LARS Information Note ~~110276~~ 091669

APPLICATION OF AUTOMATIC RECOGNITION TECHNIQUES TO EARTH RESOURCES

R. B. MACDONALD

The Laboratory for Applications of Remote Sensing

Purdue University, West Lafayette, Indiana

1976

Published in the Proceedings of the Second Annual Earth Resources Aircraft Program Status Review, NASA Manned Spacecraft Center, Houston, Texas. September 16 - 18, 1969. Pages 23-1 to 23-27.

SECTION 23

APPLICATION OF AUTOMATIC RECOGNITION TECHNIQUES TO EARTH RESOURCES

R. B. MacDonald

Introduction

The National Aeronautics and Space Administration in cooperation with the United States Department of Agriculture and Purdue University continued to sponsor the Laboratory for Applications of Remote Sensing at Purdue throughout 1968/1969 to participate in the research and development of remote sensing techniques for application in agriculture.

LARS conducts a multidisciplinary program designed to explore critical problem areas in agriculture applications. Exemplary of these is the requirement for automatic processing capabilities. Automatic processing is essential to the design of information systems which adequately meet the needs of resource development planners and resource use managers. Thus, a key program area to LARS is devoted to developing a capability of automatically recognizing important earth resource features (Holmes and MacDonald, 1969; Fu, et.al., 1969).

A successful approach to this problem must include consideration of collection schemes which produce data which are generally amenable to automatic information extraction routines. Past research conducted at Purdue/LARS has established spectral radiance data to be of this nature (LARS Volume II, 1967; LARS Volume III, 1968). In 1968 and 1969 LARS continued to explore and to develop this approach in multiple investigations applied to situations involving the naturally occurring materials such as:

- vegetation
- soils
- water

In support of these investigations, LARS has organized programs in the following research areas:

- Biogeophysical
- Measurements
- Data Processing
- Requirements and Applications
- Aerospace Systems Analysis

Biogeophysical research programs are conducted by teams of application scientists. Teams have been organized to deal respectively with problems of vegetation, soils, and water. The primary objectives of this research are the correlation of remotely sensed measurements to physical and chemical properties of radiating scenes.

The results of research at LARS to apply advanced remote sensing and processing techniques to agricultural situations are discussed in this paper.

Data Collection Missions

To meet the LARS 1968-69 objectives, data collection missions were scheduled over Tippecanoe County at particular times throughout the agricultural season. Data were to be collected over some 200 square miles along the five flight lines shown in the Purdue test sites (Figure 1). Each mission was scheduled at what was considered to be a cardinal point in the agricultural season in Indiana. Multispectral measurements together with multichannel photography were to be collected by the University of Michigan aircraft in April, June, August, October, and December. The aircraft was to collect data from altitudes of 3000 and 5000 feet and at various times of the day.

The data were fundamental inputs for LARS investigations of the 'automatic recognition' problem over larger geographical areas. Additionally, analysis of the data were to be performed to further delineate various features which could be identified and to quantify the results.

The larger quantities of data collected in each mission were essential for a more valid verification of performance accuracies.

Analysis of Soils Data

Major objectives of the April mission were to examine fields of winter wheat and areas of fresh tilled soil. In this period, fields are for the most part either in a freshly tilled condition or are in winter wheat. The April mission was unavoidably delayed until May. At the time of flight oats and other green vegetation were present in addition to winter wheat. However, a significant percentage of the area flown was in a tilled stage and much valuable data of the soils was collected. A principal objective of our soils scientists was to evaluate possibilities for soils properties mapping on the basis of spectral measurements. This project was based on previous work accomplished with data collected over central Indiana in the spring of 1967.

It is appropriate to discuss here for a moment some of the aspects of soils which the agriculturist is generally interested in. Principal components of soils are:

- minerals (clay, silt, sand)
- organic matter

- air
- water

Important characteristics of a particular soil include:

- aeration
- water holding capacity
- supply of decomposing organic matter
- available plant nutrients

Elements essential to plant growth are shown in Table 1. Some of the minor elements are needed in very small amounts. Manganese deficiencies were found on soybeans in Tippecanoe County during July even though the plants needed less than 5 ppm of this element.

Table 1 Macro and Minor Elements Needed for Plant Growth

Macro Elements	Minor Elements
. Calcium	. Copper
. Magnesium	. Manganese
. Potassium	. Zinc
. Nitrogen	. Iron
. Phosphorus	. Boron
. Sulfur	. Molybdenum

Knowledge of the organic matter content of soils is of particular importance because it:

(1) indicates soil fertility and nutrient availability

(2) indicates soil drainage

(3) affects soils moisture

LARS soil scientists selected test regions along flight lines within the 500 square mile Tippecanoe County test site. Conventional soils maps were prepared for the selected regions to provide ground truth information (Figure 2). It should be noted that the soils maps are based on soil profiles as well as on surface conditions. Soils types, slopes, and state of erosion were identified. An area covering 60 acres and referred to as the Dieterle Farm (shown within the rectangle in Figure 2) was selected for more detailed ground truthing. Two hundred and four soil samples were collected on a 100' x 100' grid (Figure 3) for detailed laboratory analysis. Chemical analysis defined such items as organic matter content, iron concentrations, mineral composition, etc. for each of the samples. Spectral data collected by the aircraft over this area in the May mission was subjected to computer aided analysis.

Investigators determined that the soils could be divided into many (10-20) categories on the basis of their spectral categories with soils characteristics of interest. The requirement and applications area, however, indicates that such a large number of categories is not practical for general use. It would appear that five to seven categories would be most practical for agricultural purposes.

Each of the 204 laboratory samples was given a 'Munsell' color designation. The samples fit into seven groups on the basis of their Munsell designations. Training samples were then selected from the spectral data for each of the seven groups. Statistics obtained from the spectral training sets were used to automatically classify all points of the Dieterle Farm (Figure 4). Typewriter symbols were selected to identify each group. Additionally the groups were colored to provide the reader with a clear delineation.

The results can be compared to the 'ground truth' map as shown in Figure 5. The reader will note that the computer-selected groups compare favorably with the soil types on the soil map. LARS investigators believe the computer derived map to be an excellent description of the color properties of the surface soils. Investigators are planning to develop a conventional soils map of an adjacent area with the aid of a computer derived spectral map of the surface soils. The results are to be evaluated on the basis of time savings and improved accuracy.

An attempt was also made to associate organic matter content of the surface soils with spectral characteristics after noting the variation of color of the surface soil samples. A soil collage was assembled from actual surface soils samples taken from the 204 points and compared to the soils map (Figure 6). Classification analysis was then conducted to separate the surface soils into one of five categories on the basis of organic matter content. The categories represented mean values 1.5%, 2%, 2.5%, 3.5%, and 5.0% organic matter. A sixth category was designated for the ditch plus 'spoil'. Analysis of training samples indicated that of the available spectral bands from 0.4 to 3 microns, the best four wavelength bands for classification were:

0.62 - 0.66 microns

1.00 - 1.40 microns

1.50 - 1.80 microns

2.00 - 2.60 microns

The classification results using a combination of these wavelength bands are shown in Figure 7. Classification accuracies as deduced from lab analysis of the 204 sample points together with field checking were estimated to be in the neighborhood of 90 percent.

During these studies the effects of mis-registration on classification accuracies were noted. Channels of spectral data from 1 to 3 microns were not spatially registered with the channels from .4 to 1 microns. The effects of such mis-registration are indicated in Figure 8. The result of this was a spectral measure of the ditch material in one set of channels coupled to a spectral measure of the field in the second set. An 'adultrated' response was measured and mis-classifications resulted. The effects of this on classification results are illustrated in Figure 9. The area represented as ditch and spoil was clearly more accurately classified with the registered data on the left.

Soil scientists at LARS are currently investigating the effects of high iron content on organic matter delineations. Figure 10 shows a Brazilian soil which is high in iron compared to a high organic Indiana soil type. The objective of these efforts is to assess the effect of this and other conditions on the detection of organic matter in surface soils.

<u>Vegetation Research</u>

A second major effort at LARS is directed at developing a capability for the automatic recognition of floristic and physiognomic characteristics of natural and cultural vegetation, i.e., identification of species and of physical appearance of the vegetation of an area (Sinclair, 1968; Swain and Fu, 1968). A large number of activities can be served by vegetation maps. Application of vegetation maps include the following:

- scientific investigations
- geographical research
- climatic records
- pedology and geology
- agriculture
- forestry
- land management and planning
- commerical, engineering, and fiscal interests

While vegetation maps are used for a variety of purposes one should not conclude that there are as many types of vegetation maps as there are applications. Careful analysis of the various applications lead one to conclude that it is reasonable to standardize such maps in the future. As an example, one could decide on three basic map types (Kuchler, 1967): (1) a physiognomic type, (2) a floristic type, and (3) a physiognomic floristic combination.

Such maps are generally concerned with illustrating a characteristic of the cultural and natural vegetation of a region. In agriculture the user is primarily concerned with important species of cultural vegetation of an area. LARS has thus far concentrated efforts on developing a capability to recognize cultural species important to agriculture in the corn belt, i.e., corn, soybeans, and wheat and to discern important physical and chemical characteristics of the vegetation of these species.

It should be recalled that missions were flown in late June and July 1966 for similar purposes. Since data was collected over small areas (4-8 square miles) in 1966 a major objective of the 1969 missions was to work with data collected over much larger areas. Thus, LARS researchers planned to utilize the June 1969 mission data to reassess their capability to detect wheat over large areas prior to its harvest. LARS also wanted to reassess the capability to identify corn and soybeans in late June, July, and August over larger areas, i.e., Tippecanoe County. At this time only partial analysis of the data collected in these missions has been completed. Therefore, the results which can be reported herein must be considered to be of a preliminary nature. The analysis has thus far been hampered by the lack of color photography collected at the time of flight for use in analysis routines. LARS expects to be receiving these before long and will upon its receipt be able to improve results through better training sample selection and evaluation of classification results.

As an example of this, 1969 is known to be a year in which weeds have been especially numerous; wet weather early in the season prevented cultivation operations. In the case of mature wheat it is now believed that 5 to 10 percent of the remote sensing units will prove to have been correctly classified as green vegetation as opposed to mature wheat. LARS investigators observed that some 5-10% of the area within wheat fields were classified as green vegetation. If these are determined to be correct classifications, the accuracy of wheat recognition will increase by that amount. Analysis of color photography is to provide such answers in the near future. The reader should note that as new recognition schemes are developed, data from former missions are subjected to additional analysis and new results achieved (Huang, 1969).

The results of the analysis of June 1966 data and June 1969 data for identification of winter wheat is shown in Table 2.

Table 2 Automatic Identification of Winter Wheat in June for Two Different Years

Date	Wheat Resolution Samples	Wheat Samples Correctly Identified	Percent Accuracy
5-28-66	2736	2540	92.9
5-25-69	7859	6719	85.5

The results are encouraging. Good accuracies have been achieved over considerably larger areas in 1969. Further analysis described above is expected to improve the recognition accuracies above the 85 percent figure.

The results of analysis of 1966 and 1969 spectral data to identify corn and soybeans are shown in Table 3. The accuracy of the 1966 data was improved using a Per Field Classifier. The accuracy of the identification of corn was increased from 78 to 95 percent, for soybeans from 75 to 82%, and for row crops from 90 to 97%. With the Per Field Classifier use is made of the probability density function of the field based on spectral measurements of the remote sensing units (RSU's) within the field, the other analysis techniques used in this work are based spectral measurements obtained on a per point basis. It is expected that similar increases in accuracies will be realized with the 1969 data upon similar treatment. The reader should recall the analysis of the 1969 data is not complete and is only of a preliminary nature.

Table 3 Automatic Identification of Corn and Soybeans for Different Years

Crop	Total Fields Tested	Fields Correctly Classified 1/	Percent Correctly Classified	Total RSU's Tested	% RSU's Correctly Classified
			July 26, 1966		
Corn	55	43	78.0	11486	68.0
Soybeans	44	33	45.0	13781	69.0
Row Crops	99	90	90.0	-----	----
			June 25, 1969		
Corn	87	55	63.2	-----	----
Soybeans	66	28	42.4	-----	----
			August 5, 1969		
Corn	121	109	90.0	11507	84.7
Soybeans	78	64	82.0	6736	84.4

1/ Fields were termed correctly identified when more than 60 percent of the Remote Sensing Units (RSU's) were identified within each field.

Upon complete analysis of the data collected during 1969, LARS researchers hope to establish when and with what accuracy the major species can be identified during the growing seasons. Preliminary analysis of the late May data indicates that headed wheat may be discriminated from non-headed wheat. These first missions were scheduled in 1969 at what researchers estimated to be cardinal points for determining certain key factors. LARS expects that analysis of these data will provide valuable guidance for scheduling missions in 1970.

LARS investigators conducted a study to identify certain tree species and to detect beetle infested Ponderosa Pines. The results of these exploratory studies were promising and are to be pursued in future efforts. The principal investigator in this study was Joseph C. Bell, Jr. of the U. S. Forest Service while a special student at Purdue University under the direction and supervision of LARS.

The results of one study shows that conifers and deciduous trees could be separated and further grouped into light and dark categories (Figure 11).

Much of the analysis of the vegetation data being collected in 1969 remains to be done and is to be accomplished throughout the winter months.

Water Research

Research investigations of an exploratory nature were conducted to further evaluate the capability to automatically detect surface water bodies and to make water quality assessments.

As LARS has reported in the past, one of the simpler categories of natural materials to identify is surface water bodies (Figure 12). All areas covered by water have been accurately identified and mapped. LARS has customarily experienced recognition accuracies of better than 95%. Stream intersections where each stream has a varying sediment load (Figure 13) appear to be typical examples which can be detected by automatic mapping techniques. A portion of the White River in the vicinity of Indianapolis, Indiana (Figure 14) was automatically mapped. Computer aided analysis has delineated seven spectral categories of water. The original objectives of this particular data mission did not include this analysis and, consequently, ground truth was not collected to define existing water properties and meaning cannot be directly attached to the various categories which were discriminated in this case. LARS intends to pursue this in future studies.

Other Research

In addition to these efforts LARS has conducted exploratory research to assess the feasibility of accomplishing the following:

- Classification of certain engineering characteristics of soils in cooperation with the joint Highway Research Project at Purdue University.
- Identify certain geologic features with data collected in central Indiana.
- Categorization of natural materials in Yellowstone National Park in cooperation with U. S. Geological Survey.
- Classification of non-agricultural wildland resources for inventory purpose in cooperation with the Forestry Remote Sensing Laboratory.

The objectives of the engineering soils studies included:

- Investigation of variables influencing the radiation characteristics of natural materi als in the spectral inverval from about 0.4 microns to 15 microns.
- Investigation of computer generated engineering soils maps based on digitized multispectral radiance measurements.

Data was collected in the spring of 1967 by the University of Michigan aircraft over a 70 mile long test site extending from Indianapolis, Indiana to Bedford, Indiana. A portion of the test site was in a limestone quarry area (Figure 15) leaving various geologic and vegetative features. The principal investigator, M. G. Tanguay, was generally able to:

(1) Classify land surface strips in terms of vegetation, soils, and water. Maps were obtained that distinguished and automatically classified up to seven soil types utilizing multispectral measurements.

(2) Delineate unique soil conditions on individual graphic printout maps, thus emphasizing the distribution of adverse soil conditions.

(3) Illustrate, for the first time, a relationship between spectral response of a soil and its land form and parent material.

(4) Relate a type of soil and its source location which was about a quarter a mile away.

This research is to be carried on in future projects due to the successful results received to date.

Uses of automatic recognition schemes utilizing multispectral measurements were investigated with data collected over Yellowstone National Park. The project was conducted in cooperation with H. W. Smedes and K. L. Pierce of the U. S. Geological Survey, Denver, Colorado.

Objectives of the project included evaluation of these techniques to identify and map different natural terrain types, and analysis of aircraft collected data to study the effectiveness of the proposed Earth Resources Technology Satellite (ERTS) data channels. The investigation generally revealed that highly satisfactory terrain maps portraying geologic-geomorphic units and rock-soil-vegetation association units could be produced (Figure 16).

Although attempts were not made to simulate coarser resolution of the ERTS sensors, the computer-generated maps indicated a good overall performance that compared closely with a map compiled through use of four channels selected by the computer.

It suffices to say that the results of these exploratory studies were highly encouraging and in-depth investigations are to be conducted in these areas in the future.

REFERENCES

1. Fu, K. S., D. A. Landgrebe, T. L. Phillips, 1969, "Automatic Reduction of Remotely Sensed Earth Resource Data," IEEE Proceedings Volume 57, Number 4, April.

2. Holmes, R. A. and R. B. MacDonald, 1969, "The Physical Basis of System Design for Remote Sensing in Agriculture," IEEE Proceedings, Volume 57, Number 4, April.

3. Huang, T., 1969, "Per Field Classifier for Agricultural Applications," LARS Information Note 060569, Laboratory for Agricultural Remote Sensing.

4. Kuchler, A. W., 1967, Vegetation Mapping, Ronald Press Company, New York, New York.

5. Laboratory for Agricultural Remote Sensing (LARS), 1967 Remote Multispectral Sensing in Agriculture. Volume II (Annual Report), Purdue University Agricultural Experiment Station, Research Bulletin 832.

6. Laboratory for Agricultural Remote Sensing (LARS), 1968, Remote Multispectral Sensing in Agriculture. Volume III (Annual Report), Purdue University Agricultural Experiment Station, Research Bulletin 844.

7. Sinclair, T. R., 1968, "Pathway of Solar Radiation through Leaves," Master's Thesis, Purdue University.

8. Swain, P. H. and K. S. Fu, 1968, "On the Application of a Non-parametric Technique to Crop Classification Problems," Proceedings of the National Electronics Conference, Chicago, Illinois.

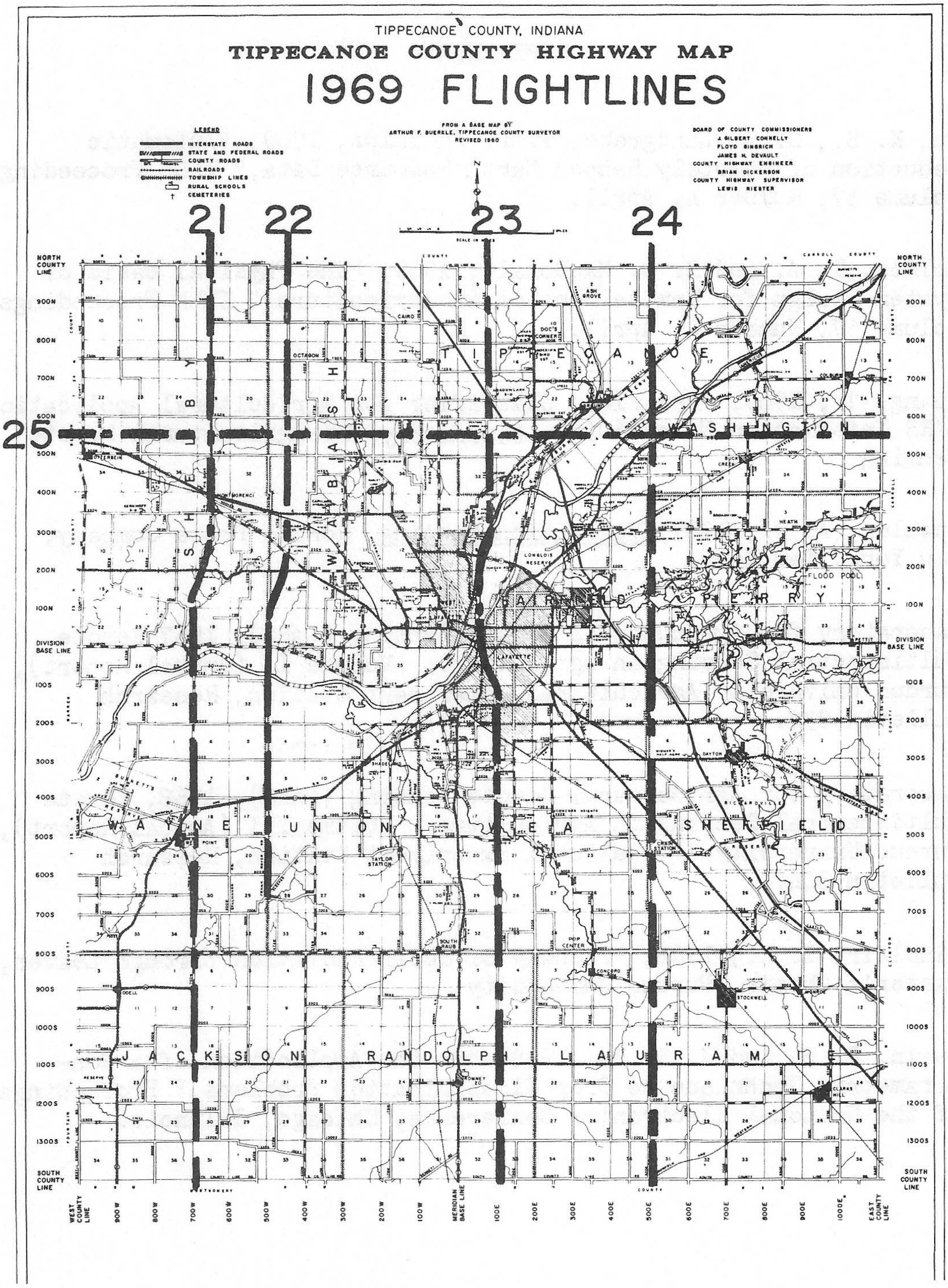


Figure 23-1.- 1969 flight lines over the Tippecanoe County Test Site.

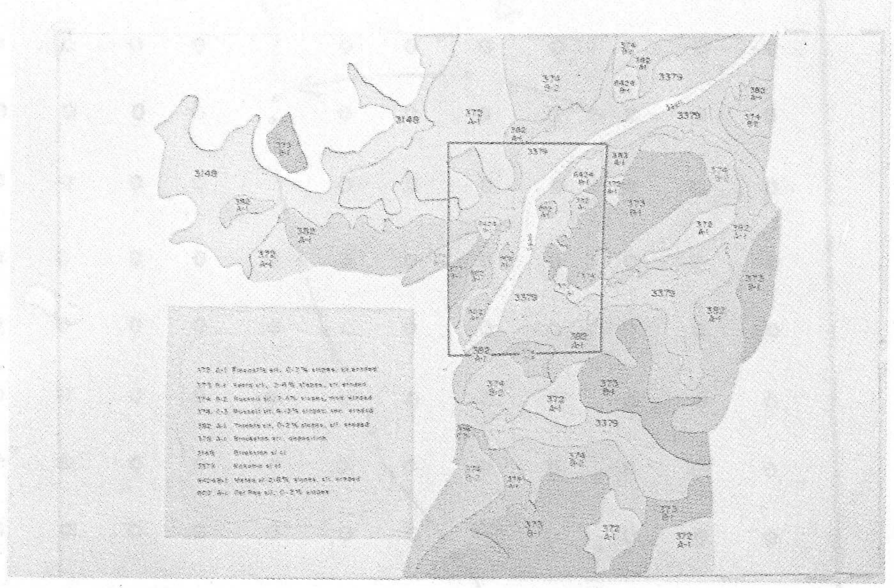


Figure 23-2.- Soils "ground truth" map of the Dieterle Farm and surrounding area.

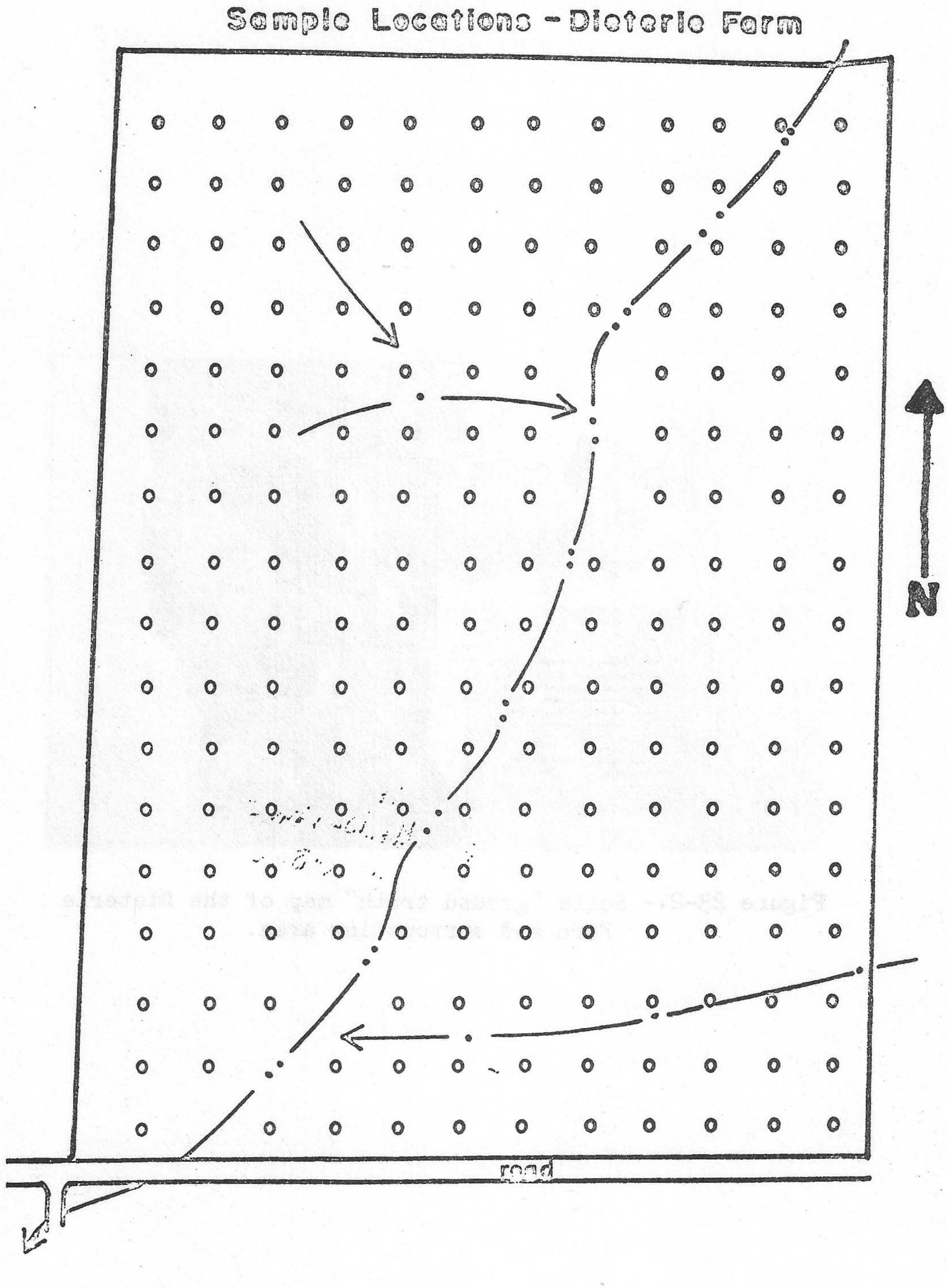


Figure 23-3.- Map of the Dieterle Farm showing locations where surface soil samples for laboratory analysis were obtained.

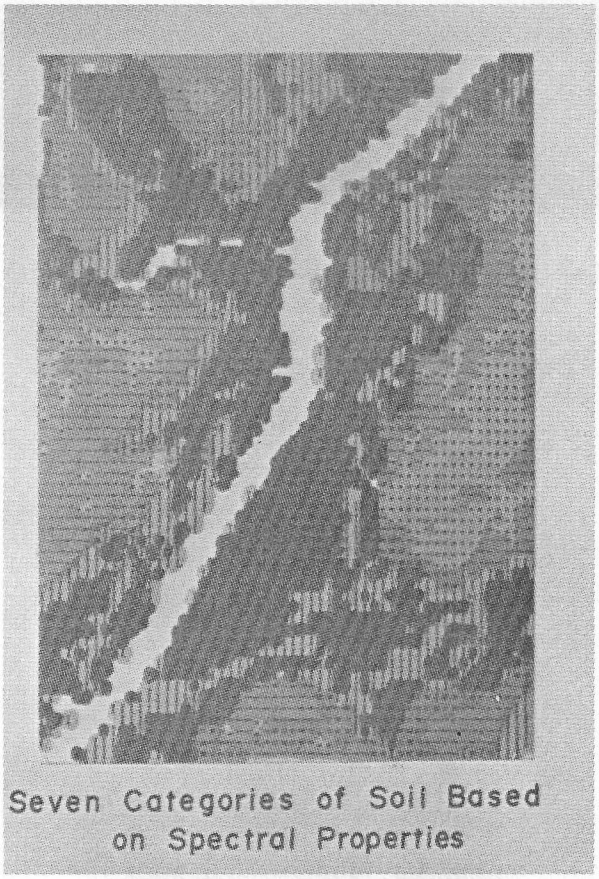
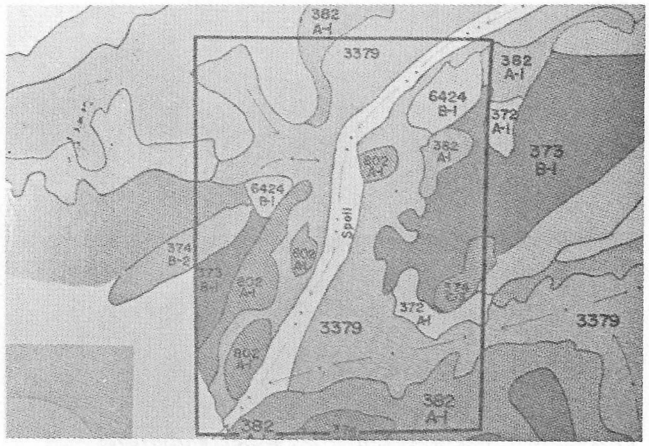
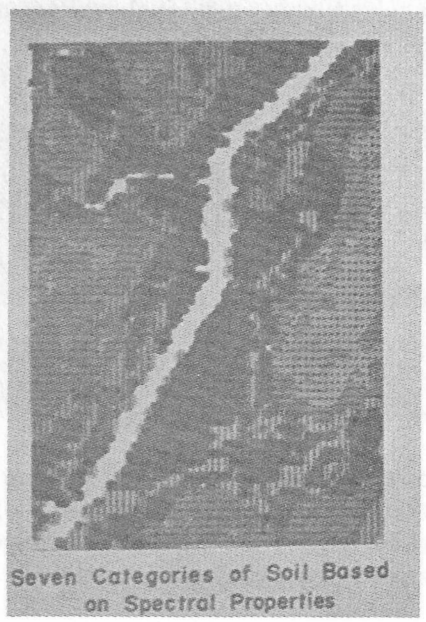


Figure 23-4.- Spectral Soils Map of the Dieterle Farm showing seven soil categories based on spectral properties measures by aircraft scanner.

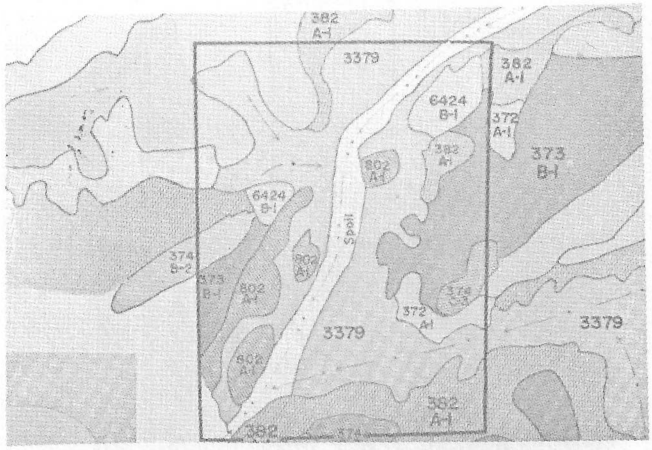


(a) Dieterle Farm area (rectangular field) showing detailed soils types as mapped in the field by a Soil Scientist.

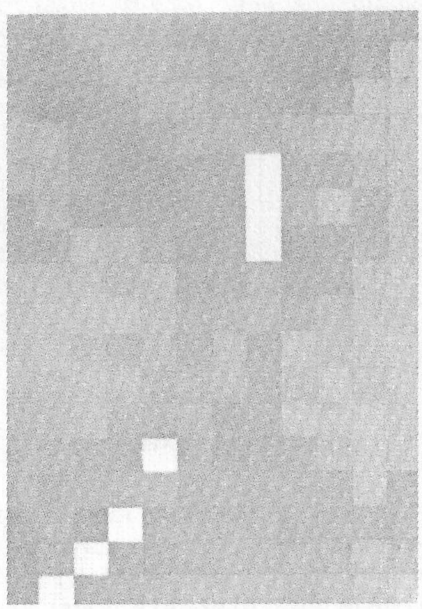


(b) Spectral soils map derived automatically by computer.

Figure 23-5.- Comparison of "ground truth" and "automatic" map.



(a) Soils map of Dieterle Farm.



(b) Soil collage compiled from the 204 soil samples taken from the Dieterle Farm.

Figure 23-6.- Comparison of "ground truth" soils map and soil collage.

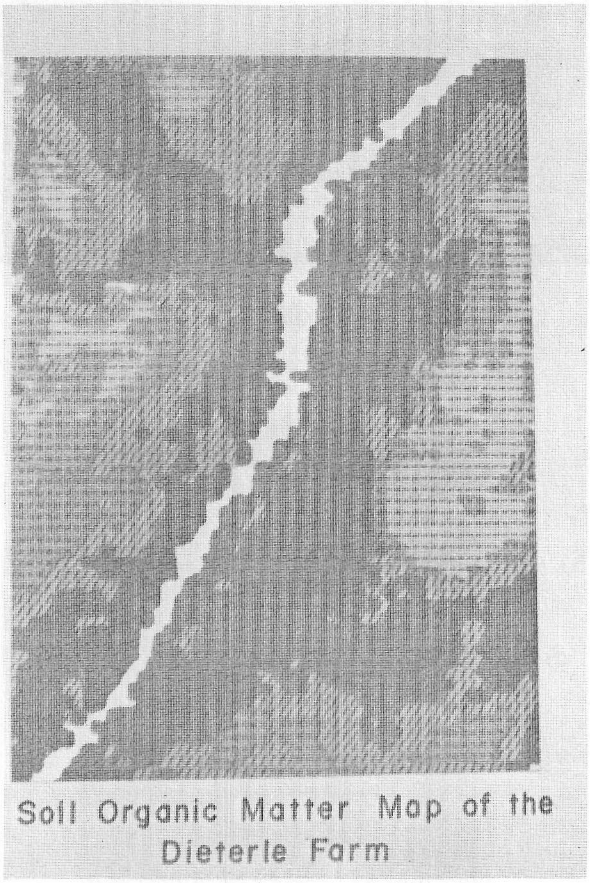


Figure 23-7.- Automatic map of organic matter content derived in conjunction with laboratory analysis of soil samples from the Dieterle Farm and spectral measurements near sample locations.

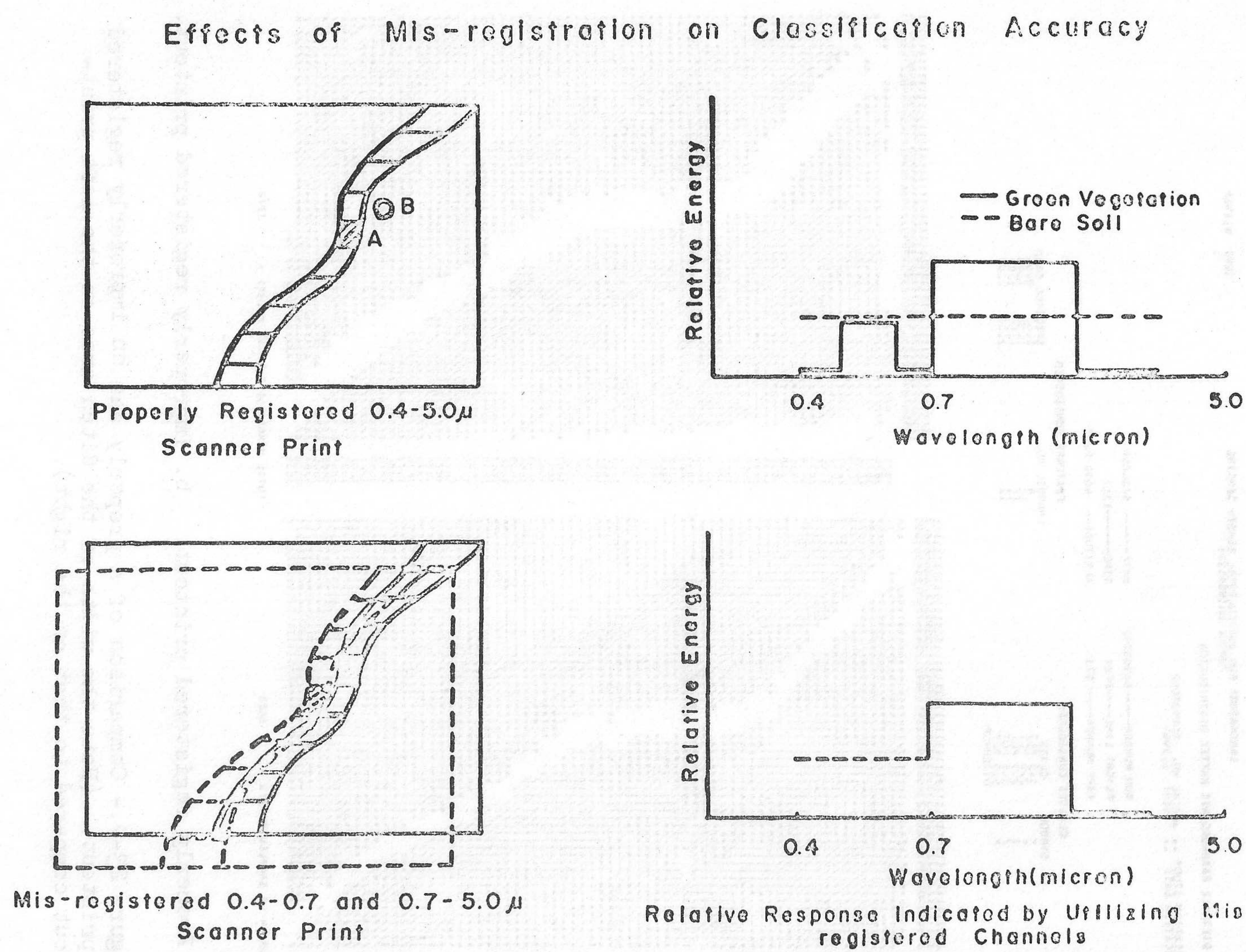
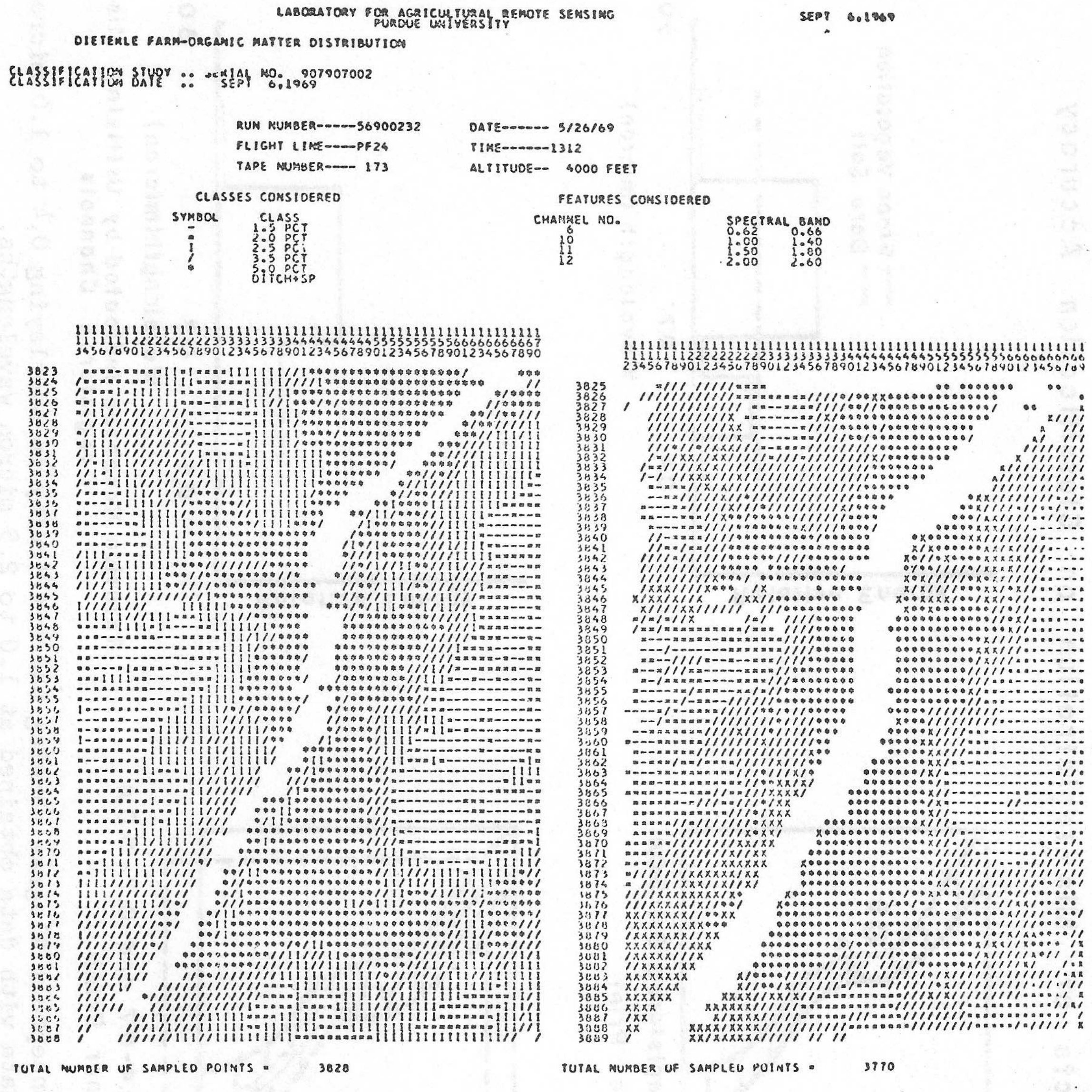


Figure 23-8.- Conceptual diagram showing the difficulty in overlaying 0.4 to 1.0 micron data with data obtained at 1.0 to 2.5 micron wavelengths.



a. Properly registered printout. b. Improperly registered printout.

Figure 23-9.- Comparison of a properly and an improperly registered printout. (Note the width of the ditch area on the left printout compared to that on the right).

(a) Brazilian soils having high iron content.

(b) Indiana soils with a high percent of organic matter.

Figure 23-10.- Comparison of soils high in organic matter and soils which are high in iron.

Figure 23-11.- An aerial photograph and corresponding computer print-out over a forested area. Research has been successful in differentiating conifers and deciduous species into separate groupings.

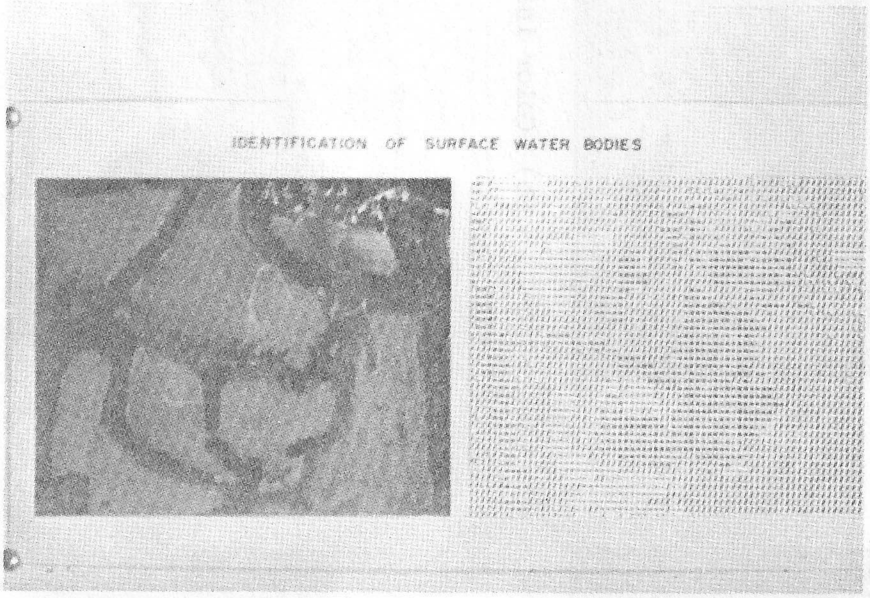
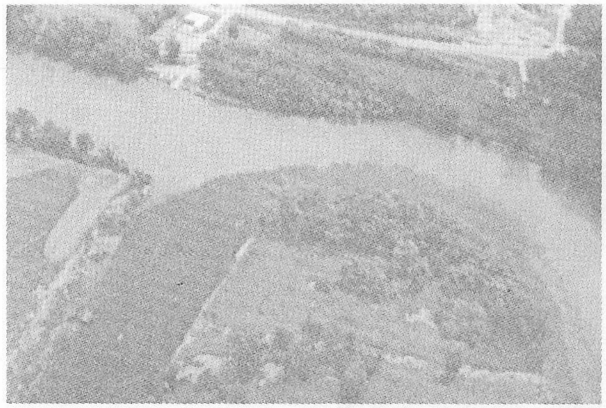
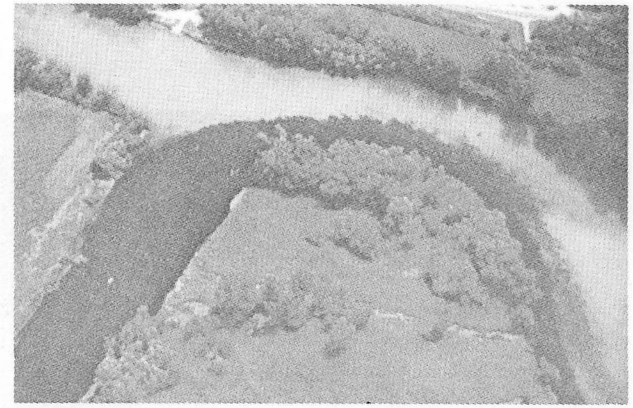


Figure 23-12.- An aerial photograph and computer print-out of an area in Southern Indiana, most of which is underwater. Experienced interpreters have difficulty in accurately mapping the water bodies on the photo. The computer, however, easily identifies water due to its spectrally different characteristics from surrounding soil and vegetation.

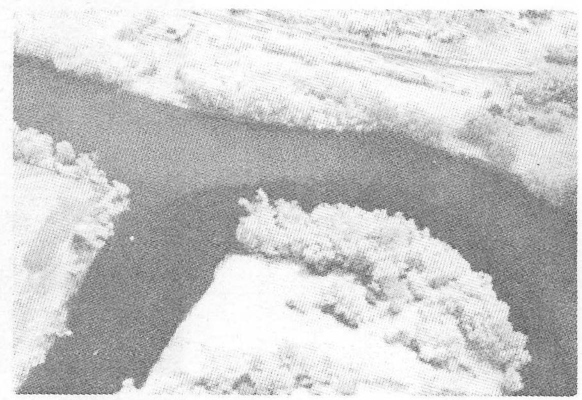


(a) Conventional color.



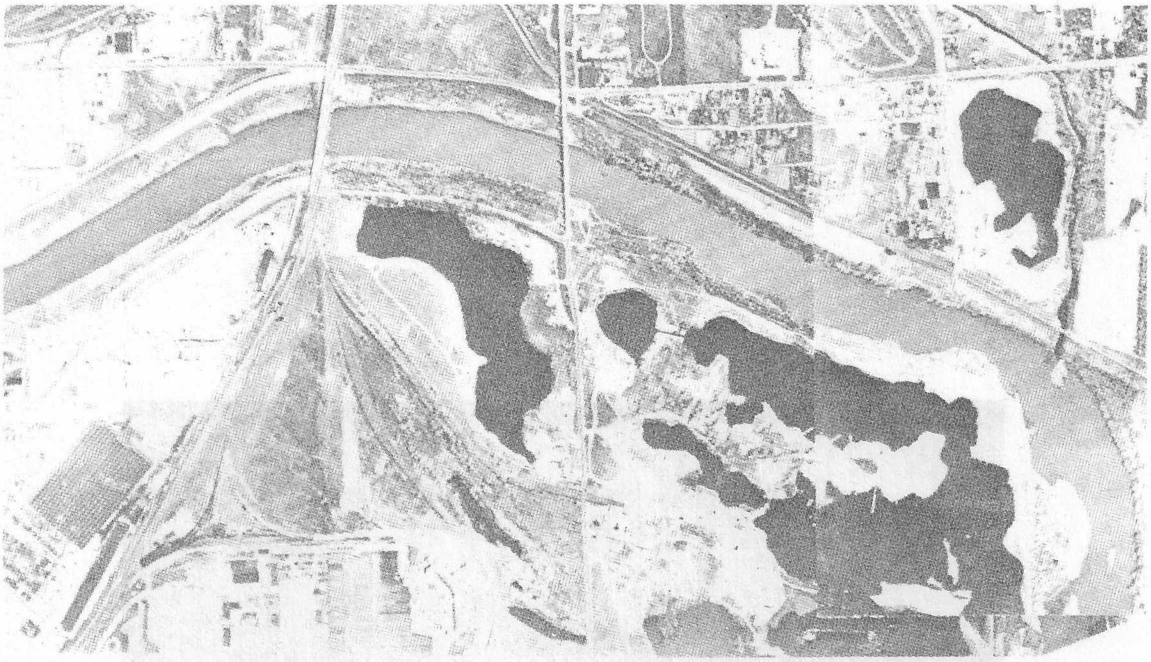
(b) Color infrared.

(c) Panchromatic black and white.

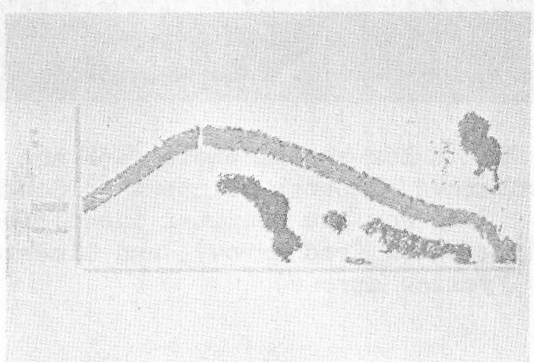


(d) Black and white infrared.

Figure 23-13.- Four film type comparison showing the junction area of two rivers.



(a) Black and white panchromatic photograph taken near Indianapolis, Indiana, showing the White River and flooded quarry pits.



(b) Artificially enhanced computer printout of the same area showing the water areas. The computer separated the water into seven distinct categories.

Figure 23-14.- Comparison of aerial photography and computer printout for hydrologic purposes.

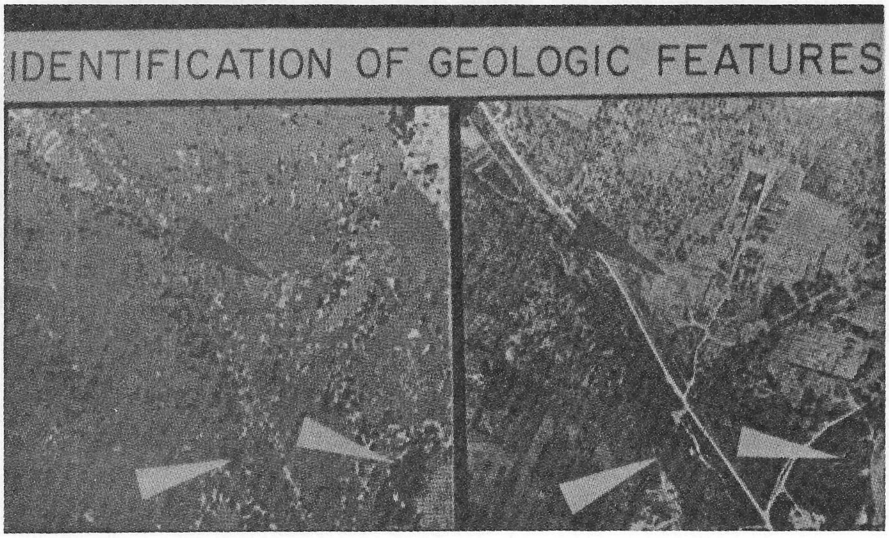
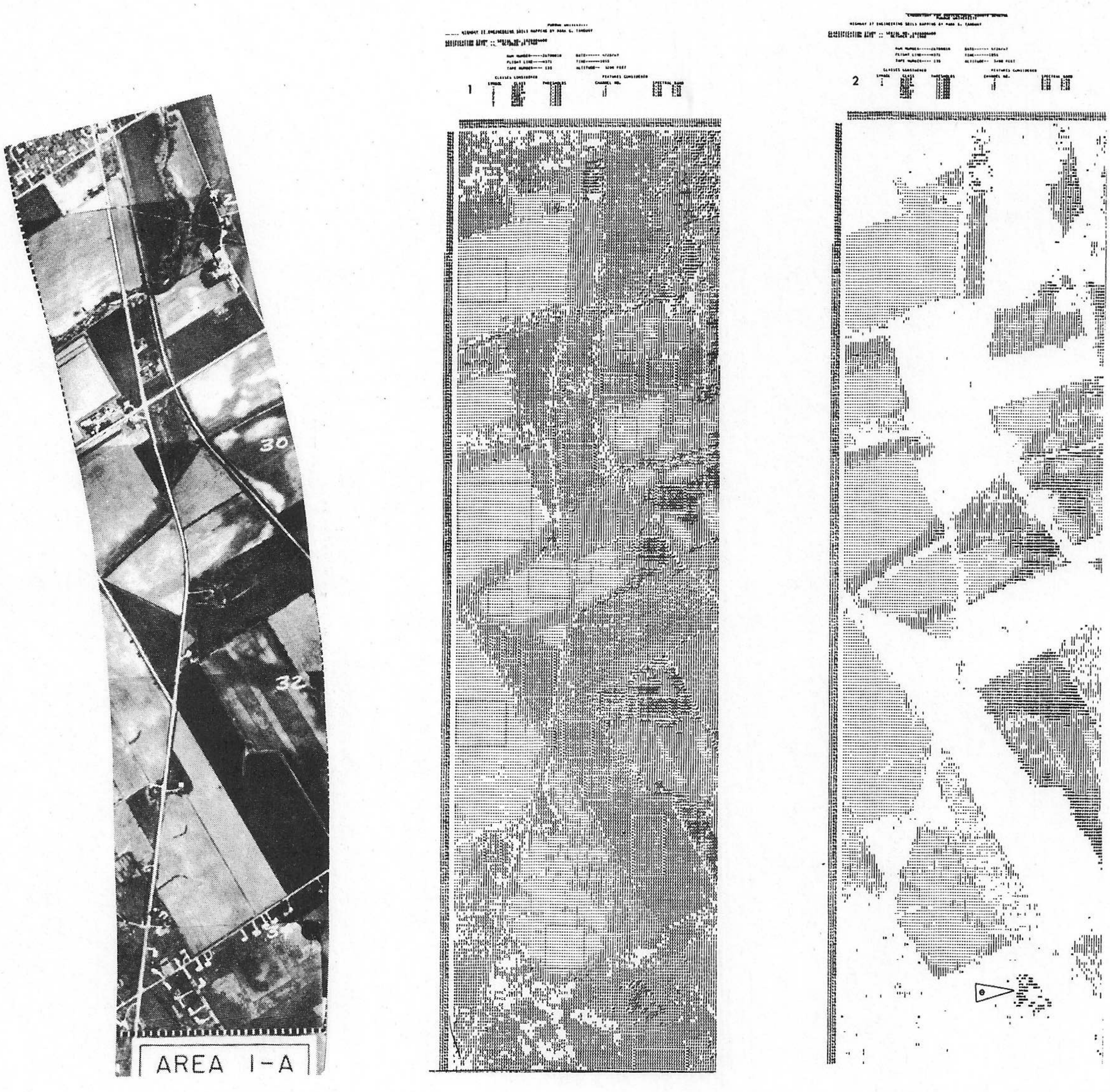


Figure 23-15.- Color aerial photograph and an artificially enhanced computer printout in Southern Indiana. Note how the computer identifies areas of standing water (red arrow), and limestone outcrops (yellow arrows).

Example Classification Results
from an Engineering Soils Study



LARS/Purdue

Figure 23-16.- Geologists have been successful in using multispectral techniques in defining and classifying geologically interesting features as shown by the comparison of the aerial photograph and artificially enhanced computer printout of a portion of Yellowstone National Park.