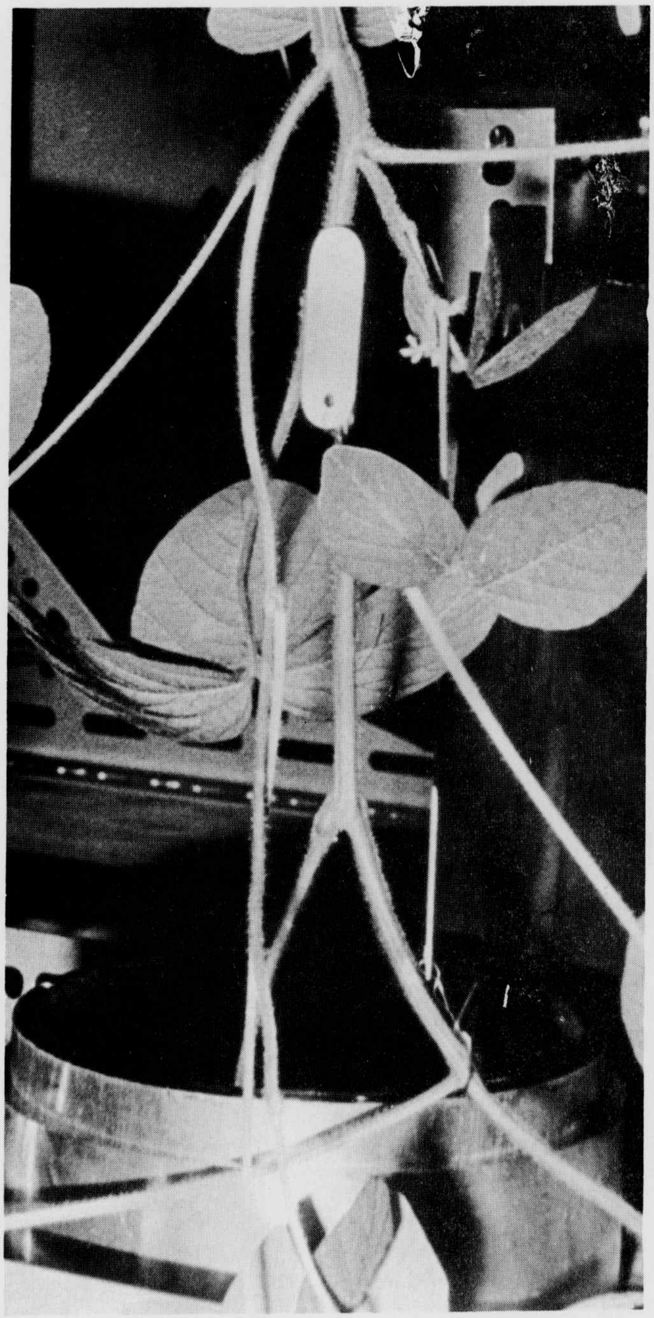


Figure 74. A Soybean Plant Beneath the Scattering Chamber with Measured Leaf in the Leaf Mount

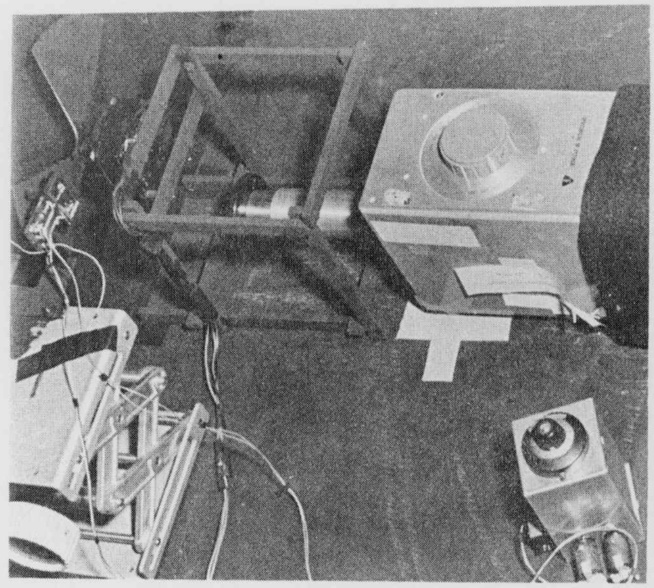


Figure 75. The Monochromator and Bean Chopper

First, the symmetry properties of the leaves were investigated. Soybean leaves were found to have C_∞ rotational symmetry about the leaf plane normal, while corn leaves were found to have C_2 rotational symmetry about the leaf plane normal, consistent with the fibrous structure of this grass leaf. Second, a statistical consistency experiment was carried out to see if several leaves from several plants could be used to make a complete set of runs. This was essential because it was nearly impossible to avoid some wear and tear on a single leaf from a single plant for the length of the data runs. Thus, it was necessary to establish that several leaves from the same general location on several plants would exhibit acceptably small variations in scattering characteristics. Discrepancies ranged as high as 20 percent while most data agreed within about

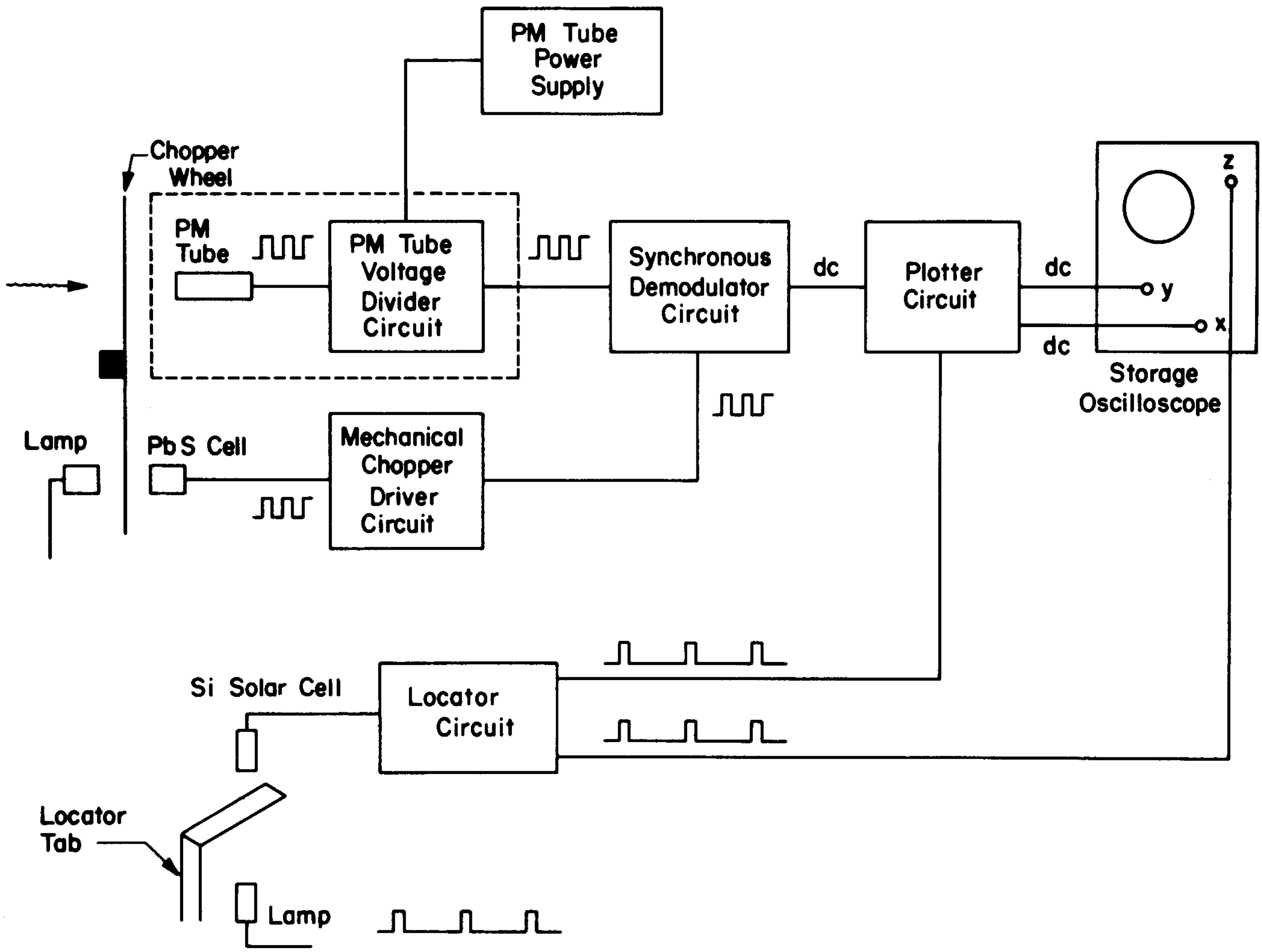


Figure 76. Electronic System Block Diagram

10 percent. Standard deviations ran between 0.1 and 2.5 except for one case at 3.1 and one case at 3.9.

A typical result for soybean leaves from the main data is shown in Figure 78. These curves have been corrected from those shown in the previously cited research[1/] in the wavelengths above 750 nanometers. An incorrect transmission curve on a "neutral" density filter used in calibration was employed in data reduction. The length of a polar vector to any point on a scattering curve, as measured with the

1/ ibid.

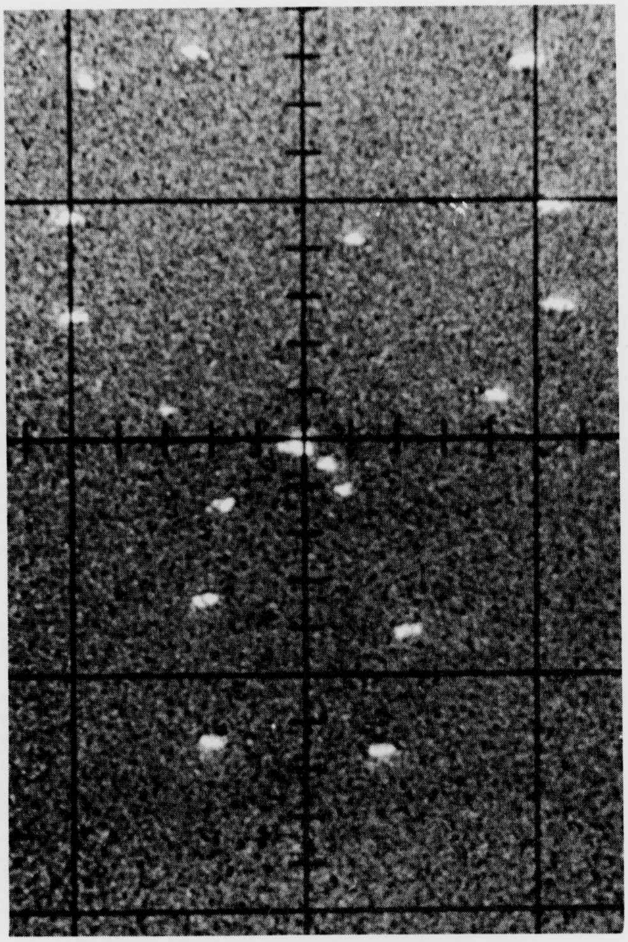


Figure 77. A Typical Polar Intensity Plot on the Storage Oscilloscope Screen for One Chosen Wavelength and One Chosen Angle of Incidence

scale shown and multiplied by the appropriate REFL or TRANS number, yields either $\rho'\cos(\theta_{coll})$ for reflection or $\tau'\cos(\pi - \theta_{coll})$ for transmission. Bidirectional reflectance and transmission distribution functions are ρ' and τ' respectively. Similar data for soybean leaves at 60° angle of incidence are shown in Figure 79. Figure 80 shows data for corn leaves at 30° angle of incidence with vertical midvein orientation. Finally, the bidirectional transmission distribution function for soybean leaves at $\theta_{coll} = 180°$ and normal incidence is shown in Figure 81, directly

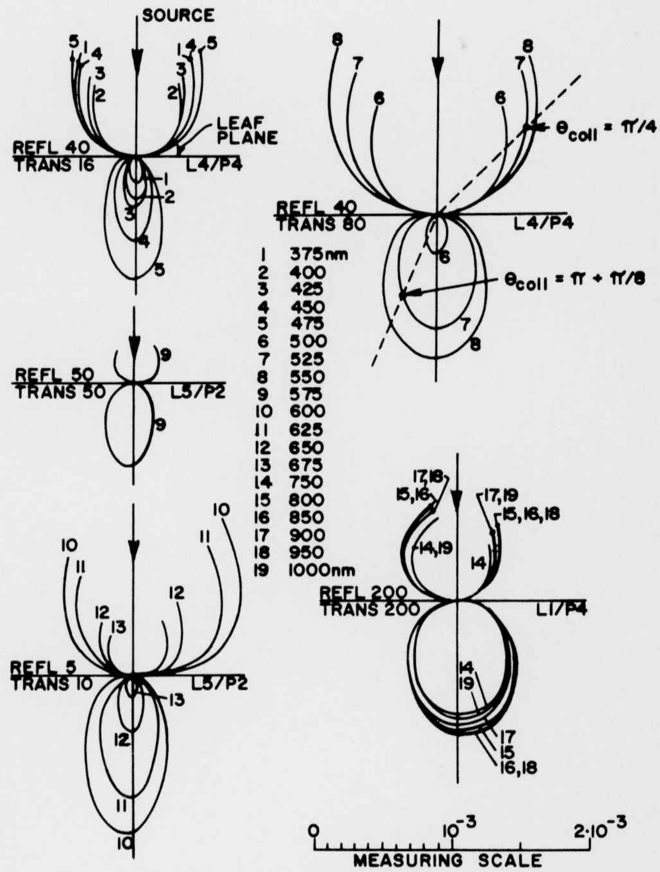


Figure 78. Polar Plots of Soybean Leaf $\rho'\cos\theta_{coll}$ and $\tau'\cos(\pi - \theta_{coll})$ for top incidence at $\theta_{inc} = 0°$. θ_{coll} is the detector collection angle, measured clockwise from the leaf normal.

from the data in Figure 78.

There are two main theoretical developments that have a bearing in modeling the scattering of light in

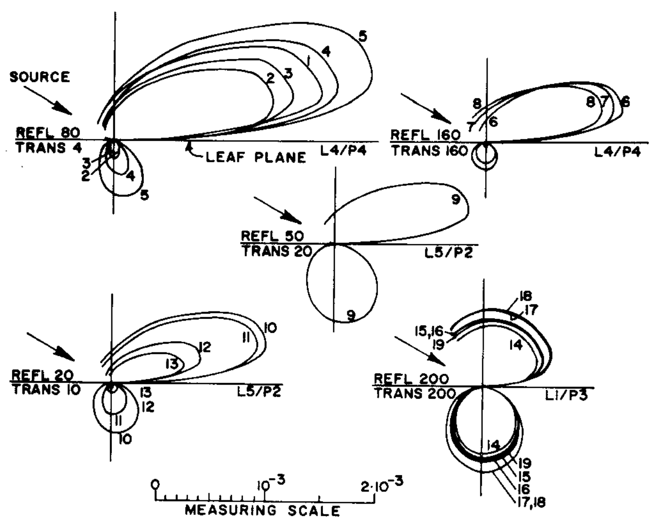


Figure 79. Polar Plots of Soybean Leaf $\rho' \cos \Theta_{coll}$ and $\tau' \cos (\pi - \Theta_{coll})$ for top Incidence at $\Theta_{inc} = 60^\circ$. See Figure 78 for Wavelength Code.

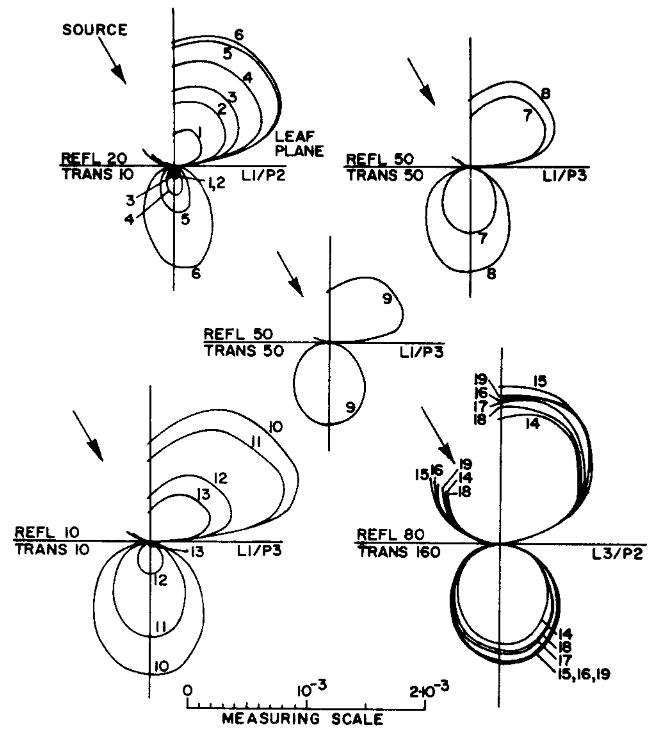


Figure 80. Polar Plots of Corn Leaf $\rho' \cos \Theta_{coll}$ and $\tau' \cos (\pi - \Theta_{coll})$ for Top Incidence and Vertical Midvein Orientation at $\Theta_{inc} = 30^\circ$.

leaves. The first is the general radiative transfer problem described in detail by Chandresekhar.1/ The second is the problem of scattering from rough surfaces described by Beckmann and Spizzichino.2/ Models that are tractable in radiative transfer theory are much too simple to be applied to the leaf. Structures of characteristic dimensions ranging from $\lambda/100$ to 100λ are neither sufficiently disorderly to be treated in a statistical way, nor sufficiently orderly to be treated with the simplifications of periodic structures. This pertains to prediction of the spatial distribution of reflected and transmitted power from the leaf.

An approximate approach to the general radiative transfer problem,

1/ S. Chandresekhar, Radiative Transfer, Dover Publications, Inc., New York (1960).

2/ P. Beckmann and A. Spizzichino. The Scattering of Electromagnetic Waves from Rough Surfaces, The MacMillian Co., New York (1963).

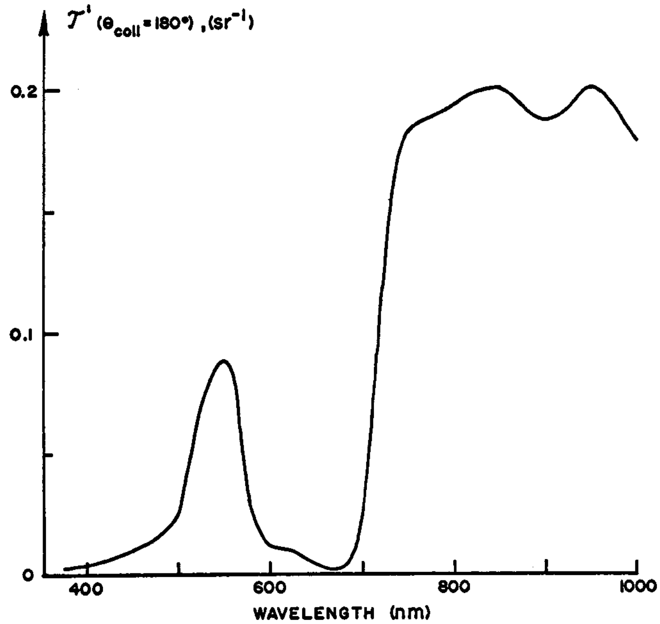


Figure 81. Bidirectional Transmission Distribution function τ' at $\Theta_{coll} = 180^{\circ}$

the Kubelka-Munk theory,[1/] can yield scattering and absorption parameters from total reflectance and transmittance data, which can then be qualitatively related to leaf structure.[2/] It is tempting to try to model the spatial leaf scattering on the basis of the Kubelka-Munk theory, but this requires an assumption on the angular distribution of radiant flux within the leaf. If the customary ideal diffuser model in which the radiant flux is uniformly distributed about any point in the material is assumed, no matter how close to the surface, then the scattering patterns for both reflection and transmission should be Lambertian. This is obviously not the case seen in the data. The reflection curves are clearly non-Lambertian for off-normal incidence and show the kind of specular reflection expected from spatially irregular dielectric-dielectric interfaces in accord with the familiar Willstatter-Stoll theory of light scattering in leaves. While most of the transmission curves are near Lambertian even for angles of incidence of 60°, a careful look shows a sharpening toward the beam direction, indicating incomplete scattering of the incident beam.

The absorption spectrum typical of total diffuse reflectance measurements made with an integrating sphere spectrophotometer is still present in the highly non-Lambertian "specular" reflected power in Figure 79. This strongly indicates that the reflected power has penetrated the waxy leaf cuticle into the absorbing interior cells with only slight air-cuticle reflection. The data presented for this report represent a small sample of the entire experiment. It was generally true that in the non-absorbing near infrared wavelength range from 750 to 1000 nanometers the scattering was more "specular" in character for corn leaves than for soybean leaves, particularly at 60° angle of incidence. Soybean and corn leaf cross-sections were shown in the previous annual report[3/] and

1/ W. W. Wendlandt and H. G. Hecht, Reflectance Spectroscopy, Interscience Publishers, New York (1966).

2/ V. I. Myers and W. A. Allen, "Electrooptical Remote Sensing Methods as Nondestructive Testing and Measuring Techniques in Agriculture," Appl. Opt., Vol. 7, p. 1819, 1968.

3/ Laboratory for Agricultural Remote Sensing, Remote Multispectral Sensing in Agriculture, Purdue University, Agricultural Experiment Station, Research Bulletin No. 844, 1968.

some discussion of structure effects on spectral properties was given. It would appear from the present data that cell form and packing do have a bearing on the light scattering; there are more cell wall surface elements in the corn leaf within some reasonable bounds on parallelism with the leaf plane than there are in the soybean leaf with its rod-like palisade cell structure.

Further research efforts will concentrate on the interpretation of the leaf scattering data. In addition, an engineering task of repackaging the electronics of the apparatus will be initiated so that use of the apparatus by the life scientist will involve a minimal amount of dial-turning.

NATURAL SCENE EMISSIVE RADIANCE SPECTRA

The LARS field van system was used extensively in the 1968 growing season to record emissive spectra from natural scenes with the Block 195T Michelson interferometer spectrometer. Data processing and plotting techniques were improved so that outputs from field scenes would be the actual scene radiance, $N_{\lambda sc}$ (W/cm^2-sr-μm) or equivalent blackbody temperature, T_{BB}, defined by the following equation:

$$N_{\lambda sc}(\text{measured}) = \frac{1.19 \times 10^4}{\lambda^5 \left[\exp(14{,}388/\lambda T_{BB}) - 1\right]}$$

Examples of these outputs are shown in Figures 82 and 83.

The typical routine of calibration and data acquisition in the field is:

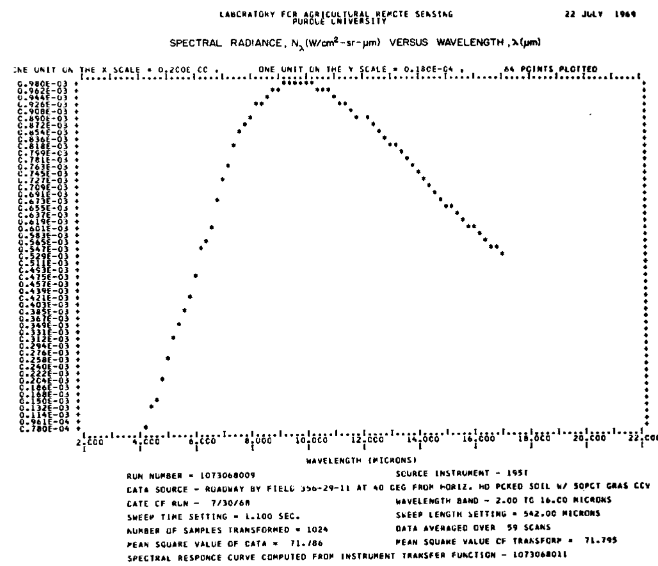


Figure 82. A Typical Scene Spectral Radiance Graph from Block 195 T Spectrometer Data.

(1) Interferograms are recorded with the Block 195T optical head in the field van viewing a conical blackbody standard heated to a thermistor-measured temperature between 40° and 50° C. The instrument cavity temperature is also measured by an internal thermistor. When these interferograms are averaged and the inverse-transformed, the resulting instrument response, $R_{std}(\lambda)$, is proportional to the difference between the blackbody standard radiance, $N_{\lambda std}$, and the instrument blackbody radiance, $N_{\lambda i}$. Therefore, the formula is:

$$R_{std}(\lambda) = K_{\lambda} \, (N_{\lambda std} - N_{\lambda i})$$

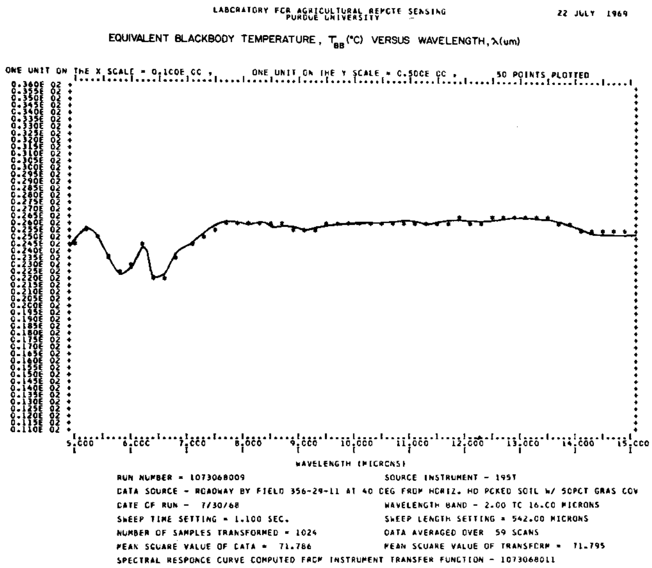


Figure 83. A Typical Equivalent Blackbody Temperature Graph from Block 195 T Spectrometer Data. Interpolation Between 1/2°C Points is Made From Tabular Data

where K_λ is the instrument throughout, the basic function determined by this measurement. Thus, the result is:

$$K_\lambda = \frac{R_{std}(\lambda)}{N_{\lambda std} - N_{\lambda i}}$$

(2) Scene spectra interferograms are recorded with the Block 195T optical head viewing the scene from the cherry picker bucket. The instrument cavity temperature is recorded for each run. When these interferograms are inverse-transformed, the result is an instrument response for the scene, $R_{sc}(\lambda)$. Therefore the formula is:

$$R_{sc}(\lambda) = K_\lambda \, (N_{\lambda sc} - N_{\lambda i})$$

$$= \frac{R_{std}(\lambda)}{N_{\lambda std} - N_{\lambda i}} \, (N_{\lambda sc} - N_{\lambda i})$$

where $N_{\lambda sc}$ is the scene spectral radiance. $N_{\lambda i}$ in the numerator is the instrument cavity blackbody radiance at the instrument temperature during the scene measurement. $N_{\lambda i}$ in the denominator is the instrument cavity blackbody radiance at the instrument temperature during the calibration run in (1).

(3) Additional calibration runs are recorded on tape during the day, so that a calibration run is at the beginning and end of each reel of analog data tape.

Results from several scenes taken both during and between flight missions are presented in Figures 84 through 89. Except for scenes containing a large portion of soil, the equivalent blackbody temperature curves are usually flat within 1°C from about 8 to 13 micrometers. Restrahlen structure is present in the soil scene spectra in the silica range from about 8.5 to 9.4 micrometers. Data consistency comparisons indicate the absolute accuracy of the equivalent blackbody temperature curves to be approximately ± 1°C while the peak-to-peak noise level is less than .25°C.

PRISM FIELD SPECTRORADIOMETER

A Perkin-Elmer Model 98 monochromator is being modified for field spectroscopy. The adjustable slit system of the original equipment is replaced by a fixed entrance aperture,

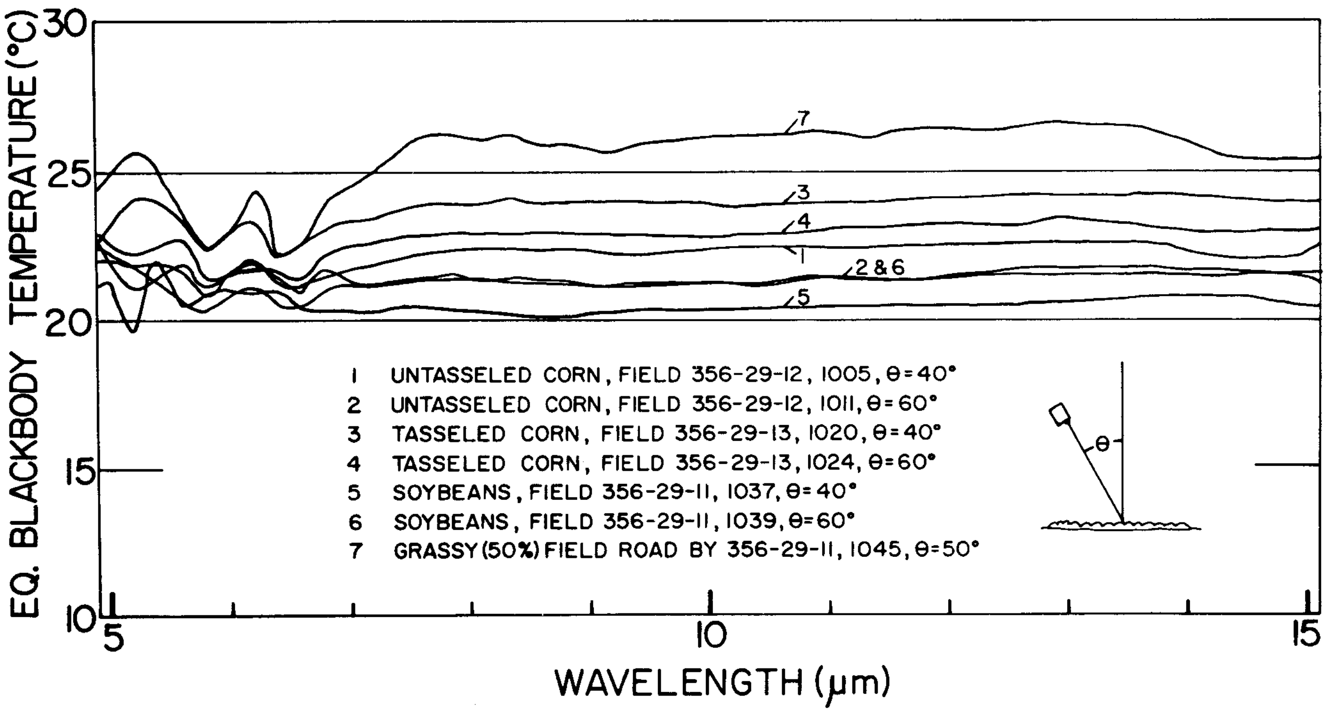


Figure 84. Equivalent Blackbody Temperature Spectra from Field Scenes on July 30, 1968, the Day of a NASA Flight Mission

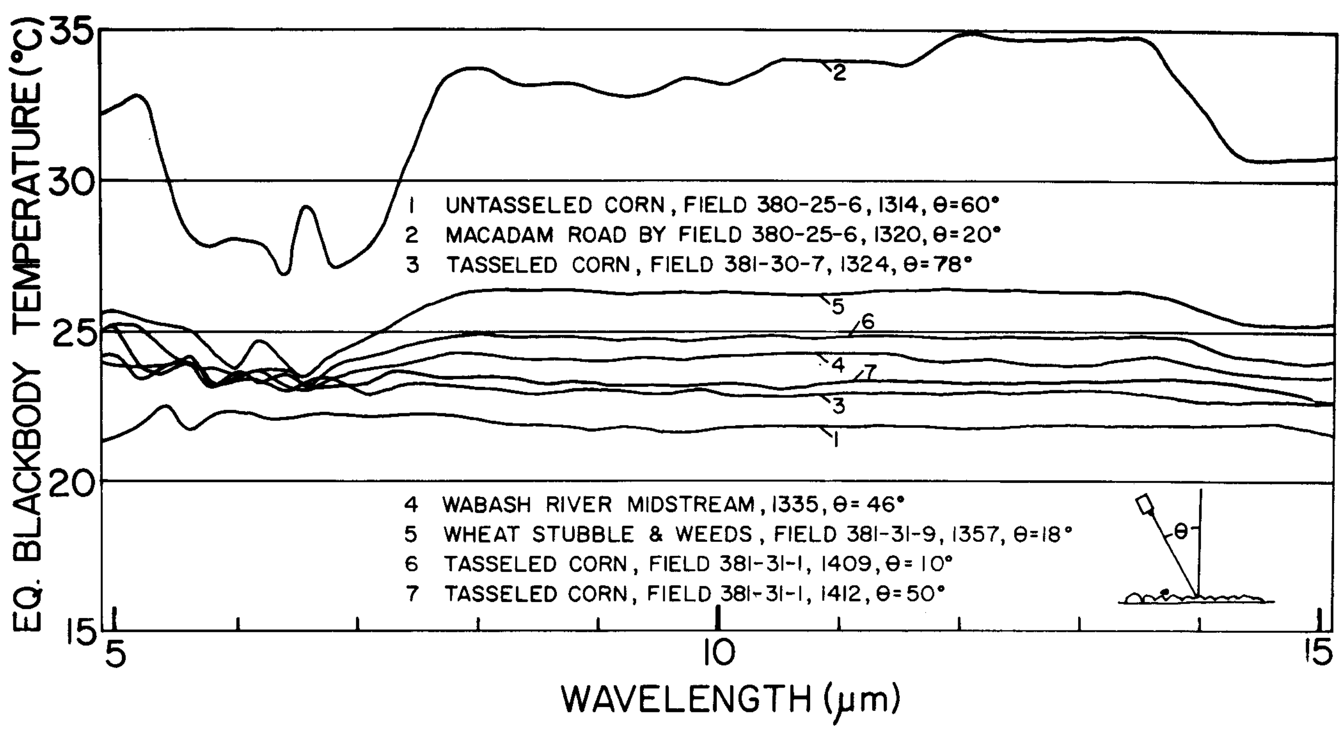


Figure 85. Equivalent Blackbody Temperature Spectra from Field Scenes on July 30, 1968, the Day of a NASA Flight Mission

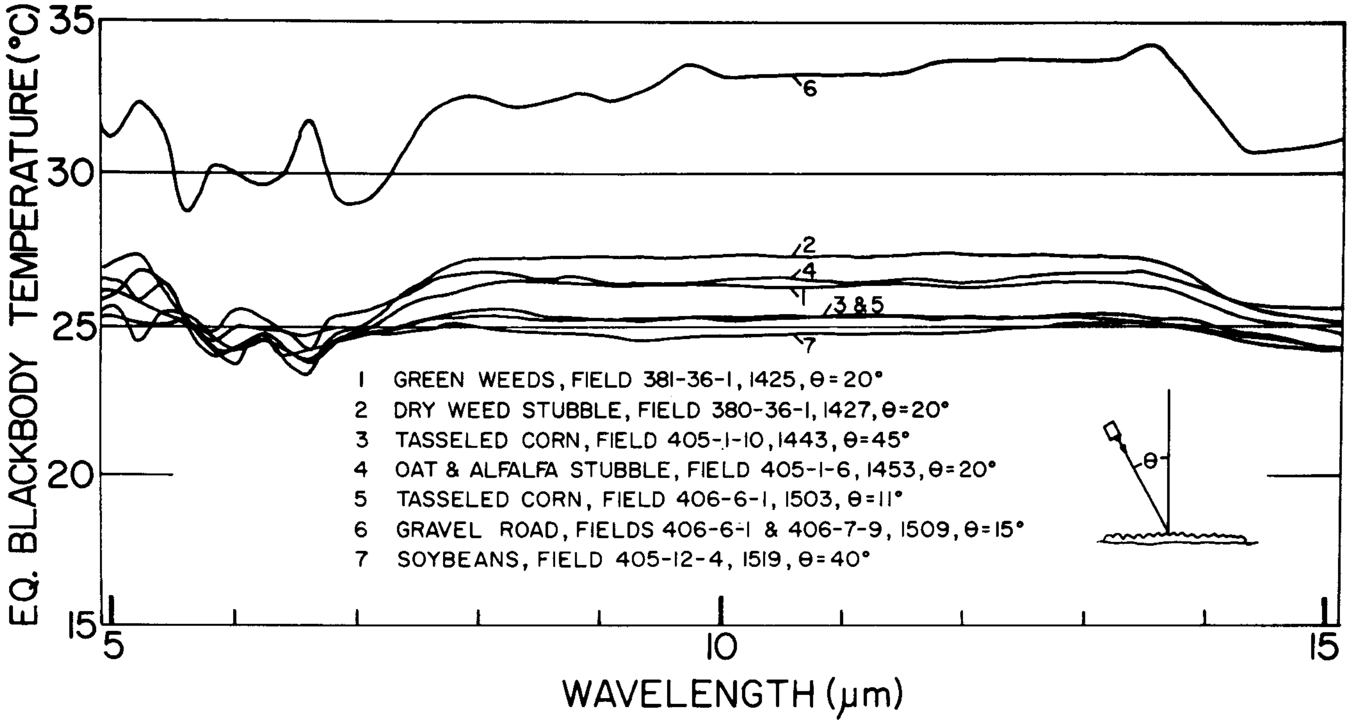


Figure 86. Equivalent Blackbody Temperature Spectra from Field Scenes on July 30, 1968, the Day of a NASA Flight Mission

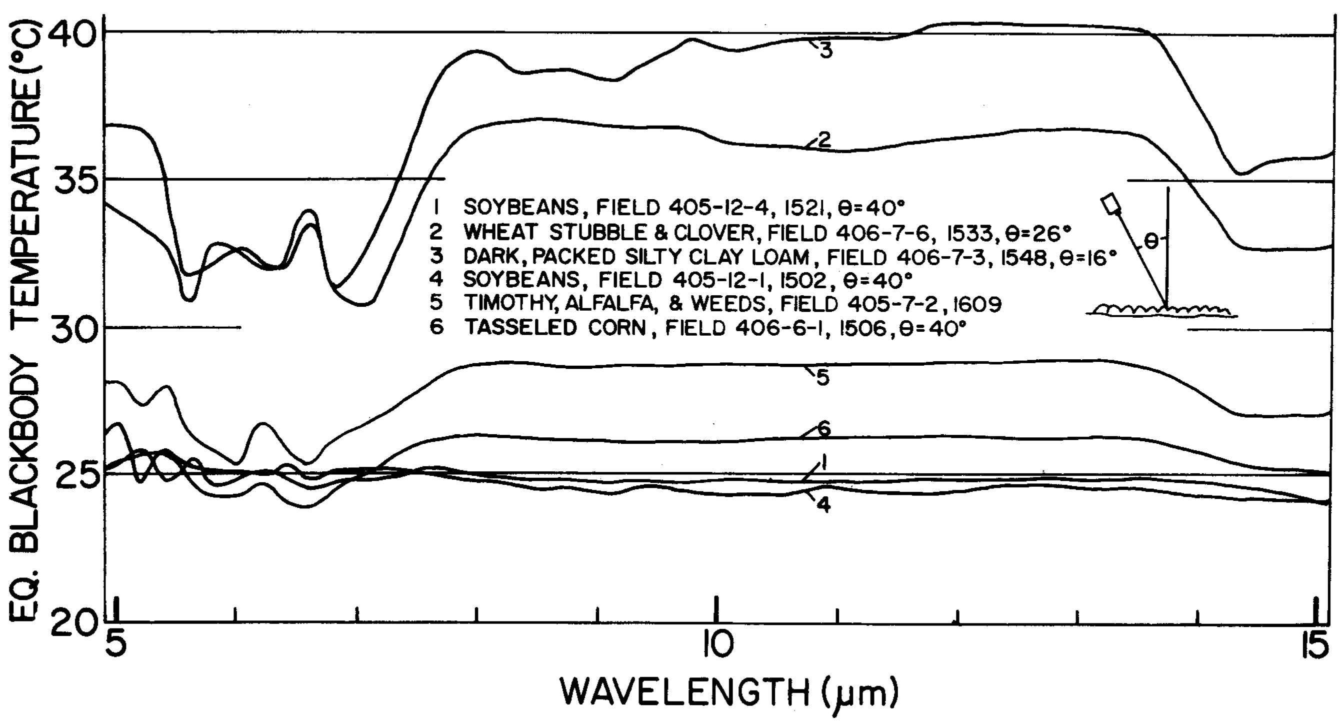


Figure 87. Equivalent Blackbody Temperature Spectra from Field Scenes on July 30, 1968, the Day of a NASA Flight Mission

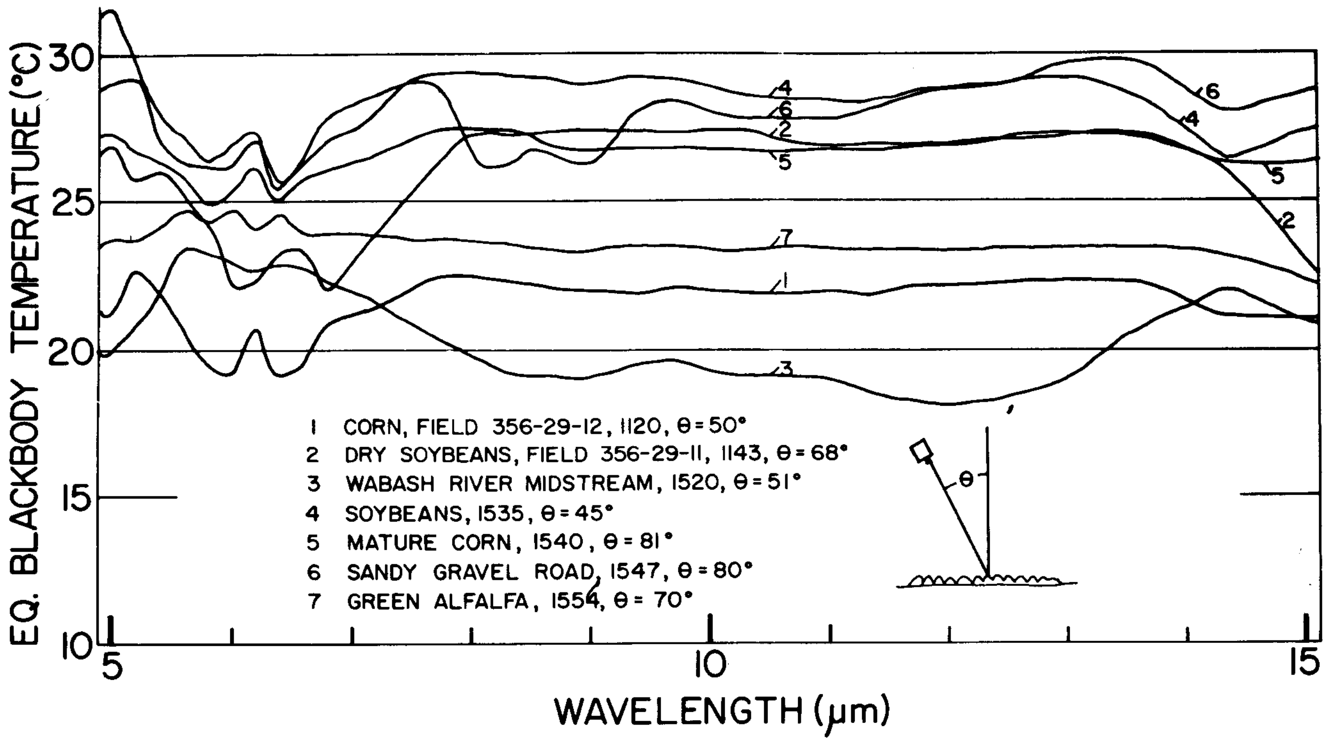


Figure 88. Equivalent Blackbody Temperature Spectra from Field Scenes on September 26, 1968, the Day of a NASA Flight Mission

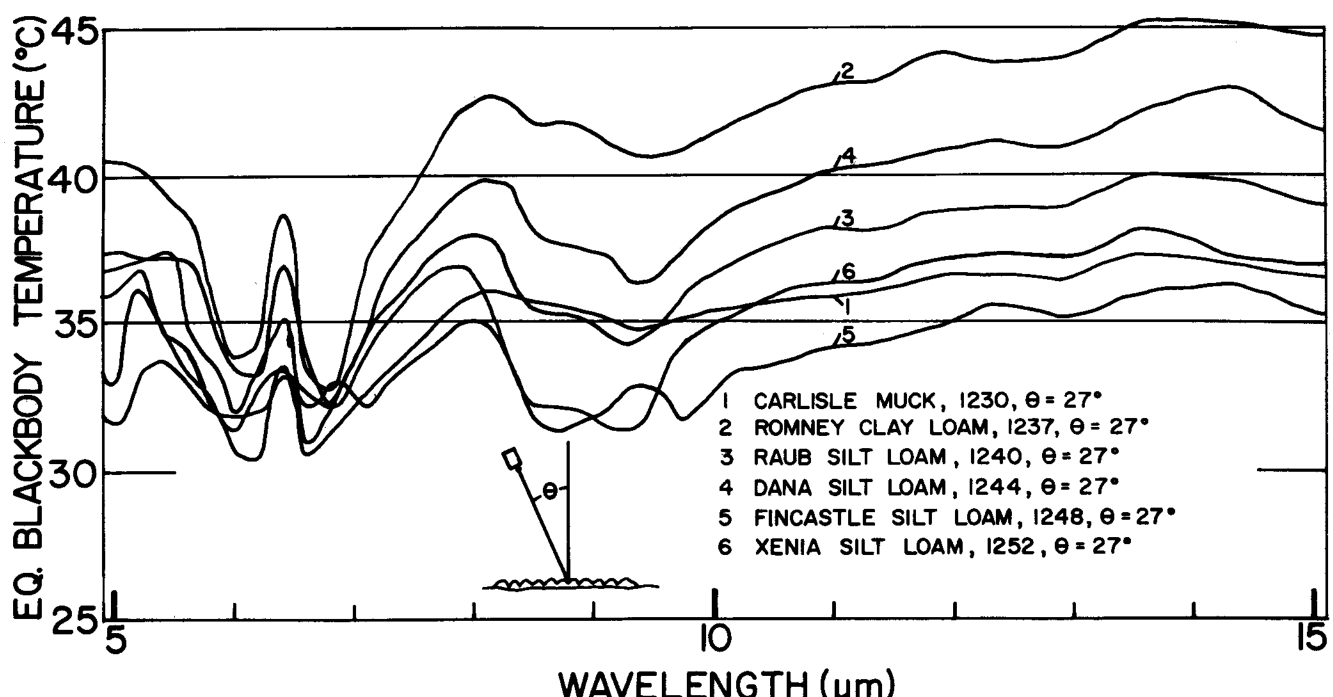


Figure 89. Equivalent Blackbody Temperature Spectra from Experimental Soil Plots, September 26, 1968

1 millimeter square. The exit slit is defined by an aperture plate covering two fiber optics bundles and by a PbS detector with an active area of 1 millimeter square. The two fiber optics bundles transmit to a photomultiplier tube and a silicon photovoltaic detector, respectively. The exit aperture and detector mount is indicated by the arrow in Figure 90. A preamplifier chassis is mounted next to the detector mount and behind the black sheet metal screen from the collimating and dispersing optical components.

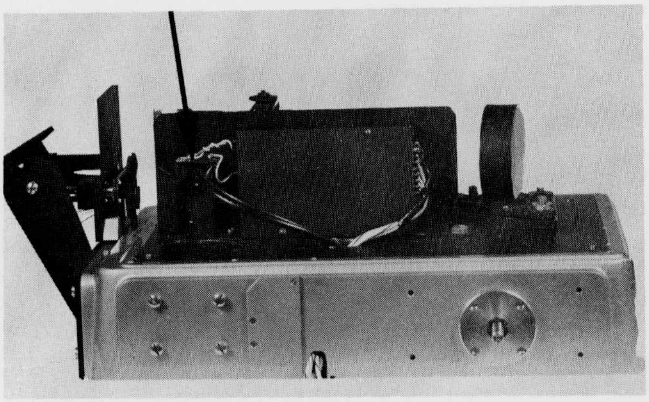


Figure 90. Exit Aperture and Detector Mount Indicated by Arrow

Figure 91 shows the chopped motor, chopper wheel, and plane mirror mount on the object side of the entrance aperture of the monochromator. This entire portion will be enclosed in a cover box with one viewing port out the bottom for gathering scene radiance. An optical reference pick-up for the synchronous demodulation has been constructed and will be mounted around the chopper blade. Figure 92 shows a close view of the Littrow system, entrance aperture, exit aperture, chopper wheel and motor, and the plane mirror mount. Preliminary tests of the system indicate good signal-to-noise ratio performance on natural scenes.

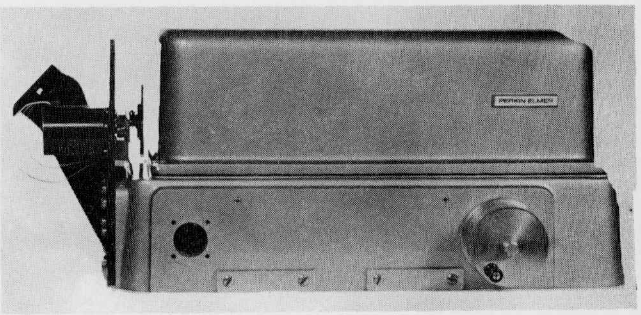


Figure 91. Object Side of Entrance Aperture

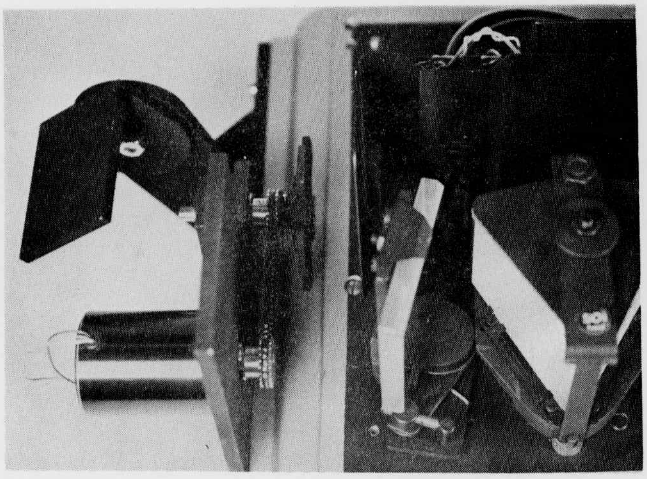


Figure 92. Close-Up View

A wavelength scanning motor with a reversing logic system was installed, tested, and discarded because of large current transient upset of the detector and preamplifier system during motor reversal. A continuous unidirectional motor-cam system to rock the Littrow mirror arm back and forth is in the design stage.

BLOCK POLARIZATION SPECTROMETER

One further test was performed on the Block P-4 Polarization Interferometer

Spectrometer to check the motion integrity of the Soleil compensator prism mount. A front-surface mirror on a microscope slide was glued to the prism mount and a helium-neon laser beam was directed at the mirror as the prism was driven back and forth. The reflected beam was observed on white paper at 4 meters from the prism mount. The beam spot on the paper bounced erratically over a 1 by 10 centimeter rectangle, indicating a rocking of the prism about its main bearing shaft of about 0.7°. This measurement was made without any attempt to stabilize the prism with a guide channel as discussed in the previous LARS annual report.

It is clear that any further designs of such an instrument must consider prism motion integrity. The instrument will be returned to the cognizant lending NASA laboratory.

CHAPTER 6

NASA DATA EVALUATION

INTRODUCTION

In 1968 flight missions were conducted at three selected times of the growing season. In addition to their high value for recording ground truth, the color and color infrared films have been studied to determine:

(1) The utility of the different film types to identify crop species.

(2) The utility of different film types to determine crop conditions.

(3) The utility of different photo scales.

(4) The utility of different film types to determine soil type and conditions.

(5) The utility of good-quality color and color infrared photography in the analysis of multispectral scanner data.

PHOTOGRAPHIC REQUIREMENTS

LARS/Purdue requested NASA to obtain simultaneously color and color IR film over test site 44 (Figure 93), Tippecanoe County, Indiana. This involved having the flight lines (Figure 94) in Tippecanoe County be overflown during three specific time periods in the 1968 growing season.

Past experience in using these emulsions has indicated the utility of each type for identifying various earth surface features.

The NASA convair was able to obtain complete photographic coverage of each flightline for the following missions and dates:

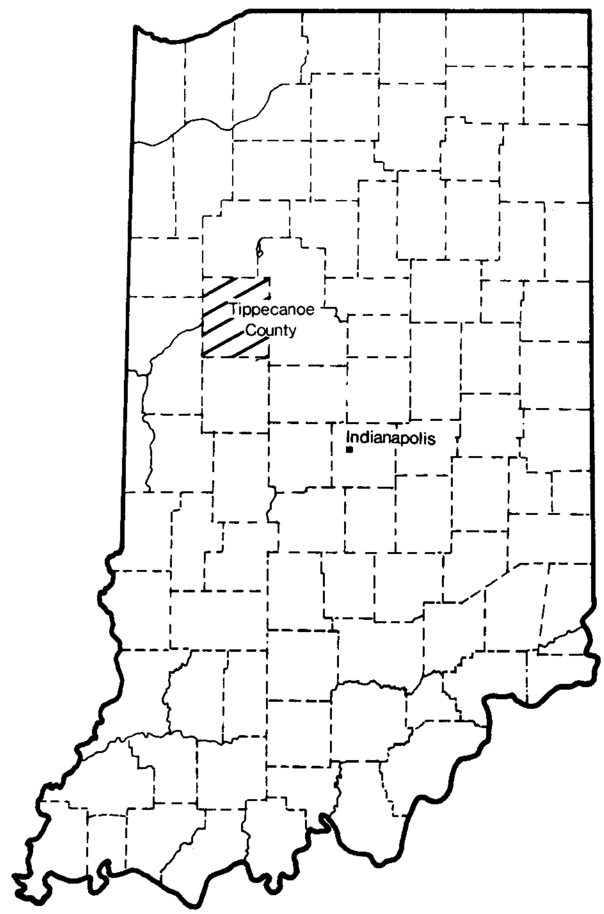


Figure 93. Test Site 44, Tippecanoe County, Indiana

- Mission 72, May 6, 7
- Mission 74, June 18, 19
- Mission 77, July 30.

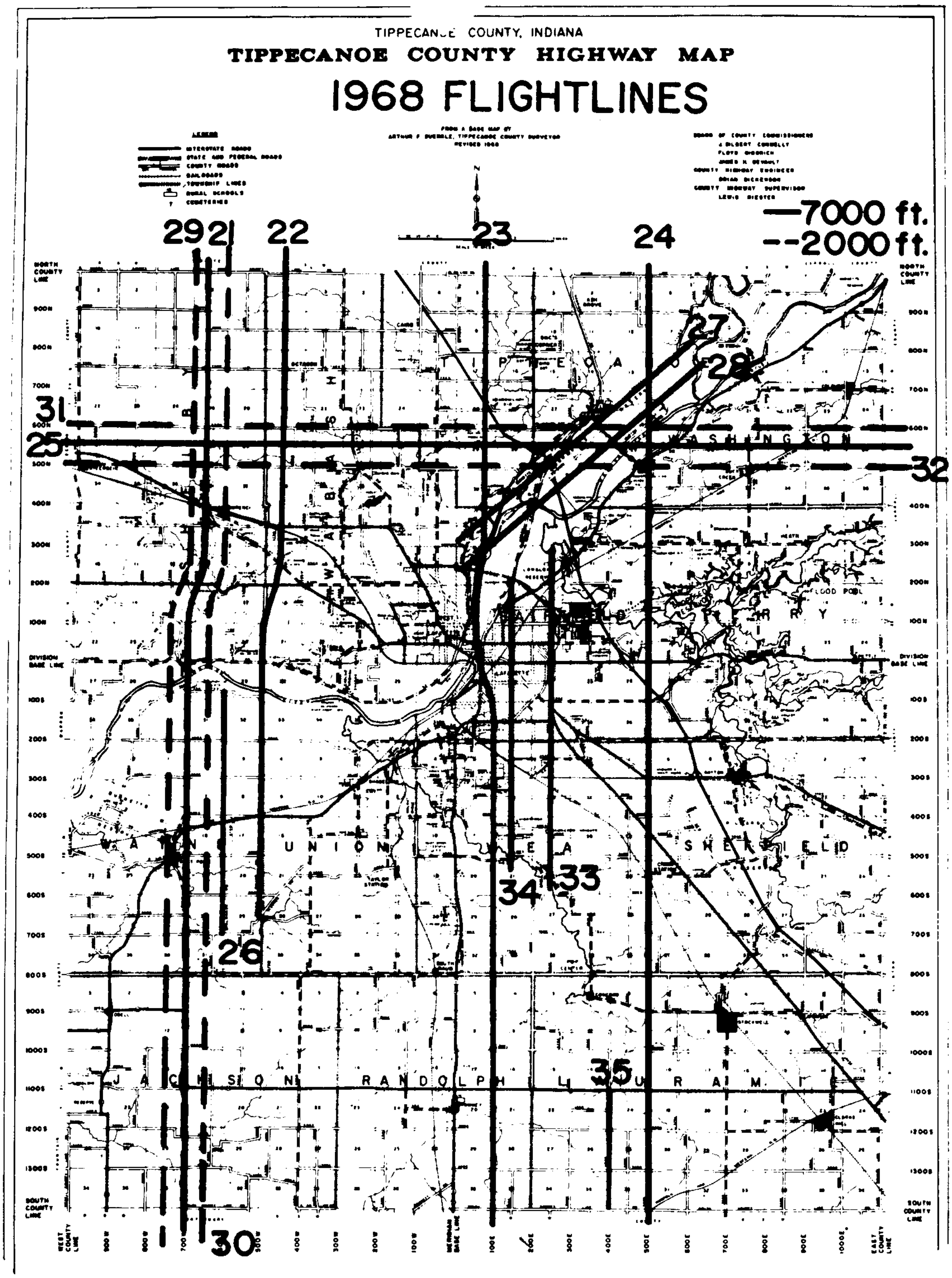


Figure 94. Flightlines Flown by NASA Aircraft in 1968

PHOTOGRAPHIC ANALYSIS

MISSION 72

Primary considerations for Mission 72 were mapping soils and fields of winter wheat.

Soil Mapping

In mid-spring much of the agricultural land in central Indiana is or has been plowed. Vegetative cover is limited to pastures, winter wheat, hay crops and naturally occurring vegetation.

Contrasts between soil types are markedly exhibited in freshly overturned soils. Soil boundaries are distinctly separable on both film types. The soil boundaries are not distinct between soils of similar colors in fields that were fall-plowed. Figure 95 illustrates the differences between spring-plowed (E-2) and fall-plowed (E-3) fields.

Vegetation Mapping

Pasture, winter wheat, hay crops and natural vegetation comprised the majority of ground cover during Mission 72.

Areas of natural vegetation (such as areas of decidious and coniferous tree species around point F in Figure 95 are easily identified with either film type by their characteristic tones and textures. Winter wheat (A in Figure 95) appears to be a fine mat of

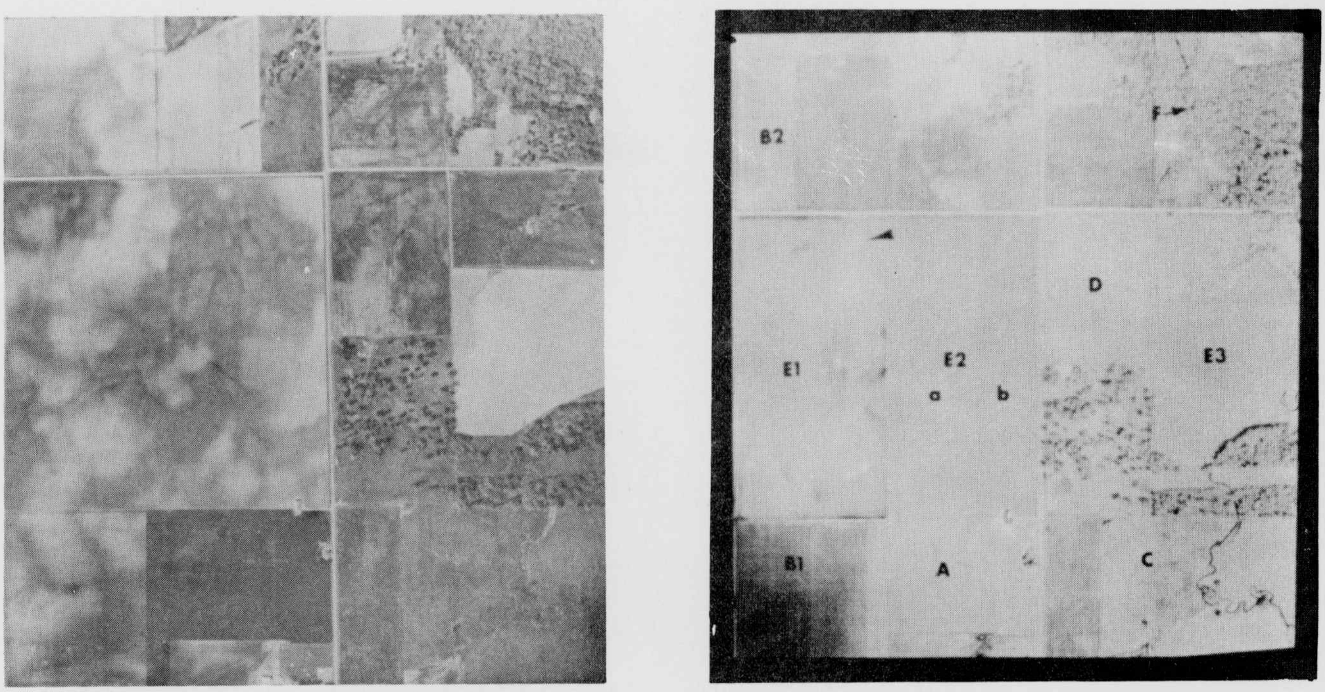


Figure 95. Mission 72 (May) Color and Color Infrared Data Showing (A) Winter Wheat, (B) Oats, (C) Pasture, (D) Hay Crops, (E) Soils, and (F) Stream

continuous uniform tone and texture. Pastures (C) are mottled due to large amounts of dead, dry vegetation remaining from the previous year's growth. Red clover and alfalfa hay crops (D) appear similar to wheat fields but are not as uniform. Hay crops usually have a mixture of grass and legume species and therefore appear mottled.

Oats, although planted and growing, were not of sufficient height or ground cover to be identified on the color photographs. However, on the color infrared film, distinct rows of growing plants could be identified. Fields B1 and B2 in Figure 95 are good examples of the enhancement effect of color infrared film. The relative case of detecting oats on the color infrared, even in field B2, makes this film type for mid-spring crop surveys.

MISSION 74

The primary objective of Mission 74 was to identify corn and soybeans at an early stage of growth.

Heavy rains 40 hours prior to the overflight proved Mission 74 to be extremely useful to soil scientists. Soil drainage patterns were enhanced even though there was some vegetative cover. Soil Conservation Service personnel and LARS researchers were interested in the ease of locating drainage tile lines, some of which had been installed 20 years before the overflight. Figure 96 compares two fields both of which were overflown during Mission 72 and Mission 74. The underground tiles indicated by drainage patterns (arrows in field E2) are distinct on the Mission 74 photography. There is no indication of their existence on the Mission 72 data (upper set of aerial photos).

Ponded fields were easily detected on the color infrared film. The heavily silted standing water appears light blue on the color infrared film. Figure 97 contains portions of color and color infrared aerial photos along the Wabash River in Tippecanoe County. Points A1 and A2 represent low lying flooded areas which are easily detected on the color infrared film. The pond at B appears black and can be considered to be silt free. The drainage ditch along C appears to be at capacity, note the ease of identifying this feature on the color film.

Due to the heavy rains and serious flooding, the primary considerations for flying Mission 74 were not completed. Corn planted before May 15 was easily separated from soybeans because of a higher percentage of ground cover. Corn planted after June 1 was especially difficult to distinguish from soybeans. The effects of the heavy rain had a conflicting effect on the separation of these crops. All soils had a darker tone due to moisture.

Figure 98 is an aerial view of the same area seen in Figure 95 but taken during Mission 74. It is interesting to note the change that has occurred over the one and a half month time span. Winter wheat (A) is close to maturity, indicated by the brown tone on the color film. Oat fields (B1 and B2) are easily distinguished during this time in June. The hay field (D) is matured and is being harvested. Pasture (C) is still separable from other green vegetation due to its mottled appearance.

Areas E1, E2, and E3 have all been planted. E1 and E3 have been planted to soybeans; E3 was planted earlier than E1. Note the ease of detecting planted crops on the color infrared film. Corn has been planted in field E2. Note the similarity in both film types of corn (E2) and early planted soybeans (E3).

In both sets of photography, a stream can be seen in the pasture (C). Following this stream through the farm woodlots is difficult on color film. The streams path (F) can be more easily followed on the color infrared film.

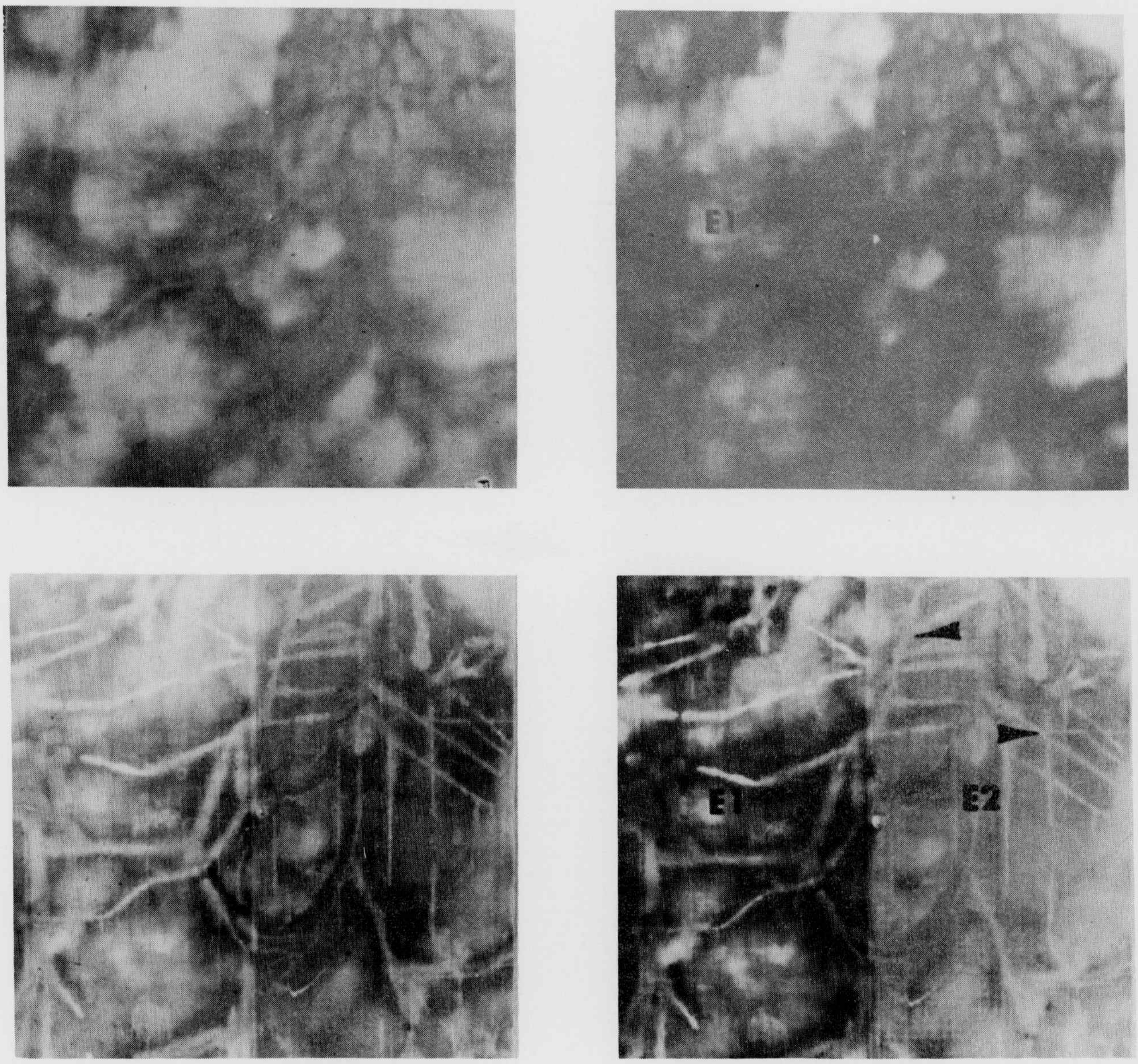


Figure 96. Comparison of the May and June Color and Color Infrared Photography. Note the drainage tile in the lower left of the June photograph.

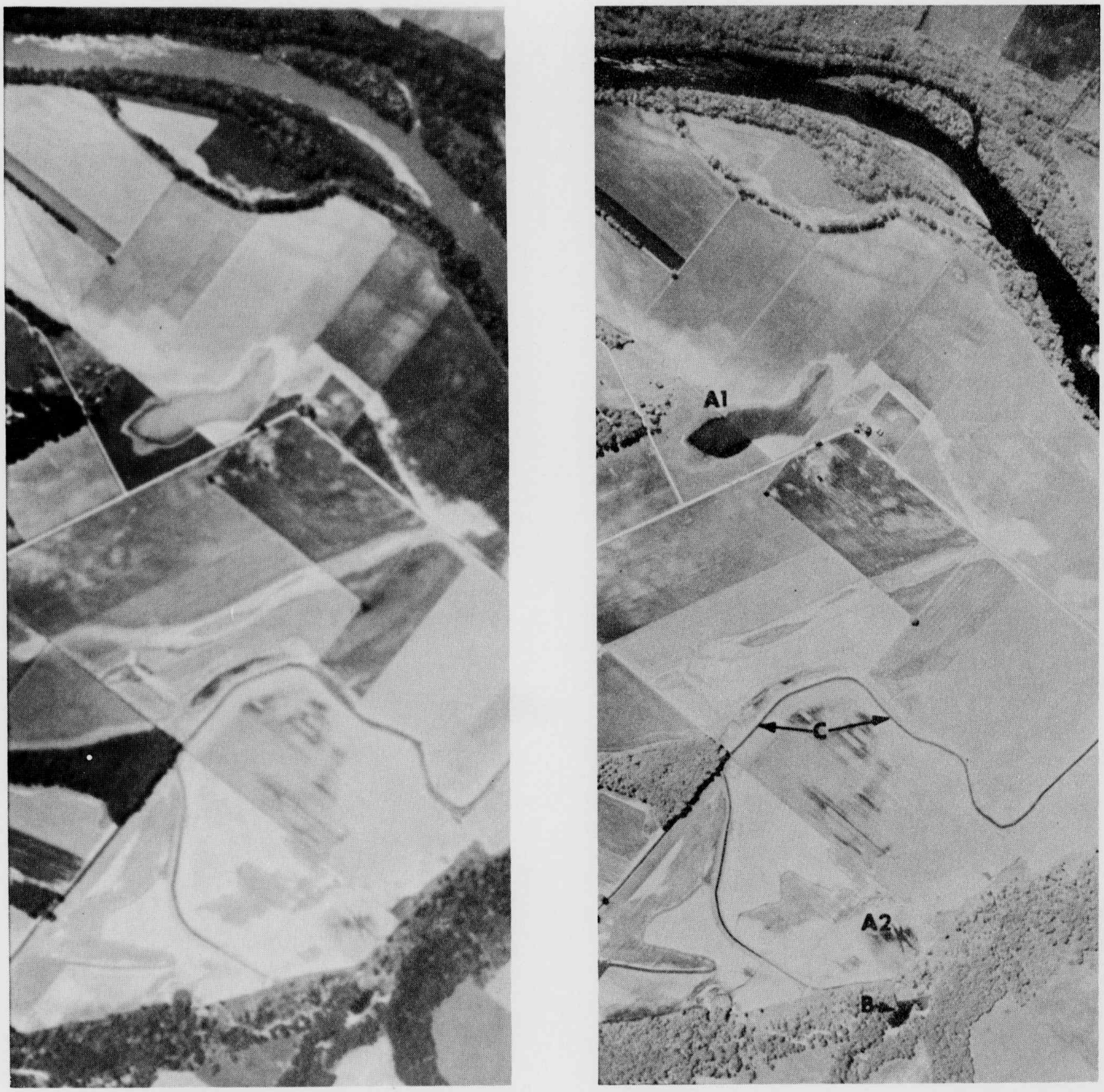


Figure 97. Bodies of standing water are easily detected on the color infrared photo (points A, B, and C) while not as apparent on the color photo taken at the same time. The arrow (upper left) points out a hot spot caused by sun reflecting from the water surface.

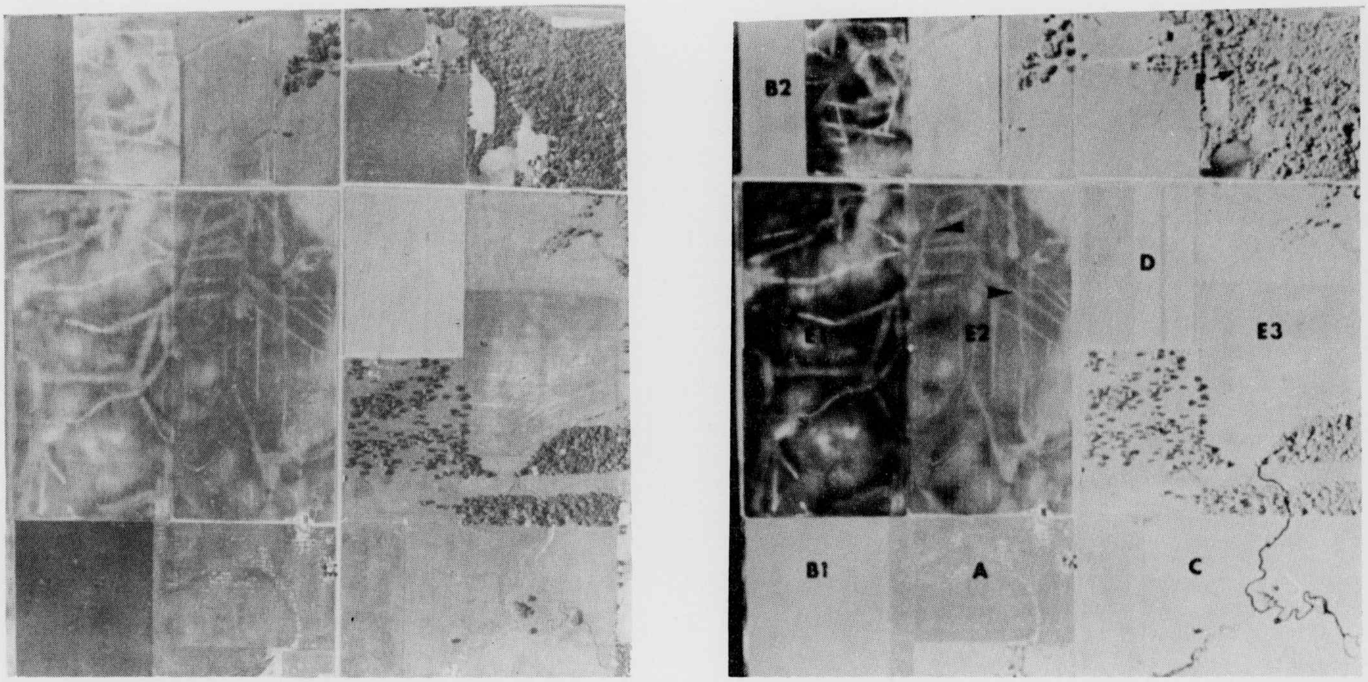


Figure 98. Mission 74 (June) Color and Color Infrared Photography Showing (A) Maturing Winter Wheat, (B) Oats, (C) Pasture, (D) Hay Crop Being Harvested, (E1 and E3) Fields Planted to Soybeans, (E2) Planted Corn and (F) Stream

MISSION 77

Mission 77 was the only mission during 1968 from which LARS was able to obtain both photographic coverage and multispectral scanner imagery on the same date. The primary goal of Mission 77 was to differentiate corn and soybeans during a mature stage in their growing cycles. It was also of interest to see if small grain stubbles could be identified from hay crops.

In 1967 LARS attempted to separate corn and soybeans on scanner data obtained in mid-July and found that species separation at this time was difficult and inaccurate. Ground cover for both species was similar and ground truth photography was not available. By postponing the overflight until late July, it was hoped that the corn would be tasseled. Tasseled corn should be spectrally different from soybeans and therefore, easier to differentiate.

Heavy spring rains required that

some corn fields be partially or fully replanted. Later initial planting dates caused variation in tasseling dates. The replanted corn fields were not tasseled on July 30. Partial or completely non-tasseled fields cause difficulty in accurately classifying the scanner data. However, since aerial photography was obtained on the same day, researchers could locate problem areas to improve final classification of the scanner data. The utility of color and color infrared films as a record of ground conditions at the time of scanner overflights proved to be extremely important.

Figure 99 illustrates the same area as Figures 95 and 97. Sequential photography of this nature is useful in monitoring the development of crops throughout the growing season. The aerial photography illustrated by Figures 95, 97, and 99 was taken approximately at one and a half month intervals. It is interesting to note the crop and field development from early May (Figure 95) to late July (Figure 99).

The field at A which was winter wheat has since been used as pasture. The vegetation is probably volunteer species of clover and grasses growing

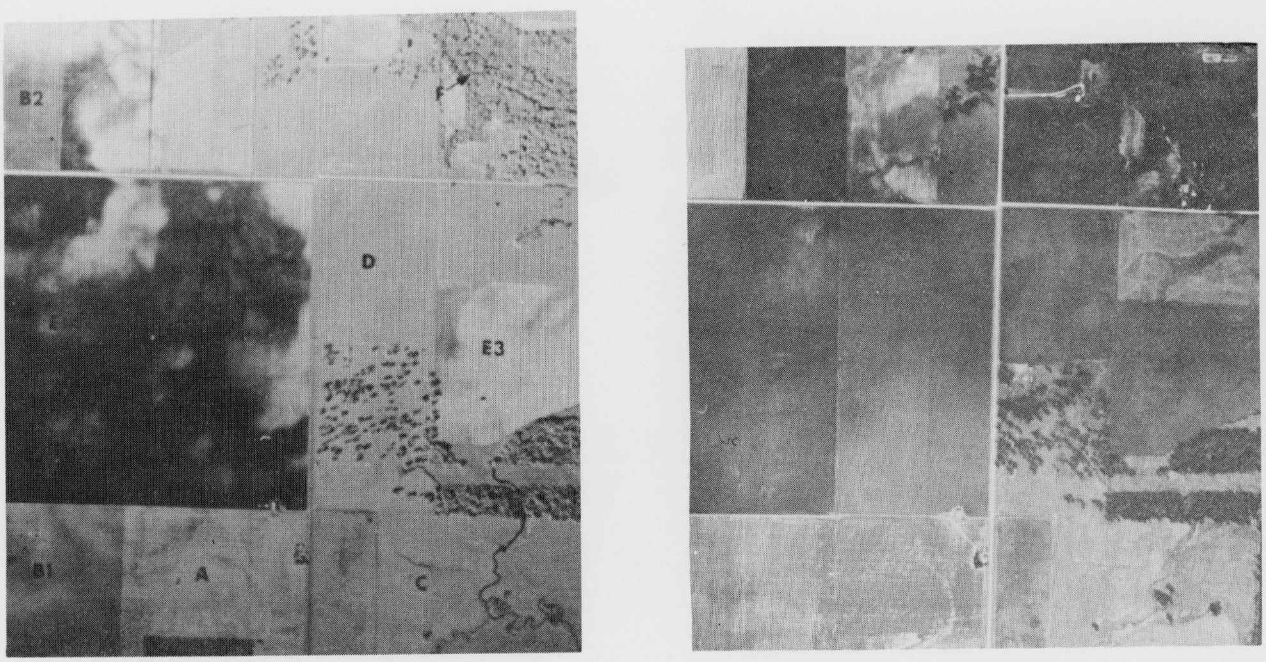


Figure 99. Mission 77 (July) Color and Color Infrared Photography Showing (A) Wheat Stubble and Green Undergrowth, (B) Maturing Oats, (C) Overgrazed Pasture, (D) Hay Crop, (E1 and E3) Soybeans, (E2) Corn Field with Different Varieties--a and b, and (F) Stream

up through the wheat residues.

The oats in fields B1 and B2 have been harvested. Close examination will indicate the path of the combine during harvest.

The pasture (C) contains much dead or dying vegetation. The mottled appearance in this field may be an indication that this pasture has been overgrazed. In a well managed pasture the extent of dead vegetation, as indicated by the mottled appearance on the aerial photos, would not be so great.

Soybean fields (E1 and E3)are ready for harvest. A tractor pulling a picker can be seen along the boundary of fields E1 and E2 indicating their maturity. The effects of the heavy spring rains can be seen in field E1. The portion of the field indicated by the arrow is poorly drained. The poor growth of soybeans in this area indicates that this area was probably flooded soon after planting.

The corn field (E3) exhibits a distinct boundary between a and b. Soil patterns, especially noticeable on the color infrared film are visible in a but not in b. Also portions of a and b are tasseled and the row width appears constant. The difference in appearance is probably due to different varieties of corn in a and b. The variety in b is faster growing and, therefore, tassles sooner than the variety in a. Not much variation can be seen in corn field (E2) in Figure 97. This supports the assumption that the differences between a and b (Figure 99) are due to variety not planting date since the field seems to have been planted at about the same time. Without access to sequential photography, a mistake in interpretation could have easily been made. The nature of the NASA 1968 missions have been invaluable as ground truth aids in such interpretative instances.

Even though there is considerable ground cover during this time of year, the stream along Point F (Figure 99) can still be easily traced on the color infrared film. The capability of this film to accurately differentiate bodies of water make it an envaluable aid in ground water studies.

CROP SPECIES IDENTIFICATION STUDIES

An experiment was conducted with the Mission 77 photography to determine which altitude and film type could best be used to identify crop species. The test was conducted in a 7.5 square mile area along Flight Line 25. In this area 59 fields of soybeans were located. Nine of these fields had not been identified by the ground truth personnel because the fields were located in the center parts of the sections. On the 1:14,000 scale color photography, positive identification was made on only 48 soybean fields. Six more were questionable, three were misidentified, and two could not be even tentatively identified with the color photography. On the 1:14,000 scale color infrared photography, 56 of the 59 fields were positively identified. There were no fields listed as unknown; one field was listed as questionable and two fields were misidentified. Using the 1:4,000 scale, either color or color infrared photography, 58 of the 59 fields were positively identified and one was listed as being a questionable identification.

In the same area, a total of 60

corn fields were examined. Of these, 10 had not been identified by the ground truth personnel because they were located in the center areas of the sections and were not accessible. Forty-seven of the corn fields were positively identified using the color photography, 1:14,000 scale. Of the remaining 13 fields, one had been misidentified, three were listed as questionable but believed to be corn, and nine fields could not be even tentatively identified so were listed as unknown. On the color infrared photography, 1:14,000 scale, 58 of the 60 fields were positively identified. The remaining two fields were listed as questionable. There were no fields listed as not identified or misidentified. Using the 1:4,000 scale photography, either color or color infrared, allowed 59 of the 60 fields to be positively identified, while one field was listed as being questionable.

Hay, pasture, and diverted acres (land retired by the farmer for government payment) were very difficult to positively identify on any scale or film type. In some cases, fields seeded with oats in the spring had been harvested by this time of year allowing an undergrowth of alfalfa or red clover to be visible. It was difficult to positively distinguish between fields of oat stubble with an undergrowth of alfalfa from recently mowed and harvested hay fields. In cases where there was no undergrowth of alfalfa or red clover, stubble fields of winter wheat or oats were readily identifiable on both film types and scales. Thus, at this time of the growing season, three general agricultural categories can be readily distinguished. These are (1) small grains, either wheat or oats; (2) row crops, either corn or soybeans; and (3) forage areas, either hay, pasture or diverted acres. Corn fields were in various stages of tasseling. Those which were completely tasseled could be readily differentiated from the soybeans. In many cases, those that had not tasseled were quite difficult to distinguish from soybeans, particularly on the color photography at 1:14,000 scale. In general, the color infrared photography taken at 1:14,000 scale was more satisfactory for identification of corn and soybeans than the comparable color photography. However, for positive identification of the crop types and species a lower altitude is recommended for future flight missions.

GENERAL CONSIDERATIONS

The LARS staff was pleased with the rapid turn-around time of the NASA labs in returning the aerial films. This is especially important as an aid in classifying scanner data, as in Mission 77.

Image quality was generally good, although color shifts were noted from mission to mission.

Vignetting appeared during all missions. It becomes a serious problem if extensive interpretation is considered along the film edges and corners.

Tilt caused some problems in estimating crop cover percentages. Varing scale factors encountered in a tilted frame can cause errors in acreage measurements and other quantative measures.

The results can be summarized as follows:

(1) Color photography is the easiest to use for purposes of crop species identification.

(2) For fields with a low percentage of vegetative cover, color infrared photographs are superior to color photographs in determining the presence of vegetation. It is extremely difficult to differentiate between completely bare soil and a low percentage of vegetative cover on color photographs.

(3) Color infrared photography is more useful than color photography for determining the presence of bare soil as opposed to dry, dead vegetation in spot locations within a canopy of green vegetation.

(4) Conditions of crop health and maturity can best be studied using both color and color infrared photographs in combination. Color photographs are easier and more reliable for interpretation purposes, but subtle differences in spectral reflectance that would be overlooked on color photographs are often more apparent on color infrared photographs.

(5) Crop canopy conditions can be grossly misinterpreted at view angles away from nadir. This is largely a function of row direction in conjunction with look angle. Crop height and row width also play a part. Precise limitations on allowable look angle have not yet been determined.

(6) Scales of 1:14,000 are very useful for aiding in the analysis of multispectral scanner data. Many crop species identifications can be made from photography at this scale. However, 1:4,000 scale photographs are required in many instances for positive identification. Therefore, both 1:14,000 and 1:4,000 scale photographs are useful for many purposes of crop species and condition identification.

(7) Color photographs are useful for purposes of soil type and condition classification. However, in many instances, color infrared photographs were required to differentiate (a) between bare soil and dead vegetation, or (b) between completely bare soil and fields containing recently germinated crops.

(8) Color and color infrared photographs were equally useful in locating fields containing drainage tiles and in mapping out the location and pattern of the tiles. On the 1:4,000 scale photographs, inferences could be drawn about the age of the tile system in many instances.

(9) NASA photographic coverage is most useful if obtained over the same flight lines several times throughout the growing season.

CHAPTER 7

AIRCRAFT DATA ANALYSIS

INTRODUCTION

Previous reports[1/][2/] have indicated that the approach for analyzing multispectral data which is under research at LARS is a type of multivariant analysis referred to as pattern recognition and that it is desirable to determine the circumstances of existence of "spectral signatures" of various materials rather than to attempt to determine precisely what the signatures are. The latter is desirable due to the large number of experimental variables (many of which are not measurable directly) which are active during any given aircraft or spacecraft data gathering activity.

The procedure followed is to use a pattern recognition scheme to evaluate the degree of existence of a unique signature for a given surface material under given circumstances. More specifically, the accuracy with which data can be assigned to a desired category by the recognition scheme reflects the degree of existence of a unique "signature" for that category. It is the purpose of this chapter to present some results obtained during the year by this process. However, before proceeding with the details of the procedure used and the results, some additional aspects of this approach will be pointed out.

The following steps must be followed in testing for the existence of a signature. Data is gathered under appropriate circumstances using the aircraft system and at the same time ground truth in the form of photography and literal information is collected over the flightline. An exhaustive but nonoverlapping set of classes of surface cover types are defined for the flightline in such a way that the set contains the particular class or classes of interest. Next training samples which are descriptively typical of each class are located in the aircraft data utilizing ground truth from only a small part of the flightline. After the classification is run on the whole flightline, ground truth from the remainder of the flightline is used to evaluate the classification and to determine the overall accuracy.

This same procedure, excluding the evaluation and accuracy determination portion and the extensive ground truthing it requires, would offer many advantages as a mode of operation for an operational system. At the time of the satellite pass for data gathering, a small area would be ground truthed to finalize the class set and determine training areas. Satellite data from these areas would be used to train a classifier; this classifier would presumably be valid over a much larger area. This procedure might be referred to as an extrapolation mode of operation in that the classifier merely extrapolates information it is given about a small area to a much larger area.

1/ Laboratory of Agricultural Remote Sensing, "Remote Multispectral Sensing in Agriculture," 1967. Vol. 2 (Annual Report), Purdue University Agricultural Experiment Station Research Bulletin 832, Lafayette, Indiana.

2/ Laboratory of Agricultural Remote Sensing, "Remote Multispectral Sensing in Agriculture," 1968. Vol. 3 (Annual Report), Purdue University Agricultural Experiment Station Research Bulletin 844, Lafayette, Indiana.

Thus, it is perhaps more precise to say that the classifier compares unknowns with knowns rather than to say it identifies unknowns.

There are a number of important advantages to this approach. One is that it does not require the existance of a "signature bank" together with associated maintenance, calibration, storage and retrieval problems. Further, calibration of the sensor output to the point of determining scene radiance in known units is not necessary; only a sensor system of fixed but perhaps unknown overall sensitivity is required. Also since training data is drawn from within the data to be classified, variables such as sun angle, season-to-season considerations, microclimatic variations, and the like are automatically normalized.

As a result of this concept, the aircraft data analysis research had the dual role of determining the circumstances of the existence of unique signatures and developing suitable means for the rapid training of classifier devices. To test the current training procedure relative to this concept, an "operational analysis" experiment was conducted in which it was attempted to obtain analysis results as soon after the flight as possible. This experiment together with other results are described in this chapter.

DISCUSSION OF DATA ANALYSIS RESULTS

FLIGHTLINE C-1, JULY, 1966

Five wavelength bands in the 0.4 to 1.0 micrometer range were selected to classify this flightline. A summary of the accuracy of the classification of the training classes is shown in Table 7. Seven different classes of data were used in this study. The table shows the total number of samples or remote sensing units (RSU) and the correct number of RSU's which were identified for each class. A breakdown of how the RSU's were classified is also shown so that the classes being confused with one another can be visually determined.

The overall performance on training fields for this flightline (2753 RSU's) was 95.1 percent correct recognition; for test fields (17,233 RSU's), the overall performance was 77.1 percent (Table 8). These results represent above average performance on training data and somewhat above average on the test data. This is probably due to the uniformity of the soils and therefore the uniformity of crops along the flightline.

The overall test data performance was lowered because of a tendency to classify soybeans as corn. Soybeans comprise approximately one third of the test data. This misclassification is fairly common in late July when corn and soybean plants are very green and exhibit a large degree of uniform ground cover.

No attempt was made to discriminate between hay (HAY), pasture (PT), and stubble (SUB) which comprised the "Mixed" test category. Such discrimination involves rather subtle distinctions which are very difficult to make accurately on the relatively small amount of available training data. Often in such a situation, the training samples may be accurately classified but generalization to the test samples may be very poor. This was true in this case.

Table 7. Classification Summary of Training Classes for C-1 Flightline, July, 1966

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	CN1
	CN2
	STUB
	PT
	HAY1
	HAY2
	RYE
	BS
	WATER

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
7	0.55	0.58
9	0.62	0.66
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TRAINING CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	SOY	CORN	HAY	STUB	PT	RYE	BS	WATE	THRS
				NO OF SAMPLES CLASSIFIED INTO								
1	SOY	704	95.9	675	29	0	0	0	0	0	0	0
2	CORN	735	95.1	35	699	1	0	0	0	0	0	0
3	HAY	191	97.9	0	0	187	4	0	0	0	0	0
4	STUB	530	88.3	0	0	20	468	3	39	0	0	0
5	PT	288	99.7	0	0	1	0	287	0	0	0	0
6	RYE	160	97.5	0	0	0	4	0	156	0	0	0
7	BS	120	100.0	0	0	0	0	0	0	120	0	0
8	WATE	25	100.0	0	0	0	0	0	0	0	25	0
	TOTAL	2753		710	728	209	476	290	195	120	25	0

OVERALL PERFORMANCE = 95.1

AVERAGE PERFORMANCE BY CLASS = 96.8

During the last week in July, the categories of hay, pasture, and stubble have visible similarities. The wheat and oat stubbles are beginning to appear green from the undergrowth of red clover and alfalfa which was seeded with these crops to improve the soil fertility. Some stubble fields are also being pastured. The pastures are greening up with new growth since the cattle or sheep are grazing. The hay crops have recently been mowed and new growth has just begun.

FLIGHTLINE C-2, JULY, 1966

A slightly different combination of five wavelength bands in the 0.4 to 1.0 micrometer range was selected to classify the C-2 flightline. Overall performance on training fields (Table 9) for this flightline (2401 RSU's) was 96.7 percent; for test fields (13,121 RSU's) the overall performance was 81.1 percent (Table 10). These results were again slightly above the levels observed in other agricultural areas.

Table 8. Classification Summary of Test Classes for C-1 Flightline, July, 1966

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	CN1
	CN2
	STUB
	PT
	HAY1
	HAY2
	RYE
	BS
	WATER

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
7	0.55	0.58
9	0.62	0.66
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TEST CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOY	CORN	MIX	BARE	WATE	THRS
1	SOY	5866	69.5	4079	1662	117	5	3	0
2	CORN	3718	85.6	507	3182	21	3	5	0
3	MIX	7483	78.4	1032	441	5867	111	32	0
4	BARE	138	94.9	2	0	5	131	0	0
5	WATE	28	96.4	0	1	0	0	27	0
	TOTAL	17233		5620	5286	6010	250	67	0

OVERALL PERFORMANCE = 77.1

AVERAGE PERFORMANCE BY CLASS = 85.0

As noted in Figure 10 flightline C-2 is located near flightline C-1. Therefore, the uniformity of the soils and crops as explained earlier are probably causing these high results. As discussed previously, the greatest classification problem was with the corn and soybeans.

An interesting feature of this flightline was a cloud shadow over a sizable portion of the area where there were primarily soybeans, pasture, and stubble fields. Analysis showed that the spectral response for a cover type varied markedly depending upon whether there was a cloud shadow or not. For this reason, separate subclasses were established for the two situations in order to avoid distinctly bimodal statistical distributions. The results showed that there were no other difficulties in classification due to the cloud shadows by the use of the above technique. This would appear to be a useful technique for analyzing data collected on partly cloudy days.

FLIGHTLINE C-1, SEPTEMBER, 1966

Four wavelength bands in the 0.4 to 1.0 micrometer range were selected to classify this flightline. No appreciable difference in the separability resulted from an additional

Table 9. Classification Summary of Training Classes for C-2 Flightline, July, 1966

CLASSES CONSIDERED

SYMBOL	CLASS
	WATER
	TREES
	SOY
	CORN
	STUB
	FRG
	SOY CS
	STUB CS
	PAST CS

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
7	0.55	0.58
10	0.66	0.72
11	0.72	0.80
12	0.80	1.00

CLASSIFICATION SUMMARY BY TRAINING CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOY	STUB	WATE	TREE	CORN	FRG	PAST	THRS
1	SOY	877	97.3	853	0	0	1	19	0	4	0
2	STUB	439	97.0	2	426	0	0	0	2	9	0
3	WATE	54	100.0	0	0	54	0	0	0	0	0
4	TREE	80	98.7	1	0	0	79	0	0	0	0
5	CORN	614	94.6	32	0	0	1	581	0	0	0
6	FRG	175	98.3	0	3	0	0	0	172	0	0
7	PAST	162	96.3	0	6	0	0	0	0	156	0
	TOTAL	2401		888	435	54	81	600	174	169	0

OVERALL PERFORMANCE = 96.7

AVERAGE PERFORMANCE BY CLASS = 97.5

wavelength band than those used in the previous analysis. Overall performance on training fields for this flightline (3952 RSU's) was 85.5 percent (Table 11); for test fields (15,993 RSU's) the overall performance was 77.9 percent (Table 12). The reduced accuracy for the training fields was due to the difficulty encountered in discriminating between pasture (PT), and hay fields (HAY). Most of the hay fields had recently been mowed which gave them the same appearance as the over-grazed pasture fields. The pasture fields had become over-grazed due to a droughty growing season.

The test fields results were about average; no effort was made to discriminate between pasture (PT), hay (HAY), and stubble (STUB) fields which were grouped as mixed category (MIX). However, the soybean test fields were confused with the mixture category. This was probably due to the early varieties of soybeans which had already begun to turn a yellowish brown color similar to the classes in the mixture category.

On the basis of these results, it is reasonable to conclude that mid-September is not a good time of the year to accurately discriminate between

Table 10. Classification Summary of Test Classes for C-2 Flightline, July, 1966

CLASSES CONSIDERED

SYMBOL	CLASS
	WATER
	TREES
	SOY
	CORN
	STUB
	FRG
	SOY CS
	STUB CS
	PAST CS

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
7	0.55	0.58
10	0.66	0.72
11	0.72	0.80
12	0.80	1.00

CLASSIFICATION SUMMARY BY TEST CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO WATE	SOY	CORN	STUB	PAST	TREE	THRS
1	WATE	107	98.1	105	1	0	1	0	0	0
2	SOY	3771	87.7	0	3306	349	21	59	36	0
3	CORN	3839	70.1	0	1084	2692	12	39	12	0
4	STUB	1975	94.2	0	11	7	1860	97	0	0
5	PAST	3429	78.0	0	87	12	652	2676	2	0
	TOTAL	13121		105	4489	3060	2546	2871	50	0

OVERALL PERFORMANCE = 81.1

AVERAGE PERFORMANCE BY CLASS = 85.6

classes of vegetative ground cover. With a given cover type, a wide range of conditions such as maturity and plant moisture are observed at this time of the year. This makes it quite difficult to define homogenous pattern classes. Therefore, under such conditions, accurate classification of these classes could not be expected.

FLIGHTLINE C-1, JULY, 1968

For this flight date, the C-1 flightline is part of a total flight line called PF21. The entire PF21 flightline was flown, but it was determined that the upper part of the flightline would be difficult to use because of cloud conditions. The alternating sunlight and cloud shadows caused severe variations in the data making it difficult to analyze.

Visible and Near Infrared (0.4-1.0 micrometer)

Four wavelength bands were selected by the $SELECT processor for this classification from the 0.4 to 1.0 micrometer range. Overall performance on training fields for this classification (3891 RSU's) was 86.8 percent (Table 13); for test fields (7,135 RSU's) the overall performance was 82.8 percent

Table 11. Classification Summary by Training Classes for C-1 Flightline, September, 1966

CLASSES CONSIDERED

SYMBOL	CLASS
	WATER
	SOY
	CN1
	CN2
	STUB
	PT
	HAY

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
7	0.55	0.58
9	0.62	0.66
12	0.80	1.00

CLASSIFICATION SUMMARY BY TRAINING CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO CORN	WATE	SOY	STUB	PT	HAY	THRS
1	CORN	1303	93.2	1214	5	62	3	15	4	0
2	WATE	62	100.0	0	62	0	0	0	0	0
3	SOY	688	81.8	77	3	563	21	16	8	0
4	STUB	800	87.2	8	0	75	698	17	2	0
5	PT	457	69.6	5	1	8	10	318	115	0
6	HAY	642	81.6	2	15	7	0	94	524	0
	TOTAL	3952		1306	86	715	732	460	653	0

OVERALL PERFORMANCE = 85.5

AVERAGE PERFORMANCE BY CLASS = 85.6

(Table 14). When no attempt was made to discriminate between stubble (STUB), pasture (PAST), and diverted acres (DA), the performance of the training fields improved to 96.1 percent. These results and problems encountered are similar to those reported for the July 1966 data.

Visible and Near Infrared (0.4-1.8 micrometer)

The same study area was reclassified using wavelength bands for a wider range of reflective energy. The $SELECT processor was used to select three visible and two reflective infrared wavelength bands for this classification from the 0.4 to 1.8 micrometer range.

The overall performance on the training fields (4,303 RSU's) was 91.9 percent (Table 15); on the test fields (7,135 RSU's) the overall performance was 83.9 percent (Table 16). When no attempt was made to discriminate between stubble (STUB), diverted acres (DA), and pasture (PAST), the performance on training fields was 95.5 percent.

Additional training samples of corn and stubble were added for this classification but were not used in the previous

Table 12. Classification Summary of Test Classes for C-1 Flightline, September, 1966

CLASSES CONSIDERED

SYMBOL	CLASS
	WATER
	SOY
	CN1
	CN2
	STUB
	PT
	HAY

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
7	0.55	0.58
9	0.62	0.66
12	0.80	1.00

CLASSIFICATION SUMMARY BY TEST CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO WATE	SOY	CORN	MIX	THRS
1	WATE	69	87.0	60	1	3	5	0
2	SOY	4595	55.2	10	2538	803	1244	0
3	CORN	3407	78.8	22	595	2686	104	0
4	MIX	7922	90.6	114	344	290	7174	0
	TOTAL	15993		206	3478	3782	8527	0

OVERALL PERFORMANCE = 77.9

AVERAGE PERFORMANCE BY CLASS = 77.9

study. Therefore, direct comparison between the results of the two training groups should not be made. The test samples were from the same locations in both studies and no significant difference can be seen from the use of the wavelength bands from 1.5 to 1.8 micrometers in this classification.

Visible, Near Infrared and Thermal Infrared (0.4-2.6 and 8-14 micrometers)

Two visible and three infrared wavelength bands were selected for use in this classification by the $SELECT processor from the 0.4 to 2.6 and 8 to 14 micrometer range. This included the use of a thermal infrared wavelength from 8 to 14 micrometers. The overall performance on the training fields (3,352 RSU's) was 93.6 percent (Table 17); on test fields (5,258 RSU's) the performance was 86.4 percent (Table 18). When no attempt was made to discriminate between stubble (STUB), diverted acres (DA), and pasture (PAST) fields, the performance on training fields was 96.1 percent.

The field of view of the scanner used to collect the thermal infrared wavelength band data is considerably narrower than the field of view of the scanner used to collect the other wavelengths. Therefore, the set of training and test samples used in this classification were somewhat different than those used for the previous two analyses. For this reason, the results cannot be effectively compared. The results from all three studies are shown in Table 19. No definite conclusions regarding the value of the thermal infrared data can be drawn on the basis of this limited set of analysis results. When stubble, diverted acres, and pasture were combined

Table 13. Classification Summary by Training Classes for C-1 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	CORN
	ST1
	ST2
	DA1
	DA2
	PT1
	PT2
	RIVER
	TREES

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
4	0.72	0.80
7	0.55	0.58
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TRAINING CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOY	STUB	D.A.	PAST	CORN	RIVE	TREE	THRS
1	SOY	846	95.3	806	1	0	23	16	0	0	0
2	STUB	651	73.9	1	481	113	54	2	0	0	0
3	D.A.	297	71.0	5	39	211	38	4	0	0	0
4	PAST	582	82.6	20	30	38	481	13	0	0	0
5	CORN	1367	91.5	67	14	10	19	1251	0	6	0
6	RIVE	84	100.0	0	0	0	0	0	84	0	0
7	TREE	64	96.9	0	0	0	0	2	0	62	0
	TOTAL	3891		899	565	372	615	1288	84	68	0

OVERALL PERFORMANCE = 86.8

AVERAGE PERFORMANCE BY CLASS = 87.3

as a class, the performance of the training samples for all three analyses are very similar.

FLIGHTLINE PF25, JULY, 1968

Only scan lines 1 to 1,100 inclusive were analyzed for this study because of the severe limitations caused by cloud patterns over the length of the flightline. Because of the unavailability of suitable calibration procedures at present to eliminate this problem, no further analyses were made over the remainder of this flightline. The portions studied were analyzed using different combinations of wavelength bands in the visible, near infrared, and thermal infrared wavelength bands.

Visible and Near Infrared (0.4-1.0 micrometers)

Five wavelength bands were selected in the 0.4 to 1.0 micrometer wavelength range. Overall performance on training fields (2,307 RSU's) was 92.5 percent correct (Table 20); for test fields, (8,912 RSU's) the overall performance was 74.7 percent (Table 21). The three classes used in this study were soybeans, corn, and a mixture. The mixture class contained the stubble,

Table 14. Classification Summary of Test Classes for C-1 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	CORN
	ST1
	ST2
	DA1
	DA2
	PT1
	PT2
	RIVER
	TREES

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
4	0.72	0.80
7	0.55	0.58
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TEST CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOY	CORN	MIX	RIVE	TREE	THRS
1	SOY	2485	81.6	2028	120	233	0	104	0
2	CORN	2212	70.2	420	1553	227	0	12	0
3	MIX	2290	95.2	101	8	2181	0	0	0
4	RIVE	84	100.0	0	0	0	84	0	0
5	TREE	64	96.9	0	2	0	0	62	0
	TOTAL	7135		2549	1683	2641	84	178	0

OVERALL PERFORMANCE = 82.8

AVERAGE PERFORMANCE BY CLASS = 88.8

pasture, and hay crops. In reviewing the samples that were classified incorrectly, it can be seen that the soybeans were not only confused with corn but were also confused with the mixture category. This is partially explained by the fact that some soybeans had to be replanted due to flooding conditions. When the beans were replanted, they were drilled instead of being planted in rows. These late planted beans probably looked very much like hay crops, especially red clover. Also, it can be seen that the mixture category was equally confused between soybeans and corn.

Visible and Near Infrared (0.4-2.6 micrometers)

One visible and three infrared wavelength bands were selected from 0.4 to 2.6 micrometer range by the $SELECT processor for classification from a wider wavelength band region. The overall performance on training fields was 92.7 percent correct (Table 22); for test fields, overall performance was 73.7 percent (Table 23). These results did not differ appreciably from the previously obtained results. This would indicate that expanding the wavelength region into the infrared from 1.0 to 2.6 micrometers had no effect on increasing the accuracy of classification results.

Visible, Near Infrared, and Thermal Infrared (0.4-2.6 and 8-14 micrometers)

Table 15. Classification Summary by Training Classes, C-1 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	CN1
	CN2
	ST1
	ST2
	DA1
	DA2
	PT1
	PT2
	RIVER
	TREE

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
4	0.55	0.58
6	0.66	0.72
8	0.80	1.00
10	1.50	1.80

CLASSIFICATION SUMMARY BY TRAINING CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOY	CORN	STUB	D.A.	PAST	RIVE	TREE	THRS
1	SOY	846	94.2	797	14	1	2	32	0	0	0
2	CORN	1799	94.3	61	1696	20	4	16	0	2	0
3	STUB	639	89.2	0	5	570	52	12	0	0	0
4	D.A.	297	80.8	7	2	17	240	31	0	0	0
5	PAST	582	88.1	16	10	11	32	513	0	0	0
6	RIVE	84	100.0	0	0	0	0	0	84	0	0
7	TREE	56	100.0	0	0	0	0	0	0	56	0
	TOTAL	4303		881	1727	619	330	604	84	58	0

OVERALL PERFORMANCE = 91.9

AVERAGE PERFORMANCE BY CLASS = 92.4

Two visible wavelength bands, two near infrared wavelength bands, and the thermal infrared wavelengths were selected from 0.4 to 2.6 and 8 to 14 micrometer range to classify the data from the study area. For this classification, the overall performance on training fields (2,805 RSU's) was 94.2 percent correct (Table 24); on test fields (4,659 RSU's), the overall performance was 73.7 percent (Table 25).

The field of view of the scanner collecting the thermal infrared data was restricted as mentioned previously. Therefore, these results cannot be strictly compared with the previous results for this flightline. A summary of the results are given in Table 26. No definite conclusions concerning the value of the thermal infrared data can be assessed even when reviewing this data without the previous limitations.

THE "OPERATIONAL ANALYSIS" EXPERIMENT

In planning for the data flight of July 30, 1968, it was decided that a number of LARS personnel and the LARS data processing equipment would be dedicated to processing the newly received data on a simulated operational basis. This was to give an idea of the time required to obtain usable analysis

Table 16. Classification Summary of Test Classes, C-1 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	CN1
	CN2
	ST1
	ST2
	DA1
	DA2
	PT1
	PT2
	RIVER
	TREE

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
4	0.55	0.58
6	0.66	0.72
8	0.80	1.00
10	1.50	1.80

CLASSIFICATION SUMMARY BY TEST CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOY	CORN	MIX	RIVE	TREE	THRS
1	SOY	2485	67.1	1667	603	164	0	51	0
2	CORN	2212	90.1	108	1994	106	0	4	0
3	MIX	2290	95.0	15	100	2175	0	0	0
4	RIVE	84	100.0	0	0	0	84	0	0
5	TREE	64	98.4	0	1	0	0	63	0
	TOTAL	7135		1790	2698	2445	84	118	0

OVERALL PERFORMANCE = 83.9

AVERAGE PERFORMANCE BY CLASS = 90.1

results by means of the current LARS aircraft data processing system.

A test analog tape was received by LARS at noon on July 30 so that minor modifications in the LARS system could be made to accomodate recent changes made by the University of Michigan in the format of the scanner analog tape. Preliminary necessary adjustments to the LARS system were performed by 2:00 p.m. the same day. However, analysis later of the test tape data showed that the locations of the "C-1" and "C-2" calibration pulses on the analog tape had been changed; this change had not been reported. Necessary changes were made and all final checks on the data processing system were completed by 5:00 p.m., July 31.

The beginning time of the operational analysis experiment was 10:30 p.m., July 31, 1968, at which time the data tapes (duplicates of the original tapes) were received by LARS. Initial tests on the digitized data indicated significant skew errors in the analog data. These may have been introducted in the tape duplication process. The LARS aircraft data reformatting program was modified to minimize the skew problem. Digitization and reformatting of all data received for flight lines PF21,

Table 17. Classification Summary of Training Classes, C-1 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	CN1
	CN2
	ST1
	ST2
	DA1
	DA2
	PT1
	PT2
	RIVER

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
4	0.55	0.58
6	0.66	0.72
9	1.00	1.40
10	1.50	1.80
12	8.00	14.00

CLASSIFICATION SUMMARY BY TRAINING CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOYB	CORN	STUB	D.A.	PAST	RIVE	THRS
1	SOYB	846	94.7	801	22	3	1	19	0	0
2	CORN	1285	95.3	29	1224	17	7	8	0	0
3	STUB	564	92.6	0	6	522	22	14	0	0
4	D.A.	297	87.5	1	5	13	260	18	0	0
5	PAST	348	91.4	12	2	6	10	318	0	0
6	RIVE	12	100.0	0	0	0	0	0	12	0
	TOTAL	3352		843	1259	561	300	377	12	0

OVERALL PERFORMANCE = 93.6

AVERAGE PERFORMANCE BY CLASS = 93.6

PF25, PF35 were completed in six hours. An hour later, pictorial printouts for PF21 and PF25 flightlines, which were to be analyzed, were available. A total of 7 hours (approximately 12 man-hours) had been used for the data handling procedures.

The data were judged to be of good quality with the exception of the skew problem. A significant degree of variation in the overall illumination of the flightlines (see previous data analysis results section) was evident.

The most time consuming portion of the data analysis involved the coordination of the ground truth data with the aircraft scanner data. This process requires finding and outlining field and other feature boundaries on the pictorial printouts and tabulating and card-punching boundary coordinates for input to the analysis programs. For the sections of PF21 and PF25 flightlines to be analyzed, approximately 8 hours (32 man-hours) were required to go from pictorial output to punched card field boundaries. An additional 16 hours (32 man-hours) were required to obtain accurate data classification within a few percentage points of the final results, which were reported in the previous data analysis results section. Accuracy of the results was similar to that achieved previously with aircraft data analyzed at LARS.

Table 18. Classification Summary of Test Classes, C-1 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	CN1
	CN2
	ST1
	ST2
	DA1
	DA2
	PT1
	PT2
	RIVER

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
4	0.55	0.58
6	0.66	0.72
9	1.00	1.40
10	1.50	1.80
12	8.00	14.00

CLASSIFICATION SUMMARY BY TEST CLASSES

				NO OF SAMPLES CLASSIFIED INTO				
	CLASS	NO OF SAMPS	PCT. CORCT	SOY	CORN	MIX	RIVE	THRS
1	SOY	1816	68.6	1246	448	122	0	0
2	CORN	1756	94.0	28	1651	77	0	0
3	MIX	1674	97.6	11	29	1634	0	0
4	RIVE	12	100.0	0	0	0	12	0
	TOTAL	5258		1285	2128	1833	12	0

OVERALL PERFORMANCE = 86.4

AVERAGE PERFORMANCE BY CLASS = 90.1

In summary, the total clock time required to obtain these results was approximately 31 hours, corresponding roughly to 76 man-hours of effort. Approximately 6400 acres of total ground area were classified.

CONCLUSIONS

The data analysis experiments carried out during the period covered by this report aided significantly in defining a workable and systematic procedure for applying the currently available hardware and software at LARS to the analysis of multispectral aircraft scanner data. The $SELECT processor has proven a useful tool both for determining subclasses within the data and for selecting data channels to be used for classification. While the classification results obtained during this period have not improved dramatically over those reported earlier, this could have been expected since research has not been concentrated in this direction. General conclusions cannot be drawn regarding seasonal and yearly agricultural variability from the available quantity of results.

Efficiency in obtaining results has been substantially improved. This

Table 19. Comparison of Results for Area C-1, July 1966

Performance	Wavelength Bands: 3 Visible 1 Reflective IR	3 Visible 2 Reflective IR	2 Visible 2 Reflective IR 1 Thermal IR
Training	86.8	91.9	93.6
Test	82.8	83.9	86.4
"Combined" training[a/]	96.1	95.5	96.1

a/ Training samples of stubble, diverted acres, and pasture combined before classification.

Table 20. Classification Summary by Training Classes, PF25 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	SOY3
	CN1
	CN2
	CN3
	MIX1
	MIX2
	MIX3

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
4	0.72	0.80
7	0.55	0.58
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TRAINING CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO: SOY	CORN	MIX	THRS
1	SOY	810	90.2	731	13	66	0
2	CORN	973	94.0	47	915	11	0
3	MIX	524	93.3	32	3	489	0
	TOTAL	2307		810	931	566	0

OVERALL PERFORMANCE = 92.5

AVERAGE PERFORMANCE BY CLASS = 92.5

Table 21. Classification Summary of Test Classes, PF25 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	SOY3
	CN1
	CN2
	CN3
	MIX1
	MIX2
	MIX3

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
1	0.40	0.44
4	0.72	0.80
7	0.55	0.58
10	0.66	0.72
12	0.80	1.00

CLASSIFICATION SUMMARY BY TEST CLASSES

				NO OF SAMPLES CLASSIFIED INTO			
	CLASS	NO OF SAMPS	PCT. CORCT	SOY	CORN	MIX	THRS
1	SOY	4411	67.6	2980	814	617	0
2	CORN	3786	84.7	513	3207	66	0
3	MIX	715	65.3	123	125	467	0
	TOTAL	8912		3616	4146	1150	0

OVERALL PERFORMANCE = 74.7

AVERAGE PERFORMANCE BY CLASS = 72.5

Table 22. Classification Summary of Training Classes, PF25 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	SOY3
	CN1
	CN2
	CN3
	MIX1
	MIX2
	MIX3

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
2	0.62	0.66
9	1.00	1.40
10	1.50	1.80
11	2.00	2.60

CLASSIFICATION SUMMARY BY TRAINING CLASSES

				NO OF SAMPLES CLASSIFIED INTO			
	CLASS	NO OF SAMPS	PCT. CORCT	SOY	CORN	MIX	THRS
1	SOY	1936	92.8	1796	47	93	0
2	CORN	966	91.3	62	882	22	0
3	MIX	842	94.3	24	24	794	0
	TOTAL	3744		1882	953	909	0

OVERALL PERFORMANCE = 92.7

AVERAGE PERFORMANCE BY CLASS = 92.8

Table 23. Classification Summary of Test Classes, PF25 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	SOY3
	CN1
	CN2
	CN3
	MIX1
	MIX2
	MIX3

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
2	0.62	0.66
9	1.00	1.40
10	1.50	1.80
11	2.00	2.60

CLASSIFICATION SUMMARY BY TEST CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOY	CORN	MIX	THRS
1	SOY	3275	72.7	2380	428	467	0
2	CORN	3273	80.1	415	2622	236	0
3	MIX	686	48.1	157	199	330	0
	TOTAL	7234		2952	3249	1033	0

OVERALL PERFORMANCE = 73.7

AVERAGE PERFORMANCE BY CLASS = 67.0

Table 24. Classification Summary of Training Classes, PF25 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	SOY3
	CORN
	MIX1
	MIX2

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
4	0.55	0.58
6	0.66	0.72
9	1.00	1.40
11	2.00	2.60
12	8.00	14.00

CLASSIFICATION SUMMARY BY TRAINING CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOY	MIX	CORN	THRS
1	SOY	1235	97.3	1202	30	3	0
2	MIX	503	98.0	10	493	0	0
3	CORN	1112	88.9	98	25	989	0
	TOTAL	2850		1310	548	992	0

OVERALL PERFORMANCE = 94.2

AVERAGE PERFORMANCE BY CLASS = 94.8

Table 25. Classification Summary of Test Classes, PF25 Flightline, July, 1968

CLASSES CONSIDERED

SYMBOL	CLASS
	SOY1
	SOY2
	SOY3
	CORN
	MIX1
	MIX2

FEATURES CONSIDERED

CHANNEL NO.	SPECTRAL BAND	
4	0.55	0.58
6	0.66	0.72
9	1.00	1.40
11	2.00	2.60
12	8.00	14.00

CLASSIFICATION SUMMARY BY TEST CLASSES

	CLASS	NO OF SAMPS	PCT. CORCT	NO OF SAMPLES CLASSIFIED INTO SOY	CORN	MIX	THRS
1	SOY	2298	77.5	1780	186	332	0
2	CORN	1804	73.7	373	1329	102	0
3	MIX	557	58.0	71	163	323	0
	TOTAL	4659		2224	1678	757	0

OVERALL PERFORMANCE = 73.7

AVERAGE PERFORMANCE BY CLASS = 69.7

Table 26. Comparison of Results for Flightline PF25, July, 1968

	Wavelength Bands		
Performance	4 Visible 1 Near IR	1 Visible 3 Middle IR	2 Visible 2 Middle IR 1 Thermal IR
Training	92.5	92.7	94.2
Test	74.7	73.7	73.7

is essential if large volumes of remote sensing data are to be processed in a timely manner. The simulated operational analysis experiment indicates definite progress in this direction. As a result more extensive analysis experiments are anticipated in the future.

The data analysis efforts indicate that some important problems remain, however. Improved classification results are still a significant objective and one which will require research in several areas. New approaches are necessary in defining classes and subclasses on a multispectral basis and in developing and applying advanced pattern recognition techniques such as nonparametric methods. Such techniques may provide better statistical models for the data than the Gaussian (multivariate normal distribution) assumption now employed. Another problem is the relatively slow process of outlining and tabulating field boundaries for input to the analysis programs. This process greatly needs to be automated. Possibly the digital display system which has been proposed by LARS would aid to solving the problem. Some additional work on automating the analysis of the $SELECT processor output should yield further efficiencies in the overall data analysis procedure.

CHAPTER 8

SUMMARY AND CONCLUSIONS

APPLICATIONS AND REQUIREMENTS

Again during this work period efforts were expended to better define the information requirements of the various sectors of the agricultural complex. During this period special emphasis was given to defining the information needs of various members of agrobusiness. This included discussions with representatives of farm implement manufacturers, seed producers and fertilizer and chemical manufacturers. It was concluded that there is a general need for information and advisory services based on modern technology that will provide accurate and timely measurements of agricultural situations over relatively large geographic areas.

A general opinion was that to be useful such systems should provide two categories of information: (1) current information for short range planning, (2) information to facilitate long range planning and decision making. The need for information about the percent of agricultural land seeded at selected times in the spring is an example of information necessary for short range planning to assist in product transportation. An example of long range planning information would be data pertaining to drainage characteristics of poorly mapped agricultural land areas.

LARS personnel concluded that there is need for a general educational program of the capabilities of remote sensing for industry. Also, it was concluded that it would be extremely beneficial to establish a pilot experimental information and advisory service program operating over a geographic region. Such a pilot program would serve as (1) a basis for a demonstration of the techniques developed in research programs, (2) an experimental testing ground for research techniques, and (3) facilities for educational and training programs. Additionally, LARS believes that such a pilot program would be extremely valuable in providing guidelines for future research.

LARS plans to develop the description of such a program in cooperation with representatives of the agrobusiness community during the next work period. Such a program should be the collaberative effort of representatives of research, private industry and federal and state government personnel.

DATA PROCESSING

A number of projects of interest were conducted in Data Processing programs during this period. A project of major significance was instituted to develop a capability to automatically make congruent, resolution element by resolution element, the data gathered through different sensor aperatures. Such a capability is required before researchers can begin to use all the spectral channels currently available in the University of Michigan multispectral scanning sensor. Equally important is the requirement for this capability in order to investigate the utility of temporal variations in multispectral measurements. This would make possible the utilization of time histories of multispectral measurements in the identification of important earth resource conditions.

Further importance for this capability is established through the need to utilize multispectral data collected at widely different points in the electromagnetic spectrum and therefore with different sensors. With this capability, measurements at radar frequencies could be simultaneously processed with data collected in the visible portion of the spectrum.

LARS continued to concentrate on research aimed at establishing criteria for the selection of training samples and for the selection of separable categories of various agricultural scenes. Since it is the opinion of LARS personnel that future operational automatic data processing will be based on "at the time" training on the basis of data from ground truth areas, it is important that efficient training procedures utilizing minimum ground truth data be established.

LARS data processing programs concentrated on developing analysis procedures resulting in extremely short turn around time. The objective was to be able to analyze airborne data quickly enough to permit field checking before conditions were drastically altered. LARS demonstrated that a 25 linear mile flightline could be analyzed within 48 hours after data was made available.

To further facilitate its research capabilities, considerable efforts were devoted to the design of a visual digital display permitting a better interface between the researcher and the digital computer facility. A hardware design for this system was finalized in cooperation with assistance from various segments of industry. It is believed that the visual digital display will greatly facilitate the research at LARS and prove to be extremely valuable in future operational systems.

In summary, the data processing capability at LARS was improved to permit the processing of data collected over tens of square miles in an accurate and timely basis. LARS personnel believe that their program has reached a point where such a capability is essential to the conducting of meaningful research in future periods. This capability makes it extremely desirable to have magnetic tapes and photographic data as soon as they are obtained.

BIOGEOPHYSICAL

Biogeophysical research programs reached several milestones. A flightline 70 miles in length was flown from Indianapolis, Indiana to Bedford, Indiana. One of the objectives of this experiment was to investigate the effects of geographic distance on classification accuracies using training samples collected from a single point in the flightline. Training samples were selected for categories of vegetation, bare soil, and surface water. For these categories, the classification accuracies averaged better than 90 percent. Since instrumentation drift was not corrected for and variations in insulation energy were ignored, it is expected that categories more spectrally similar would have been noticeably affected. These effects will be considered in future research projects. LARS considers the automatic processing of data over such a large geographic area to be a major milestone in its program.

The identification of important agricultural species continued to be a major objective of the Biogeophysical programs. A four mile by one mile area of row crops was correctly classified to an accuracy of 92.4 percent. The overall performance of 89.7 percent was realized with 27,000 samples of which 15,000 samples were row crops.

Soil scientists were able to distinguish up to six distinct soil categories based on their spectral characteristics. This was accomplished with measurements collected over areas free of any great amount of surface cover. The spectral radiation characteristics of surface soil were seen to vary with different moisture levels. Results tended to indicate that total reflectance is more drastically changed than is relative reflectance within the 4000 angstrom to 10,000 angstrom region. Further research is planned to relate surface soil moisture with spectral response patterns.

An additional milestone for LARS was the analysis of data collected over widely different geographic areas. The "Moon Lake" area in Texas was analyzed with 67 percent accuracy for categories of bare soil, water and trees. Similarly data collected over California was classified into categories of soil, immature rice, safire, and water with a 98 percent accuracy for training samples. Ground truth at hand at the time of this analysis was not adequate for more extensive investigations. However, LARS is of the opinion that results obtained over Indiana are generally representative of results that should be expected within other geographic regions.

Laboratory studies of leaf and soil radiance characteristics continued to be an important part of the LARS program. The effects of pigments on reflective characteristics, the effects of leaf moisture on leaf reflectance and the effects of different moisture contents on soil reflectances were studied. The spectral characteristics of sand, silt and clay were studied at various moisture levels.

However, LARS believes strongly that there continues to be a great need for field spectrometry in its research. A continuing absence of an adequate field spectrometer prevents the collection of insitu measurements of the macroscopic affects of plants and their backgrounds. Biogeophysical research programs will continue to work with measurements researchers to define desirable characteristics of a future field instrument. It is hoped that such an instrument covering the spectral range from some 4,000 angstroms to 15 microns can be developed by industry for the research program.

It is believed that such an instrument will be found to be an extremely useful instrument in many different situations - research and operational.

MEASUREMENTS

Measurements programs continued their efforts to develop equipment and procedures for collecting radiance measurements in the laboratory and field. In the absence of funding for the development of a field spectrometer by industry, LARS engineered an in-house project aimed at converting a Perkin Elmer model 98 monochrometer into a 3 channel field spectrometer. Three spectral bands between 4,000 angstroms and 2,000 nanometers is a design goal for this instrument. The three bands are to be alterable in the laboratory between field experiments. In future research, it is expected that this instrument will permit the collection of more quantitative measurements of the radiation characteristics of vegetation under stress. These measurements will be compared to those collected with color infrared film and filtered black and white photography.

Graduate study projects culminated in the development of a laboratory sensor to permit the investigation of the geometric characteristics of the radiation patterns of plants in a more natural state. In this study the radiation characteristics of potted plants were measured.

FLIGHT MISSIONS

A certain number of missions were conducted over the Purdue test site by other NASA aircraft. Photographic data from these missions were utilized to study the utility of different film filter types to identify crop species, crop conditions, soil types at different scale factors. This research greatly complimented the automatic recognition research utilizing multispectral measurements.

LARS in the future hopes to investigate the utilization of photographic emulsions as to measure multispectral energy. One of the objectives of the research with photography collected by NASA was to develop techniques for selecting training samples. This involves identification of a universal set of categories from a small amount of photography. Multispectral measurements from areas so specified can be utilized to train the computer to automatically classify a majority of the data.

LARS recommended that certain future experiments be conducted from space. It was recommended the multispectral measurements be collected from primary earth surface coverings which naturally occur in large area sizes. This would include jungle vegetation, range land, sand, fresh water, salt water, and so forth. In addition, it was recommended that spectral measurements of wheat be collected where this crop covers large areas of farmland. Four such regions were recommended for future experiments.

SUMMARY

Generally, the research capabilities of LARS were improved during this period and certain milestones were reached in the application of remote sensing to agricultural needs. Plans have been completed for future missions to be flown routinely over a 20 percent area sample of the 500 square mile Tippecanoe County test site. Future research is to identify the effects of such geographic size on automatic recognition capabilities.

LARS looks forward to developing a capability adequate to support NASA's first Earth Resources Technology Satellite Programs and Skylab.