022877

QUARTERLY REPORT

Reporting Period: December 1, 1976 to February 28, 1977
Contract Number: NAS9-14970
Title of Investigation:

Research in Remote Sensing of Agriculture,
Earth Resources, and Man's Environment

Principal Investigator:

D. A. Landgrebe
Laboratory for Applications of Remote Sensing
Purdue University
West Lafayette, Indiana 47906

Submitted to:

NASA Lyndon B. Johnson Space Center
Attn: Jon D. Erickson, Mail Code SF3
Houston, Texas 77058
Mark for: Contract NAS9-14970

2.1 Test of Boundary Finding/Per Field Classification (ECHO)

I. Background

During the first two quarters of FY77 efforts concentrated largely on defining the problems to be addressed, developing a plan of attack and securing the appropriate resources to proceed towards the projected accomplishments of documented FORTRAN programs implementing the ECHO classification approach, evaluations of these ECHO algorithms over LANDSAT, aircraft and simulated Thematic Mapper data, and products provided to JSC for the evaluation of the utility of ECHO object maps to LACIE Analyst Interpreters. Comparisons of ECHO and pixel-at-a-time classification accuracies are to be made, and improvements are to be made to the algorithms, as required.

To these ends, a detailed implementation plan was developed; the ECHO project staff was familiarized with the ECHO concept and the existing software; users of the research software were consulted; the desired configuration and I/O for the delivery software were selected and upgrade plans were devised for both the supervised and non-supervised processors; the existing software was altered to simulate the anticipated delivery software; problems with the ECHO analysis approach were identified and examined; the LANDSAT aircraft and simulated Thematic Mapper data sets available at LARS were surveyed and candidate sets selected for the experiments; the experimental design was revised and the experiments and software implementation begun.

During the third quarter, substantial progress was made toward the project objectives. Work concentrated on the finalization of the experimental design, the software implementation, running the experiments for the evaluation, software documentation and development of the ECHO User's Guide.

II. Major Activities

A. Experimental Plan. During the third quarter, the experimental design was revised and informally discussed with JSC. The purposes of the experiments are to establish guidelines for the parameter settings of both the supervised and non-supervised processors, to assess the effective-

ness of the extraction and classification of homogeneous objects, and to compare the performance of the ECHO processors to the performance of the per point classifier.

In discussion of the experimental design on January 28, 1977, M. Trichel (JSC/EOD) asked about the possibility of two additions to the experiments. First, he asked that two additional methods of classification performance be considered. The first method would be to measure classification accuracy on a random sample of pixels of fields selected by the analyst. Due to lack of adequate ground truth and to the level of personnel resources which would be required, we have concluded that it is not feasible to add this method of assessment to the experimental plan. The second method would be to assess classification accuracy on exhaustive fields; i. e., on "wall-to-wall" fields in the data set. Again, ground truth limitations for some data sets will prevent the use of this method; on others (the LACIE SRS segments) the "extended field center pixels" (discussed in the Experimental Plan) are essentially exhaustive fields over a portion of the area.

The second suggestion by M. Trichel was the addition of multitemporal analyses to the plan. If time and resources permit, it is our intention to make an informal assessment of the performance of the ECHO processors in multitemporal analysis. But we do not expect to be able to make a sufficiently detailed assessment to support firm and quantitative conclusions about the multitemporal performance of ECHO. Therefore, we have not added this to the formal experimental plan.

In mid-February, the Experimental Plan was submitted to JSC for review.

A catalog of the 22 sets selected for the experiments was attached as an appendix to the Experimental Plan as transmitted to JSC.

B. Software Development. Software implementation has proven to be a somewhat larger, more complex task than was initially anticipated. When it became evident that the February 8 milestone for completion of the restructured software implementation would not be met, resources were shifted, placing additional personnel on the programming tasks. The status of the delivery versions of the processors is as follows:

1. Supervised ECHO

   The coding for the production of the intermediate tape has been written and checked out. The coding enabling the processor to begin with an intermediate tape and proceed to a results tape has been written. The coding for producing the classification directly from the raw input data is still in the debugging stages.

2. Non-supervised ECHO

   The coding for the non-supervised ECHO software has been written and compiled. The program is still in the debugging stage, however.

A comparison of the classification results produced by running the supervised coding in two phases (with an intermediate tape) for both the delivery version and for the research software has been made. The delivery version does duplicate the results of the existing software in the two-phase case. Results produced by the non-supervised processors and the single phase version of the supervised processor have not yet been compared to the results from the existing research software.

C. Experiments. During the third quarter, the experiments on the supervised processor for the eleven LANDSAT data sets were completed. The experiments utilizing the Simulated Thematic Mapper data sets were begun. Over 360 ECHO classifications were produced during this period. Classification accuracies, variability and CPU times were recorded as outlined in the Experimental Plan. These results are being punched onto computer cards so that they can be analyzed by statistical programs. Typical graphs of the classification accuracy measured on the basis of field center pixels, expanded test fields, and training fields as well as CPU times and average variability for the Graham County, Kansas LACIE/SRS test site are presented in Figures 1 through 5. These results are from the supervised processor with cell width of two pixels.

Figures 1, 2, and 3 demonstrate that, for Graham County, the supervised ECHO classifier achieves a higher classification accuracy than the per point classifier and is not highly sensitive to the cell selection

and annexation parameters. Both the classification average variability (discussed below) and the CPU time required to perform the classification decrease as the cell selection parameter is increased (cell splitting decreases).

Figure 3 demonstrates that as the tendency to annex increases, the training field performance increases. This result is expected, as increasing the annexation parameter increases the likelihood that the pixels which comprise the test field will be identified as homogeneous objects. It can also be noted that the effects of annexation increase as the cell selection variable increases. This is due to the fact that annexation is attempted only between adjoining cells which are homogeneous. When the cell selection parameter is lower, more cells are split and hence, fewer opportunities for annexation exist.

Test and training fields for Graham County were selected using the ground truth for the LACIE/SRS sites as provided by NASA. Alternating fields for each class in the ground truth were designated as test fields and training fields. "Expanded test fields" were produced by expanding the "field center" test fields one pixel line or column in each direction. These expanded fields approximate the actual field size and position.

Average variability is a measure of the rate of change from one information class to another. It should reflect the degree to which ECHO reduces the "salt and pepper" effect which is sometimes present in per point classifications. Average variability is calculated by systematically selecting 50 lines of the classified area, counting the number of information class changes and dividing by the number of opportunities for class changes:

Let: NS = The number of classified pixels/line

NCC = The number of class changes over the 50 selected lines.

Then:

Average variability = NCC/(50*(NS-1))

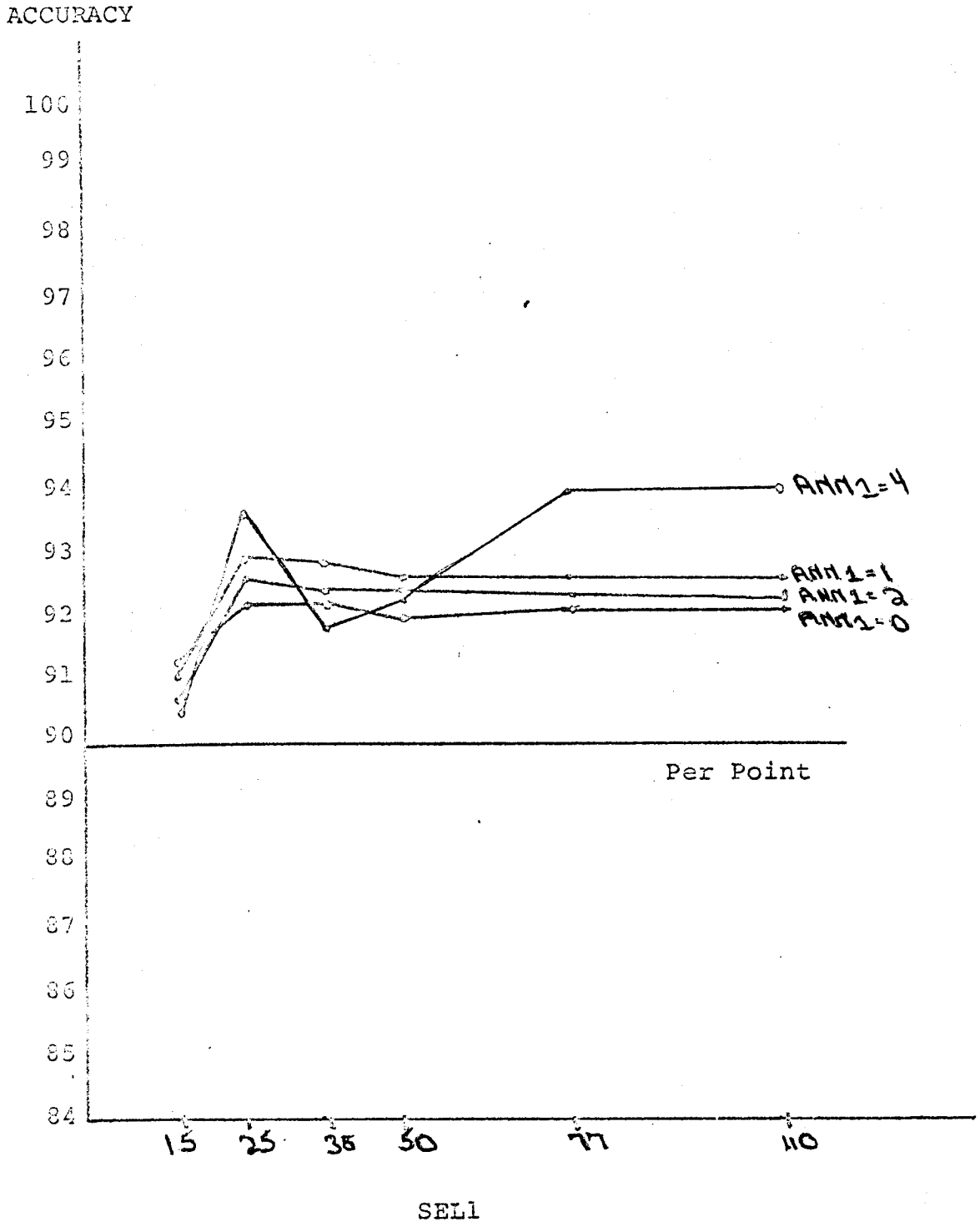
D. Software Documentation. Though the bulk of the software documentation will be produced upon completion of the programming tasks, a detailed flow chart of the restructured supervised processor has been produced.

E. User's Guide Development. The Introduction to the User's Guide was drafted during this quarter. The Introduction provides a general overview of the ECHO concept and a brief description of the difference between the supervised and non-supervised ECHO processors.

FIGURE 1

2 X 2 GRAHAM COUNTY FIELD CENTER PIXEL PERFORMANCE

SUPERVISED ECHO



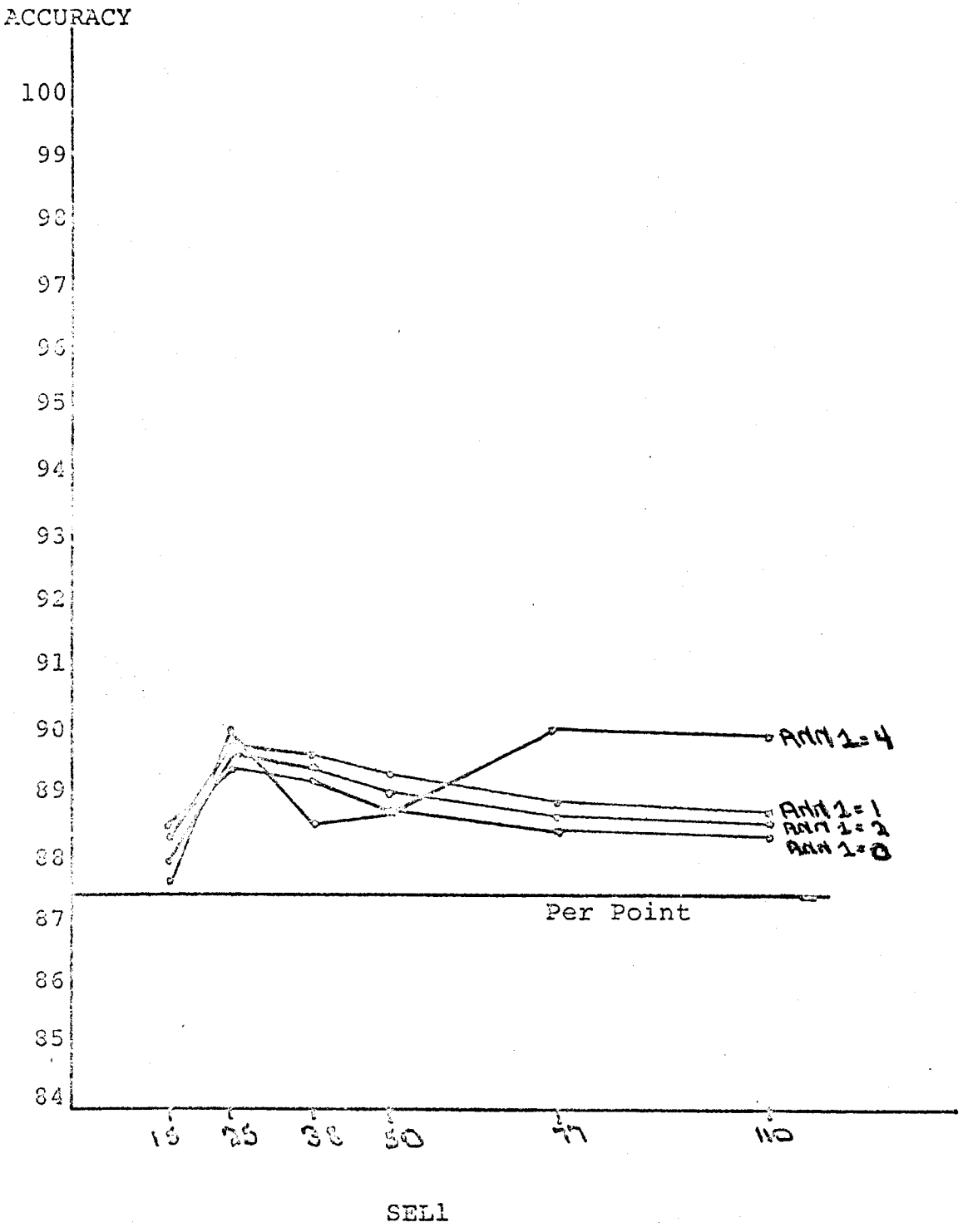
ANN1 = THE ECHO ANNEXATION PARAMETER

SEL1 = THE ECHO CELL SELECTION PARAMETER

FIGURE 2

2 X 2 GRAHAM COUNTY EXPERIMENTAL FIELD PERFORMANCE

SUPERVISED ECHO



ANN1 = THE ECHO ANNEXATION PARAMETER

SEL1 = THE ECHO CELL SELECTION PARAMETER

FIGURE 3

2 X 2 GRAHAM COUNTY TRAINING FIELD PERFORMANCE

SUPERVISED ECHO

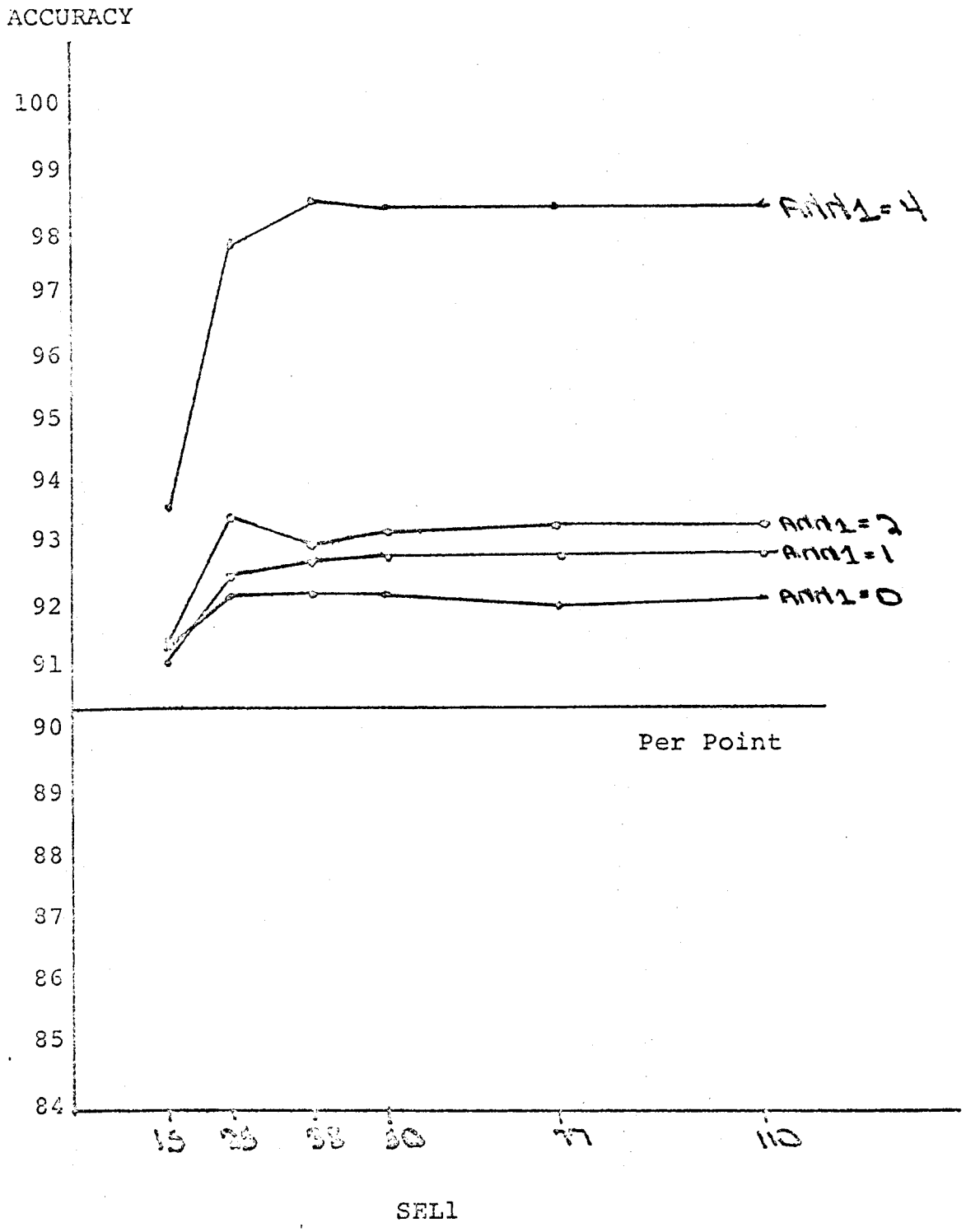


FIGURE 4

2 X 2 GRAHAM COUNTY CPU TIME

SUPERVISED ECHO

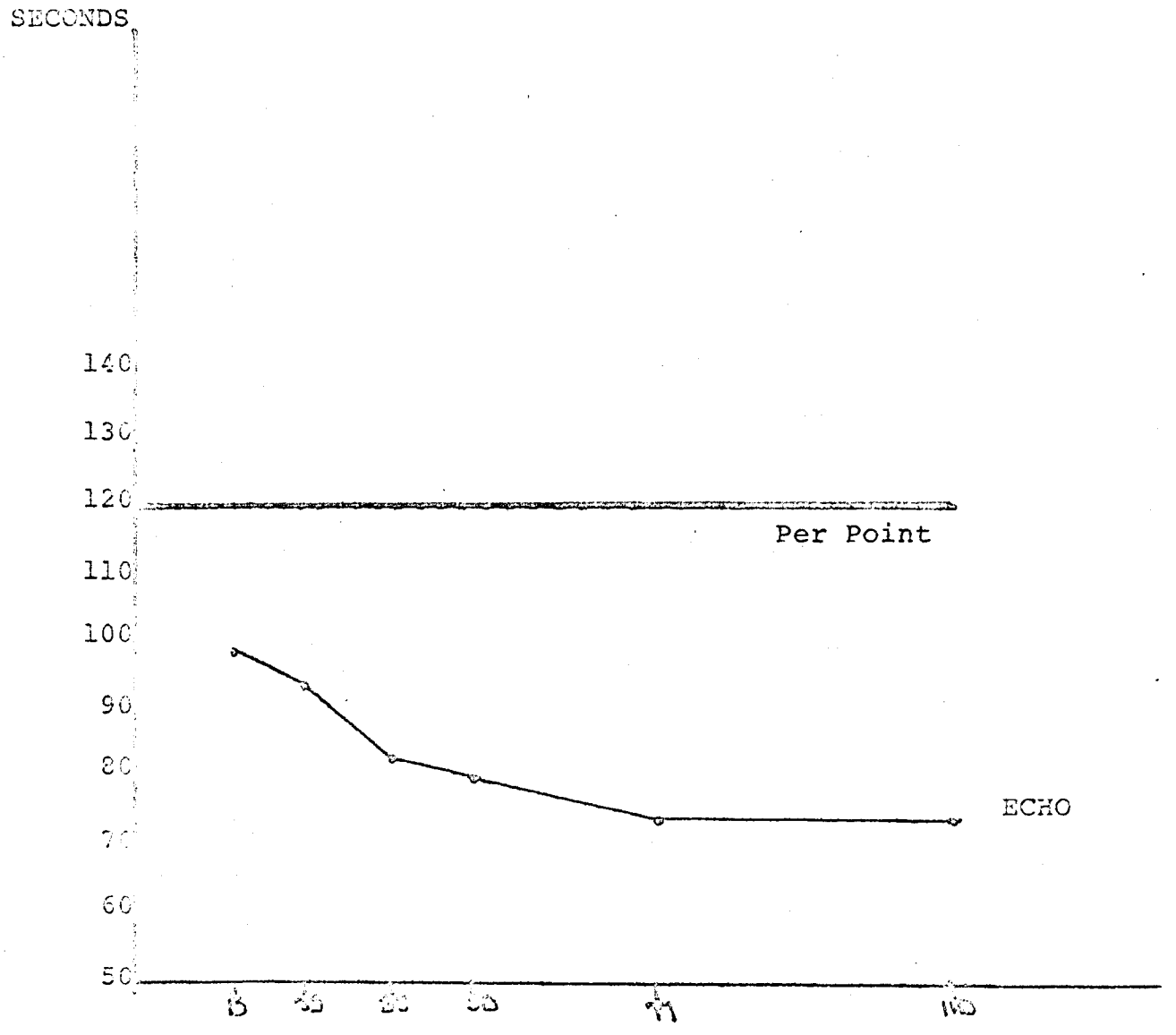
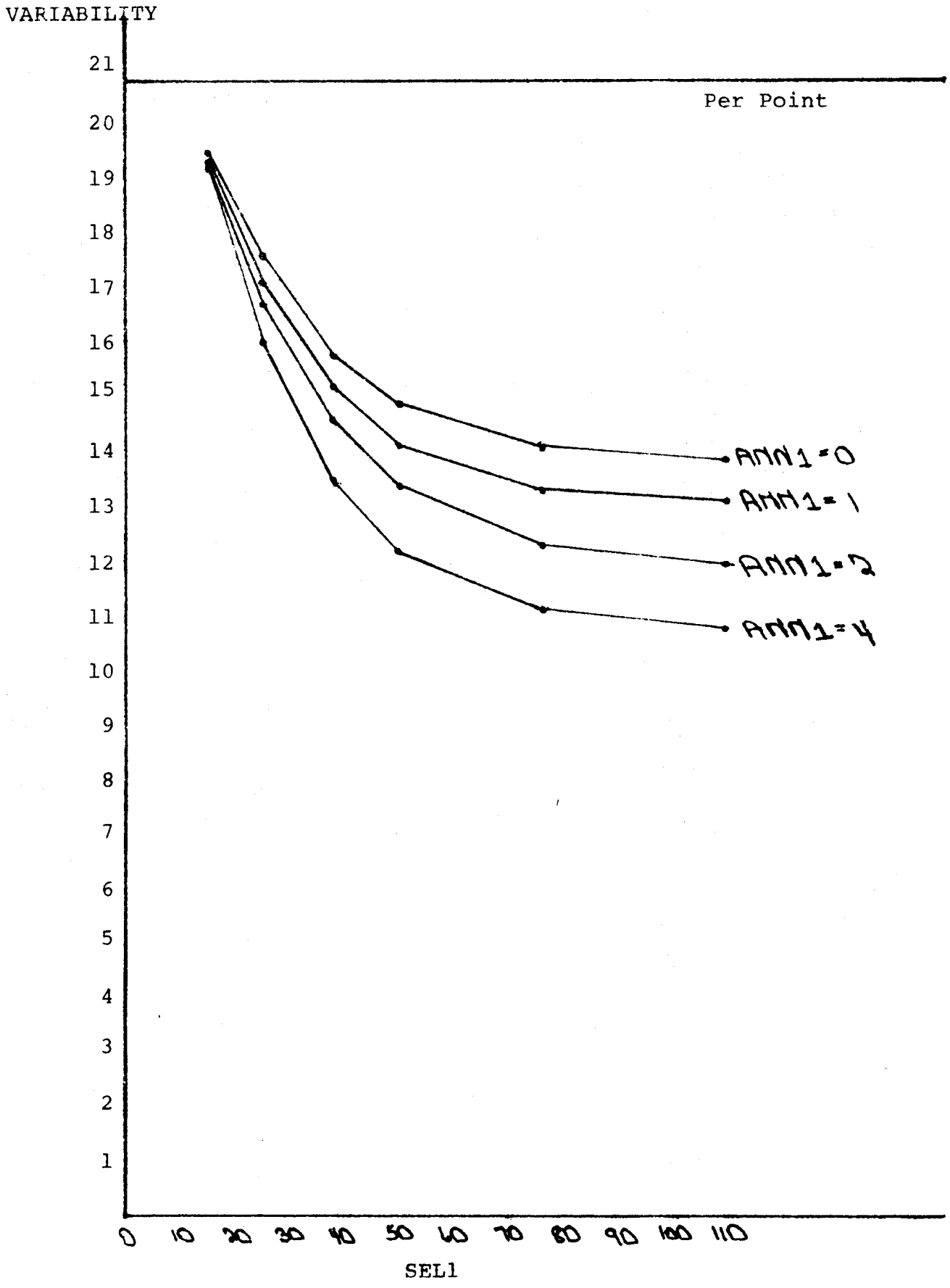


FIGURE 5

2 X 2 GRAHAM COUNTY VARIABILITY

SUPERVISED ECHO



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. Progress has been made in all three subtasks during this quarter, including a major modification in approach in subtask 2.2a, refinement of the advanced digitization methods in subtask 2.2b, and completion of the sample selection program in 2.2c. A revised implementation plan for task 2.2a was prepared and submitted to the contract monitor for approval.

During this quarter three members of the LARS staff supporting task 2.2 participated in a Large Area Crop Inventory Experiment (LACIE) Project Review. The LACIE Project Review was held at Johnson Space Center, Houston, Texas, from January 26 through January 28, 1977.

2.2a Stratification by Machine Clustering

As described in the second quarterly review, a reconsideration of our approach to partitioning has resulted after assessment of results obtained to date. A revised implementation plan was discussed with M. Trichel on January 28, 1977. His suggestions were considered and an amended implementation plan (February 7, 1977) has been submitted for formal approval.

The revised implementation plan calls for initiation of a more intensive investigation into the fundamental nature of the apparent strata in LANDSAT imagery. Such a study is essential to the development of the quantitative measures needed for an effective machine partitioning method.

The revised study has been limited to two frames of Landsat data due to time and personnel limitations. At present the choices have been narrowed down to scenes: 5428-16053, 5428-16060, and 2416-16362. The first two scenes were acquired on June 20, 1976, and the third was acquired March 13, 1976. These frames were chosen so that: a) digital data was available in-house at LARS; b) the number of LACIE segments in the area covered was maximized; c) the date of data acquisition was suitable with respect to the crop calendar; and d) static stratifications developed by the USDA for LACIE and by UCB were available for the area.

Task 1.1*-Preparation of Landsat Data. The full-frame digital MSS data for these LANDSAT frames was already available through LARSYS, but the MSS data for the LACIE segments was not received by LARS from

*Numbers refer to 2.2a Subtasks

JSC until the end of November. The LACIE segments to be used in conjunction with the full frames are being reformatted first and will be available the first week in March.

Task 1.2- Receipt of Landsat Imagery. Since these frames cover LACIE segments, the masters for the imagery exist at Salt Lake City. The imagery has been ordered and should be available at LARS early in March. At that time the reformatted segments will be located in the imagery (subtask 1.5) and two of the three frames will be chosen for stratification on the basis of the number of LACIE segments in the frame.

Task 1.3-Enhancement of Full-Frame Images. Comparison of the "standard" color composite produced with the digital display at LARS, the Landsat false color imagery, and LACIE PFC products has indicated major differences among these products. Since these images, with additional enhanced images, are needed for this study, it is important to be able to reproduce images consistently. Duplication of JSC Production Film Converter (PFC) false color imagery has been attempted with very good success. The present 'standard' LARSYS algorithm for generating CRT gray level imagery involves setting 16 levels over the whole data range into bins of equal probability within each channel. Channels 4, 5, and 7 are then displayed on the CRT screen and exposed on color Ektachrome film through blue, green, and red filters, respectively. The time of exposure is 6/15 second through the blue and the green filters and 16/15 second through the red filter at an f-stop of 9.5.

In order to generate imagery similar to the PFC imagery, data from LACIE segment 1111 in Rice County, Kansas, was histogrammed and

equal range levels were selected covering the data range of ± 3 standard deviations from the mean in each of the channels 4, 5, and 7. These levels for each channel were displayed on the CRT and photographed, using their respective color filters, at the standard exposure settings mentioned above, as recommended by the LARSYS Display Users' Guide.

Visual comparison of the resultant imagery to the JSC imagery indicated that the image was too dark and too green. To resolve these problems a series of test transparencies were produced in which f-stop and exposure length for individual filters were varied. The table below lists the parameters and how they were varied for each test as well as comments about the image quality.

	TEST *	1	2	3	4	5	6	7	8	9	10
	f	8	8	8	8	8	9.5	9.5	9.5	9.5	9.5
EXPOSURE*	BLUE	6	8	8	12	8	6	8	12	8	8
	GREEN	6	6	4	4	4	6	6	4	4	4
	RED	12	12	12	12	8	12	12	12	12	8
	COMMENT	1	2	1	3	1	1	4	5	5	5

*Exposures in fifteenths of a second.

1. too green
2. yellow green
3. little bluish; very close to PFC product
4. greenish yellow
5. too yellow

Test exposure number 4 was selected as closest to the imagery we were trying to duplicate. Although the transparency had more blue than the JSC image, the bright bare areas showed distinctive soil patterns and there appeared to be more crop condition discriminability.

A decision soon will be made regarding the enhancement techniques to be used in this stratification study. The techniques being considered initially are principal components, the "Tasselled Cap" transformation, and ratioing. These techniques have previously been implemented on the LARS computer system.

Task 1.4-Partitioning. The Kansas portion of the partitioning of the United States, developed by USDA personnel at JSC and to be used in the sample allocation for Phase III of LACIE, was delivered to staff of LARS during the LACIE Project Review, January 26-28, 1977. These strata have been digitized for quantitative assessment. The static stratification of Kansas developed by UCB in FY76 also will be digitized and evaluated. The portions of these partitions covering the frames in Kansas will have been digitized by the end of February.

Beginning the first week in March, three LARS analysts will stratify the two frames chosen at the completion of subtask 1.2. They will be using both the Landsat imagery and the digital display facilities at LARS.

During this quarter, much of the personnel and resources for Task 2.2a were directed to Objective 1, the modified approach to stratification which has been described above. As a result, there has been no significant progress on the subtasks associated with Objective 2, the physical factor analysis. To date the difficulties with the temperature variables described in the second quarterly report have not been completely corrected and so the regression analysis is not completed.

2.2b Digitization and Registration of Ancillary Data

Work in the third quarter was concentrated on Tasks 6 and 7 which involve development of advanced map digitizing methods. Tasks 2, 3, 4 and 5 were completed in the last quarter and no discussion of these is included here. Task 8 concerns the evaluation of coordinate systems for ancillary data storage and Task 9 is the determination of resources needed for ancillary data digitization. Tasks 6 and 7 are discussed in detail below with discussion of the limited activity on Task 8 and 9 following. Tasks 7, 8 and 9 will be continued through the fourth quarter. Task 6 was completed on schedule.

Tasks 6 and 7 - Advanced Digitization Technique Development. An advanced technique in which map digitization is performed by a flying-spot scanner was investigated in the second quarter of CY77. It was found that the color-coded features on a land use map of Kansas can be reasonably accurately recovered by simple maximum likelihood classification of the scanned color map photograph. There is nearly no error in polygon-centers but numerous misclassifications are observed along field boundaries. In this quarter, methods of improving boundary accuracy were studied.

Use of a Neighbor-Aware Classifier. The land use map consists of color-coded features which are separated by boundaries drawn in black ink, thus boundaries appear to the scanner as a mixture of colors which are easily misclassified. If the classifier is allowed to consult neighbor pixels, accuracy may be improved. The LARS ECHO classifier is a strengthened maximum likelihood classifier whose decision logic is partly based on the likeliness that pixels in a cell belong to the same population.

Classification results using ECHO with 4-pixel cell size show slight but not satisfactory improvement of classification of boundaries. Further adjustment of the selection parameter (the threshold deciding that pixels within a cell come from the same class) still does not yield sufficiently improved results.

Use of an Edge-Detector Post-Processing Approach. An alternate approach to improve boundary recovery is to clean up misclassifications after the maximum likelihood classification. A problem associated with this type of approach is that merely looking at a pixel at a time is not sufficient to tell that it is a misclassified pixel or not. For the purpose of cleaning the boundary, it is assumed that misclassification occurs mainly along boundaries. To locate a boundary, a simple derivative-type edge-detector is used. The algorithm of such a detector is presented at the end of this discussion.

One-dimensional Cleaning Process. The following discussion considers a situation in which a boundary is north-south or vertical, and the direction of cleaning is west-east or horizontal. If no misclassification has occurred at the boundary, there should be a clear-cut stepwise transition of pixel class, thus the output of the edge-detector should show a jump of one pixel width at the boundary. However, if some pixels are misclassified to some other classes, other than those on both sides of the boundary, the output of the detector should show consecutive jumps of total width longer than one pixel. The objective of cleaning is to make such a "thickened" boundary "thin". To do this, we re-assign the pixels on the "thickened boundary" to the classes of the pixels before and after the jumps according to some criterion. For example, a simple half-and-half re-assignment may serve this purpose. The figure below

is an illustration of the cleaning process and the algorithm is listed at the end of this section.

```
                              1     2
                              ↓     ↓
            scan line:        AAAAAAABECCCCDDDDD
edge-detector output:         100000011100010000
       cleaned line:          AAAAAAAACCCCCDDDDD
```

Illustration of the 1-dimensional cleaning process. A "thickened" boundary occurs at position 1, pixel B is re-assigned as A, pixel E is re-assigned as C. A "thin" boundary occurs at position 2; no re-assignment takes place.

Cleaning Process for the Classification Map. It is obvious that the 1-dimensional cleaning process is effective only when the boundary is not parallel to the direction of cleaning. But the classification map requires a 2-dimensional cleaning process. However, a direct extension of the 1-dimensional method to a 2-dimensional case is handicapped by some problems, such as detection of parallel-running edges; thus a more complex decision scheme is required. A simple but effective method to avoid these difficulties is to apply the 1-dimensional cleaning processing horizontally and then vertically. A major disadvantage of this method is that it requires the interchange of line and column and may be difficult to accomplish if the size of the map is too large.

Accuracy and Resolution. It should be noted that the cleaning process is nothing more than a re-assignment of misclassified pixels based on an estimate of the positions of boundaries. Because the color-coded features are well separated by distinct boundaries that can still adequately reappear after maximum likelihood classification, the estimate

of boundary position by the edge detector is very likely to be sufficiently accurate. It is subsequently expected that pixel re-assignment of boundaries is reasonably accurate. Even with this, the uncertainty of the position of the cleaned boundaries is not greatly affected, hence resolution remains unimproved. As a matter of fact, resolution is expected to be lowered because the re-assignment may move the boundaries slightly. How much a boundary can possible be moved depends on a parameter L in the 1-dimensional cleaning process. This parameter is the length of consecutive pixels of the same class that are regarded as correctly classified. For example, if L=2, then in the direction of cleaning, a new class one pixel in length is considered correctly classified and will not be re-assigned. Resolution is lowered accordingly.

Results and Discussions. A small area of 190 lines by 211 columns from the maximum likelihood classification map was processed with the 1-dimensional cleaning process horizontally and then vertically. Results show that boundaries are more distinct and many of the misclassifications at the boundaries are removed. Furthermore, it is found that some other misclassified pixels in field centers--especially those corresponding to marks and lettering on the map are removed. The general classification quality is significantly improved.

A drawback that has been observed is that certain misclassifications are enhanced, especially those with large areas; it is expected that increasing the parameter L in the cleaning process may remove this problem. Since the sampling interval is only half of the diameter of the spot size of the scanner, and the parameter L is set to 2, there is no loss in resolution power. A comparison of the original per-point classification of the land use map and the cleaned version is presented in Figures 1 and 2. Significant improvement in edge and interior cleanness is apparent.

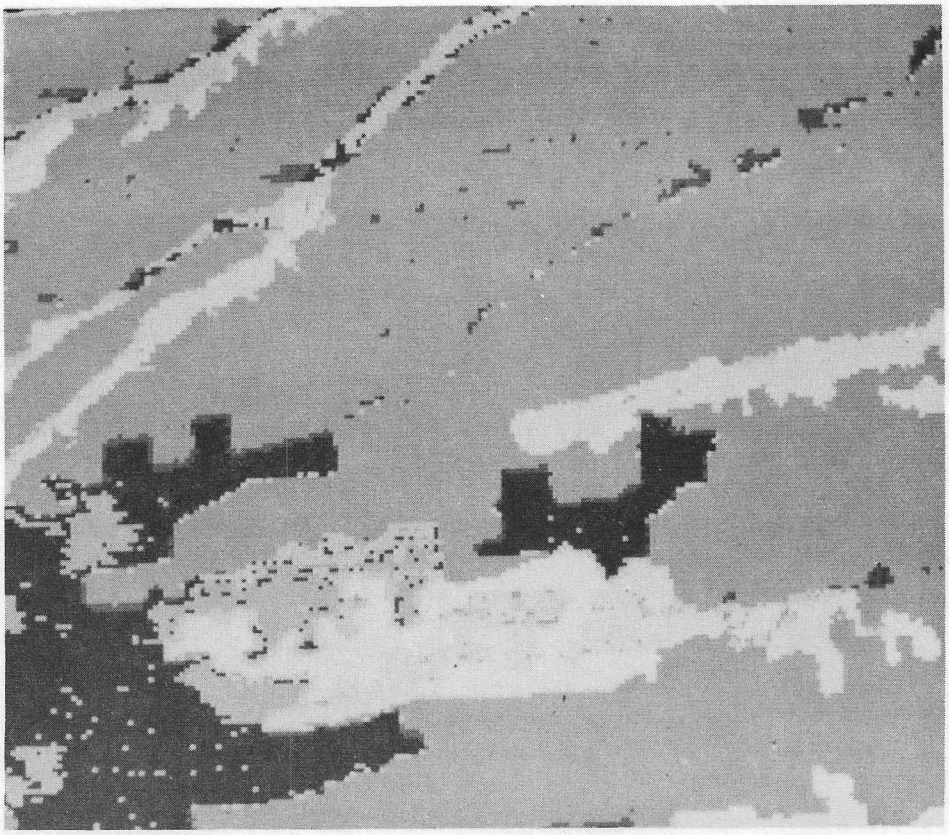


Figure 1. Per-Point Classification of Digitized Land Use Map of Kansas

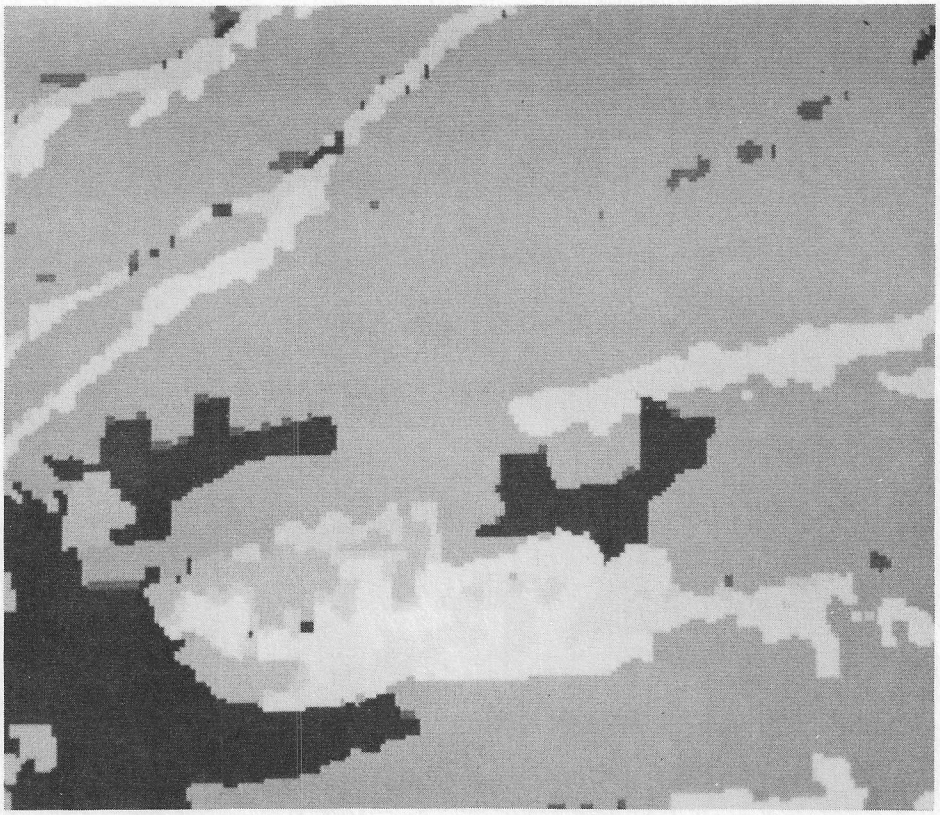


Figure 2. Post-Processed Classification Using Cleaning Algorithm to Remove Errors

The method implemented and evaluated appears to show promise for improving the efficiency of color coded map digitization. The activity planned for the last quarter will be to further test the method of cleaning and document the work fully.

Algorithm for a Simple Edge-Detector

Input:	a line of pixels of length NN in array PIC
Output:	the derivative of above in array IG of dim NN
Step 1:	Set IG(1)=1 (a convention)
Step 2:	Do step 3 from K=2 until NN
Step 3:	If PIC(K) differs from PIC(K-1), then IG(K)=1, otherwise IG(K)=0

The 1-Dimensional Cleaning Algorithm

Step 1:	Define L=length of consecutive pixels that are regarded as correctly classified
Step 2:	Do step 3 through step 10 from first line to last line
Step 3:	Get output from the edge-detector in IG
Step 4:	From IG, search a consecutive block of zeros of length longer than L-1 pixels
Step 5:	If search is unsuccessful, go to step 6, otherwise go to step 7
Step 6:	(only left side of a boundary found) re-assign pixels according to left side of the boundary. Go to step 10.
Step 7:	Test width of boundary. If equals 1, go to step 9 (i.e., no re-assignment)
Step 8:	(a "thickened" boundary found) If left side of boundary lies outside the line, re-assign pixels according to right side of the boundary. Otherwise re-assign according to half-left-half-right criterion.
Step 9:	Test end of line. If yes, go to step 10. Otherwise, go to step 3.
Step 10:	Continue.

Task 8 - Data Base Coordinate System Study. Further consideration was given to the question of pixel versus polygon storage methods for ancillary data and a plan was established for completing this task in the fourth quarter. The approach will be to study the details of the polygon digitizing and grid conversion process followed in CY76 for soil, land use, and temperature/precipitation map digitizing and use this information to evaluate the grid and polygon methods.

Task 9 - Evaluate Resources for Digitization of Ancillary Data. Further data on costs for map digitizing was accumulated during the quarter but no analysis was performed. This task will be focused on in the fourth quarter and completed by May 31, 1977.

2.2c Crop Inventory Using Full-Frame Classification

Task 1 - Classify Full-Frames. The required classifications which were completed by October 1 were all performed by a systematic random sample of pixels dispersed throughout a county using at least every fourth line and every fourth column of data (a one-sixteenth sample). Twenty-three counties were classified with a one-sixteenth sample and 56 counties were classified with a one-quarter sample (every other line and every other column of data). Due to an error recently discovered in the original classifications, only 79 rather than 80 counties were included.

Task 2 - Develop Alternate Sampling Plans. For the segment size of one pixel, the number of segments to be selected will be 120,393 rathethan 1,926,288 as given in the last quarterly report. Due to the constraint that some counties have available classifications including only one-sixteenth sample, it was decided that all segments chosen would be sampled on this basis to attain consistent precision among the segments. By taking a one-sixteenth sample of eachof the segment sizes (5x6, 4x4, and 2x2 nm), the total number of pixels sampled will range from about 115,000 to 120,000, the variation due to round off error because each segment is not an exact multiple of 4 columns and 4 rows in size. Taking a one-sixteenth sample of pixels as well yields the new figure of 120,393.

Task 3 - Locate Sample Segments and Compile Classification Results. A computer program to carry out the sample selection procedure has been written, debugged, and documented. This program is capable of handling any segment size and number of segments, but will be revised for selecting the pixel samples to improve the efficiency. It is anticipated that

operational sample selection and calculation of the wheat area estimates will begin before the end of this quarter. Wheat area estimates will be calculated as described in "The Use of Landsat Data in a Large Area Crop Inventory Experiment (LACIE)" by R.B. MacDonald, F.G. Hall, and R.B. Erb, reprinted from Proceedings of Symposium on Machine Processing of Remotely Sensed Data, June 3-5, 1975, Purdue University, West Lafayette, Indiana, pp. 1B-10 through 1B-11.

The computerized method of sample selection utilizes whenever possible the LACIE procedures (as outlined in LACIE-00200, Volume IV: Crop Assessment Subsystem (CAS) Requirements, Level III Baseline [Reference CCBD #III-0001, Dec. 16, 1974], pp B1-B6) for the allocation and location of sample segments. However, the procedure for allocating the number of segments per county was adjusted slightly to force the state total to be correct.

The LACIE constraints specify that a segment should not be selected if it falls entirely within a nonagricultural area. The nonagricultural areas defined by this investigation were not as carefully outlined as they were by LACIE, but did not include areas of less than 2x2 miles. The rationale for this arises from results found in another investigation: in wheat area estimation for Kansas, only eight nonagricultural areas were excluded, but estimates of high accuracy and precision were obtained. Nonagricultural areas, including urban areas, reservoirs, and Federal lands, were defined from Kansas county maps by the use of a table digitizer at LARS and excluded from consideration as sample segments.

The LACIE constraints on the geographic proximity of segments were followed for the 5x6 nm segments, but were adjusted for the other segment sizes in the following manner:

Segment Size	Rectangle Considered	Segments Allowed in Extended Rectangle E-W	N-S
5x6	10.5 x 12	4	8
4x4	8.4 x 8	6	10
2x2	4.2 x 4	12	20

Constraints of this type will not be feasible to consider for the single pixel sample size.

At the present time, it is anticipated that sample selection and compilation of results will have begun before the end of the third quarter. The calculation of sampling errors should be completed by the end of March and comparisons of the results achieved by mid-April.

2.3 Field Measurements For Remote Sensing of Wheat

Field Measurements activities during this quarter have included: project leadership and coordination, data processing and evaluation, management of the data library, and data analysis.

Project Leadership and Coordination

Four representatives from Purdue/LARS attended the Field Measurements Project Review held at NASA/JSC on December 6-8. The discussion during the meeting centered on developing ways to strengthen the Field Measurements Project and to help the LACIE project. Inputs were received from the LACIE Field Measurements Project Manager, ERIM, and Texas A&M for changes to the draft project plan. However, LACIE data requirements for research, test, and evaluation to which the field measurements project can respond have not yet been received.

At the end of the quarter the data collection schedule for 1977 was prepared by NASA/JSC and LARS.

Data Processing and Coordination

During this quarter significant progress was made in completing the processing of 1974-75 and 1975-76 Field Measurements Project Data. The status of data is summarized in Tables 2.3-1 to 2.3-4.

Landsat MSS Data. Landsat collected over the test sites that were received during this quarter have been processed. The registrations of the 1975 Landsat-2 data collected over the Finney County and Williams County intensive test sites were completed. The two registrations include four good dates for Finney County, Kansas and two good dates for Williams County, North Dakota. See Table 2.3-1 for a listing of dates included.

Aircraft Scanner Data. Twenty-seven flightlines representing six dates of aircraft multispectral scanner data were converted from universal format to LARSYS format. Also, one date, the 6/20/76 MSS data collected over Hand County, South Dakota, was received to find if any of the data would be usable.

FSS Data. Processing and evaluation of the 1975-76 FSS data in the field averaged format was completed. The 1975-76 FSS data will be ready for distribution and analysis after it has been recalibrated for the spectral reflectance of the calibration panel determined by the truck spectrometers during 1976. (At present the 1975-76 FSS data is calibrated using the data collected in September 1975).

Preparation of the 1976 spectral reflectance tables for the helicopter calibration panel are almost complete. The tables are being prepared using the data collected by the two truck-mounted spectrometer systems, the Purdue/LARS Exotech 20C and the NASA/JSC FSAS systems, over the helicopter panels. As soon as the tables are completed, the 1976 FSS data will be recalibrated and the 1976 single scan format FSS data will be processed.

Overall processing of FSS data went smoothly. The most difficult and time consuming aspect is verification of entry of ground observation and other ancillary data problems which turned up involved field number changes during the middle of the 1976 growing season. This resulted in mis-correlation of ground observation data from the JSC tape-to-film logs and the ASCS data sheets.

LARS Spectroradiometer Data. During the second quarter, an analysis showed that the procedure being used to correct LARS Spectroradiometer data for analog tape recorded "WOW", experienced during field operation

was not adequate. A review of the procedure showed that the parameter associated with the analog reproduce unit was inaccurate. The solution required redigitization of the data with modified reproduce unit calibration inputs.

During January, data were redigitized and, as of February 16, all dates were reformatted and ready for analysis except August 1 and August 6, 1976. The reformatting was achieved using the solar port reflective calibration procedure. An ongoing analysis indicates that a modification to this procedure could improve data accuracy significantly. A decision is pending on implementation of the revised calibration procedure. Re-reformatting the data would not seriously delay processing completion, two weeks is estimated. Several small data sets were reformatted using the modified procedure to support the analysis task.

Processing of the 1976 Landsat-band radiometer crop canopy modeling data collected at Williston, North Dakota was completed and sent to ERIM and Texas A&M.

JSC FSAS Data. Modifications to library entry software necessitated by additions to the header record of the tape received from JSC were completed. All FSAS data received as of February 16, 1977 have been stored in the library and distributed. This includes data collected at the Garden City, Kansas Agriculture Experiment Station between March 30 and June 12, 1976. No problems were encountered in handling FSAS data.

Table 2.3-1. Dates Included in 1975 Intensive Test Site Landsat-2 Data Registrations

Finney County, Kansas	Williams County, N. Dakota
3/20/75	7/20/75
6/15/75	8/15/75
7/06/75	
8/11/75	

Table 2.3-2. 1975 Data Processing/Reformatting as of February 16, 1977.

Instrument/Data Type	Completed & In Library	In Processing
Landsat MSS		
Whole Frame CCT	20 Frames	0
ITS Multitemporal Registration		0
Aircraft MSS	19 Dates-149 Runs	0
Helicopter Mounted Field Spectrometer (FSS)		
Field Averages	19 Dates-2251 Runs	0
Individual Scans	19 Dates-35,000 Runs	0
Truck Mounted Field Spectrometer		
FSAS	6 Dates-65 Runs	0
Exotech 20C	24 Dates-1577 Runs	0
Exotech 20D	45 Dates-645 Runs	0

Table 2.3-3. 1976 Data Processing/Reformatting as of February 16, 1977.

Instrument/Data Type	Completed & In Library	In Processing	At JSC
Landsat MSS			
Whole Frame CCT	61 Frames	0	N.A.
ITS Multitemporal Registration	0	3+	1 Dat
Aircraft MSS	16 Dates-97 Runs	0	1 Date
Helicopter Mounted Field Spectrometer (FSS)			0 Date
Field Average	7 Dates-391 Runs	20 Dates-1743 Runs*	
Individual Scans	7 Dates-10,000 Runs	20 Dates	
Truck Mounted Field Spectrometer			
FSAS	17 Dates-253 Runs	0	4 Dates
Exotech 20C	0	19 Dates-2700 Runs	N.A.

\+ We are reviewing whether the data registered for LACIE would be appropriate.

* Data has been reformatted; will be in library when recalibrated.

N.A. Not Applicable

Table 2.3-4. 1977 Data Processing/Reformatting as of February 16, 1977.

Instrument/Data Type	Completed & In Library	In Processing	At JSC
Landsat MSS			
Whole Frame CCT	14	0	N.A.
Aircraft MSS	2 Dates-10 Runs	1 Date	2 Dates
Helicopter Mounted Spectrometer (FSS)			1 Date
Field Average	0	4 Dates	
Individual Scans	0	4 Dates	
Truck Mounted Field Spectrometer			
FSAS	0	0	3 Missions

N.A. - Not Applicable

Data Library

Computer tapes containing all of the 1974-75 FSS single scan format data were distributed to researchers at ERIM and Texas A&M University. Also, the 1976 North Dakota crop canopy modeling data were distributed to researchers at ERIM. All data received from JSC during the quarter have been cataloged and appropriate data distributed to Texas A&M and ERIM.

The preliminary versions of Volumes 2 and 3 of the Field Measurements Data Catalogs for the 1975-76 and 1976-77 crop years, respectively were distributed to interested users.

Data Analysis

During this quarter analyses of field measurements data for agronomic and geometric objectives were pursued.

Agronomic Data Analysis. The statistical analyses of this investigation are being pursued using both the Landsat bands and the proposed thematic mapper bands. Data is being analyzed throughout the growing season using individual dates and combinations of dates from several LACIE biophases. This approach will permit comparisons to be made concerning the effectiveness of different wavelength bands and different observation times on the detection and discriminability of crops and crop conditions.

Analysis of the truck-mounted spectrometer data taken during 1974-75 at the Garden City, Kansas and Williston, North Dakota experiment stations is now well underway. The spectral separability of wheat and other small grains is being assessed by analysis of variance and discriminant analysis. The effects on spectral response of various experimental treatments such as irrigation, fertilization, and crop variety are being examined by

statistical techniques including analysis of covariance with time as a covariate when time appears to influence spectral response. During the next quarter, the same objectives will be continued as well as additional analyses relating spectral reflectance to agronomic measurements using regression and correlation analysis.

Analysis of the S-191 data taken at the Finney County, Kansas Intensive Test Site during 1974-75 was begun. Several difficulties were encountered in the initial stages of the analysis, particularly inconsistencies in field identification information.

Currently several analyses are underway with the helicopter data including examination of changes in the spectral response of a crop over the growing season (Figures 2.3-1 to 2.3-6) and variation in spectral response within a field. The examination of these differences which began with visual interpretation of data plots, will be quantified by statistical analysis.

Geometric Analyses. A major study of measuring the geometric characteristics of crop canopies was completed this quarter. The procedure and results are briefly summarized below; a complete report of the work will be submitted in early March.

A measurement technique is needed which is capable of providing timely information concerning the geometric characteristics of a vegetative canopy, the location and orientation of its foliage. Such data is required as input to many models for the radiation regime in a canopy. Therefore, this work:

1) proposes such a technique, designated as the "laser technique",
2) demonstrates the feasibility of the technique, and
3) offers suggestions for the implementation of the technique.

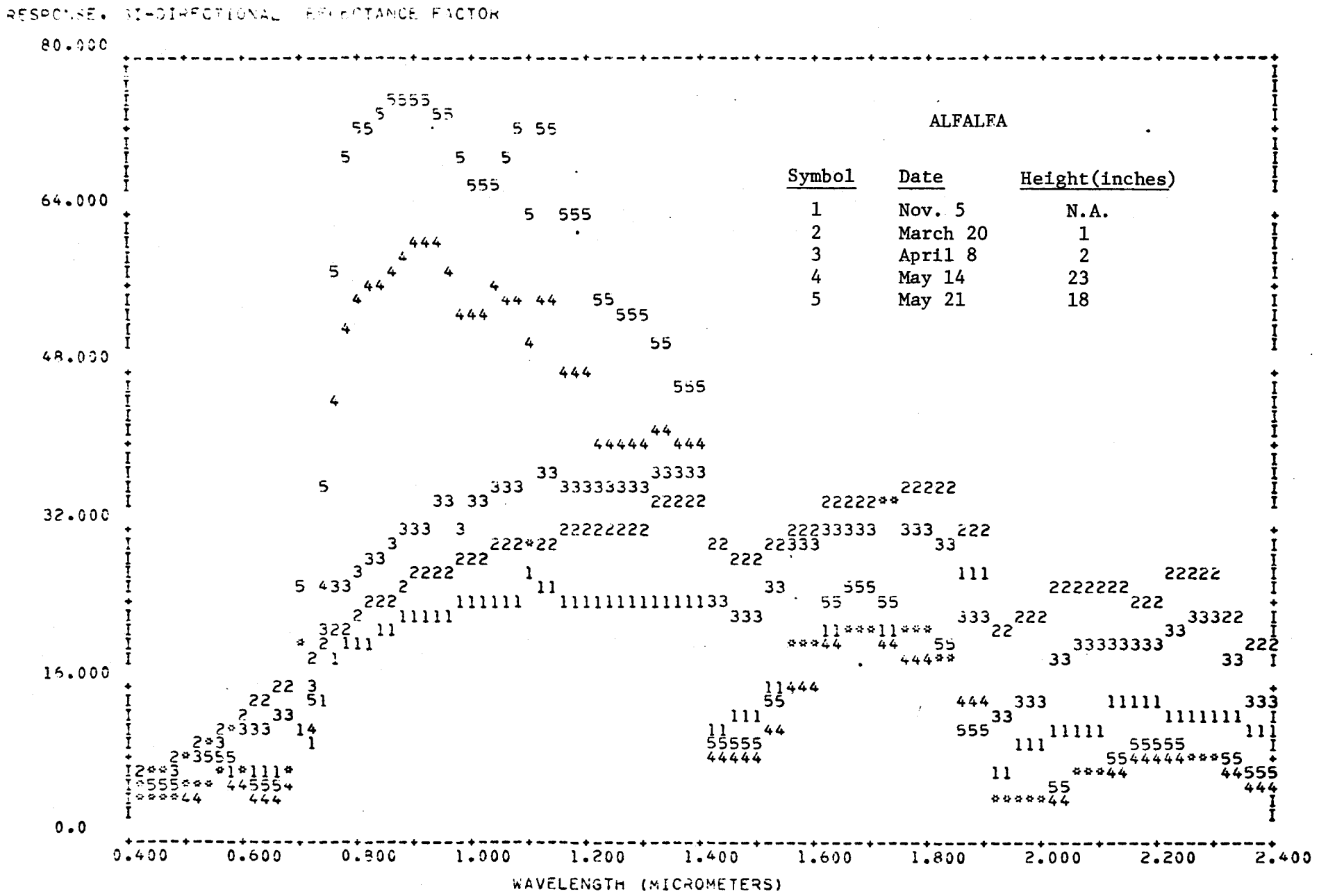


Figure 2.3-1. Spectral reflectance of alfalfa on five dates between November 5 and May 21.

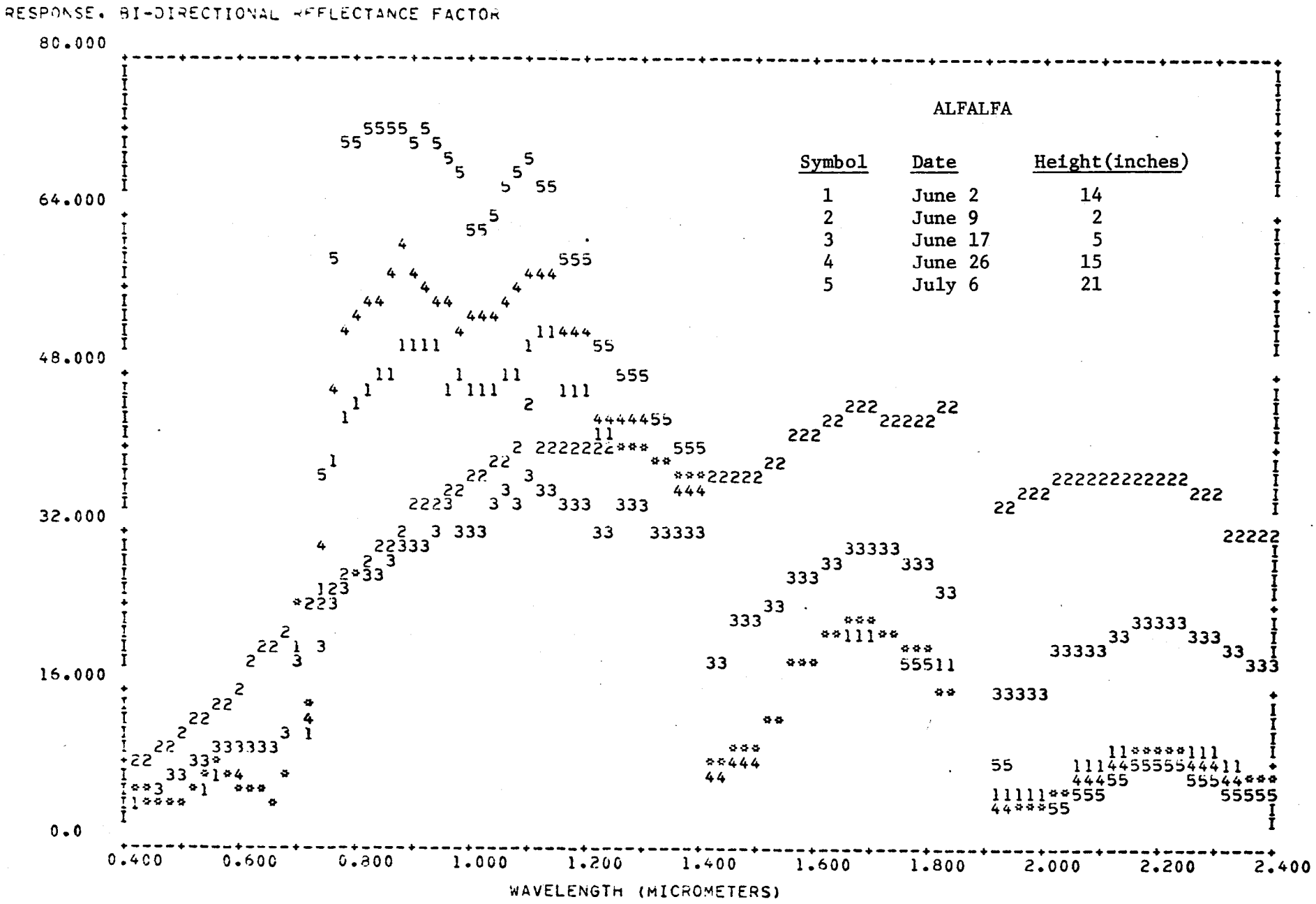


Figure 2.3-2. Spectral reflectance of alfalfa on five dates between June 2 and July 6.

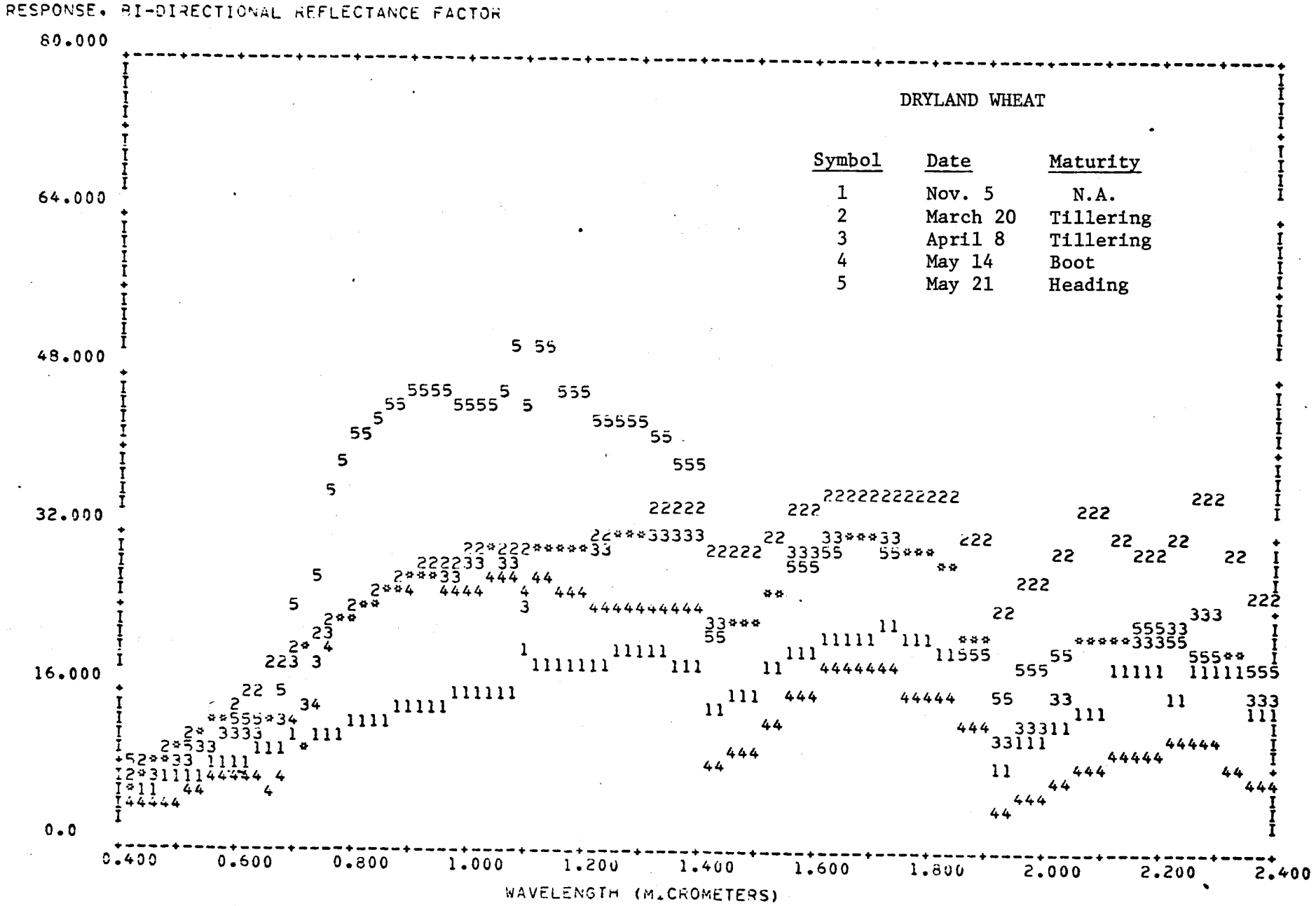


Figure 2.3-3. Spectral reflectance of dryland wheat on five dates between November 5 and May 21.

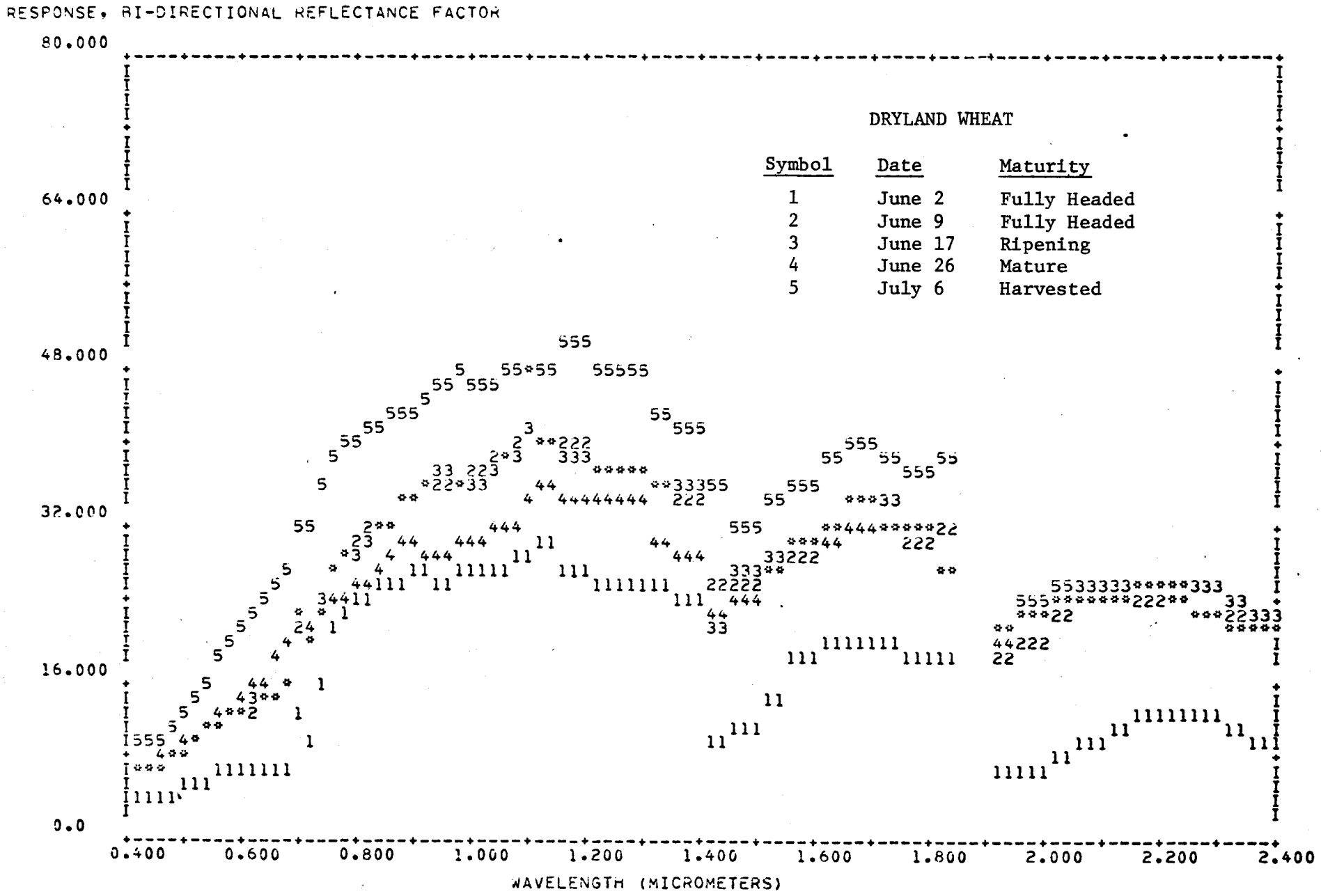


Figure 2.3-4. Spectral reflectance of dryland wheat on five dates between June 2 and July 6.

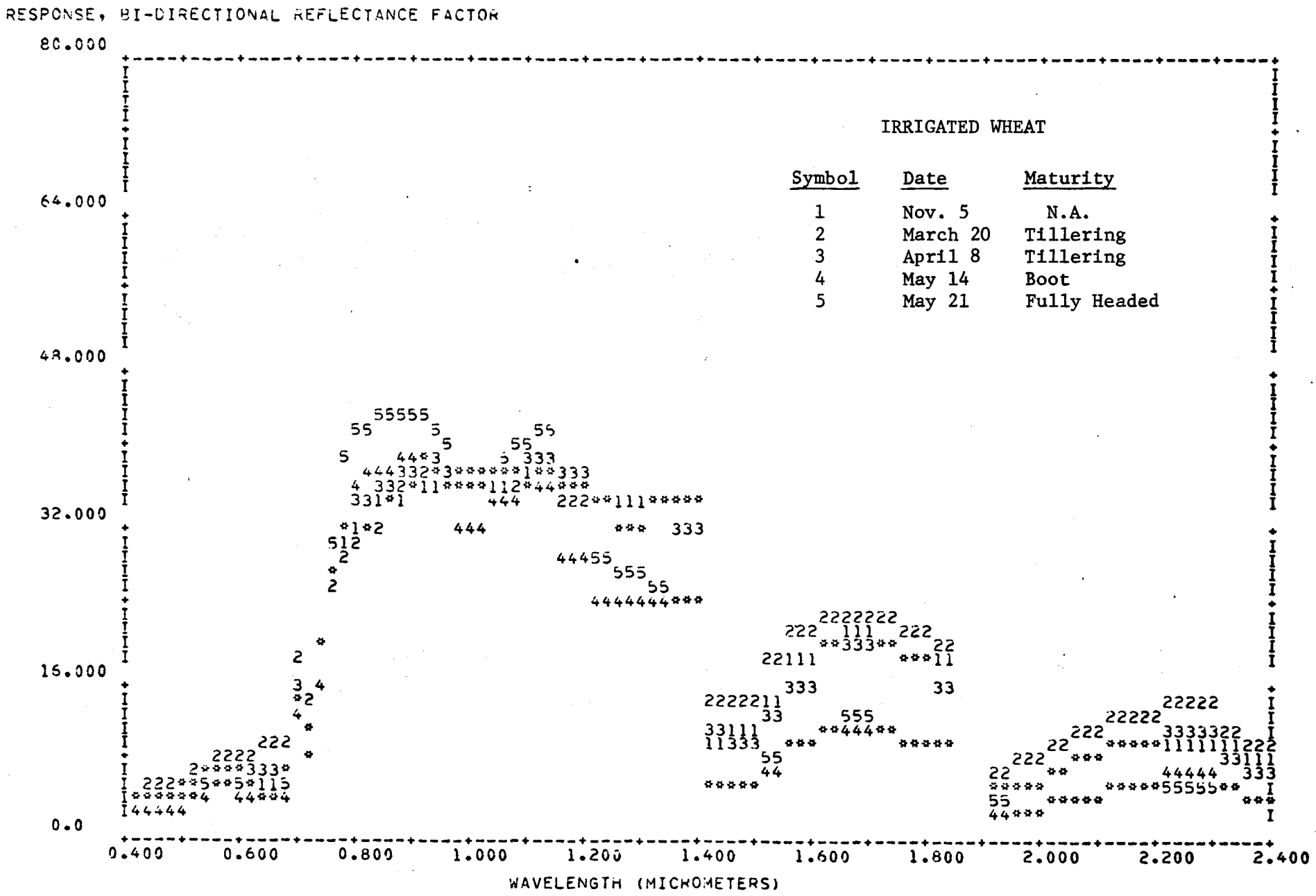


Figure 2.3-5. Spectral reflectance of irrigated wheat on five dates between November 5 and May 21.

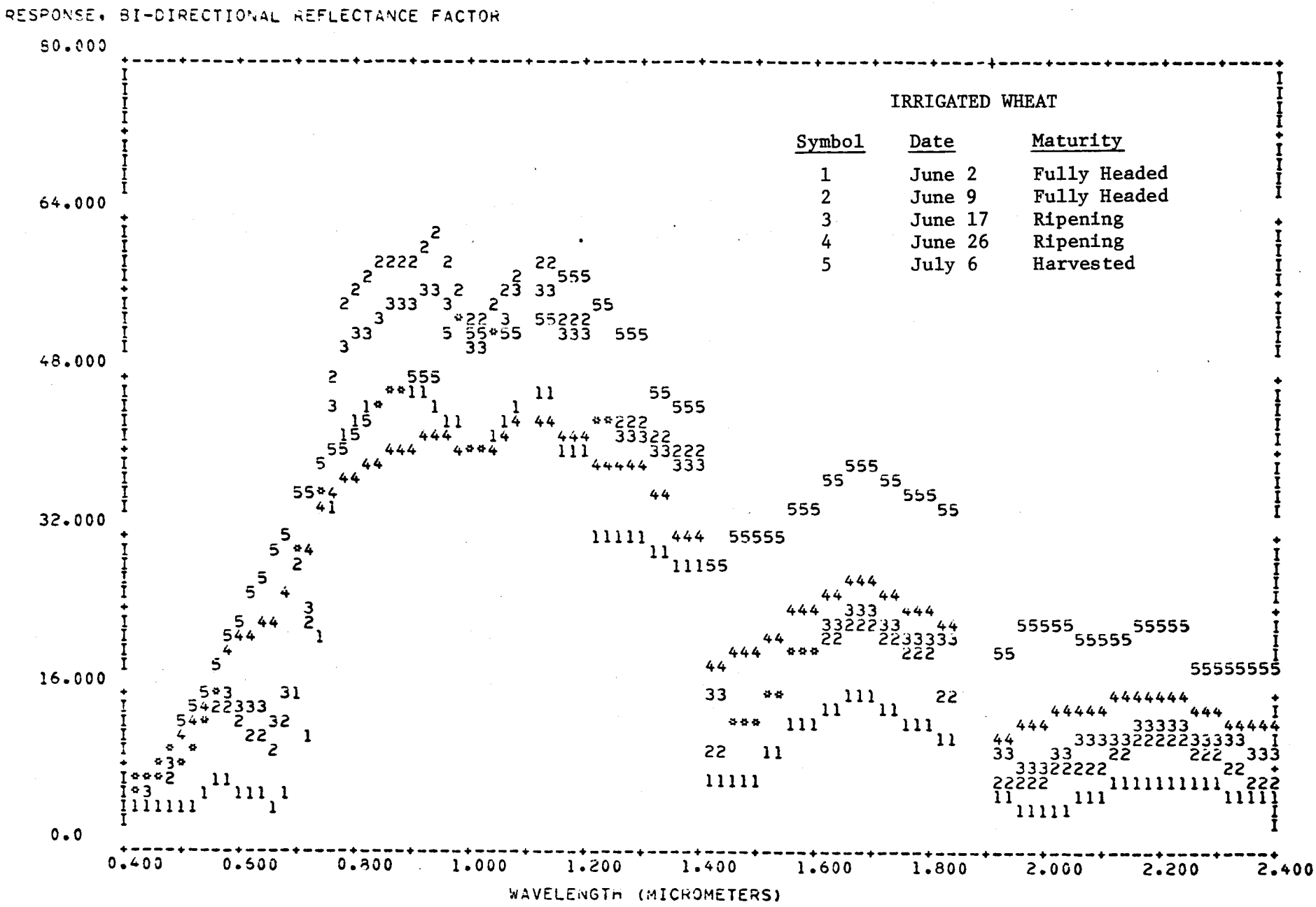


Figure 2.3-6. Spectral reflectance of irrigated wheat on five dates between June 2 and July 6.

Basically, the method, a variant of the point quadrat method, involves aiming a collimated light beam of very small cross section at a canopy and measuring the height at which the beam first hits a component of the canopy. Lasers are particularly well-suited to provide the small, intense beam required.

Several kinds of information can be obtained using the laser technique. Two are examined. First, the interception of solar power by the canopy is investigated as a function of solar zenith angle (time), component of the canopy, and depth into the canopy. Second, the projected foliage area, cumulative leaf area, and view factors within the canopy are examined as a function of the same parameters.

Feasibility of the proposed method is verified using data obtained from two vegetative crop canopies, wheat (Triticum sestivum L.) and corn (Zea mays L.). The key aspects of the technique and results are illustrated in Figure 2.3-7 to 2.3-11. Two systems are proposed that are capable 1) of describing the geometrical aspects of a vegetative canopy and 2) of operation in an automatic mode. Either system would provide sufficient data to yield a numerical map of the foliage area in the canopy. Both systems would involve the collection of large data sets in a short time period using minimal manpower.

Plans for Next Quarter

Processing of the 1975-76 FSS, Exotech 20C, FSAS, and aircraft MSS data will be completed early in the quarter and the data will be distributed to researchers at Texas A&M University and ERIM. Processing will begin on the Fall 1976 FSS data. The Field Measurements Project for 1977 with the corrections and additions will be printed and distributed.

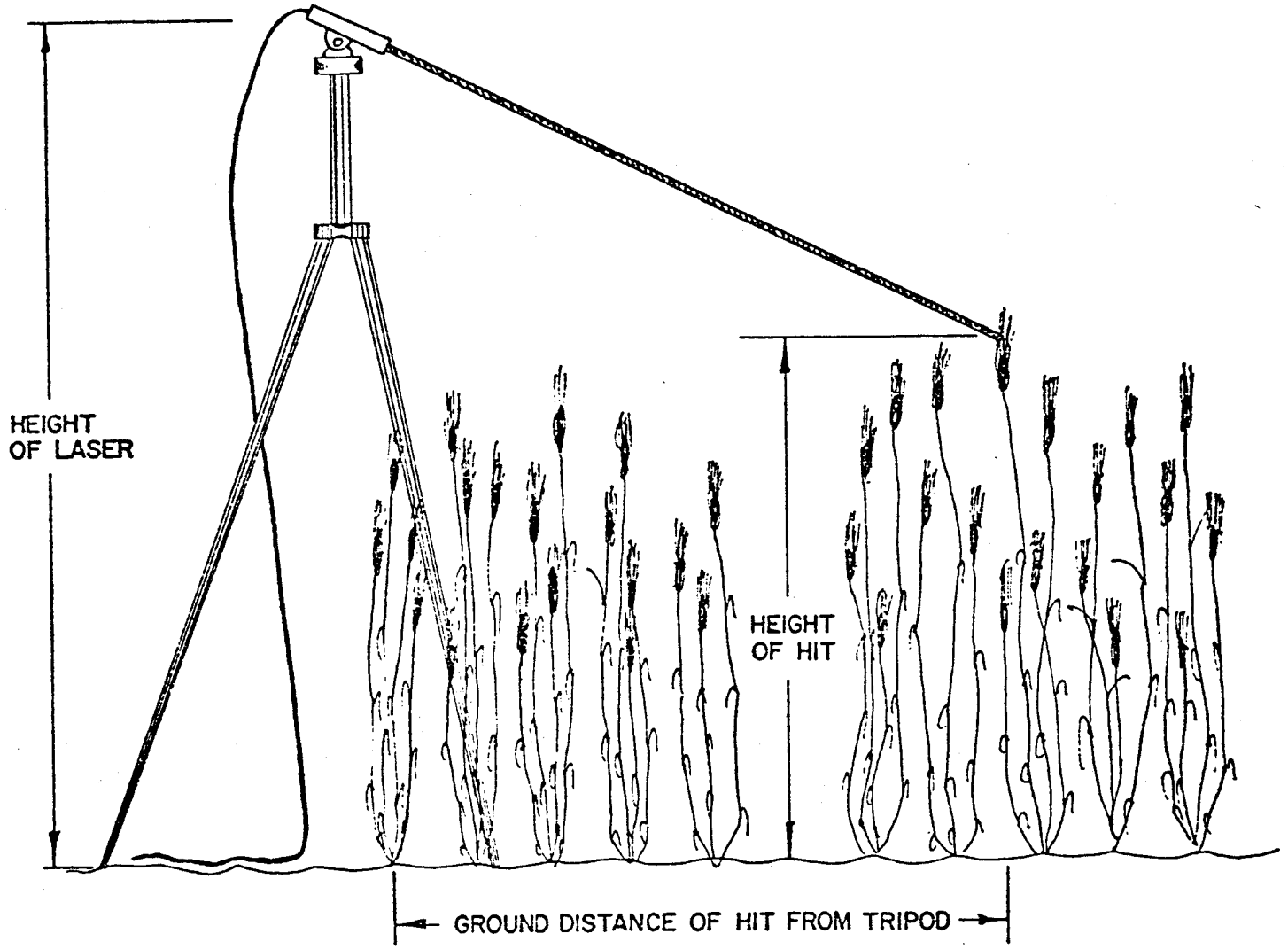


Figure 2.3-7. The laser technique was implemented in a wheat canopy. The height of the first impact of the laser beam with the canopy (a 'hit') was measured for each hit as well as the ground distance of the hit from the laser.

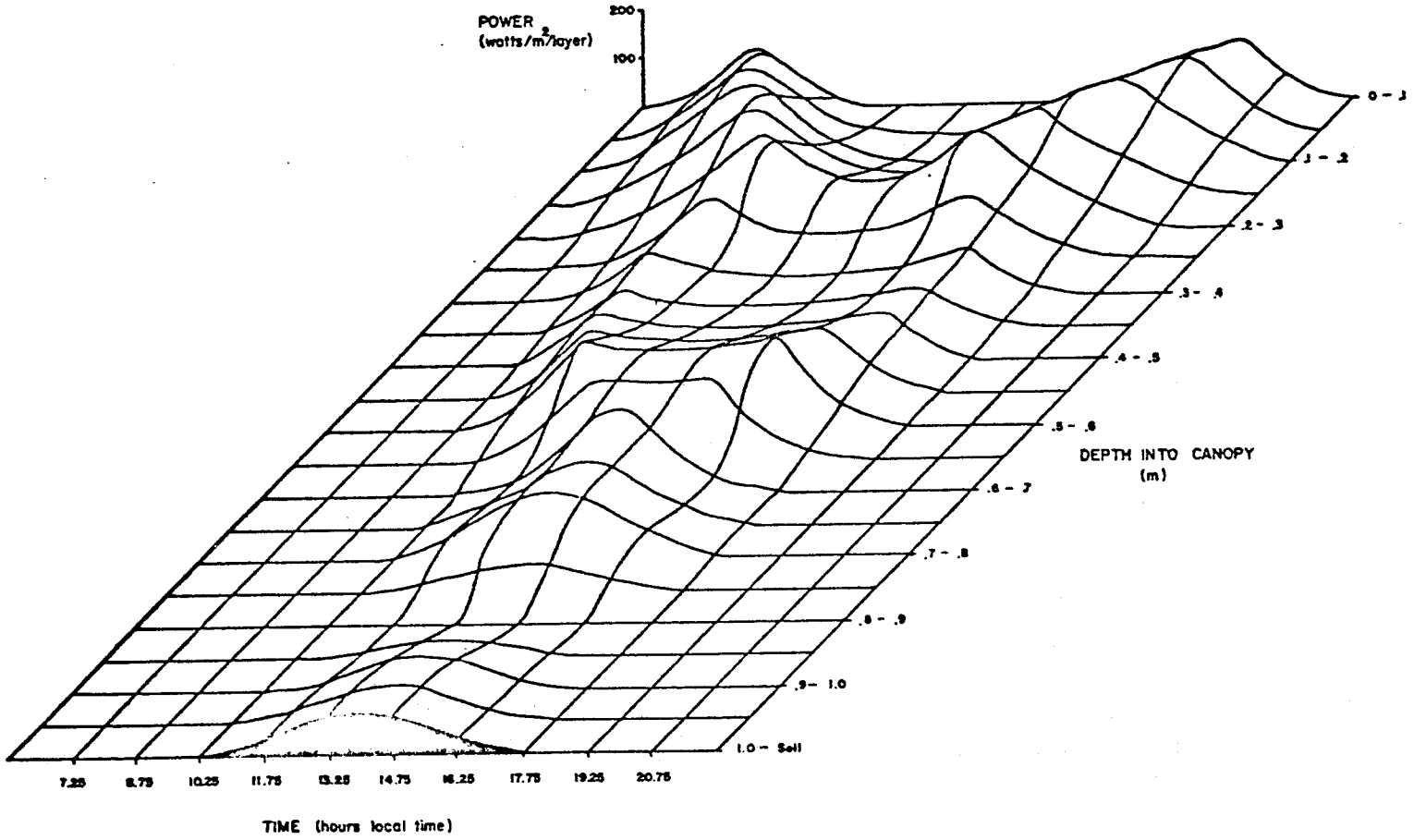


Figure 2.3-8. Solar flux, ABik, intercepted by foliage (or soil) in each layer of the wheat canopy for 30 July 1975.

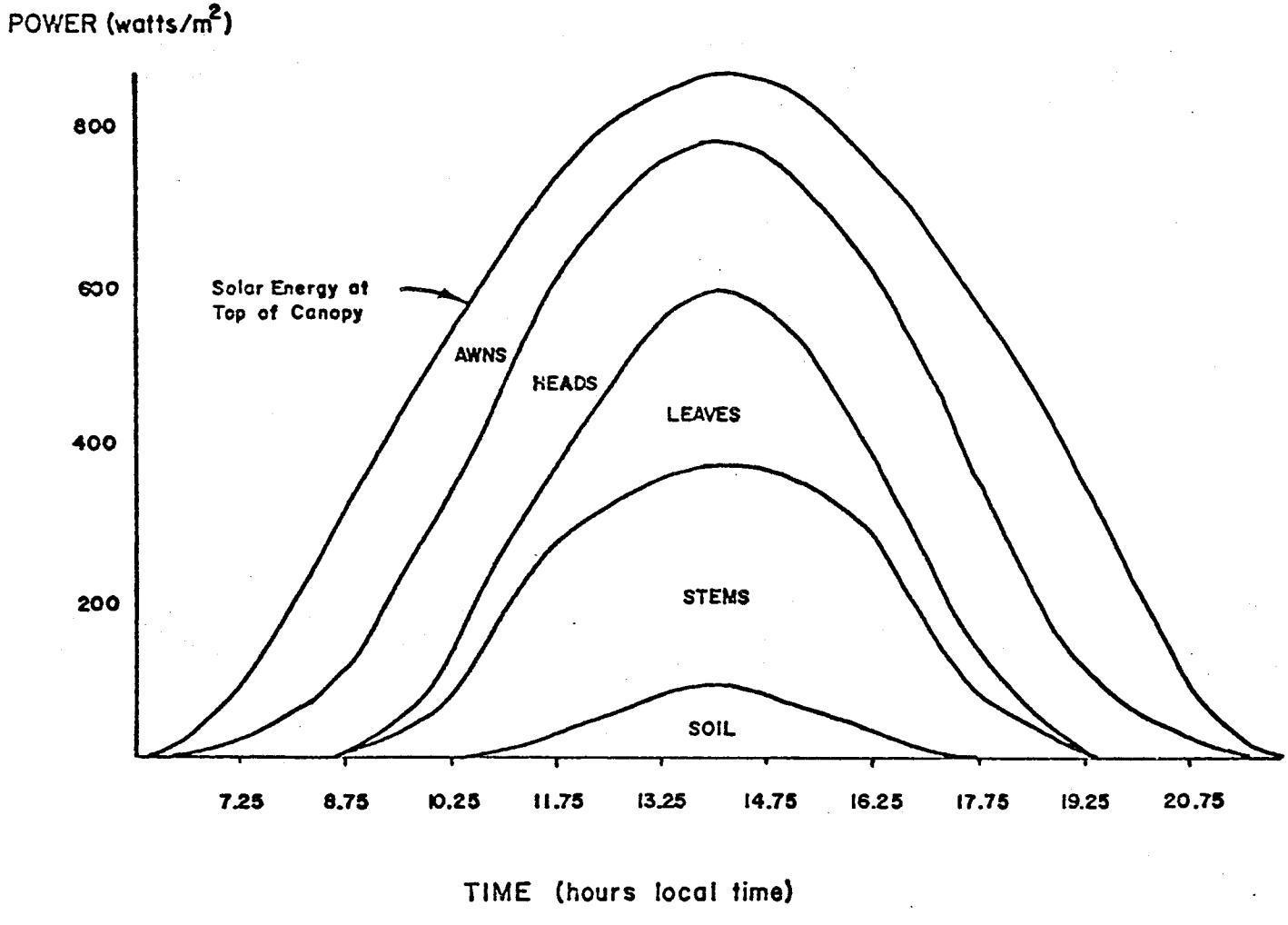


Figure 2.3-9. Solar flux, ABikl, intercepted by each component in each layer in the wheat canopy for 30 July 1975.

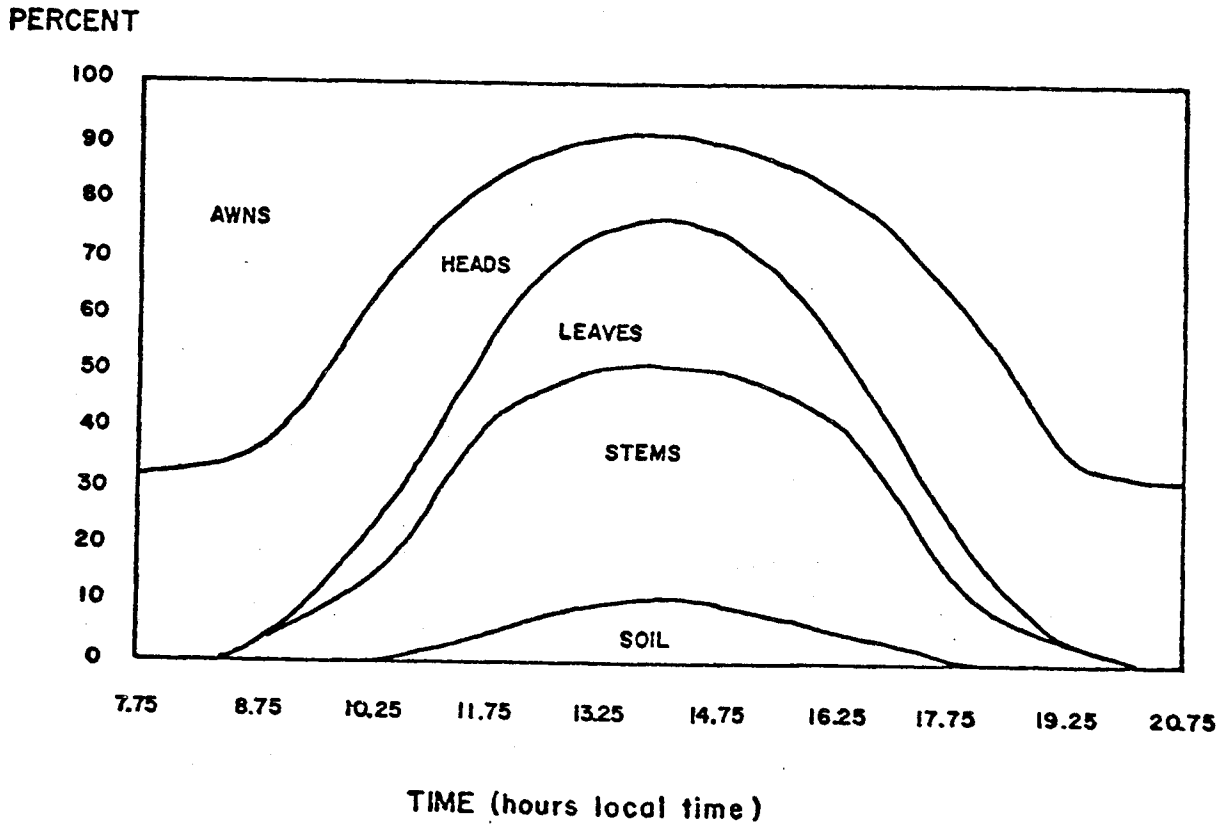


Figure 2.3-10. Percent of solar flux intercepted by each component of the wheat canopy for 30 July 1975.

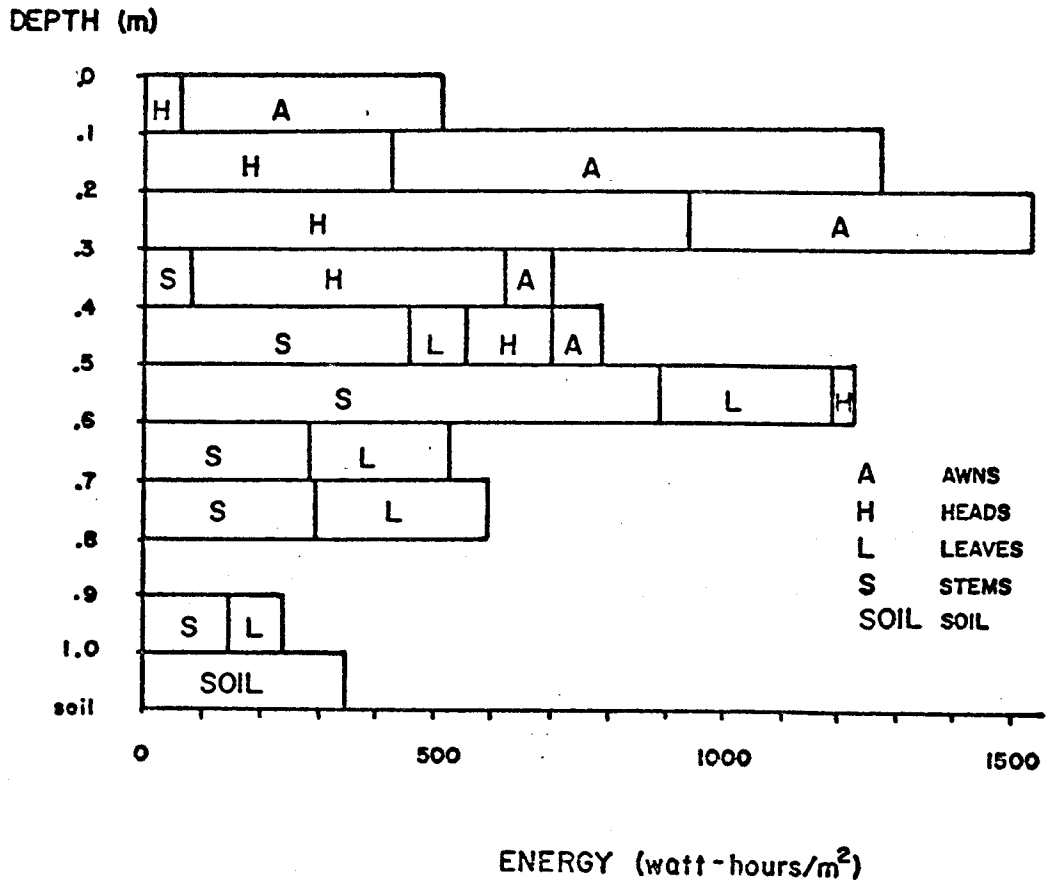


Figure 2.3-11. Solar energy intercepted by components of each layer of the wheat canoy for 30 July 1975.

Also, during this next quarter, planning and preparation for this summer's data collection activities by LARS at Williston, North Dakota will be completed. At present the plans are to meet with the FSAS system and crew in Garden City, Kansas for the May 17-20 Mission and then go to Williston, North Dakota to begin data collection at the end of May. The purpose of the meeting in Garden City, Kansas is to obtain data for further correlation of the FSS, FSAS, and Exotech 20C sensors.

During the next quarter, analysis of the spectral discriminability of different crops and of crops from bare soil will be investigated using analysis of variance and discriminant analysis. In addition, regression and correlation analyses of the available agronomic variables including maturity stage, canopy height, percent ground cover, and stand quality will be performed. Similar procedures and objectives will be pursued with the data from the North Dakota and South Dakota Intensive Test Sites for both years of data.

2.4 Scanner System Parameter Study

Background

The scanner parameter study was initiated in CY76 to conduct research on analytical methods of multispectral sensor system design. It was planned that basic 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 without the use of simulation techniques. 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 CY76.

Shortly into CY76, the direction of the study was changed by 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 CY76 and the results were reported in the June 1976 final report and in Information Note 110976. Research on the analytical model program was resumed in CY77. Progress has been made in the three basic areas defined for the study.

With respect to the implementation plan five tasks were active during the quarter. Spectral data base development (Task 2) was completed in the quarter and Task 3 Statistical Model Definition was essentially begun. The core of the study is Task 4, in which the scene spectra are analyzed with the goal of defining spectral sampling functions to optimize sensor perfor-

mance. Tasks 6 address the problem of information extraction and the goal is to develop numerical methods of classification error prediction. Task 7 addresses the other parameters to be modeled (e.g., spatial sampling function, signal-to-noise ratio) and is the activity which will integrate the elements of the study into a set of programs and procedures for scanner system analysis. Progress on these tasks is discussed in the following sections.

Task 2. Spectral Data Base Development (completed December 31, 1976)

The activities included in this task relate to the acquisition, reformatting, preprocessing, error checking, cataloging and averaging of field spectrometer data for use in defining the scene model. A large number of spectra are needed for characterizing example classes. Work was completed in the quarter to carry out preprocessing of an adequate number of field spectra to support the scanner study. Data collected from Finney County, Kansas, Hand County, South Dakota, and Williams County, North Dakota, was selected.

A sufficient number of spectra have been processed and Task 4 has made use of these spectra for defining statistics for example classes for algorithm test purposes. The approach chosen was to concentrate on the five information classes studied in the thematic mapper simulation in CY76 which will enable comparison of results from the theoretical models and simulation results. The five classes are: 1) unharvested wheat; 2) harvested wheat; 3) grasses/pasture; 4) fallow; 5) other (corn/oats). These are the classes used in the thematic mapper simulation using MSDS aircraft scanner data from Williams County, North Dakota, on August 15, 1975. Figure 2.4-1 contains a plot of typical field spectra for these five classes. The wavelength range being studied is

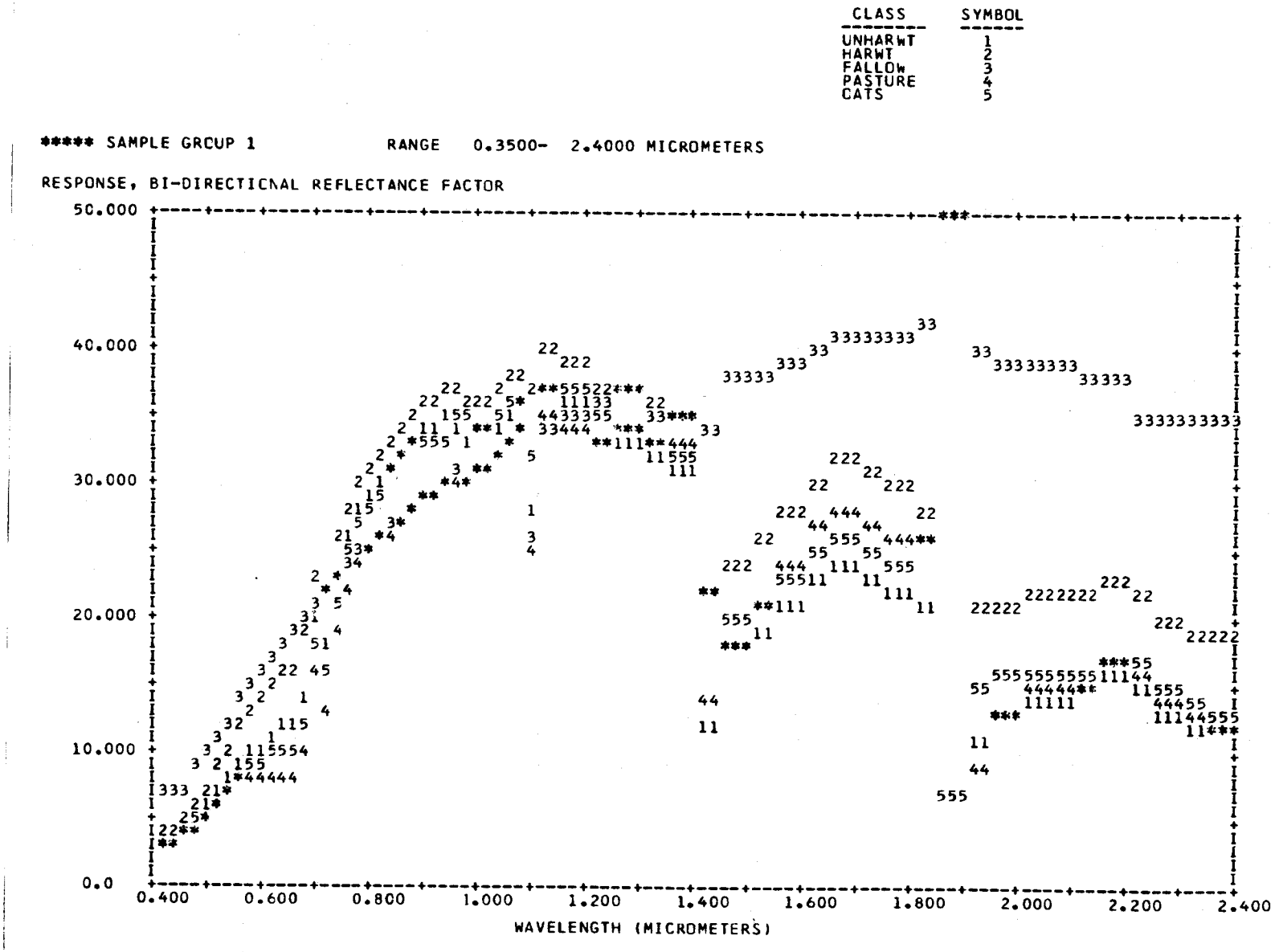


Figure 2.4-1 Example field spectrometer data for the five informational classes being used in the study.

from .4 to 2.4 micrometers which is represented by 200 values. Three water absorption bands can be seen in the data at about 1.1, 1.42, and 1.85 micrometers. A large number of these spectra are now available to the other tasks in the study via the accumulated data base and the LARS EXOSYS field spectrometer data handling system.

Task 3. Statistical Model Definition

This task has evolved as a component of the optimum basis function work (Task 4) and therefore will not exist as a separate task. Task 4 includes two studies of optimum representation of the spectral response of the scene and statistical models (Karhunen-Loeve, mutual information, entropy and others) are being developed in these activities. Since the evolution in the viewpoint toward Task 3 evolved in the third quarter it will be assumed that this task is complete on February 28, 1977, and the essence of the task continued in Task 4 for the duration of the study.

Task 4. Optimum Basis Function Definition

The work on this task is proceeding well and some output from the software generated is presented in this report. Two approaches continue to be pursued to define scanner basis or wavelength sampling functions: 1) Karhunen-Loeve expansion approach; and 2) Information Theoretic approach. Approach (1) has proceeded to the point of processing field spectrometer data for the five information classes and estimation of the K-L basis functions. Work on (2) is still in the theory development stage. Further discussion on these two activities is presented below.

Karhunen-Loeve Representation

The theory of K-L expansions is well established and we are applying it here to the case of an ensemble of reflectance/emission spectra from selected classes of areas of the earth's surface. The K-L representation in the case considered here is based on a set of orthogonal basis functions, which are derived from the covariance matrix of the discrete process which is being modeled.

Particular emphasis in this quarter was directed toward developing an upper bound on the ability of a remote sensing system to discriminate between two classes given a set of representative sample functions from each class. The random processes consist of a set of sampled functions of wavelength for each class taken by a field spectrometer. The probability density function associated with the random process is assumed to be Gaussian.

In order to obtain expressions for the probability density function for each class, each random process is represented by a series expansion in terms of orthonormal basis functions ϕ_n.

$$(1) \qquad X_1(\lambda) = \sum_{n=1}^{N} \alpha_n \phi_n \qquad \text{for class 1}$$

$$X_2(\lambda) = \sum_{n=1}^{N} \beta_n \phi_n \qquad \text{for class 2}$$

Furthermore, the coefficients α_n and β_n should be uncorrelated.

$$(2) \qquad E[x_i x_j] = 0 \qquad i \neq j$$

$$E[\beta_i \beta_j] = 0$$

The coefficients are given by

$$\alpha_i = \phi_i^T X_1(\lambda)$$

$$\beta_i = \phi_i^T X_2(\lambda)$$

Consider the received waveform at the scanner, $\gamma(\lambda)$.

$$\gamma(\lambda) = \sum_n \gamma_n \phi_n$$

and

$$\gamma_i = \phi_n^T \gamma(\lambda)$$

The probability density function of the received signal given that the scanner is looking at class 1 is

$$p(\underline{\gamma}|X_1) = \frac{1}{(2\Pi)^{N/2} \prod_{i=1}^{N} \gamma_{1i}} \exp\left\{ -\tfrac{1}{2} \sum_{i=1}^{N} \frac{(\gamma_i - \alpha_i)^2}{\gamma_{1i}^2} \right\}$$

where γ_{1i} is the variance of the i^{th} component. The corresponding density function for class 2 is

$$p(\gamma/X_2) = \frac{1}{(2\Pi)^{N/2} \prod_{i=1}^{N} \gamma_{2i}} \exp\left\{ -\tfrac{1}{2} \sum_{i=1}^{N} \frac{(\gamma_i - \beta_i)^2}{\gamma_{2i}} \right\}$$

From these density functions it is straight forward to design a classifier to discriminate between the classes. The criterion for performance is the

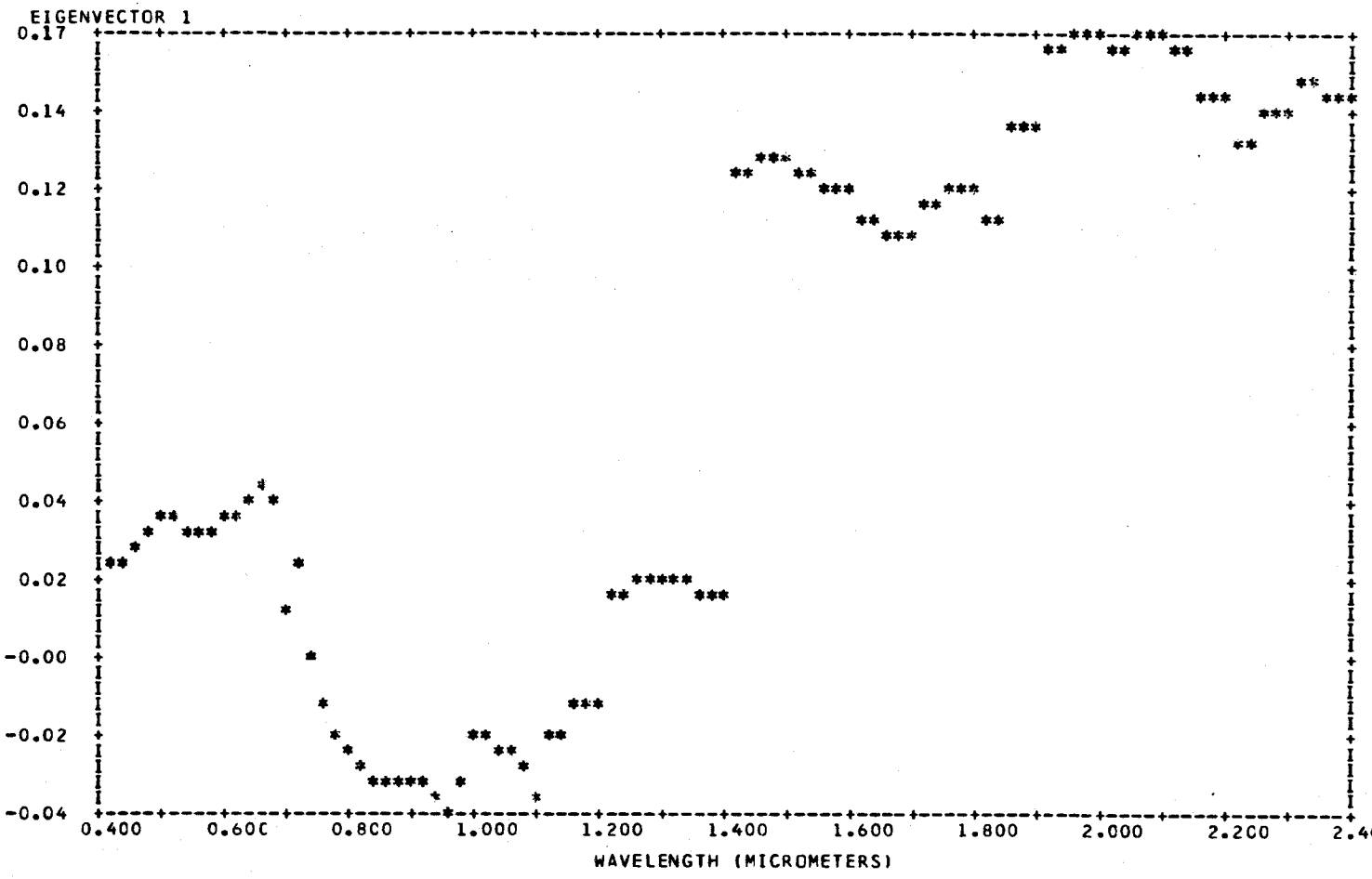
probability of misclassification. The misclassification probability can be used to compare the performance of a candidate system with that of the optimum system using K-L computed basis functions.

To demonstrate what some of the orthonormal basis functions will look like, some representative functions have been computed for four classes: spring wheat; Durham wheat; fallow; and pasture. The first four functions corresponding to the four largest eigenvalues in the Karhunen-Loeve expansion are plotted for the Durham wheat class in Figure 2.4-2. The functions are plotted against wavelength over the range from .4 to 2.4 micrometers. Care must be taken when comparing graphs as the scale is different for each. In theory one should be able to reconstruct the original waveform from a linear combination of the basis functions with appropriate coefficients. The accuracy of the reconstruction will improve as more basis functions are used. The purpose of using such a representation is to convert an infinite dimensional process into a finite dimensional process which is adaptable to numerical solution.

In the fourth quarter both classification error bounds and probability of correct classification using the numerical method of Task 6 will be computed. Both the K-L basis function and the rectangular basis functions from the thematic mapper study will be used and results compared. This will produce quantitative results for the main thrust of the scanner parameter study. Significant improvement in performance of the optimum basis function approach will have potential impact on advanced system designs.

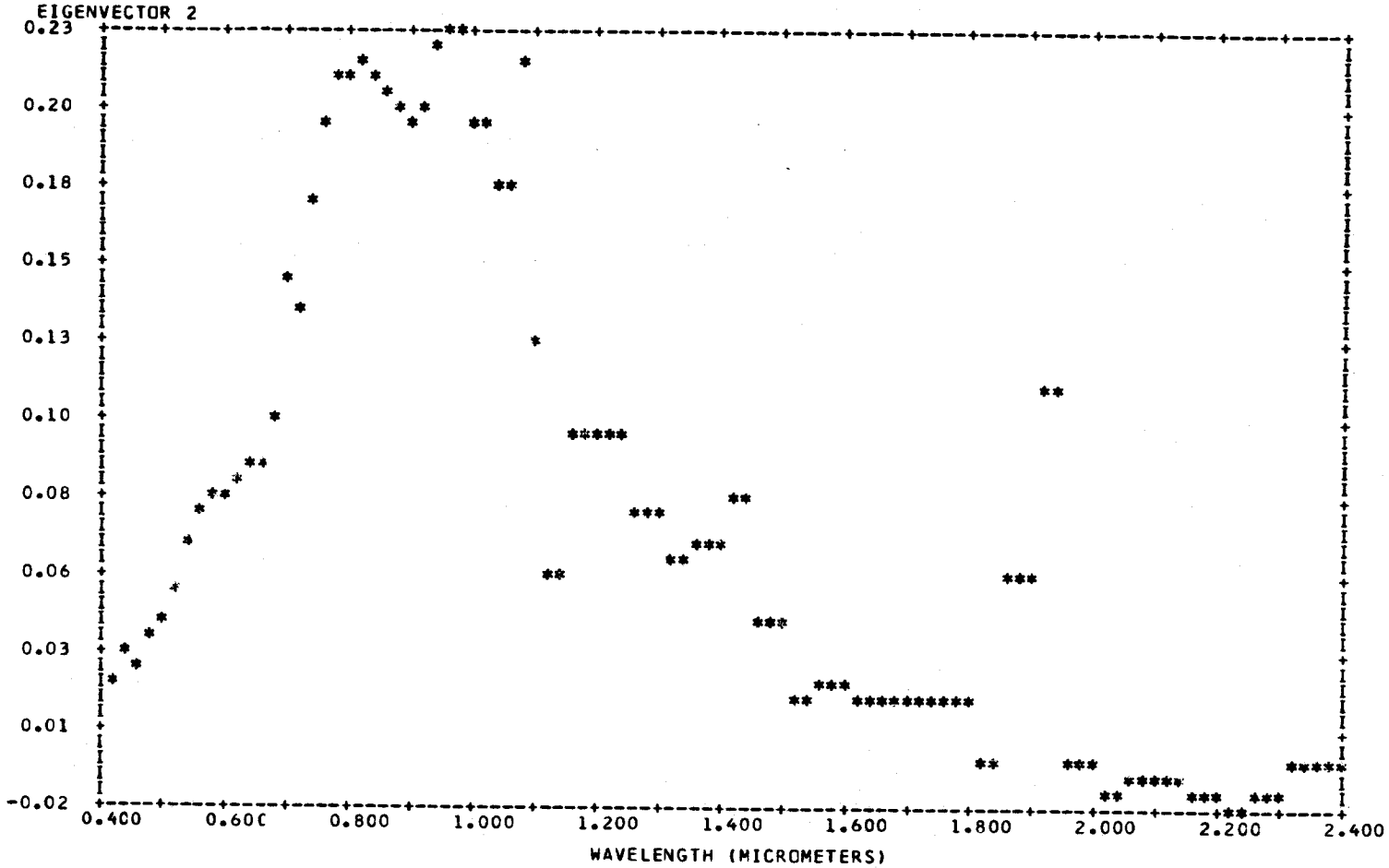
Information Theoretic Approach

In the previous quarterly report, the information theoretic approach to spectral band selection was outlined. It was found that one is required to evaluate the equivocation $H(c/\underline{x})$. That is, evaluation of



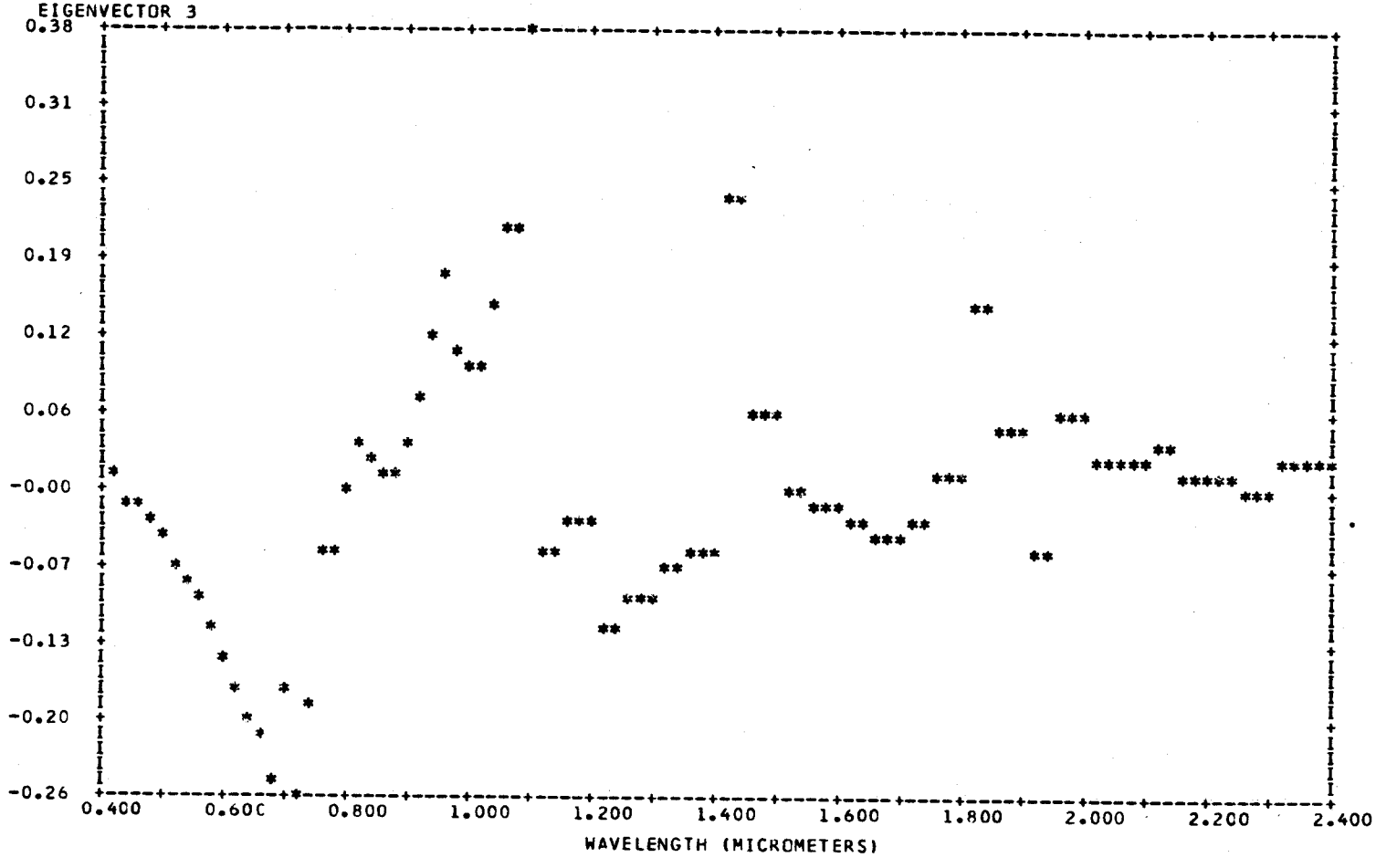
0.0196	0.0225	0.0261	0.0288	0.0321	0.0324	0.0288	0.0284	0.0311	0.0332
0.0349	0.0378	0.0432	0.0374	0.0067	0.0215	-0.0024	-0.0182	-0.0274	-0.0312
-0.0343	-0.0366	-0.0377	-0.0378	-0.0359	-0.0378	-0.0410	-0.0442	-0.0359	-0.0265
-0.0265	-0.0278	-0.0297	-0.0357	-0.0422	-0.0241	-0.0241	-0.0158	-0.0158	-0.0158
0.0103	0.0103	0.0170	0.0170	0.0170	0.0152	0.0152	0.0142	0.0142	0.0142
0.1251	0.1251	0.1282	0.1282	0.1282	0.1267	0.1267	0.1199	0.1199	0.1199
0.1123	0.1123	0.1100	0.1100	0.1100	0.1180	0.1180	0.1219	0.1219	0.1219
0.1140	0.1140	0.1368	0.1368	0.1368	0.1597	0.1597	0.1653	0.1653	0.1653
0.1607	0.1607	0.1629	0.1629	0.1629	0.1581	0.1581	0.1483	0.1483	0.1483
0.1352	0.1352	0.1438	0.1438	0.1438	0.1495	0.1495	0.1468	0.1468	0.1468

Figure 2.4-2a First K-L basis function for field spectrometer data for a Durham wheat sample.



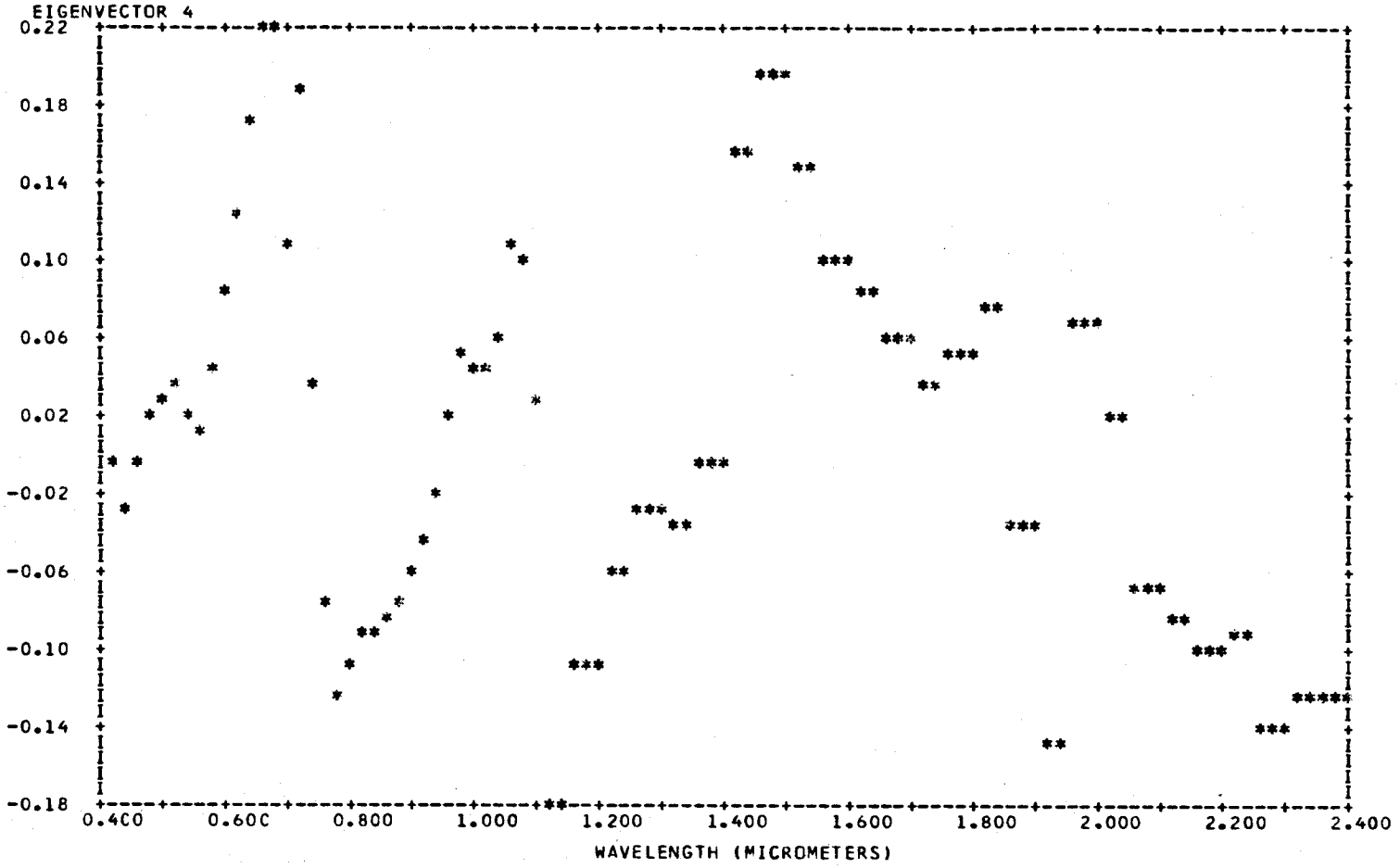
0.0199	0.0307	0.0259	0.0333	0.0412	0.0494	0.0620	0.0734	0.0765	0.0793
0.0833	0.0864	0.0884	0.1042	0.1444	0.1345	0.1720	0.1951	0.2087	0.2072
0.2123	0.2110	0.2054	0.2009	0.1932	0.2001	0.2202	0.2243	0.2259	0.1950
0.1950	0.1751	0.1751	0.2159	0.1283	0.0551	0.0551	0.0968	0.0968	0.0968
0.0992	0.0992	0.0730	0.0730	0.0730	0.0565	0.0565	0.0614	0.0614	0.0614
0.0787	0.0787	0.0377	0.0377	0.0377	0.0156	0.0156	0.0202	0.0202	0.0202
0.0151	0.0151	0.0149	0.0149	0.0149	0.0136	0.0136	0.0134	0.0134	0.0134
-0.0024	-0.0024	0.0555	0.0555	0.0555	0.1133	0.1133	-0.0058	-0.0058	-0.0058
-0.0140	-0.0140	-0.0106	-0.0106	-0.0106	-0.0086	-0.0086	-0.0120	-0.0120	-0.0120
-0.0162	-0.0162	-0.0136	-0.0136	-0.0136	-0.0065	-0.0065	-0.0022	-0.0022	-0.0022

Figure 2.4-2b Second K-L basis function for field spectrometer data for a Durham wheat sample.



0.0012	-0.0227	-0.0227	-0.0389	-0.0555	-0.0704	-0.0825	-0.1029	-0.1248	-0.1496
-0.1792	-0.1978	-0.2123	-0.2543	-0.1744	-0.2597	-0.1859	-0.0655	-0.0612	-0.0076
0.0327	0.0130	0.0020	-0.0034	0.0280	0.0613	0.1191	0.1658	0.1001	0.0876
0.0876	0.1409	0.2115	0.2016	0.3778	-0.0647	-0.0647	-0.0395	-0.0395	-0.0395
-0.1248	-0.1248	-0.1004	-0.1004	-0.1004	-0.0746	-0.0746	-0.0666	-0.0666	-0.0666
0.2262	0.2262	0.0501	0.0501	0.0501	-0.0104	-0.0104	-0.0285	-0.0285	-0.0285
-0.0340	-0.0340	-0.0450	-0.0450	-0.0450	-0.0308	-0.0308	0.0054	0.0054	0.0054
0.1462	0.1462	0.0400	0.0400	0.0400	-0.0662	-0.0662	0.0587	0.0587	0.0587
0.0116	0.0116	0.0138	0.0138	0.0138	0.0248	0.0248	0.0064	0.0064	0.0064
-0.0003	-0.0003	-0.0048	-0.0048	-0.0048	0.0179	0.0179	0.0117	0.0117	0.0117

Figure 2.4-2c Third K-L basis function for field spectrometer data for a Durham wheat sample.



-0.0050	-0.0305	-0.0058	0.0123	0.0255	0.0287	0.0119	0.0068	0.0402	0.0834
0.1244	0.1704	0.2227	0.2181	0.1058	0.1827	0.0337	-0.0786	-0.1266	-0.1174
-0.0965	-0.0937	-0.0879	-0.0843	-0.0628	-0.0487	-0.0235	0.0197	0.0462	0.0423
0.0423	0.0595	0.1064	0.1010	0.0217	-0.1826	-0.1826	-0.1171	-0.1171	-0.1171
-0.0685	-0.0685	-0.0318	-0.0318	-0.0318	-0.0376	-0.0376	-0.0120	-0.0120	-0.0120
0.1520	0.1520	0.1939	0.1939	0.1939	0.1431	0.1431	0.0994	0.0994	0.0994
0.0810	0.0810	0.0564	0.0564	0.0564	0.0338	0.0338	0.0497	0.0497	0.0497
0.0726	0.0726	-0.0425	-0.0425	-0.0425	-0.1576	-0.1576	0.0658	0.0658	0.0658
0.0123	0.0123	-0.0704	-0.0704	-0.0704	-0.0933	-0.0933	-0.1042	-0.1042	-0.1042
-0.0966	-0.0966	-0.1431	-0.1431	-0.1431	-0.1296	-0.1296	-0.1303	-0.1303	-0.1303

Figure 2.4-2d Fourth K-L basis function for field spectrometer data for a Durham wheat sample.

$$H(c/\underline{x}) = -\sum_{j=1}^{K} \int_{X_1} \cdots \int_{X_m} p(c_j,\underline{x}) \log [p(c_j/\underline{x})] dx_1,\ldots dx_m$$

$$= \sum_{j=1}^{K} p(c_j) \int_{X_1} \cdots \int_{X_m} p(\underline{x}/c_j) \{\log [p(x/c_j)p(c_j)] -$$

$$-\log [p(\underline{x})]\} dx_1,\ldots dx_m$$

It is seen that evaluation of an m-multiple integral of a function of the mixture density $p(\underline{x})$ is required. Such an integral is, in general, not easy to evaluate since an amenable functional form for $p(\underline{x})$ is not available. Thus even an accelerated search technique for the optimum $(X_1,\ldots X_m)$ would require excessive computation time. It is thought that an approach which obviates this problem has been found.

Let the spectral response be considered as a portion of a stochastic process in wavelength. It can be shown that the mutual information between two wavelengths λ_1 and λ_2 is given by:

$$I(\lambda_1,\lambda_2) = \tfrac{1}{2} \int_{\lambda_1}^{\lambda_2} h(\lambda,\lambda) d\lambda.$$

The expression $h(\lambda,\lambda)$ is the impulse response of the optimal Weiner filter for estimating the spectral response of the scene from the observed spectral process.

Thus the problem has been placed in a form which requires the evaluation of a single integral rather than a multiple integral.

The Weiner filter response $h(\lambda,\lambda)$ can be determined from the nature of the spectral stochastic process and the assumed noise disturbing the observation. Current efforts are being directed towards identification, selection, and validation of adequate models for the spectral random process. Data from field measurement data tapes are being used to identify the models.

A dynamic stochastic model for spectral data should be developed by the end of the next quarter. The modeling techniques are similar to those used in time series analysis approaches. It is also hoped that evaluation of mutual information as a function of spectral bandwidth will be initiated. It is expected however that model identification, selection and validation will consume most of the quarter.

Task 6. Numerical Methods for Classification Error Evaluation

Introduction:

The classification accuracy in pattern recognition problems is generally regarded as the most important piece of information in evaluating the performance of the classifier. In all but very simple cases, this quantity is not available. This statement is specially true if multiple-populations are involved. In this case the only method available so far has been the Monte-Carlo technique.

In LARSYS, a maximum-likelihood (Bayes) classifier is employed, and classification is performed point by point followed by a ratio estimate of probability of error (ratio of misclassified samples to total number of samples). In the previous quarterly report, a method was explained that takes a mathematical approach toward estimating the classification accuracy and does not require any knowledge of sample values of each field. We will briefly review the method.

Classification Accuracy Estimate as a Multi-Dimensional Integral:

The assumptions currently used in LARSYS implementation is adhered to, i.e., Gaussian density function for each class, also assumed is knowledge of

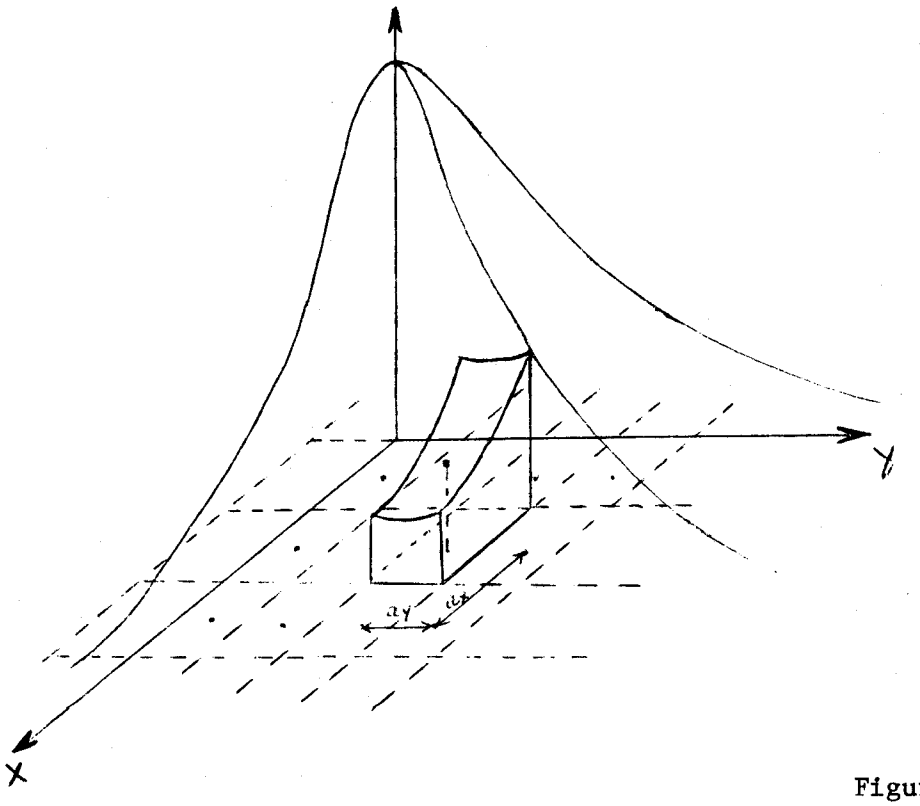
the statistics of each class. Let X, Σ, μ and Pi be the sample value, class covariance matrix, mean vector and a priori probability of each class respectively. Then, the classification is performed on the following basis:

$$\text{Compute } W_i = (\underline{X}-\underline{\mu}_i)^T \Sigma_i^{-1} (\underline{X}-\underline{\mu}_i)+\ln|\Sigma i|-2\ln P_i \qquad \text{Eq.1}$$

$$\text{choose class i if } W_i < W_j \text{ for all } j\neq i \qquad \text{Eq.2}$$

Inequality in Eq.2 defines a region in N-dimensional space. The probability of correct classification is the proper class density integrated over this hypervolume. This method quantizes the space using a binomial approximation to a normal random variable.

The current results correspond to an equi-distance grid along each coordinate axis, therefore resulting in quantization of, for example, a two-dimensional space into rectangles. Then the integration is performed over each area element and the results are summed (see Figure 2.4-3):



$$a_y = \frac{2}{\sqrt{n}}\sigma_y$$

Figure 2.4-3

$$PCC = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_n}\int_{i_1-\frac{a_1}{2}}^{i_1+\frac{a_1}{2}} f(x_1)dx_1 \int_{i_2-\frac{a_2}{2}}^{i_2+\frac{a_2}{2}} f(x_2)dx_2\cdots\int_{i_m-\frac{a_n}{2}}^{i_m+\frac{a_n}{2}} f(x_n)dx_n$$

where n is the number of points along each coordinate axis. This procedure is repeated for each class and then weighted by a priori probabilities to obtain overall classification accuracy.

Preliminary Results:

The thematic mapper simulation study for determining the classification accuracy for various parameters considered a total of 25 classes. For ease of implementation, 5 subclasses from the original 25 classes were chosen (selecting the one largest subclass from each of the 5 information classes) and the thematic mapper simulation classification was performed again. The results are shown in Table 2.4-1.

	Resolution(M)			
Subclasses	30	40	50	60
Unharvested Wheat	98.3	97.8	98.8	100
Harvested Wheat	99.5	100	100	100
Fallow	97.2	98.6	97.6	98
Pasture	98.8	99.5	99.3	100
Corn/Oats	98.9	100	98.2	97.1
Overall	98.2	99.1	98.5	99.1

Table 2.4-1 Training Field Classification Accuracy Obtained through Simulation

Table 2.4-2 shows the size of the subclasses in relation to the size of entire class (number of points). Channels used are 2, 6, and 8 (0.59-0.64, 1.53-1.62, 0.77-0.88 micrometers).

	Resolution (M)			
Subclasses	30	40	50	60
Unharvested Wheat	403/1231	271/997	260/632	86/414
Harvested Wheat	240/1089	160/371	204/402	128/270
Fallow	1315/1895	811/1069	658/658	343/435
Pasture	1471/2260	1178/1438	288/446	288/522
Corn/Oats	88/153	70/130	56/111	46/81

Table 2.4-2 Number of Points in Each Subclass with Total Number of Points in the Class.

For comparison purposes, the same subclasses were then processed by the new processor. First, look at the performance of this algorithm. It is obvious that the quality of the estimate obtained by the multiple integration is directly related to the fineness of the grid, in fact, as $n \to \infty$, we expect to converge to the true classification accuracy. How rapid the convergence is, can be an important element. Figure 2.4-4 shows the probability of correct classification versus grid size for 4 different resolutions. As it is seen, the accuracy stabilizes for a rather small grid of size $n = 8$ and beyond. Since the computation time varies as n^d where d is the dimensionality of the data, it is important to be able to keep n as small as possible.

The reason behind this small grid size can be attributed to the fact that the outer limits of the grid extends to $\sqrt{n}\ \sigma$, where σ is the variance along a particular coordinate axis. And since the normal density functions amplitude is quite negligible at anything beyond 3σ, small values of n could

provide accurate estimates. Of course, the case may be slightly different for other choices of classes. Table 2.4-3 shows the classification accuracy versus grid size.

	GRID SIZE (number of cells along each coordinate)					
Resolution	4	8	12	16	20	24
30	92.71	95.07	94.93	95.14	94.99	95.20
40	96.20	99.20	99.52	99.57	99.58	99.58
50	94.56	96.17	97.20	97.34	97.35	97.35
60	94.53	97.70	98.01	97.97	98.10	97.99

Table 2.4-3 Classification Accuracy versus Grid Size and Resolution.

The comparison of overall classification accuracy obtained via these two methods is in order. In Table 2.4-4 both estimates are shown. Figure 2.4-5 is a plot of overall accuracy versus IFOV for both the thematic mapper simulation and numerical integration cases.

Interpretation of the Results:

Since the computation time in the integration algorithm is proportional to n^d, it is important that the convergence be attained for small values of n. Figure 2.4-4 shows that this is indeed the case for reasons explained previously.

The overall classification accuracy obtained using the new processor compares favorably with the simulation results. Table 2.4-4 and Figure 2.4-5 show that overall estimate of the probability of correct classification is generally within one percentage point of simulation results. The greatest

Resolution (M)

	30		40		50		60	
	Siml.	Est.	Siml.	Est.	Siml.	Est.	Siml.	Est.
Unharvested Wheat	98.3	93.08	97.8	98.97	98.8	93.06	100	95.23
Harvested Wheat	99.5	86.79	100	99.80	100	98.35	100	98.40
Fallow	97.2	99.43	98.6	99.72	97.5	98.65	98	98.56
Pasture	98.8	99.39	99.5	99.94	99.3	99.61	100	99.91
Corn/Oats	98.9	97.00	100	99.44	98.2	97.03	97.1	97.73
Overall	98.2	95.14	99.1	99.57	98.5	97.34	99.1	97.97

Table 2.4-4 Comparison Between Simulation and Estimation Results.

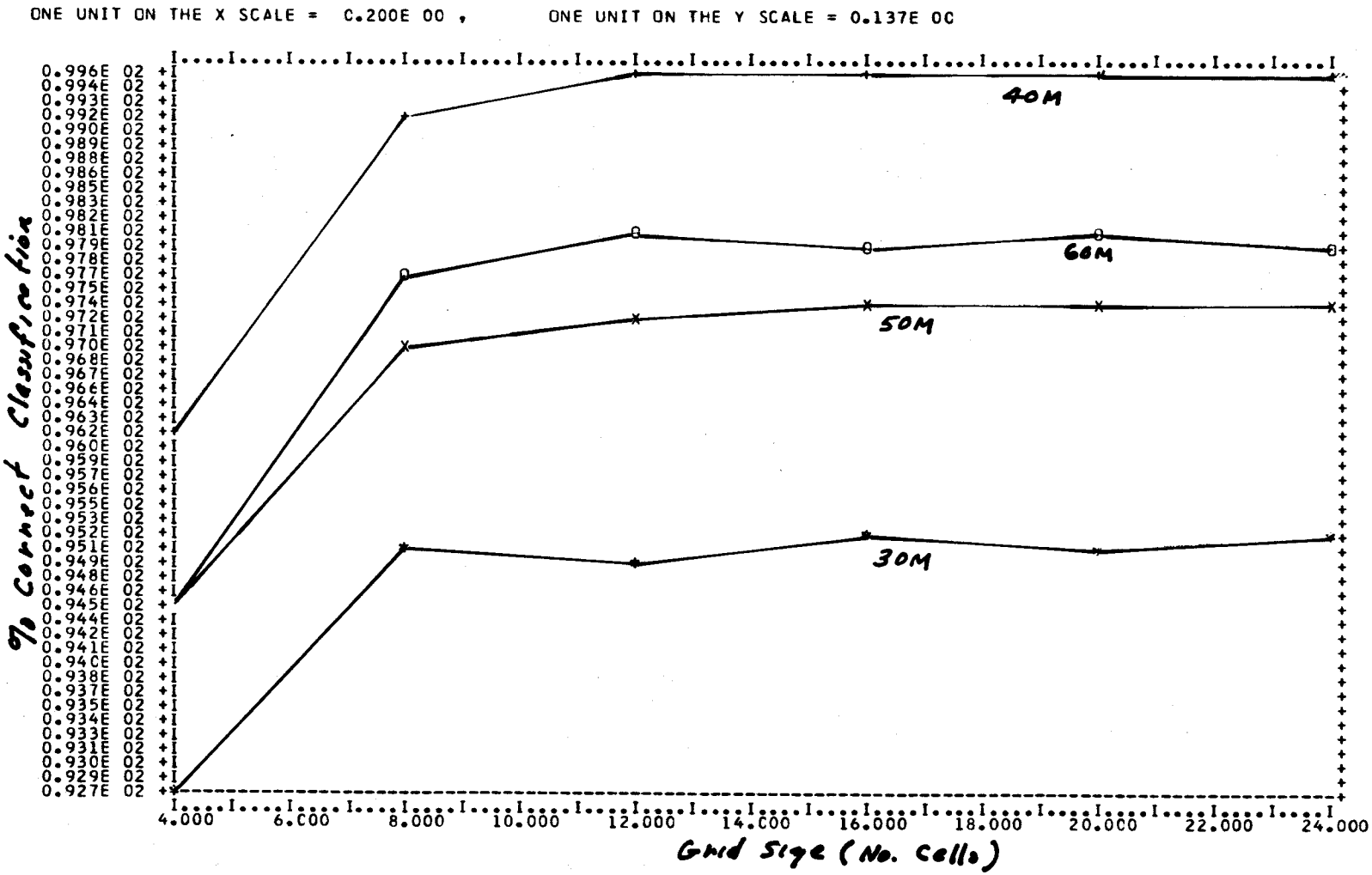


Figure 2.4-4 Probability of correct classification estimate as a function of grid size.

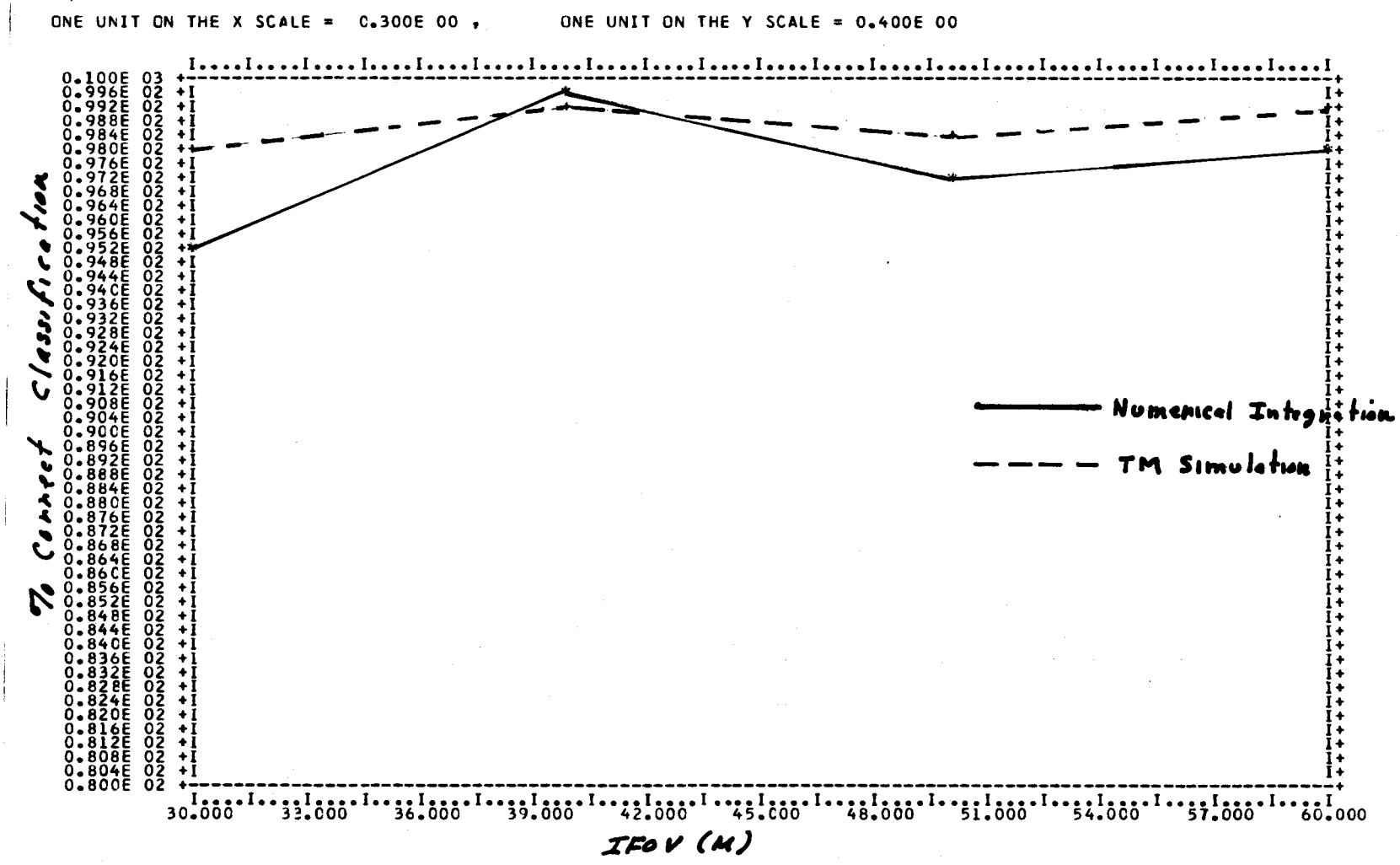


Figure 2.4-5 Comparison of thematic mapper simulation results and numerical integration results for the five information classes.

difference is observed for 30_m data which is 3 percentage points.

Conclusions:

The results presented here show that this theoretical approach to the estimation of the performance of a Bayes classifier is feasible provided the dimensionality of the data is moderate (≤ 5). Moreover, the algorithm is independent of actual field sizes and sample values. This estimate is almost unbiased. Some upper and lower bounds are available for this estimate, but the effort is directed toward a more comprehensive study of the quality of the estimate.

Task 7. Design Multispectral Scanner Model

The scanner parameter study includes consideration of IFOV (Instantaneous Field of View) effects, signal-to-noise ratio, sampling and quantization effects in addition to the spectral band functions. The scanner model task encompasses the generation of models for these other parameters and the combination of all the component models into a system for predicting scanner performance. Methods continue to be considered for modeling the spatial characteristics of a scene and representing the effect of an IFOV function on the spectral data from the scene. The goal here is to generate a statistical representation for various scene types which will enable evaluation of IFOV functions without having to use convolution of real imagery with the candidate functions. Similarly sensor and data system noise models were considered which will permit explicit evaluation of noise effects without simulation.

Specifically the IFOV will be modeled by transforming the statistics of the high resolution pixels covered by an IFOV window to the sensor output

statistics. An experiment to be carried out in the fourth quarter will use the six meter MSDS data from the thematic mapper study. The spatial covariance matrix will be computed for each of the information classes and the output statistics for a Gaussian IFOV will be explicitly computed. Similarly the noise will be represented by the computation of the statistics of a sum of class random variables and noise random variables. These experiments will be reported on in the final report.

2.5 Transfer of Computer Image Analysis Techniques

Change in Implementation Plan. In a letter dated December 20, 1976 JSC requested a change in the implementation plan for Task 2.5 to reflect a) the fact that JSC will not participate in the proposed instructor training or evaluation of training materials, and b) that LARS will estimate the performance, cost, and schedule for installation (from go-ahead) of a tape drive on the JSC LARS terminal. LARS' agreement to these changes was transmitted to Dr. Jon Erickson by letter on January 14, 1977.

Major Activities. The major activities carried out under this task during the period December 1, 1976 through February 28, 1977 were:

- Continued Support of the JSC Remote Terminal
- Development of a set of options for up-grading the JSC terminal
- Evaluation of Education and Training Materials
- Development of Education and Training Materials

These major activities reflect the sub-tasks described in the implementation plan and its modifications as mentioned above. The work accomplished in each of these areas during the past quarter is described below.

Support of the JSC Remote Terminal. Support of the JSC/LARS remote terminal includes maintaining at the LARS computer facility the necessary hardware and software, providing computer services, providing hardware/software system consultation, and developing concepts for an improved remote terminal system.

Hardware support consists of maintaining communication modems at LARS and a port on the 3705 communication system. Software support includes programming the 3705 to recognize the JSC terminal, CP-67 and LARSYS software, and maintaining a virtual machine dedicated to the JSC terminal.

During the past quarter Mr. John Sargent of NASA/JSC was designated as our official contact for the JSC remote terminal in place of John Cornwell. Monthly computer usage reports are now being sent to him. Frank Ravet reported two data runs which did not have as many lines as listed in the runtable; the runtable is being corrected. He also sent two tapes representing half a LANDSAT frame, which were reformatted into LARS' tape library. A system crash on February 6 caused a few of LARS' computer ID's to lose information stored on their private disks, including one of the JSC ID's. As long as the users occasionally store this information on a back-up computer tape, they can easily recover the lost information. On February 9, possible phone line noise was reported to be keeping the typewriter terminals from communicating properly. This has cleared up and the situation will be monitored.

Computer usage for the period November 1976 through January 1977 was as follows (Information for February is not yet available).

Month	CPU hours used	Hrs. terminals attached
November	6.0 hours	264 hours
December	9.3 hours	305 hours
January	5.7 hours	231 hours

Options for Upgrading the JSC Terminal. In response to JSC's request to investigate methods for upgrading the JSC remote terminal to facilitate the transmission of data sets from Houston to LARS, a systems

analysis was undertaken and three approaches to the problem were investigated. Briefly the three approaches are:

A - Use existing Houston 2780 terminal to transmit small data sets

B - Procure enough hardware and software to read and transmit data tapes

C - Procure enough hardware and software to functionally replace existing 2780 terminal and read and transmit data tapes in addition.

A summary of relative capabilities, cost range, and estimated implementation times is given by the following table:

	Approach A	Approach B	Approach C
Pixels per day capability (4 channel)	100K	15M	15M
Hardware cost range (exclusive of telecommunications line)	0	$35K-$50K	$70K-$85K
Time to implement	1 mo.	8 mos.	8 mos.

An oral presentation of these three approaches was made to Mr. Donald Hay (JSC) by Mr. Terry Phillips (LARS) at JSC during the week of January 24, 1977. In the course of the ensuing discussion it was decided to pursue option C in more detail.

Option C is a natural outgrowth of work reported earlier to provide enhanced remote terminal access to the LARS system. At conclusion of Phase I of the current work, a prototype terminal, based on the PDP-11/34, will be operational at LARS. A possible follow-on Phase II would be to acquire similar hardware to be installed at Houston, and to

add industry-standard magnetic tape hardware to the Houston installation. The additional software effort beyond the Phase I development would be minimal.

Depending on amount of data to be transmitted and relative priorities, the question of sharing the telecommunications link with other use of it may or may not be a problem. It would be advisable to upgrade the link to 9600 bps.

Some of the important design considerations of this option are:
a) Throughput considerations. Throughput is limited by the data rate of the telecommunications line and modems. The following table shows throughput for two possible data rates. The figures include reasonable allowance for errors requiring retransmission of some data blocks and for turnaround between blocks. They do not assume that any data compression techniques are used. (Such techniques have the potential to reduce transmission time by one-third or more).

	4800 bps	9600 bps
Actual data throughput rate	512 bytes/sec	976 bytes/sec
Time to transfer 1 million bytes	.54 hour	.28 hour
Time to transfer ¼ Landsat frame	4.1 hour	1.4 hour
Time to transfer full Landsat frame	16.4 hour	8.6 hour

b) Hardware requirements for the Houston terminal. A possible set of hardware is given by the following list:

DEC PDP-11/34 System with terminal, software, disks
Tape Unit and Control
Communication Interface Hardware
2780 Emulator Software
Card reader and punch
Electrostatic printer/plotter & interface

At this stage of development of improved remote terminal capabilities at LARS, it is still too early to give very good answers to several important questions relative to the JSC terminal. For example, the PDP-11/04 CPU could likely be sufficiently powerful that it could be substituted for the 11/34 at some cost savings. The card punch capability is quite expensive. After development of our prototype, it is envisioned that the card punch capability will not be required or else a very low volume, low cost device will suffice. If the card punch capability were not required at Houston, then a card reader would be available for $5.1K. Because of these unanswered questions and uncertainties with respect to the availability of some of the proposed hardware it is not possible to determine a firm cost figure for the proposed hardware. However, a cost range of between $70K and $85K has been determined.

c) Software at Houston. In addition to installing software developed for the LARS improved remote terminal at JSC, enhancement would be made so that data could be read from tape and converted to card images for transmission to LARS.

d) Hardware at LARS. No additional hardware is required.

e) Software at LARS. The software requirements at LARS is the same for all three options. It is only necessary to reformat data received as card images into LARSYS-format runs.

f) Lead-time requirements. The major consideration is the 6-month delivery time currently being quoted for the mini-computer hardware.

Evaluation of Education and Training Materials. As indicated in the change in implementation plan, JSC is not participating in the evaluation of education and training materials previously developed. Thus the only evaluation of materials during the past quarter was that done at LARS in conjunction with the development of new materials. These efforts are described in the next section.

Development of Education and Training Materials. During the past six months there has been a shift in emphasis with regard to the development of education and training materials under SR&T support. Effort on materials dealing with fundamental concepts (the FOCUS series) has been declining. Effort on materials related to current SR&T application tasks (Simulation Exercises) has remained relatively constant and there has been a significant increase in effort in developing materials to support new analysis techniques (ECHO case study). This reorientation of priorities seems consistent with changes in JSC's requirements for technology transfer and compliments the effort to upgrade the JSC remote terminal to make it a more effective tool for the test and evaluation of new analysis procedures.

a) Materials completed. The simulation exercise entitled "Determining Land Use Patterns Through Man-Machine Analysis of Landsat Data" was published as LARS Information Note 070676 near the beginning of this reporting period. This document is designed to show the reader typical steps in the analysis of remotely sensed data for determining land use patterns. The format of this document was designed to enhance

the transfer and understanding of remote sensing technology to a particular segment of the user community - in this case, persons interested in land use management.

Also completed this past quarter were three additional titles to the FOCUS series. Each FOCUS is a two-page foldout consisting of a diagram or photograph and an extended caption of three to four hundred words treating a single concept. Their titles and brief descriptions are listed below.

REFORMATTING LANDSAT DATA

David M. Freeman

Landsat data tapes are supplied in a format that is efficient for data collection and transmission. To obtain data in a format efficient for data processing, reformatting is commonly done. Changing the data to one widely accepted format, LARSYS format, involves merging the four separate subimages the data is supplied in into one continuous image, and reordering the data samples.

THE MULTIBAND CONCEPT

James D. Russell

Objects that appear similar when photographed with black-and-white film often can be differentiated if the portion of the spectrum the film is sensitive to is broken into several more narrow wavelength bands by using appropriate filters in conjunction with the film. As more film and filter combinations are used more objects previously appearing similar can be differentiated. This concept of using multiple wavelength bands to differentiate and identify ground cover types has led to the development of multispectral scanning instruments.

SNOW COVER MAPPING

Shirley M. Davis

In many parts of the world, melting snow is collected in reservoirs and is the major source of water. Ground based observations of snow extent and depth are normally used to efficiently manage these reservoirs. However, the advent of satellite-borne sensor systems has allowed demonstration of the ability to obtain detailed snow measurements using spectral data.

These three additions to the FOCUS series completes the planned activities in this area.

b) Materials under development. The beginnings of a revision of a portion of the LARSYS Educational Package was described in the last quarterly progress report. Drafts of revisions of the hardware dependent units (Unit III, Demonstration of LARSYS on the 2780 Remote Terminal and Unit IV, the 2780 Remote Terminal: A "Hands-On Experience") were made to accomodate a Data 100 Printer-Card Punch-Card Reader Unit and an IBM 2741 typewriter or CRT typewriter terminal. During the December 1-February 28 quarter the level of effort on these revisions was reduced significantly in order that resources could be devoted to upgrading of the JSC terminal and the ECHO case study. However, drafts of the revised materials were tried by several new users of the LARSYS software system.

In an effort to provide education and training materials on ECHO classification almost simultaneously with the release of the ECHO software, two instructional development activities are now underway. One activity is the redesign of the case study format so that the materials are 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 activities stratigically placed through the materials. Appropriate computer printouts will be available at key points simulating a batch machine operation.

The analysis procedure being used to develop the new format is the Landsat per point classification of Unit VII of the LARSYS Educational Package. Testing the instructional approach with familiar content (i.e., an analysis example which has already been thoroughly documented) will help assure the effectiveness of the instructional format because we have a baseline against which we can judge its effectiveness. A draft of the Landsat per point classification case study using the new format was prepared in December. During January and February it was revised based upon suggestions of experienced analysts and a student tryout version was completed. It will be tested by a group of students in a graduate remote sensing course in March.

The development of the ECHO case study is proceeding along the same sequence as the per point classifier example described above but several steps behind. Thus we will be able to take into account any format weaknesses uncovered during the tryout of the per point classification materials. During the past quarter LARS instructional developers have been working with the analysts assigned to task 2.1 to learn and document ECHO classification procedures. A preliminary outline of the ECHO case study has been written. Each step in the analysis will be discussed at three levels: an overview describing the purpose, objectives and techniques used; a second level dealing with the specifics of a sample analysis; and, a third level which will allow the student to carryout his own ECHO analysis. An initial draft of the overview level has been completed for each step of the analysis.

Plans for Next Quarter. Support of the JSC remote terminal will be continued as planned. Work will continue on the proposal to up-grade the JSC terminal so as to provide tape drive capability. A report detailing the steps required to implement the system described above as option C will be prepared.

The only evaluation of education and training materials will be that associated with the new case study format and the ECHO case study.

While no additional titles to the FOCUS series are planned, an information note compiling all titles in the series will be issued. Revisions of LARSYS Educational Package materials will continue on a low priority basis. A high priority will be placed on the development of the ECHO case study.

2.6 Large Area Crop Inventory Design

Several Purdue faculty members from the Agronomy and Agriculture Economics Departments who have experience in international agriculture have been consulted and inputs/recommendations for the rationale for selecting areas and crops for inventory received. A data tape containing data for 1960 to 1975 on the acreage, yield, production, export, and import of all cereal grain crops on a country basis has been obtained from USDA/ERS. The data are being summarized to determine the ranking and trends in grain production, imports and exports, of approximately 25 countries.

Five LARS staff attended the LACIE review at JSC in January and information received at that meeting is being evaluated. In addition, previous remote sensing experiments of wheat and other crops are being reviewed to determine the potential and problems for a multicrop inventory.

An oral presentation, including recommendations for areas and crops for inventory, review of the LACIE design and results with regard to their applicability to other areas and crops, and assessment of the "state of the art" in relation to a multicrop inventory system will be made at the SR&T Quarterly Review at JSC, March 1-4.

2.7 Forestry Applications Project

Background

The forestry applications task is working toward a preliminary design for forest inventory utilizing Landsat data and computer-aided analysis techniques. Recommendations are being prepared for a first level of design intensity which identifies the critical inventory variable of acreage. Figure 1 is a schematic outlining the anticipated levels of information in the preliminary design.

This work is in response to our objective of developing inventory methods using remote sensing techniques for large area forest and rangeland surveys. During the initial phase of this task our major emphasis was directed toward a review of the state-of-the-art of inventory design and the requirements of the 1974 Renewable Resources Planning Act. Our attention was recently focused by the Forestry Applications Program (JSC) on Forest Survey, the most highly developed of the renewable resource survey systems.

The major activity during this reporting period has been in determining how remote sensing, specifically computer-assisted analysis of Landsat data can input into Forest Survey. From a review of the Resources Planning Act we have determined that area is the most directly measurable estimate that Landsat analysis can provide. Furthermore, for the six resource systems defined by the RPA, only area of range, water and forest can be directly determined from Landsat data.

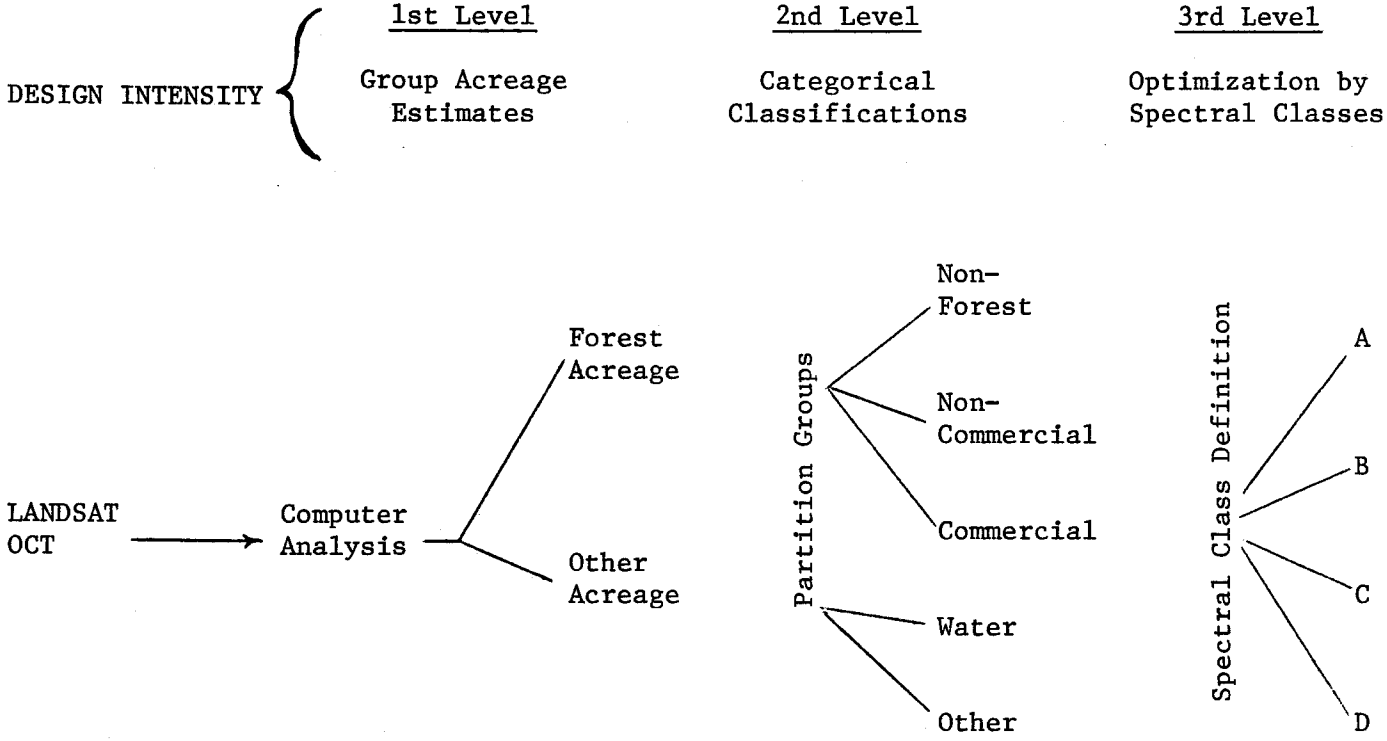


Figure 1. Schematic flow anticipated in design definition.

Current Survey Situation

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

The sample points thusly selected are further categorized as being:

a) commerical forest by type-size-density

b) non-commercial forest

c) non-forest with or without trees

d) water

A subset or sample of the points identified as forest are ground checked. Extensive information is collected about the merchantable as well as the growth characteristics of only the commercial forest plots. This ground information is then expanded back over the entire sample area to give an estimate of the total forest resource of the site.

Appraisal

The above example assumes that the initial estimate of acreage from the aerial photo interpretation was correct. Because survey rarely enjoys the luxury of new aerial photos (new being dated within a year of the survey), acreage estimates must be adjusted to reflect changing land use patterns. Since Landsat data is available on a

timely basis, changes in land use trends would be identified on this data and allocation of the photo and field samples should therefore be improved although not optimized. The basic assumption would be that the acreage estimate provided from the Landsat data analysis is within the accuracy requirements for the survey.

Hypothesis

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

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

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

Test

A test will be preformed which compares the areal estimates of forest versus all other classes as determined by Landsat analysis to Forest Survey estimates. Comparison will be made on a county-by-county

basis for selected counties in three or four states. Counties are being selected from a 191 county area that has previously been analyzed at LARS. Statistical tests will be preformed to determine if there is any significant difference between the Landsat estimates and the published survey figures.

In addition to preforming a comparison of areal acreage estimates for the classifications, we will review the data which comprise the forest classes to determine if any relationships exist between spectral response and the following variables:

1) date of data collection
2) relative reflectance per channel versus
    a) number of species associations expected
    b) general soils association
    c) geomorphology
    d) climate

If any relationships exist between spectral response and the nature of the ground cover then these may be used to define the number or the range of spectral classes that would make up the forest component of any scene. Such relationships will aid in describing guidelines for the class selection process of the design.

Another important activity which will occur involves an estimation of the time required for the preliminary acreage classifications. In addition to using a timely data set, timely information must be provided from the classification procedure if any anticipated benefit to Survey is to be expected.

Summary

In developing Landsat inventory guidelines we will:

A. Test to determine if acreage estimates from machine-assisted Landsat analysis fall within defined Forest Survey Standards of:
   1. An allowable error of 3%/million acres for commercial forest, and
   2. An allowable error of 10%/million acres for non-commercial forest

B. The Landsat survey boundary conditions should fit the following criteria:
   1. Overall survey efficiency must be equal to or better than existing surveys with regards to:
      a. timeliness of data
      b. timeliness of acreage estimates
   2. Physical limitations of survey
      a. classification results must be repeatable for large areas
      b. data through put should not be a limiting factor
   3. A priori inputs based on ancillary information should be used to:
      a. improve speed of classification
      b. improve classification accuracy

The above information will be inputs to the preliminary design procedures which will be included in the final report.

No section 2.8

2.9 Interpretation of Thermal Band Data

Major Activities. This report will serve as a general update concerning activities to date. Material from the first two quarters, some previously unreported, is included for clarity.

In each of the three Missions undertaken during the Summer 1976, a separate report was prepared providing details on the objectives, experiment design, procedures, and organization of the data records. To expedite the completion of these reports, no attempt was made to analyze the observations or prepare conclusions on the work.

In this present report, our aim is to summarize what appears to be the most significant accomplishments of the Missions. The amount of data collected is far beyond our capability to reasonably study so it was necessary to sort out primary effects that are likely to impact our understanding of thermal phenomena in canopies and influence the direction of future studies.

The report presents in a fairly complete manner, the procedures used to determine the temperature distribution in vertical and horizontal places from the thermal scanner imagery. The method to calibrate the imagery is very important and must be carefully done. The comparison of this temperature distribution with thermistor measurements of the canopy air temperature as a function of elevation and of the soil surface are in good agreement. The conclusions drawn from this study indicate the thermal scanner is a powerful tool for observing thermal phenomena in the canopy and relating such observations to what would be remotely sensed.

The primary objective of the experiments to be described is to obtain simultaneous measurements of the temperature distribution in the wheat canopy by three methods: thermal scanner, thermistors, and a radiation thermometer.

Figure 2.9-1 shows the instrumentation layout at Plot 124 which was typical of the Test Plots 7, 8, 9 and 10 that were studied also. A limited amount of data was taken on the Commercial Field test site located adjacent to the UNO-Williston Agricultural Test Station.

Four thermistor stakes were placed between the canopy rows in the second and eighth rows. Each pair of row stakes was placed at the right and left edges of the scene as viewed by the thermal scanner. Cables to each stake were laid behind the next row and then toward the canopy edge to connect the digital printer via a junction box. The thermistor probes were located at the canopy heights shown in Table 2.9-1. Thermistor #1 was placed 2 to 3 centimeters into the soil at each stake location; the advantage gained by covering the thermistor is to eliminate direct solar irradiance which could cause systematically high temperature readings of the soil surface. During the observations on the Small Grain Test Plots #7 through #10, only one thermistor stake per plot was used since all three plots were scanned at the same time.

The radiation thermometer, Barnes Model PRT5, was placed on a tripod near and to the side of the canopy. For the majority of measurements, the PRT-5 was aimed at the front center of the wheat canopy, and connected to a digital data printer such that readings were recorded each time a thermistor stake was printed. The standard procedure was to take a set of thermistor and PRT-5 readings approximately every 10 minutes while the thermal scanner was recording. In several cases, the thermistor data was recorded every 15 to 20 minutes through periods when the thermal scanner was not operating. The Data Summary tables in each of the Mission Reports give the PRT-5 and thermistor data observations for each plot and day on which experiments were performed.

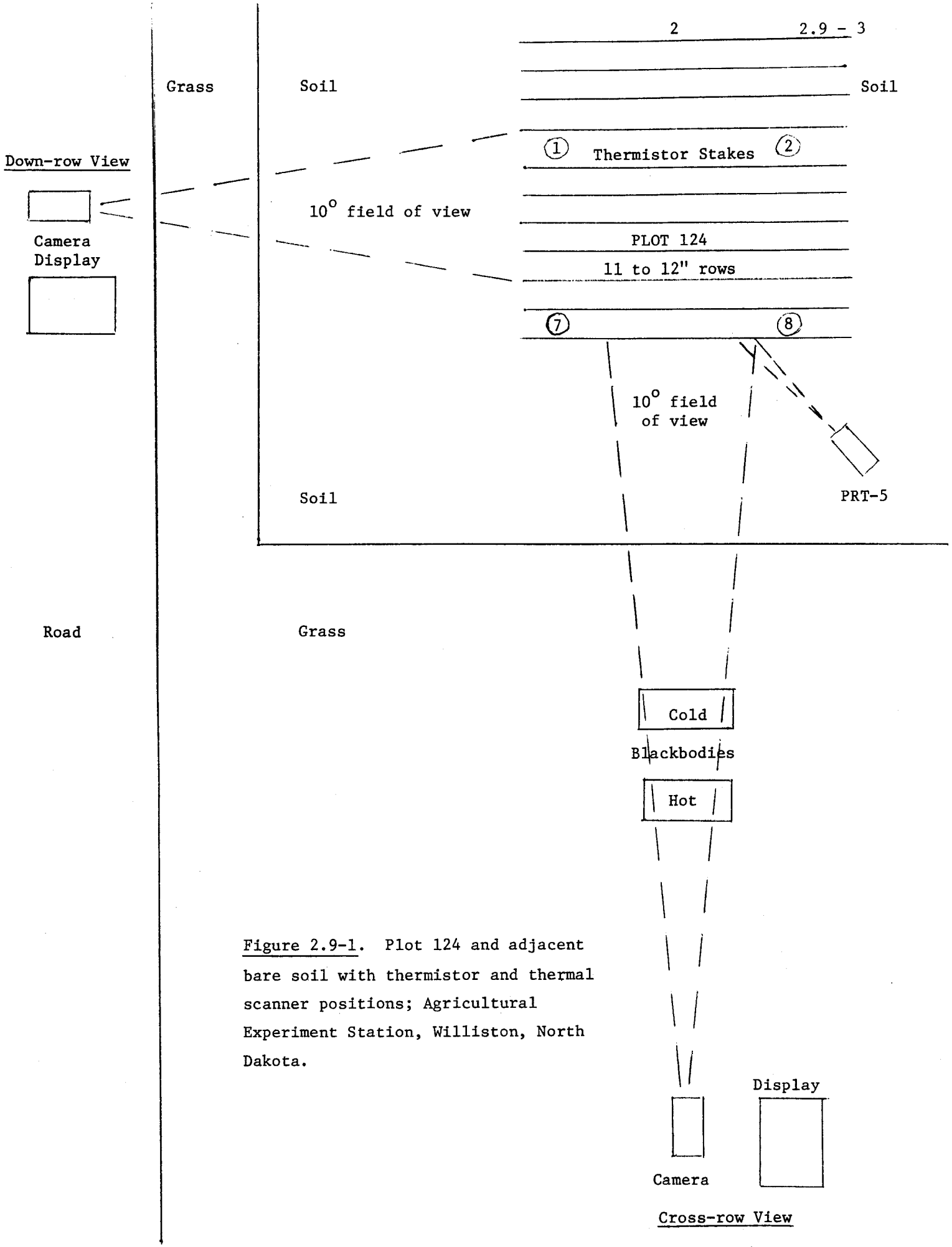


Figure 2.9-1. Plot 124 and adjacent bare soil with thermistor and thermal scanner positions; Agricultural Experiment Station, Williston, North Dakota.

Table 2.9-1. Thermistor Height Placement

Thermistor Identification (#)

Elevation above Soil Surface(cm)

Plot No.	#1	#2	#3	#4	#5
124	70	5	25	50	100
7	65	5	25	50	90
8	75	5	27	53	100
9	45	5	18	36	75
10	57	5	23	46	85
Commer.	90	5	30	60	105

The thermal scanner was used in a number of positions to provide a variety of views of the wheat canopy. Figure 2.9-1 shows the position of the thermal scanner to obtain imagery of the side view of the canopy. The imagery of the side of the canopy included the hot and cold blackbodies in the lower right and left corners respectively of the field of view. This arrangement provided imagery from which the vertical temperature profile, showing variations from the soil to the top of the canopy, could be discerned.

The thermal scanner was relocated at the left side of the canopy to obtain down-row views as shown in Figure 2.9-1. On three occasions, the elevated platform of the Hi-Ranger folding a boom truck was available to position the thermal scanner 6 to 9 meters directly above the canopy. The overhead views were taken over the thermistor stakes in order that air temperature profile data would be available for analysis.

The imagery obtained directly from the thermal scanner must be processed using calibration information to obtain quantitative temperatures in the scene. The direct calibration of the thermal scanner using blackbodies at known temperatures in the field of view has proven to be the most reliable method. The entire set of thermal imagery for the wheat experiments were recorded on video tape and calibration of imagery during playback requires considerably more effort because of the automatic gain control feature in the video recorder.

The most direct approach to image calibration is to utilize reference temperatures in the scene in the form of hot and cold blackbodies of near unity emissivity which are at very stable temperatures. The data summaries in each of the Mission Reports gives these reference temperatures throughout each day of the experiments.

Calibrating a thermal image consists of matching these temperature references to the signal value or intensity of the two blackbody images in the scene.

The problem in the processing of imagery stored on the video tapes comes from two sources. The video tape recorded is AC coupled to the thermal scanner which allows the zero setting or apparent origin to drift. The video tape recorder has an automatic gain control feature permitting the scale factor of the recorded signal to also drift. To complicate matters further, power line noise causes jumps and spikes which create similar disturbances in the origin and scale factors. Because of these effects, the quality of the imagery was at times poor and so that in our analysis it was necessary to be selective.

The calibration procedure used varied with the type of display. The most direct approach is to photograph (usually Polaroid black/white for this analysis method and positive/negative transparency for digitization as discussed below) line scans from the thermal scanner display. The line so displayed is a trace of signal intensity along a selected line through the image. The line scans are matched to an arbitrary grid to obtain numerical values, usually placing the origin at the cold blackbody temperature. Measuring the hot blackbody signal intensity by the grid gives the scale factor (^{o}C/grid unit size) where the cold blackbody temperature becomes the origin. A simple linear temperature calibration is adequate in most cases. The temperatures of other objects in the scene are found by measuring the signal intensity against the same grid, multiplying by the scale factor, and adding the cold blackbody temperature. In our initial field observation only one blackbody, usually the cold one, was placed in the scene; this provided the level of the temperature origin and we relied on the scanner display unit sensitivity setting

to give the scale factor. This was a satisfactory procedure for preliminary studies but will not provide reliable, consistent results as can be obtained when using two blackbodies in the scene.

Digitized thermal imagery is calibrated in an analogous fashion. A transparency of a thermal image, obtained by photographing the thermal scanner (video playback) display screen, is converted by an optical digitizer into an array of numbers representing intensity. The intensity value of the cold blackbody is used as the origin, and the scale factor is derived from the difference between hot and cold blackbody values and their known temperatures. This method was used in the digitized imagery to obtain the data shown in Figure 2.9-1 and Figure 2.9-2 of the September 1 - November 30, 1976 Quarterly Report.

The body of data collected during the three Missions is relatively large and includes a variety of scenes and ambient conditions. The immediate objective of our first analysis study was to demonstrate methods for determining temperature profiles in a typical canopy.

A number of side views of the wheat canopy were recorded from cross-row, down-row and 30^{o} angle positions. It became apparent early in the work that the cross-row view would produce a high quality measurement of the thermal profile of the canopy. Thermal imagery will be given in the Final Report. The scanner was located 7.6 meters from the front of the canopy such that the scene is approximately 1.3 meters wide and 1.3 meters high. The wheat canopy occupied the middle 0.8 meters of the scene with the cool sky above and warm soil in the foreground below.

Processing functions in the scanner display unit allow the selection and plotting of signal amplitude for a single scan line. Using only the cold blackbody for reference, a temperature scale can be associated with signal amplitude,

thereby producing the temperature profile. The scale is in terms of radiance temperature, not yet corrected for secondary effects of emissivity and background radiance.

The shapes of these profiles are similar to those obtained from thermocouple measurements others of wheat leaves on a canopy and similar to thermistors sensing canopy air temperatures as will be subsequently discussed. The peak of the temperature profile represents the very hot soil in the foreground which is fully sunlit in front of the canopy. Soil surface temperature within the canopy is best measured from overhead views rather than attempting to infer this temperature from the profile in the region between the foreground and the canopy.

The direct display of temperature profiles from the line scan features of the thermal scanner is quick and easy but suffers shortcomings as discussed previously. This approach toward obtaining temperature profiles does not account for the significant local variation of temperature that occur in the canopy. It would be preferable to average the profile, horizontally, across the width of the scene. Such a space averaged profile is more stable and probably more representative of thermal phenomena in the canopy than any individual line temperature profile would be.

The thermal imagery is in a better form for quantitative analysis when a digital format is used. A digital image consists of a two dimensional array of integers corresponding to the film density across the image photographed from the scanner display screen. The imagery was taken at 100 lines resolution; at 7.6m viewing distance a pixel in the image corresponds to approximately 10mm square on the canopy scene. The digitizer was set for the spatial resolution resulting in digital images of 102 rows by 110 columns. The film density at

each point is converted to integers in the 0 to 255 range; that is a grey scale with 256 levels. Due to the limited opacity in the transparency, the actual range was limited from 48 to 255.

Some of the data were presented in the last Quarterly Report. The general feature of the profile is that the top of the canopy is close to ambient air temperature, and the base is 1.5 to 2.0 C warmer than the adjacent air. This is the expected profile for wheat under minimum moisture availability.

The spatial radiance temperature profile and radiation thermometer temperatures have not been corrected for emissivity and background radiance. Assuming the canopy emissivity is 0.97, this correction alone will raise the temperature profile approximately 0.3 to 0.5^{o}C.

The side view thermal imagery can be expected to be very useful in crop research. Much of the available research in evapo-transpiration attempts to relate a single radiance temperature to the target crop. The examples presented here should make it obvious that thermal profiles are more appropriate and are likely to reveal more information about energy utilization and distribution.

The overhead thermal imagery taken from the Hi-Ranger folding boom truck provides a vertical view, as from the aircraft scanner, but with greater detail.

The predominant effect in these overhead views is the strong appearance of the hot soil between rows. The radiance temperature variations from row-to-row are small with respect to the soil radiance temperature. The implication here is quite important to interpretation of aircraft thermal imagery. The aircraft spatial resolution is not high enough to identify rows. Crop and soil are blurred together to give a radiance temperature which is higher than that of the crop.

Our analysis of thermal imagery has demonstrated that initial impressions about what a radiation thermometer or aircraft scanner detects is not necessarily the complete picture.

It is not a simple matter to interpret the thermal image side view of a canopy. A crop canopy does not have a solid or plane front surface. The image is due to radiation form the first row, a lesser contribution from the second row, and so on. An experiment was designed to evaluate the two dimensional effects present in thermal image of the canopy. That is, the objective was to determine the canopy depth (volume) which has an influence on the thermal image and hence the infrared temperature. Images of a small blackbody just in the front of the canopy and one placed several rows deep were analyzed. Line scans were used to evaluate the apparent temperature of the small blackbody versus the row depth. Line scans behind successive rows showed the progressive disappearance of the blackbody behind the wheat rows.

This experiment with the small blackbody was repeated locating the body in the lower, middle and upper third of the wheat canopy. The apparent radiance temperature of the small blackbody is then plotted as a function of row depth.

Each set of temperatures for the low, middle, and high portions of the canopy data sets are fitted to exponential extinction curves. The curve representing the thickest portion of the canopy. The extinction curve for the highest section in the canopy decays toward that mean temperature slowly, indicating a much thinner area around the heads and stalks. This variation in canopy thickness implies that a thermal image includes most of the first six rows near the top of the image, but only the front two or three rows near the base. The thermal image hence represents not a front, one-dimensional

place, but some depth averaged radiant temperatures of the canopy. This leads to the questions then of whether the front rows are significantly cooler (or hotter) than the temperatures within the canopy and will be discussed at the end of this section.

An extinction coefficient can be determined from the cold blackbody apparent temperatures observed as various row depths within the canopy according to the relation

$$T(n) = T_b e^{-kn} + T_c[1-e^{-kn}]$$

where T(n) is the apparent temperature of the small blackbody when located in the nth row. The temperature of the blackbody observed both radiometrically and with a mercury thermometer is T_b. The canopy radiance temperature is T_c, that is, the temperature determined from the thermal scanner imagery. The extinction coefficient is k and has the units of $(\text{row})^{-1}$; since the row spacing is known it would be easy to convert this form of the extinction coefficient to the more usual reciprocal length units. It is convenient to use the optical depth, N, defined as the inverse of the extinction coefficient which has physical interpretation of representing the mean depth or rows that are most significant in the extinction process.

The extinction coefficients in the small blackbody experiment are listed in Table 2.9-2. This extinction coefficient should not be confused with a similiarly named term used in sunlight penetration models as it is unique to this experiment. From a physical point of view we expect that the vertical variation in the extinction coefficient should bear a close relation to the vertical distribution of biomass. Such agronomic information was not obtained concurrent with the observations because of the extensive effort that must be required. However, a photographic method could be utilized wherein the

Table 2.9-2. Horizontal Canopy Extinction Coefficient

Location in Wheat Canopy	Extinction Coefficient $(rows)^{-1}$	Optical Depth (row)
High Third	.34	2.9
Middle Third	.43	2.3
Low Third	.79	1.3

image of succeeding deep rows with a white background would provide a measure of the distribution.

Technical Problems and Solutions. Development of a radiation model for prediction of canopy radiance temperature in terms of radiance temperature was delayed until data reduction procedures, discussed above, were clearly established. Development of the model will resume in the next quarter.

Plans for the Next Quarter. Data reduction will be completed and work on the radiation model will continue. The Final Report will be prepared.

2.10 Super Site Data Management

Major Activities. The quarter's activities include implementation of some of the verification procedures to verify and document the quality of the LACIE Field Measurements data. The objective of this work is to begin to formulate written material which can be used by researchers using the data to ascertain what procedures were used to collect the data and what environmental and system conditions were at the time data was collected.

FSS (S191-H) Spectrometer Data. Software has been implemented for an evaluation of the FSS data using the calibration panel data. The calibration panel data for each date will be normalized with respect to sun angle. There are usually three or four observations of the calibration panel during each mission. A large variation in the normalized data can indicate possible problems in the data for a particular mission.

This information is being incorporated with irradiance strip charts when available to access how good the mission day was from a meteorological standpoint.

A study of the FSS data to date does reveal that there is a 2-6% offset at .68μm. This possible problem has been noted as one for further review. (A similar problem occurs in the Exotech 20C data, they may be related).

Aircraft Multispectral Scanner Data. Information concerning the verification of the scanner data is being compiled. Software to further evaluate the data will begin to be implemented this next quarter. The information being completed at present includes histograms of data (check A-D converter) and reviews of imagery. The software to be implemented will include quatitative checks on the A-D converter system, scanner noise, and line-to-line and channel-to-channel synchronization.

Truck-Mounted Spectrometers. Verification of the Exotech 20C data was mainly concerned with evaluating how well the algorithm to correct for the tape recorder WOW in the 1976 data works. The information indicates that the algorithm works very well. A document is being compiled which discusses the tape recorder WOW problem, the algorithm used to correct the data, and information illustrating how effective the correction was.

Also, during this quarter the wavelength calibration of the Exotech 20C spectrometer was examined (it has not changed in the past four years), the performance of the solar port on the spectrometer (used to measure incident radiation) was evaluated, and the offset between the silicon and lead sulfide detector data was studied. The offset problem is discussed further in the next section.

Spectrometer Data Correlation. During this quarter work was begun in correlating the truck-mounted spectrometers. The truck spectrometers include the Earth Resources Laboratory (ERL) Exotech 20D, the NASA/JSC FSAS (VISS) and the Purdue/LARS Exotech 20C. All three truck-mounted spectrometers have collected data over the five gray panels which are deployed at the intensive test sites for the aircraft and helicopter systems. The data collected by the spectrometers over the gray panels can be used to compare the spectral measurements of the different spectrometers.

Plots of the spectral measurements are shown in Figures 2.10-1 to 2.10-5 for gray panels #1 to #5, respectively. Gray Panel #1 is the lightest and Gray Panel #5 is the darkest. Each plot represents the average of all spectral measurements by a given spectrometer for a given year. The variation in the measurements about the average was the largest for the ERL Exotech 20D Panel #1 measurements, a standard deviation around 6 percent (bidirectional reflectance factor - BDRF). The other variations were much smaller - standard deviations less the 2 percent (BDRF). See Table 2.10-1.

Preliminary analysis of the data shown in Figures 2.10-1 to 2.10-5 indicate that there is a large variation in the different spectrometer measurements of panel #1. The variation may be caused in part by a variation in the actual reflectance of the panel, since the panel is washed every six months or so. Some of the instruments will measure the panel before it is washed and some after it is washed. Information concerning the washing dates are known and work is planned this next quarter to quantitatively study the change in reflectance from washing to washing. Also, the margin of error associated with the 1975 Exotech 20D data is probably larger than that for the other instruments, since an optical misalignment occurred in the middle of their data collection at the Garden City, Kansas Agriculture Experiment Station during 1975. A review of the gray panel data indicates that the misalignment may have occurred around June 1, since before this date the gray panel measurements are consistent.

The preliminary analysis of the gray panel data also indicates that the Purdue/LARS Exotech 20C BDRF data for 1975 may be two to five percent high in the 0.4 to 1.3μm range and the data for 1976 may be two to five percent high in the 0.4 to 0.7μm range. An offset is present in the Exotech 20C field data around 0.7μm. This is the position where the silicon detector data ends and the lead sulfide detector data begins. All efforts to date to find the cause have been unsuccessful. With the information available at the time the 1975 data were processed, everything indicated that the data from 0.7 to 1.3μm was low. Therefore, the 1975 data in this range were adjusted to match the silicon detector data at 0.7μm. The information available now indicates that this may have been a wrong conclusion; the silicon detector data may be too high. The 1976 data were not "offset" adjusted; therefore, only the 0.4 to 0.7μm data appear high. Work is underway presently to prove or disprove that the reflectance measured by the instrument is too high from 0.4 to 0.7μm.

FSS (S191-H) data is available at present for only Gray Panel #3. The data is very consistent with that from the other spectrometers in the region from 0.7 to 2.4μm. The FSS data is lower in the 0.4 to 0.7μm range. No reason is known at this time. Only one sample of gray panel data is available at this time. More FSS gray panel data are needed for evaluation before any significant statements can be made.

Except for the cases mentioned above, the gray panel data collected by the different spectrometers are very similar. Efforts to quantify how similar the data are are shown in the correlation plots in Figures 2.10-6 and 2.10-7 and the linear regression analyses given in Table 2.10-2.

The correlation plots and the linear regression analyses include data for six wavelengths scattered throughout the wavelength spectrum being measured. If the data from the instruments were correlated perfectly, the points would fall on a straight line ($r^2 = 1$ for regression analysis), have a slope of one ($b = 1$) and pass through the origin ($a = 0$).

The information indicates that the gray panel data from the instruments are highly correlated. The r^2 values are high - around .999 for the FSAS vs Exotech 20C comparison and .970 to .999 for the Exotech 20D vs Exotech 20C comparison. The gray panel spectrometer data appears to be the most closely correlated in the 0 - 30 percent (BDRF) range, i.e. Gray Panels #2 to 5. Efforts are underway to further improve the spectrometer correlation. These are namely, 1) determine reason for offset in Exotech 20C data, 2) determine reason(s) for larger variation in data collected by different spectrometers over Panel #1 than that in the data for the other four gray panels, 3) planned activity in mid-May to have the three spectrometers, the Exotech 20C, FSAS, and FSS, in Finney County, Kansas at the same time for direct instrument comparison experiments.

Further Information Needed. It would be very helpful to have any other available S191-H gray panel data to further develop the correlation of the S191-H and the truck-mounted spectrometers.

Summary and Conclusions. The analysis to date of the correlation of the 1974-76 spectrometer data is not complete; therefore, no final conclusions can be made in this report. The task of developing and following procedures to calibrate and correlate the data from different spectrometers in a field environment is very difficult. This task is being accomplished and for the first time quantitative information is available.

The analysis results so far in comparing four spectrometers of two different types (filter wheel and interferometer), collecting data in locations hundred of miles apart (Kansas and North Dakota) and at different times of the year are very encouraging. The reflectance data for the four gray panels used to calibrate the aircraft scanner data has a correlation better than .99 and agrees to within 4% of value. The reflectance data for panel #1 is within 10% of value for the FSAS and Exotech 20C data at two different sites with the panel on two different platforms. Further examination for systematic differences in past and future data will include a side by side comparison of instruments and a re-examination of the procedures used during observation of panel #1, the helicopter panel.

Table 2.10-1. Variation in Spectrometer Measurements of Reflectance of Gray Panels.

Year-Instrument	Gray Panel+ #1	#2	#3	#4	#5
1975 Purdue/LARS Exotech 20C	3.6++	1.7	1.4	0.9	1.2
1975 ERL Exotech 20D	6.0	1.5	1.5	0.7	0.7
1976 Purdue/LARS Exotech 20C	1.1	1.1	1.2	0.5	0.6
1976 NASA/JSC FSAS (VISS)	1.9	*	*	*	*

+ Panel #1 is the brightest. Panel #5 is the darkest.

++Largest standard deviation in units of bidirectional reflectance factor (percent) in the data from 0.5 to 1.4μm.

* Data for Panels #2 to 5 were obtained from Will White over the phone. He indicates that this data (collected 10/76) may need to be updated with a more recent calibration table.

Figure 2.10-1. Comparison of Spectrometer Measurements of Gray Panel #1.

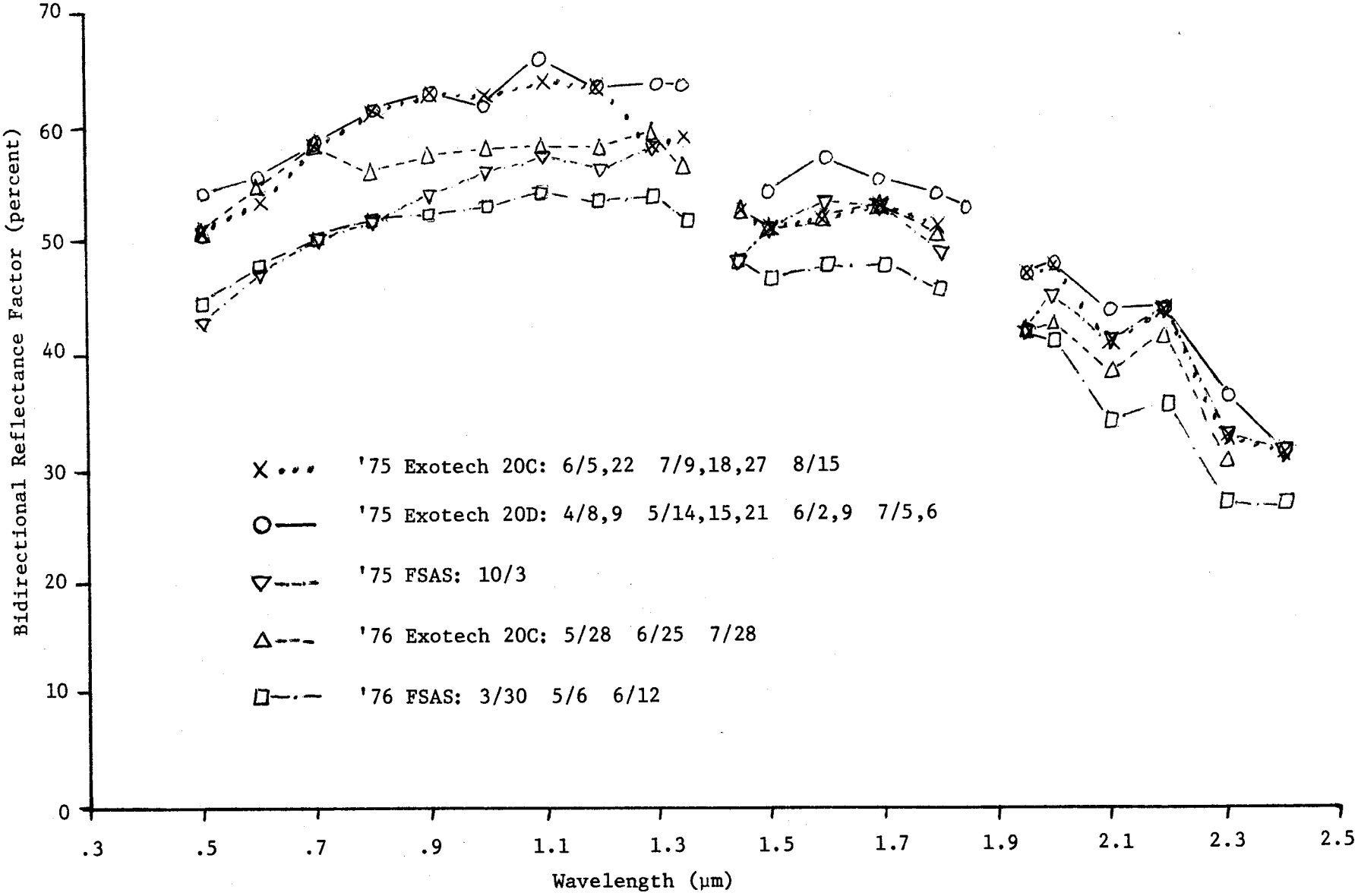


Figure 2.1--2. Comparison of Sepctrometer Measurements of Gray Panel #2.

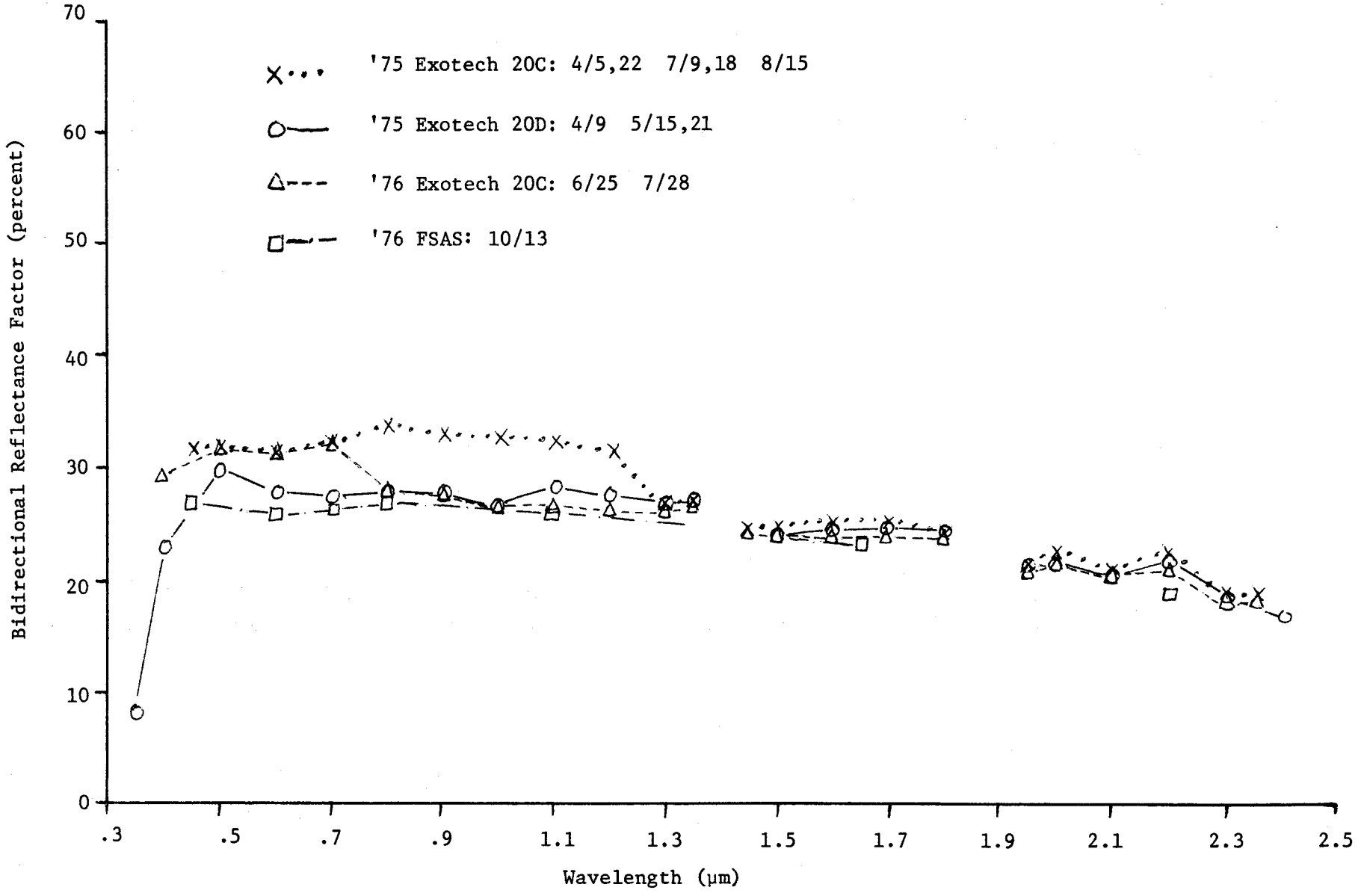


Figure 2.10-3. Comparison of Spectrometer Measurements of Gray Panel #3.

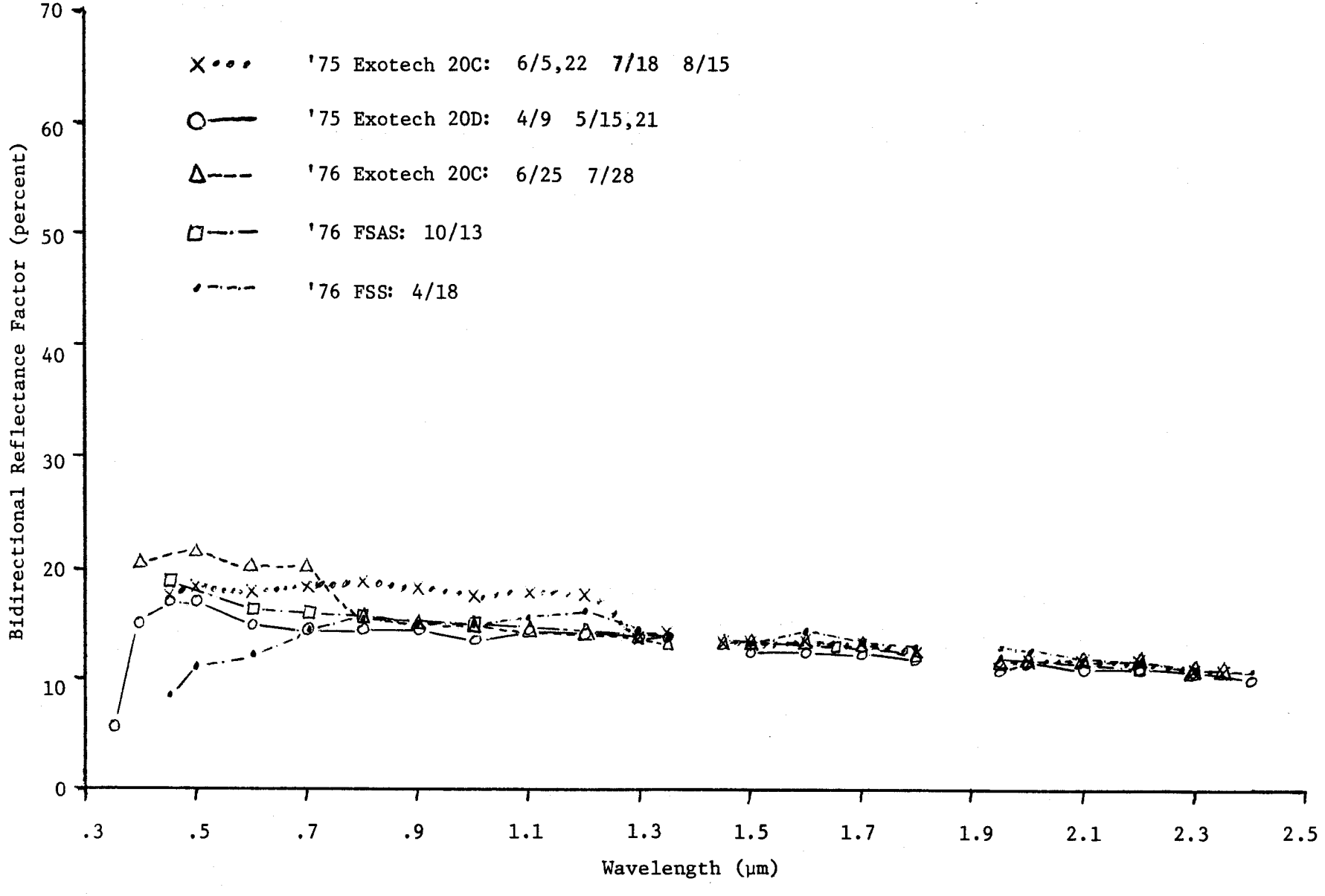


Figure 2.10-4. Comparison of Spectrometer Measurements of Gray Panel #4.

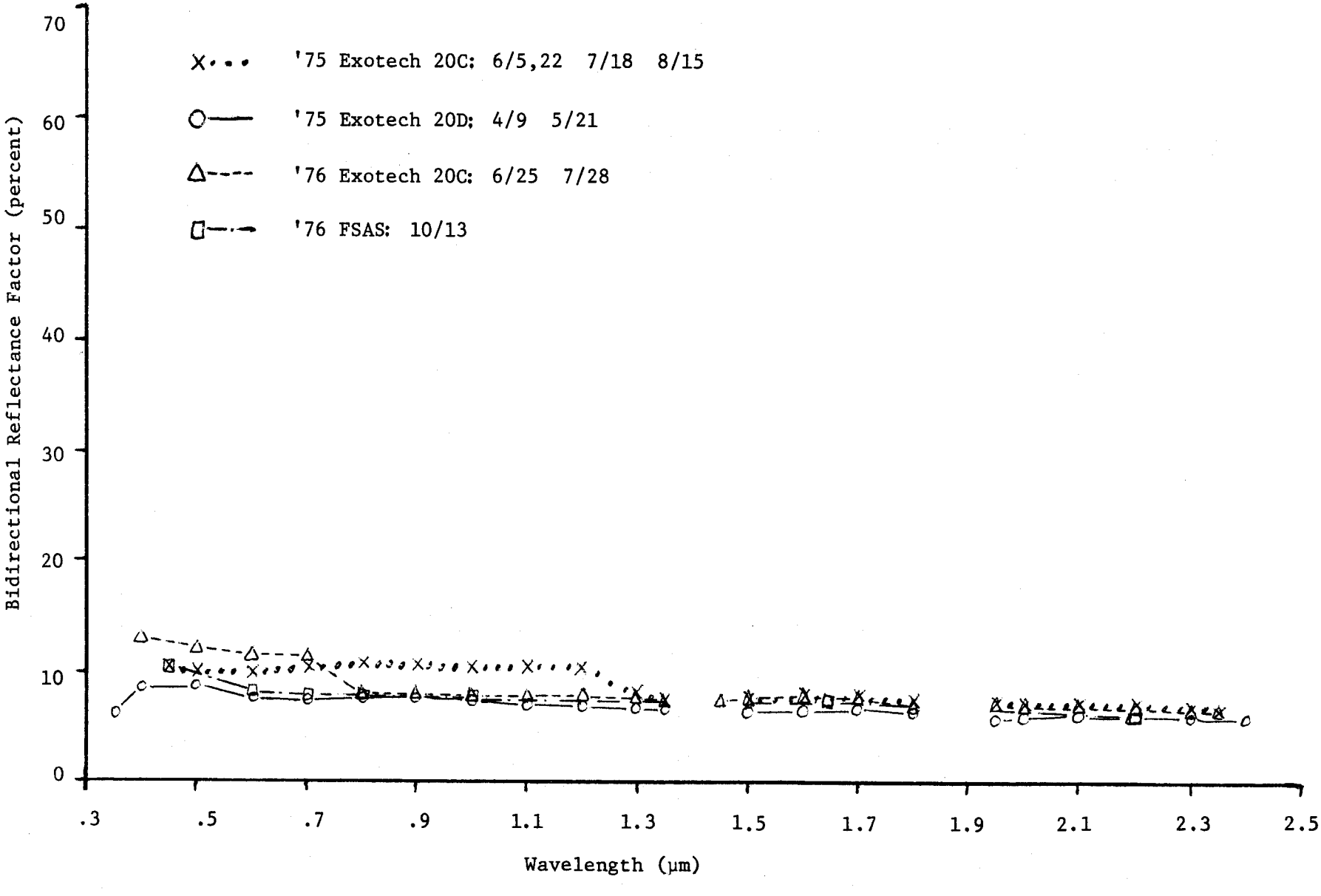


Figure 2.10-5. Comparison of Spectrometer Measurements of Gray Panel #5.

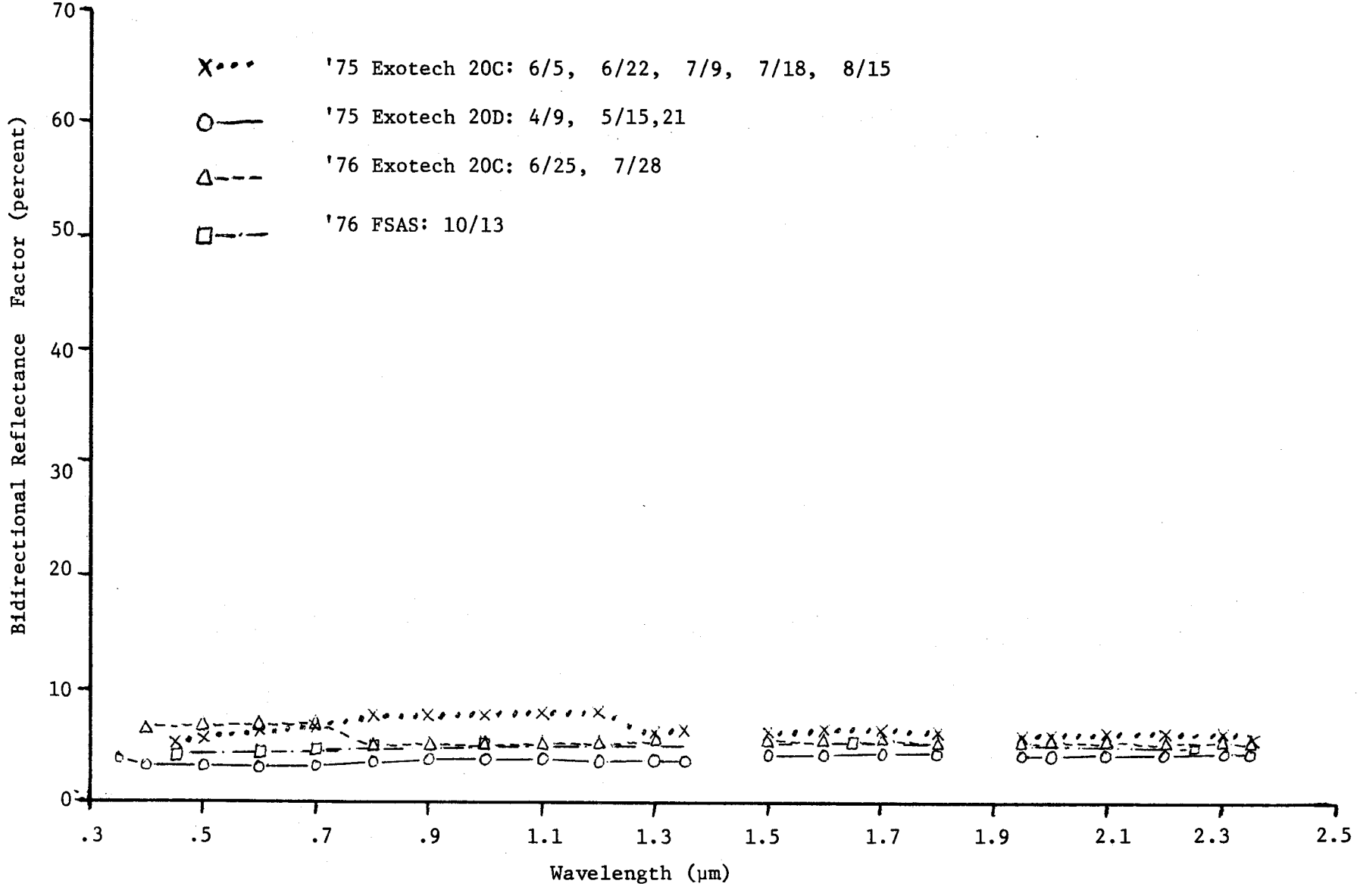


Figure 2.10-6 Correlation Plots of 1975 Gray Panel Spectrometer Data. ERL Exotech 20D vs Purdue/LARS Exotech 20C.

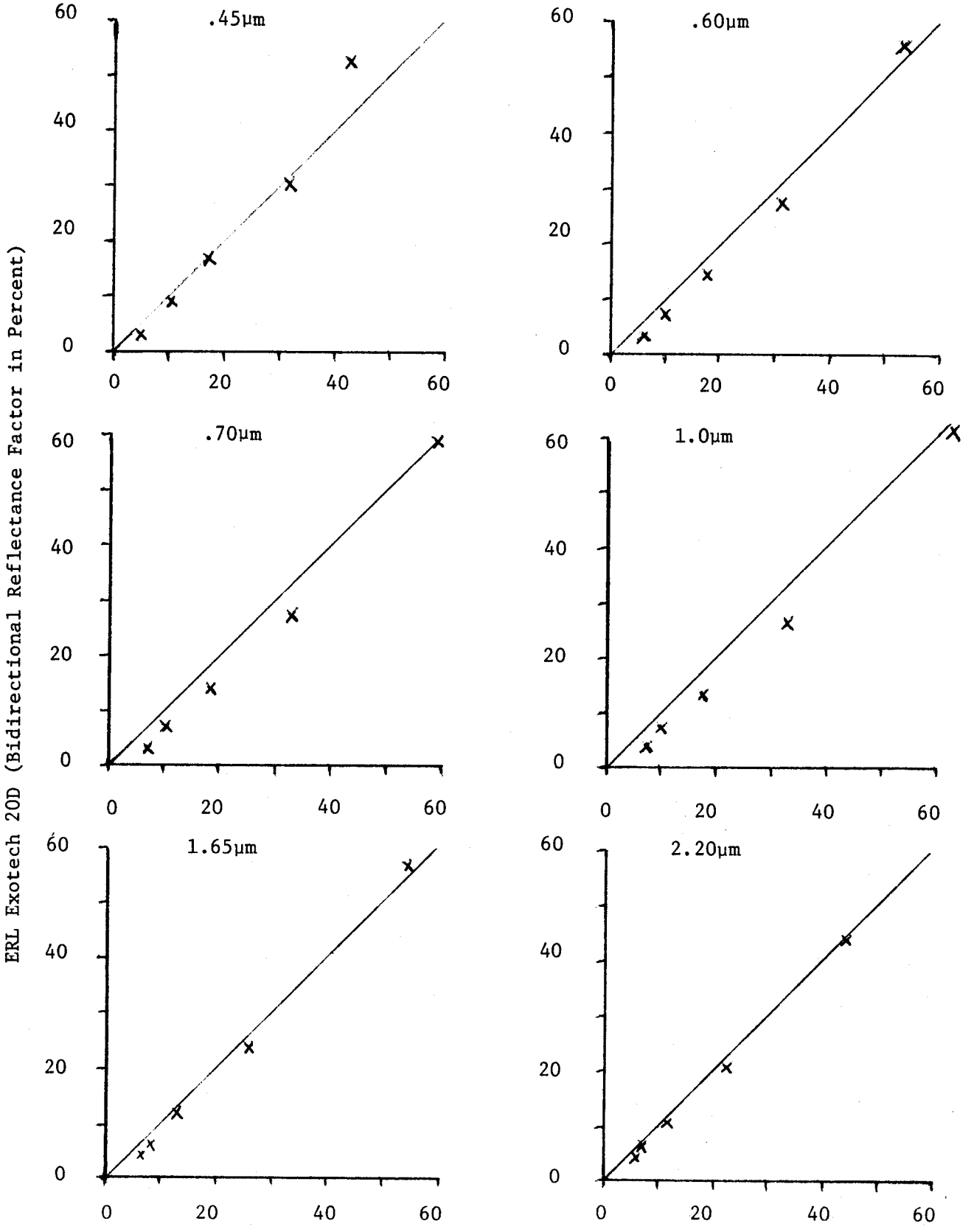


Figure 2.10-7. Correlation Plots of 1976 Gray Panel Spectrometer Data. NASA/JSC FSAS (VISS) vs Purdue/LARS Exotech 20C.

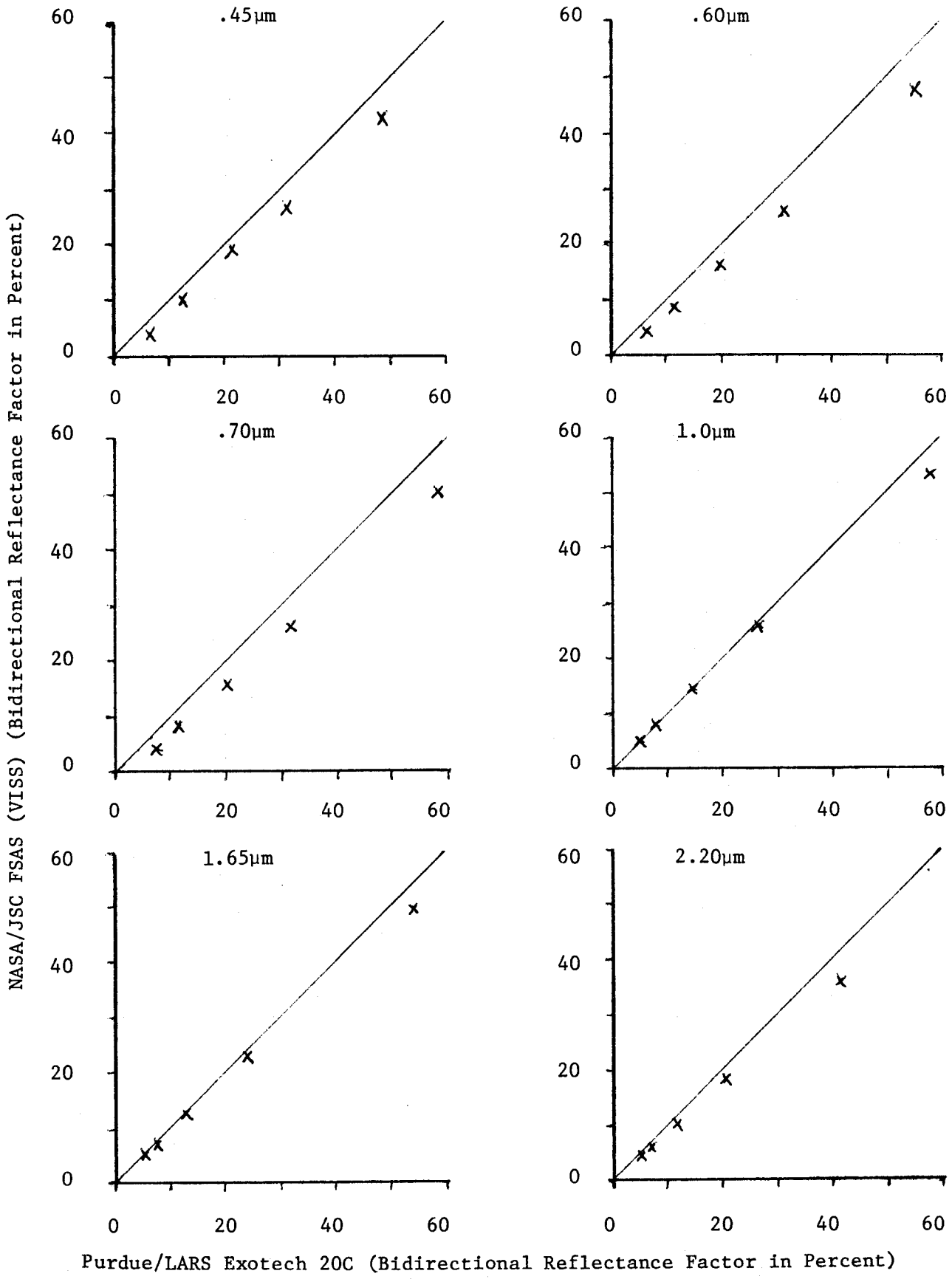


Table 2.10-2. Preliminary Linear Regression Analyses of Spectrometer Gray Panel Data*.

BDRF+ Range (percent)	Regression Coefficients	Wavelength (μm) 0.45	0.60	0.70	1.0	1.65	2.20
		ERL Exotech 20D vs Purdue/LARS Exotech 20C (1975 Data)					
0-60	r^2	0.972	0.995	0.996	0.995	0.999	0.998
	a	-4.6	-4.5	-4.6	-4.7	-2.5	-1.7
	b	1.260	1.111	1.054	1.035	1.083	1.035
0-30	r^2	0.994	0.999	0.999	0.998	0.998	0.995
	a	-1.6	-2.7	-2.8	-2.5	-1.7	-1.0
	b	1.018	0.980	0.935	0.888	1.012	0.964
		NASA/JSC FSAS(VISS) vs Purdue/LARS Exotech 20C (1976 Data)					
0-60	r^2	0.998	1.000	1.000	0.999	1.000	0.999
	a	-1.5	-1.9	-1.8	1.3	0.9	0.5
	b	0.915	0.898	0.882	0.900	0.903	0.849
0-30	r^2	0.994	1.000	1.000	0.999	0.999	0.999
	a	-1.6	-1.5	-1.5	0.7	0.3	0.1
	b	0.919	0.870	0.862	0.950	0.949	0.892

*Regression Equation: $y = a+bx$

+Bidirectional Reflectance Factor

2.11 Soil Classification and Survey

Spectral Relationships. Recent results of soil spectral studies indicate that more attention to soil-forming factors is necessary for future research on the relationships between reflectance and physical/chemical properties of soils. During the past decade several limited studies have been conducted on soil reflectance. Condit (1,2) studied the reflectance in the 0.32 to 1.0μm spectral range of 160 American surface soils in both wet and dry conditions. He concluded that the spectral curves he examined could be classified into three general types with respect to the shapes of their curves. No attempt was made by Condit to relate the spectral measurements to other characteristics of the soils other than color.

Shields, et al. (10), investigating the quantitative effects of moisture and organic matter on soil color, obtained spectral curves in the near ultraviolet and visible region of the spectrum of two Great Groups: Chernozemic (Mollisol soil order) and Gray Wooded (Alfisol soil order). They found that the relationship between soil color and the organic carbon content of the Chernozemic soils differed significantly from that of the Gray Wooded soils.

At LARS Montgomery (8) 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 properties 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 NASA-funded 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.

Although a great body of knowledge has been accumulated about the characteristics of soils as they are influenced by the soil-forming factors of climate, parent material, relief, biological activity, and time, there is very limited knowledge about the potential for identifying and characterizing soils by digital analysis of multispectral radiation from surface soils.

Recent studies using computer-implemented pattern recognition analysis of Landsat multispectral data have indicated that spectral separability of surface soils can be used to delineate soils having different internal drainage characteristics (9), differences which were not visually discernible in the field or on black and white aerial photographs. Other studies have revealed that surface soils appearing the same in the field in Jasper County, Indiana, were spectrally separable into two soils derived from different parent materials, one from glacial till, the other from lacustrine deposits (4).

Several other investigators (3,5,6,7,11) have used multispectral data to delineate and map meaningful soils differences, but as yet there is limited understanding of the effects which annual mean precipitation, annual mean 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 significant effects of soil background on crop identification and yield predictions.

When the proposal for this task (2.11 Soil Classification and Survey) was developed a year ago the results and conclusions of Montgomery's research were not available. In the light of his conclusions it became most important to exercise great care and statistical integrity in the selection of soil samples for further study of the problem as outlined in the proposal. A replacement for Dr. Montgomery was accomplished on 1 September 1976 and work began on sample selection and implementation of the research outlined in the proposal.

In order to achieve the research objectives a valid statistical sampling of soils of the United States must be obtained from:

(1) Three significantly different mean annual precipitation zones

(2) Three significantly different mean annual temperature zones

(3) Major parent material regions

(4) The ten soil orders, the highest level or categorization of soils according to the U.S. System of Soil Taxonomy.

During recent months we have made steady progress in studying and selecting soils to meet the above criteria. The soils to be included in this study are being selected from a list of 1,377 Benchmark Soils which have been identified, studied and described in detail by the Soil Conservation Service, U.S. Department of Agriculture.

When the proposal was written, it was also assumed that cooperation with the Soil Conservation Service would continue to the extent that this Agency would supply LARS with samples of soils selected for this study. As soils were selected for the study, it became apparent that many of the desired soils would have to be resampled since the supply of many specific samples in the SCS Regional Laboratories was depleted or quite old. It became necessary to delay the acquisition of soil samples until the 1977 field work season when SCS technicians will be engaged in field survey and can easily obtain the desired samples and provide detailed profile descriptions.

The consequences of this delay are that we have concentrated on a field study of the spectral variations from soils in Indiana.

Outdoor Exotech Experiment. Considerations of sun elevation and equipment availability require that any experimentation on soil spectral properties in a field condition be conducted from late April to early May 1977. An experimental design has been developed which will allow the investigation of the effects of organic surface residue (corn stover) and soil moisture content on the reflectance of two soils which differ greatly in surface soil color, organic matter content, and natural drainage.

Russell silt loam, a light colored soil developed under forest vegetation, and Chalmers silty clay loam, a dark colored soil developed under tall grass prairie, will be studied. Plots will receive treatments of saturation with water (a) several hours before and (b) one day before reflectance measurements are made. Organic residue to corn stover will be applied at a rate of 10 metric tons per hectare on half of the plots. Four treatment combinations will result for each soil: 1) moist bare soil, 2) drier bare soil, 3) moist soil with corn stover residue, and 4) drier

soil with corn stover residue. Surface roughness will be minimized to reduce shadow effects, and readings with the Exotech Model 20-B spectroradiometer will be made as close to solar noon as possible to reduce sun angle effects. Measurements will be made on successive days to verify the repeatability of the responses. Surface soil samples collected at the time of Exotech readings will be analyzed for moisture content, organic matter content, and other physical/chemical properties.

It is hoped that the results of this experiment can help explain some of the soil spectral differences that are commonly observed in Landsat data from late April and early May when organic residue and moisture content may result in different spectral responses from the same soil mapping unit.

Analysis of Landsat Data. Four test sites in Tippecanoe County, Indiana were selected for evaluation and comparison of the use of computer-implemented analysis of multispectral data from aircraft and Landsat scanners to delineate meaningful soils boundaries. These sites represent soils formed under deciduous hardwood forests on glacial till, soils formed under tall prairie grasses, and soils formed in a transition zone between the prairie and forest soils.

Multispectral analysis is proceeding well with the first two test sites. Soils of Location A were formed under deciduous hardwood forests on glacial till. Location B is in a transition zone between soils developed under deciduous hardwood forests and those developed under tall prairie grasses. The following soils are included in the test areas:

Location A (25 hectares)

Kokomo silty clay loam	Typic Argiaquoll
Brookston silty clay loam	Typic Argiaquoll
Metea silt loam	Arenic Hapludalf
Toronto silt loam	Udollic Ochraqualf
Del Rey silt loam	Aeric Ochraqualf
Fincastle silt loam	Aeric Ochraqualf
Xenia silt loam	Aquic Hapludalf
Russell silt loam	Typic Hapludalf

Location B (40 hectares)

Reeseville silt loam	Aeric Ochraqualf
Celina silt loam	Aquic Hapludalf
Crosby silt loam	Aeric Ochraqualf
Brookston silt loam	Typic Argiaquoll
Brookston silty clay loam	Typic Argiaquoll
Ragsdale silty clay loam	Typic Argiaquoll
Toronto silt loam	Udollic Ochraqualf

Only the Brookston and Toronto soil series were common to the two test areas.

The aircraft multispectral scanner (MSS) data (spatial resolution = 4.68m^2 or 0.0004ha) were acquired on 25 June 1969 at an altitude of 1,300m. Landsat-2 data (spatial resolution = 4,500m^2 or 0.45ha) were obtained on 20 June 1976 at an altitude of 915km.

A standard soils map of the two test sites (A and B) was prepared as a part of the ground observation task. A sampling grid plan was used to select a training set for a supervised classification of the aircraft MSS data. The Landsat MSS data were analyzed by both nonsupervised (or cluster) and supervised pattern recognition techniques.

Accuracy of each of the classifications was evaluated by comparing them with ground observation data. The spectral patterns revealed in the classifications from aircraft and satellite MSS data resembled the general patterns of the soils of the conventionally prepared soil maps. The

spatial resolution of the aircraft scanner was adequate to recognize each soil mapping unit in both test sites. However, the limited spatial resolution of the Landsat data made it difficult to delineate those soil features with widths less than the spatial resolution of the scanner (approximately 70 meters).

For Location A a comparison between the boundaries of the Russell and Xenia soils of the soil map and the delineations on the computer-generated spectral map revealed slight differences in mapping unit boundaries. A field check of these differences suggested strongly that the spectral delineations for the Russell and Xenia series were more accurate than the mapping unit boundaries on the conventional soil map.

Kokomo silty clay loam was correctly classified and Toronto silt loam was reasonably well delineated on the spectral classification except in some transitional zones. In general Fincastle silt loam soils were identified and mapped spectrally the same as on the conventional soil map.

In summary for test sites A and B those soils patterns which were broad enough to exceed the spatial resolution of the Landsat scanner were delineated very well by spectral analysis. Although it was impossible and perhaps invalid to make a detailed comparison between the multispectral data from the aircraft and from Landsat, analysis results of the data from both scanners suggest that the variations in the signal from the aircraft scanner may be more difficult to explain than those of the satellite scanner.

Comparison of spectral analysis results for test sites C and D will be completed in March. Although a spectral map of the entire county has been produced, precision correction of two dates (9 June 1973 and

20 June 1976) is being done to provide two new spectral maps of the county for the purpose of delineating meaningful soils boundaries and for further comparisons with the results from aircraft scanner data and conventional soil maps. The two dates selected represent relatively wet and dry soil conditions and were obtained when most of the cultivated soils in Tippecanoe have little or no surface cover.

References

1. Condit, H. R. 1970. The spectral reflectnace of American soils. Photogrammetric Engineering 36:955-962.

2. Condit, H. R. 1972. Application of characteristic vector analysis to the spectral energy distribution of daylight and the spectral reflectance of American soils. Applied Optics 11:74-86.

3. Gausman, H. W., A. H. Gerberman, C. L. Wiegand, R. W. Leamer, R. R. Rodriguez, and J. R. Noriega. 1975. Reflectance differences between crop residues and bare soils. Soil Sci. Soc. Amer. Proceed. 39:752-755.

4. Kirschner, F. R. 1977. Personal communication. Laboratory for Applications of Remote Sensing. Purdue University.

5. Kristof, S. J. and A. L. Zachary. 1974. Mapping soil spectral features from multispectral scanner data. Photogrammetric Engineering 40:1427-1434.

6. Leamer, R. W., V. I. Myers, and L. F. Silva. 1973. A spectroradiometer for field use. Rev. Sci. Instrum. 44:611-614.

7. Mathews, H. L., R. L. Cunningham, J. E. Cipra, and T. R. West. 1973. Application of multispectral remote sensing to soil survey research in southeastern Pennsylvania. Soil Sci. Soc. Amer. Proceed. 37:88-93.

8. Montgomery, O. L. 1976. An investigation of the relationship between spectral reflectance and the chemical, physical and genetic characteristics of soils. Ph.D. Thesis. Purdue University.

9. Peterson, J. B. 1976. Personal communication. Laboratory for Applications of Remote Sensing. Purdue University.

10. Shields, J. A., E. A. Paul, R. J. St. Arnaud, and W. H. Head. 1968. Spectrophotometric measurement of soil color and its relationship to moisture and organic matter. Canadian J. Soil Sci. 48:271-280.

11. Westin, F. C. and C. J. Frazee. 1976. Landsat data, its use in a soil survey program. Soil Science Society of America Journal 40:81-89.

Information Notes Issued
During Period
December - February

091576 Description and Evaluation of a Bidirectional Reflectance Factor Reflectometer by D. P. DeWitt and B. F. Robinson.

The Note describes the LARS reflectometer for making bidirectional reflectances factor measurements on large area (30x30cm) samples in the 0.38 to 2.5μm spectral region. This reflectometer simulates field measurement conditions for studying the effects of solar zenith angle and viewing direction on remote sensing observation of targets. Results for typical sample surfaces - paints, soil and cloth - are presented and discussed.

The research reported in this paper was sponsored by NASA under Contract NAS9-14016 and NAS9-14970.