

10

Made available under NASA sponsorship  
In the interest of early and wide  
semination of Earth Resources Survey  
Program information and without liability  
for any use made thereof."

STIF

Special

E7.4-10573

CR-138678

ABSTRACT

ORSER-SSEL Technical Report 9-74  
THE PENN STATE ORSER SYSTEM FOR PROCESSING AND ANALYZING ERTS AND OTHER  
MSS DATA  
G. J. McMurtry, F. Y. Borden, H. A. Weeden, and G. W. Petersen

The Office for Remote Sensing of Earth Resources (ORSER) of the Space  
Science and Engineering Laboratory (SSEL) at The Pennsylvania State University  
has developed an extensive operational system for processing and analyzing  
ERTS-1 and similar multispectral data. The ORSER system was developed for  
use by a wide variety of researchers working in remote sensing. Both photo-  
interpretive techniques and automatic computer processing methods have been  
developed and used, separately and in a combined approach. A Remote Job  
Entry (RJE) system permits use of an IBM 370/168 computer from any compati-  
ble remote terminal, including equipment tied in by long distance telephone  
connections. An elementary cost analysis has been prepared for the processing  
of ERTS data.

Original photography may be purchased from  
EROS Data Center  
10th and Dakota Avenue  
Sioux Falls, SD 57198

(E74-10573) THE PENN STATE ORSER SYSTEM  
FOR PROCESSING AND ANALYZING ERTS AND  
OTHER MSS DATA Interim Report  
(Pennsylvania State Univ.) 86 P HC  
N74-28817  
Unclas  
CSCL 05B G3/13 00573

Reproduced by  
NATIONAL TECHNICAL  
INFORMATION SERVICE  
US Department of Commerce  
Springfield, VA. 22151

PRICES SUBJECT TO CHANGE

88

**Interim Report**

ORSER-SSEL Technical Report 9-74

**THE PENN STATE ORSER SYSTEM FOR PROCESSING AND ANALYZING ERTS  
AND OTHER MSS DATA**

**G. J. McMurtry, F. Y. Borden, H. A. Weeden, and G. W. Petersen**

**ERTS Investigation 082  
Contract Number NAS 5-23133**

**INTERDISCIPLINARY APPLICATION AND INTERPRETATION OF ERTS DATA  
WITHIN THE SUSQUEHANNA RIVER BASIN**

**Resource Inventory, Land Use, and Pollution**

**Office for Remote Sensing of Earth Resources (ORSER)  
Space Science and Engineering Laboratory (SSEL)  
Room 219 Electrical Engineering West  
The Pennsylvania State University  
University Park, Pennsylvania 16802**

**Principal Investigators:**

**Dr. George J. McMurtry  
Dr. Gary W. Petersen**

**Date: June 1974**

## MCMURTRY, BORDEN, WEEDEN, AND PETERSEN

### Computer Processing Facilities

Automatic data processing equipment utilized in the ORSER system is primarily located at The Pennsylvania State University Computation Center. The principal computer is the IBM System 370 Model 168, consisting of a main frame and attached devices for input and on-line storage. Users may have access to the computer in any of three ways: central and remote high speed dispatch points operated by the Computation Center; slow speed Remote Job Entry (RJE) terminals using IBM 2741, Tektronix 4010, or similar remote terminals supported by the user or by the Computation Center; and intermediate speed remote batch terminals, such as the IBM 2780, supported by the user or the Computation Center. The ORSER processing system for MSS data was developed for use with any of these entry points. ORSER investigators use RJE terminals for most developmental work. Bulk output for final runs is directed from an RJE terminal to any of the high speed terminal sites. No program card decks need to be input, as the MSS data processing programs are kept in library files. Files for building control information or for storing output are available to the user. MSS data is input from magnetic tapes which, along with user-owned working tapes, are managed by the Computation Center.

Non-University users as well as Penn State users may join the system. Any compatible terminal may be used to process data via the RJE system, including equipment at non-University Park locations tied in by long distance telephone connections.

### The Digital Data Processing System

A standard digital tape format was designed within which all known MSS sources can conveniently be placed. More than one file per tape is allowed as well as a continuation of a file to another tape. Within the file, four kinds of records exist: (1) identification records, (2) table of contents records, (3) MSS response records, and (4) history records. Each MSS response record consists of a complete scan line. Each scan line is numbered and scan lines are always in ascending order in a file. A working file will usually contain one or more small parts of the whole data set. The table of contents is particularly useful in such cases in avoiding costly searching for data which are not present in the file.

The system is couched in a multivariate framework. Each observation, identifiable by scan line and element number, consists of a vector with as many components as there are channels. At present, each vector is composed of just MSS

## PENN STATE ORSER SYSTEM FOR PROCESSING MSS DATA

response values. It is anticipated, however, that the vectors will be augmented by transformed scanner data, or by additional (nonscanner) data such as topographic information.

Although the system is not in a conversational mode, where the user and the system dynamically interact during processing, the preparation of the control specifications by a user operating from an RJE terminal is conversational. Each program accepts input control specifications, processes the MSS data according to the specifications, and outputs the results. For non-RJE operation, control specifications are made and entered into the system by punched cards. All control specifications on the RJE are identical in format to the corresponding punched cards. Typical turn-around time from an RJE terminal is less than two minutes, while for card deck submission it is about five minutes.

The programs discussed here are all operational and are documented at the user's level.<sup>1</sup> Although many other programs<sup>2</sup> are used, those discussed here illustrate the general approach to the processing of MSS tapes.

The digital tape processing system for MSS data described here is regularly run for production and has been extended to meet the needs of various related projects. The system was designed to be easily augmented, typified by the addition of a number of supervised and unsupervised analysis and classification algorithms. The general procedure to be employed for a previously unstudied area or type of target will be presented and illustrated here.

The first step is to select the particular targets and area of interest, primarily using maps. Consultation of the catalogues of imagery<sup>3</sup> and digital tapes<sup>4</sup> will indicate what data are available and their quality. Tapes corresponding to the selected scenes are chosen and the areas of useful data are specified. The data for these areas are then produced as subsets on separate tapes, using the SUBSET program. It is likely that this step has already been done in the process of cataloguing and storing ERTS tapes by ORSER, in which case the appropriate library subset tapes would be selected directly. Subsetting accomplishes:

- 1) reduction of a large data set to small subsets containing only the data of interest;
- 2) reformatting of ERTS or other MSS data to the common ORSER format;
- 3) doubling of data tape density from 800 b.p.i. to 1600 b.p.i. thus doubling the data rate in subsequent use;

## MCMURTRY, BORDEN, WEEDEN, AND PETERSEN

- 4) reduced computing cost; and
- 5) reduced turn-around time for subsequent runs because of reduced run time. (This also results in higher priority run assignments.)

A run is then made with the NMAP program to show the overall pattern of the data. This program is written to map element brightness, using all channels or any subset of channels. The norm of each multivariate vector is taken as the measure of brightness. The norm is then converted to a percentage of the maximum possible value. This value is translated to the mapping symbol for the percentage range within which it falls. The process is repeated for every element in every scan line in the data blocks specified by the user. Output from the NMAP program consists, then, of a brightness map. This is similar to a gray-scale map, but NMAP does not employ expensive and time-consuming techniques such as overprinting. Careful choice of mapping symbols, however, will result in maps with readily distinguished degrees of brightness.<sup>5</sup> These maps are useful for initial target location, verification of general location, etc. It is interesting and important to note that the program requires no a priori knowledge of target spectral signatures or other characteristics.

The UMAP program is run next in order to identify areas of local spectral uniformity. Each element is compared with its near neighbors using the euclidean distance between spectral signatures as the measure of similarity or dissimilarity. If the largest distance is smaller than a value specified by the user, then the symbol for uniformity is assigned to that element. One or more categories of uniformity can be mapped according to distances specified by the user. All elements with distances from their neighbors greater than those specified are mapped as contrasts. The map output shows the pattern of uniformity and contrasts from which the user can designate coordinates of training areas for supervised classifiers. It may also be used to determine high contrast boundaries between uniform areas.

Signatures and associated statistics are next obtained by the use of the STATS program, which computes the multivariate statistics for one or more training areas obtained from UMAP or similar output. The user designates for each identifiable category, a training area by line and element coordinates and the program computes the statistics for all of the data which fall within the boundaries. The mean and standard deviation vectors for each category are found, and the correlation and variance-covariance matrices are computed as well as the

PENN STATE ORSER SYSTEM FOR PROCESSING MSS DATA

eigenvalues and eigenvectors of these matrices. Frequently histograms for selected channels are also computed.

When most of the target categories have been identified by training areas, a classification run is made using the classifier or classifiers deemed most appropriate for the mix of targets under consideration. A variety of supervised classification programs are available, including parametric and non-parametric classifiers with either linear or quadratic discriminant functions. Preprocessing before classification is also possible, using programs for normalization, principal components, canonical analysis, etc. The output of these programs is in the form of a digital character map, with each category of classification represented by a unique symbol. Digital character maps are useful primarily as working maps for the user in the analysis of MSS data. They are inherently distorted in the length-to-width relationship because of the fixed number of lines and characters per inch of high-speed printer output.<sup>5</sup>

The LMAP program, yielding output on the CalComp plotter or the RJE terminals (Tektronix 4010) with graphic displays, is intended for the production of distortion-free, finished copy, line maps. There are three main advantages to line maps when compared to character maps: (1) orthographic maps to a selected scale can be made, (2) photographic overlays can be prepared for these maps (this is quite important in the comparison of classification results with corresponding imagery), and (3) legible maps for publication purposes can be prepared.

An example of the use of the programs described above is given in Figs. 1 through 8. The MSS data used for this analysis came from ERTS-1 scene 1028-15295, scanned on August 20, 1973. This is an area northeast of Clearfield, Pennsylvania, on which U.S. Route 80 and the West Branch of the Susquehanna River cross. The location of the test site is shown in Fig. 1, which was taken from two 7 1/2 minute USGS quadrangle maps. The right hand side of this figure is from a map printed in 1959, before Route 80 was constructed, while the left hand side is from a 1971 map. Figures 2 and 3 show map output for NMAP and UNMAP, respectively. The strip of low brightness in Fig. 2 follows the river, as does the blank (non-uniform) area shown in Fig. 3. Basic statistics for the "stripmine" category, obtained by the STATS program are shown in Fig. 4. Statistics for the desired categories, as obtained from a series of sample sites, are input to a classification program. Figure 5 shows the output from the DCLASS program which classifies according to a minimum euclidean distance algorithm. In this case, only two general categories are represented by



PENN STATE ORSER SYSTEM FOR PROCESSING MSS DATA

BLOCK SPECIFICATIONS

BEGINNING LINE 1030
ENDING LINE 1059
BEGINNING ELEMENT 2470
ENDING ELEMENT 2525
LINE INCREMENT 1
ELEMENT INCREMENT 1

2470|2475|2480|2485|2490|2495|2500|2505|2510|2515|2520|2525|
1030 |XXXXXXXXXXXXXXXX-XXXX-XXXXXXXXXX-XXXXXXXXXXXX---XXXXXXXXX|
1031 |XXXXXXXX--XX-X- XXXXXXXXXXXXXXX-XXXXXXXXXX---XXXXXXXXXX|
1032 |XXXXXXXX-XXXXX- -XXXXXXXXXXXX-XXXXXXXXXX---XXXXXXXXXX|
1033 |XXXXXXXX---XXXX- -XXXXXXXXXX-XXXXXX---XXXXXXXXXX|
1034 |XXXXXXXX-XXXXXX- -XXXXXX- -X-XXXXXXXXXX|
1035 |XXXXXXXXXXXXXXXX-XXXX- -X-XXXXXXXXXX-XXX|
1036 |X---XXXXXXXXXXXXXXXX- -XXXXXXXXXX-XX---XX|
1037 |X---XXXXXXXXXXXXXXXXXXXXXXX-XXX- -XXXXXX-XX-X---XX-|
1038 |X-XXXX- -XXXXXXXXXXXXXXXX-X- -XXXXXXXXXX- -X|
1039 |XX---XX-XXXXX-XXXXXXXXXXXXXXXXXX- -XXXXXXXXXX- -X|
1040 |XXXXXXXX-XXXXX-X- -XXXXXXXXXXXXXXXXXXXX- -XXXXXX- -X|
1041 |XXXXXX-X-XXXXXX- -XXXXXXXXXXXX- -XX- -XXXXXXXXX|
1042 |XXXXXX-XXXXXX-XXXX- -XXX- -XXXXXXXXXXXXXXXX|
1043 |XXXXXX-XXXXXXXXXXXX- - -XXXXXXXXXXXXXXXX-XX|
1044 |XXXXXX-XX-XXXXXXXXXX- -XXXXXXXXXXXXXXXX-|
1045 |XXXXXX- -X-XXXXXXXXXXXXXXXXXX- -XXXXXX-XXXXXXXX|
1046 |XXXXXXXX-XXXXXXXXXXXXXXXXXX- -XXXXXXXX-XXXXX|
1047 |XXXXXX-XX- - -XXXXXXXXXXXX-XXXXXXXXXXXXXXXX-|
1048 |XXXXXX-XXXX-XXXXXX- -XXXXXXXXXX-XXXXXXXXXX- -X-|
1049 | -XXXXXXXXXXXXXXXXXX XXXXXXXXXXX-XXXXXXXXXXXXXXXX-XXXXX|
1050 | -XXXXXXXXXXXXXXXXXX- -XXXXXXXXXXXXXXXX- -XXXXXX|
1051 |XXXXX- -XXXXXXXXXXXXXXXX- -XXX-XXXXXXXXXX- -XXXXXXXXXX|
1052 |XXXXX- -XXXXXXXXXXXXXXXX- -X-X-XXXXXX- -XXXXXXXXXX|
1053 |XXX- -XXXXXXXXXXXXXXXX- -XXXX-X-XXXXX- -XXXXXXXXXXXXXXXX|
1054 |XXXXX- -XXXXXXXXXXXXXXXX- -XXXXX-XXXXXX-XX-XXXXXXXXXX|
1055 |XXXXX- -XXX-XXXXXXXXXX- -XXXXX-XXXXXX- -XXXXXXXXXX|
1056 |XX-XX- -XXXXX-XXXXXX- -XXXXX-XXXXXXXXXXXXXXXXXX|
1057 |XXXXXXXX-XXXXXXXX-XXXXXX- XXX- -XXXXXXXXXX-XXXXXXXXXX|
1058 |XXXXXXXXXXXXXXXXXXXXXXXXXX- -XXX- -XXXXXXXXXX-XXXXX- -X|
1059 |XXXXXXXXXXXXXXXX-XXXXXX- -XXXXXXXXXXXXXXXX-XXXXX-XX-XXX|

2470|2475|2480|2485|2490|2495|2500|2505|2510|2515|2520|2525|
CLASS : 14.0 CLASS - : 20.0 CLASS X : 100.0

Figure 2: Brightness map (NMAP).

MCMURTRY, BORDEN, WEEDEN, AND PETERSEN

BLOCK SPECIFICATIONS

BEGINNING LINE 1030
ENDING LINE 1059
BEGINNING ELEMENT 2470
ENDING ELEMENT 2525
LINE INCREMENT 1
ELEMENT INCREMENT 1

Table with columns for line numbers (1030-1059) and element numbers (2470-2525). The cells contain alphanumeric characters representing data points for a uniformity map.

CLASS 0 : 5.0 CLASS - : 8.0
CLASS : 15.0 CLASS + : 100.0

Figure 3: Uniformity map (UMAP).

PENN STATE ORSER SYSTEM FOR PROCESSING MSS DATA

CHANNELS USED : 1 2 3 4

MEANS AND STANDARD DEVIATIONS FOR GIVEN CHANNELS

29.55	26.86	30.78	13.83
3.29	4.80	5.21	3.79

VARIANCE-COVARIANCE MATRIX

10.82			
14.91	23.09		
0.78	0.23	27.16	
-3.27	-5.68	18.00	14.33

CORRELATION MATRIX FOR GIVEN CHANNELS

1.00			
0.94	1.00		
0.05	0.01	1.00	
-0.26	-0.31	0.91	1.00

EIGENVALUES COMPUTED FROM CORRELATION MATRIX.

EIGENVALUES WITH THEIR ASSOCIATED PERCENTAGES:

2.21	1.70	0.06	0.03
55.4	42.4	1.4	0.8

EIGENVECTORS:

0.5227	-0.4645	-0.7059	-0.1129
0.5416	-0.4368	0.6347	0.3362
-0.3600	-0.6257	0.2327	-0.6402
-0.5376	-0.4453	-0.2115	0.6814
DET. =	0.7050-02		

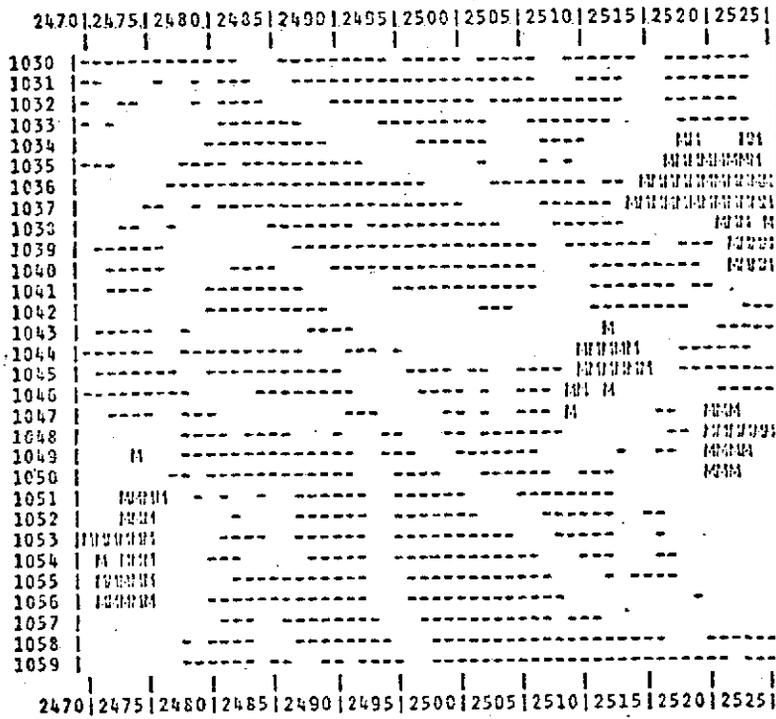
SAMPLE CATEGORY : STRIPMINE

Figure 4: Statistics of sample areas obtained by STATS for strip-mine category.

MCMURTRY, BORDEN, WEEDEN, AND PETERSEN

BLOCK SPECIFICATIONS

-----  
 BEGINNING LINE 1030  
 ENDING LINE 1059  
 BEGINNING ELEMENT 2470  
 ENDING ELEMENT 2525  
 LINE INCREMENT 1  
 ELEMENT INCREMENT 1



SYMBOL M : STRIPHINE      SYMBOL - : VEGETATION

Figure 5: Classification map from DCLASS using signatures obtained by STATS.

PENN STATE ORSER SYSTEM FOR PROCESSING MSS DATA

symbols; unclassified elements are left blank. LMAP output using data from the DCLASS program is shown in Fig. 6.

It frequently happens that a sample target is not of sufficient size or area to lend itself to categorization using the STATS program. Such targets may be linear features such as streams, or a series of small scattered features which are not large enough to be represented as uniform areas by UMAP. In such cases, these areas are defined for analysis by an unsupervised classifier which develops its own set of spectral signatures and statistics using a clustering algorithm. The map output of one such program, DCLUS, is shown in Fig. 7. A comparison of Fig. 7, with the DCLASS output of Fig. 5 reveals that DCLUS was able to map some features which could not be mapped by DCLASS with STATS signatures. Map output from DCLASS, using signatures from the DCLUS statistics, is shown in Fig. 8 in LMAP form.

The approach employed for change detection or where a temporal dimension is involved is similar to the approach for non-temporal analyses in many respects. The major difference is in the establishment of permanent training areas for analysis and classification. These areas must be selected and specified more carefully and with more refinement than when the temporal dimension is not of interest.

Hybrid Analysis

After separate analyses of ERTS-1 data by photointerpretation alone and by computer processing of MSS digital data without the assistance of photointerpretation, it became apparent that each method had shortcomings which might be overcome if the methods were combined. When applying photointerpretation techniques to ERTS imagery, in only a few cases could a feature be uniquely determined by this method alone. The use of U2, C130 and C54 imagery has been found to improve these interpretation results, but photointerpretive techniques have not been completely satisfactory as a single means of analysis. Computer differentiation of areas from scanner data is far superior to that done by the human eye. Computation of areas from the digital data makes delineation of these areas unnecessary and is far more accurate than planimetric methods at the scale of ERTS MSS imagery. Since the end result of processing ERTS-1 data is a map, the automated processes of thematic mapping by computer is the efficient way to go. However, "ground truth" is the key to correct signatures for this mapping. Underflight data and photointerpretation of underflight photography, as well as of ERTS imagery, are vital

## MCMURTRY, BORDEN, WEEDEN, AND PETERSEN

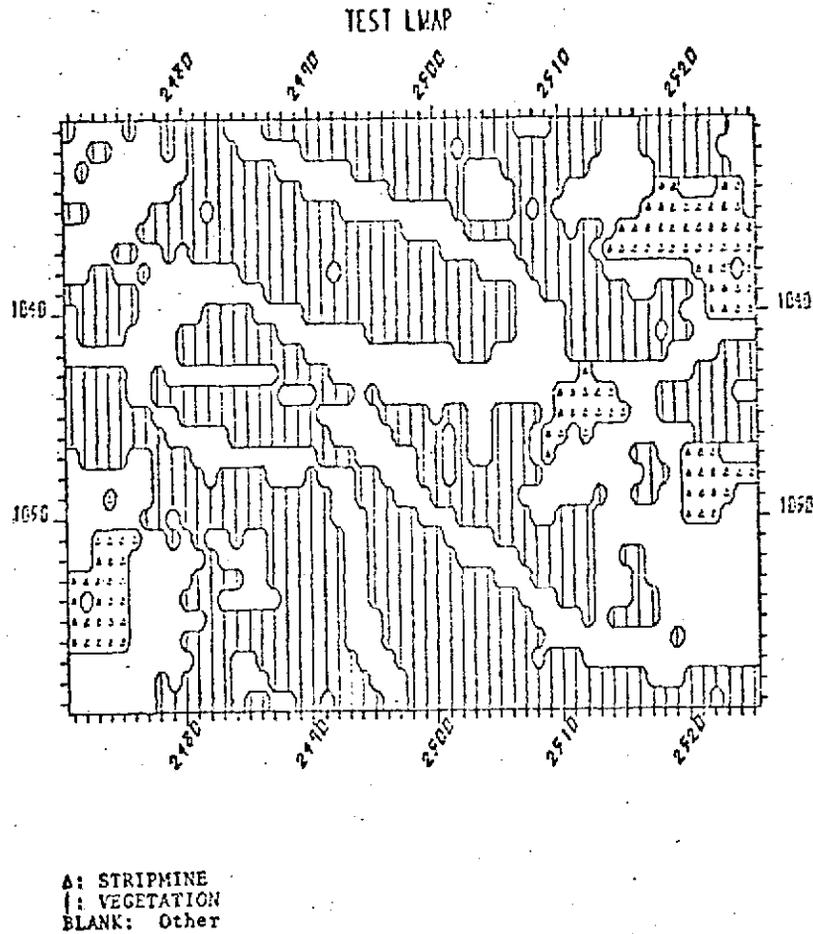
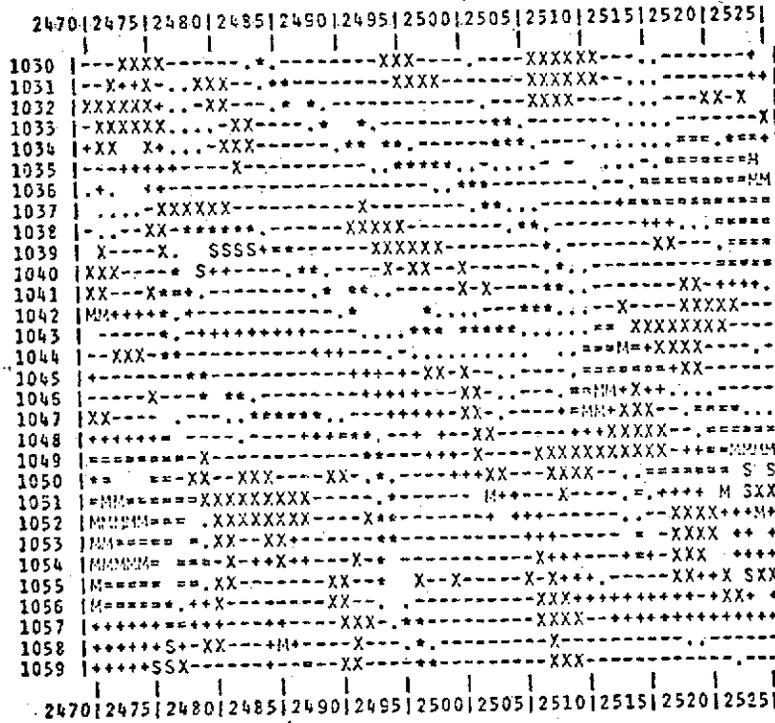


Figure 6: Preliminary classification map of the Clearfield area (LMAP).

BLOCK SPECIFICATIONS

BEGINNING LINE 1030
ENDING LINE 1059
BEGINNING ELEMENT 2470
ENDING ELEMENT 2525
LINE INCREMENT 1
ELEMENT INCREMENT 1



CHANNELS USED : 1 2 3 4
INITIAL CRITICAL DISTANCE : 8.5

Figure 7: Signature map by DCLUS.

MCMURTRY, BORDEN, WEEDEN, AND PETERSEN

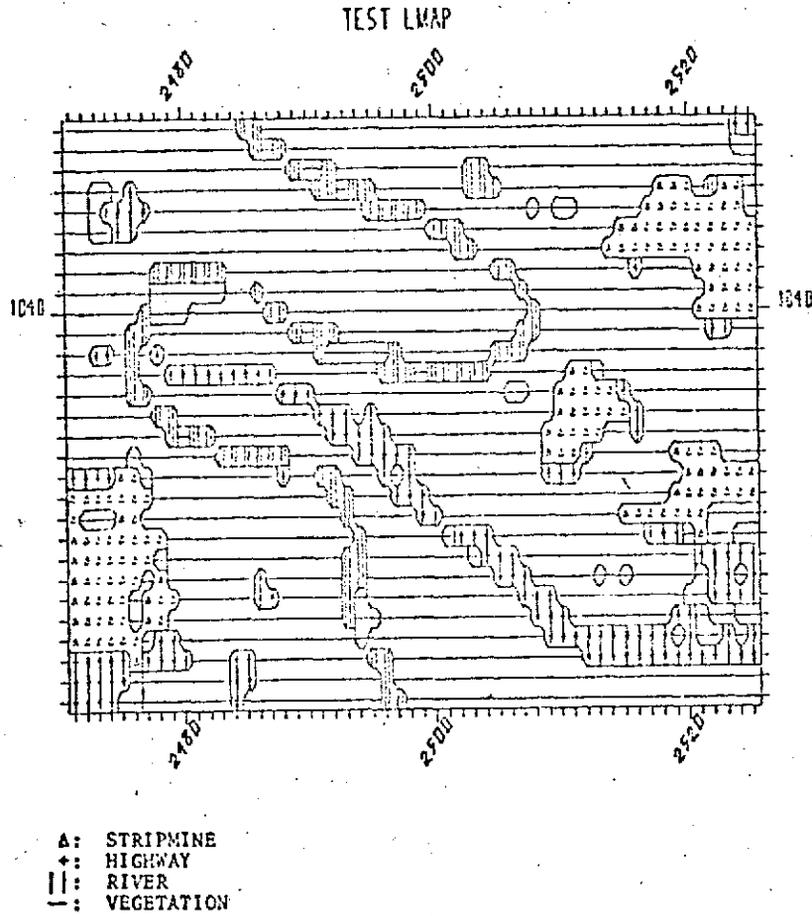


Figure 8: Classification map of the Clearfield area (LMAP).

## PENN STATE ORSER SYSTEM FOR PROCESSING MSS DATA

links leading to valid signatures for the thematic map. A marriage of these two disciplines, photointerpretation and computer processing, is essential for maximum utilization of ERTS-1 data. Thus ORSER investigators evolved a method of ERTS MSS data analysis referred to as the "hybrid approach" and shown in Fig. 9. Table 1 provides an explanation of the steps shown in Fig. 9. The letters in Table 1 respond directly to those in Fig. 9. This method involves intimate interaction of the computer analyst and the photointerpreter, using aircraft photography for comparison with the computer output. A Bausch and Lomb Zoom Transferoscope is frequently used for this comparison.

### Applications

The applications objectives of the ORSER interdisciplinary investigation using ERTS data are grouped into three major categories: (1) geology and hydrology; (2) inventory of natural resources and land use; and (3) environmental quality. Specific results obtained to date include a study of land use, discrimination between types of forest resources and vegetation, detection of previously unknown geologic faults and correlation of these with known mineral deposits and ground water, mapping of mine spoils in the anthracite region of eastern Pennsylvania, mapping of strip mines and detection of acid mine drainage effects in Central Pennsylvania, agricultural land use mapping, and detection of gypsy moth infestation.<sup>6,7</sup>

### Cost Analysis

Two major components of the cost for analyzing and interpreting ERTS-1 digital data are computing cost and personnel cost. Computing cost can be partitioned into the cost for spectral signature identification and the cost for production processing after signatures have been identified. The major personnel cost is associated with the development of signatures, since remote sensing analysts and interpreters are required for this phase. In production, much less personnel time is required, although analysts and interpreters remain closely involved in evaluating the products.

In the Susquehanna River Basin test site in Pennsylvania, two characteristics dominate the analysis and interpretation of ERTS-1 data. These are the diversity of targets and the areal smallness of target units. Compared with other areas where these characteristics are less pronounced, signature

Table 1 Explanation of Figure 9

PRELIMINARY PROCEDURES

- A. Determine scan line and element limits.
- B. This becomes the working tape.
- C. Identify clouds.
- D. Review scene for definable boundaries..

FIRST LEVEL MAPPING

- A. Collaboration of photointerpreter and computer mapper. Select easiest targets first. Choose spectrally homogeneous items with positive geographic locations. Select replications in widely separated areas.
- B. Identify training areas on NMAP and UMAP.
- C. Check for uniformity on UMAP. Attempt to find a large number of like elements. Loop A, B, and C until a sufficient number of training areas are identified.
- D. Review statistical characteristics of defined targets.
- E. Make first run on classification map.
- F. This is a verification step. Project U2 image onto computer map. Identify satisfactory classifications. If some areas lack definition, redefine training areas.

SECOND LEVEL MAPPING

- A. Attempt to identify items outside training areas.
- B. Define items not subject to definition by training areas. These might be linear features or stream channels. Add these to the list of signatures and continue.
- C. This a recycle, with smaller training areas and more weight placed on cluster analysis.

THIRD LEVEL MAPPING

- A. Review the classification categories originally defined as desirable. If present map output is unnecessarily refined, combine some groups.
- B. Some categories will require broadened spectral parameters. A series of successive approximations will be required to define these units. The resulting training areas will be less spectrally homogeneous.
- C. Requires collaboration of the photointerpreter and the computer mapper.
- D. Establish limiting goal.

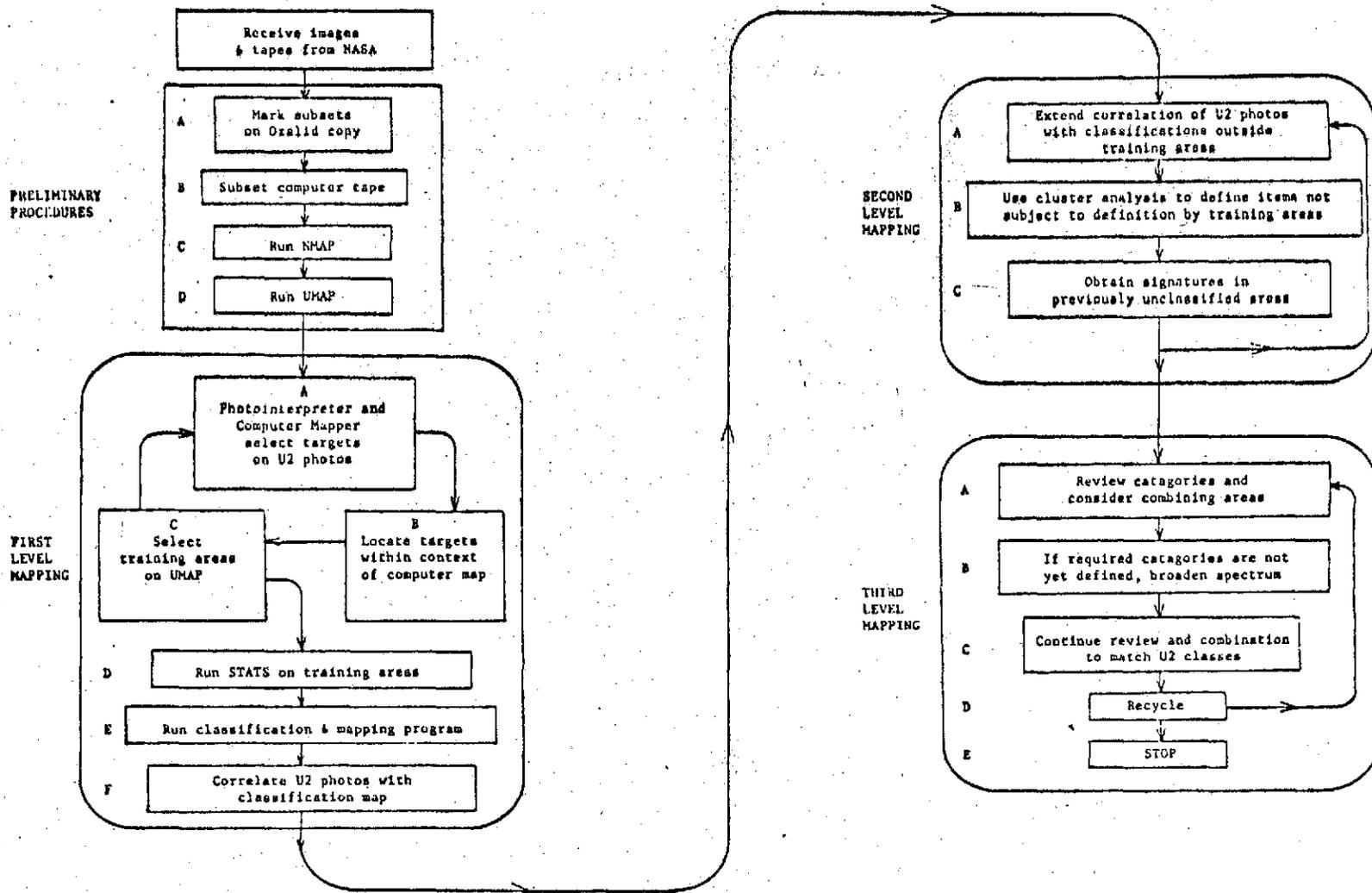


Figure 9: Flow diagram for the hybrid approach to ERTS data processing.

## MCMURTRY, BORDEN, WEEDEN, AND PETERSEN

identification presents a greater challenge and is therefore likely to be more costly. In addition, the cost for signature identification is contingent upon the nature of the particular problem to be solved and therefore it is difficult to categorically specify this cost.

The ORSER MSS computer analysis methods emphasize the minimization of computation costs by being designed for signature identification based on short computer runs on small subsets of ERTS-1 data. In the computation cost evaluations provided here, the computer run costs are based on the standard rates charged at the Computation Center of The Pennsylvania State University. (For reference purposes, the job processing rate for the IBM 370/168 at Penn State is 10¢/second, high speed output records cost 6¢/100 records, and RJE terminal connections cost 3¢/minute.)

Data for personnel and computation cost were obtained for a typical ERTS scene analysis. For the identification of 22 signatures judged necessary to meet the analysis objectives, the personnel cost was \$400 and the computation cost was \$600, a total of \$1000. Using the 22 signatures, mapping of a full ERTS-1 scene cost \$560, based on a cost of \$0.043 per square mile. Considering signature identification cost plus full-scene processing cost, the cost per square mile was \$0.12. For subsequent scenes of the same area, signature identification cost would be expected to be substantially less because, in the first time through, a great deal of personnel time and computer cost are spent in familiarization and learning for the specific area and targets. Much of this work does not have to be repeated in subsequent analyses of the same area.

Data have been accumulated for the computation costs of running different programs. For subsetting a complete ERTS-1 digital tape, the cost has averaged \$0.032/square mile. The cost of running mapping and classification programs has been found to be dependent on the number of signatures as well as the area to be mapped. The dollar cost per square mile (C) as a function of the number of signatures (S) has been found to adhere to the following formula:

$$C = S (2.56) 10^{-3}$$

For 15, 30, and 60 signatures, this cost is \$0.038, \$0.075 and \$0.154, respectively. These costs are for the production of digital character maps. If line maps are desired, additional costs are encountered in the use of the Cal Comp plotter.

## PENN STATE ORSER SYSTEM FOR PROCESSING MSS DATA

The computer time required for running mapping and classification programs is also dependent on the number of signatures and the area to be mapped. Typically, one complete ERTS scene can be classified using eight categories in approximately 45 minutes.

The above calculations of analysis and processing costs have not taken into account the cost of developing the digital computer processing system. The system was developed so that it could be easily used for processing ERTS-1 as well as any other satellite or airborne platform MSS digital data. The system development has not been financially supported by the ERTS-1 project, although it has received partial NASA support through a sustaining University grant. Extension and modification of the system has been partially supported in the ERTS-1 project.

## References

1. "ERTS and Aircraft Multispectral Scanner Digital Data Users Manual," ORSER/SSEL Technical Report 11-73, May 1973, The Pennsylvania State University, University Park, Pa.
2. "Program Descriptions," ORSER/SSEL Technical Report 10-73, May 1973, The Pennsylvania State University, University Park, Pa.
3. Bolling, N.B., "Storage and Retrieval of ERTS and Underflight Imagery," ORSER/SSEL Technical Report 12-73, May 1973, The Pennsylvania State University, University Park, Pa.
4. Borden, F.Y., Lachowski, H.M., and Applegate, D.N., "Catalogues for Remote Sensing Digital Data Tapes," ORSER/SSEL Technical Report 18-73, May 1973, The Pennsylvania State University, University Park, Pa.
5. Wilson, D.A., and Ackley, R.E., "Gray Scale Printing and Photographic Reproduction of Large Area Computer Generated Maps," ORSER/SSEL Technical Report 3-74, February 1974, The Pennsylvania State University, University Park, Pa.

## MCMURTRY, BORDEN, WEEDEN, AND PETERSEN

6. Williams, D.L., and Turner, B.J., "Computer Analysis and Mapping of Gypsy Moth Defoliation Levels in Northeastern Pennsylvania Using ERTS-1 Data," Third Annual Remote Sensing of Earth Resources Conference, March 1974, University of Tennessee Space Institute, Tullahoma, Tenn.
7. McMurtry, G.J., and Petersen, G.W., "Interdisciplinary Applications and Interpretations of ERTS Data within the Susquehanna River Basin," Type II Report, Contract NAS 5-23133, October 1973, NASA Goddard Space Flight Center, Greenbelt, Maryland.

ABSTRACT

ORSER-SSEL Technical Report 8-74  
INTERDISCIPLINARY ANALYSIS AND INTERPRETATIONS OF EARTH RESOURCES TECHNOLOGY  
SATELLITE (ERTS) DATA  
G. J. McMurtry

The objectives of this work have been to ascertain the usefulness of ERTS data in the areas of natural resources and land use inventory, geology and hydrology, and environmental quality. Specific results include a study in land use, discrimination between types of forest resources and vegetation, detection of previously unknown geologic faults and correlation of these with known mineral deposits and ground water, mapping of mine spoils in the anthracite region of eastern Pennsylvania, mapping of strip mines and acid mine drainage in central Pennsylvania, agricultural land use mapping, and detection of gypsy moth infestation. Both manual photointerpretive techniques and automatic computer processing methods have been developed and used, separately and in a combined approach.

**Interim Report**

ORSER-SSEL Technical Report 8-74

INTERDISCIPLINARY ANALYSIS AND INTERPRETATIONS OF EARTH RESOURCES  
TECHNOLOGY SATELLITE (ERTS) DATA

G. J. McMurtry

ERTS Investigation 082  
Contract Number NAS 5-23133

INTERDISCIPLINARY APPLICATION AND INTERPRETATION OF ERTS DATA  
WITHIN THE SUSQUEHANNA RIVER BASIN

**Resource Inventory, Land Use, and Pollution**

Office for Remote Sensing of Earth Resources (ORSER)  
Space Science and Engineering Laboratory (SSEL)  
Room 219 Electrical Engineering West  
The Pennsylvania State University  
University Park, Pennsylvania 16802

**Principal Investigators:**

Dr. George J. McMurtry  
Dr. Gary W. Petersen

**Date:** June 1974

20

# INTERDISCIPLINARY ANALYSIS AND INTERPRETATIONS OF EARTH RESOURCES TECHNOLOGY SATELLITE (ERTS) DATA <sup>1</sup>

George J. McMurtry

Associate Professor of Electrical Engineering  
Office for Remote Sensing of Earth Resources (ORSER)  
Space Science and Engineering Laboratory (SSEL)  
The Pennsylvania State University  
University Park, Pennsylvania 16802

## ABSTRACT

The objectives of this work have been to ascertain the usefulness of ERTS data in the areas of natural resources and land use inventory, geology and hydrology, and environmental quality. Specific results include a study in land use, discrimination between types of forest resources and vegetation, detection of previously unknown geologic faults and correlation of these with known mineral deposits and ground water, mapping of mine spoils in the anthracite region of eastern Pennsylvania, mapping of strip mines and acid mine drainage in central Pennsylvania, agricultural land use mapping, and detection of gypsy moth infestation. Both manual photointerpretive techniques and automatic computer processing methods have been developed and used, separately and in a combined approach.

2780, supported by the user or the Computation Center. The ORSER processing system for MSS data was developed for use with any of these entry points. ORSER investigators use RJE terminals for most developmental work. Bulk output for final runs is directed from an RJE terminal to any of the high speed terminal sites. No program card decks need to be input, as the MSS data processing programs are kept in library files. Files for building control information or for storing output are available to the user. MSS data is input from magnetic tapes which, along with user-owned working tapes, are managed by the Computation Center.

Non-University users as well as Penn State users may join the system. Any compatible terminal may be used to process data via the RJE system, including equipment at non-University Park locations tied in by long distance telephone connections.

## INTRODUCTION

The Office for Remote Sensing of Earth Resources (ORSER) was established by the Space Science and Engineering Laboratory (SSEL) at The Pennsylvania State University to encourage interdisciplinary research activities involving remote sensing. For nearly two years, a group of nine faculty members, along with approximately twenty graduate students, from six departments in three colleges (Agriculture, Earth and Mineral Sciences, Engineering) of the University have been working together in ORSER/SSEL on the analysis of remote sensing data obtained by the Earth Resources Technology Satellite (ERTS). The geographical area being investigated is that part of the Susquehanna River Basin which lies in Pennsylvania. The general objectives of this work are grouped into four major categories: (1) geology and hydrology; (2) inventory of natural resources and land use; (3) environmental quality; and (4) digital processing and pattern recognition.

## Digital Data Processing

Various supervised pattern recognition programs are available, including parametric and non-parametric classifiers with either linear or quadratic discriminant functions. Signatures and associated statistics for use in the classifiers are obtained by a program which computes the multivariate statistics for one or more training areas. It frequently happens that training targets are not of sufficient size or area to lend themselves to categorization using the supervised classifiers directly. In such cases, these areas are defined for analysis by an unsupervised classifier which develops its own set of spectral signatures and statistics using a clustering algorithm. Preprocessing before classification is also possible, using programs for normalization, principal components, canonical analysis, etc.

The output of these programs is in the form of a character (or digital) map, with each category of classification represented by a unique symbol. However, the digital character maps are inherently distorted in the length-to-width relationship because of the fixed number of lines and characters per inch of high-speed printer output. The LMAP program, yielding output on the CalComp plotter or the RJE terminals with graphic displays, is used for the production of distortion-free, finished copy, line maps.

In a typical output the percent coverage of the analyzed scene is indicated for each category. Such data is readily converted to acreage or other units as desired for inventory and survey purposes.

The programs mentioned above are all operational and are documented at the user's level. Detailed descriptions of ORSER programs currently available may be found in ORSER/SSEL Technical Report 10-73.<sup>1</sup> The general analysis procedure employed is described in reference (2).

## ORSER DATA PROCESSING SYSTEM

ORSER has developed an extensive operational system for processing and analyzing ERTS-1 and similar multispectral data. The system was developed for use by a wide variety of researchers working in remote sensing. These users represent many disciplines and have a wide range of experience and skill in photointerpretation and computer usage.

### Computer Processing Facilities

Automatic data processing equipment utilized in the ORSER system is primarily located at The Pennsylvania State University Computation Center. The principal computer is the IBM System 370 Model 168, consisting of a main frame and attached devices for input and on-line storage. Users may have access to the computer in any of three ways: central and remote high speed dispatch points operated by the Computation Center; slow speed Remote Job Entry (RJE) terminals using IBM 2741, Tektronix 4010, or similar remote terminals supported by the user or by the Computation Center; and intermediate speed remote batch terminals, such as the IBM

<sup>1</sup> Presented at the 9th International Symposium on Remote Sensing of Environment, April 15-19, 1974, Willow Run Laboratories, Environmental Research Institute of Michigan, Ann Arbor, Michigan.

Hybrid Analysis

After separate analyses of ERTS-1 data by photo-interpretation alone and by computer processing of MSS digital data without the assistance of photointerpretation, it became apparent that each method had shortcomings which might be overcome if the methods were combined. Thus ORSER investigators evolved a method of ERTS MSS data analysis referred to as the "hybrid approach". This method involves intimate interaction of the computer analyst and the photointerpreter, using aircraft photography for comparison with the computer output. A Bausch and Lomb Zoom Transferscope is frequently used for this comparison.

RESULTS

The applications objectives of the ORSER interdisciplinary investigation are grouped into three major categories: (1) geology and hydrology; (2) inventory of natural resources and land use; and (3) environmental quality. Specific results obtained to date are discussed briefly below.

Geology and Hydrology

The detection of lineaments and fracture patterns using ERTS imagery has been the major focus of the geological work. Examination has shown these fractures to be strongly correlated with mineral and groundwater resources.

In 1972, as part of a project to locate major features -- lineaments -- that have engineering applications, a map was published on which only six major lineaments were shown. Less than a year later ORSER prepared a mosaic of ERTS imagery which provided the means by which more than 50 new lineaments could be recognized that have geologic and engineering applications in groundwater exploration, foundation stability studies, mining, etc. Still more exciting is the fact that enlargements of individual ERTS frames reveal the presence of more than a hundred additional lineaments which are 10 to 50 miles long or longer. The scale of ERTS is of vital importance because few if any of these larger scale features are apparent on conventional 1:20,000 scale aerial photographs.

A most important spin-off of the lineament map is its potential application to the location of ore deposits. Using ERTS imagery, close correlation has been observed between lead and zinc ore deposits and the Mount Union-Tyrone lineament in Pennsylvania. Seven areas of mineralization are known to be located along this lineament. ORSER investigators are consequently studying other lineaments for potential indications of mineral and water deposits.

Inventory of Natural Resources and Land Use

Significant results have been attained in mapping land use, agricultural croplands, forest resources and vegetative cover.

Land Use Mapping. Several applications of land use mapping using ERTS data have been made. In most cases, thematic maps generated by computer processing are the final product. The types of land use that are classified and mapped depend upon the geographical location and the types in question.

When mapping relative large areas such as major river basins and watersheds, a relatively few categories may be mapped. In mapping large portions of the Susquehanna River Basin, Schuylkill River Basin, and northwest Pennsylvania, for example, ORSER has used

seven categories of land use as defined by the Soil Conservation Service: water, urban, disturbed (e.g., strip mines, quarries, etc.), pasture, cropland, forest, and other.

In smaller heterogeneous areas, classification of categories has been performed in greater detail. Examples of categories mapped include golf courses, suburban, industrial, shopping centers, etc. Such detailed classification is obviously more difficult and requires considerable ground truth and aircraft data for support and verification.

Agricultural Land Use Mapping. Three agricultural sites selected for computer analysis include two in Pennsylvania and one in Montana. The three sites, taken together, represent a broad range of soils, soil parent materials, climate, modes of agricultural operations, crops, and field sizes.

In the production of agricultural land use maps in two sites in Pennsylvania, five major categories of agricultural land use have been delineated using a cluster analysis on the digital data. Two forest types, water areas, and two agricultural categories (grasses and bare soil with corn stubble), were mapped over an area approximately twelve by eighteen miles. In addition, areas dominated by shales were distinguished from areas dominated by sandstones. Dendritic drainage characteristics of the Allegheny Plateau are clearly shown.

To examine the comparative difficulties of mapping agricultural areas in regions of small, irregular features such as found in Pennsylvania, ORSER investigators classified a portion of Hill Country, Montana, where fields are larger and more uniform.

Survey of Forest Resources and Vegetative Types. The goal of this project was to determine the extent to which it is possible to discriminate between coniferous and non-coniferous forest vegetation using ERTS-1 data, under typical Pennsylvania conditions of intimate mixtures of these two vegetative types. The test area chosen is a part of The Pennsylvania State University Experimental Forest in Stone Valley, Huntingdon County. This area includes the 70 acre University dam and the surrounding forest land, comprising approximately 4500 acres. The forest is managed, well mapped, and continuously inventoried. It is an excellent area for study because of the comprehensive ground truth.

Six categories were mapped, along with an unclassified category. These six are hardwoods, shaded hardwoods, conifers, hemlock-hardwood mixture, fields and water. Shaded hardwoods are those on the northwest aspect of steep slopes where species composition is confounded with the lower incident radiation at the mid-morning time of the ERTS-1 overpass. Identification of species composition is partially a function of aspect but the spectral characteristics and lower incident radiation are not easily separable for ERTS-1 data. The data for this analysis was composed of October 11, 1972 data merged with January 6, 1973 data to yield eight channels of data for analysis. Coniferous vegetation was successfully mapped where it occurred in blocks of five acres or more and comprised the bulk of vegetation in those blocks. This was possible in both summer and winter scenes. Merging of winter and summer seasons made it possible to differentiate hardwoods with coniferous understory from hardwoods on the one hand and conifers on the other. Discrimination between coniferous species on the basis of spectral characteristics alone does not appear very promising; however, where a particular species is associated with another vegetation type, discrimination

is possible. Clear-cut areas of more than 20 acres in size were easily detected by computer processing.

Classification and mapping of major vegetative cover types is also being investigated in an attempt to distinguish between forest and open vegetative areas; and in the open area, to distinguish between herbaceous vegetation, scrub, and small saplings. It is anticipated that these distinctions will lead to mapping of various game covers and estimations of carrying capacities for different game species in a given area. They would also be useful in evaluating land use changes over a period of time by comparison of current maps with succeeding imagery at intervals of perhaps a year or so.

#### Environmental Quality

ORSER investigators have mapped strip mines and detected the effects of acid mine drainage waters from ERTS scenes. Coal refuse in the anthracite mining areas of Pennsylvania has been effectively mapped from ERTS-1 data. Damage to vegetation by air pollution and by insects, in particular the gypsy moth, is being investigated. The groundwork for a study of the environmental effects of atomic power plants has been laid. It is expected that several of these plants will come into operation in the next few years and monitoring of their effects can be begun.

Strip Mines and Acid Mine Drainage. An area along the West Branch of the Susquehanna River which contains old and new stripped areas and numerous examples of acid mine drainage effects was investigated using ERTS data. Classification by cluster analysis provided the best definition of stripped areas and related features. Trenched areas, recent workings, and partially vegetated peripheral zones were identified. Sections of the river classified as strip mines were discovered to represent refuse from nearby stripping that had been dumped along the banks. Areas of dead or dying vegetation caused by acid drainage from strip mines were distinctly classified by cluster analysis. While visual examination of ERTS imagery revealed the location of larger strip mines, only by digital analysis was it possible to distinguish subclassifications and details in these mines.

As in most investigations, aerial photography was of great importance in this study. In particular, U2 color infrared photography is used in the location of targets and very extensively in the verification of results.

Mapping of Anthracite Refuse. Deep mine refuse is a serious problem throughout the anthracite coal region. Deep mines are still active, old refuse is still being reprocessed, and strip mining has increased. Developing an up-to-date inventory and monitoring these activities are feasible by using satellite MSS data.

An area including the Southern and Middle fields of the Anthracite Coal Region of eastern Pennsylvania was studied. Using a supervised classification program, all known coal refuse piles and silt basins were mapped as either refuse or silt, in a preliminary study. Water bodies, cities and towns, and four-laned highways, where they traversed forested areas, were also mapped correctly. The use of a cluster analysis algorithm in a more detailed study resulted in further differentiation of coal refuse banks and silt basins.

A total of 59 distinct spectral signatures representing 13 different mapping categories were defined, including several signatures each for categories of water, refuse, silt, strip mines, industry, towns, roads, brush, vegetation, and swamps.

Insect Damage to Vegetation. The damaging effects of insect infestation on large stands of forests and vegetation are well known. Defoliation by gypsy moth in recent years has exhibited a geometric increase in affected acreage. In 1972, 404,000 acres were affected in Pennsylvania.

Two areas of eastern Pennsylvania have been mapped for gypsy moth defoliation using an ERTS scene of July 8, 1973. The forested areas of interest were mapped as healthy, moderately defoliated, or heavily defoliated. Areas which had been sprayed with Dylor were readily recognized and mapped. Low altitude aircraft photography provided by the U.S. Forest Service was utilized extensively and was very helpful in verifying the results of computer classification. The study of mapping defoliated areas is continuing with emphasis on classification of degrees of defoliation, i.e., percentage defoliation.

Air Pollution Damage to Vegetation. Computer mapping using ERTS data was accomplished for an area surrounding a zinc smelting plant. Various degrees of air pollution damage to vegetation were readily mapped, with effects including: complete absence of vegetation in close proximity to the plant; only scrub and stunted growth of trees and no undergrowth; normal forest but somewhat stunted growth; and healthy vegetation and forest located some distance from the plant.

#### Temporal and Spatial Transference of ERTS Spectral Signatures

Spectral signatures were developed for various vegetative and water targets in a forested area near the East Branch Reservoir in northcentral Pennsylvania, using data from ERTS scene 1028-15295 for August 20, 1972. These signatures were then used to classify targets in the Stone Valley Experimental Forest for scene 1045-15243 of September 6, 1972. These signatures were initially used without success. The data for the two scenes were then recalibrated to match means and variances by channel and modulo scan line number. After this recalibration, excellent classification of the Stone Valley area was achieved using data from the East Branch area. Since the signatures used for classification were developed from ERTS data collected 17 days earlier and in an area approximately 150 miles away from Stone Valley, it appears that transference of spectral signatures in both time and space is feasible.

#### REFERENCES

- (1) "Program Descriptions," ORSER/SSEL Technical Report 10-73, The Pennsylvania State University, University Park, Pa., May 1973.
- (2) McMurry, G. J., Borden, F. Y., Weeden, H. A., Petersen, G. W., "The Penn State ORSER System for Processing and Analyzing ERTS and Other MSS Data", Third Annual Remote Sensing of Earth Resources Conference, University of Tennessee Space Institute, Tullahoma, Tennessee, March 1974.

## ABSTRACT

ORSER-SSEL Technical Report 7-74

A METHOD OF SPECIFYING REMOTELY SENSED UNITS FOR SOIL SAMPLE POINTS

G. A. May, G. W. Petersen, F. Y. Borden, and D. N. Applegate

Three sites were selected from Berks (shale-derived), Duffield (limestone-derived), and Penn (siltstone-derived) soils. A uniform training area was selected within each of these sites. The multispectral scanner data within each training area were analyzed by computer programs.

Scan line and element numbers (RSU) from the multispectral scanner data were assigned sampling locations in predetermined training areas. Each training area was visually located in the proper agricultural field. Three points, identifiable by scan line and element number in the multispectral scanner data, were located in the field. A base line was drawn between two of these points and the third point used as a check in the surveying and calculations. Features that are permanent and easily identifiable by a scan line and element number in the multispectral scanner data were selected for constructing the base line in the field.

Soil samples were collected within each training area. The angle and distance of each sampling point from the constructed base line was determined by a stadia rod. The angles and distances were input to an RSU identification program that outputs the scan line and element number (RSU) of each sampling point. Results from this program indicated that a high percentage of soil sampling locations were within previously designated training areas and the duplication of soil samples within an RSU occurred at only one site.

**Interim Report**

ORSER-SSEL Technical Report 7-74

A METHOD OF SPECIFYING REMOTELY SENSED UNITS FOR SOIL SAMPLE POINTS

G. A. May, G. W. Petersen, F. Y. Borden, and D. N. Applegate

ERTS Investigation 082  
Contract Number NAS 5-23133

INTERDISCIPLINARY APPLICATION AND INTERPRETATION OF ERTS DATA  
WITHIN THE SUSQUEHANNA RIVER BASIN

Resource Inventory, Land Use, and Pollution

Office for Remote Sensing of Earth Resources (ORSER)  
Space Science and Engineering Laboratory (SSEL)  
Room 219 Electrical Engineering West  
The Pennsylvania State University  
University Park, Pennsylvania 16802

**Principal Investigators:**

Dr. George J. McMurtry  
Dr. Gary W. Petersen

**Date:** June 1974

A METHOD OF SPECIFYING REMOTELY SENSED  
UNITS FOR SOIL SAMPLE POINTS<sup>1</sup>

G. A. May, G. W. Petersen, F. Y. Borden, and D. N. Applegate<sup>2</sup>

Office for Remote Sensing of Earth Resources  
Space Science and Engineering Laboratory  
The Pennsylvania State University  
University Park, Pennsylvania

ABSTRACT

Three sites were selected from Berks (shale-derived), Duffield (limestone-derived), and Penn (siltstone-derived) soils. A uniform training area was selected within each of these sites. The multispectral scanner data within each training area were analyzed by computer programs.

Scan line and element numbers (RSU) from the multispectral scanner data were assigned sampling locations in predetermined training areas. Each training area was visually located in the proper agricultural field. Three points, identifiable by scan line and element number in the multispectral scanner data, were located in the field. A base line was drawn between two of these points and the third point used as a check in the surveying and calculations. Features that are permanent and easily identifiable by a scan line and element number in the multispectral scanner data were selected for constructing the base line in the field.

Soil samples were collected within each training area. The angle and distance of each sampling point from the constructed base line was determined by a stadia rod. The angles and distances were input to an RSU identification program that outputs the scan line and element number (RSU) of each sampling point. Results from this program indicated that a high percentage of soil sampling locations were within previously designated training areas and the duplication of soil samples within an RSU occurred at only one site.

1. INTRODUCTION

Multispectral scanner (MSS) data collected from areas on the earth's surface are specified as remotely sensed units (RSU's). The area on the ground represented

<sup>1</sup>Authorized for publication on February 28, 1974, as paper no. 4636 of the Journal Series of the Pa. Agr. Exp. Sta., University Park, Pa.

<sup>2</sup>Graduate Assistant, Associate Professor of Soil Genesis and Morphology, Associate Professor of Forestry, and Research Assistant, The Pennsylvania State University, University Park, Pa., respectively.

by such a unit depends upon the scanner specifications and altitude of the airborne vehicle in which the scanner is mounted. Objects or areas on the earth's surface often need to be directly related to specific RSU's in the MSS data to allow correlation between MSS data and ground data. The objective of this study was to determine if a procedure could be developed to allow soil sampling within an RSU on the ground and to relate this point sample to a specific RSU in the MSS data.

## 2. MATERIALS AND METHODS

### 2.1. AIRCRAFT DATA COLLECTION

Multispectral scanner data were collected over Lancaster, Lebanon, and Berks Counties in southeastern Pennsylvania by an optical-mechanical scanner mounted in an aircraft operated by the Environmental Research Institute of Michigan. These data were collected on May 15, 1969, between 12:40 and 1:25 p.m., EDT, at an altitude of 1,000 meters above the terrain. The scanner system had a resolution of 3 milli-radians and a scanning angle of 40° on each side of the nadir (fig. 1). At an altitude of 1,000 meters above the terrain, each RSU can be represented as a square 9 m on a side at the nadir. The reflected radiation was recorded in 16 spectral bands and digitized. Thirteen of these bands were used in this study, ranging from violet at .40-.44 μm to reflected infrared at 2.00-2.60 μm.

Color photographs along the flightline were taken at the time of the MSS data collections.

### 2.2. SELECTION OF UNIFORM BARE SOIL TRAINING AREAS

With the use of available soil maps (1, 4, 5), three sites were selected for study. The first site consisted primarily of Berks (shale-derived) soils, the second of Duffield (limestone-derived) soils, and the third of Penn (siltstone-derived) soils.

The MSS data collected over each of these three sites were used to produce a digital representation of the relative brightness of the area. The computer-generated maps were produced by computing the norm for each RSU of the data. Geometrically, the norm is the length of the 13-channel vector for each RSU (2). This norm is computed and transformed into the percentage of the maximum possible value for the norm and then translated into a mapping symbol. The brightness maps were produced by the IBM 370 model 165 using a computer program called NMAP (2). Identifiable features on these reflectance maps were then related to common identifiable features on the photograph. In this manner it was possible to delineate bare soil areas on the computer-produced reflectance maps for each of the sites.

From these bare soil areas it was necessary to select training areas that were uniform in their spectral response. To locate these uniform training areas, a uniformity program and a cluster analysis program were run on the MSS data for each site. The uniformity program, or UMAP (3), determines areas of local uniformity and non-uniformity within a block of MSS data. The data used by this program were normalized. This removes the differences in intensity of each RSU and thus only the differences between relative reflectances are important.

UMAP compares the 13-channel vector of a specific RSU with each of the 13-channel vectors of three neighboring RSU's (fig. 2). In Figure 2, RSU-1 is compared to RSU-2, RSU-3, and RSU-4; and RSU-2 is compared to RSU-3. The angular separation between these pairs of vectors is determined. If the angular separation between a pair of these vectors is sufficiently small, the vectors are geometrically close together and therefore the two RSU's are similar. If the angle is large, the two RSU's are dissimilar.

The maximum angular separation of the four pairs of vectors is transformed into a percentage of the maximum possible distance between any two vectors and is then compared to a uniformity limit assigned by the user. If the maximum angular separation is less than the specified uniformity limit, it will be mapped as uniform.

This was one technique used to locate uniform training areas within Berks, Duffield, and Penn soils. An example of the UMAP computer output is shown in Figure 3. It can be seen from this figure that all RSU's within the Duffield training area were mapped as uniform. The uniformity limit used to produce Figure 3 was 3.3.

The computational methods to determine uniformity were performed on the spectral signatures of randomly chosen soil samples taken from RSU's on the ground within this training area. The maximum angular separation between each of the chosen RSU's and their respective neighbors was below the 3.3 limit.

Another technique utilized to obtain uniform or homogeneous training areas was to perform cluster analysis on the MSS data for each site. This was done using the cluster analysis program DCLUS (6). This program samples the MSS data and groups or clusters RSU's with similar spectral responses. Uniform areas designated by the uniformity program were compared to uniform areas designated by the cluster analysis program. Bare soil areas that were designated as uniform by both programs were chosen as training areas.

### 2.3. LOCATING AND SAMPLING OF SOILS

Information from the aerial photography, multispectral data, and visual observations were correlated at each site to enable location of the preselected training area within the proper agricultural field. These were then identified on computer digital output maps. A sampling site deemed to be the corner of the training area was selected for the initial sampling site. The distance between the center of two RSU's was assumed to be 9 m. Thus, the diagonal of an RSU was 12.7. A rope, with these distances indicated, was used to determine approximate sampling locations between adjacent RSU's.

At each of the three sites, soil samples were collected and the angle and distance from a predetermined base line were measured by a stadia rod for each sampling location. Each of these sampling locations was considered to be a different RSU.

### 3. RESULTS AND DISCUSSION

Easily identifiable and permanent features such as single trees, cross roads, or sharp field boundaries were located on the aerial photograph and also on a brightness map of each site. Location of these features on the brightness map identified individual RSU's by scan line and element number. Two RSU's were located in the field and a base line drawn between them. A third RSU was used as a check in surveying and calculations.

Figure 4, a photograph of the limestone area, and Figure 5, a brightness map of the same area, illustrate the three points and the training area in a portion of a corn field. Point one is a single tree in this field and the transit was set up at this point. Point two is the corner of the corn field. The line connecting points one and two is the base line. Each sampling point was recorded by the angle and distance from this base line. Point three is one of three smaller trees located in the corn field and is used as a check. Notice that all three points were easily observable on the photograph and could be assigned a scan line and element number with a high degree of accuracy.

The angle and distance of each sampling location was determined from this base line. These values were then used to determine the RSU from which the soil sample was collected. The following procedures were used to make these determinations utilizing the geometric relationships in Figure 6.

The base line is determined by the points (P<sub>1</sub>, P<sub>2</sub>), and the angle and distance of any sampling point, P<sub>3</sub>, from this base line is θ and d<sub>2</sub>, respectively. Given θ and d<sub>2</sub>,

$$d_3 = d_2 \sin \theta,$$

and

$$d_4 = d_2 \sin (90^\circ - \theta).$$

The coordinates of P<sub>4</sub> are

$$X_4 = \frac{(X_2 - X_1) d_4}{d_1} + X_1$$

and

$$Y_4 = \frac{(Y_2 - Y_1) d_4}{d_1} + Y_1.$$

Then the coordinates of  $P_3$  which correspond to the RSU scan line and element number can be determined from the following two equations:

$$d_3^2 = (X_3 - X_4)^2 + (Y_3 - Y_4)^2$$

and

$$(Y_3 - Y_4) = \left[-\frac{1}{m} (X_3 - X_4)\right],$$

where  $m$  is the slope of the base line.

The scale of the photography for this flightline was used to determine the length and width of an RSU on the ground. The width of each RSU changes when progressing along a scan line due to the rotation of the scanner. As the scanner rotates toward and away from the nadir, the angle at which the scanner looks at the earth's surface and collects information changes. The altitude of the airplane, the nadir element, and the scanning angle were used to correct the length of each RSU along a scan line.

For each sample, the scan line and element number from the multispectral scanner data were output by the RSU identification program.

The success of this procedure is shown in Figure 3. In this training area of Duffield soils, 73 samples were collected. As shown in this figure, all but 6 of the soil sampling locations fell within the designated training area. No attempt was made to determine the position of a soil sample within an RSU.

At both the Berks and Penn sites, a base line was drawn and a third point determined to allow for utilization of this same sampling procedure. All 64 soil samples collected at the Berks site were within the training area and 72 soil samples out of the 81 taken from the Penn site were within the Penn training area.

#### 4. LITERATURE CITED

1. Ackerman, F. S. 1970. Soil Survey Berks County, Pennsylvania. USDA Soil Conservation Service, Washington, D. C. 121 pp.
2. Borden, F. Y. 1971. NMAP Program Description. Tech. Report 2-71. Office for Remote Sensing of Earth Resources, The Pennsylvania State University, University Park, Pa. 7 pp.
3. Borden, F. Y. 1971. UMAP Program Description. Tech. Report 4-71. Office for Remote Sensing of Earth Resources, The Pennsylvania State University, University Park, Pa. 14 pp.
4. Carey, J. B. 1956. Soil Survey Lancaster County, Pennsylvania. USDA Soil Conservation Service, Washington, D. C. 132 pp.
5. Mathews, H. L. 1972. Application of Multispectral Remote Sensing and Spectral Reflectance Patterns to Soil Survey Research. Ph.D. dissertation, The Pennsylvania State University, University Park, Pa. 110 pp.
6. Turner, B. J. 1972. Cluster analysis of multispectral scanner remote sensor data. Remote Sensing of Earth Resources, Vol. I. Edited by F. Shahrokhi. Tennessee Space Institute, University of Tennessee, Tullahoma, Tenn. pp. 538-549.

#### 5. ACKNOWLEDGMENT

The authors wish to express their appreciation to Dr. Harold T. Rib, Federal Highway Administration, Washington, D. C., for making multispectral remote sensing data available for this study.

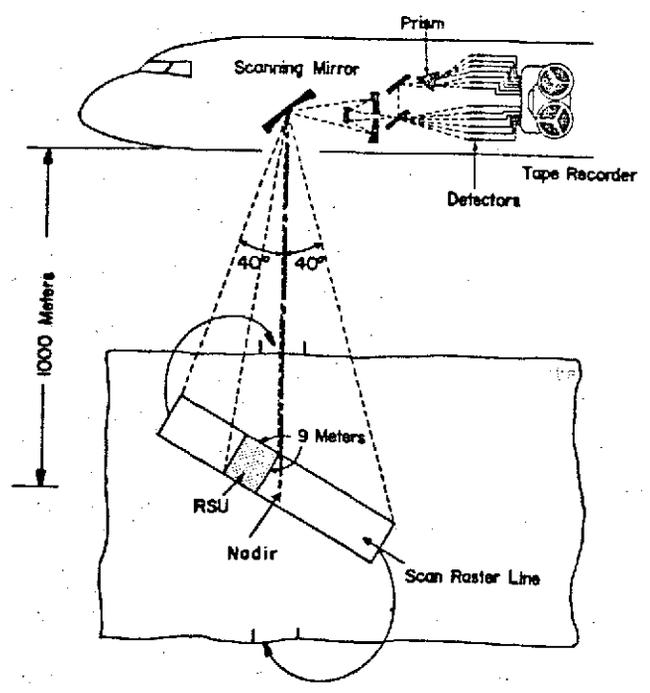


Fig. 1: Schematic representation of the 13-channel scanner collecting multispectral data reflected from the earth's surface.

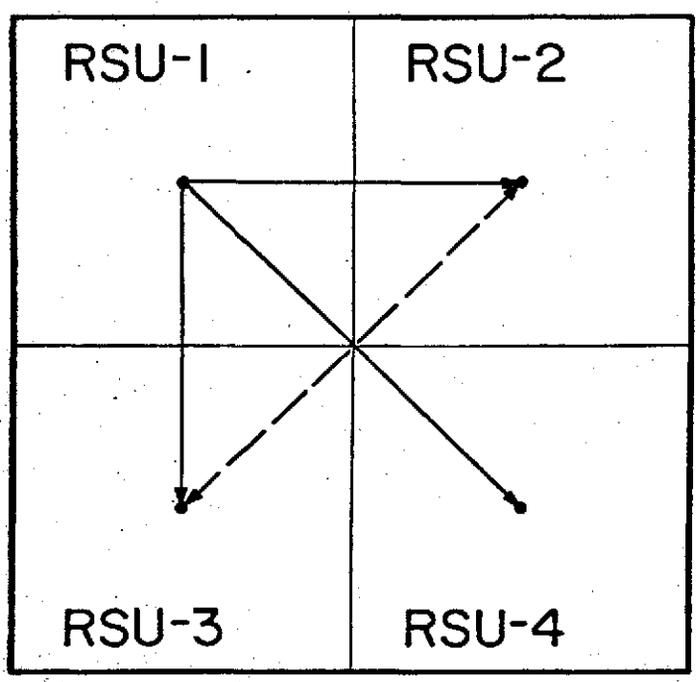


Fig. 2: Diagrammatic illustration of how UMAP compares the spectral responses of adjacent RSUs.



Fig. 3: UMAP output indicating the training area of uniform spectral response within the limestone study site.

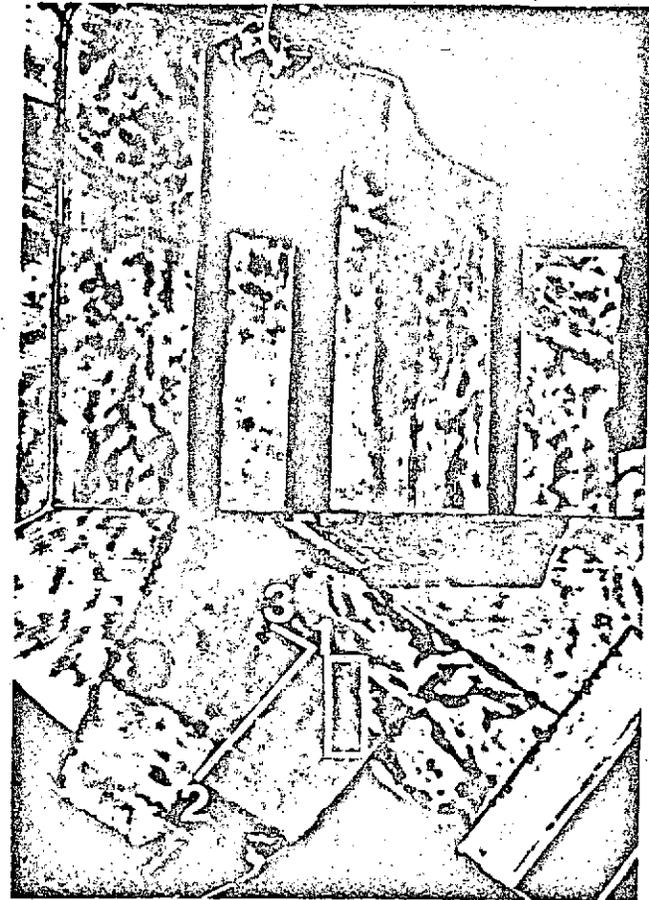


Fig. 4: Areal view of the Duffield soil.

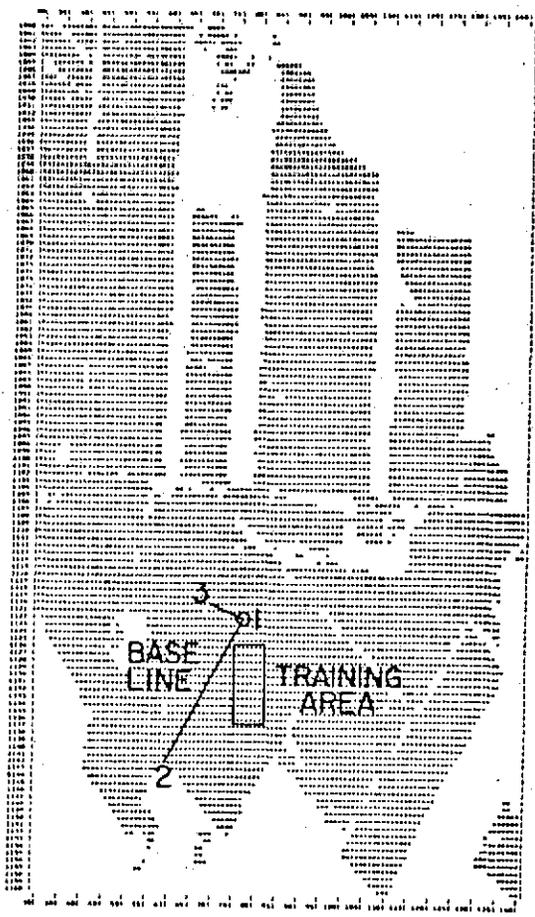


Fig. 5: Delineation on the brightness map of individual remote sensing units within areas of Duffield soils.

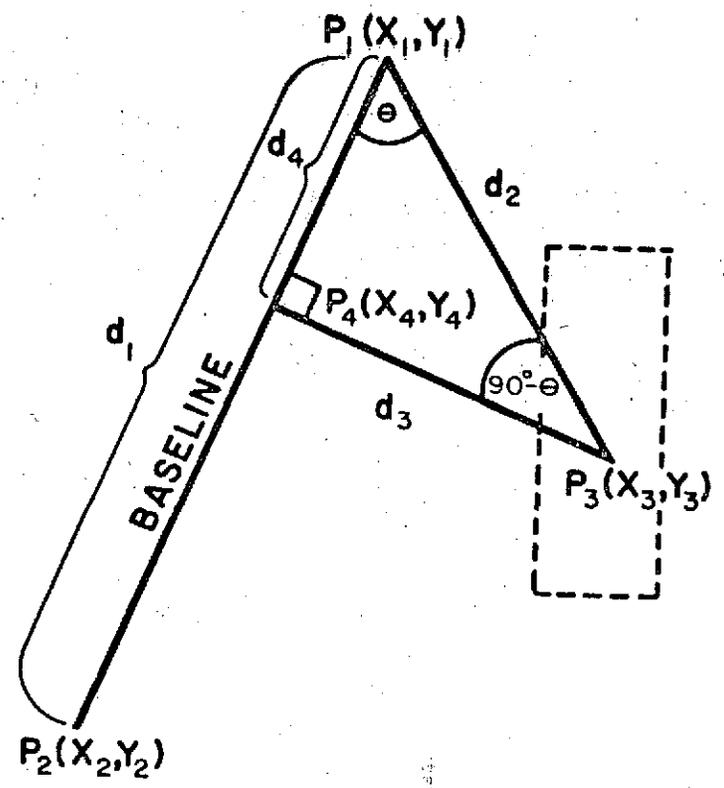


Fig. 6: Diagram of the geometry used to determine the RSU from which a soil sample was collected.

ABSTRACT

ORSER-SSEL Technical Report 5-74  
COMPUTER ANALYSIS AND MAPPING OF GYPSY MOTH DEFOLIATION LEVELS IN NORTHEASTERN  
PENNSYLVANIA USING ERTS-1 DATA  
D. L. Williams and B. J. Turner

Gypsy moth caterpillars defoliated 860,000 acres of Pennsylvania's woodlands in 1973. Heavy defoliation for two successive years can result in significant tree mortality. Present methods of mapping defoliation extent can be improved by supplementation with ERTS-type data. Computer analysis of an ERTS-1 scene taken at the time of defoliation indicates that at least two levels (heavy and moderate) of defoliation can be detected and mapped. Current limitations in using ERTS-type data, due to timing and cloud cover, could be overcome if future satellite systems result in more frequent coverage.

**Interim Report**

**ORSER-SSEL Technical Report 5-74**

**COMPUTER ANALYSIS AND MAPPING OF GYPSY MOTH DEFOLIATION LEVELS  
IN NORTHEASTERN PENNSYLVANIA USING ERTS-1 DATA**

**D. L. Williams and B. J. Turner**

**ERTS Investigation 082  
Contract Number NAS 5-23133**

**INTERDISCIPLINARY APPLICATION AND INTERPRETATION OF ERTS DATA  
WITHIN THE SUSQUEHANNA RIVER BASIN**

**Resource Inventory, Land Use, and Pollution**

**Office for Remote Sensing of Earth Resources (ORSER)  
Space Science and Engineering Laboratory (SSEL)  
Room 219 Electrical Engineering West  
The Pennsylvania State University  
University Park, Pennsylvania 16802**

**Principal Investigators:**

**Dr. George J. McMurtry  
Dr. Gary W. Petersen**

**Date:**

COMPUTER ANALYSIS AND MAPPING OF GYPSY MOTH  
DEFOLIATION LEVELS IN NORTHEASTERN  
PENNSYLVANIA USING ERTS-1 DATA

D. L. Williams and B. J. Turner\*

Office for Remote Sensing of Earth Resources  
Space Science and Engineering Laboratory  
The Pennsylvania State University  
University Park, Pennsylvania

Abstract

Gypsy moth caterpillars defoliated 860,000 acres of Pennsylvania's woodlands in 1973. Heavy defoliation for two successive years can result in significant tree mortality. Present methods of mapping defoliation extent can be improved by supplementation with ERTS-type data. Computer analysis of an ERTS-1 scene taken at the time of defoliation indicates that at least two levels (heavy and moderate) of defoliation can be detected and mapped. Current limitations in using ERTS-type data, due to timing and cloud cover, could be overcome if future satellite systems result in more frequent coverage.

Financial support for this work was provided by McIntyre-Stennis funds through research project 2025 of the Pennsylvania Agricultural Experiment Station and by the Office for the Remote Sensing of Earth Resources, The Pennsylvania State University.

The authors wish to record their gratitude to Mr. Doman Roberts of Applied Technology Applications Corporation for supplying ground truth from the Pa. Bureau of Forestry and the U. S. Forest Service and valuable suggestions. We wish also to thank ORSER personnel for advice and assistance.

\*Graduate Assistant in Forest Resources and Assistant Professor of Forest Management, respectively, School of Forest Resources.

Introduction

The woodlands of Pennsylvania fall under the broad classification of "eastern hardwoods," with oaks and hickories predominating in many areas. The loss of part, or all, of these woodlands due to the rapid spread and increase in populations of defoliating insects has become a topic of major concern in recent years. The death of trees affects stream flow, increases fire and erosion hazards, reduces land and recreational values, and destroys wildlife habitats. The need to salvage commercial timber stands, where heavy mortality has occurred, has disrupted plans for maintaining a continuous flow of timber products from these forests. It has resulted in substantial losses from the necessity to harvest immature stands. Defoliated acreage in Pennsylvania increased by 113 percent over 1972 for a total of about 860,000 acres in 1973.<sup>1</sup>

Results of studies by Nichols<sup>2</sup> on oak mortality in Pennsylvania show the effects of defoliation. He found that two consecutive years of 60 to 100 percent spring defoliation caused mortality. Only rarely did mortality result from one heavy stripping preceded by several years of moderate defoliation. Continuous moderate defoliation did not kill trees, but it started their decline. A year of moderate defoliation reduced radial growth by 20 to 30 percent. But one year of heavy spring defoliation resulted in a 40 to 70 percent growth reduction. Thus, the mortality of trees in future years, as a result of defoliation now, depends largely on the percentage of leaves removed during infestation.

The gypsy moth (Porthetria dispar) is probably the most destructive defoliation insect threatening or currently attacking forests and shade trees throughout most of the Eastern United States. A single defoliation can kill white pine, spruce, and hemlock, while two defoliations are sufficient to kill most hardwoods. Most of the damage is done by late June, at which time the caterpillars enter a resting or pupal stage. Most hardwood trees produce a new crop of leaves about three weeks after they have been defoliated. This process

GYPSY MOTH DEFOLIATION

greatly reduces the energy reserves of the trees and, if heavy defoliation occurs again in the next year, significant mortality may result.

Much research has been done in an attempt to eradicate the gypsy moth, but the ban of DDT in the 1960's rendered the task of complete eradication an economic impossibility. At best, the spread can be held in check, and the population decreased, in the near future. Therefore, some means of monitoring this forest insect, and its resulting damage, is needed. Present methods of monitoring involve the extensive use of aircraft with either airborne mapping, or photography with subsequent photo-interpretation and mapping. Airborne mapping is cheaper, but less accurate due to orientation problems. Furthermore, it would be desirable to do this monitoring on a frequent (perhaps semi-annual) basis and, at the same time, hold down costs. Preliminary results of computer runs, using programs that process data received from the ERTS-1 satellite, seem to suggest that these data offer this desirable combination.

The woodlands of Pennsylvania offer the unique situation of widespread similarity of tree species at differing degrees of infestation. The northeastern part of Pennsylvania has been attacked for several years and offers varying degrees of defoliation. Traveling westward, one passes through a spectrum of variations in time (years) of occupancy and degrees of defoliation. In the southwestern part of the state, one finds woodlands that have not yet felt the stress of gypsy moth attack. This may change in the next few years unless intensive research results in major breakthroughs.

Procedure

It was decided to concentrate initially on the forests of the northeastern portion of Pennsylvania. There were two major factors which brought about this decision: 1) vast acreages of varying degrees of gypsy moth defoliation and 2) availability of ground truth information in the form of color IR aerial photography at a scale of 1:6000 (taken July 18, 1973) and data gathered in the field by the

D. L. WILLIAMS AND B. J. TURNER

Pennsylvania Bureau of Forestry (B.O.F.) and the U. S. Forest Service (U.S.F.S.).

The acreage and severity of gypsy moth defoliation reaches a peak from mid-June through the first few days of July. If the defoliation has only been light to moderate in past years, the deciduous trees will respond with a new canopy of leaves by mid- to late-July. For this reason, the timing of aerial coverage and ground support data is extremely important.

It was very fortunate that ERTS-1 coverage of northeastern Pennsylvania on July 8, 1973, fell on a cloud-free day. ERTS-1 scenes 1350-15183 and 1350-15190 were chosen for computer analysis.

The first step in processing was to develop work tapes containing the digital data from those portions of the scenes where gypsy moth damage was known to have occurred. There were two aids used in deciding which scene portions would make good study areas. One aid was the Forest Pest Report,<sup>1</sup> supplied by the B.O.F. It gave a county-by-county breakdown of total defoliation for both 1973 and 1972. This information was derived from aerial reconnaissance. The other aid was the ERTS-1 images themselves. Image reproductions of MSS band 7 showed defoliation as dark gray. Healthy vegetation appeared as lighter shades of gray, while black represented water bodies. (Fig. 1) The SUBSET program developed by the Office for Remote Sensing of Earth Resources (ORSER) at The Pennsylvania State University was used to transfer the desired data from the original NASA ERTS tapes.<sup>3, 4</sup> The SUBSET program allows the user to select only data he needs from a tape and put them on a work tape. It also converts the tape format to a standard ORSER format.

Once the desired portions were subsetted, proper location and orientation on the tape, pertaining to the areas of interest on the ground, had to be achieved. The NMAP program was used for this purpose. This program uses all channels of the work tape to map element brightness based on these channels. Output from the program consisted of a brightness map. Certain symbols were assigned to

GYPSY MOTH DEFOLIATION

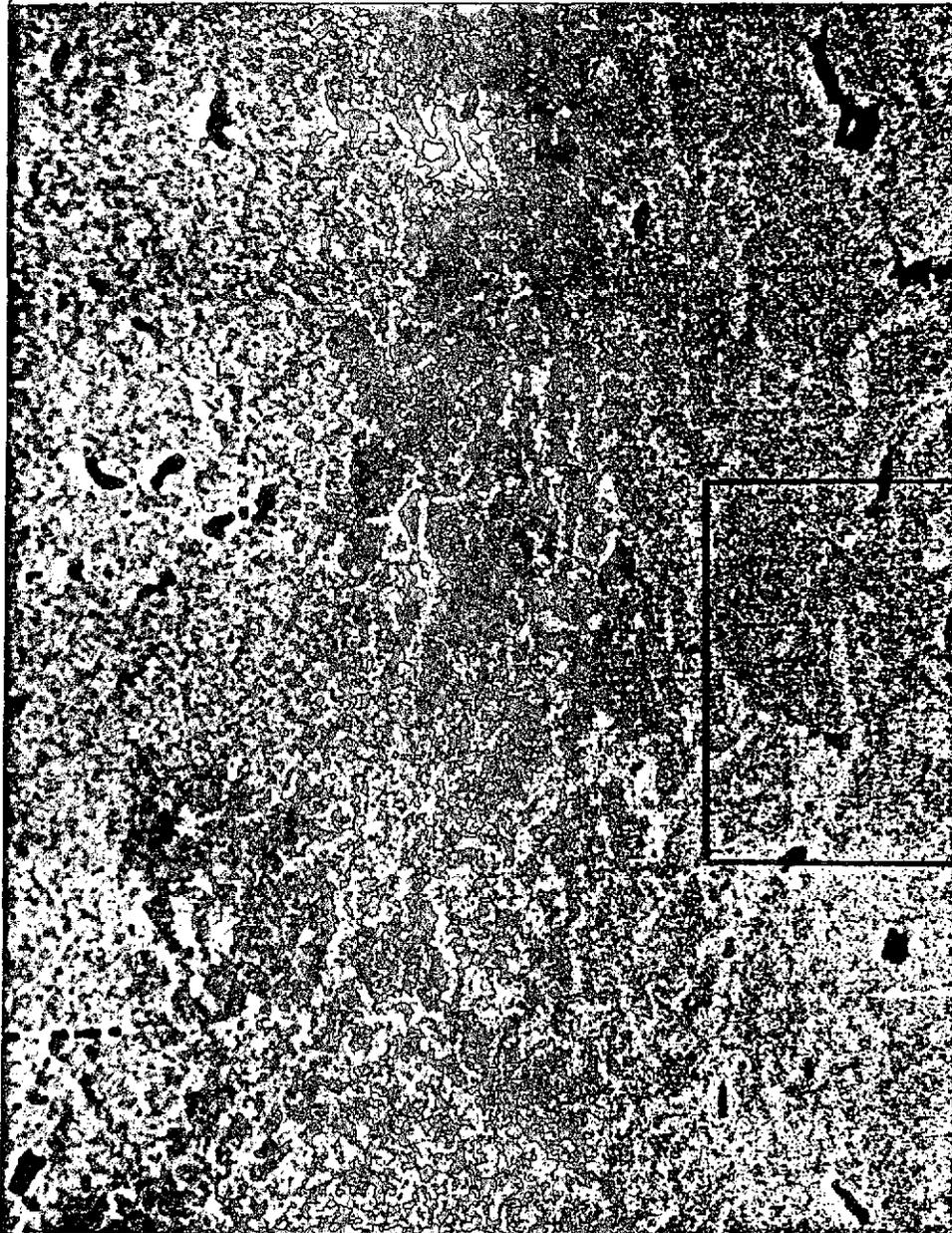


Fig. 1 A blowup of a portion of MSS band 7 from ERTS-1 scene 1350-15183, July 8, 1973. Water appears black, defoliation appears as darker shades of gray, and healthy vegetation appears as lighter shades of gray. The area inside the box includes Dylox spray plots represented in other illustrations.

various percentage intervals of brightness and the patterns of symbols were helpful in interpretation. Water bodies were extremely helpful for orientation purposes due to their low brightness, uniformity, and distinctive spectral signature. Fortunately, northeastern Pennsylvania has numerous lakes and swamps large enough in size to show up on ERTS-1 images. A complete supply of U.S.G.S. 7.5 minute topographic maps of the study areas proved to be invaluable in substantiating accurate identification of water bodies.

After the proper areas had been subset and orientation within these areas achieved, the combined tasks of classification and character mapping were undertaken. Two different paths were available at this point--the supervised or unsupervised classification procedures. The unsupervised classification, or cluster analysis, procedures were chosen, with the intent of using the supervised procedures at a later date to check for uniformity as well as variability in the spectral signatures obtained from cluster analyses.

The majority of spectral signatures used throughout the project were developed using the DCLUS program. In using this program, one specifies the corner coordinates of the target area(s) to be processed, the number of sample points to be initially chosen, and the initial critical clustering distance. Output consists of a list of spectral signatures corresponding to the mean vector length and standard deviation. A character is assigned to each cluster spectral signature and a character map of the specified target area(s) is output.

This method of developing spectral signatures proved to be quite adequate. With the help of the color IR aerial photography, and other supporting ground truth information, classification of these spectral signatures into categories was relatively easy. It was readily apparent that defoliated trees had lower responses in MSS channels 6 and 7, corresponding to the severity of defoliation. Initially, ten categories were delineated and used as input in the more sophisticated classification program DCLASS.

GYPSE MOTH DEFOLIATION

DCLASS allows the user to input spectral signatures obtained from DCLUS or certain other ORSER programs. The distance of separation between an element vector and each of the category vectors is used to assign the element to the category for which it has the smallest distance of separation. If there is no category to which the element can be assigned, it is classified as "other."

The standard deviations, in distance units, from DCLUS were used as a guide for setting the critical distance in DCLASS. An initial critical classification distance of 8.0 was used. Mapping symbols were assigned to each category and character maps were generated. Further refinements in category specification were made possible as familiarity with the target areas increased. Eventually, a total of 31 categories were separated through continued use of DCLUS and DCLASS.

It should be noted that in addition to character maps, DCLASS outputs frequency distributions for all categories and distances of separation between all pairs of spectral signatures. Of the 31 signatures, there were 9 for lakes and swamps, 10 for healthy vegetation, 5 for defoliated forests, and 7 for miscellaneous categories. The frequency distribution table indicated that the signatures accurately identified 98 percent of the study area, leaving only 2 percent unclassified.

The table of distances of separation revealed that the distance among spectral signatures within the set of five categories for heavy defoliation and within the set of ten categories for healthy vegetation were much less than the distances between category sets. Within the five categories for heavy defoliation, the average distance of separation was 2.89, with a range of values from a low of 0.90 to a high of 4.20. Within the ten categories for healthy vegetation, the average distance of separation was 4.10, with a range of values from 0.80 to 9.00. The higher average and greater range of values for healthy vegetation were not surprising, as one would expect greater heterogeneity in spectral signatures due to the vast number of differing deciduous and coniferous species found in the mixed forests of Pennsylvania. On the other

hand, one would expect better overall similarity among spectral signatures for denuded woodlands, as the forest floor and bare branches would reflect roughly the same amount of radiation regardless of tree species.

With this in mind, it was decided to average the 5 spectral signatures for heavy defoliation to get one overall signature for this category. The same was done with the 10 spectral signatures for healthy vegetation. Thus, the total number of categories was trimmed from 31 to 11. This was done to cut the costs of generating character maps, without having to sacrifice accuracy in the process.

DCLASS, with 11 categories, generated a character map which closely resembled previous maps when all 31 categories were specified. The greatest change occurred in the percentage that was unclassified. It increased from 2 percent to 7 percent. However, the distance of separation between the category for heavy defoliation and that for healthy vegetation was 12.7. When the initial critical classification distance of 8.0 was slightly increased to 10.0, the percentage unclassified reduced from 7 to 3 percent.

As a first approximation to obtaining a category for moderate defoliation, a spectral signature midway between that for heavy defoliation and healthy vegetation was obtained by averaging their spectral signatures, thus differing from each by 6.3 units. This signature was refined after comparing character maps with estimates of defoliation on ground photos.

### Results

Table 1 gives a list of the 12 categories delineated within the study area, including a summary of the symbols assigned to each category and their spectral signatures. It should be noted that channel numbers 1, 2, 3, and 4 correspond with MSS bands 4, 5, 6, and 7. The signatures of particular interest are those for heavy defoliation (10), healthy forest (11), and moderate defoliation (12). Although these three categories show similar responses in channels 1 and 2, significant

GYPSY MOTH DEFOLIATION

Table 1 Mean spectral signatures and mapping symbols derived from DCLASS output

Category name	Number	Symbol	Limit
Water	1	=	10.0
Wetlands	2	-	10.0
Swamp1H20	3	=	10.0
Lake-edge	4	-	10.0
Siltwater	5	=	10.0
Swamp2H20	6	=	10.0
Swamp3H20	7	=	10.0
Swamp4H20	8	=	10.0
Swamp5H20	9	=	10.0
Heavy Defol.	10	@	10.0
Healthy Forest	11	I	10.0
Moderate Defol.	12	+	10.0

Unnormalized Category Specifications

		Channels			
		1	2	3	4
1	=	27.83	19.00	17.25	5.21
2	-	33.00	24.00	31.00	14.00
3	=	32.51	22.77	24.67	9.90
4	-	32.50	22.75	32.75	15.00
5	=	35.60	24.93	21.57	6.74
6	=	33.32	24.26	26.71	11.31
7	=	31.63	21.46	22.33	8.61
8	=	33.02	22.98	21.02	5.23
9	=	29.03	19.22	23.67	9.25
10	@	34.08	25.47	41.30	21.08
11	I	34.79	24.61	51.47	28.52
12	+	34.44	25.04	46.38	24.80

differences in response can be seen in channels 3 and 4, corresponding to the degree of defoliation. "Healthy forests" is considered to have 0-30 percent defoliation, "moderate defoliation" 30-55 percent, and "heavy defoliation" 60-100 percent defoliation.

One training area encompassed three experimental plots aerially sprayed with Dylox, and was of particular interest throughout the investigation.

These plots were easily identifiable on ERTS-1 image 1350-15183 and on character maps due to their peculiar shapes and orientations. Figure 2 is a sketch of these plots derived from B.O.F. maps and color IR aerial photography. These plots not only offered a good reference point, but they were most helpful in the development of spectral signatures as all defoliation levels were represented. Those areas receiving maximum spray application were relatively unharmed and healthy, while the surrounding unsprayed forest land was generally heavily defoliated. Certain areas within and around these plots, where the application of the spray was not as heavy, showed moderate degrees of defoliation.

A character map produced by the classification program DCLASS from the spectral signatures given above shows the three sprayed plots in the lower half of the figure. (Fig. 3) The darkened areas represented by the "@" symbols indicate heavy defoliation, while the "+" symbols represent moderate defoliation, and the "I" symbols represent

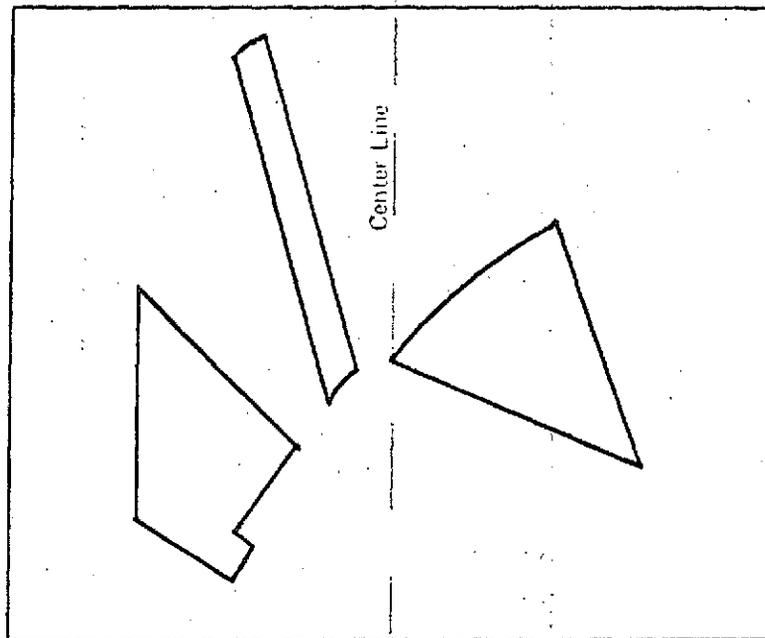


Fig. 2 Sketch of Dylox spray plots to show shape and orientation.

### GYPSY MOTH DEFOLIATION

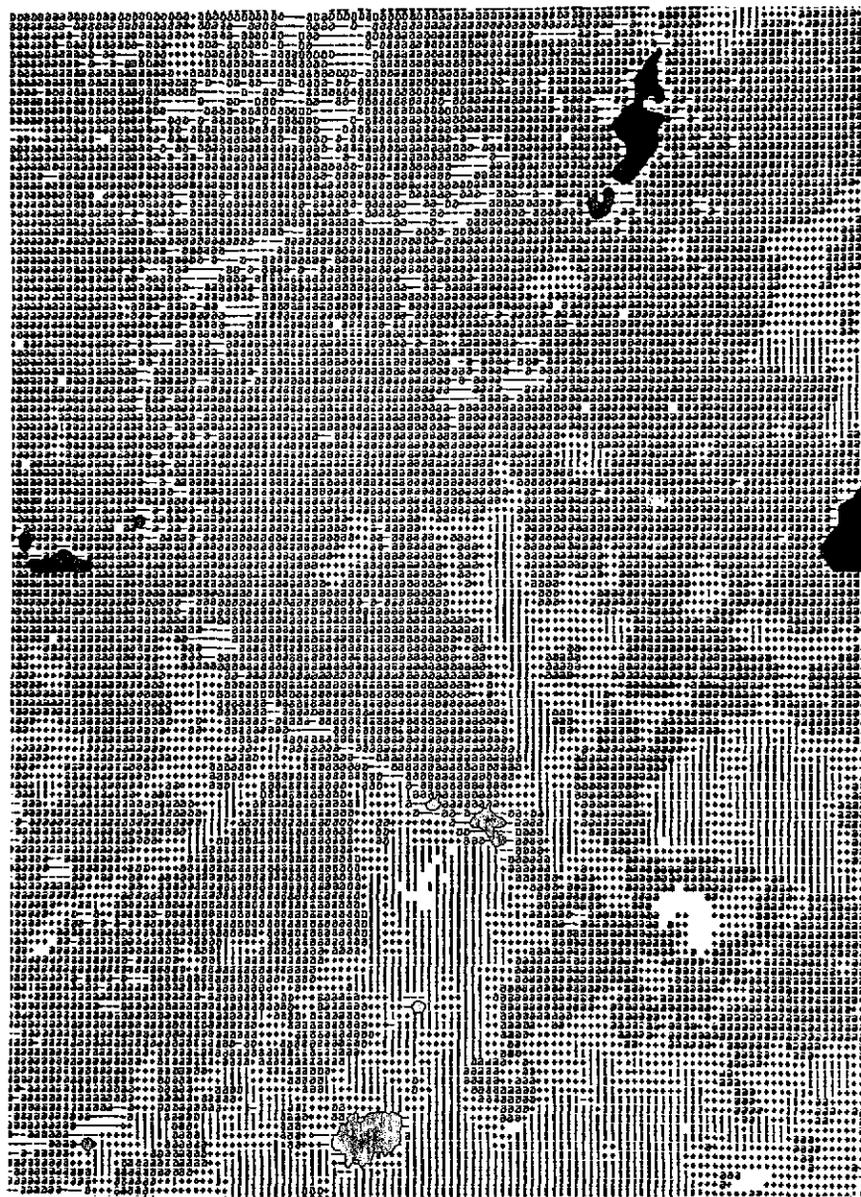


Fig. 3 Character map derived from DCLASS using spectral signatures given in Table 1. Dylox spray plots appear in lower half.

Legend Heavy Defoliation @  
 Moderate Defoliation +  
 Healthy Forest Land I  
 Swamps -  
 Water "dark"

healthy forest land. The black areas are water bodies and the "-" symbols represent swamps. Character maps such as these are useful as work maps for the user in his analysis of MSS data. They are inherently distorted in the length to width relation because of the fixed number of lines and characters per inch on line printers. The LMAP program developed at ORSER was designed to overcome this problem by allowing the user to prepare a line map to any desired scale from any character map.

Figure 4 represents the same area drawn by a CalComp plotter from output generated by the LMAP program. It results in a map more pleasing to the eye and easier to interpret. Here the vertical lines represent healthy forest land, the double horizontal bars represent moderate defoliation, the blank area represents heavy defoliation, the "+" symbols represent water, the "X" symbols represent unclassified areas, and the slashes represent swamps.

#### Discussion

The results indicate that ERTS-1 data can be used to discriminate between defoliated and healthy vegetation in northeastern Pennsylvania, and that digital processing methods can be used to accurately map the extent and degree of defoliation. The question that now arises is how well this method compares with existing methods for monitoring defoliation extent.

As was indicated earlier, a critical factor in any defoliation detection and mapping scheme is timing. All data must be collected after the peak of feeding by the larvae and before the refoliating trees become indistinguishable from undefoliated trees. At most, this period is two or three weeks.

Present methods of mapping rely mainly on the extensive use of light aircraft with an observer experienced in the local geography (usually a forester) directly mapping the extent of defoliation onto topographic maps. These data have been supplemented with aerial photography along sample strips and ground observation data on sample plots.

### GYPSY MOTH DEFOLIATION

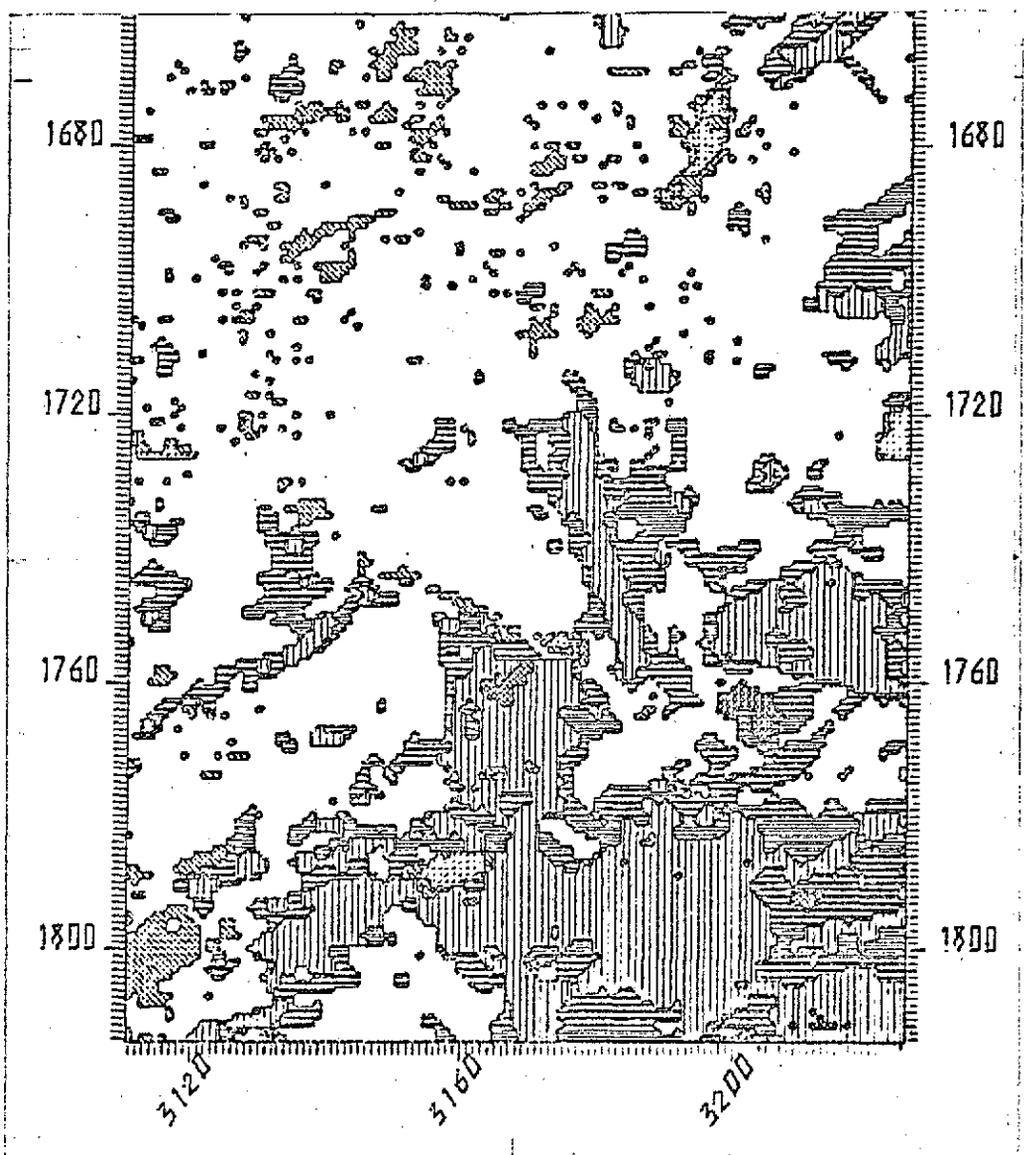
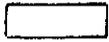
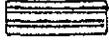
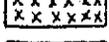
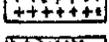
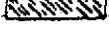


Fig. 4 Line map of same area as Fig. 3, drawn to specified scale using the LMAP program and the CalComp plotter.

Legend	Heavy Defoliation	
	Moderate Defoliation	
	Healthy Vegetation	
	Unclassified	
	Water	
	Swamps	

The present method has several limitations which can be overcome by using satellite-derived data. Because of the necessity of covering vast areas in a short time span, the same observer cannot be used throughout, so differences between observers' interpretations of defoliation level can be substantial. This may be compounded by the change in the amount of refoliation which can occur over the time necessary to map the total area. Maps produced from satellite data are not subject to operator bias and substantial areas can be monitored in a single pass; e.g., a third of Pennsylvania was covered on one day.

In the present method, the precision of mapping depends on the extent of the observer's knowledge of local geography and his ability to relate this to what he sees from the air. Fatigue and air discomfort may also affect the observer's judgments. He is rarely able to achieve more than generalized locations of the boundaries, which may be sufficient for obtaining total acreage figures but may be inadequate for planning salvage or spraying operations. ERTS-1 data, on the other hand, are not affected by these types of human failings and boundaries can be mapped much more precisely. In addition, area estimates are obtained as a by-product of mapping.

Despite these several advantages, there are serious problems associated with relying solely on ERTS-type satellites. These limitations are associated with the frequency of passes. The satellite covers a given area only every 18 days, which is roughly the length of the critical feeding stage in the life cycle of the gypsy moth. There is therefore the possibility that a pass will not occur when defoliation is at its optimum detection stage. In addition, chances are high in this part of the country that cloud cover will ruin coverage on a given overpass. While it would not be sensible to place sole reliance on ERTS satellites for defoliation monitoring, these present limitations would be obviated if future satellites had shorter periods, giving more frequent coverage.

While cost comparisons are hazardous at the present experimental stages of this new technology,

GYPSY MOTH DEFOLIATION

it appears that computer mapping of defoliated regions could be done for roughly 25 cents/square mile covered. It would thus seem likely that on an operational basis, mapping of defoliation levels from satellite-derived data could compete favorably cost-wise with existing methods.

There are a few problems yet to be solved. One of these is to discriminate between defoliated trees and trees killed as a result of previous defoliation. This can be done by comparing data obtained at the time of defoliation with data from a pass later in the summer. Another problem which may arise in the future is discriminating between defoliation caused by gypsy moth and that caused by other defoliating insects. At present, this is not a difficulty since the areas of attack by spring defoliators are separated geographically.

References

<sup>1</sup>Nichols, J. O., Forest Pest Report--July 1973, Pa. Dept. of Environmental Resources, Bureau of Forestry.

<sup>2</sup>Nichols, J. O., The Gypsy Moth in Pennsylvania, Pa. Dept. of Agr., Misc. Bulletin No. 4404, 1961.

<sup>3</sup>Borden, F. Y., "A Digital Processing and Analysis System for Multispectral Scanner and Similar Data," Remote Sensing of Earth Resources, Vol. 1, 1972, pp. 481-507.

<sup>4</sup>Borden, F. Y., Applegate, D. N., Turner, B. J., Lachowski, H. M., and Hoosty, J. R., ERTS and Aircraft Multispectral Scanner Digital Data Users Manual, ORSER-SSEL Tech. Report 11-73, May 1973, The Pennsylvania State Univ., University Park, Pa.

ORSER-SSEL Technical Report 5-74  
The Pennsylvania State University  
March 1974

ORSER-SSEL Technical Report 4-74  
SATELLITE AND AIRCRAFT MULTISPECTRAL SCANNER DIGITAL DATA USER'S MANUAL  
F. Y. Borden, D. N. Applegate, B. J. Turner, H. M. Lachowski, and  
J. R. Hoosty

Programs and their control cards in current use by the Office for Remote Sensing of Earth Resources (ORSER) at Penn State are described. These include programs of the following types: 1) a program to obtain information about the contents of a tape from the tape internal label, 2) a subsetting program, 3) programs for creating brightness and uniformity maps, 4) a program to obtain basic statistics for user-defined small blocks of data, 5) supervised classifiers which classify data from a set of user-specified spectral signatures according to the angle or distance of separation, 6) unsupervised classifiers which develop their own set of spectral signatures using a clustering algorithm, 7) a program using the method of canonical analysis to derive an orthogonal transformation which maximizes category separability on as few axes as possible, 8) a classification and mapping program based on the ratio of two selected channels of data, 9) a program which merges data from two different passes of ERTS over the same area, 10) a program which compares, element by element, two digital classification maps of the same ground area, 11) a series of classifiers employing the linear discriminant function, and 12) a series of classifiers employing the quadratic discriminant function. The digital aircraft multispectral scanner data tape format used by ORSER is also described. Instructions and examples of computer processing of both ERTS and aircraft multispectral scanner data, using cards or the Remote Job Entry terminal, are included.

(This report is an updated version of ORSER-SSEL Technical Report 11-73.)

ABSTRACT

ORSER SSEL Technical Report 3-74  
GRAYSCALE PRINTING AND PHOTOGRAPHIC REDUCTION OF LARGE AREA  
COMPUTER GENERATED MAPS  
A. David Wilson and Richard E. Ackely

Computer generated maps from ERTS-1 digital data have a nominal scale of 1:20,000 which is suitable for analysis or planning purposes. However, this scale is much too large for presentation or publication. With the method described in this paper, photographic reduction of a 7 by 7 foot high-speed printer generated computer map has proven practical. With careful selection of symbols and control of the ink intensity of the printer, individual mapping symbols (each representing a pixel size of 1.3 acres) can be identified on an 8 by 10 inch glossy photographic print of a 9000 square mile area. Thus the method retains the resolution of the original ERTS data while representing areas of regional importance in a size convenient for publication. The techniques used are available in any well equipped photographic studio.

Interim Report

ORSER-SSEL Technical Report 3-74

GRAY SCALE PRINTING AND PHOTOGRAPHIC REDUCTION OF LARGE AREA  
COMPUTER GENERATED MAPS

A. David Wilson and Richard E. Ackley

ERTS Investigation 082  
Contract Number NAS 5-23133

INTERDISCIPLINARY APPLICATION AND INTERPRETATION OF ERTS DATA  
WITHIN THE SUSQUEHANNA RIVER BASIN

Resource Inventory, Land Use, and Pollution

Office for Remote Sensing of Earth Resources (ORSER)  
Space Science and Engineering Laboratory (SSEL)  
Room 219 Electrical Engineering West  
The Pennsylvania State University  
University Park, Pennsylvania 16802

Principal Investigators:

Dr. George J. McMurtry  
Dr. Gary W. Petersen

Date: February 1974

GRAYSCALE PRINTING AND PHOTOGRAPHIC REDUCTION  
OF LARGE AREA COMPUTER GENERATED MAPS

A. David Wilson and Richard E. Ackley\*

Computer generated maps from ERTS-1 digital data have a nominal scale of 1:20,000 which is suitable for analysis or planning purposes. However, if the area selected for mapping exceeds several square miles, this scale is much too large for presentation or publication. Therefore, some form of reduction becomes necessary. XEROX or standard photographic reduction onto multilith mats has been useful for reducing a single high-speed printer page to publication format. With the method described here, photographic reduction of a 7 by 7 foot high-speed printer generated computer map has proven practical. On an 8 by 10 inch photographic print of a 9000 square mile area, the individual pixel (1.3 acres) mapping symbol was identifiable. Although individual pixels cannot be identified on Multilith copies of such a print, the process reduces large maps to a convenient publication format.

Map Preparation

The selection of symbols for use in mapping an area is very important. The symbols should be selected for their "gray scale" value. Previous trials at gray scale mapping have indicated that only five gray tones are differentiable, utilizing standard printing techniques. However, to preserve identity within similar categories, they can be mapped with different symbols of similar intensity. In the figure, an area from Montana, the categories of RANGE, PRAIRIE, and CUTBANK were mapped as R, P, AND C, respectively. These three letters have similar intensities, and appear as the same gray tone on the photographic reduction. A series of darkest to lightest symbols might be arranged in the following order: M X # \$ s \*  
+ = f / - ' .

It should be noted that on the IBM Model 1403 printer the upper-lower case print train will often print faster than the standard upper case train. This is because the upper case (QN) train has 13 symbols that are represented only once on the train, whereas the upper-lower case (TN) train has each of the 120 graphics represented three times. The map should be printed on the unlined side of the paper. Use of 20 pound unlined paper enhances the appearance of the finished product.

The entire map must be printed at the same intensity. A new ribbon often has excess ink, causing two problems: the excess ink may rub off onto the paper, giving it an overall gray cast, and a slight shift in gray tone of the symbols occurs as the excess ink is used. In the figure, the white bands across the width of one page are the result of "dry ends." The ribbon was only several hours old, but it had been in use continuously and the end of the ribbon had not been allowed to replenish itself.

---

\*Photographer, Still Photo Section, University Division of Instructional Services, The Pennsylvania State University.

Most high-speed printers have the option of printing at either six or eight lines to the inch. The number of characters is fixed at ten to the inch. The vertical scale at eight lines per inch is 1:18,700, whereas the fixed horizontal scale is 1:22,600, resulting in expansion in the vertical direction. If six lines per inch spacing is chosen, the vertical scale becomes 1:24,900, yielding only a slight compression in the vertical direction when compared to the horizontal scale. Thus the use of eight lines per inch results in a map that is more representative of the true areal distribution of features. The somewhat distorted scale of such a map, however, makes it essential that both horizontal and vertical scales be indicated.

The figure is a Multilith copy of a photo-reduction of a high-speed printer map of north central Montana, from September 13, 1972 ERTS-1 scene 1052-17452. This map, of lines 1000 to 1600 and elements 1615 to 2324, was composed of 49 pages of computer output. It measured approximately 7 by 7 feet. Adjacent strips were trimmed and cemented together with rubber cement into three strips of two page widths and one strip of three page widths. These four strips were hung on the studio wall overnight to permit the perforation creases to hang out. After matching adjacent strips, double-sided tape was used to anchor the free edges. It was not practical to join all of the strips together to obtain an absolutely flat map. In the future, a vacuum easel, constructed from pegboard framed with two-by-fours and sealed with masonite, will be used to mount the maps.

Photographic Reduction

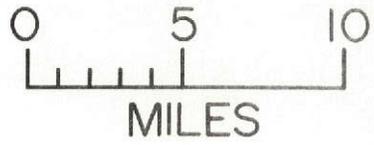
The map was photographed using a 4 by 5 inch studio camera and Kodak's Kodalith film, which is ideally suited for this type of work. The nature of this film permits only a very narrow exposure latitude, hence, care must be taken to evenly illuminate the map. The film was processed in Kodak D-11 developer, rather than Kodalith developer, permitting a limited gray scale to develop in this normally very high contrast film.

The image on a 4 by 5 inch negative represents an approximately 30X linear reduction from the original. As it is the starting point of many possible applications, it is essential to obtain a good negative in order to utilize the method described here to full advantage.

Positive slides were made by placing the 4 by 5 inch negative over an evenly illuminated light source and copying it onto 35 mm Kodalith film, again processing in D-11 developer. Negative slides were obtained by contact printing the positive slides onto Kodak fine grain release film, processed according to Kodak's data sheet. The figure was prepared by standard Multilith photo-transfer techniques.

It is theoretically possible to tilt the negative and/or the paper in the enlarger to obtain a print with the linear distortion removed. Acetate overlays could be prepared, by opaque OZALID processes, from prints enlarged to the proper scale. Overlays could also be prepared by printing onto photographic film.

MONTANA 8 LAND USE SEPT 13



Summary

The procedure described here can be used advantageously to reduce large scale maps produced by computer analysis of ERTS-1 digital data. Slides and large photographic prints can be obtained quite easily. Products suitable for publication are readily obtainable. The advantage of this method of reduction lies in its capability for retaining resolution of the original ERTS data while mapping areas of regional importance. The photographic techniques used in this procedure are available in any well equipped photographic studio.

---

ORSER-SSEL Technical Report 3-74  
The Pennsylvania State University  
February 1974

ABSTRACT

ORSER-SSEL Technical Report 2-74  
GEOLOGIC INVESTIGATIONS OF THE GREAT VALLEY IN PENNSYLVANIA  
D. Krohn and D. P. Gold

ERTS-1 imagery and computer compatible tapes of scene 1080-15185, October 11, 1972 were analyzed to test the effectiveness of discriminating geologic features. The Great Valley of southeastern Pennsylvania was chosen as a training area for several reasons: 1) it encompasses five contrasting lithologies in three physiographic provinces, 2) the background geology is well known, 3) aircraft data is available, and 4) it is currently undergoing rapid urban growth and development.

Geologic features were successfully identified visually on enlargements to a scale of 1:250,000 of the imagery in the four channels as follows: 1) diabase ridges were differentiated from folded sandstone ridges, 2) the Cambro-Ordovician carbonates were partially separated from the upper Ordovician Martinsburg shale, 3) lineaments in the valley were observed to trend from north to northwest, and 4) the trace of the main thrust fault bounding the displaced carbonate sequence was seen.

Using the standard ORSER sequence of tape processing, character maps from a euclidean distance classification program were generated. Spectral signatures were mainly derived from a cluster analysis technique, although a few larger areal features, such as water bodies, could be defined by a homogeneous training area.

Comparison of the digital character maps with the image enlargements indicated that spectral signature responses derived from a mean are sensitive to geologic features when these are enhanced by areal land use and vegetation patterns. Linear geologic features are difficult to enhance with computer techniques, but are readily identified on all scales by visual analysis of image enlargements.

**Interim Report**

**ORSER-SSEL Technical Report 2-74**

**GEOLOGIC INVESTIGATIONS OF THE GREAT VALLEY IN PENNSYLVANIA**

**D. Krohn and D. P. Gold**

**ERTS Investigation 082  
Contract Number NAS 5-23133**

**INTERDISCIPLINARY APPLICATION AND INTERPRETATION OF ERTS DATA  
WITHIN THE SUSQUEHANNA RIVER BASIN**

**Resource Inventory, Land Use, and Pollution**

**Office for Remote Sensing of Earth Resources (ORSER)  
Space Science and Engineering Laboratory (SSEL)  
Room 219 Electrical Engineering West  
The Pennsylvania State University  
University Park, Pennsylvania 16802**

**Principal Investigators:**

**Dr. George J. McMurtry  
Dr. Gary W. Petersen**

**Date: January 1974**

# GEOLOGIC INVESTIGATIONS OF THE GREAT VALLEY IN PENNSYLVANIA

D. Krohn and D. P. Gold

The area of study for this project is the Great Valley, the easternmost extension of the Appalachian Valley and Ridge province. In southeastern Pennsylvania the Great Valley traverses a wide arc from north to east. The particular area of concern was from South Mountain (Carlisle) to the Reading Prong (Reading). In this region the Valley is about 10-15 miles wide. The Great Valley was selected for an area of study for the following reasons:

1. Well-defined boundaries. The forested quartzite ridge of Blue Mountain to the north and the forested diabase ridge of the Triassic basin enclose and stand in sharp contrast to the cultivated land in the valley.

2. Geographic Reference Points. Two major rivers, the Susquehanna and the Schuylkill, act as geographic benchmarks for locating specific areas within the valley. This is especially important in working with computer compatible tapes.

3. Variety of Rock Units and Structural Elements. Including the two boundary areas, there are five major lithologies to be differentiated in the Great Valley area. From north to south these are as follows:

<u>Area</u>	<u>Lithology</u>
The Valley & Ridge Province	(1) Tuscarora Formation (orthoquartzites)
The Great Valley	(2) Autochthonous Martinsburg Formation (graywackes)
	(3) Allochthonous Martinsburg Formation (shales)
	(4) Cambro-Ordovician Carbonates (limestones & dolomites)
The Triassic Basin	(5) Newark Formation (red beds & diabase)

Also within the Valley are a variety of structural features, such as thrust faults, normal faults, and overturned folds.

4. Well-documented ground-based Geology. This area has been studied in detail by the Pennsylvania Topographic and Geologic Survey.

5. Availability of ERTS-1 Data. Several scenes of this area are relatively cloud free; also both the ERTS-1 images and the corresponding ERTS-1 tapes of that scene are available in the ORSER library.

6. Availability of Underflight Data: U2 flights (flown at 65,000 ft) as well as C130 flights (flown between 5000-15,000 ft) have been flown over areas of the Great Valley recently. Patterns and features mapped or noted in ERTS images can thus be checked with higher resolution photography.

7. Personal Knowledge. The authors have a good working knowledge of the topographic and cultural features of the area.

Methodology

ERTS-1 image 1080-15185 from 11 October 1972, was selected as the scene for analysis of the Great Valley because, as of the end of December 1973, it had not been exceeded in quality. Photographic enlargements were made from 4 by 5 inch negatives of approximately one quarter of the area of the 9 by 9 inch transparencies in each of the four channels. These enlargements, made on 16 by 20 inch paper, were on a scale of approximately 1:250,000, which represents an upper practical scale limit for terrain and geologic analysis.

One set of enlarged images centered around the city of Reading and covered a portion of the Great Valley from Allentown in the east to Harrisburg in the west. They included a large portion of the Reading Prong as well as much of the Triassic Basin. The second set of enlarged images centered around the city of Carlisle, and covered the portion of the Great Valley from Lebanon to the east to Hagerstown (Maryland) to the southwest. South Mountain was included in this set of enlargements. There was approximately 25 percent overlap between these two sets.

The first step in the photointerpretive procedure was to prepare an overlay of the topographic, hydrologic and cultural features in the

60

area by projecting a USGS 1:100,000 topographic map on the enlarged ERTS image with a Saltzman Projector. Major rivers were found to be the best identifying feature; highways were too indefinite and mountain boundaries were too vague. The final overlay included the major drainage systems, cities, forest areas and highways. The drainage pattern, in particular, was useful in later analysis. This overlay was used as a guide to orientation in the geologic interpretation of the scene.

The standard ORSER digital data processing system\* was used in conducting a preliminary digital analysis of a portion of the Great Valley, using the digital data tapes from the same scene interpreted visually.

### Results

The features discerned on the October 11 ERTS scene of the Great Valley are summarized on Table 1 and discussed below.

#### Lithologic Separations

The contrasting boundary lithologies of the Great Valley, the quartzite ridge of Blue Mountain to the north, and the diabase ridge to the south, were easily recognized. Because they are heavily forested, they appear dark on channels 4 and 5, and their appearance on channels 6 and 7 was due probably to a low sun angle. Blue Mountain and the other fold mountains appeared as a single continuous bright line. The diabase ridge to the south had a much broader and dimpled appearance compared to the fold ridges. On channel 7 this area has a swirled appearance which may represent flow swirls from an intrusive center.

#### Lineaments

Joints, fracture traces and lineaments are three expressions of linear features at different scales. A joint is a fracture, usually perpendicular to bedding, which does not show any displacement. Joints very commonly appear as patterns of intersecting sets on a scale inches up to a thousand feet. Fracture traces are the linear surface expression

---

\* See ORSER-SSEL Technical Report 9-73, Chapter 2.

Table 1: Geologic Features Interpreted from Enlargements of the October 11 ERTS Image

Feature	Channels
Contrasting topographic areas that are emphasized by differences in tone and vegetation and define physiographic boundaries.	4, 5, 6, & 7
Broad areal patterns of consistent texture that seem to represent areas of similar lithologies. These areas are most readily discriminated by the change in texture at their boundaries. Vegetation tends to enhance the patterns.	4 & 5 (mainly)
Linear features (lineaments) trending north-northwest that transect through the Great Valley and other physiographic provinces.	4, 5, 6, & 7
Linear features that parallel the strike direction in the Great Valley. These appear as lines on the images in contrast to merely an alignment of features. These features appear to correspond to lithologic boundaries or fault traces.	6 & 7
Anomalous areas of high reflectance. These areas may indicate water-saturated soils, but have no apparent geologic affinity.	6 & 7

of joints, or zones of joint concentration, as seen on the ground surface or from low-level aerial photography; they are on a scale of a thousand feet to a mile. Lineaments are linear features from one to several hundred miles in length that are discernible from aerial photography or small scale maps. They are recognized by the linear alignment of a series of discontinuous features. These features seem to be related to the regional tectonic pattern; their intensity and points of intersection have a significant effect on ground-water movement.

Examination of ERTS-1 images reveals many lineaments at all scales that transgress not only physiographic provinces but also rocks as old as the Precambrian and as young as the Triassic. Sixty lineaments were plotted directly on the photographic enlargements.

In channels 6 & 7 three sets of characteristic directions for the lineaments were observed. In the eastern section of the Great Valley around Lebanon, the dominant direction was N 10°E. From the Susquehanna River westward, two intersecting sets of lineament patterns appeared, one heading approximately N 15°W, the other N 70°W. In channels 4 & 5, besides the N 10°E, N 15°W, and N 70°W, a N 45°W direction was also apparent.

Other Linear Features

A dark band striking west approximately bisects the Great Valley. It extends westward from east of Lebanon, southwest to Harrisburg, crosses the Susquehanna River and terminates in a "Y" just south of the Conodoquinet Creek, a total distance of about 35 miles. This feature differs from the lineaments because it is not an alignment of unrelated objects, but actually appears as a dark continuous line, particularly on channel 7. This line corresponds geometrically to the carbonate-shale boundary in the middle of the Lebanon Valley sequence, an interpretation which is supported by the differences in texture and tone on either side of this linear feature. Closer analysis from ground truth and U2 data complicates this apparently simple explanation. A major east-west artery, Route 422, with a string of houses and small villages and towns coincides with this line. A railroad also parallels this highway. West of Annville, a series of elongated limestone quarries

falls in this linear zone. On the U2 infrared imagery these quarries are shown as distinct black lines, and a stereoscopic examination reveals an escarpment, just north of the linear feature. This difference in relief can be explained by the relative erosional resistance of the shales compared to the carbonates. Because the escarpment lies to the north, a continuous shadow of the escarpment is not recorded on the images.

The system of faults is complex. In the vicinity of Annville, fault contacts appear as straight lines, indicating steeply dipping fault planes. To the west, where allocthonous carbonates and the autocthonous carbonates units overlap, the fault trace is highly curved and complex, indicating inclined fault planes. Thus there are two areas of carbonate lithology and two areas of shale lithology that should be detected on the image. A dark linear feature is visible in the U2 infrared imagery which corresponds on a topographic map to a low ridge known as Chambers Hill. The reasons for this dark response along the fault trace vary. In places the linear pattern is only an expression of cultural features on the ground surface, e.g., the quarries or the highway. Bedrock control is expressed in its westward extension through Chambers Hill and across the Susquehanna River.

One of the major thrust faults recognized on geologic maps of this area is the Yellow Beeches. This fault is exposed along the edge of Chambers Hill and extends northward into the Martinsburg formation. On the ground the fault trace in the Martinsburg formation appears as a "distinct textural break and a possible lithologic change running approximately through the middle of the Martinsburg belt with more disordered structures to the south."<sup>1</sup> Such a change in texture is visible on the channel 7 ERTS-1 image, near Harpers, where the tonal pattern changes across a straight boundary. A change in texture is also visible on the infrared images of the U2 underflight.

---

<sup>1</sup>MacLachlan, D. B. and Root, S. I. (1966) Comparative Tectonics and Stratigraphy of the Cumberland and Lebanon Valleys: Guidebook for the 31st Annual Field Conference of Pa. Geologists, Pa. Topo. & Geol. Survey, p. 78-79.

64

### High Reflectance Areas

Several anomalous areas of high reflectance were noted on channel 7 imagery near the town of Hershey. A few miles to the south, the diabase ridge separates into two arms and their pattern on channel 7 appears swirled. Whether these anomalous areas are related to the diabase or are merely artifacts of the image is unknown at this time.

### Conclusions

Several conclusions can be drawn from the examination of enlargements of ERTS-1 images of the Great Valley:

1. Geologic features that are apparent as gross changes on topography and vegetation are easily recognized. Such features readily define physiographic provinces.
2. Geologic features that are expressed as escarpments or differences in relief are identifiable subject to interpretation.
3. Primary geologic structures are not visible unless they are separated by a distinct topography or tonal pattern.
4. A series of linear features -- tentatively defined as lineaments -- are readily visible at all scales.

---

ORSER-SSEL Technical Report 2-74  
The Pennsylvania State University  
January 1974

ABSTRACT

ORSER-SSEL Technical Report 1-74

ANALYSIS AND APPLICATION OF ERTS-1 DATA FOR REGIONAL GEOLOGICAL MAPPING

D. P. Gold, R. R. Parizek, and S.S. Alexander

The advent of ERTS presents an unparalleled opportunity for the analysis of a large body of data on a small scale, gathered under consistent conditions. The objectives of this study were two-fold: regional geologic mapping and the study of linear and transgressive features.

The major mapping criteria were boundaries or interfaces and linear transgressive features (lineaments). Visual image interpretation, and computer techniques using both supervised and unsupervised classification programs, have been used. On a mosaic of channel 7 ERTS-1 images, it was found that bedrock structures show up well, especially in mid-winter scenes, even where not accentuated by topography or vegetation. On channel 5 imagery at a scale of 1:250,000, the contacts of some lithologic boundaries in eastern Pennsylvania can be placed with the accuracy of 400 meters with respect to the 1960 Geologic Map of Pennsylvania. Generally, little bedrock is sufficiently well exposed to exhibit a direct spectral response. Most contacts are reflected indirectly, from soil, vegetation, and land use patterns.

Numerous lineaments have been identified on ERTS images and mosaics, using the photointerpretive approach - - a technique found superior to machine processing because of the variation of lineament expression along strike and the ease of confusion with man-made linear features easily identified by eye. A rose diagram of lineament orientation frequency for a mosaic of Pennsylvania has been constructed, as well as a length-frequency histogram. Three scales of lineaments have been identified and there seems to be an inverse relationship of length to abundance and density. The lineaments transgress rocks of all ages in Pennsylvania and they are not obscured by Pleistocene glacial drift. It is suspected that lineaments overlie fracture zones and prevail to depths corresponding roughly to the same order as their length. If this is true, lineaments represent an important control in the evolution of landscape, localize groundwater movement, are significant in oil and gas migration and leakage, and are of major importance in engineering analysis. Large scale lineaments tend to pass through watergaps and it has been observed that well yields in four such gaps have been nearly double those of wells located in the same formations adjacent to the watergap. Close correlation has been found between five metallic ore deposits and a major lineament in Pennsylvania.

Lineaments, readily seen for the first time on ERTS imagery, represent a significant new class of structural element, one which may generate a resurgence of studies of global fracture patterns and may tie into plate tectonics.

**Interim Report**

**ORSER-SSEL Technical Report 1-74**

**ANALYSIS AND APPLICATION OF ERTS-1 DATA FOR REGIONAL GEOLOGICAL MAPPING**

**D. P. Gold, R. R. Parizek, and S. S. Alexander**

**ERTS Investigation 082  
Contract Number NAS 5-23133**

**INTERDISCIPLINARY APPLICATION AND INTERPRETATION OF ERTS DATA  
WITHIN THE SUSQUEHANNA RIVER BASIN**

**Resource Inventory, Land Use, and Pollution**

**Office for Remote Sensing of Earth Resources (ORSER)  
Space Science and Engineering Laboratory (SSEL)  
Room 219 Electrical Engineering West  
The Pennsylvania State University  
University Park, Pennsylvania 16802**

**Principal Investigators:**

**Dr. George J. McMurtry  
Dr. Gary W. Petersen**

**Date: January 1974**

ANALYSIS AND APPLICATION OF ERTS-1 DATA  
FOR REGIONAL GEOLOGICAL MAPPING

D. P. Gold, R. R. Parizek, and S. S. Alexander

Painstaking synthesis over many years has provided regional, state, national, and global geologic maps. However, the synthesis of features on one scale does not guarantee that a larger feature will necessarily be apparent on the smaller scale map generated. Artifacts in mosaicing, poorly known scaling laws, and inconsistent conditions (variable sun angle, albedo, seasons, etc.) of data collection are more likely to obscure than enhance subtle features. The advent of ERTS however, presents an unparalleled opportunity for the analysis of a large body of data on a small scale, gathered under consistent conditions. This is the ideal situation for the discovery and study of large scale features and their geologic significance.

Objectives

The initial objectives of this study are two fold:

- 1. Regional geologic mapping: a few well known areas have been studied in detail, with a view toward application of the results to larger areas.
- 2. Linear and transgressive features apparent on ERTS and all scales of aircraft imagery are being studied. Again, for smaller scale linear features, initial study has been concentrated on a few well known areas.

Because we have a considerable body of ground truth data for establishing correlations, and because there is a theoretical basis for possibly linking linear features through scale, we feel that objective number two will have the best and quickest economic pay-off in terms of hydrogeology, engineering and environmental geology, and ore-deposits.

65

## Methodology

The methodology is continually developing as we gain experience in interpreting the imagery and learn to recognize the real and significant signals. The criteria used in mapping like areas (visual similarity in tone, spatial patterns, or texture) are classified according to whether they represent a direct or indirect manifestation of the bedrock condition. In forested areas such as Pennsylvania, a knowledge of the indirect indicators is important for geologic interpretations even though their relationship to the bedrock conditions may not be understood. The main mapping criteria used are:

1. Boundaries or interfaces that separate areas of different tone, texture, or pattern. Whereas irregular boundaries generally result from differences in land use (arable land versus forests), smooth and regular boundaries commonly reflect geologic control, especially where layered rocks are involved. Combinations of these two relationships (e.g., forest cover over untillable rocky areas) enhance contrast and interpretability, if the correlation can be made and the cause identified from ground truth data. For example, the diabase sills in eastern Pennsylvania show up best on channel 5 because their overlying forest cover stands in contrast to the surrounding cultivated fields.

2. Linear transgressive features (lineaments) that show as narrow bands of contrasting tone and topography, or that displace areas of like tone or pattern. These are generally long features (five to several hundred km long). Some morphologically represent the alignment of wind and water gaps, and others represent the surface expression of features such as dikes, faults, and zones of fracture concentration without any apparent displacement.

Three criteria were used to select geographic areas for detailed evaluation: 1) the availability of good ground truth data, 2) the presence of underlying rocks with good lithologic contrast, and 3) the presence of abundant faults. The areas chosen are:

1. the Traissic Basin of eastern Pennsylvania and associated diabase sills

- 69
2. the Anthracite Basin around Scranton
  3. the Precambrian inliers of the Reading Prong
  4. the Martic Line north of Philadelphia
  5. transgressive long lineaments through Mount Union and Tyrone
  6. transgressive lineaments in the South Mountain area
  7. short and intermediate lineaments, for orientation and density comparisons, in different parts of the State.

The choice of a contained problem in a restricted area is important at this stage, not only to keep down digital processing costs but also to keep the computer-generated maps to a manageable size (each character represents approximately 1.1 acre).

The evaluation is being conducted in three stages:

1. Primary correlations are sought by projecting the available ground truth geologic boundaries onto enlargements (1:250,000) of ERTS imagery in all available channels. Anomalies are checked out first on other maps (agricultural, soils, highways, topographic, etc.), and followed with a field check where necessary.

2. Selected areas are mapped by computer, using cluster analysis and the digitized data, with the parameters controlled (supervised) from training areas.

3. This is followed by unsupervised computer mapping (i.e., free from known geologic biases) to bring out any latent features that might have geologic significance.

### Geologic Structures

A mosaic of ERTS-1 images of Pennsylvania was prepared using the first available cloud-free scenes. No attempt was made to obtain uniform tone or alter individual frames. Rather, the frames were simply spliced together to permit an early study. Physiographic and structural provinces are displayed spectacularly on this mosaic, and the resolution achieved in some images enables one to trace formation contacts for hundreds of

70

kilometers. Bedrock structures show up well, especially on channel 7 in midwinter scenes, even where not accentuated by topography or vegetation. On channel 5 imagery at a scale of 1:250,000, the contacts of some lithologic boundaries in eastern Pennsylvania (see Figure 1) can be placed with the accuracy of 400 meters (1/4 mile) with respect to the 1960 Geologic Map of Pennsylvania. Much of this error may be a result of transferring the boundaries from ERTS imagery to the base map. The mapping of superimposed structural features such as faults, particularly along the northern end of the Triassic Basin, was disappointing. While the margin of the Reading Prong could be traced from the tonal and land use variations, the geologically mapped faults were not everywhere apparent. Vegetation enhancement over the Triassic diabase sills and dikes rendered mapping both simple and accurate. Fracture and drainage patterns, and tonal variations, serve to distinguish certain rock types but generally little bedrock is sufficiently well exposed to exhibit a direct spectral response. Most contacts are reflected indirectly, in the condition and type of overlying soil, vegetation, and land use.

#### Linear Features

Perhaps the most encouraging and unexpected characteristic of ERTS imagery is the number, distribution, and patterns of unspecified linear features which can be seen. Geologists have long recognized the presence of straight to slightly curved linear features on the earth's surface. These vary in size from a few to tens of meters, for systematic and nonsystematic joints, to "lineations" tens of kilometers long that have no obvious field expression and are visible only on airphoto mosaics. To distinguish among features recognizable on aerial photographs, Lattman<sup>1</sup> defined a "fracture trace" as a "natural linear feature consisting of topographic (including straight stream segments), vegetation, or soil

---

<sup>1</sup>Lattman, L. H. (1958) "Technique of Mapping Geologic Fracture Traces and Lineaments on Aerial Photography," Photogram. Engineering, Vol. 24, p. 568-576.

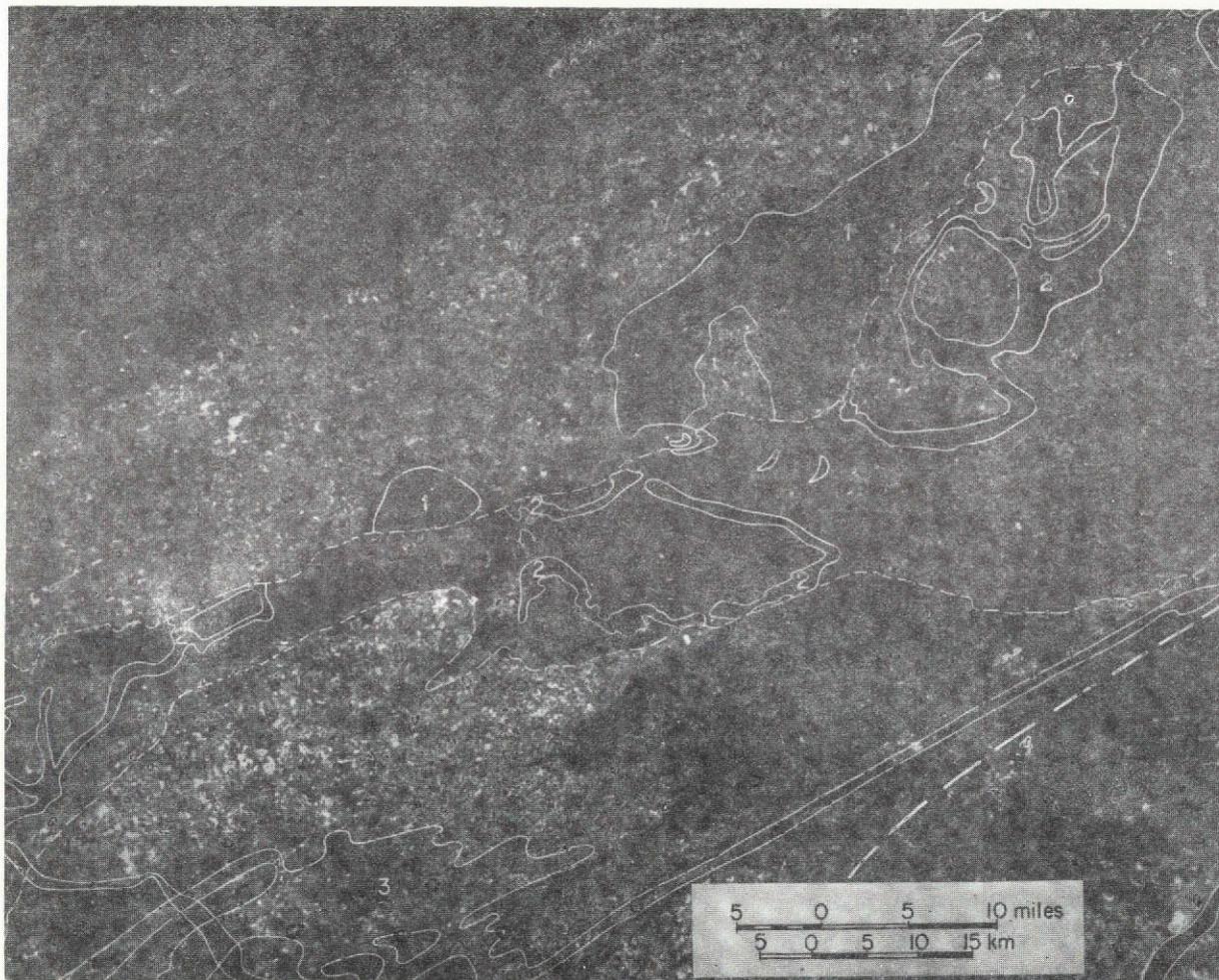


Figure 1: ERTS image of an area east of Harrisburg, Pennsylvania, with geologic features outlined. 1) Precambrian rocks of the Reading Prong, 2) the Triassic Basin (dashed lines) and diabase sills, 3) the Conestoga formation, and 4) the Martic Line in the Piedmont province. (Enlargement of ERTS-1 image 1080-15185, channel 5, October 11, 1972.)

tonal alignments, visible primarily on aerial photographs and expressed continuously for less than one mile." Those greater than one mile (1.6 km) he termed "lineaments." A considerable amount of work has been done with joint traces, and joint orientation studies in the field, and more recently (since 1957) with fracture traces. Using ground based and aircraft data as well as that from ERTS, we now recognize at least six scales of linear features. While there must be a mechanism linking those between the joint and fracture trace scales, theory suggests that the same mechanism may link all scales.

Many lineaments that transgress regional structural grain and also physiographic province boundaries have been discovered from visual examination of ERTS-1 MSS bulk-processed images. ORSER geologists, who have mapped some lineaments and portions of others from high quality aerial photo-mosaics and from low-flying aircraft, and who have studied aerial photographs and conducted field work for years in central Pennsylvania, did not recognize the trend of many of the major lineaments of the State until ERTS images became available. Linear features are detectable on ERTS images and mosaics, as well as on underflight photography, as lines or bands defined by an alignment of valleys, of wind and water gaps, and by straight stream segments or a linear change in image tone. Many are subtle features, especially in areas of low relief, and the tonal contrast may vary along their length. They may be enhanced in mountainous areas by a low sun angle. Shorter linear features (several meters to several kilometers in length) are best seen on aerial photography or ERTS enlargements. When features 15 or more kilometers long are being mapped, enlargement may render the longer lineaments (over 100 kilometers) less obvious. Lineament mapping from ERTS digital tape data is cumbersome because of the variation of lineament expression along strike (e.g., between ridges and valleys, or even from one valley to another) and because of the difficulty of distinguishing lineaments from man-made effects (such as roads) and artifacts of data collection (such as scan lines). Our work to date has been concentrated on visual approaches, because the human eye is more sensitive for mapping these features than any machine processing tried so far.

73

A first attempt at a "megalineament" map of Pennsylvania on a 1:1,000,000 mosaic of channel 7 images has been completed (see Figure 2). A rose diagram of lineament orientation frequency for this map has been constructed (Figure 3), as well as a length frequency histogram (Figure 4). The necessary computer programs to analyze the orientation and length of linear features and to compare these in different cells have been developed<sup>1</sup>. Similar analyses on intermediate and short lineaments in a test area east of Harrisburg are being performed to try and establish whether or not lineament lengths are part of a continuum or have a stepped distribution. Orientation comparisons will be made with areas of known folds and faults in an attempt to determine if a single or divergent pattern exists in various parts of Pennsylvania. This will be tested around the major flexure in the Appalachian folded mountain chain. Because lineaments are sensed as a trace on the surface and are perceptible as subtle changes in tone or contrast that may vary along their length, lineament mapping in this test area is being duplicated, using the same images and two operators (Parizek and Gold), to determine the reproducibility of the technique. Both orientation and density data are being recorded.

Three scales of lineaments have been recognized: 1 to 5 miles (1.6 to 8 kms), 5 to 50 miles (8 to 80 kms), and 50 miles (80 kms) to a few hundred miles (kms) long. Lineaments on a subcontinental scale are anticipated. Little is known about their length, frequency, and relationship to fracture traces and joints, but there appears to be an inverse relationship of length to abundance and density. Lineaments mapped to date appear to have a consistent relationship to fold and fault axes, are generally straight features, appear to cut across physiographic provinces, and are not influenced by faults. Their linear nature regardless of topography suggests they are the surface expression of near-vertical fractures or fracture zones. They transgress rocks of all ages in Pennsylvania and blankets of Pleistocene glacial drift do not obscure them. These deep-seated features must imprint themselves

---

<sup>1</sup>See Podwysocki, M. H., "Computer Applications of Fracture Trace Analyses," and "Computer Algorithms for Use in Fracture Trace and Lineament Analysis," NASA X Documents 644-73- (in press).

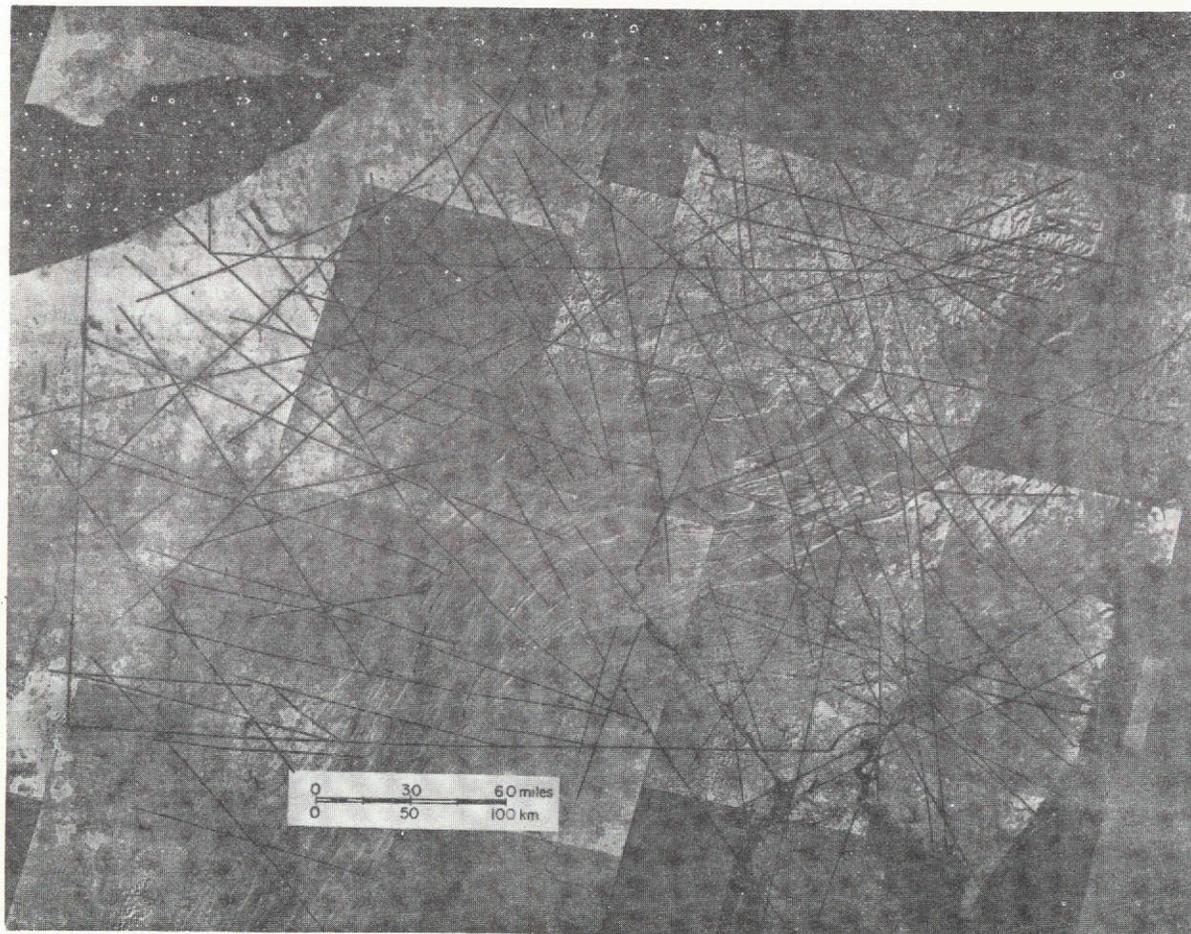


Figure 2: Megalineament map of Pennsylvania, plotted on a mosaic base of 1972 channel 7 ERTS-1 images.

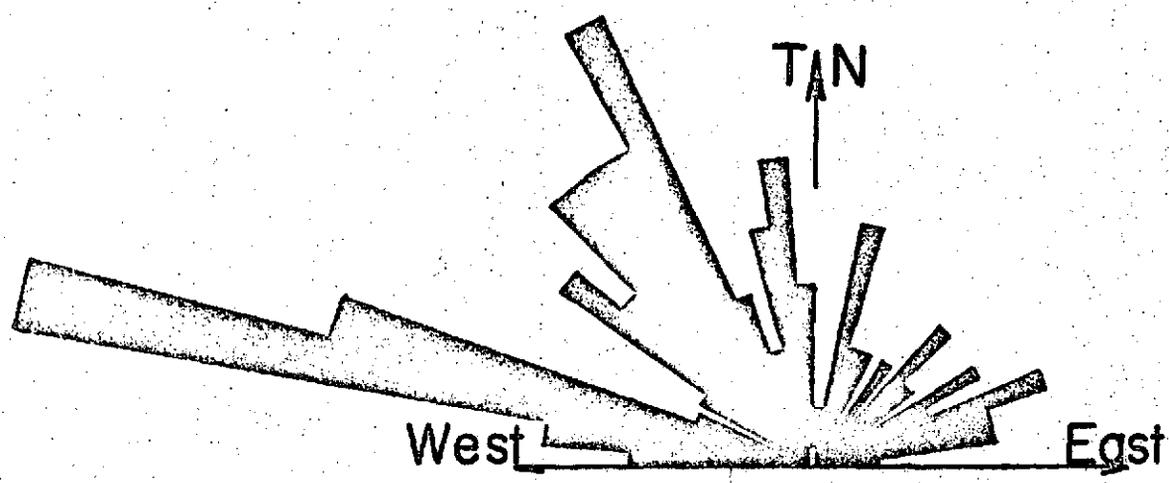


Figure 3: Rose diagram of long lineaments plotted on the megalineament map of Pennsylvania.

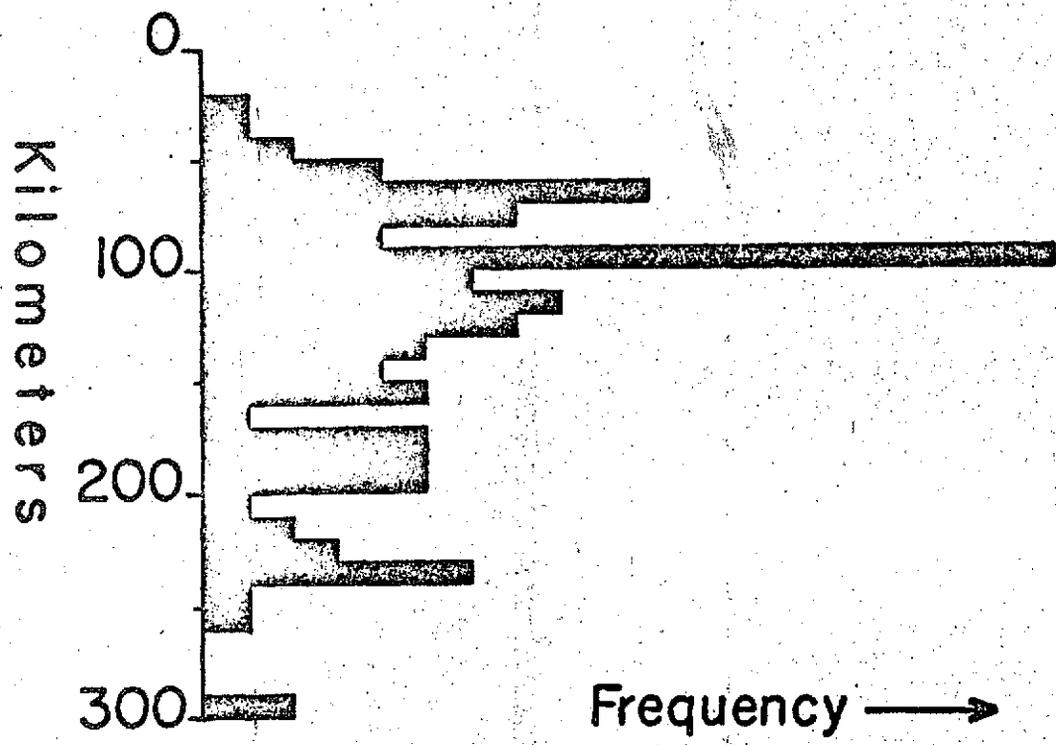


Figure 4: Histogram of long lineaments plotted on the megalineament map of Pennsylvania.

76

on younger deposits in a systematic manner and are themselves inherently old. Major lineaments, for example, can be traced from Precambrian metamorphic and igneous rocks through overlying Paleozoic sedimentary rocks and into the down-faulted sediments and diabase sills of the Triassic Basin. They must represent either rejuvenated fractures, a "tectonic inherence" from the underlying crustal rocks; or a recently imposed fracture system, as might be expected from stresses associated with a drifting lithospheric plate.

Some "lineaments" are actually fault traces, in that lateral offsets in individual rock layers can be established. Many major faults previously mapped in Pennsylvania that are transverse to regional stratigraphic or structural strike appear on ERTS images. Many previously unmapped faults also should become obvious with further study. Recently, the probable motion deduced from first arrival directions of both the P and S waves for an earthquake near Philadelphia, was combined with lineament directions mapped from ERTS-1 imagery to locate the most probable fault plane. The seismic data were compiled by Dr. Shammus of the Geophysical Section of the College of Earth and Mineral Sciences, at Penn State, in co-operation with seismologists from the Lamont Geophysical Observatory of Columbia University, New York.

The unexpected density of short and intermediate length lineaments mapped on 1:250,000 scale of ERTS-1 channel 7 images in the test area east of Harrisburg (Figure 5) suggests a scaled-up version of fracture traces. If the lineaments overlie fracture zones (as is suspected) and prevail to depths corresponding roughly to the same order as their length (as implied by theory), then these implications are important to:

- 1) stream and river control and the evolution of landscape;
- 2) ground-water movement;
- 3) oil and gas migration and leakage;
- 4) underground gas storage; and
- 4) engineering foundation exploration, analysis, and design.

Part of our research thrust is in the three-dimensional characterization of these lineaments. Unfortunately, obtaining pertinent information at depth on a feature of this scale is difficult and may take years of careful synthesis of data on sinkhole development; drilling and pumping records; and poor ground conditions in mines, quarries, and highway construction projects.

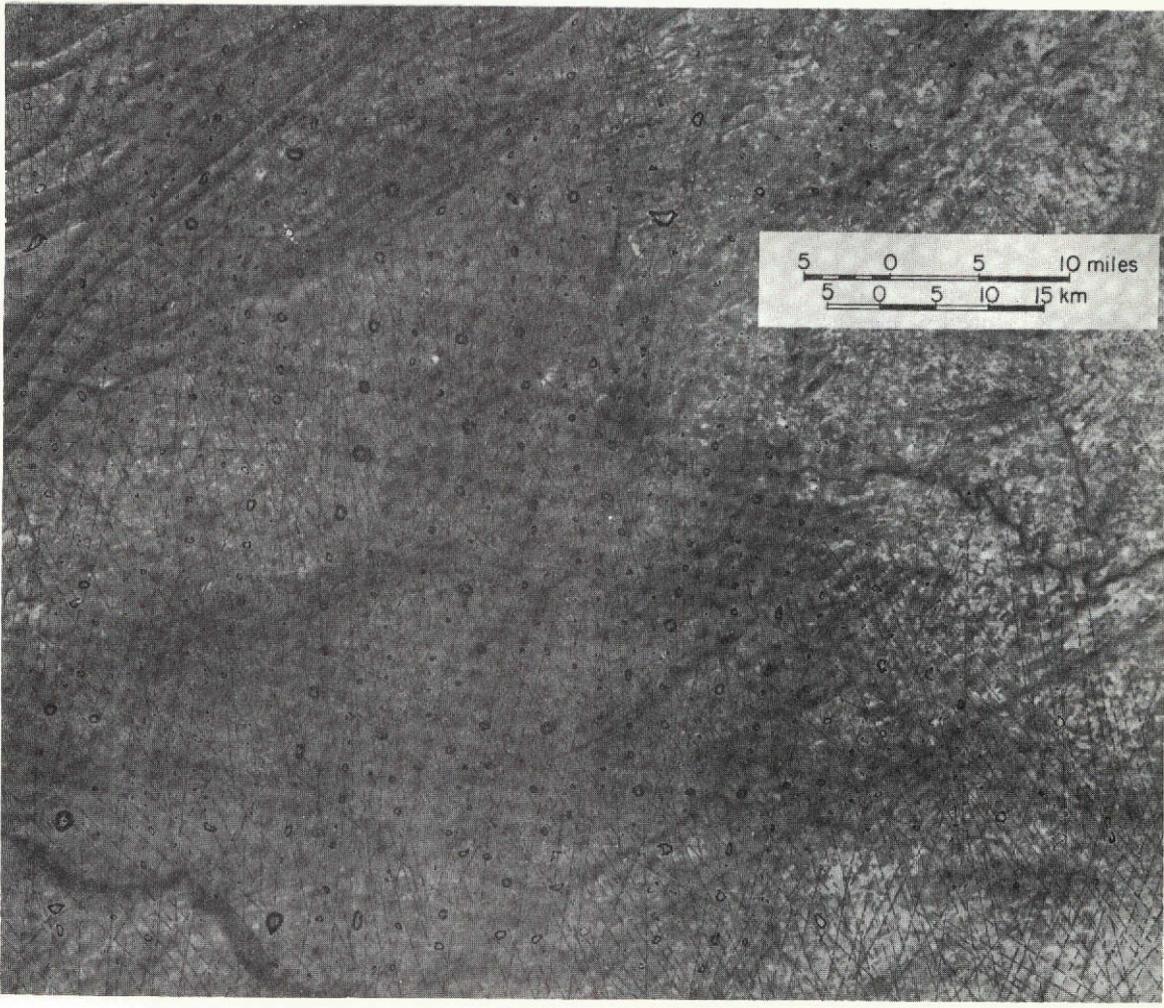


Figure 5: Short to intermediate lineaments crossing the Valley and Ridge, Great Valley, Triassic Basin, and Piedmont structural provinces east of Harrisburg, Pennsylvania. The Susquehanna River appears at the lower left and the Schuylkill River in the northeastern half of the figure. Reading is located on the Schuylkill near the center. (Enlargement of ERTS-1 image 1116-15192, channel 7, November 16, 1972.)

## Conclusions

### Geologic Structures

The scale relationships of geologic information available on data sensed at different altitudes has not been investigated, as originally proposed, because of the lack of underflight data in areas where good ERTS imagery first became available. However, we have looked at the problem of scale as it affects lineaments and fractures in other areas. We suggest there is a link in mechanism between joints; fracture traces; and short, intermediate, and long lineaments; and that their identification is related to scale. A study of this link is a major objective of one aspect of our ongoing research. We plan also to systematically map lineaments and fold axes of Pennsylvania on a scale of 1:250,000. A tectonic map of the State will then be developed, with the aid of existing structural data provided by the Pennsylvania Topographic and Geologic Survey.

### Ore Deposits

A most important spin-off of the lineament map is its potential application to the location of ore deposits. Using ERTS imagery, close correlation has been observed between metallic ore deposits and the Mount Union-Tyrone lineament in Pennsylvania (Figure 6). Five metallic sulfide deposits are known to be located along this lineament. The bedrock conditions are currently being investigated, and underflight photography along this feature has been requested. In the Ebensburg area, an investigation is being conducted to determine the quality of coal and mining conditions relative to short lineaments observed on the ERTS imagery. The groundwork has been laid to procure quality control data from the Bethlehem Steel Corporation underground mines near Ebensburg. We have started a program to investigate mining problems and the possible relationships between lineaments and coal distribution and quality on the Allegheny plateau, near Clearfield. A similar study is planned near the major lineaments cutting the anthracite field in the Wyoming basin, near Scranton and Wilkesbarre.

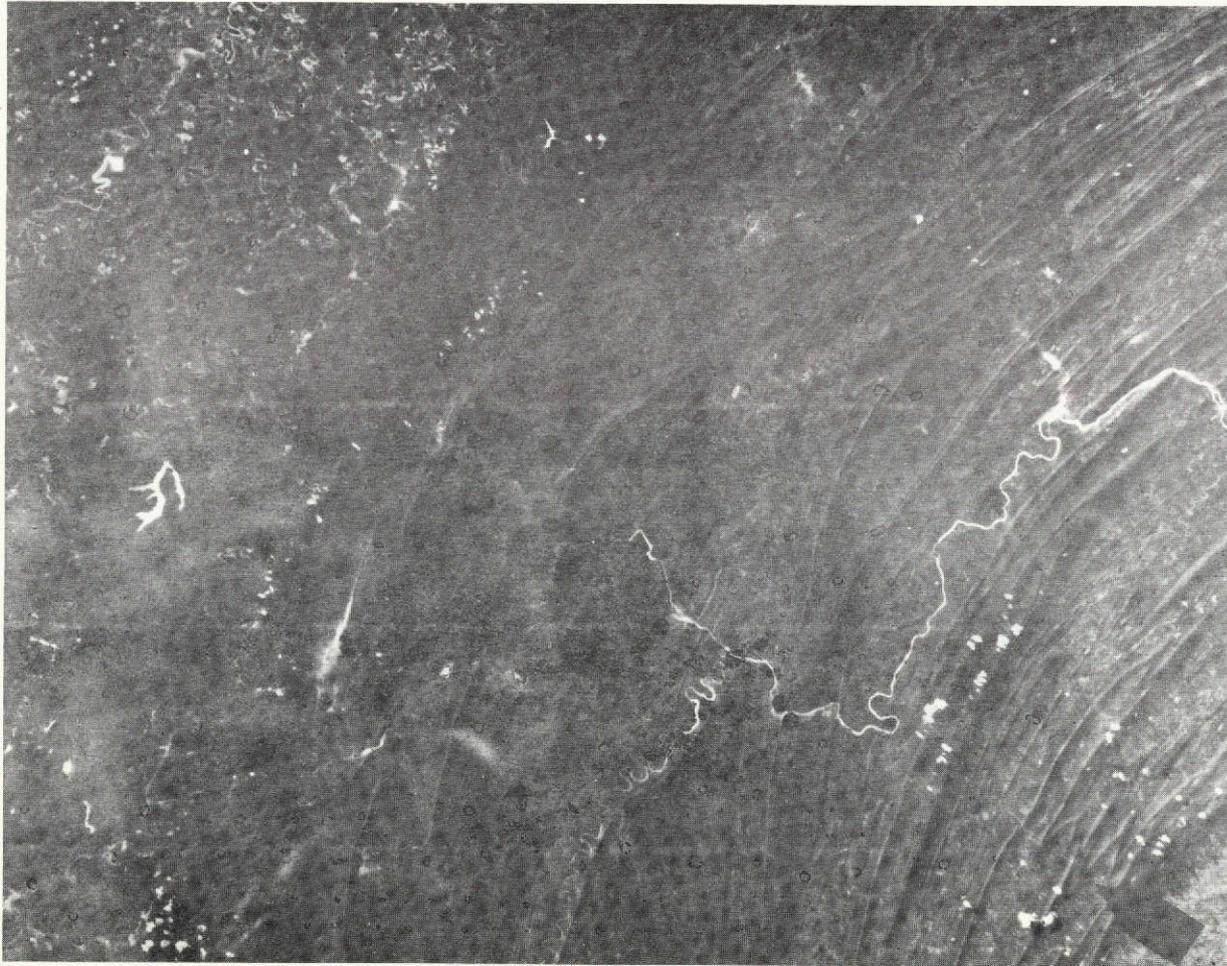


Figure 6: Major lineament between Mount Union and Tyrone, Pennsylvania, which correlates closely with known metallic ore deposits. Lead and zinc deposits are shown as dots. Note that this lineament also localizes the Juniata River along this stretch. (Negative enlargement of ERTS-1 image 1045-15243, channel 7, September 6, 1972).

Ground-water

Locations of known ground-water sources are being correlated with fracture traces and lineaments observed on ERTS imagery. Investigations at Penn State since 1961 have established the significance of the use of fracture trace mapping techniques to locate highly productive water wells drilled in carbonate rocks. Many of these areas were otherwise very poor water producers. Investigations in areas of siltstone, shale, coal, and sandstone sequences are showing a similar relationship of fracture traces to well yields<sup>1</sup>. In preliminary studies, Parizek has found an association of well yields and water gap locations, and it has been frequently observed that large scale lineaments tend to pass through water gaps. For each of the four watergaps under study, well yields have been found to be nearly double those of wells located in the same formations adjacent to the water gap. Various geologic and hydrologic factors other than lineaments could account for the yield differences observed, and these must be isolated and studied. However, success to date indicates that ERTS data will be a highly productive source of information for linear mapping on a regional scale that should aid in ground-water exploration.

Aeromagnetic Intensity

A preliminary examination of the relationship of lineaments observed on the ERTS mosaic to such aeromagnetic intensity maps as are available for Pennsylvania, revealed a distinct difference between the correlation in the western and the eastern portions of the State. In western Pennsylvania, where the lineaments appear to connect magnetic lows, the magnetic anomalies are thought to be basement controlled. In the eastern portion of the State the anomalies are highly concentrated and tightly intertwined, with no visible relationship between lineaments and the magnetic patterns.

---

<sup>1</sup>Brown, R. L. and R. R. Parizek (1971) "Shallow Ground-water Flow Systems Beneath Strip and Deep Coal Mines at Two Sites, Clearfield County, Pennsylvania," Special Research Report, SR-84, Coal Research Section, The Pennsylvania State University, University Park, Pa.

Continuing Research

ERTS imagery provides a timely base for mapping lineaments, a new class of structural element, and the ground follow-up (geological and geophysical probing for the three-dimensional aspect) should aim to characterize these features, not only to facilitate the development of a genetic classification but also to assess their economic potential and utility. The ERTS program has already spurred the development of a theory to account for the size and frequency of "linear" features on all scales<sup>1</sup>, and these should be refined so that dynamic analyses of stress distribution are possible. We foresee in the study of linear features from ERTS data a resurgence in studies of regmatic shears and global fracture patterns and a possible tie into plate tectonics. More important is the potential use of lineament mapping in mineral and groundwater exploration. The ERTS programs provides a new research tool which improves on the resolution of features that were previously rather obscure. Hardware from the ERTS program provides us with such a new tool.

---

<sup>1</sup>This theory was offered by Moody and Hill in 1956 ("Wrench Fault Tectonics," Bull. Geol. Soc. America, Vol. 67, p. 1207-1246) and is being developed by Gold, Parizek, and Alexander.

---

ORSER-SSEL Technical Report 1-74  
The Pennsylvania State University  
January 1974

82

THE PENN STATE ORSER SYSTEM FOR PROCESSING AND  
ANALYZING ERTS AND OTHER MSS DATA <sup>1</sup>

G. J. McMurtry, F. Y. Borden, H. A. Weeden, G. W. Petersen  
Office for Remote Sensing of Earth Resources (ORSER)  
The Pennsylvania State University

Abstract

The Office for Remote Sensing of Earth Resources (ORSER) of the Space Science and Engineering Laboratory (SSEL) at The Pennsylvania State University has developed an extensive operational system for processing and analyzing ERTS-1 and similar multispectral data. The ORSER system was developed for use by a wide variety of researchers working in remote sensing. Both photointerpretive techniques and automatic computer processing methods have been developed and used, separately and in a combined approach. A Remote Job Entry (RJE) system permits use of an IBM 370/168 computer from any compatible remote terminal, including equipment tied in by long-distance telephone connections. An elementary cost analysis has been prepared for the processing of ERTS data.

Introduction

The Office for Remote Sensing of Earth Resources (ORSER) of the Space Science and Engineering Laboratory (SSEL) at The Pennsylvania State University has developed an extensive operational system for processing and analyzing ERTS-1 and similar multispectral scanner (MSS) data. The ORSER system was developed for use by a wide variety of researchers working in remote sensing. These users represent many disciplines and have a wide range of experience and skill in photointerpretation and computer usage.

Interpretive techniques which are used in the system include computer processing, visual image interpretation, and a hybrid technique which closely integrates photointerpretive techniques with computer analytic procedures.

---

<sup>1</sup> Presented at the Third Annual Remote Sensing of Earth Resources Conference, March 25-27, 1974, The University of Tennessee Space Institute, Tullahoma, Tennessee.