Annual Technical Summary Report

NASA Contract NAS9—14970

June 1, 1976 - May 31, 1977

D. A. Landgrebe, Purdue University

Principal Investigator

J. D. Erickson, NASA/JSC

Technical Monitor

Submitted by

The Laboratory for Applications of Remote Sensing

Purdue University   West Lafayette, Indiana

1977

NAS9-14970
T-1314/5
MA-657T

CR-151451

Annual Technical Summary Report

(E77-10183) [APPLICATIONS OF REMOTE
SENSING] Annual Technical Summary Report, 1
Jun. 1976 - 31 May 1977 (Purdue Univ.) 61 p
HC A04/MF A01   CSCL 05B

N77-33553

G3/43

Unclas
00183

NASA Contract NAS9-14970

June 1, 1976 - May 31, 1977

D. A. Landgrebe, Purdue University
Principal Investigator

J. D. Erickson, NASA/JSC
Technical Monitor

Submitted by

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

1977

Preface

Contract NAS9-14970 contained a series of twelve tasks on related problems. They varied greatly in the magnitude of effort required. This report provides a brief summary on the results obtained in each.

Figures are included where it is possible to display and example results graphically. Figures are not necessarily referred to in the text.

A complete report of results obtained in each task is contained in the Final Technical Report, Volumes I, II and III.

TABLE OF CONTENTS

TECHNICAL SUMMARIES

2.1 TEST OF BOUNDARY FINDING/PERFIELD CLASSIFICATION

One might wish to adopt a new classification technique for one of three reasons:

- It provides a greater accuracy than the currently implemented technique.
- It is less expensive than the current technique.
- The results it produces are easier to use and interpret than those currently produced.

The classification algorithms discussed in this report have shown promise in all three areas. They are referred to by the acronym ECHO* and incorporate a combination of boundary finding and perfield classification.

Contemporary classifiers catagorize remotely sensed data by comparing the spectral measurements from each feature of each point to class statistics, computing a likelihood or discriminant function value associated with each class, and assigning the point to the class with the largest discriminant function value. Each point is classified individually on the basis of its spectral measurements alone. One premise of this technique is that the objects of interest are large in comparison to the size of a point. If this were not so, a large proportion of points would be composites of several classes, making statistical pattern classification unreliable since pre-specified categories would be inadequate to describe actual states of nature. From this premise it follows that objects are represented by arrays of points, and that a statistical dependence exists between consecutive points. Contemporary classifiers fail to exploit this statistical dependence between adjacent points when assigning classes.
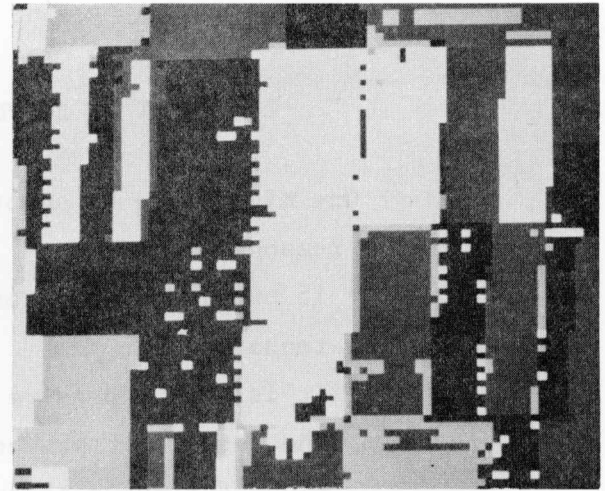
The ECHO processors benefit from spatial information by first aggregating into groups points whose spectral responses are not significantly different in a statistical sense, and then applying a maximum likelihood classification rule to these homogeneous groups. Homogeneous

---

*ECHO stands for Extraction and Classification of Homogeneous Objects.
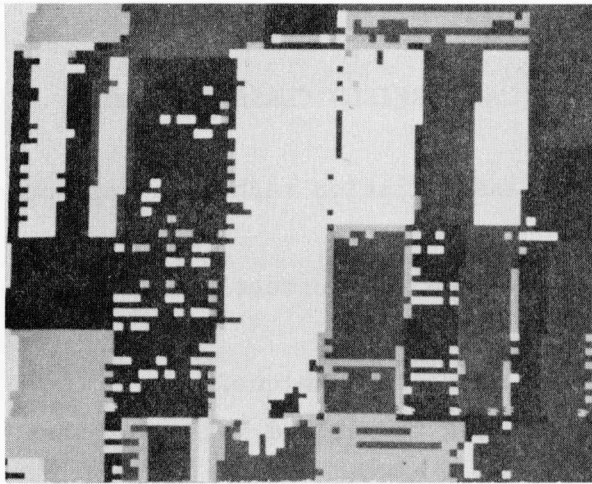
	Perpoint		Low Homogeneity Threshold
	87.5%	Field Center Pixel Performance	90.0%
	7.3%	RMS Proportion Error	2.5%
	1200 Seconds	Computation Time	754 Seconds
	34.8%	Classification Variability	29.9%
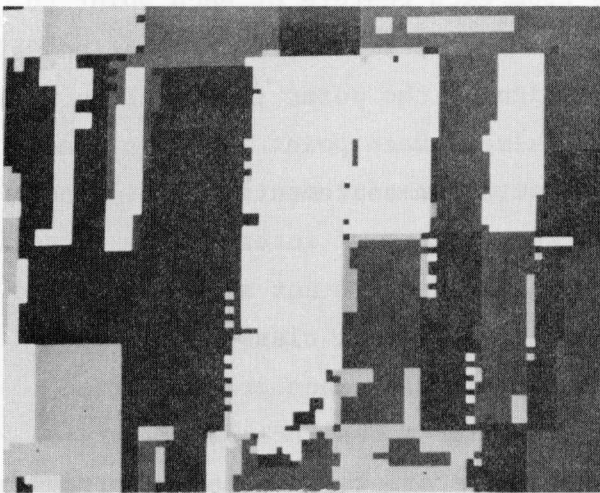
	Medium Homogeneity Threshold		Large Homogeneity Threshold
	91.4%	Field Center Pixel Performance	90.7%
	2.5%	RMS Proportion Error	2.8%
	494 Seconds	Computation Time	370 Seconds
	22.5%	Classification Variability	18.6%
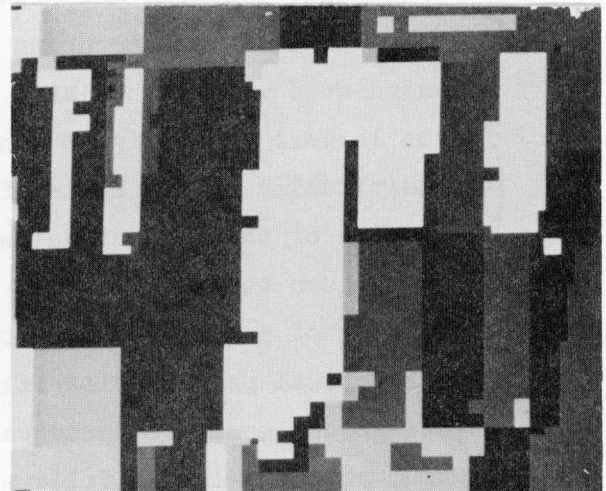
Figure 2.1-1 This figure shows a comparison between a conventional perpoint classifier and ECHO for several different cell homogeneity thresholds. The images show only a small portion of a complete test area over which the performance figures were determined.
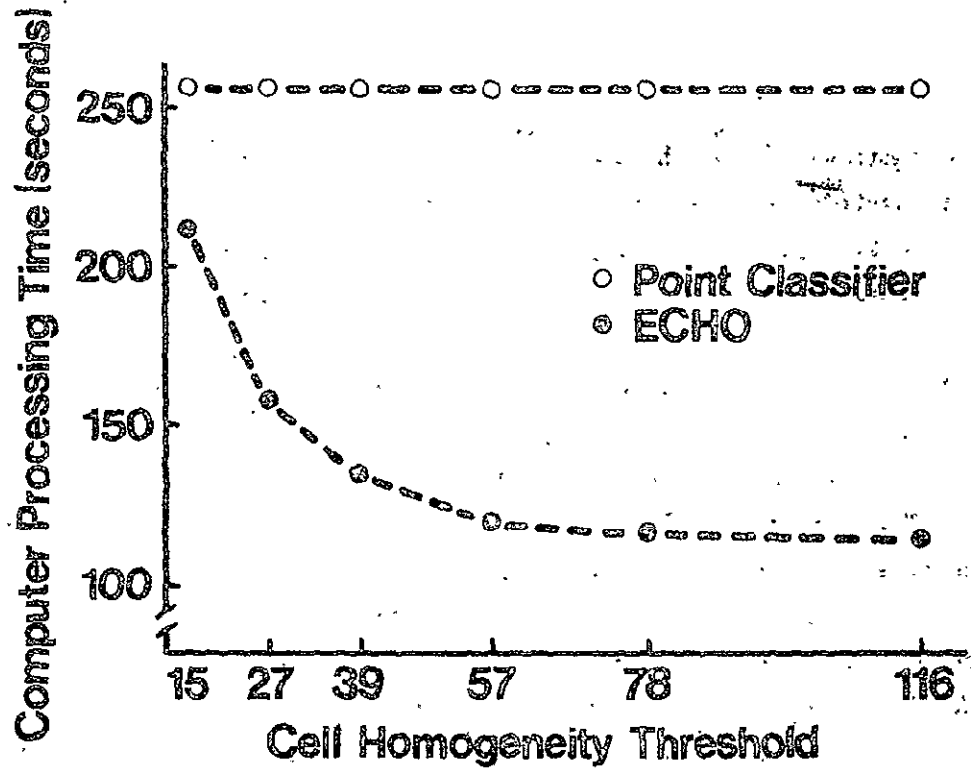
2.1-2. This figure shows one example of how the processing time for ECHO varies with the cell homogeneity threshold. It also provides a comparison of ECHO with the conventional point classifier. These results were obtained on Landsat data from Kansas. Results with multispectral aircraft scanner data showed similar trends.

objects are identified in a three step process. First, cells are formed
by systematically partitioning the data into N by N sized blocks of
pixels. The statistics of each cell are then compared to a homogeneity
criterion. Points which do not comprise homogeneous groups are classi-
fied on a point-by-point basis, just as contemporary classifiers classify
all points. Statistics of adjoining homogeneous cells are then compared
to annexation thresholds. Adjoining cells which appear to belong to the
same statistical population are combined into a single object.

Two separate ECHO algorithms have been developed at Purdue/LARS[1].
The first, Supervised ECHO, makes use of pre-specified class statistics
to identify homogeneous objects. The second, Nonsupervised ECHO, iden-
tifies homogeneous objects without the use of class statistics. Conse-
quently, those objects identified by the Nonsupervised algorithm may be
used to aid in the training process or for other processing tasks not
directly associated with classifications.

Objectives of the 1976-1977 Task

The objectives of the FY77 work were to:

1) Deliver to NASA documented FORTRAN programs implementing the
ECHO algorithms, and to make improvements to those algorithms, where
appropriate.

2) Systematically test the algorithms on MSS data for agricultural
regions as observed by the LANDSAT satellites, aircraft scanners, and on
the simulated Thematic Mapper data.

3) Provide products enabling the determination of the utility of
the object maps to a LACIE Analyst Interpreter in the selection and
labeling of training fields.

A. Documented FORTRAN Programs

Program listings and program abstracts for the 2200 lines of code
for the Supervised ECHO processor were prepared. Certain reorganization
of the Supervised routine was accomplished to allow for both computa-
tional efficiency and user convenience.

The Nonsupervised processor was reprogrammed in a two phase approach
allowing the analyst to utilize the object map produced by the first
phase of the approach as a training aid. The second phase catagorizes

the homogeneous objects identified by the first phase in conjunction with user specified class statistics. One important characteristic of the nonsupervised processor is that it identifies homogeneous objects and calculates the mean and covariance matrices for each homogeneous objects prior to specification of class-statistics. It therefore can be used as a spectral-spatial clustering algorithm.

B. Tests of the ECHO Algorithm

The ECHO algorithm was tested over 10 LANDSAT, 8 simulated Thematic Mapper and 3 aircraft data sets. LANDSAT data sets were drawn from the 1974 LACIE/SRS test sites and the CITARS data sets. Simulated Thematic Mapper and aircraft data sets were selected from within those data sets with appropriate ground truth which were available in the Purdue/LARS data library.

Six variables were chosen as measures for evaluation of the ECHO classification results.

C. Results

The results of the tests on the Supervised ECHO processor indicated that the results of the ECHO processor at the optimum settings for object identification were more accurate than the results generated by a perpoint classifier, that ECHO required less computer time to perform the classification, and that ECHO results contained less "salt and pepper" than the perpoint results.

In addition, the effects on the six evaluation variables of varying the ECHO partitioning parameters (cell width, cell homogeneity threshold, and cell to field annexation parameter) were recorded for each of the 21 data sets separately and for the data sets combined into groups by scanner type (LANDSAT, Thematic Mapper and aircraft). On both a data-set-to-data-set and a scanner-type basis, the three partitioning parameters had statistically significant effects.

D. Provide Products for LACIE/AI Evaluation.

Magnetic Tapes were chosen as a means of transmitting the object maps produced by the Nonsupervised processor to NASA/JSC for evaluation. The tape format was the NASA/JSC Universal format.

E. Current Status

The Supervised ECHO processor and associated documentation are transferred to NASA/JSC in the Final Technical Report. The Nonsupervised processor has been extensively revised but is not yet ready for delivery to JSC. Results of the evaluations of the Supervised processor are included in Final Technical Report.

Future work to develop and assess training techniques exploiting the spectral-spatial clustering characteristics of the Nonsupervised process or would be desirable, because it may eventually prove to be of great benefit, reducing the amount of analyst time required for training and increasing the accuracy of the training statistics.

REFERENCE

1. Kettig, R. L. 1975. Computer Classification of Remote Sensed Multispectral Image Data by Extraction and Classification of Homogeneous Objects. Ph.D. Thesis, School of Electrical Engineering, Purdue University, West Lafayette, Indiana.

2.2 STRATIFICATION OF SCENE CHARACTERISTICS

Task 2.2, Stratification of Scene Characteristics, contains three subtasks: 2.2a. Stratification by Machine Clustering, 2.2b. Digitization and Registration of Ancillary Data, and 2.2c. Crop Inventory Using Full-Frame Classification.

2.2a Stratification by Machine Clustering

The work on the stratification by machine clustering task has progressed in two stages in the past year. First, from June 1976 to January 1977, previous work at LARS was evaluated and a new approach to the scene stratification problem was formulated. Secondly, from January 1977 to May 1977, the new approach was implemented and investigated on a trial basis. A brief summary of these activities is as follows:

During FY76 a straightforward machine stratification algorithm which utilized spectral characteristics directly observable in the imagery had been implemented and tested. In the so called dynamic stratifications produced, when the partitioning procedure placed two segments in different strata, this assignment was generally correct; however, when segments were assigned to the same stratum by the partitioning procedures, they did not always seem to come from the same stratum. We concluded that the clustering methods used in FY76 did not successfully partition the scene into regions sufficiently homogeneous to assure adequate non-local classification performance. While these results did not rule out the idea of using machine clustering methods for dynamic stratification, they suggested that more knowledge of the nature of stratification is needed if an effective machine method is to be developed.

A more intensive investigation into the fundamental nature of apparent strata in LANDSAT imagery was deemed essential to the development of the quantitive measures needed for an effective machine partitioning method. It has been observed that an experienced data analyst develops an ability to visually stratify a scene. However, it had not been determined whether this ability is sufficiently precise for achieving the necessary classification accuracy. It was suspected that

the analyst utilizes scene attributes that are both spectral and spatial in nature, but is was not clear how these attributes should be quantified and applied in dynamic stratification. Therefore, an hypothesis was posed that the strata an analyst determines from imagery are related to classification accuracy; this hypothesis was to be tested and became the basis for the new approach to the scene stratification problem taken at LARS. Essentially the quantitative properties of analyst-defined strata were investigated to determine if a quantitative measure could be found to define strata.

In seeking support for the above hypothesis, a new approach was formulated and implemented. Using this approach, partitions defined by three Purdue analysts on two frames of LANDSAT data covering a portion of the 1975-1976 LACIE segments, and these were then compared with the static partitions developed by the University of California/Berkeley (UCB), and the partitions used by segment selection in LACIE Phase III[3]. Three image enhancement algorithms were implemented and applied to the two LANDSAT scenes to see if the enhanced data aided the analysts in defining partitions. Attributes of the partitions were measured, including spectral response and class structure. The partitions were evaluated with respect to these attributes and long-term climatic variables.

It would be unwise to draw any permanent conclusions about the whole matter of dynamic scene stratification since only limited tests on actual data and limited efforts to establish a sound basis to other circumstances were a part of this task. However, the following observations on the results obtained are important in determining what the next steps relative to scene stratification should be.

A. The three stratification approaches produced strata varying in the degree of detail (average area per stratum) in the following order:

   1) LACIE Phase III partitions-relatively more general.

   2) Dynamic strata generated at Purdue-intermediate in detail.

   3) Static strata generated at UCB-the most detailed.
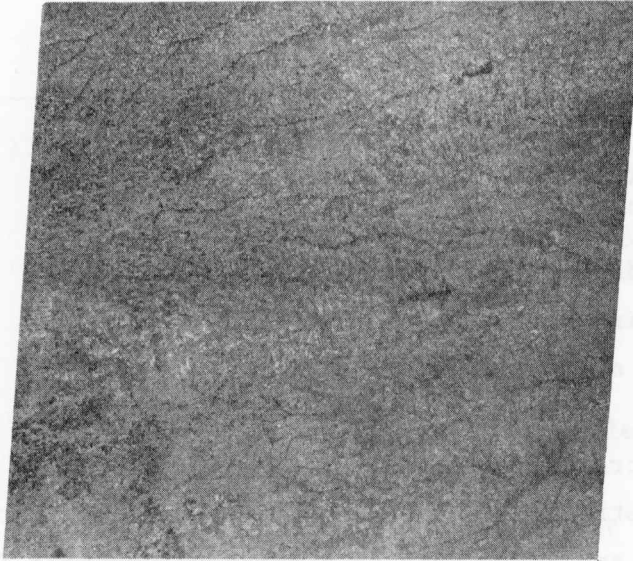
Landsat scene 5428-16053, band 5,
of Southwest Kansas on June 20,
1976, shown for reference.
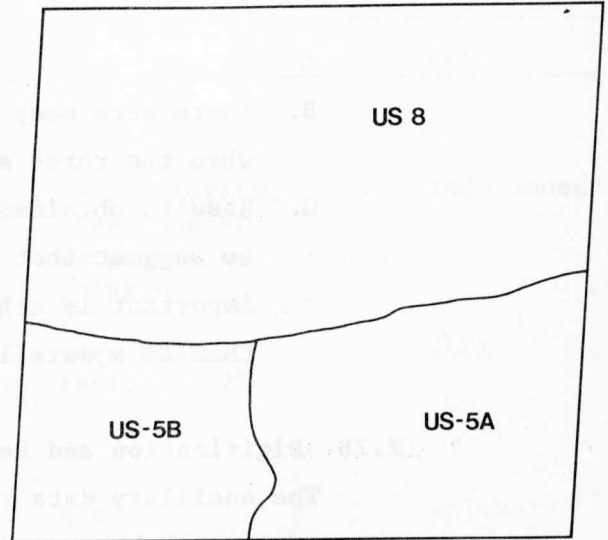
LACIE Phase III partitioning
of the reference scene.
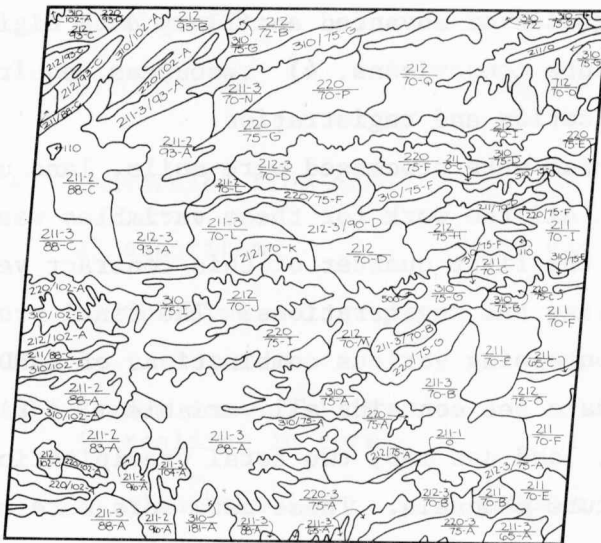
Static factors partitioning of
the reference scene by UCB.
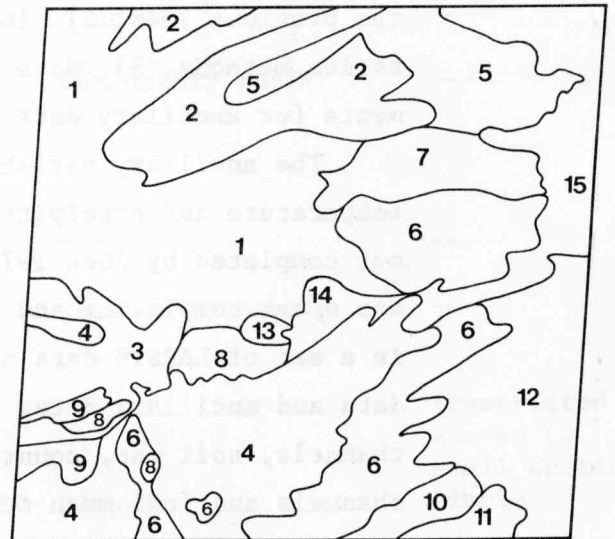
Dynamic factors partitioning of
the reference scene based upon
Tasselled Cap transformed data.

Figure 2.2a-1. This figure shows examples of three different stratifications
of a Landsat data set. Note that they vary in detail and are not always
consistent with one another at all points (See Table 2.2a-1); the indications
are that the sampling plan and training procedures used are more dominant
in determining classification performance than is stratification.

B. There were many points of agreement and of disagreement
    when the three stratifications were cross-compared.

C. Results obtained in subtask 2.2c and reference 1 tend
    to suggest that training procedures may be much more
    important in achieving precise, accurate classifications
    than is a detailed, accurate scene partitioning.

2.2b Digitization and Registration of Ancillary Data

The ancillary data registration task was established to support
strata research by providing ancillary data as additional channels in
remote sensing data sets. Several tasks were pursued during the year in
support various aspects of ancillary data handling. Four primary tasks
were defined: 1) completion of ancillary data registrations begun in
the previous year, 2) investigation of advanced ancillary data digiti-
zation methods, 3) data structure comparisons, 4) resources require-
ments for ancillary data digitization and registration.

The ancillary variables which were processed were soils, land use,
temperature and precipitation. All the work for these variables was
not completed by June 1976 and the first quarter of this contract year
was spent completing and verifying the registrations. The final product
is a set of LARSYS data sets containing various combinations of LANDSAT
data and ancillary data. One data set contains all variables: 4 LANDSAT
channels, soil map, county map, land use map, two total precipitation
channels and four mean temperature channels. These data sets were used
by strata researchers to correlate spectral strata with ancillary
variables.

The advanced digitization method task developed a film scanner
based method of digitizing certain color polygon maps. The map is
photographed, scanned, and digitized through blue, green, and red filters
producing a three channel data set. LARSYS classification techniques are
then applied to reproduce the polygons. Due to noise problems in scan-
ning and classification, errors occur in the result and an error correc-
tion scheme was implemented. Evaluation of a test case shows very close
agreement between the original map and the scanner digitized version.

		Polygon	Line-Compression	Grid
1)	Space efficiency			
	a) general dependence	on complexity and size of polygons	on complexity, size, and arrangement of polygons	independent
	b) efficiency for large area	extremely high typical less than 1%	very high typical 1-2%	100%
2)	Possibility of residing in core memory	excellent	good	unlikely
3)	Conversion to grid storage			
	a) algorithm	complex	simple	-----
	b) CPU time	lengthy	reasonably short	-----
4)	Conversion from grid			
	a) algorithm	not tested	simple	-----
	b) CPU time	not tested	reasonably short	-----
5)	Random accessibility if residing in core			(restricted access to small sections)
	a) algorithm	complex	simple	simple
	b) CPU time	lengthy	reasonably fast	fast

Table 2.2b-1 This figure provides a comparison of storage methods of physical maps after digitization. The three methods compared are: storage of polygon information directly; storage in pixel by form; and storage after using a simple data compression scheme on lines of data has been used.
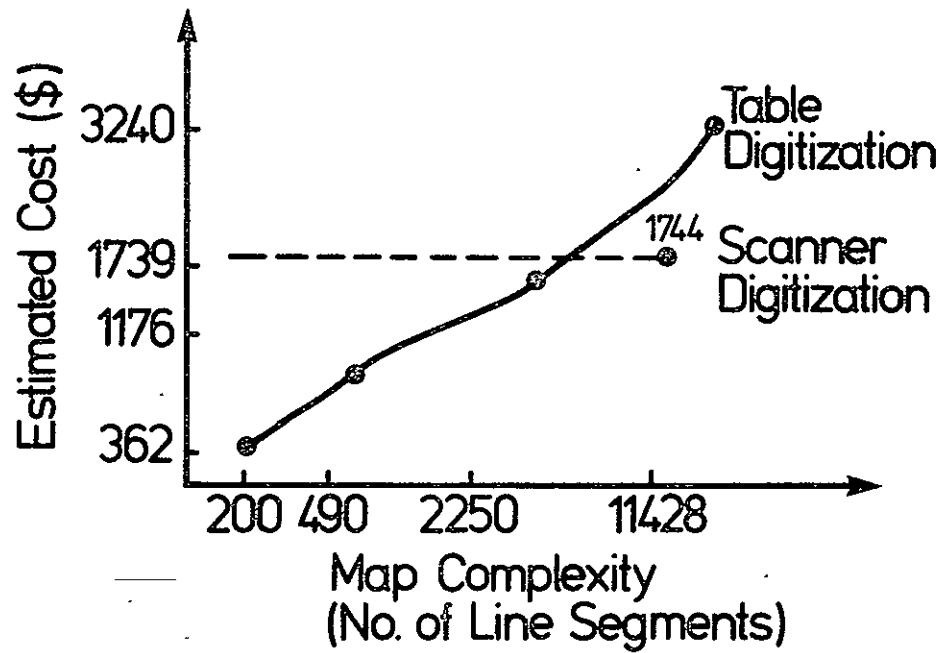
Figure 2.2b-1 This figure provides one illustration of how the costs of
table digitization and scanner digitization of ancillary data in map form
vary as the complexity of the map increases.

The last two tasks concern data structures and costs. A method of storing ancillary data using line compression is compared with polygon and grid formats. The compressed line format is shown to be a good compromise between polygon (efficient but difficult to randomly access) and grid (easy to access but costly) formats. Resource comparisons were made between manually digitized and registered maps and the automatic scanner digitization method. It is shown that for complex polygon maps the scanner method is more efficient whereas for simpler maps the manual method is cheaper. The costs are for one test case; however, the result does provide an effective comparison of methods. The complexity measure used was total number of line segments in the map.

2.2c Crop Inventory Using Full-Frame Classification

Wheat area estimates have been made by LACIE for the state of Kansas utilizing a sample segment segment approach. The procedure used has been to select 84 5x6 nm segments in Kansas, allocated to the counties in proportion to their historical area of wheat, and to classify the segments to make county estimates of the area of wheat which are then aggregated to make crop reporting district and state estimates. Studies need to be made to determine the accuracy and precision of the LACIE method of estimation and compare it to alternatives. The objective of this task was to compare the sampling errors associated with crop inventory methods using full-frame classification and the LACIE approach of sample segment classification.

Full-frame classifications of 80 counties in Kansas into wheat and non-wheat made in another investigation were used in this study[1]. The full-frame classifications were considered to have negligible sampling error and were repetitively sampled to simulate alternative sampling plans.

The experimental design consisted of three parts: 1) comparison of sampling error achieved by using a fixed total sample size (number of pixels), but varying the segment size and number of segments, 2) calculation of the number of segments of a given size required to achieve the sampling error of the LACIE-like allocation of 84 5x6 nm segments, and 3) comparison of the area estimates achieved by all sampling methods.

Part One. LACIE procedures for determining the allocation
(number) and location (geographic placement) of segments were followed
whenever possible and applicable [4]. The selection of sample segments
was computer-implemented, permitting a large number of segments to
be chosen with little personnel time, facilitating choice of any seg-
ment size or number of segments, and permitting statistical tests
of precision.

Four segment sizes were used: 5x6 nm, 4x4 nm, 2x2 nm, and individual
pixels. Eight replications were drawn for each sampling plan and esti-
mates of the wheat area were calculated [5]. Figure 2.2c-1 illustrates
the variability of each plan, showing that estimates obtained using
smaller segment sizes are more precise.

Two sampling errors, each based on a subset of four replications,
were computed for each plan. Statistical analyses showed that the stan-
dard deviations of the estimates differed at the 8% level of signifi-
cance and that the 5x6 nm segments had a significantly higher variance
than any of the other schemes.

Part Two. The objective of this part of the analysis was to cal-
culate the number of segments of each segment size required to achieve
the sampling error found from the 5x6 nm segments, assuming stratified
sampling with each county as a stratum [2]. The number of 5x6 nm seg-
ments which should have been required to achieve the fixed error was
calculated as a check; the result of this check was unreasonable and
indicated that the assumptions necessary for this calculation appar-
ently were not met. This approach was not possible because insuffic-
ient 5x6 nm segments were taken in a single replication to calculate
the variance of the area estimates and the restrictions on location of
segments prohibit the use of replications to estimate the variance.
Furthermore, it is not clear that these restrictions could be fully
taken into account in any theoretical calculation. Thus, this part was
not pursued further.

Part Three. Analysis of variance showed a difference at the 5%
level of significance among the area estimates found using the various

Figure 2.2c-1 This figure compares the wheat area estimates (in hectares) under different sampling schemes. The size of the sample segments was varied inversely with the number of segments in such a way that the total number of pixels used was constant. The results show a very large variance and a bias when the segment size gets as large as 5x6 nautical miles.

sampling plans. Multiple range tests showed the estimates obtained by sampling 5x6 nm segments were different from the other estimates; the other area estimates did not differ from one another.

The results of this investigation are well illustrated in Figure 2.2c-1. The area estimates found by the use of 5x6 nm segments cover a much larger range of values and thus have a larger variability than any of the other segment sizes. The estimates become more and more precise as the segment size decreases and more segments are taken. The estimates obtained from the 5x6 nm segments were different from the estimates obtained by the other segment sizes, indicating a bias in the 5x6 nm sampling scheme.

The estimates achieved using the 5x6 nm segments have the least precision of any sampling scheme tested. However, to assess the implications of this result for the LACIE sampling scheme, additional factors must be considered. In order to fully evaluate the scheme, the method of training and classification which would be used in conjunction with a sampling plan must be considered. Also, even though the precision of choosing more but smaller segments may be higher, this gain in precision must be weighed against the costs of sample selection and classification.

<div align="center">REFERENCES</div>

1. Bauer, M.E. and Staff. Crop Identification and Area Estimation over Large Geographic Areas Using Landsat MSS Data. Final Report of NASA Contract NAS5-20793, LARS Technical Report 012477. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana.

2. Cochran, W.G. Sampling Techniques. Second Edition. John Wiley and Sons, Inc., New York, 1963, p. 96.

3. Hay, C.M., and R.W. Thomas. Application of Photointerpretative Techniques to Wheat Identification, Signature Extension, and Sampling Strategy. Space Sciences Laboratory Series 17, Issue 33. University of California at Berkeley. May, 1976.

4. LACIE-00200, Volume IV: Crop Assessment Subsystem (CAS) Requirements, Level III Baseline (Reference CCBD # III-0001, Dec. 16, 1974), pp. B1-B6.

5. MacDonald, R.B., F.G. Hall, and R.B. Erb. "The Use of Landsat Data in a Large Area Crop Inventory Experiment (LACIE)." Proceedings, Symposium on Machine Processing of Remotely Sensed Data, June 3-5, 1975, Purdue University, West Lafayette, Indiana, pp. 1B-10 through 1B-11. (IEEE Catalog No. 75 CH 1009-9-C).

2.3 LACIE FIELD MEASUREMENTS FOR REMOTE SENSING OF WHEAT

The overall objective of the LACIE Field Measurements project is to acquire, process, and distribute to researchers fully annotated and calibrated multitemporal sets of spectral measurements along with agronomic and meteorological data which will serve as a data base for investigations in (1) quantitatively determining the temporal-spectral charac-teristics of wheat and other crops, (2) defining future sensor systems, and (3) developing advanced data analysis techniques. Test sites for the project are located in Finney County, Kansas; Hand County, South Dakota; and Williams County, North Dakota. The data collection activities are summarized in Figure 2.3-1.

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

The LACIE Field Measurements data base is undoubtedly the largest of its type now available for research purposes. It is unique in its comprehensiveness in terms of sensors and missions over the same sites throughout the growing season. The calibration of all multispectral data to a common standard is also unique. Finally, the kind and quantity of supporting agronomic and meteorological data is uncommonly great compared to most remote sensing experiments.

Purdue/LARS is a focal point for the project with responsibility for the following tasks: technical coordination and leadership, data acquisition, data processing and reformatting, data evaluation and verification, maintenance of data library and data distribution, and data analysis. These tasks are described further in the following sections.

Technical Coordination

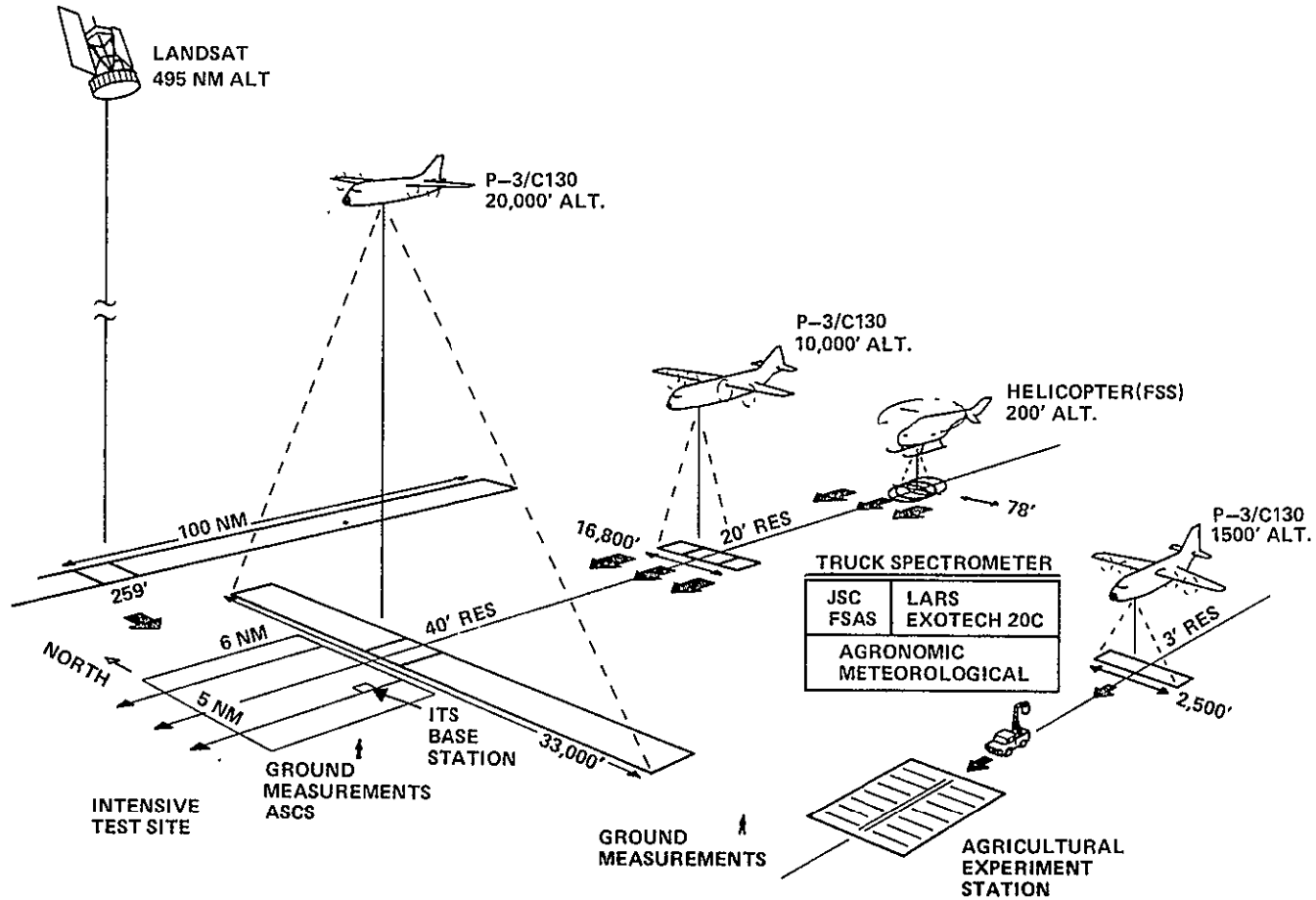A very important activity of this phase was preparation of the

Figure 2.3-1 The Figure provides a pictorial synopsis of data gathered as a part of the LACIE Field Measurements Project.

project plan for 1976-77. However, there is communication between NASA/JSC and Purdue/LARS and the other participating organizational units on all phases of the project on a continuous basis. In addition, LARS personnel have participated in several project meetings at NASA/JSC and Purdue/LARS.

Data Acquisition·

Four two-week data acquisition missions were performed by Purdue/ LARS at Williston, North Dakota between May 25 and August 31. The missions covered the planting-premergence to mature-harvested stages of spring wheat development. Data were collected using the Purdue/LARS spectroradiometer system for: (1) 60 spring wheat and other small grain plots, plus 8 plots of other crops and cover types, (2) a study of the reflective and emissive characteristics of wheat canopies as a function of view and solar angle, (3) a study of the polarization characteristics of wheat canopies as a function of view and solar angles, (4) measurement of the reflectance of canvas panels used for calibration of the FSS and aircraft scanner data. In addition, data were periodically obtained with a tripod-mounted radiometer in support of LACIE/SR&T wheat canopy modeling studies.

Detailed agronomic data including leaf area index, dry matter production of leaves, stems, and heads, soil moisture, and grain yield were obtained. Supporting meteorological measurements included: cloud cover and type, barometric pressure, total irradiance, wind speed and direction, air temperature, and humidity.

Data Processing and Reformatting

The objective of this subtask is to process and reformat the data so that it will be available to analysts in standard formats and units. Calibration to bidirectional reflectance is a key element of the acquisition and processing of all spectrometer data. These same units can also be applied to the aircraft and Landsat scanner data.

During this contract year the processing of all 1974-75 data was completed. Processing of 1975-76 data is 90 percent complete. Processing of 1976-77 data acquired in the fall of 1976 is currently underway. The status of data processing is shown in Table 2.3-1.

Instrument/Data Type	1974–75 Data Completed & In Library	1975–76 Data Completed & In Library	In Processing	At JSC	1976–77 Data Completed & In Library	In Processing	At JSC
Landsat MSS							
Whole Frame CCT (Frames)	20 Frames	61	0	N.A.	14	0	N.A.
Aircraft Scanner (Dates/Runs)	19/149	16/97	0	1	2/10	1/	2/
Helicopter Mounted Field Spectrometer (Dates/Runs)							
Field Averages	19/2,251	27/2,193	0		0	4/	1/
Individual Scans	19/35,000	27,40,000	0		0	4/	1/
Truck Mounted Field Spectrometer (Dates/Runs)							
FSAS	6/65	23/322	0		0	0	3/
Exotech 20C	24/1,577	14/1,307	5/1,361		–	–	–
Exotech 20D	45/645	–	–		–	–	–

Table 2.3-1 This Table shows the amount of data gathered, by instrument and the data library status as of May 31, 1977. Data catalogues are available which provide specific details on the data in the library.

Data Evaluation and Verification

The objective of this task is to understand and document the quality of the spectral measurements. It is particularly important that the data are calibrated so that meaningful comparisons can be made among sensors, sites, and dates. Data evaluation begins with sensor checks made before going to the field and then following defined procedures during all data acquisition. During data processing checks are routinely made of data quality. In the case of scanner data these checks include review of illumination conditions during data collection, review of histograms of each wavelength band, and review of imagery. Evaluation and verification of spectrometer data includes examination of illumination conditions at time of data collection, a calibration uniformity test, and examination of sample spectra. Reports of data quality will aid analysts in selecting data for analysis and aid in interpretation of results.

Additional verification of data has been conducted as part of Task 2.10, Supersite Data Management and will be described there. In May 1977 all sensors will be at Garden City, Kansas to acquire data at the same time over common targets (calibration panels) to quantify more fully the correlation between instruments.

Data Library

All data listed in Table 2.3-1 are in the Field Measurements Data Library located at and maintained by Purdue/LARS. In addition, supporting data acquired at the time of each mission including 9-inch color and color IR aerial photography, 70-mm color photography from the helicopter, ground level photography of plants, agronomic observations and measurements, and meteorological and atmospheric measurements are in the library. Catalogs of the data were prepared this year to aid analysts in selecting data for analysis.

Data Distribution

Data from the spectrometer systems (FSS, FSAS, and Exotech 20C), plus canopy modeling data, and associated supporting data have been routinely supplied to investigators at Texas A&M University (Harlan) and the Environmental Research Institute of Michigan (Malila). In addition, selected data have been provided to Goddard Institute for

RESPONSE, BI-DIRECTIONAL REFLECTANCE FACTOR

JULY 10, 1975

WILLIAMS CO., N.D.

X   WHEAT
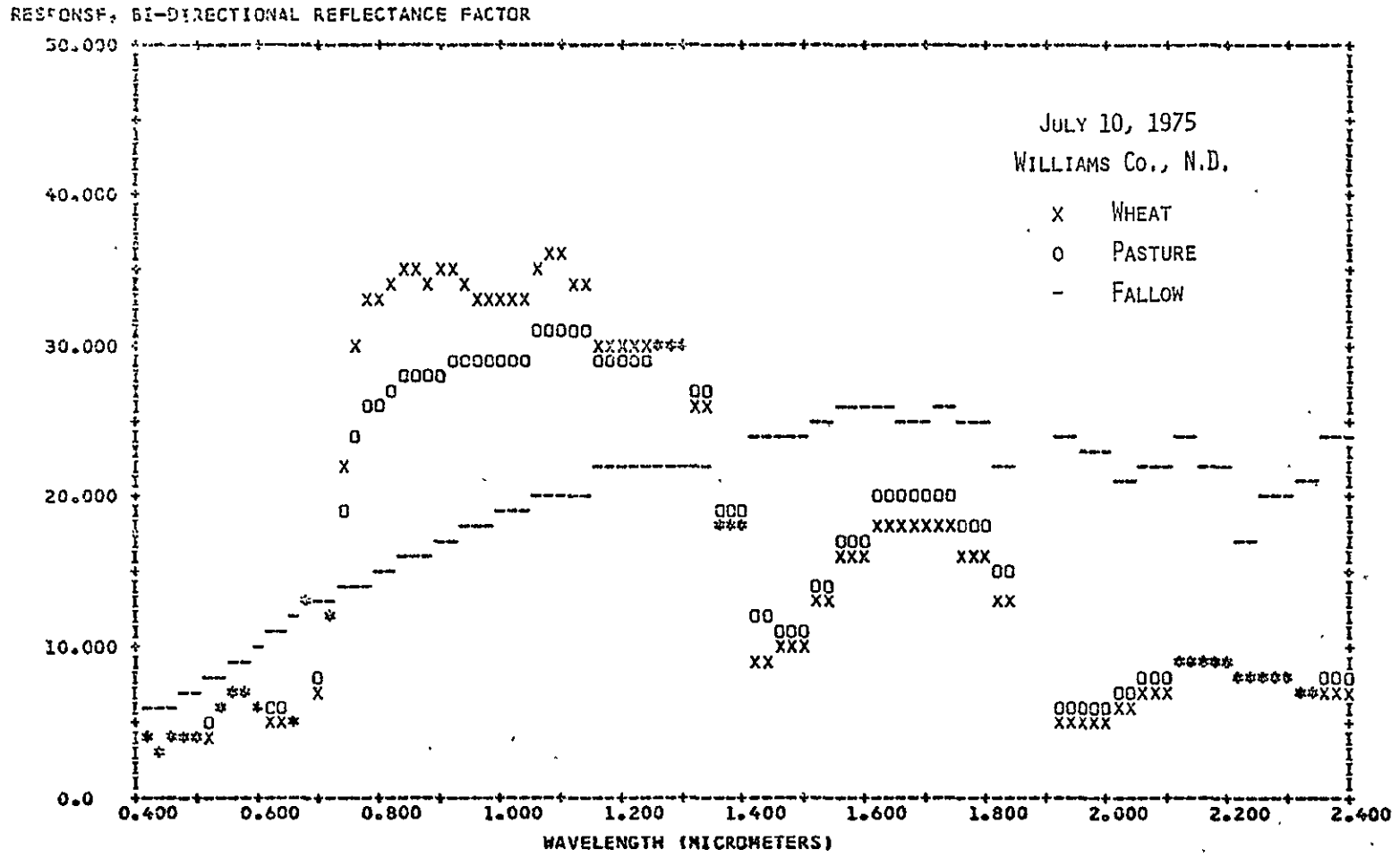
O   PASTURE

—   FALLOW

WAVELENGTH (MICROMETERS)

Figure 2.3-2 This Figure shows in graphical form examples of the spectra in the Data Library. There are over 75,000 such spectra in the library and they can be made available in several formats including digital tape. Supporting agronomic and meteorological data are also available.

Space Studies (Unger), Goddard Space Flight Center (Barker/ de Gasparis) and Kansas State University (Kanemasu). The library facility is prepared, upon approval of NASA/JSC/EOD and the Earth Resources Program Office, to prepare data sets for other interested investigators for the cost of reproduction.

Data Analysis

Analysis of data for two projects was completed this year. The technical report, "A Laser Technique for Characterizing the Geometry of Plant Canopies," by Vanderbilt, Silva, and Bauer was submitted during the third quarter. The method, a variant of the point quadrant method, involves aiming a laser at the canopy and measuring the height at which the beam hits a component of the canopy. Two kinds of information, location and orientation, describing the geometric characteristics of canopy foliage can be obtained using the method. These kind of data are required as input to many models of canopy radiation. The feasibility of the technique was tested for wheat and corn canopies and recommendations are made for implementation of an automated version of the method.

A report of the results of the second analysis describing the multispectral reflectance of a wheat canopy as a function of time of day and azimuth and zenith view angles is currently being prepared. Additional analyses of the spectrometer data acquired over the intensive test site fields and agriculture experiment station plots is currently underway. The objective of these analyses is to quantify the relation of various agronomic variables, including stage of maturity, leaf area index, biomass, height, and percent cover to the multispectral reflectance of the canopies.

2.4 . SCANNER SYSTEM PARAMETER STUDY

The scanner parameter study was initiated to conduct research on analytical methods of multispectral sensor system design. It was planned that basic signal representation and information theoretic approaches would be taken to determine optimum performance levels achievable in a given environment against which to test performance of actual or modeled systems. Scanner system modeling techniques were to be developed which would permit explicit evaluation of scanner performance based upon actual field data. The study includes representation of the scene and information extraction process to provide models of the environment in which the scanner would operate. Thus three main activity areas were defined (scene, scanner, and information extraction modeling) and research was begun in June 1975.

Shortly into the first year the direction of the study was temporarily changed at the request of the sponsor. A numerical simulation study of certain proposed thematic mapper parameters was requested using NASA 24 channel scanner data as input. This activity required almost all the project resources through the end of the first year and the results were reported in the Final Technical Report of Contract NAS9-14016 in June 1976 and in Information Note 110976. The results were presented before the 1976 Symposium Machine Processing of Remotely Sensed Data and appears in the July 1977 issue of the IEEE Transactions on Geoscience Electronics. Research on the analytical model program was resumed in June 1976. Progress has been made in the three basic areas defined for the study.

Scene Modeling

In order to provide sample test data against which to design the scanner and performance estimation models a scene data base task was carried out. A restricted set of classes and data types were chosen to fit within the resources of the study. The wheat mapping problem was chosen as the context and the data to be used was selected from the LACIE Field Measurements Data Library. In order to compare results

directly with the thematic mapper study results the same classes were chosen. These were: wheat, harvested wheat, fallow, grass/pasture and other (corn-oats). The primary sensor from which data is being utilized is the helicopter borne S-191 spectrometer, however, EXOTECH 20C and 24 Channel aircraft scanner data are also being used. The William County, N.D. site was selected as the test area.

The activities in this task included the selection, reformatting, preprocessing, error checking, cataloging and averaging of field spectrometer data for use in defining the scene model. A limited number of spectra for each class were chosen to keep the dates of acquisition close together. Once a suitable set of spectra were chosen the means and covariance matrices were computed for each class. This work has been completed and the statistical representations for each of the test classes are routinely being used for the scanner modeling study. After scanner modeling has reached a satisfactory point of development, additional data sets of this same nature but based upon other field conditions will be needed.

Scanner Parameter-Modeling

This task is the major thrust of the study and the approach in formulating the project is that a complete (theoretically optimal) representation of class spectral characteristics would be used to establish the ultimate degree of spectral separability of classes. Generalized wavelength sampling functions derived from the ensemble of spectral data comprising the scene model are used as basis functions. Individual spectral responses from the scene are then represented by a finite weighted sum of these basis functions. Two approaches were pursued in choosing basis functions. The first is based on the use of the Karhunen-Loeve (K-L) representation and the other is an information theory approach which utilizes the concept of mutual information.

Progress was made in developing models using both approaches. Algorithm selection and coding has proceded to the point that example sets of optimum basis functions have been computed for the informational classes cited above and evaluation of their performance relative to LANDSAT and thematic mapper rectangular wavebands is in progress. No

final results are yet available; however, preliminary indications using error bound estimators are that significant improvement in separability is achieved using the optimum basis functions.

The second key parameter, spatial sampling, was also modeled during the year. A spatial transformation was implemented which uses as input a high resolution data set. For any given spatial sampling function (e.g., Gaussian, triangular, rectangular, etc.) the output spatial covariance matrix can be computed. The spatial statistics of the 24 channel scanner data from Williams County, N.D. was used during the year for first tests. An exponential model for covariance appears to fit sufficiently well and parameters were estimated. Effects of the spatial sampling function on classification accuracy can be evaluated with this model.

Error Estimation

A numerical model has been developed and tested for estimating the probability of classification error for several classes and channels. Direct solutions for more than two classes are generally not available. The method commonly used is to classify a sample of data using a given method and estimate the error from the results. In the numerical model developed here the probability of error is computed directly from the multivariate Gaussian density functions for each class using the maximum likelihood decision rule. Evaluations showed that the predicted overall classification accuracy for all the five classes using three channels was within 3% of the results from the thematic mapper simulation for the same classes and channels. Further testing and evaluation is needed to verify the model, however it appears now that it will be an effective tool in evaluating alternative scanner parameter sets.

Summary

During this contract year, after completion of the last portions of the thematic mapper simulation including thorough reporting of the results in the literature, the initial versions of all three needed components to an analytical system for comparing scanner system designs were devised. A trial ensemble of spectra was constructed from the LACIE Field Experiment data base; this ensemble will be used as an

initial test set for testing and debugging of the other components.
The initial software implementing algorithms for two spectral and one
spatial optimal sampling schemes was devised. And an initial procedure
for projecting the classifier performance achievable with such optimal
sampling schemes was devised. During the coming year these techniques
will be verified and refined as needed. It is currently expected that
after that the entire system could begin to be used to test a wide
variety of parameter sets. These could ultimately include various
numbers, shapes and locations of spectral bands; various spatial reso-
lutions and sampling schemes; different signal-to-noise ratios; uni-
and multitemporal data from different geographic locations and times
of the year; data from combinations of sources such combined optical
and radar data; and class sets relevant to various specific application
problems.

2.5 TECHNOLOGY TRANSFER

Objective and Approach

The objective of this task is to develop and test hardware, educational materials and training programs for the transfer to users of the principles and techniques of the analysis of earth resource data. The approach being taken centers around a computer terminal located at JSC and connected by means of a communications link to the Purdue/LARS computing facility. This remote terminal concept incorporates a careful blending of hardware and software capability supported with specially designed training materials and user oriented operating pro- procedures. As the concept has grown and developed over the past five years the relative emphasis among these four components has varied according to current needs and requirements as shown below.

Year	Major Activities	Emphasis
1972	Installation of remote termi-nal	Hardware, training, pro-cedures
1973	LARSYS 3.0 release, Revised training package	Software, training
1974	LACIE Analyst training	Training
1975	LARSYS 3.1 release, LANDSAT Case Study	Software, training
1976	Simulation exercises, train-ing procedures documented.	Training, procedures
1977	Terminal upgrade, ECHO Case Study	Hardware, new algorithm training

This contract year has seen an increasing emphasis on upgrading the hardware capability of the JSC/LARS terminal as well as the development of training materials to test and evaluate new classification algorithms.

Accomplishments

Efforts on this task resulted in three primary accomplishments:

2780 Remote Terminal Support
Data 100 Installation Plan
ECHO Classifier Case Study

2780 Remote Terminal Support .

Support of the JSC/LARS.2780 remote terminal includes maintaining at the LARS computer facility the necessary hardware and software, providing computer services and providing system consultation. During the last quarter of the contract year communications with remote terminal users increased markedly. This resulted in devising improved procedures to handle routine ID changes, changes in disk storage space and reformatting requests. At the request of a JSC terminal user a PL/I compiler was made available to JSC and has been successfully used. Computer usage for the year has been averaging 9.88 hours of CPU time and 325 hours of terminal attach time per month.

Data 100 Installation Plan

In response to NASA's request (December 20, 1976), LARS estimated the performance, cost, and schedule for installation of a tape drive capability on the JSC/LARS terminal. Three approaches were studied: using existing hardware to transmit small data sets; procuring enough hardware and software to read and transmit data tapes; and, procuring enough hardware and software to functionally replace the existing 2780 terminal and read and transmit data tapes in addition. These options were discussed with JSC personnel during the week of January 24, 1977 at which time it was agreed to pursue option C in more detail. This investigation lead to the preparation of a plan for the installation of a Data 100 Model 78 remote terminal at JSC. The plan, which was submitted on April 1, 1977, discussed the benefits and purposes, milestones and hardware budget. The key milestones in the plan are: hardware specifications, project start, hardware installation, software plans, training plans, software and training implementation and evaluation.

Rapid response to JSC's request for terminal upgrading was possible in part because LARS has been engaged in a low level effort to develop a concept for a significantly improved remote terminal. This terminal concept is based around a minicomputer supported by special purpose and general purpose peripheral equipment. Under this concept, known as the LARS Intelligent Terminal for Earth Resources (LITER), the large central computer would be logically subordinate to the remote

terminal and would serve as a resource for those processing steps requiring large computational capabilities. Recent progress on LITER has been minimal due to the low level of funding.

ECHO Case Study

An analysis case study is being developed in order to provide training in the use of the supervised ECHO classification algorithm. The case study is scheduled to be completed within three months after the release of the ECHO software. Unlike the case studies developed for 2780 remote terminal training, the ECHO case study is being written so that it is not dependent upon a particular hardware configuration. This will give the case study wider utility and will facilitate the test and evaluation of the ECHO processor as it might be implemented on JSC's computer system. The new format essentially simulates a batch mode computer environment. The case study is divided into a number of steps corresponding to the analysis sequence. The student will interact with a tutor at the conclusion of several activities stratigically placed throughout the materials. Appropriate example computer printouts_will be available at key points simulating a batch machine operation.

Other education and training materials completed during the contract. year included 6 additional titles to the FOCUS series and an analysis Simulation Exercise.. The new FOCUS titles are: Multispectral-Multitemporal Concept, LARSYS Version 3.1, Regional Land Use Inventories, Reformatting LANDSAT Data, The Multiband Concept and Snow Cover Mapping. Each FOCUS consists of a diagram or photograph and an extended caption of three to four hundred words treating a single concept. LARS Information Note 052977, "The FOCUS Series: A Collection of Single-Concept Remote Sensing Educational Materials" is a compilation of all titles in the series.

The simulation exercise entitled "Determining Land Use Patterns Through Man-Machine Analysis of LANDSAT Data," LARS Information Note 070676, is designed to show the reader typical steps in the analysis of remotely sensed data for determining land use patterns.

Recommendations

The remote terminal concept has proven to be successful for the transfer of proven computer image analysis techniques. It is potentially valuable for the research, test and evaluation of newly researched analysis algorithms and procedures. It is recommended that the use of the JSC/LARS remote terminal for research, test and evaluation efforts be increased and that the development of more powerful terminal systems (LITER) be continued.

2.6 LARGE AREA CROP INVENTORY DESIGN

The objective of this task was to develop a multicrop inventory design for wheat, corn, soybeans, rice, and/or other economically important crops. The approach was (1) to review the LACIE design and results to assess its adequacy for multicrop applications and (2) to develop a plan with recommendations for selection of crops and areas, defining system performance goals, and data and research requirements.

As part of the first part of the approach, LARS staff participated in LACIE reviews at NASA/JSC. In addition, results of other related investigations have been examined. Our initial conclusions from these efforts is that the LACIE design is not optimum, nor adequate for multicrop applications. In particular, the more complex environments which will be encountered in multicrop applications will require more advanced procedures and data. Most LACIE-related investigations have had limited objectives and it is difficult to apply their results to other situations.

In February, a presentation was made at NASA/JSC describing our findings concerning the distribution and production of major crops around the world. A rationale for selecting important grain production areas for inventory (rather than particular crops) was presented. A limited study of the variation in area, yield, production, imports, and exports of major by countries has been initiated, but has not been completed.

Beginning in March, at the request of NASA/JSC/EOD, the resources of the project were devoted to developing the requirements of a global food and fiber information system. With this change, major staff changes were made bringing new people from the Purdue School of Agriculture to the task. Organization and plans for this effort have been largely completed during the April-May period. This effort will be directed at an in-depth assessment of the information requirements for crops, forests, and rangelands of the world.

2.7   FORESTRY APPLICATIONS PROJECT

The FAP task activities during this past contract year have been
directed toward investigating broad scale forest inventory designs.
This task has been responsive to the perceived needs envisioned for
the Forest Service under the mandate of the 1974 Renewable Resources
Planning Act.   A broad objective for this work involves the potential
application of computer-aided LANDSAT analysis, combined with ancillary
data as input into the development of inventory designs for forest and
rangeland resources included in the 1.6 billion acres covered by the
RPA.

This FAP task activity has been narrowed somewhat in order to
make better use of available resources and expertise.   Specifically,
the task has pursued more detailed work in forest inventory, especially
that aspect of inventory involving Forest Survey (Renewable Resource
Evaluation) as conducted by the Forest Experiment Stations.   Along
with a narrowed task objective, the approach has been simplified.   Rather
than highlighting empirical methods to inventory designs we have stressed
the more practical approaches.   This direction was pursued because:

1)   we understand how the perceived user's need can be met
     working in the practical realm, and

2)   this approach allows us to better understand the data
     collection, analysis and use cycle which is often ignored
     by more highly sophisticated approaches.

As a result of our initial studies, we are proposing that a hier-
archical series of classification intensities be considered.   The amount
of detailed information which the LANDSAT analysis would provide would
vary with each level of classification.   The least detailed classifica-
tion would provide a breakdown of acreage for a given area based on a
forest/non-forest breakdown.   The maximum classification would contain
information relative to forest conditions.

Similarly under this scheme the level of ancillary or ground data
collection would vary inversely to the level of classification detail.
At the maximum level of classification detail a minimum of ground data

collection would be anticipated. For all levels of classification intensity the use of Forest Survey permanent plot field data is being considered. Recommendations embodied in this tasks final report will detail the proposed schemes and obvious advantages to the Forest Service.

2.8 REGIONAL APPLICATIONS PROJECT

Objective

The overall objective of the Regional Applications Project was the development of a computer procedure for (1) inventorying and. categorizing Texas coastal zone environments (based upon the association between Landsat multispectral data and known ground truth information) and (2) monitoring and identifying significant alterations in these environments.

Approach

Previous research by Purdue/LARS involved the classification of Landsat MSS data to produce an inventory of environmental units within the Texas Coastal zone. For the 1976/77 contract year, the emphasis was placed on development, refinement, and documentation of a change detection procedure through the following efforts: (1) Determine the extent to which detectable seasonal changes in the coastal zone classifications could be identified. (2) Determine the extent to.which detectable permanent-changes in the coastal zone classifications could be identified. (3) Determine methods for delineating between seasonal and permanent variations for specific coastal zone features. (4) Register detected change classes to 1:24,000 USGS quadrangles. (5) Define significant changes. (6) Transfer the change detection procedure with software, algorithms and user documentation to NASA/JSC.

Accomplishments

The accomplishments under this task, a six-month investigation terminated by the sponsor on 30 November 1976, are summarized as follows.

Of the four change detection techniques investigated during the past contract year (post classification comparison, delta data, spectral/ temporal, and layered spectral temporal), the post classification comparison technique was selected for further development. This was based upon test performances of the four change detection methods, straightforwardness of the procedures and the output products desired.
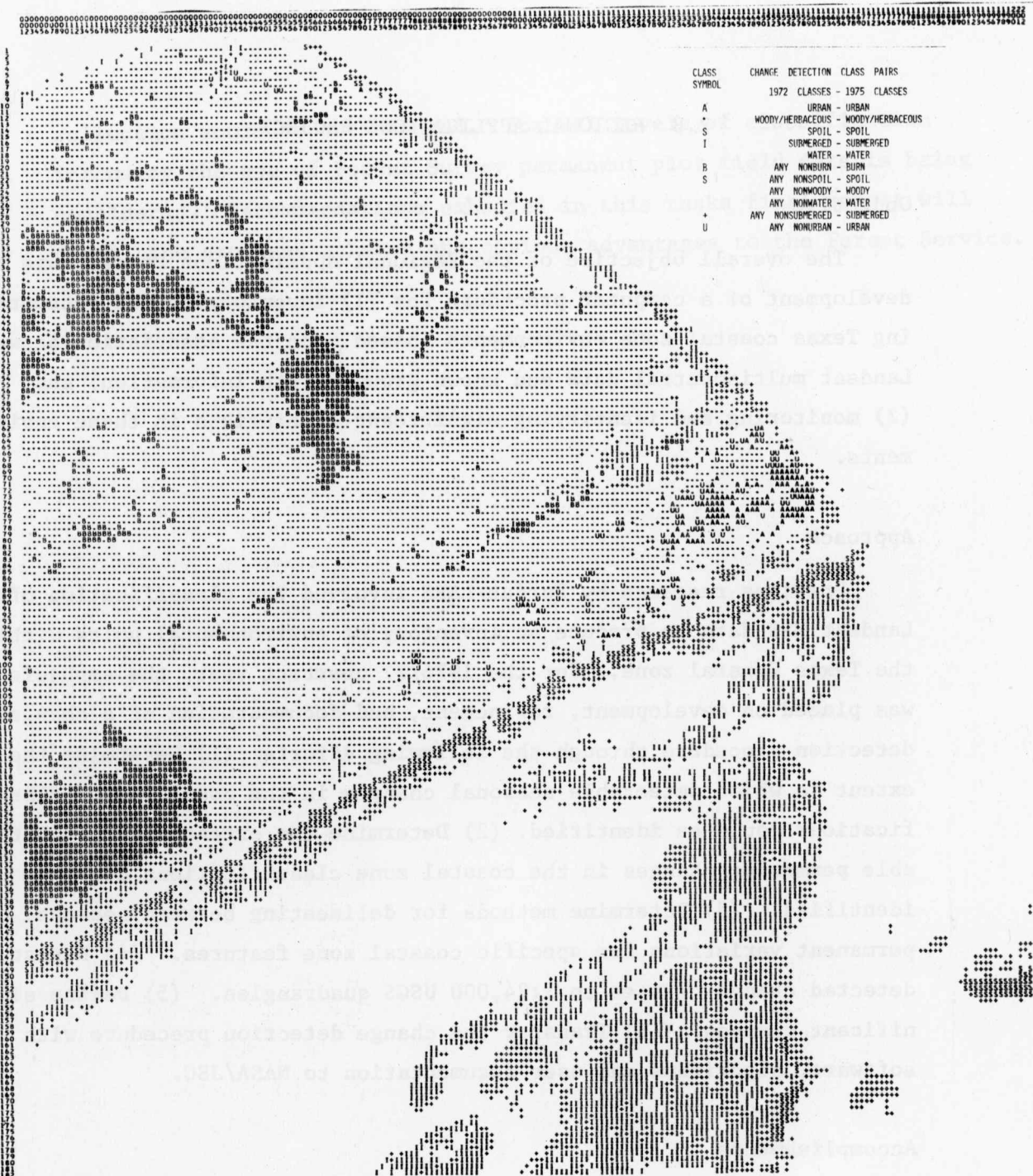
Figure 2.8-1 Change detection classification results map of the Port O'Connor, Texas quadrangle. These results were obtained by classifying Nov 1972 and Feb 1975 Landsat data using the post classification change detection processor.

A standaridized "modified" unsupervised classification procedure for analyzing the Texas Coastal Zone data was compiled. Single date classifications were completed for the three quadrangle areas for the February 25, 1975 date to be used as the base data to which classifications developed from the later three dates were to be compared.

An additional task involved the modification of the Spectral Environmental hierarchy developed by Johnson Space Center/Lockheed Electronics Corporation. The original hierarchy, developed using photointerpretive information such as texture, spatial distribution, subjective cultural interpretations and spectral response, did not adequately describe the Landsat data since the Landsat multispectral scanner records data predominantly spectral in nature. The final version of the modified spectral hierarchy was not completed.

Another area of involvement consisted of investigation of methods for improving the geometric accuracy of the registered coastal zone quadrangles. The mathematical functions tested were the affine and the collinearity geometric corrections. The results of these methods were compared to the least squares bi-quadratic method used in the original corrections. Results indicated that the affine geometric correction function is the most accurate of the two tested in areas where control points are not well distributed over the area to be corrected.

The RAP simulation study was completed during this period. This study consists of a documented procedures package for classifying the Pass Cavallo quadrangle in the Texas Coastal Zone.

Recommendations

There is a great need for quantitative change detections and procedures, not only for monitoring coastal zone changes, but in many application areas. It is recommended that the development of such techniques be resumed. For this effort to be successful, the collection of adequate reference data coincident with the desired Landsat overpasses is required. Without coincident reference data, performance criteria cannot be properly evaluated.

Even though the post classification comparison change detection technique was selected for this portion of the study, it is felt that the layered spectral/temporal change detection procedure holds great promise for monitoring change. This technique should be reevaluated when an appropriate data set (coincident reference and Landsat data) has been assembled.

2.9 THERMAL BAND INTERPRETATION

The objective of the thermal band interpretation project was to
develop a model that describes the heat transfer from wheat and other
crop canopies. The procedure used involved the acquisition of experi-
mental data during three stages of wheat growth at the Williston,
North Dakota test site. The data collected were that necessary to
support a simplified radiation model that was developed during the
course of the project. The data were then reduced and analyzed to
test the efficacy of the model in predicting the effective radiant
temperature of the canopy as a function of canopy structure.

The canopies chosen for study were winter wheat and spring wheat
at the Williston, North Dakota intensive test site. Wheat fields on
the North Dakota State University Agricultural Experiment Station as
well as wheat fields near the Station were used as sources. A specially
modified rapid-scan thermal infrared scanner was used to obtain contin-
uous radiant temperature profiles in the wheat canopies. Additionally,
data on a corn canopy in Tippecanoe County were also obtained. Ther-
mistors mounted at 5 different heights in the canopy were used to
acquire air temperatures within the canopy. Soil temperatures and
air temperatures above the canopy were also obtained. A radiation
thermometer was used to measure the radiant temperature data in critical
portions of the canopy. Calibrated blackbody references were used to
standardize all radiant temperature data obtained in the canopy. Pro-
cedures described in Information Note 120776 were used to obtain data
necessary to calculate view factors in the canopy. Data were obtained
during June, July, and August with the June data obtained on North
Dakota winter wheat, the July data obtained on a mature green, fully
headed spring wheat canopy, and the August data obtained on a ripe spring
wheat canopy. In addition to the thermal profile data, thermal image
data were obtained with both horizontal and vertical views of the canopy
which were subsequently digitized for conventional level slicing analy-
sis.

The radiant temperature profile data obtained by the thermal scan-
ner was carefully calibrated using the blackbody references present in

the field of view of the thermal scanner. The air temperature profile data that was obtained simultaneously by the thermistor arrays positioned in the canopy was carefully correlated with the radiant temperature profile data. Fast scan thermal scanner data was obtained in image format and the effects of canopy motion upon the apparent radiant temperature of the canopy were studied. Spectral-reflective and spectral-thermal were also obtained on the canopies.

A simplified radiation model was developed to correlate the geometric characteristics of the canopy with the radiant temperature viewed from various angles above the canopy. The model utilizes the radiant temperature profile along with the view factors described above as input data. The model yields the apparent radiant temperature of the top of the canopy (assuming the emittance of the canopy to be equal to unity) as a function of viewing angle. Complete agronomic data concerning the yield of the canopy as well as the leaf area index of the canopy were also available to enable correlation of the view factor data with crop canopy yield.

The following specific conclusions were obtained during the contract period.

1. When a wind blows over the canopy an insignificant change in the radiant temperature of the plants themselves seems to occur. The change in the apparent radiant temperature of the canopy is due almost entirely to the changing exposure of the heated soil and the ratio of soil to plants that occurs during the wind induced motion of the canopy.

2. The steepness of the radiant temperature profile from the top of the canopy to the bottom of the canopy is directly related to the vigor of the canopy. The air temperature profile in most cases can be placed in close correspondence with the radiant temperature profile of the canopy.

3. Insufficient data were acquired to geometrically characterize the canopy to an extent necessary for a full test of the radiation model. A special experiment to take place during the summer of 1977 will be directed toward acquiring an adequate set of geometric characterization data that will permit the calculation of accurate view factors.

4. There is very little spectral character in the thermal data of full canopies.
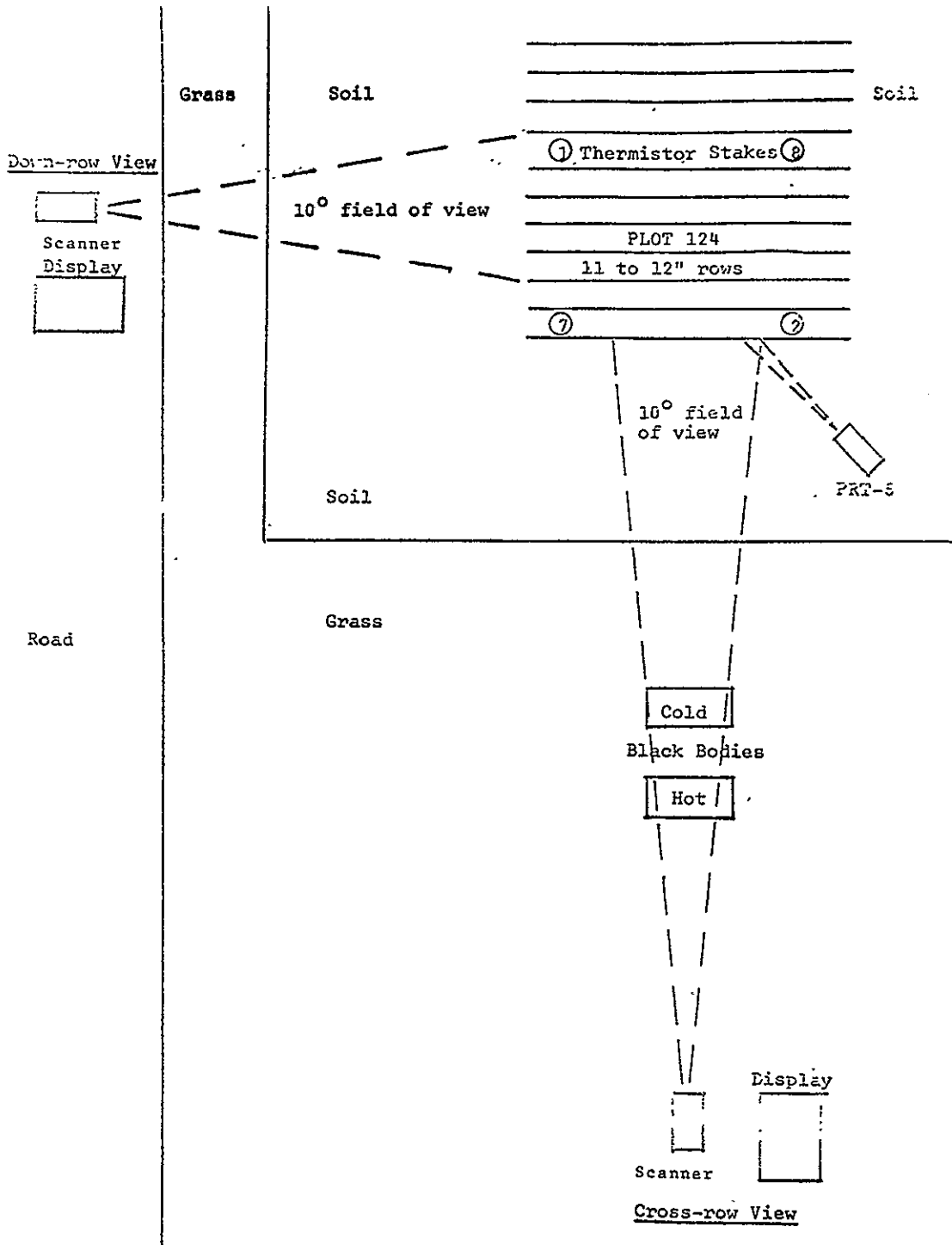
Figure 2.9-1 As a part of deriving a thermal model for a wheat canopy, data was gathered with an imaging thermal scanner. Shown here is the instrument arrangement used.
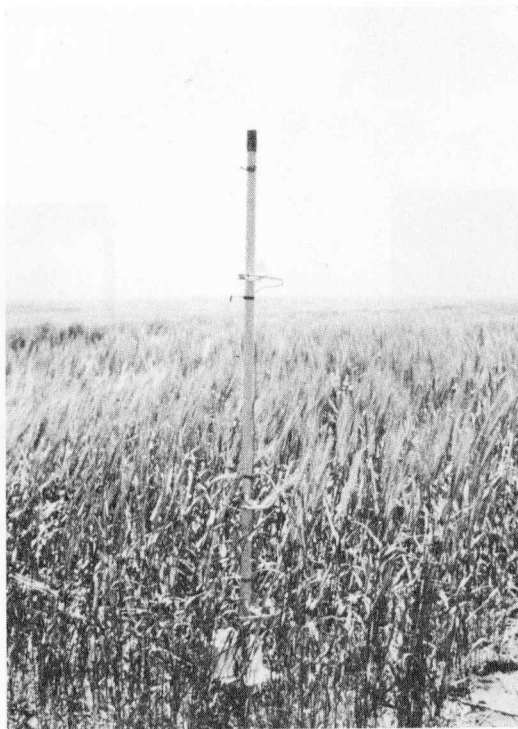
(A) Thermistor Stake in Wheat Canopy

(B) Arrangement of Instrumentation

Figure 2.9-2   Photograph of the Thermistor Stake and Instrumentation Arrangement
(Plot #124, 17 June 1976)
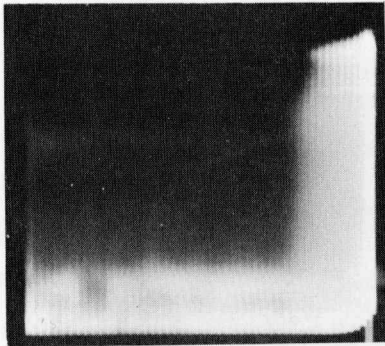
(A) Thermal Image,
    full sunlight
    (local time 1425)

(B) Line Scan,
    full sunlight
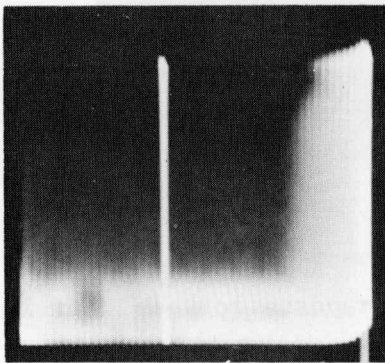
(C) Thermal Image,
    thick cloud cover
    (local time 1428)
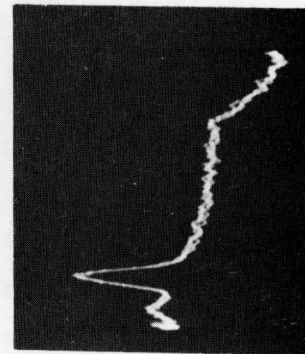
(D) Line Scan,
    thick cloud cover

Figure 2.9-3   Side Thermal Images and Line Scans
(Plot #124, 17 June 1976)

(A) Canopy Side View
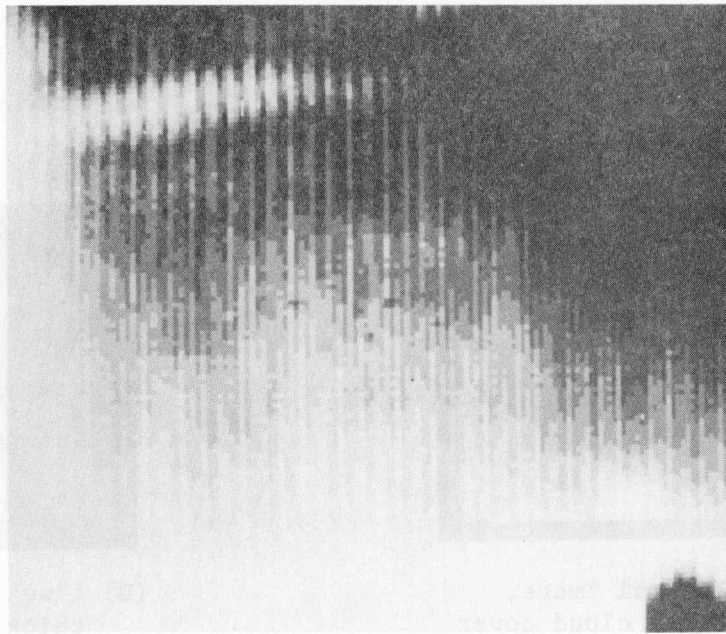
(B) Digitized Thermal Image

Figure 2.9-4 Photograph and Digitized Thermal Image, Side View of Canopy
(Small Grain Test Plot #7, 10 July 1976, local time 1801)

2.10 SUPER SITE DATA MANAGEMENT

The primary objective of this task has been to perform evaluation
and verification of the LACIE Field Measurements data. The task
requires quantitative evaluation of each of the major sensors, fol-
lowed by documentation of the results. The documentation will be
used by researchers who use the field measurements data to learn what
procedures were used to collect the data and what environmental and
system conditions existed at the time of data collection. The following
sections describe the accomplishments and current status of the effort
for each type sensor.

Landsat MSS

Landsat MSS imagery is examined for cloud cover over the test
sites; data with extensive cloud cover are rejected. The data is also
checked for such problems as bad scan lines.

Aircraft Multispectral Scanner

Due to the volume of scanner data, it cannot be satisfactorily
evaluated by visual checks, thus software to use in evaluation of this
data has been prepared. It provides for analysis of histograms of each
channel of data to check the analog to digital conversion, calculation
of the level of scanner noise, and checks on line-to-line and channel-
to-channel synchronization. Threshold levels for problems can be set
and any questionable data flagged.

Helicopter Spectrometer (FSS/S-191)

Software has also been developed and implemented for evaluation
of the FSS/S-191 data using measurements of the calibration panel. The
calibration panel data for each of the three to five observations of
the panel during a mission are normalized with respect to sun angle.
A large variation in the normalized data can indicate possible problems
with the data for a particular mission. To quantify the variation, the
coefficient of variation of all bands (except in water absorption
regions) is calculated. Large variation would indicate cloudiness,

panel not filling the instrument field of view, instrument instability, or shadowing of panel by the helicopter. This information combined with the total irradiance strip charts will provide analysts with a very good measure of data quality.

Truck-Mounted Spectrometers

The data from the truck-mounted systems (JSC/FSAS, ERL/Exotech 20D, and LARS/Exotech 20C) are evaluated using measurements made of the five gray level panels. These data provide information on the repeatability of measurements and the correlation or agreement in measurements by each of the sensors.

Summary and Conclusions

The task of developing and performing procedures to calibrate and correlate data from different sensors in a field environment is very difficult. However, it is being accomplished and for the first time quantitative information on data quality will be available. The results to data indicate that the data from the several sensors are highly correlatable. Further verification of the instruments will be made during a side-by-side comparison over common targets at Garden City in May 1977. A "Field Measurements Spectral Data Verification Report" is currently being prepared which will discuss the procedures and results of the calibration and correlation.

SYMBOL FLT. LINE CAL TIME IFOV ZENITH COSINE ZENITH

SYMBOL	FLT. LINE	CAL TIME	IFOV	ZENITH	COSINE ZENITH
1	1	18.21.09	1:25	37.3	0.795
2	2	18.39.23	1:25	36.7	0.801
3	3	19.03.15	1:25	36.6	0.803
4	4	19.22.21	1:25	37.0	0.799

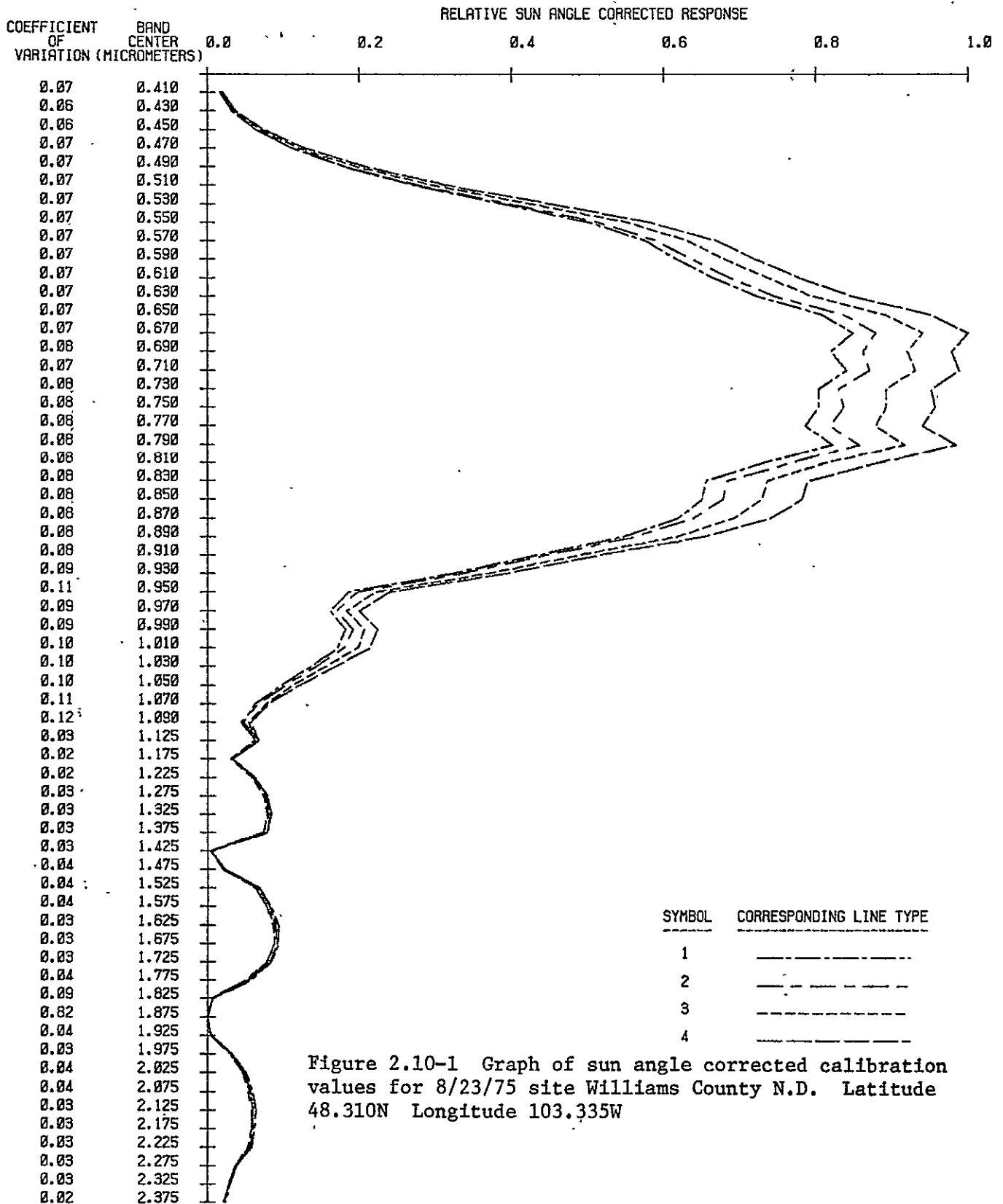AVERAGE COEFFICIENT OF VARIATION (COEF. VAR.) IS .06

Figure 2.10-1 Graph of sun angle corrected calibration values for 8/23/75 site Williams County N.D. Latitude 48.310N Longitude 103.335W

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

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

RELATIVE SUN ANGLE CORRECTED RESPONSE

COEFFICIENT OF VARIATION	BAND CENTER (MICROMETERS)
0.02	0.410
0.03	0.430
0.02	0.450
0.02	0.470
0.02	0.490
0.01	0.510
0.01	0.530
0.02	0.550
0.02	0.570
0.02	0.590
0.01	0.610
0.01	0.630
0.01	0.650
0.01	0.670
0.01	0.685
0.00	0.695
0.01	0.710
0.01	0.730
0.01	0.750
0.00	0.770
0.01	0.790
0.01	0.810
0.01	0.830
0.01	0.850
0.01	0.870
0.01	0.890
0.01	0.910
0.01	0.930
0.02	0.950
0.01	0.970
0.02	0.990
0.03	1.010
0.03	1.030
0.05	1.050
0.05	1.070
0.06	1.090
0.10	1.125
0.08	1.175
0.06	1.225
0.06	1.275
0.08	1.325
0.20	1.375
0.17	1.425
0.11	1.475
0.06	1.525
0.05	1.575
0.05	1.625
0.05	1.675
0.05	1.725
0.07	1.775
0.10	1.825
0.12	1.875
0.06	1.925
0.07	1.975
0.05	2.025
0.04	2.075
0.04	2.125
0.03	2.175
0.03	2.225
0.03	2.275
0.03	2.325
0.03	2.375

SYMBOL	CORRESPONDING LINE TYPE
1	
2	
3	

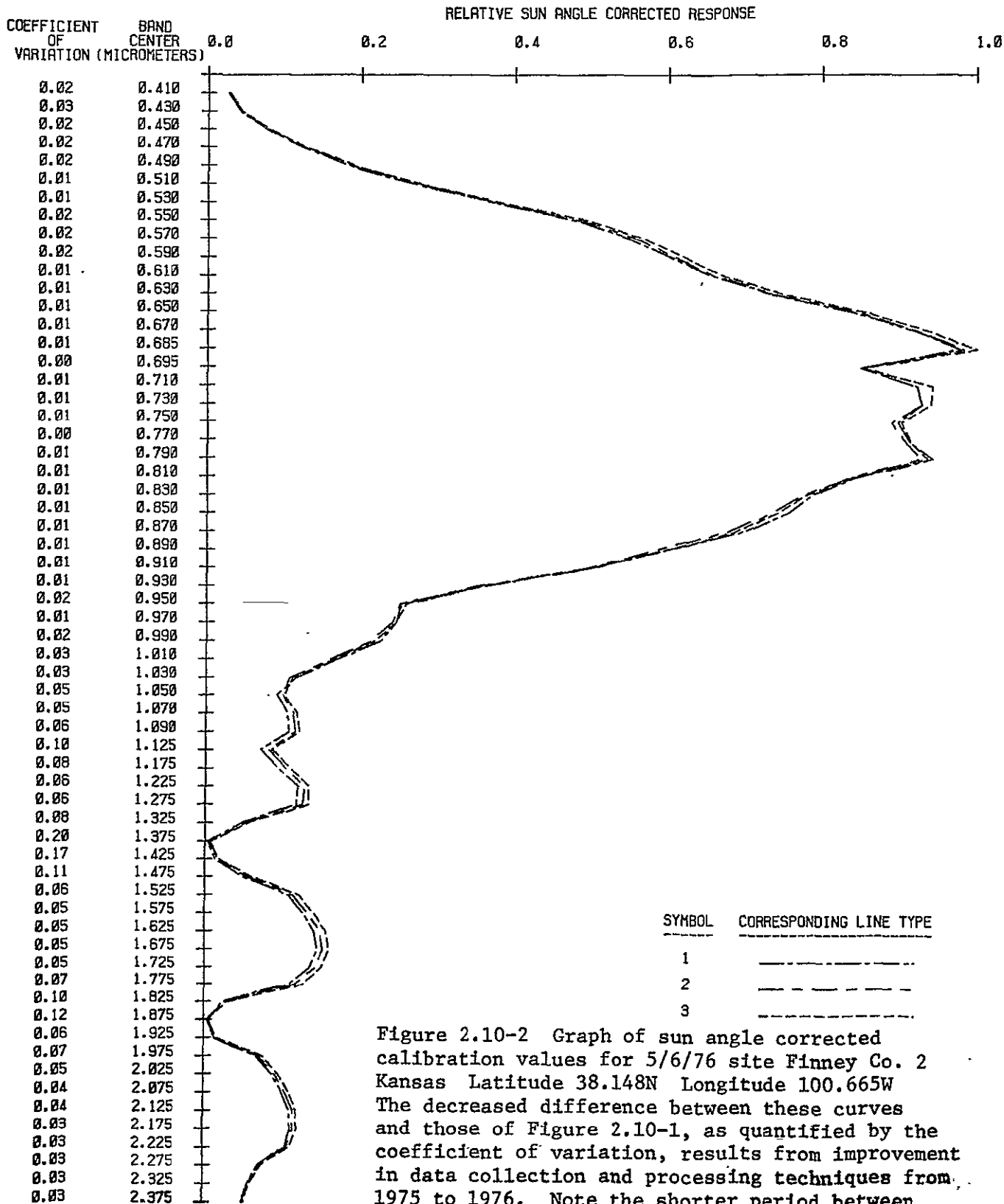Figure 2.10-2 Graph of sun angle corrected calibration values for 5/6/76 site Finney Co. 2 Kansas Latitude 38.148N Longitude 100.665W The decreased difference between these curves and those of Figure 2.10-1, as quantified by the coefficient of variation, results from improvement in data collection and processing techniques from 1975 to 1976. Note the shorter period between calibration times in the two cases.

Finney Co., Kansas		Williams Co., N. Dakota		Hand Co., So. Dak.	
Date	Ave. Coef.* Variation	Date	Ave. Coef. Variation	Date	Ave. Coef. Variation

1974-75 Crop Year

Finney Co., Kansas		Williams Co., N. Dakota		Hand Co., So. Dak.	
11/05/74	.13	6/5/75	.09		
3/20/75	.09	6/22	.08		
4/8	.09	7/10	.09		
5/14	.03	7/18	.06		Not
5/21	.12	7/27	.02		
6/2	.04	8/5	.04		Applicable
6/9	.10	8/15	.04		
6/17	.09	8/23	.06		
6/26	.04	9/1	.08		
7/6	.09				

1975-76 Crop Year

Finney Co., Kansas		Williams Co., N. Dakota		Hand Co., So. Dak.	
9/16/75	.05	5/13/76	.01	10/15/75	.06
10/3	.24	5/28	.01	10/30	—
10/21	.04	6/17	.01	11/5	.04
11/11	.03	6/25	.01	5/11/76	.03
3/18/76	.02	7/6	.03	6/1	.03
3/31	.05	7/20	.03	6/19	.05
4/18	.02	7/28	.03	7/8	.04
5/6	.03	8/9	.02	7/31	.03
6/12	.06	8/19	.08		
6/30	.07				

1976-77 Crop Year

Finney Co., Kansas		Williams Co., N. Dakota		Hand Co., So. Dak.	
9/28/76	.01			9/21/76	.02
10/15	.03				

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

Table 2.10-1 A summary of FSS Calibration Panel response variations. Tables such as this would help direct a potential data user to data quality characteristics which may be important to his research.

2.11 SOIL CLASSIFICATION AND SURVEY

Since detailed soils maps (1:25,000 or larger) are available for only a small fraction of the land area of the earth, a rare opportunity is available for the application of remote sensing technology to make a significant contribution in defining quantitatively the land use capabilities and potential productivities of the soils of the world. If dependence is placed upon conventional methods of soil survey to provide detailed land use capability maps and potential soil productivity ratings, many decades will pass before the task is completed. The race between the global population increase, food production and land deterioration suggests that the human family dare not wait for conventional methods to characterize and map the soils of the world.

If present satellite sensors and the much improved sensor systems planned for future satellites are to be used most effectively in the preparation of land use capability maps and soil productivity ratings as these relate to food production, it is crucial to define quantitatively the soil variables related to productivity which can be measured by or correlated with multispectral radiation from the surface soil. With this imperative in mind research at Purdue/LARS during the past year has been conducted to add to our understanding of the multispectral reflectance of soils as it relates to climate, physical properties of soils, chemical properties of soils, internal drainage, deterioration hazards, and potential productivity.

At LARS Montgomery recently concluded an examination of a limited number of soils selected to represent a wide range of climatic and other environmental conditions in the United States. He obtained reflectance curves over the spectral range from 0.5 to 2.38μm and studied the correlations between reflectance and 10 other physical/chemical properties of soils. Those found to have the highest correlation with reflectance are cation exchange capacity and the contents of silt, clay, iron oxides and organic matter. Silt content was the single most significant parameter of those being studied in explaining the spectral variations of soils. Results from this study indicated that the middle infrared (2.19-2.38μm) is the best spectral region of

the range studied (0.5-2.38μm) for evaluating these relationships.
Further, Montgomery found, even with his limited number of samples,
that best correlations were obtained between reflectance and physical/
chemical measurements when the samples were stratified according to
climatic region.

Two manuscripts are nearing completion and will be submitted to
technical journals to report the results of Montgomery's research.

As a continuation of this research a much larger number of soils
is being selected from a list of 1,377 Benchmark Soils which have been
identified, studied and described in detail by the U.S. Department of
Agriculture. A statistical sampling of U.S. soils will be made to
include (1) three significantly different mean annual precipitation
zones, (2) three significantly different mean annual temperature zones,
(3) major parent material regions, and the ten soil orders, the highest
level or categorization of soils according to the U.S. System of Soil
Taxonomy. The ultimate objective of this continuing research is to
provide a body of knowledge and interpretive skills which will render
remote multispectral sensing a valuable tool for mapping soils, deter-
mining land use capabilities and soil productivity ratings, identify-
ing crops and predicting yields.

In a study to relate multispectral measurements with field condi-
tions, two spectrally different soils (a dark, poorly drained Typic
Hapludalf and a light well-drained Typic Argiaquoll) have received the
following treatments: (1) surface dry, (2) surface wet, (3) surface
without cover, (4) surface with 25% cover of chopped corn stover.
Spectral measurements were obtained over the spectral range from 0.5
to 2.38μm. Surface soil samples were obtained from each plot at the
time of acquisition of multispectral data. Moisture and other physi-
cal/chemical determinations will be made for studying relationships
with reflectance.

Three different dates of Landsat data for Tippecanoe County,
Indiana are being examined and analyzed to determine the extent to
which meaningful soils delineations can be determined on different
dates. Comparisons will be made for four test sites among the

spectral maps from Landsat data, spectral maps from aircraft scanner data, and soils maps prepared by conventional means.

Several investigators at LARS and elsewhere have used multispectral data to delineate and map meaningful soils differences, but as yet there is limited understanding of the effects which mean annual precipitation, mean annual temperature, age of weathering, parent material and interactions of these factors have on the relationships between reflectance (0.5-2.4μm) and various physical/chemical properties of soils. Until a more complete understanding of these relationships is realized, the applications of remote multispectral sensing will be greatly limited as a predictive tool in (1) characterizing land use capability and soil productivity, (2) mapping and monitoring global land degradation and (3) defining the effects of soil background on crop identification and yield predictions.

2.12   IMPROVED ANALYSIS TECHNIQUES FOR MULTITEMPORAL DATA

Objectives

The objectives of this task were to develop and test algorithms
and procedures for analysis of multitemporal data sets; to evaluate the
utility of classifiers of advanced design (such as perfield classifiers
and layered classifiers) in contributing to effective multitemporal analy-
sis; and to document and transfer effective techniques to NASA.

Approach

The most straightforward way to accomplish multitemporal ana-
lysis is to extend the "conventional" multivariate classification meth-
odology to data sets formed by "stacking" or concatenating the multispec-
tral data vectors from successive acquisitions over a given site. This
presumes, of course, that the data sets are geometrically registered. This
straightforward approach is attractive because of its apparent simplicity,
and in some instances it produces excellent results, although at some
computational cost due to the higher dimensionality of the data involved
(computation is roughly proportional to the square of the total number
of channels used for classification).

However, in addition to the computational cost, there are a
number of compelling reasons to pursue alternative strategies for multi-
temporal analysis. All are directly or indirectly related to the relative-
ly limited quantity of training data typically available in the face of
the increased amount of training data actually needed for multitemporal
analysis. Thus, the analysis techniques pursued under this task sought
to minimize the need for training data while maximizing the performance
gains derived from the multitemporal analysis.

The need for increased amounts of training data arises in many ways.
To begin with, the joint effects of natural variability in the ground
scene together with differential rates of change in the scene with the
passage of time can result in an enormous profusion of "spectro-temporal"
subclasses. Adequate training data must be found for each such subclass,
and as the dimensionality of the data increases, the number of training
data points needed per class also increases. Thus there is a two-fold
effect.

The situation can become worse through the loss of reference data ("ground truth") due to various factors such as loss of the multispectral data at one or more of the acquisitions due to cloud cover or data system malfunction; and, in situations involving ground visitation, failure of the observer to make or accurately record sufficient observations at all specified sites.

Registration accuracy is also a factor in training sample availability. Training fields, which are often fairly small, must be treated as though they were even smaller in multitemporal data because of the increased edge effects due to registration errors.

In summary, although increased data dimensionality plus spectro-temporal variability in multitemporal data dictate a need for increased amounts of training data, there are many factors at work which tend to reduce the amount of training data actually available. Our approach to dealing with this dilemma is to use classification methods which permit "temporal decoupling", staged, rather than joint, use of multi-temporal data, and thereby substantially reduce the demand for training data. The results obtained before termination of the project at mid-year had begun to demonstrate the effectiveness of this strategy.

Accomplishments

The accomplishments under this task, a six-month investigation, are summarized as follows.

Research was completed on techniques for registration of multi-temporal data and a comprehensive technical report was prepared and delivered to NASA [1]. Briefly, several registration methods were evaluated relative to their geometric accuracy capability. Also an investigation of the effects of relative geometric distortion on registra-tion accuracy was conducted.

Four multitemporal data sets were assembled and two methods of multistage classification for multitemporal analysis were investigated: the Layered Classifier or decision tree classifier developed earlier [2,3], and an extension of the maximum likelihood classifier which we have formulated and called the Cascade Classifier (see Figure 2.12-1). The latter is described more fully in the annual report for the contract. The preliminary results obtained indicate that both approaches have

much to offer in the multitemporal domain. These results were obtained through a (non-funded) student project in the School of Electrical Engineering at Purdue University.

Experimental Results

Comparative analyses were performed for the Fayette County and Grant County test sites.

The Fayette County, Illinois, test site was classified into corn, soybeans, wheat, and other. The Layered Classifier was designed to use t_2 data to resolve uncertainties resulting from classification of t_1 data. The Cascade Classifier was designed using highly simplified assumptions about the prior probabilities and transition probabilities. The results were as follows:

t_1 unitemporal (June 29, 1973)	68% correct overall
t_2 unitemporal (July 17, 1973)	72% correct overall
Layered Classifier	82% correct overall
Cascade Classifier	84% correct overall

The Grant County, Kansas, test site was classified into alfalfa, corn, fallow, pasture, wheat. The same design strategies were used as in the Fayette County case. The spectral discrimination problem was clearly much more difficult (notice that the unitemporal performance is considerably below that of the Fayette County case). The results were as follows:

t_1 unitemporal (May 9, 1974)	62% correct overall
t_2 unitemporal (July 20, 1974)	55% correct overall
Layered Classifier	68% correct overall
Cascade Classifier	64% correct overall

In this case, the Cascade Classifier was apparently much more sensitive to the simplified assumptions used for the prior probabilities and transition probabilities.

In neither of the above cases was there sufficient training data available to permit multitemporal analysis by the "stacked vector" approach.

Conclusions and Recommendations

Many questions remain to be answered with respect to how to best

Make use of the information contained in the temporal dimension when multitemporal data is available. The "stacked vector" analysis approach is the most straightforward to apply, but though successful in some instances, it suffers some severe disadvantages when the availability of training data is relatively limited. It is also computationally very expensive unless subsets of the available multitemporal channels are selected.

Multistage analysis, using methods such as the Layered Classifier and the Cascade Classifier, ameliorate many of the difficulties encountered with the "stacked vector" approach. The dimensionality of each decision is minimized, and this reduces both the computation and the amount of training data required. Experimental results obtained using these classifiers demonstrate that further investigation of this staged approach to multitemporal analysis is warranted. Pursuit of this research is strongly recommended.

REFERENCES

1. Svedlow, M.C., C.D. McGillem, and P.E. Anuta, "Analytical and Experimental Design and Analysis of an Optimal Processor for Image Registration,"LARS Information Note 090776, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana 47907, September 1976 (178 pages).

2. Wu, C.L., D.A. Landgrebe, and P.H. Swain, "The Decision Tree Approach to Classification, "LARS Information Note 090174, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana 47907, September 1974 (194 pages).

3. Final Report on NASA Contract NAS9-14016, D.A. Landgrebe, Principal Investigator, June 1, 1975 to May 31, 1976, Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana 47907.