LARS Technical Report 080879

An Investigation of Analysis Techniques of Landsat MSS Data Designed to Aid the Soil Survey

Sue A. Kaminsky
Richard A. Weismiller
Byron O. Blair

Laboratory for Applications of Remote Sensing
Purdue University West Lafayette, Indiana 47906 USA
1979

AN INVESTIGATION OF ANALYSIS TECHNIQUES OF LANDSAT MSS DATA DESIGNED TO AID THE SOIL SURVEY*

Sue A. Kaminsky
Richard A. Weismiller
Byron O. Blair

August 1979

*The work described in this report was sponsored by NASA under Grant No. NGL-15-005-186.

TABLE OF CONTENTS

Page

LIST OF FIGURES

AN INVESTIGATION OF ANALYSIS TECHNIQUES OF LANDSAT MSS DATA DESIGNED TO AID THE SOIL SURVEY

ABSTRACT

In the 1930's initial use of aerial photography did much to improve the accuracy and speed of the soil survey effort. Since the 1940's laboratory and field research has provided significant information concerning the interrelationships of soil spectral reflectance characteristics and soil properties. Technological advances during this period contributed to the ability to use computer-implemented analysis of spectral data obtained from multispectral scanners.

With the launch of the Landsat earth resource satellites, investigations have determined analysis of scanner data to delineate organic matter differences, textures and drainage characteristics. A spectral classification of Jasper County, Indiana, was accomplished by computer-aided analysis of Landsat data which provided a basis for this research.

Four Landsat data point selection schemes were investigated to determine the best classification representation of the soils within the county. Of the four classifications, those considering parent material boundaries proved to be the most representative. Within these classifications drainage characteristics, soil erosion, and textures could easily be identified by spectral definitions.

A county spectral map depicting soil classes was the first of its kind, and may prove to be valuable in reducing man hours involved in conducting a county soil survey program. The classification will be printed in field sheet size printouts and accompany the aerial photograph in the soil mapping investigation. Delineation of soil drainage within parent material areas indicated the soil series present within a given area and provided information about areas not easily accessible or covered with dense crop vegetation. The use of these analysis techniques may contribute to decreasing the time and expense of a county soil survey.

INTRODUCTION

For decades soil surveys have been conducted by field investigation of soils with the aid, in recent years, of aerial photographs. Although aerial photography improved the accuracy of the survey, there was difficulty in measuring that accuracy because of subjectivity. A more objective and quantifiable approach to soil studies would benefit the county soil survey.

The first priority of the soil survey is delineation of soil boundaries into as uniform or pure groupings as possible; but as the size of average map units decreases, the cost of surveying increases (9). The surveyor is then left with the need to maximize soil map unit homogeneity and minimize the area physically covered. The problem is especially true in Indiana where an order two survey is conducted with map units of .7 to 4.5 ha delineated. Within these map units are inclusions that on first mapping are overlooked or because of time and expense are generally ignored. It has been reported that an economically feasible level of mapping leaves 30% to 40% soil inclusions within map units that is much higher than the 15% normally accepted (33).

If Landsat or some type of remotely sensed data could augment the existing use of aerial photography and field investigations, a more accurate, quantifiable and less expensive survey may result. Preliminary to this study, Landsat data analysis has yielded fruitful results in quantification, extension and location of soil map units as well as identification of such soil parameters as drainage characteristics, organic matter differences, textural differences and differing cultural practices (22,23, 42,43). Image interpretation of Landsat data has also aided in compilation of soil association maps and parent material boundaries (51). More recent studies involving Landsat digital data in association with ancillary data in the form of topographic boundaries have shown increased interpretive capabilities (49).

This study was initiated to investigate analysis techniques used for developing spectral soil information for a county soil survey. A hypothesized step process that employs Landsat imagery and digital analysis was purported to augment the conventional soil survey method. A methodology is proposed to blend conventional investigations and remotely sensed data analysis to obtain a more efficient method of conducting a soil survey.

Sampling techniques of Landsat data points were investigated to determine the satisfactory means of selecting points to be used in establishing distributions from which statistical probabilities would be calculated and spectral responses representative of *in situ* soils be classified. Parent material boundaries were identified to determine if the addition of ancillary data would provide a more accurate spectral classification. Finally, research within Jasper County was initiated for definition of the extent to which soil taxa could be interpreted through spectral evaluation.

It is the intent of this work not only to support and build on previous successful research but also to develop a methodology that could be readily used in a soil surveying process.

LITERATURE REVIEW

One of the earliest attempts at remote sensing was advocated by Arago, Director of the Paris Observatory, who encouraged the use of photography for topographic identification in 1840. Remotely sensed data are acquired from the earth's surface by measuring instruments not in contact with that surface; therefore, Arago was recommending remote sensing in the form of chemically processed photographic images. Practical application of Arago's idea was not feasible until a navigable platform was available, which came several years later in the form of the airplane. Aerial photography became not only a feasible but more accurate method of broad scale surveying (34).

Remotely sensed data, in the form of black and white aerial photography, were first employed as a base map for the soil survey in Jennings County, Indiana in 1929 (11). The poor accuracy and time involved in the use of plane tables to draw base and soil maps prompted surveyors to turn quickly to the aerial photograph for base maps. By 1938 most U.S. surveys were conducted with the black and white photograph as a base map (39). This advancement led to the present day capability of mapping approximately 20 million hectares annually at a cost of one billion dollars (16).

Military investigations of the 1930's produced color, color infrared, and false color photographic products for characterizing landscape scenes. It seemed that false color, using colors different than the actual scene, tended to enhance differences in scene objects (30). Civilian research in these areas did not begin until after World War II when investigators began adding to this military research by comparing information between black and white photos and other photographic products. From this work it was reported that soil boundaries could be more accurately differentiated on color than from black and white photography (14). Comparison of color, color infrared and black and white photography in distinguishing such parameters as soil series, land types, drainage and organic matter showed no statistically significant differences although color and color infrared were slightly more contrasting than black and white (29). Overall research at this time indicated that there was potential in using various products (12). Even though much research had been conducted with regard to photographic techniques, the aerial black and white panchromatic photograph remained the major soil mapping aid largely due to prohibitive costs of other products.

Success at discerning scene objects with aerial photography prompted analysts to investigate more sophisticated remote sensing techniques involving digitized photography, optical-mechanical scanners, and multi-images. Both multiband and multi-emulsion photography were used in initial attempts to further the applications of remote sensing. Quantization of the tonal variations of the three emulsion layers in color photography was accomplished by measurement of film densities in the three layers. Multiband images in which various filters were used in repeatedly photographing the same scene were also used to obtain unique wavelengths of data. These methods were analyzed by taking measurements with a scanning microdensitometer that measured adjacent lines in a sequence over the entire photograph. Multiband photographs were scanned as unique spectral bands that could later be analyzed as an entire set of spectral imagery (3,20).

Microdensitometer scans revealed crop spectral response to be a combination of vegetative cover and soil background reflectance. In general, soils were characterized by a high spectral response in the thermal infrared, a rather low response in the reflective infrared, and a varied response in the visible portions of the spectrum. As knowledge about the various information available across the spectrum unfolded, it was decided that scene identification was much easier where combinations of spectral wavelengths were taken.

Digitization of photography collected every two weeks at 18,000 meters (resolution ~ 1 meter) was used in a southern corn blight survey (5). When compared to conventional photo interpretive methods, greater accuracy was obtained from the digitized photography but not significantly so. Tonal variations caused by vignetting and inaccuracies in density analysis caused researchers to investigate remote sensing in forms other than chemical photography. Thirteen spectral band multispectral scanner data proved to be the best of the three methods although all bands were not needed to provide an accurate classification. The best three spectral channels generally gave better results than any two channels and in most cases four channels.

Computer technological capabilities developed in the 1960's enabled analysts to overlay optical-mechanical multi-aperture images. These 12 or 13 channel scanners were former military instruments that collected data through two apertures. Overlaying these data provided a more accurate interpretation of earth surface features than was previously afforded. By 1965 the 12-channel scanner was improved by devising a single aperture scanner capable of collecting data from .32 - 13.5 μm with internal calibration. Reflectance measurements were recorded in analog form on magnetic tape. Although the capabilities devised for overlaying multiple aperture images were no longer necessary, these same methods could also be used to overlay and analyze multidate and multi-image data.

In the late 1960's numerical statistical analysis was facilitated by computer techniques to apply 12-channel scanner data to soil investigations. Alfisols and Mollisols were found to be spectrally separable and highly correlated to field data despite shadowing and decreased reflectances of surface soils resulting from cultural practices (43). Organic matter was shown to have high correlation to spectral reflectance (20). No single array of spectral channels was found to do the most accurate identification of organic matter, but selection and number of channels did affect accuracy.

Other soil parameters such as texture, color, moisture relationships and soil type were distinguishable using numerical analysis of aircraft scanner data (7,26). Gross variations in soil features could be quickly accomplished, but extension of statistical data from one area to another did not produce accurate results (23). Some of the difficulty was due to soil variability in the subsurface or underlying horizon of the soil. Since soil series are identified by surface and subsurface properties, it was thought that mapping soil series was not feasible.

While investigations of multispectral scanner data acquired from aircraft platforms were being conducted, the interrelationships between the spectral responses of soils and their physical and chemical properties were being studied utilizing laboratory instrumentation. Results indicated such parameters as moisture, grain size, clays, and soil types could be identified with dependable accuracy (35,48). Soil particle types were identified by the relationship to their size; it was found that the larger the particle size the lower the reflectance or the less reflected light (6).

These parameters were investigated across the spectrum with instruments such as spectrophotometers that sensed ultraviolet (.25-.39 μm), visible (.39-.76 μm), and infrared (.76-5.0 μm). Of these wavelengths the infrared proved optimal in recognizing minerals of carbonates, sulfates, and silicates (21). Specifically related soil parameters were measured using a spectroradiometer which was reported to have found high correlations to spectral response with moisture, silt, clay, iron and organic matter (8). Moisture and clay content contributed to overall soil reflectance while silt correlated with reflectance in the .8-2.5 μm range (6,26).

The ability to specify these properties led to speculation about the optimal range that should be used in soil investigation. The .8-14 μm range was suggested because it contained the fundamental Si-O vibration. The red portion (.6-.7 μm) was also advocated because it contained the range which was reported to show a peak in the soil spectra (26). Much discussion arose at this time about recommendations for the type and range of sensors that should be carried on future earth resource satellites.

Varying molecular and macromolecular composition provides the ability to differentiate minerals and soils spectrally. Under controlled environmental conditions like those present in laboratory analysis, unique spectral responses of varying objects enabled analysts to identify soils, rock types or minerals.

Discrepancies in spectral responses of laboratory and environmental measurements occurred due to differences in atmospheric conditions. Within the electromagnetic spectrum energy moving at a constant velocity is measured. Spectra are classified according to wavelength of radiance and frequency of wavelength (12). When radiance travels through a heterogeneous medium, interference and scattering of that straight path of radiance occur.

Atmospheric interference with spectral response is twofold in that solar energy traveling through the atmosphere encounters ozone and oxygen absorption. Emitted radiance must travel back through the same disruptions it encountered upon entering the earth's atmosphere. In reality it becomes virtually impossible to obtain a constant or unchanging spectral response from an object in the environment. Presence of water adds further confusion by interrupting radiation throughout the entire spectrum. Specifically, water absorbs light in the infrared, scatters in the visible and scatters in the ultraviolet. Reflectance from an object in the natural environment is not only affected by the interfering atmosphere but also by distance from the target and sun intensity. In addition O and CO_2 reduce incoming radiation. Not only outside forces but also variations within the object affect the spectral response (17). By statistically evaluating spectral data comparison of unique spectral responses could aid in accounting for atmospheric and scanner variability.

The onset of the space programs provided a synoptic view of earth surface features that was not available with aircraft data. Multiband and multi-emulsion imagery were first taken from the Apollo spacecraft as it passed over Imperial Valley, California. Digitized and computer-analyzed, the photography discriminated soil textures such as separations of clays and loams (2).

The launching of Landsat-1 (ERTS) satellite in July 1972 provided the first orbiter specifically designed to monitor earth resources. An array of six detectors per spectral band (24 detectors per 4 channels) simultaneously sensed radiation from .5 to 1.1 μm over a 185 km swath. Detector output was encoded to six bits per second. The continuous data gathering was processed in frames of data representative of 33,000 km^2. Representative of .45 hectares, each pixel was a rectangular resolution element. Each Landsat frame contained approximately 7.5 million of these pixel elements per spectral wavelength band or 30 million unique responses across the entire frame. Before telemetering to receiving stations in Brazil, Canada, Italy and the USA, data are digitized on board the satellite. This information is then made available through EROS Data Center, Sioux Falls, South Dakota. Two formats can be acquired from the Center, i.e., imagery (color composite or black and white single band) and computer compatible tapes from which photographic products and/or computer analysis can be accomplished. MSS data can then be analyzed numerically or analyzed by conventional photographic interpretations.

Advantages of Landsat over previous data gathering techniques are as follows:

1) The scanner, in a near polar orbit, provides sun synchronous data collected at mid-morning to eliminate shadowing effects;

2) Since the Landsat satellite orbits the earth every 18 days and completes an earth orbit every 103 minutes, it is feasible to inventory earth features repeatedly;

3) A synoptic view of the earth's surface is provided.

Delineating soil associations, land use changes, slope and drainage patterns was possible through the synoptic view of Landsat (51). Simulated spacecraft imagery analyzed before the launch suggested that accurate definition of soil associations would be possible in grassland areas or areas with slight tree cover (36). Preliminary investigations such as these hypothesized MSS data could be used to map soil at the family level of the taxonomic classification. Detailed soil surveys appeared to be unattainable from space altitudes because of inadequate resolution and primitive interpretive techniques (47).

Information gleaned from the scanner revealed much more information than was hypothesized. Sensing portions of spectra other than visible provided an analysis tool that was previously not available on large scale mapping. Identification of water, soil and vegetation was possible by comparing responses across four spectral bands. Vegetation, for example, displays a low response in channel 2 (.6-.7 μm) due to absorption of light by chlorophyl and is highly responsive in channel 3 (.7-.8 μm) because of the reflective properties of mesophyl located in plant tissue. Whereas vegetation varies across the spectrum, soil responds rather evenly across the four data acquisition channels and water is distinguished by high response in the visible channels (.5-.6 μm, .6-.7 μm) and low responses in near infrared (.7-.8 μm, .8-1.1 μm). These general curves can be used for class identification in scene analysis. The quality and intensity of response change with soil, topography, season and chemistry and physiology of green vegetation. Soil parameters and vegetative response can be identified within a given scan by comparing response, but correlation to different scenes is not possible because of scanner calibration, atmospheric conditions, and changes in environment.

Using only visual image interpretation of simulated infrared and individual black and white band imagery, soil association maps of single counties and entire states have been created (24,41). For example, a South Dakota soil association map was produced in approximately five weeks at a cost of $.02/ha (51).

Broad scale inventories in international research have provided information for specific country problems. For example, research in India has found that salt affected soils along the Ganges River Plain can successfully be separated from nonaffected soils (37). A land use inventory produced favorable results in the Bangkok area (40). Resource mapping was also accomplished in Poland using data analysis techniques (10). Landsat data have met the needs of many developing countries which have little information or technical means of conducting a broad scale mapping of their country.

Work at the Laboratory for Applications of Remote Sensing specifically related to soils of the United States was stimulated by the discovery of a narrow strip of prairie soils running east and west for approximately 64 km in a predominantly Alfisol (timber) area across north central Indiana. This strip of Mollic soils is believed to have formed in heavy textured, poorly drained glacial debris which filled a preglacial tributary of the Teays River system. Spectral analysis of the Mollisols within the strip and surrounding areas showed the soils to have a unique reflectance when compared to the predominantly timber soils (31).

An in-depth study of spectral responses in Clinton County, Indiana, found drainage characteristics to be readily identified through Landsat analysis. Identifying drainage characteristics made it possible to associate previously investigated soil map units with spectral data which could then be quantified (due to constant resolution element size). Inclusions and location of various soils within a map unit could also be noted which could aid in map unit evaluation and verification of the need for soil complexes to be established (22). Prior to this time conventional soil maps were generally used for checking the accuracy of Landsat classifications. This research revealed the reason for previous discrepancies was due to the quantitative nature of the satellite resolution element and the subjective nature of conventional mapping techniques.

Considering the previous work with MSS data it was apparent that satellite data alone could not provide an adequate tool for land surveying. Since Landsat MSS data portray only surficial reflectance properties, widely varying soils with respect to horizonation and parent material may exhibit the same spectral properties. Ancillary data in the form of physiographic boundaries provided added information which contributed to an enhanced evaluation of the county and allowed for correlation of general soil series and spectral soil classes (48).

Previous investigations in spectral properties of soils have provided a basis from which to do further research. Many soil parameters have been identified through previous research techniques, but more advanced capabilities in the past few years have provided techniques that may add in further characterization of these parameters.

METHODS AND MATERIALS

County soil surveys, in general, follow a stepwise plan for surveying that begins at the decision to map within a county which leads to the final published soil survey. A hypothesized methodology is proposed that would augment the soil survey by introducing remotely sensed data that would aid in increasing performance and shortening duration of a survey. Figure 1 lists the conventional approach while Figure 2 is representative of a methodology that is hypothesized to augment the traditional soil survey.

Selection of Data for Investigation

County

Jasper County, located in the nortwest corner of Indiana was chosen to investigate a methodology for using Landsat multispectral scanner data in relation to soil survey procedures. The county is to initiate a soil survey in the near future which made it a viable candidate for research, and close proximity to Purdue made Jasper County readily accessible for repeated field observations relating to compilation of an ancillary data base and evaluative and correlative examinations.

Remotely Sensed Data

The remotely sensed data were of two forms, i.e., aerial photography and Landsat data. May 1976 aerial coverage of Jasper County, Indiana was taken at an altitude of 2000 m creating an approximate map scale of 1:15840. The Landsat digital data were obtained in the form of a computer compatible tape from EROS Data Center, Sioux Falls, South Dakota. The Landsat data were collected June 9, 1973 at 10:00 a.m. at an altitude of 1,087,300 meters. These data were relatively free of vegetative canopy, snow cover, interfering clouds and fog, and scanner distortions.

Specifically, scanner distortion may cause striping if detectors and associated electronics are not correctly calibrated or if any detectors are malfunctioning. If one or more of the six detectors per band are malfunctioning, striping occurs in increments of six lines. Single bad data lines can be caused by satellite tilt or malfunction of equipment.

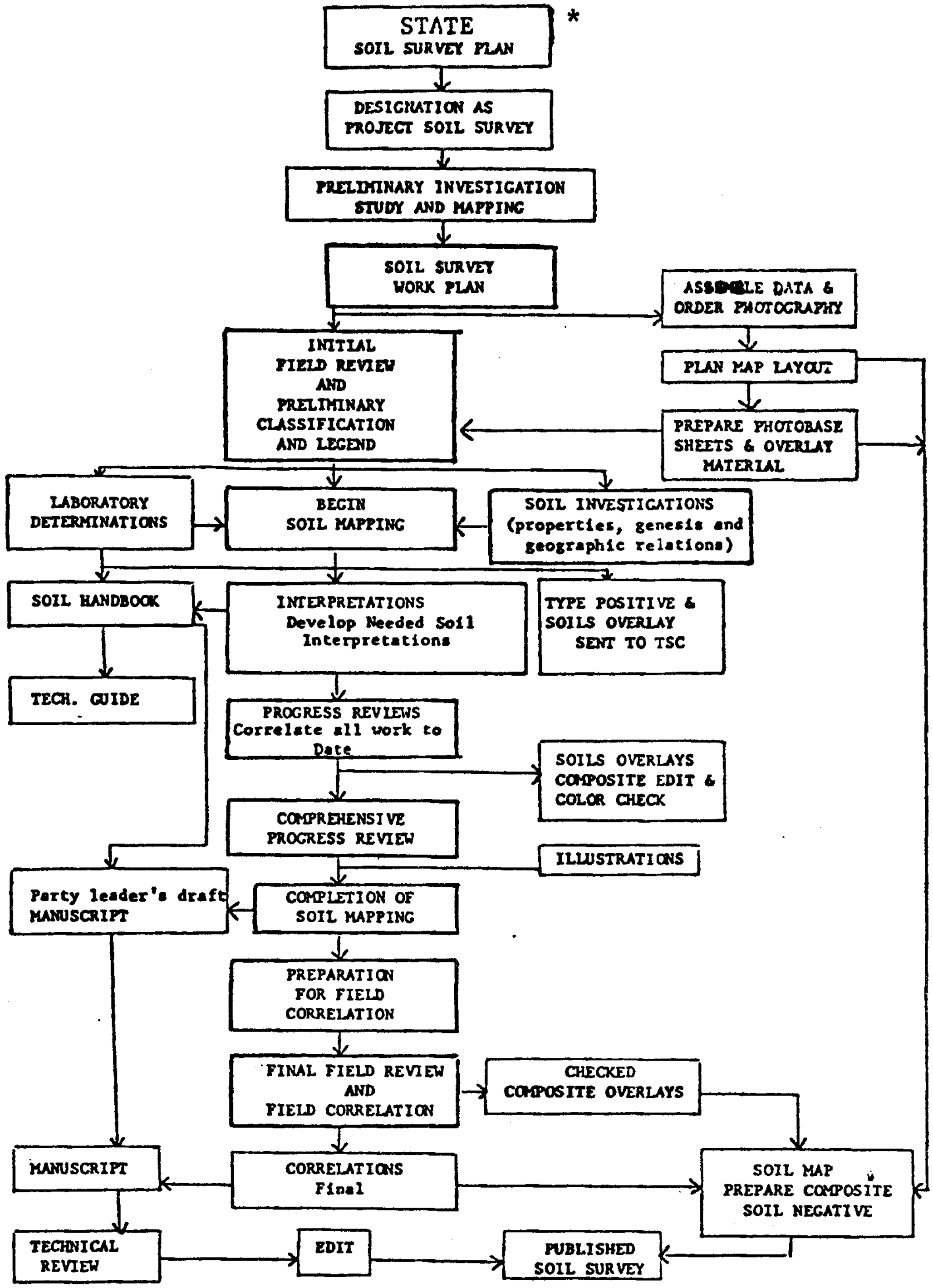


Figure 1. Conventional soil survey procedure.

*Taken from National Soils Handbook, Soil Conservation Service (Exhibit 205.3(a))

1) Selection of data for investigation
   A. County
   B. Remotely sensed data

2) Data preparation
   A. Base map
   B. Geometric registration and rectification of remotely sensed data

3) Imagery analysis
   A. Histogramming remotely sensed data
   B. Assignment of colors to density slices and compilation of map image

4) Stratification of county
   A. Geological history
   B. Stratification of parent material

5) Digital analysis of remotely sensed data
   A. Data sampling and analysis techniques
   B. Spectral classification

6) Evaluation
   A. Field observation
   B. Correlation to soil map units at randomly selected sites

7) Refinement of classification

8) Creation of map products
   A. County map
   B. Detailed soils map
   C. Parent materials map
   D. Vegetation map
   E. Erosion map
   F. Drainage map
   G. General soils map
   H. Non-agricultural interpretations
   I. Variable scales

Figure 2. Proposed county analysis procedures.

Data Analysis

Base Map

A base map consisting of a block of 85 black and white panchromatic aerial photographs was used in the registration of the Jasper County Landsat data. Known north-south roads were located on aerial photographs and used to parallel a y coordinate axis. A three-parameter linear transformation was used to block the photos to a common coordinate system which took the form:

$$\begin{vmatrix} x^1 \\ y^1 \end{vmatrix} = \begin{vmatrix} \cos\theta\sin\theta \\ -\sin\theta\cos\theta \end{vmatrix} \begin{vmatrix} x \\ y \end{vmatrix} + \begin{vmatrix} \Delta x \\ \Delta y \end{vmatrix}$$

where x^1, y^1 are transformed coordinates;
θ is the rotation angle between the coordinate system;
x, y are the coordinates to be transformed; and
$\Delta x, \Delta y$ is a linear shift in transformed coordinate space.

Halftone acetate positives were created from the photographs. The resulting transparencies were rectified and trimmed to field sheet size. Rectifying corrected for tilt, vertical aspect, which improved scale variations and crabbing (rotation). Crabbing and scale variation are caused by changes in altitude, mechanical distortions (photographic films and lenses), and rotation of the plane while in flight. The halftone acetate positives were used for comparing field soil patterns with the spectral classification of soils.

Geometric Registration and Rectification

The blocked set of aerial photographs was used as a base in the geometric registration and rectification of the Landsat data. Corresponding points between the two images (Landsat and photo block) were located by either displaying the Landsat image on a CRT screen or by cluster analysis of the digital data (described later). Groups of approximately 100 data points were clustered and specific points within the clustered areas were located on the aerial photos. A twelve parameter equation transformed the coordinates between the images by:

$$Ax,y + \Delta x,y = Bx,y$$

where A is the reference image;
B is the overlaid image; and
Δ is the transformation function.

The biquadratic function, Δ, is of the form:

$$\Delta x = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy$$

$$\Delta y = b_0 + b_1 x + b_2 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 xy$$

A least squares approach was used to solve the coefficients,

$$\underline{\alpha} = (p^T p)^{-1} p^T \delta x$$

$$\underline{\beta} = (p^T p)^{-1} p^T \delta x$$

where

$\underline{\alpha}$ and $\underline{\beta}$ are 6 x 1 coefficient vectors for Δx and Δy

p is the matrix p_{ij} of powers x_i and y.

For each checkpoint $p_{ij} = x_i^k y_i^l$ where i is the number of the checkpoint i = 1; k = 0,1,0,2,0,1; l = 0,0,1,0,2,1; for j = 1,2,3,4,5,6, respectively.

$\delta x, \delta y$ = N x 1 column vector A and B coordinates

$$\delta xi = XB_i = XA_i, \ \delta y_i = YB_i - YA_i$$

Compatible scales between the base map and the Landsat data were accomplished by expanding the Landsat data to a scale of 1:15840, the mapping scale for Jasper County. When a scale is expanded, holes or blanks appear in the data because there are not sufficient data values to show a one to one correspondence with, for example, the aerial photographs. The data values, in other words, are not of sufficient number to cover the entire area represented. Extra data points must then be created to fill the data gaps.

Expanding or rescaling of data can generally be achieved by one of three methods. Simply duplicating data lines to fit a new scale is inexpensive yet inadequate if numerical analysis will be attempted. The duplicated pixels are weighted which result in overemphasizing some data point distributions. Previous registration in Clinton County, Indiana, for example, was done by simply duplicating lines and columns which resulted in large unacceptable groupings of like data points.

A bilinear or nearest neighbor interpolation duplicates points nearest a desired sample location (1). Although data values are not altered by this approach, statistical distributions are again weighted by duplication of points. These map products are blocky in appearance because of the duplicated points. Map errors at scales of 1:24,000 were approximately plus or minus 16 meters of tolerance.

For this registration a cubic convolution resampling algorithm was used to rescale the Landsat image to 1:15840. Intermediate data values were calculated using a Lagragian third order equation that used a 4 x 4 matrix or 16 spectral points. On the curve of this equation intermediate data values were plotted and used in expanding the scale. This method had the effect of smoothing the image which could contribute to a somewhat less accurate classification but provide a higher quality map image for soil mapping. Classes that are very close spectrally could lose their distinctness because of these calculated intermediate values.

Errors are evident in all the scale expanding methods so one must choose the parameters that should not be sacrificed. However, the problem of rescaling data has now been eliminated through the use of an electrostatic magnetic dot matrix plotter. This plotter provides a means of adjusting pixel size elements on a variable matrix scale adjustment. A girded matrix containing various black and white designators can be compiled to print various textural or grey scale patterns. Instead of correcting data prior to classification, statistical analysis of original Landsat data can now be done. The finished classification can then be printed at any number of scale values.

Imagery Analysis

Creating a False Color Image

The false color composite image of three channels (.5-.6, .6-.7, and .8-1.1 μm) was categorized by spectral response into 32 prespecified levels of data (Figure 3). The data were assigned levels by a histogramming algorithm that creates a 100 bin histogram for data from each specified channel. Each bin has an equal width based on the entire range and distribution of data. An initial bin size and lower limit are first assumed. A bin number is calculated as

$$\text{bin number} = \frac{\text{data value - histogram lower limit}}{\text{bin size}}$$

A bin number falling in the range 1-100 is placed in the bin; otherwise the bin is adjusted to fit the new data range while keeping the individual bin size constant (32).

Compilation of Color Map Image

Three channels were combined to create a color composite map, generally referred to as a false color image, at a scale of 1:180,000. This image was generated to aid in stratification of the county spectra. Enlarging the image to 1:120,000 enabled the user to delineate or interpret finer detail.

Stratification of the County

Geologic History

The geology of Jasper County is quite complex. Underlying the county are tertiary and quarternary bedrock valleys formed primarily by water

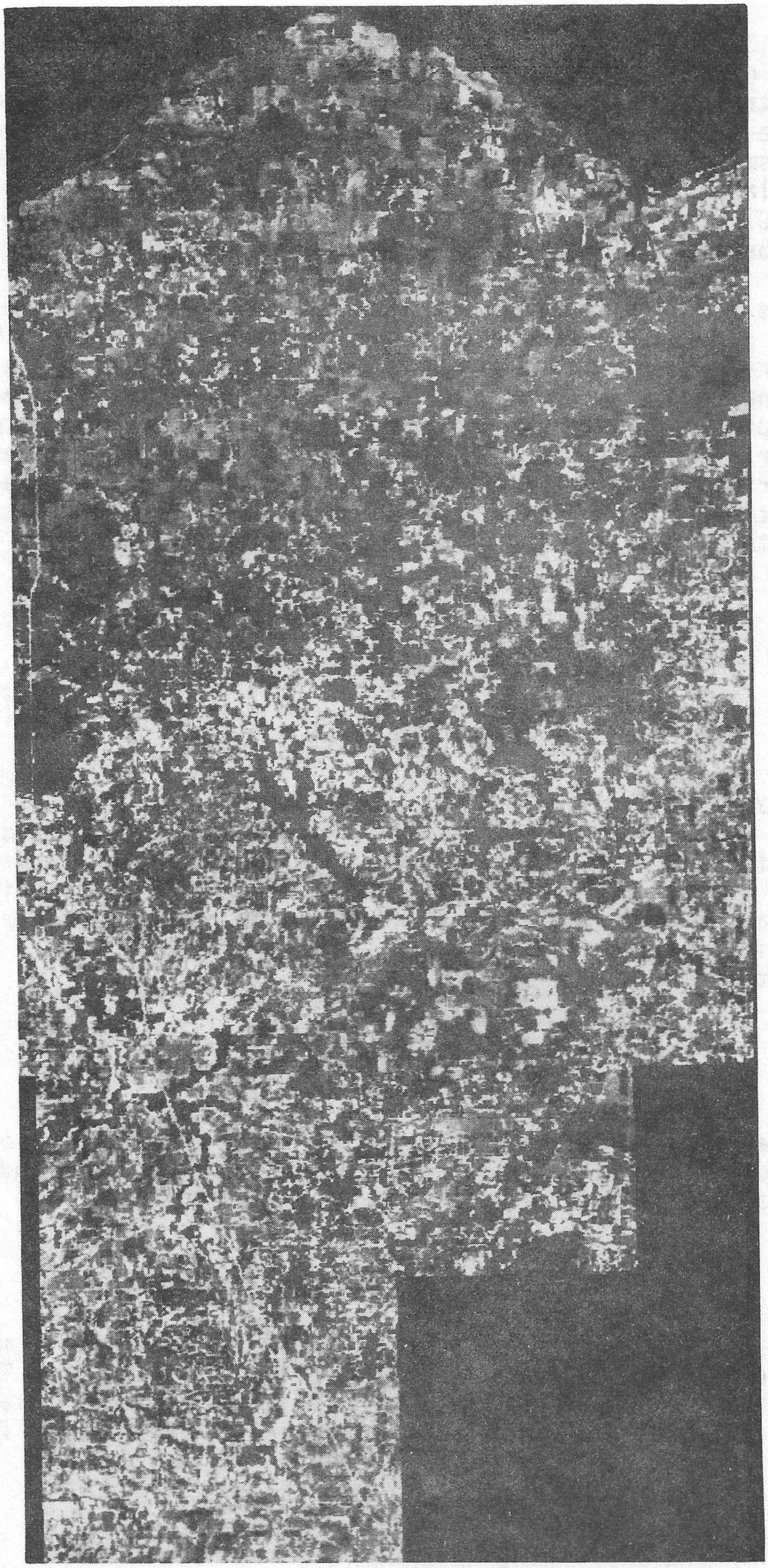


Figure 3. False color image of Jasper County, Indiana.

erosion. These valleys, initially filled by quarternary debris, were later covered by the early Kansan and Illinoian glacial deposits.

Evidence of much earlier geologic phenomena occurs to the west where coral reef domes reach within one to two meters of the surface. The reefs are thought to be a product of the Silurean or Devonian ages and are a good source of limestone. Material that accumulated to the side of the domes, most likely water deposited, is generally not of good quality which is why the smaller domes have been largely left untouched by limestone excavation.

Glacial deposits that covered all of Jasper County from the Kansan and Illinoian were obliterated by a coalesced ice sheet from the Lake Michigan and Erie glaciers of the early Wisconsin age. A thin ice extending from the Saginaw northeastern lobe covered the previous glacial activity and appears to have truncated the Marseilles moraine in the eastern portions of the county resulting in belts of kettles and intervening dunes covered by submorainic rises. Characterized by the thin ice sheet the Saginaw lobe covered low lying areas, but largely left higher elevations untouched. Present surficial deposits in the lower areas are credited to this glacier. The retreating glacier also left melt water laden with silts and clays which, when the water eventually subsided, left these lacustrine deposits.

Outwash sands were blown into parabolic and longitudinal dunes across the northern part of the county. Located under these dunes are peat areas that suggest vegetation once grew in ice block depressions left by the glaciers before being covered by (aeolean) sands. Vegetation establishing itself on the dunes gradually caused them to stabilize. After glacial activity subsided, geologic changes within the county have been in the form of drifting outwash sands and the accumulation of peat and marl in low lying areas (38).

This complex geology was considered in the compilation of a parent materials map of Jasper County. With the aid of Landsat data the area was investigated and parent material boundaries were delineated.

Stratification of Parent Materials

Training statistics are created by sampling data points and calculating a mean and covariance matrix for each unique spectral range. This set of means and covariances was used to "train" a classifier by providing a data base for calculating probabilities of remaining data points belonging to certain distributions.

Prior work in Indiana revealed uniqueness of spectral classes to be lost as training statistics were combined over a large area such as a county. As spectral classes were combined, distributions became larger and closer together. To avoid this problem, in Jasper County, a parent material map was created so that training statistics could be generated and used in specific parent materials, thus eliminating the need to extend training statistics over broad areas. Parent material delineations also provided a means of separating spectrally similar but genetically different soil classes within Jasper County.

The false color image, county and township road maps and knowledge of the geological history of the area were used to create the parent materials map of Jasper County. Initially, spectral stratification was noted on the map image and investigated through field observations. A soil auger was used to probe the horizon to determine underlying parent materials and characterize the profile. The location of calcareous till was determined by application of acid, and Munsell color charts were used to identify color boundaries for Alfisols, Mollisols and drainage characteristics. Textural boundaries were made and refined as the investigation progressed.

The completed parent materials map was transferred to a Jasper County map at a scale of 1:126,720. An electronic plane table digitizer was used to record the mechanical movement across the boundaries of parent materials on the Jasper County map. Signals from the mechanical movements were relayed through a computer hardware system that in turn punched cards with coordinate points corresponding to the boundaries. Utilizing a special software system the punched boundary coordinates were written on a separate computer magnetic tape, and these superimposed boundaries were added as a channel of information along with the Landsat precision registered data. By assigning unique values to spectral data in a parent material and recording that data in a channel, that information could be used to discriminate statistical distributions created in each parent material. For classification four channels of Landsat data rescaled to 1:15840 by a cubic convolution interpolation and precision registered to the 85 aerial photographs were used. For illustrative purposes four channels capable of generating a color composite map image were also available.

Digital Analysis of Remotely Sensed Data

Data Sampling and Analysis Techniques

Differentiating parent material boundaries made it possible to develop unique statistical distributions of the data within each delineation. Unique and subtle differences were hypothesized to be more distinct in parent material delineations than distributions developed across a whole county. Based on this hypothesis and the need to develop a better point sampling method, four techniques were devised. These techniques were designed to test the significance of parent material delineations within a statistical classification and to determine if differing sample point selections would change classification accuracies. A summary of these techniques is listed in Figure 4.

Two methods of sampling data points were used to determine which would most represent responses within the specified classification area. Subjective sampling of blocks of data was compared to systematically sampling points at specific line and column coordinates across the entire classification area. It was hypothesized that systematic sampling would more adequately represent the spectral variability of a scene rather than subjective sampling.

Another variability within the design was to limit the size of area classified. The importance of parent material delineations was tested by

	Data Point Selection	Clustering	Classification
Classification one	Subjective sampling of representative blocks of data within each parent material	Each block of data clustered requesting 13 cluster classes	18 spectral distributions used to train the Gaussian maximum likelihood classifier
Classification two	Systematic selection of data points from across the entire county (every eleventh line and column)	Clustering the entire county selecting data points every eleventh line and column. 18 cluster classes requested	18 spectral distributions used to train the Gaussian maximum likelihood classifier
Classification three	Subjective sampling used in classification one	Same cluster grouping used but groupings within parent materials kept as unique	Layered tree design used with Gaussian maximum likelihood classifier (60 classes)
Classification four	Systematic selection of data points of every fifth line and column within parent materials	Clustering within each parent material every fifth line and column. 13 classes requested per cluster	Layered tree design used with Gaussian maximum likelihood classifier (60 classes)

Figure 4. Data point selections and subsequent classifications.

classifying only within parent materials as opposed to classifying the entire county without regard to delineated boundaries.

A method was devised to evaluate the final spectral classes as to countywide performance and accuracy within specific parent materials. Compilation of soils at selected locations was hypothesized to be an adequate means of testing performance. Due to time limitations the number of test sites was limited. Quarter sections within the county were selected as test sites because they were easy to locate and randomly select. Quarter sections were numbered across the southern part of the county within each of three major parent material areas (outwash, lacustrine, till). Numbers were then randomly selected within each parent material area and corresponding quarter sections were noted on a Jasper County sections map. The 72-hectare quarter sections were then located on aerial photographs which were reproduced at 3 cm to 1 km to allow for mapping detail not generally mapped.

General Analysis Procedures

A general procedure was followed for each of the four analysis techniques. Initially, a clustering algorithm was used that established a group of spectral classes consisting of means and covariance matrices which through an interpretive process was used in statistically classifying the county. Figure 5 shows the process involved in creating training statistics for the area. A more detailed explanation of the algorithm related to the LARSYS processors follows (32).

Clustering Algorithm. This algorithm is based on distance relationships between each data point and the centers of groups of data points (4). It groups a set of vectors drawn from a spectral data tape into a number of classes that must be specified by the analyst. After the number of classes is specified, locations for these cluster centers are assigned in feature space as the cluster centers. The locations of initial centers are chosen by computing the mean vector and variance for the entire set of measurement vectors. Centers are uniformly spaced along the diagonal of a rectangular parallelepiped whose dimensions are based on the mean and square root of the variance of the data. Calculation of the Euclidean distance from each vector to each center makes it possible to assign the vectors to the clusters. The cluster centers are initially calculated by

$$C_{ij} = M_i + S_i \left[\frac{2(j-1)}{M-1} - 1 \right]$$

where M_i = mean;
S_i = square root of the variance;
i = measurements vectors; and
j = number of clusters requested.

By determination of Euclidean distances between vectors and cluster centers points can be assigned to the centers. The clustering process proceeds in a two step method. After samples are all assigned to centers, new centers are formed by calculation of the mean of assigned points. The two step iteration continues until there is no change of cluster centers from one

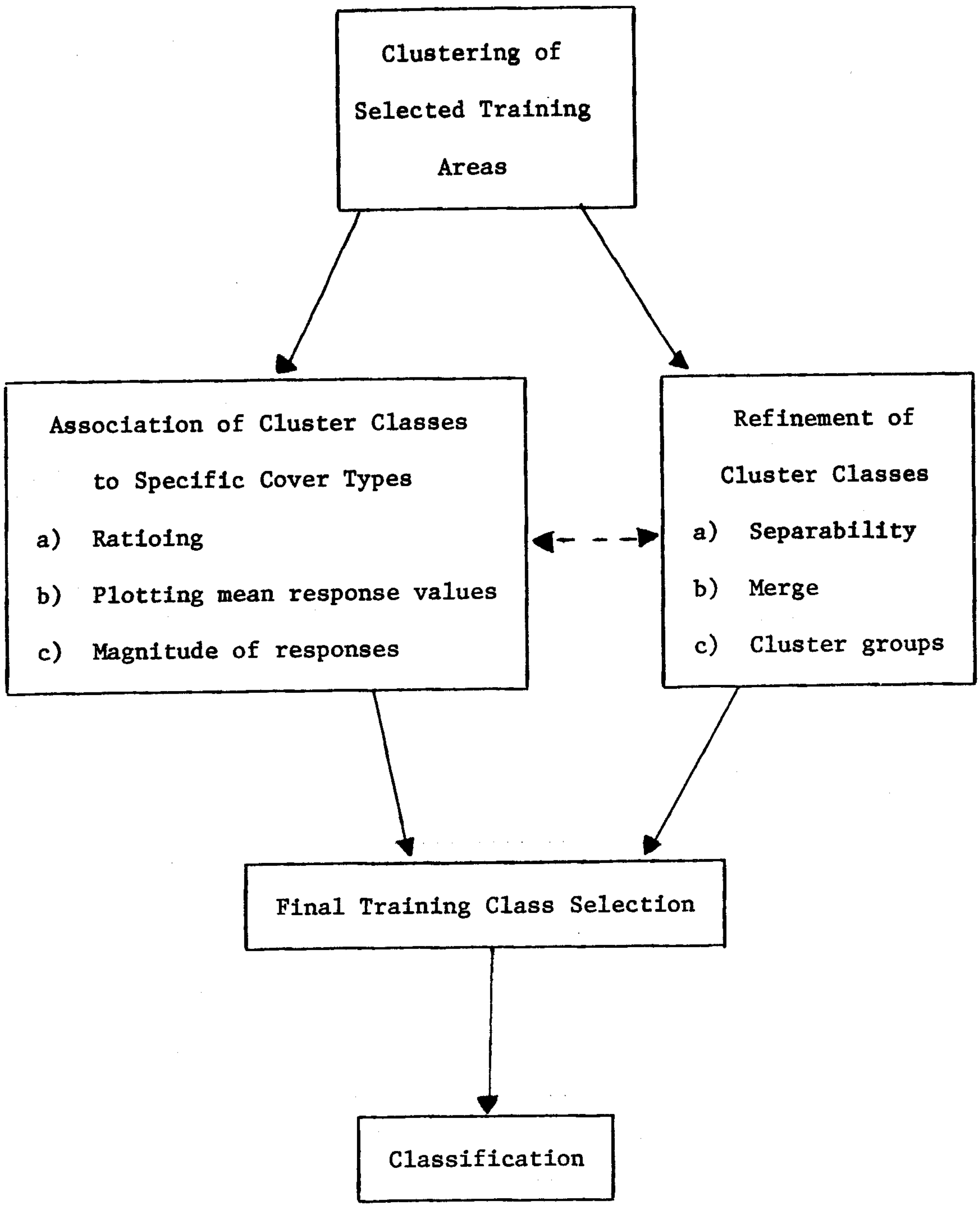


Figure 5. General analysis procedure implemented by LARSYS functions.

iteration to the next. Resulting from this program is a statistical grouping that contains mean and covariances for each channel of spectral data for each spectral class.

Association of Cluster Classes to Specific Cover Types.

Plotting Mean Response Values. Classes derived from clustering were evaluated as to their spectral properties. Identification of broad categories of vegetation, soil and water could be made by observing the mean relative responses across the four channels (Figure 6). Characteristic curves of soil, water, and vegetation make them easily identifiable.

Ratioing. Differences between vegetation and soil can also be detected by summing reflectances in the visible bands (.5-.6 μm and .6-.7 μm) and dividing by the sum of the two near infrared bands (.7-.8 μm and .8-1.1 μm). A high response in channel three and low response in channel two yield a ratio between channels of less than one that would indicate a vegetation response. Water is more responsive in the visible bands and therefore maintains ratio values over one. Response curves associated with soils generally follow an even pattern which displays values of one or more.

Magnitude of Response. When soils curves are identified, further separations between the soils can also be made by consideration of their relative magnitudes. The relative response across all four channels is summed and these magnitudes are compared in order to identify such soils as a high spectral response of a well drained soil or low responding poorly drained soil. Drainage classes and their differing responses are shown in Figure 6.

Refinement of Cluster Classes.

Merging Function. Clustering statistics developed from more than one clustering can be combined into a set of calculated means and covariances. A merging function takes all statistical classes requested and compiles classes with new calculated means and covariance matrices. Combined classes were measured for divergence and pairs of classes with low divergence values were merged into one spectral group. The processor used to merge classes together calculated a new mean of all points contained in the original classes merged and a resulting covariance matrix.

Some classes were encountered that had a spectral response representative of both soil and vegetation (Figure 6). These classes were combined with a vegetation if they were spectrally similar or left to represent a soil if the influence of vegetation was not too great.

Separability of Classes. Divergence of these cluster groupings was calculated to obtain a measure of the similarity between the classes. The measurement of similarity was defined as

$$D(ij) = \tfrac{1}{2}tr(K_i-K_j)(K_j^{-1}-K_i^{-1}) + \tfrac{1}{2}tr(K_i^{-1}+K_j^{-1})(\mu_i-\mu_j)(\mu_i-\mu_j)^t$$

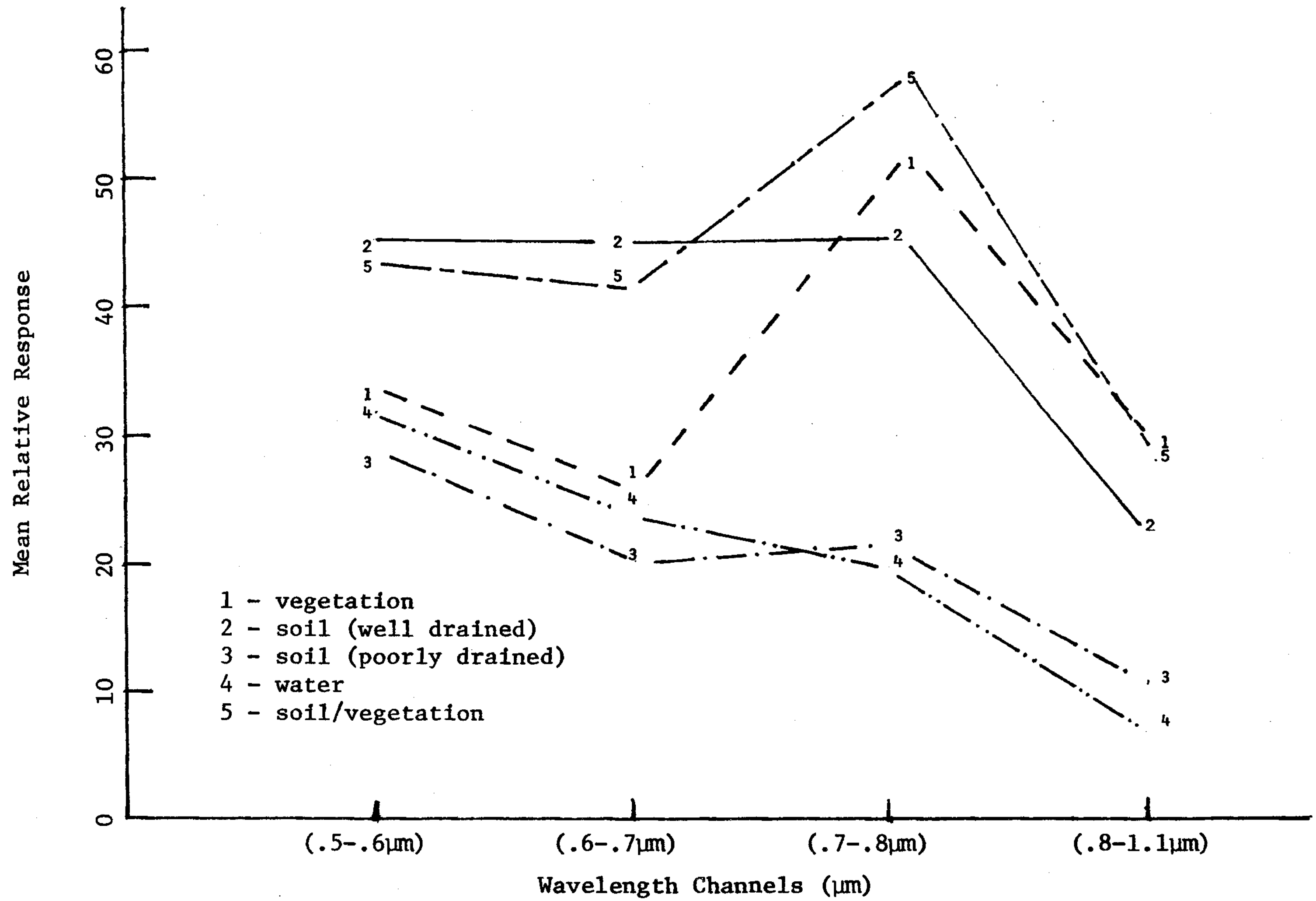


Figure 6. Representative responses of soil, water, and soil/vegetation spectral classes.

where

$K_i K_j$ = class covariance matrices;

$\mu_i \mu_j$ = class mean vectors;

tr = trace of matrix (sum of the elements in the diagonal); and

$(\ )^t$ = transpose of the matrix.

Divergence indicated the similarity of pair groupings. All possible combinations of classes provide information necessary for combining, remaining as distinct or eliminating classes (25).

The algorithm performs this divergence calculation that is then transformed.

$D_{(t)}$ = transformed divergence

$D_t = 2[1-\exp(-D/8)(i,j)]$ (45)

The transformed value has provided a better indication of spectral class separability than simple divergence.

Final Training Class Selection. Information obtained from transformed divergence measurements, ratioing, plotting, and notation of relative reflectance responses was used to define a statistical set of data points representing the area to be classified. Creation of statistical distributions most representative of the overall response is of crucial importance to correct classification. When spectral distributions are confused, the classifier will fail to separate accurately the data. Well defined sepaable distributions must be established if an accurate classification is the desired result. The classifier also assumes classes are normally distributed with mean and variance which further necessitates closely analyzing the final training statistics.

Classification. A Gaussian maximum likelihood classifier was used in all four analysis procedures. The classification algorithm

$$g_i(X) = \log p(\omega_i) - \frac{n}{2}\log(2\pi) - \tfrac{1}{2}\log|K_i| - \tfrac{1}{2}(X-M_i)^T K_i^{-1}(X-M)_i$$

where

$g_i(X)$ = a discriminant function (where X is an unknown data vector);

M_i = mean vector;

K_i = covariance matrix;

ω_i = class;

$p(\omega_i)$ = <u>a priori</u> probability of class ω_i; and

n = the number of channels

is used to classify X to class ω_i if

$$g_i(X) \geq g_j(X) \text{ for all } j \neq i$$

A priori probabilities may also be specified to be assigned to certain information classes. If a priori probabilities are not assigned, the log $p(\omega_i)$ of a certain class is set equal to zero which is equivalent to assigning equal probabilities to all classes.

If in the classification algorithm $y = (X-M_i)$ and K_{jk} is equal to a component of the matrix K_i^{-1}, the expression then becomes

$$\log p(\omega_i) - \frac{n}{2}\log 2\pi - \tfrac{1}{2}\log|K_i| - \tfrac{1}{2}[Y_1K_{11}Y_1 + Y_1K_{12}Y_2 +$$

$$\ldots Y_1K_{13}Y_3 + Y_2K_{22}Y_2 + Y_2K_{23}Y_3 + \ldots Y_3K_{31}Y_1 + Y_3K_{32}Y_2 + Y_3K_{33}Y_3 + \ldots]$$

Since $K_{21} = K_{12}$, the expression becomes

$$\log p(\omega_i) - \frac{n}{2}\log 2\pi - \tfrac{1}{2}\log|K_i| + [-\tfrac{1}{2}Y_1K_{11}Y_1$$

$$- Y_1K_{21}Y_1 - \tfrac{1}{2}Y_2K_{22}Y_2 - Y_3K_{31}Y_1 - Y_3K_{32}Y_2 - \tfrac{1}{2}Y_3K_{33}Y_3 \ldots]$$

where ω_i = a spectral class (17,28).

The discriminant value is calculated for each class and then compared to other discriminant calculations. The largest discriminant value is the class into which the data vector is classified. The results of the classification were written on a magnetic data tape which could then be accessed for displaying part or all of the area. The data tape can be read to produce the classification in the form of an alphanumeric map image, a grey scale image and/or tabular output.

The specific classification procedures deviated in the method of selecting training points for statistical analysis and in their application to the area. A description of the variations within each analysis procedure follows.

Classification One

Training sites consisting of approximately 1% of the data were chosen within each parent material with at least one training site located within each parent material. By previous field inspection and notation of transitions on the false color image training sites were selected that appeared most representative of the area. Training areas were located by a coordinate system of lines and columns which designated the appropriate data points within the county.

Eleven blocks were clustered within the outwash area; seven blocks were clustered in the rolling moraine till, seven in the outwash over till, four in the lacustrine, two in the till (Alfisol) and one in the till (Mollisol). Approximately 7,000 total points were clustered in all the areas. Ten to thirteen cluster classes were specified per block, depending on the apparent spectral variability within each area.

The resulting cluster classes were identified as vegetation, soil, water or some combination of cover types based on the previously described analysis procedures. Urban classes and related spectral responses were largely ignored because they were of minimal area in the county and were not of interest in soil characterization.

Spectral classes from all parent materials were merged together into one set of means and covariances. Through a process of merging and divergent measurements, a distinct set of spectral classes resulted. Ignoring the parent material delineations the classifier categorized each data point from the county into one spectral class developed from countywide sampling.

Classification Two

A systematic sampling of data points for compilation of training statistics characterized the second classification procedure. In systematic sampling an indication of all ranges of responses is sampled if the areas are significantly large or if the sampling increment is high enough.

A one percent sampling (eleventh line and column) approximated the size of the first sampling and produced a set of eighteen means and their associated covariances. Increments of six lines were avoided because of the possibility of error due to scanner noise, as previously described. Parent materials were not considered in the systematic sampling of data points nor in the resulting classification. Parent materials were disregarded to test if a significant increase in accuracy would occur when the areas were delineated in classification.

Classification Three

Spectral samples clustered in classification one were again used in classification three. Data points selected for training were combined only within parent material areas. These numbers of points varied with size of the area; therefore, a large area would be represented by a larger number of points. Similar spectral classes were combined if they represented the same cover type. Soil responses from other parent material areas in some cases were quite similar, but the property of the classifier made it possible to retain those classes as unique within the same classification algorithm (45).

By the use of a decision tree design each data point was not tested against all other data points in all other spectral classes but rather was tested against only those classes formed from spectral information within a particular parent material area. Sixty statistical classes were contained at the root node from which 6 stem nodes each representing a parent material projected. These nodes were equidistant from the root node; therefore, they constituted one layer within the decision tree design. Consisting of a set of spectral classes each node was used to discriminate which classes would be used within a designated parent material area. A Gaussian maximum likelihood rule was still used to classify points although the tree design was used to discriminate the number of classes used in each unique area (45).

One channel or a combination of channels could be used in the layered approach for either discriminating parent material or classifying data points. Of the 60 sets of means and covariances six classes consisted only of a fifth channel which was used as a designator of parent materials. These six classes were previously mentioned as the first layer in the classification scheme. Remaining classes of Landsat data contained in the root node were compared to each of the stem nodes. Each parent material designator specified a unique set of statistical classes to be used in classifying only that parent material. The process by which the classifier proceeded is shown in Figure 7. Stem nodes and their respective classes were prespecified in the classification program.

The ability of the classifier to use only 60 classes limited complete freedom in spectral definition. Soils were of primary importance in the investigation; thus it was decided to combine vegetation from all areas and classify using only three vegetation classes grouped from across the county. Again the classifier used approximately 1% of the data within the county to train the classifier.

Classification Four

Consideration of parent materials was integrated into the last analysis and classification. Irregular boundaries of the six parent material delineations made it extremely difficult to record all points manually within each area; therefore, a FORTRAN program was devised to locate every fifth line and column coordinate point within each parent material area. This sampling technique involved approximately a four percent sample of the county. An increment of five was used to insure adequate sampling and avoid six line noise. These line and column coordinates were used to cluster each entire parent material area. It was decided that thirteen cluster classes would be the maximum number asked for. A smaller number would not adequately represent the ground scene, and a larger number may leave some spectral classes with too few points to be considered a good statistical sampling. A separate set of means and covariances was generated for each parent material. Vegetation classes, soil, and scattered vegetation classes were identified by the same process described in the previous classifications. The four percent sampling was used in the layered design to produce a county spectral classification based on spectral probabilities from six different sets of statistical distributions. The design of the decision tree was identical to classification three except different spectral classes were used to compile the tree.

60 Class Statistics
(Classification 3,4)

Channel 5 Area 1 Outwash

Channel 1-4

Classification 3	Classification 4
1-9 soil classes	1-8 soil classes
3 vegetation classes	2 vegetation classes

Channel 5 Area 2 Rolling moraine till

Channel 1-4

Classification 3	Classification 4
1-10 soil classes	1-8 soil classes
3 vegetation classes	2 vegetation classes

Channel 5 Area 3 Outwash over till

Channel 1-4

Classification 3	Classification 4
1-8 soil classes	1-8 soil classes
3 vegetation classes	2 vegetation classes

Channel 5 Area 4 Lacustrine

Channel 1-4

Classification 3	Classification 4
1-9 soil classes	1-8 soil classes
3 vegetation classes	2 vegetation classes

Channel 5 Area 5 Alfisols (Till)

Channel 1-4

Classification 3	Classification 4
1-8 soil classes	1-11 soil classes
3 vegetation classes	2 vegetation classes

Channel 5 Area 6 Mollisols (Till)

Channel 1-4

Classification 3	Classification 4
1-10 soil classes	1-8 soil classes
3 vegetation classes	2 vegetation classes

Figure 7. Tree design used in classification procedure of Jasper County spectral soil maps.

Evaluation

Field Observation. Evaluation of classifications was accomplished by comparison of completed classifications to the mapped quarter sections. Three soil scientists comprised of one SCS soil scientist and two soil science graduate students mapped the quarter sections with the specific objective of mapping .45 ha delineations or larger. Normal mapping procedures were used to investigate the quarter sections. Each quarter section was located and position noted on the photograph. By traversing the land and taking sufficient borings to identify drainage patterns and textures, map units were delineated on the aerial photographs. Underlying calcareous till was identified by applying acid and observing if any reaction were present. The color chart was used to determine Mollic or Ochric opipedons. After investigation of surface and horizons, map units of .45 hectares or more were noted on the field sheets. Each quarter section was arbitrarily divided into three sections and mapped by one of the investigators. After the quarter section was traversed, soil characteristics were discussed and questionable areas were revisited. The final soil map was a combination of observations from all individuals. The northern part of the county was not chosen for evaluation because the distance was prohibitive in the investigation. Mapping of these quarter sections was done prior to computer analysis so bias in soil mapping could be avoided. The completed soil maps were used to evaluate the spectral classifications.

Correlation to Map Units at Randomly Selected Sites. An electrostatic dot matrix plotter was used to produce individual plots of the mapped quarter sections that would be used in the evaluative procedures. Copies of these classifications were also reproduced on acetate to enable overlaying on the photograph. All spectral classes were graphed as to their relative spectral response across the four Landsat spectral bands, and copies were provided for each analyst. Analysts were asked to compare each classification to the soil maps and rate the classifications as to their correspondence to the maps. The two most representative classifications would be used in field checking and from this the most representative classification would be chosen.

Statistical Validity. Statistical analysis of the varying soils was done to determine the necessity of parent material delineations. A selective sample of soils responses in each parent material was chosen to be compared to other soils spectra within other parent materials. A test of homogeneity was performed that would indicate how homogeneous the groupings were which would indicate the significance of considering this type of separation. The hypothesis tested the equality of variance-covariance matrices in the multivariate case where

$$H_0 = \Sigma_1 = \ldots = \Sigma_g$$

$$H_1 = H_0 \text{ is false.}$$

The procedure is a generalization of Bartlett's test in the univariate situation.

The test lets S_i denote the unbiased estimate of Σ_i for the ith group with N_i independent p-vector observations from a multivariate normal distribution with mean μ_i and variance-covariance matrix Σ_i

$$N = \sum_{i=1}^{g} N_i, \quad V_i = N_i - 1$$

and

$$c = \frac{2p^2+3p-1}{6(p+1)(g-1)}\left(\sum_{i=1}^{g} \frac{1}{v_i} \frac{1}{N-g}\right)$$

where

g = number of groups

p = number of vectors

The F value calculated by

$$C_0 = \frac{(p-1)(p+2)}{6(g-1)} \sum_{i=1}^{g} \frac{1}{v_i^2} - \frac{1}{(N-g)^2}$$

$$v_0 = \frac{v+2}{C_0 - C^2}$$

$$F = \frac{1-C-v/v_0}{v}$$

Homogeneity is rejected at the significance level α if $F>F^{\alpha}(\nu_1 \nu_0)$ where F^{α} denotes the upper percentage point of the F distribution (46).

RESULTS AND DISCUSSION

The four classifications were completed and evaluated by comparison to previously determined randomly selected sites within Jasper County. The following is a summary of the results of the methodology used in the Jasper County classification procedures.

Results of Registration

Difficulties in fitting the rectified halftone positives to the registered Landsat data prompted a registration of Jasper County using the rectified halftone positives as a base rather than the black and white panchromatic unrectified photographs. The rectified halftone positives, when used for registration, will not provide a better correlation between the two images, i.e., Landsat and halftones. However, with the use of halftones in conjunction with the Landsat classification, both must be registered to the same standard. As with the unrectified photographs, error was predicted to be no more than thirteen meters displacement.

Using rectified photographs should have eliminated some error due to crabbing and scale differences in blocking the photographs. However, because the halftones were trimmed to less than 20% overlap, difficulty is being encountered in the blocking procedure. When registration to the halftone positives is completed, the data points will be reclassified using the most accurate of the four statistical distributions.

Creation of the Parent Materials Map

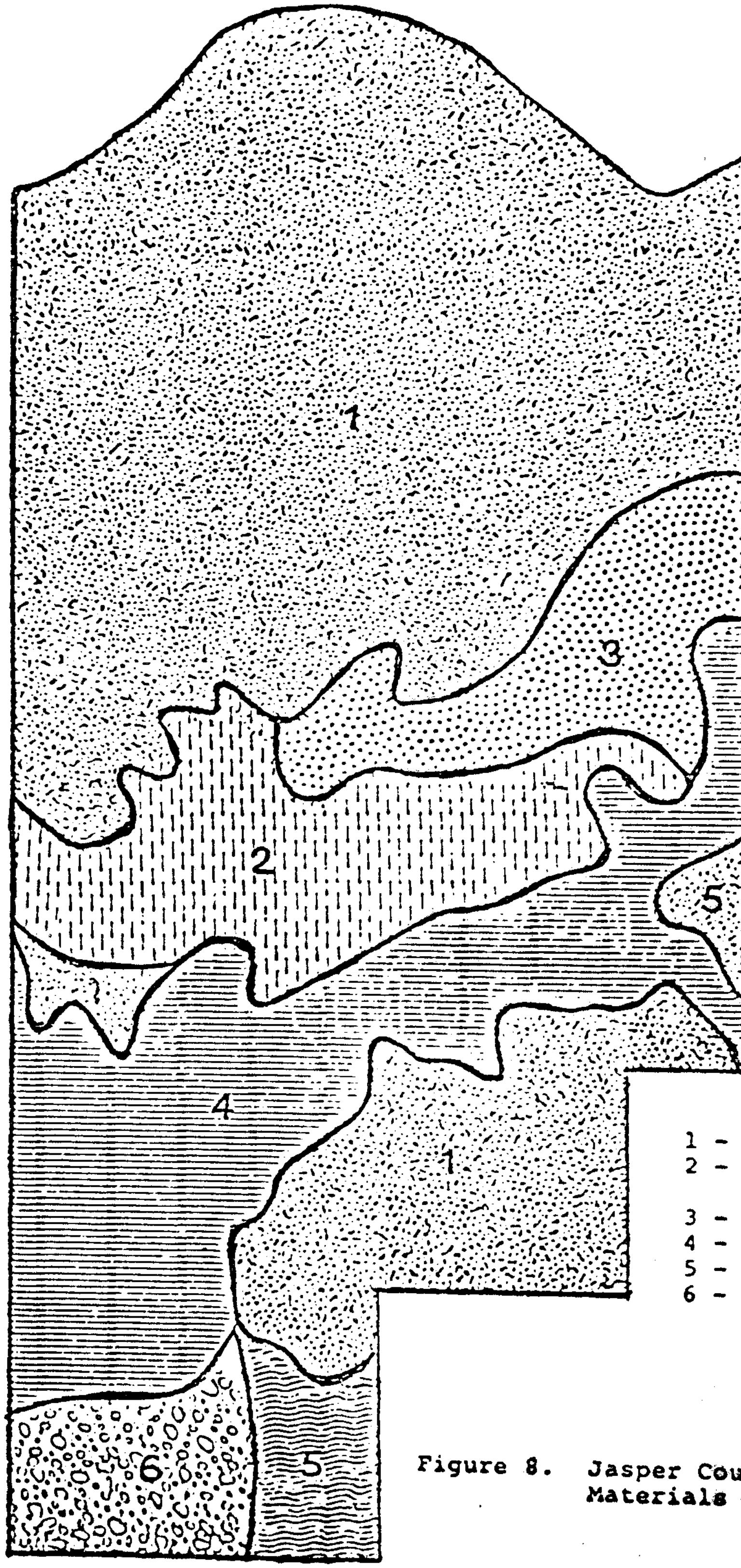
The 32 level histogrammed false color image proved to be more detailed than necessary for preparation of a parent materials map. Although the fine slicing of the spectral distribution provided more information, because of difficulty in visually interpreting the color levels (minute differences were not easily discernible) fewer defined levels would be more reasonable.

Spectral differences delineated on the false color image were field checked resulting in the identification of six parent material areas (Figure 8). To the north a primarily outwash section was identified that was characterized by windblown outwash sand ridges covered by scrub oak. The northern border of the county formed by the Kankakee River was surprisingly free of alluvial materials. The heavier sandy soils of this region are not as easily transported outside the regular bounds of the river as are the finer textured materials existing along the Iroquois River to the south. These small areas of alluvium were not of sufficient size to consider them as a designated parent material. Inclusions of organic soils also appeared in the northern and southern outwash areas. In this research the Histic areas were not sufficiently large to consider them as a separate parent material. Since they were spectrally different from other soils, they need not be considered as a separate delineation. So they were left as inclusions within other parent materials.

Glacial till, representative of the mid-north section of Jasper County, progressed across the county becoming covered by subsequent outwash deposits in the east. Rolling moraines characterized the mid section of the glacial till while to the southwest prairie influences were noted that transitioned to Alfisols in the southeast. The complexity of development was revealed by the wide variety of soils in evidence through the areas.

Between the two till areas were lacustrine deposits characterized by coarse to fine loamy textures with inclusions of finer materials located in broad low lying landscapes. Much like the Histic soils lacustrine inclusions were also scattered throughout the county.

Field Mapping of Quarter Sections. The map units were recorded on aerial photographs at 3 cm to 1 km which were evaluated by four soils analysts. Mapping of the quarter sections required approximately two weeks of field work to complete.



1 - Outwash
2 - Till-Rolling Moraine
3 - Outwash over till
4 - Lacustrine
5 - Till-Alfisols
6 - Till-Mollisols

Figure 8. Jasper County Parent Materials Map

Random selection resulted in the outwash quarter sections occurring in the same section while two other quarter sections occurred side by side in the other parent materials (Figure 9). The occurrence was advantageous in that mapping the entire section was easier by eliminating the need to travel to four different locations, but the abundance of wooded lots and pastures narrowed the area that could be used for spectral evaluation of soils. One disadvantage of MSS data is the inability to obtain soil responses through trees, or dense vegetation such as maturing crops and pastures.

The completed soil maps of the quarter sections (Figures 10-16) display a wide variety of soils. The outwash section, although primarily covered by vegetation, ranged from excessively well drained Plainfield sand to various Histic soils such as Houghton and Adrian. The lacustrine map units ranged from well drained to poorly drained soils characterized by fine sandy loam soils to silty clay textures. Till areas also ran from well to poorly drained profiles with primarily silt loam textures. In western till portions of the county bedrock occurred within 1 meter of the surface, evidence of coral reefs that were earlier discussed.

Results of Classification

Four separate classifications of Jasper County were created. Of the four, two were chosen as the most representative of county soils. The following gives an indication of each classification performance.

Classification One

All block clusters resulted in at least two vegetative responses that when tested by a divergence measurement proved to be spectrally separable. These vegetative responses generally were representative of wooded areas or crops and pastures. A wide variety of soil responses were identified, but many of these classes were eventually merged when minimal distances were found for their divergence values. Water responses were not considered of major importance. Because there were no extensive bodies of water except for scattered borrow pits along Interstate 65, their response grouped with that of poorly drained soils such as the muck areas in the north and southeast. Urban areas were not of sufficient size to consider them in a unique spectral stratification and were not of interest in characterization of soils. Urban areas were classified as vegetation because of the abundance of trees and grassed areas within cities and towns.

An overabundance of representative vegetation responses, which were not of importance, were eliminated by tolerating a lower divergence value between pairs of vegetation classes than for pairs of soils classes. That is, vegetative response classes were grouped together that were not as spectrally similar as soils classes. In this process vegetation distributions became large with wide variances which resulted in detriment to overall classification accuracy. Overall the distributions became too large and variant when classes were combined from across the county. If vegetation were left as discrete classes, correlation of soil, vegetation and combination soil-vegetation responses could be investigated.

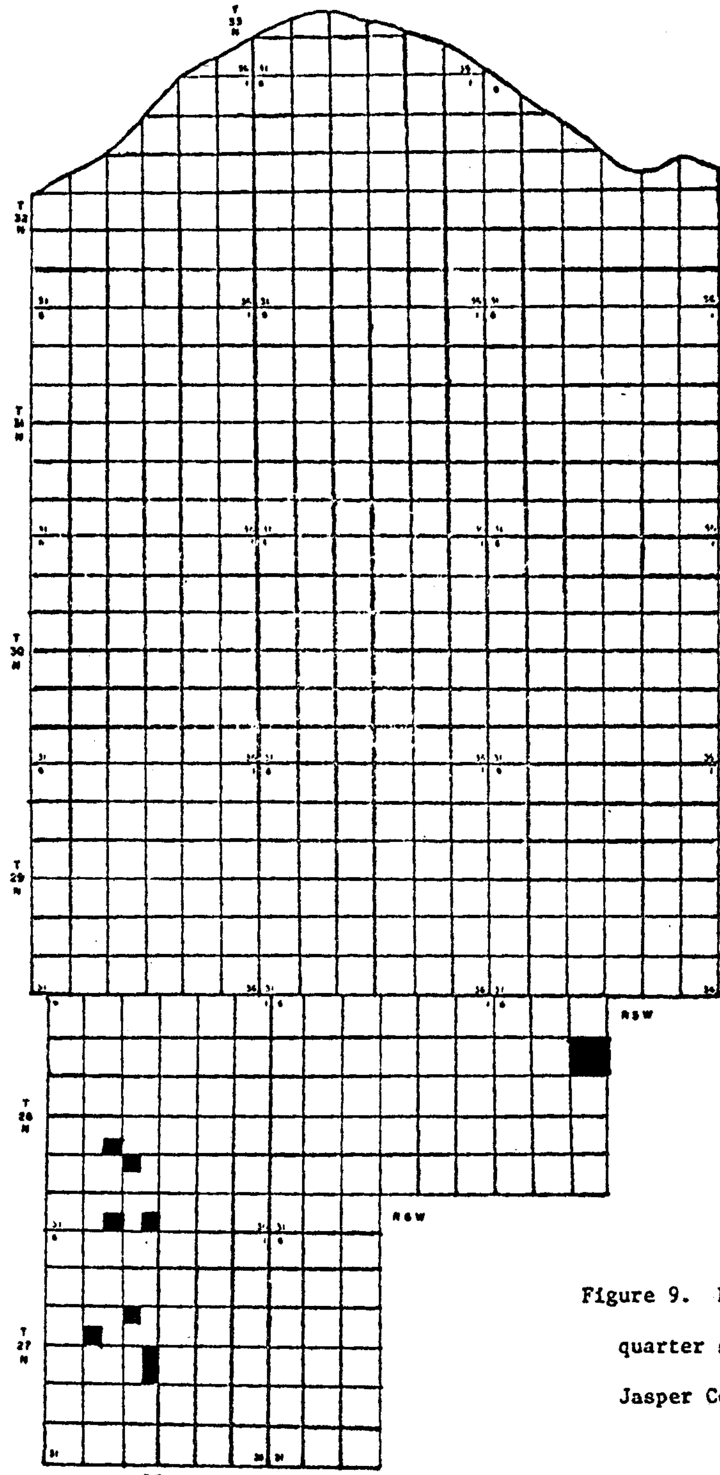
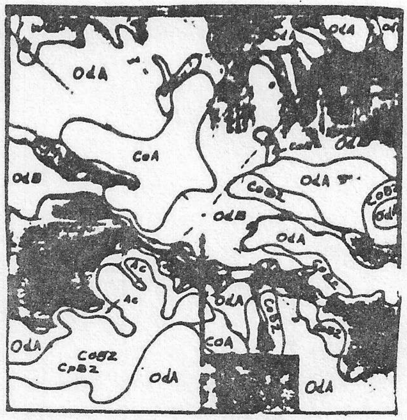
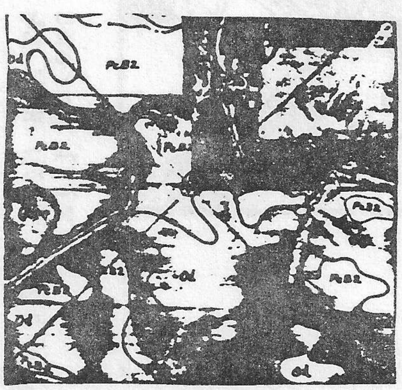


Figure 9. Randomly selected quarter sections within Jasper County.



A-C - A-C profile
Co - Corwin, well drained
Wo - Wolcott, poorly drained
0 - Bedrock 1 meter

Figure 10. Till parent material
T27N R7W SW1/4 Sec 20



Od - Odell, somewhat poorly drained
Pc - Parr, well drained

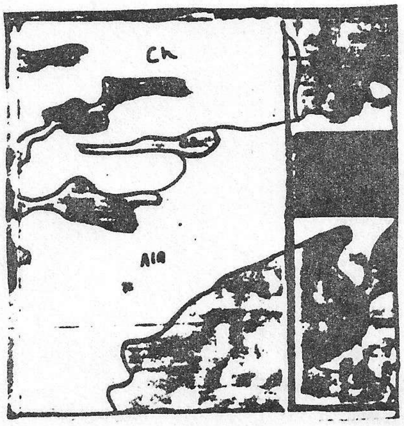
Figure 11. Till parent material
T27N R7W NW1/4 Sec 21

Cn- Conover somewhat poorly drained

Mo- Montmorenci moderately well drained

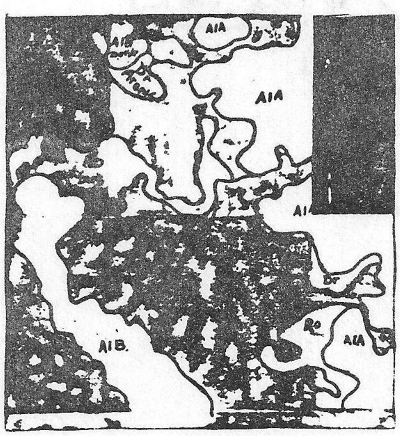
Wo- Wolcott poorly drained

Figure 12. Till parent material
T27N R7W E½ Sec28



Al - Alvin, moderately well drained
Ch - Chelsea, excessively drained
Rr - Rensselaer, poorly drained
St - Starks, somewhat poorly drained

Figure 13. Lacustrine parent material
T28N R7W SE1/4 Sec 23



Al - Alvin, well drained
Dr - Darroch, somewhat poorly drained
Ma - Mahalasville, poorly drained
Ro - Roby, somewhat poorly drained
Rr - Rensselaer, poorly drained

Figure 14. Lacustrine parent material
T28N R7W NE1/2 Sec 28

Dk - Dickinson, excessively drained
Dr - Darroch, somewhat poorly drained
Rr - Rensselaer, poorly drained

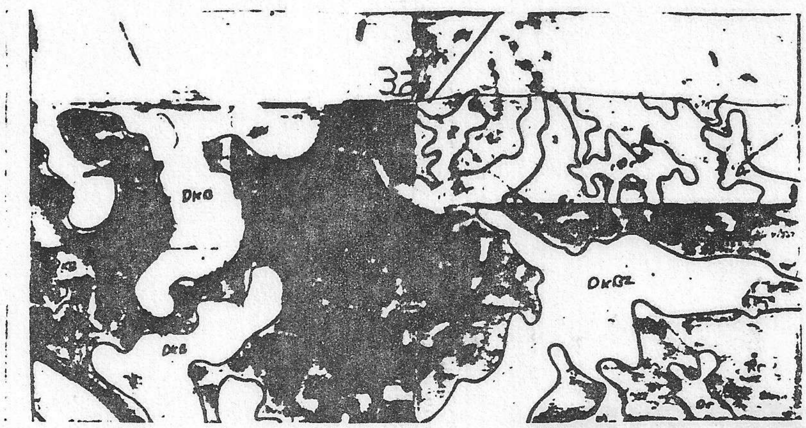


Figure 15. Lacustrine parent material
T28N R7W SE1/4 Sec 32

Ad - Adrian, very poorly drained
Ba - Brady, somewhat poorly drained
Gf - Gilford, poorly drained
Ho - Houghton, very poorly drained
Md - Maumee, poorly drained
Mr - Morocco, somewhat poorly drained
Pn - Plainfield, excessively drained

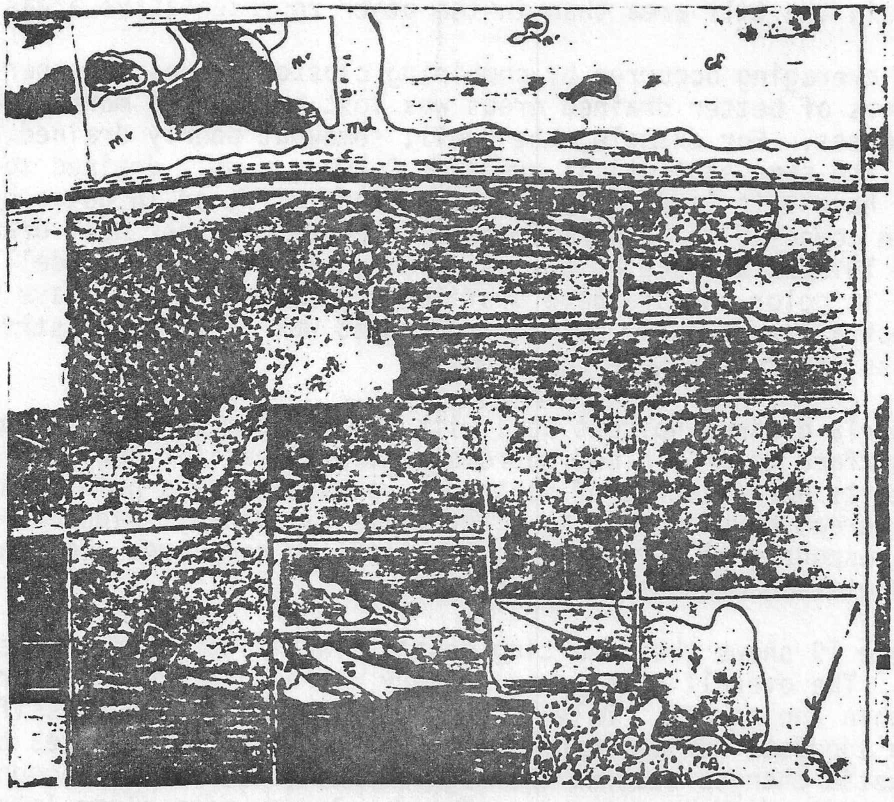


Figure 16. Outwash parent material
T23N R5W Sec 28

Over 65 statistical distributions were combined from all parent materials which resulted in 18 spectrally distinct classes. Of these 18 classes six displayed vegetation responses, nine were soils and the remaining three had characteristics of both soil and vegetation. Since these three classes had characteristics of soil, it was thought that information from these classes would contribute to identification of soils; therefore, they were considered part of the soil response group. No consideration of parent material delineations was taken either in the merging of cluster classes or in classification of the county.

Of the four analysis techniques the first spectral classification was the least representative of the mapped quarter sections. The most accurately classified of the twelve areas occurred in the till (Mollisol) parent material. Spectral soil classes were more correlated to specific map units in the till area than in the other representative areas.

Some averaging occurred by combining cluster groups together. Distinctiveness of better drained areas was lost as well as more poorly drained soil responses. For example, the Odell, somewhat poorly drained, was represented by the same statistical group as Corwin, a well drained soil. Corwin and Odell have silt loam surface textures and differed in color by 10YR2/2 for Corwin compared to 10YR2/1 for Odell. Parr, another well drained silt loam with 10YR2/2 surface color, was also confused with the Odell. Due to closeness in color and drainage profiles, these soils would have been quite close spectrally; however, soil distinctness was lost when distributions from across the county were combined.

A poorly drained Wolcott soil with a silt loam surface texture and 10YR2/1 surface color did not correlate to any specific spectral class. All map units had evidence of scattered vegetation data points which were not in as great abundance as in other classifications. Figures 17-18 show spectral responses of soils and soil-vegetation complexes along with identified vegetation classes.

Figure 19 shows the resulting county classification from the first analysis. The overall county map is very representative of general cover types within the county. Only on fine detail maps is the classification less than adequate for defining soil series. Soils differences that indicate dramatic changes such as the organic soils are easily recognized, but the more subtle differences are confused. Borrow pits along Interstate 65 are recognizable but, as stated before, were classified as a poorly drained or Histic soil.

Classification Two

A systematic clustering of prespecified lines and columns characterized the second classification scheme. By systematically sampling the entire county, data which could be overlooked by block clustering would be sampled. Unique areas of smaller than 225 ha could be bypassed since those areas are not mapped (due to the expense in establishing a soil series and the small area in relation to the county) as a soil series within a county. Those areas overlooked in a systematic sampling would not be of importance in characterization of county soils.

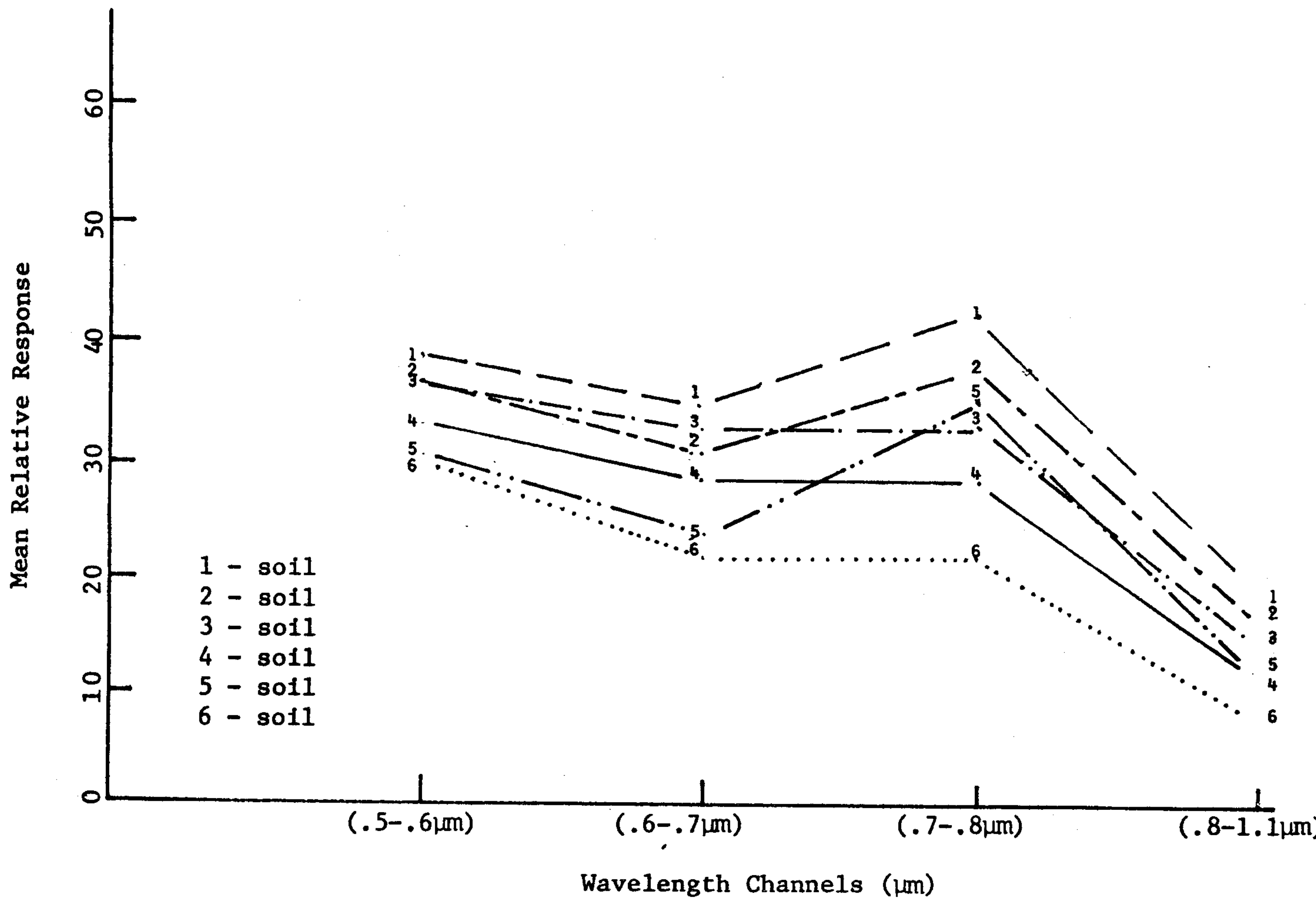


Figure 17. Soil responses for classification one.

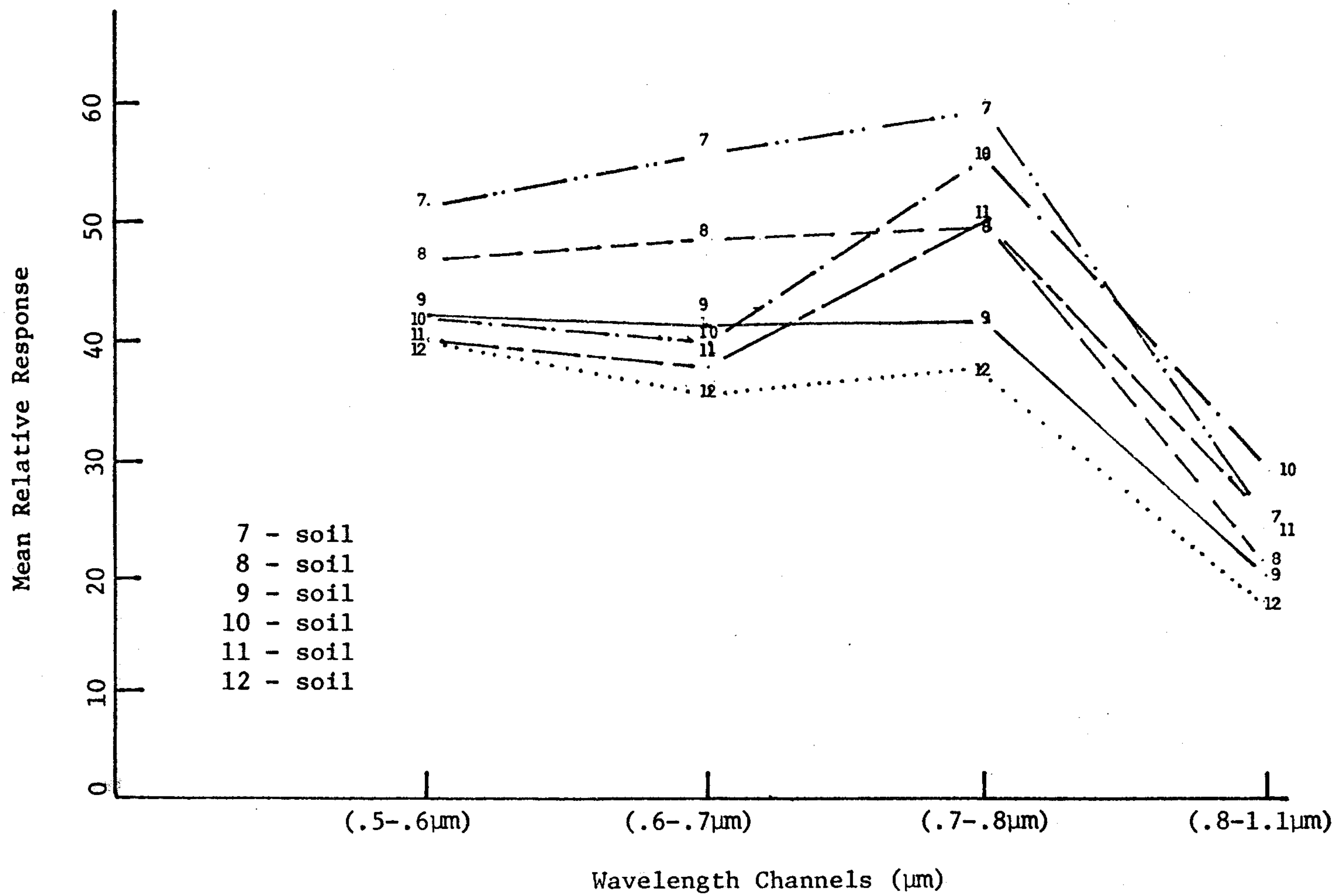


Figure 17. (Continued)

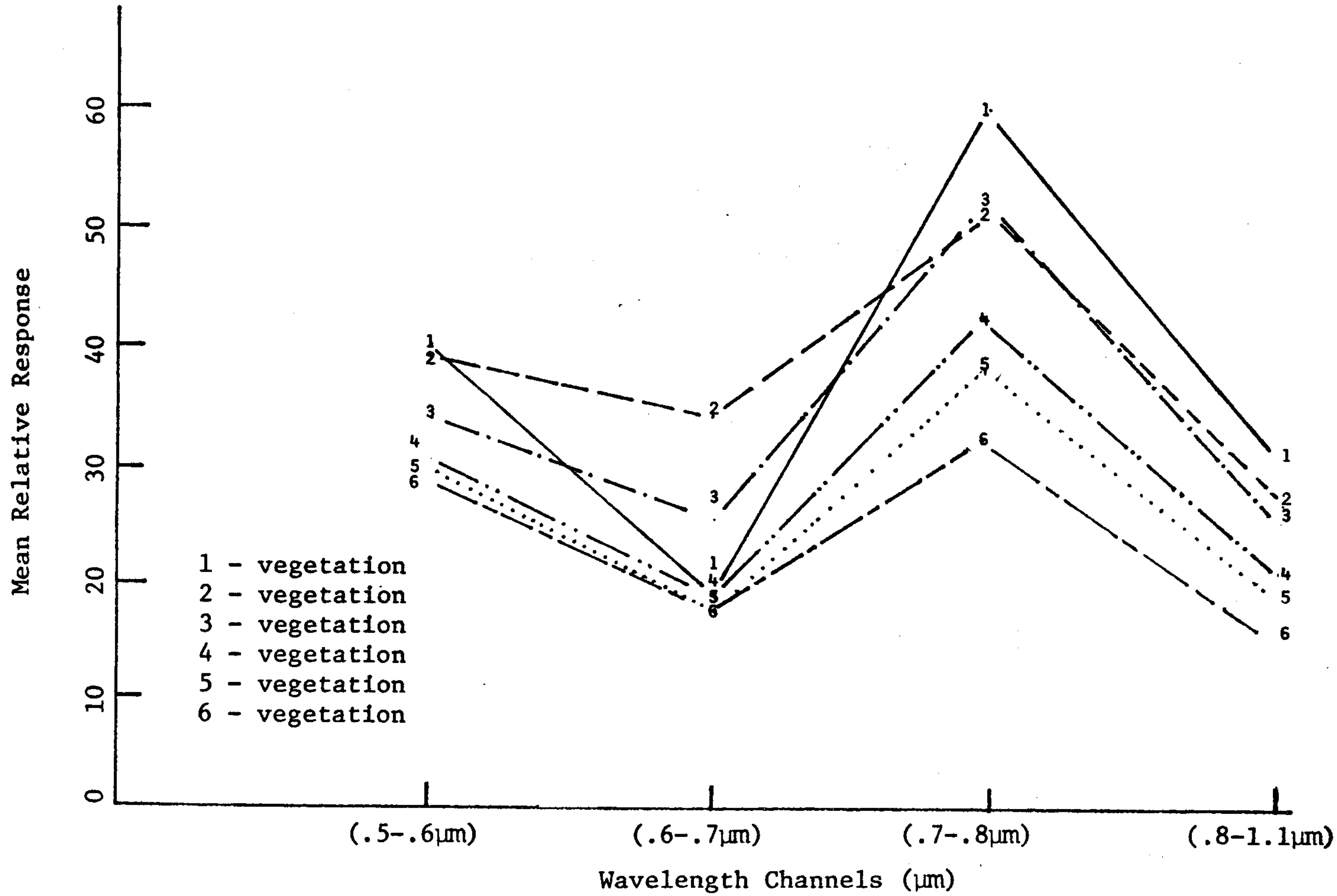


Figure 18. Vegetation responses for classification one.

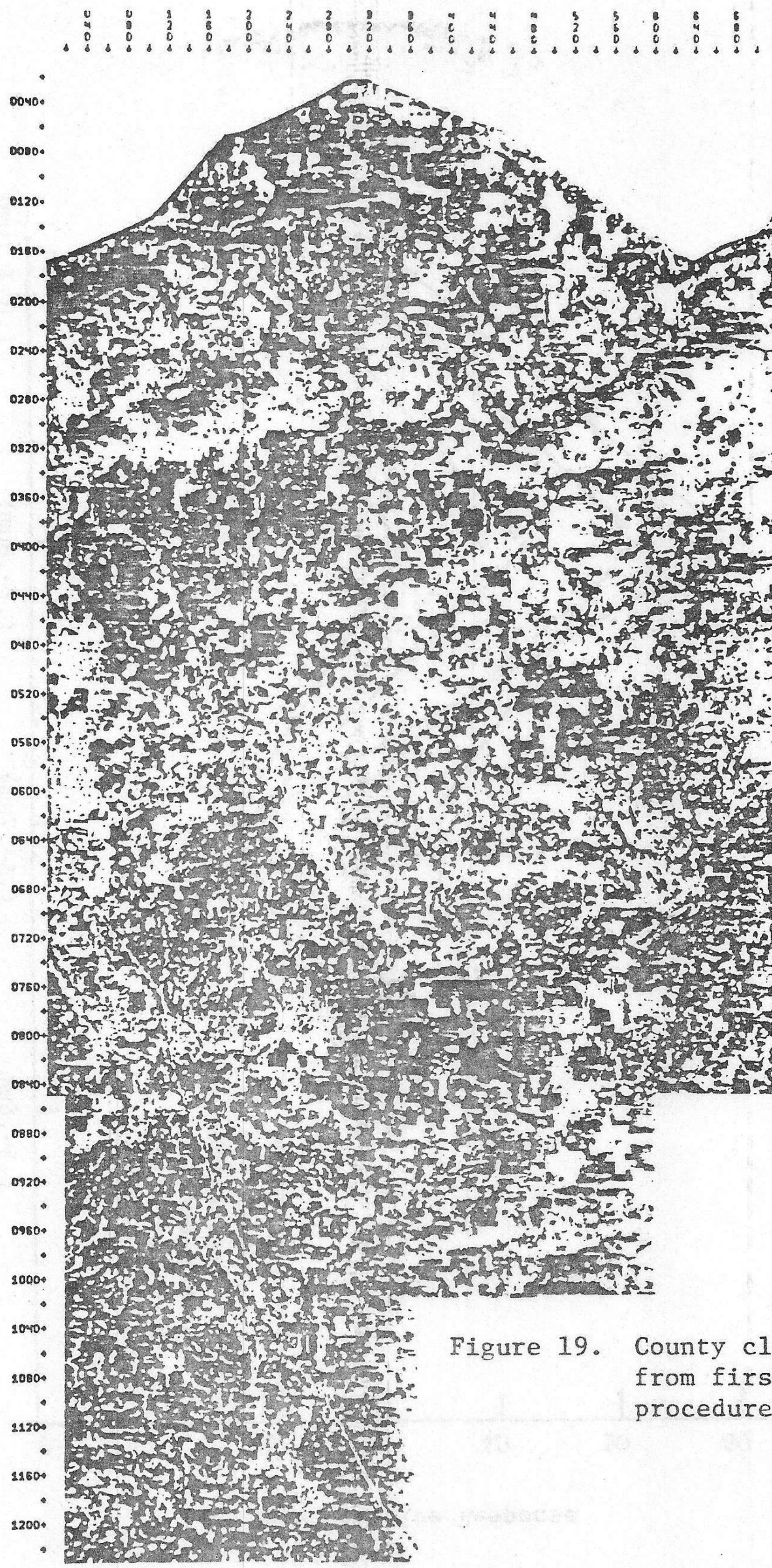


Figure 19. County classification from first analysis procedure.

Identification of cluster classes resulted in eight definite soil responses, six vegetation classes and five soils with some vegetative influence. In subjective sampling, vegetated areas such as the Jasper Pulaski State Fish and Wildlife Area, wetland areas and scrub oak areas on sand ridges were generally avoided, but systematic sampling chose points throughout the county which accounted for the increased number of soil-vegetation responses. A good definition of scrub oak, wooded areas and trees and plants along creeks and rivers was the result of the second classification because these were not avoided and could be classified with actual representative data points from the area.

Evaluation of the soil maps revealed the second classification to be more representative than the first but not of the quality displayed in the third and fourth techniques. Again, difficulties were encountered with scattered data points of vegetation appearing across the map units, but not to the extent of classification one. Odell, a somewhat poorly drained soil, was again confused with the well drained soils, Parr and Corwin.

Investigated scattered data points indicated characteristics of both soils and vegetation. In an attempt to improve the homogeneity of the map units, statistical distributions were altered by eliminating the combination soil-vegetation classes and reclassifying only those areas corresponding to the mapped quarter sections. Figures 20 and 21 show the graph of data before and after alteration. In general, data points previously classified as mixed soils and vegetation were classified as the surrounding soil response. Only classes displaying a relatively pure soil or vegetation response were used in reclassification.

The entire county was not reclassified because of the expense and because the reclassification was to be done on the rectified registration that was not yet completed. Although favorable results were encouraging, the quarter sections were not of sufficient size to infer the same results would occur across the county. In some areas elimination of classes could be of detriment to accurate scene identification by forcing data into classes not indicative of their true response nature. For example, training statistics developed specifically for a well drained soil would accurately identify that soil. However, if lighter responding erosion classes had no training statistics spectrally near, it could be classified with a light colored well drained soil. This could have been avoided had there been a unique spectral distribution designed to represent an eroded soil. Further research into identification of spectral classes should be attempted so the most influential parameters contributing to response properties can be determined. Large classification areas the size of a county necessitate careful selection of spectral responses that provide an optimal spectral design of the tract.

The second analysis provided better correlation with mapped soils which could be attributed to better defined, more evenly distributed spectral classes that were selected by a systematic approach.

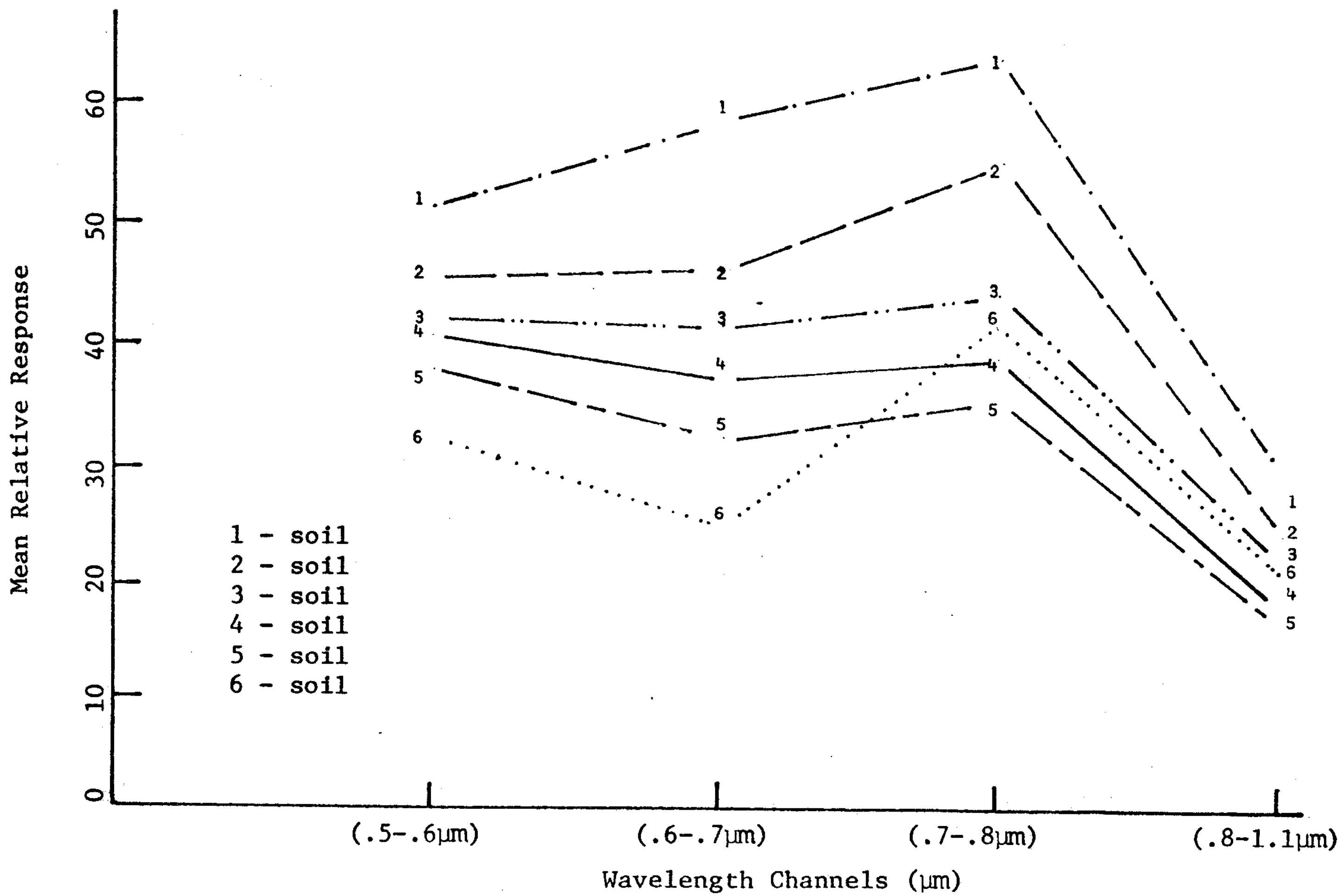


Figure 20. Soil responses of second classification.

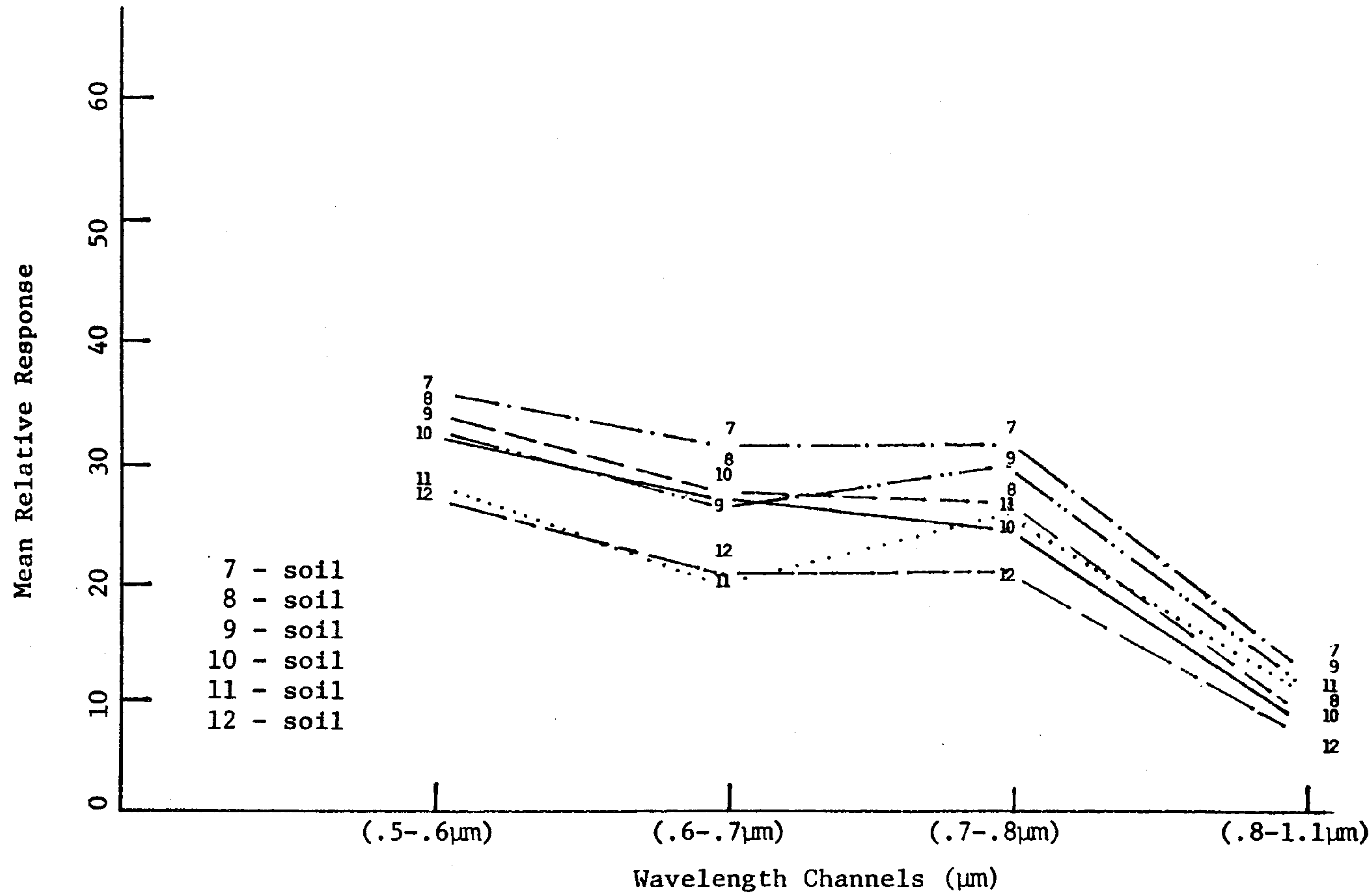


Figure 20. (Continued)

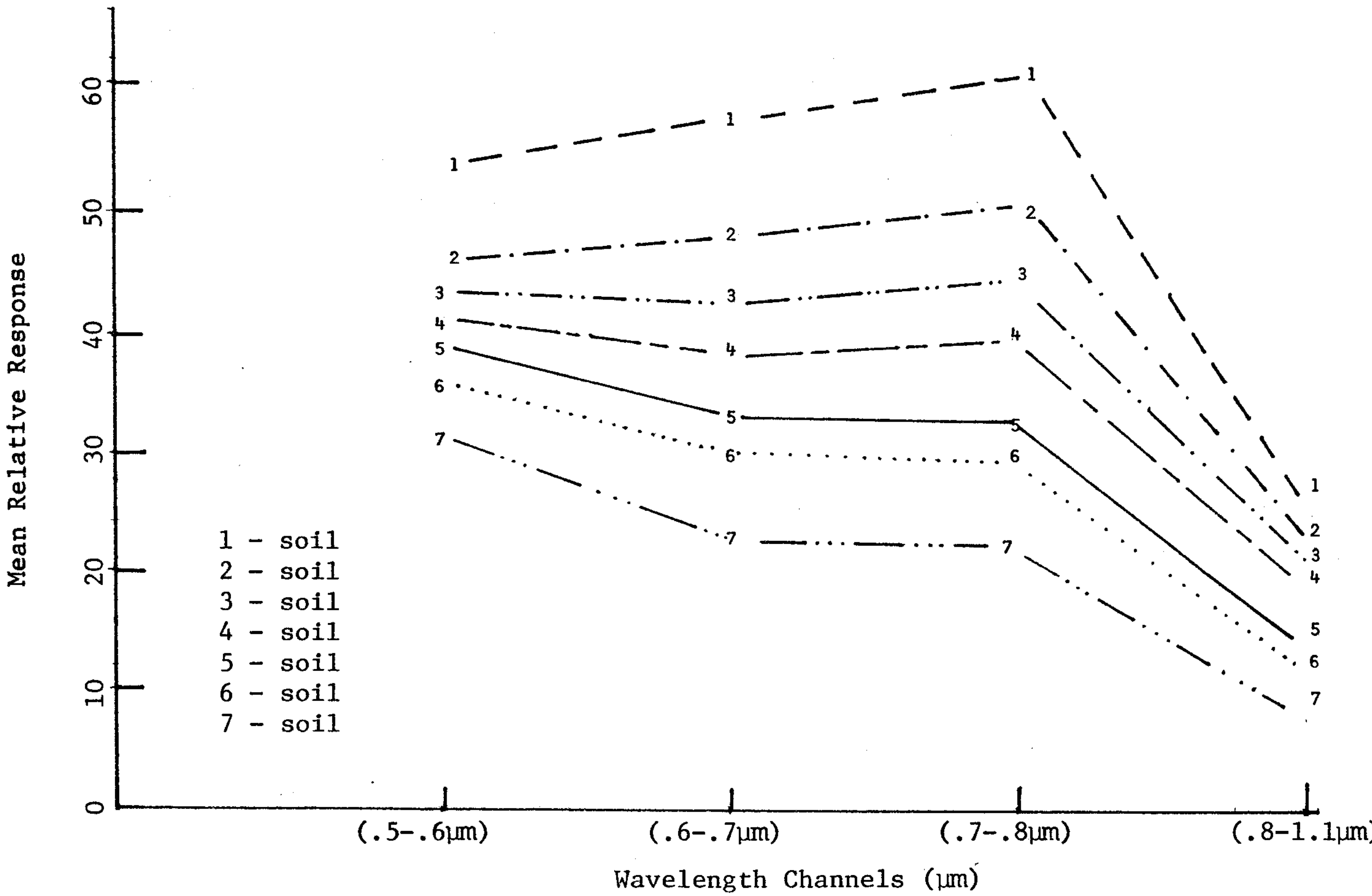


Figure 21. Soil responses of second classification after alteration of statistics.

Classification Three

By combining cluster classes only within parent materials six sets of statistical distributions were created. Each set was used to classify only within a particular parent material. The resulting set of statistical distributions contained nine soils in outwash, six soils in the rolling moraine till, seven within outwash over till, eight in the lacustrine area, eight in the till Mollisol and seven in the till Alfisol (Figure 7). Initially, vegetation was combined and a standard set of vegetation spectral distributions was used in each area. Modifications to the tree design and classifier has enlarged the capacity which will benefit future attempts at county characterization. The third analysis provided still better correlations than the first although some quarter sections displayed definite inaccuracies in soil representations.

The outwash over till parent material area produced rather unique spectral responses (Figure 22) in that only two classes displayed a characteristic soil response over the four channels while the remaining classes responded highly in channel three (indicative of vegetation) which leaves question as to whether there was a lot of scattered vegetation in the area at the time of the overpass or if these are unique soil responses that have not been encountered before. Only three of the major parent materials had been chosen to evaluate the classification so no prepared ancillary data were available to help explain these phenomena.

In the till areas Odell and Corwin were difficult to discriminate; both reflected as the lightest soil class. In some transitions to darker poorly drained soils, such as in T27N R7W Sec 20 SW¼ till, Odell responded much lower spectrally than when it was associated with Parr or Corwin. It appears that Odell, a somewhat poorly drained soil, has a wide range of reflectance. Since its drainage characteristics resemble well drained and poorly drained parameters, it may be less well drained in association with poorly drained soils and better drained when associated with well drained soils. Also, data point averaging could affect these responses.

The lacustrine shows good correlations with excessively drained soils, but evidence of inclusions within the soils was not supported with ancillary data. Areas of small inclusions could have been overlooked in the initial mapping. So, areas in question should be revisited.

The outwash areas showed good definition between spectral classes and soil series. Some variability was evidenced in separation of Brady, a somewhat poorly drained silt loam with 10YR3/1 color, and Plainfield, a well drained fine sand with 10YR4/3 color. Although Brady was separated for the majority of the map units, some pixels representative of Plainfield were integrated in the map units.

Gilford, a poorly drained sandy loam with 10YR2/1 color, was completely separated from the Maumee, a poorly drained loamy fine sand with 10YR2/0 color. Other parameters than drainage characteristics must have contributed to this spectral variability. Slight differences in texture and color could also have contributed to the ability to separate these two poorly drained

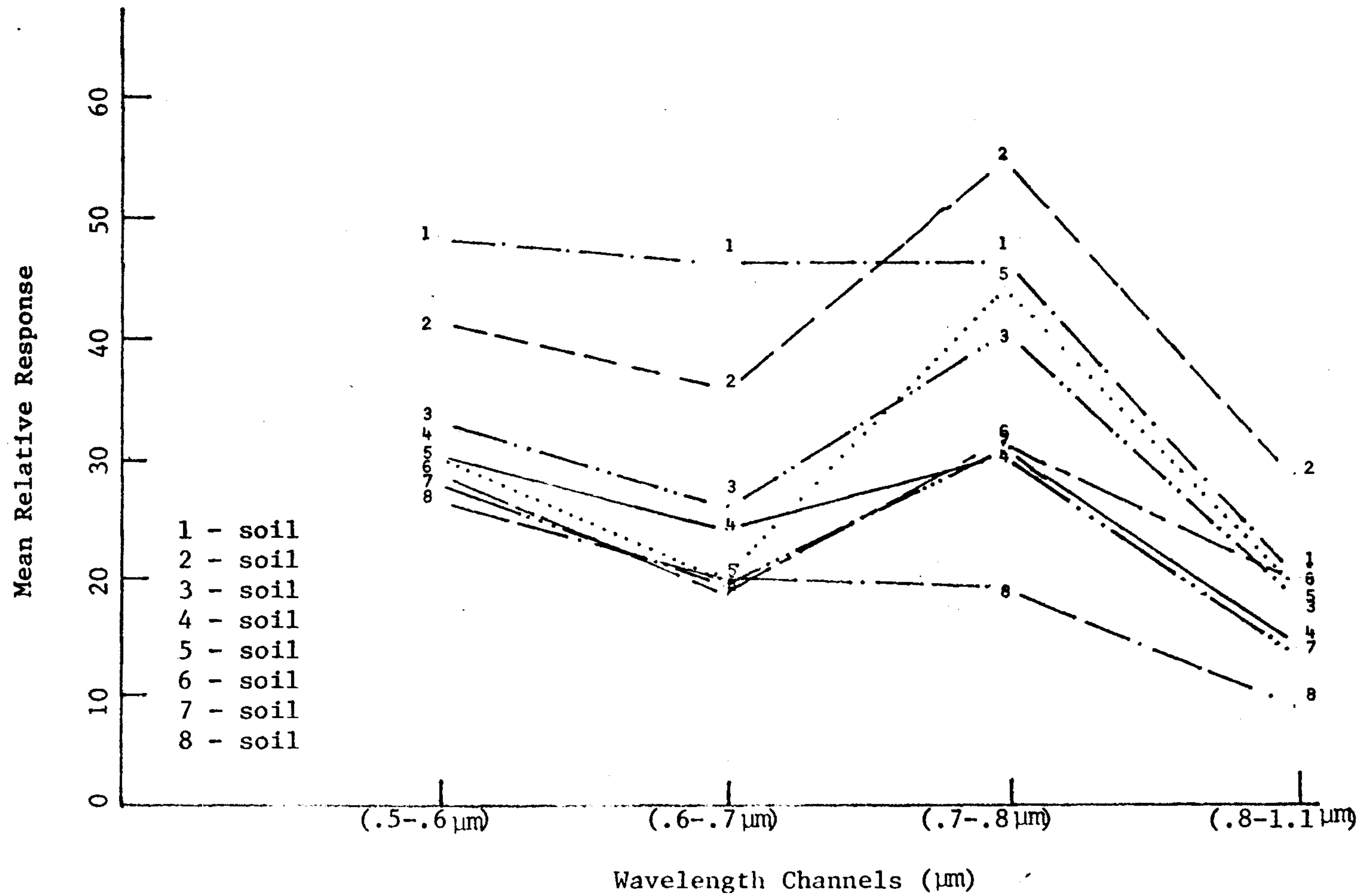


Figure 22. Soil responses in outwash over till parent material area for classification three.

soils. Maumee, for the most part, appeared in depressional wet spots and could have been categorized as a very poorly drained soil which may also have contributed to separability.

Evaluation after classification suggested that again the greatest contributor to misclassification was largely due to the influence of a combination of soil-vegetation responses.

Classification Four

The last classification proved to be the most accurate of the four analysis techniques. By clustering within parent materials four percent of the data was sampled compared to a one percent sampling in the previous analysis techniques. A larger sampling provided better definition of spectral response which resulted in a more accurate classification.

The county was statistically classified with 52 soil representations and two vegetation classes. Ten spectral classes were used in the outwash, twelve in the till (rolling moraine), ten in the outwash over till, nine spectral classes in the lacustrine area, thirteen identified in the till (Alfisols) and ten classes within the till Mollisol area.

Although overall classification four appeared more representative, misclassification was apparent in the outwash and lacustrine quarter sections. The lacustrine area, T28N R7W Sec32 SE¼, was better represented by classification three. Mixing of two spectral classes occurs in the Dickinson map units, an excessively drained fine sandy loam with 10YR2/2 surface color. Vegetation is scattered more throughout these map units in the last classification than in classification three. The outwash section, T23N R5W Sec28, which consists of large map units of Gilford and Maumee, two poorly drained soils, were not differentiated as in the previous classifications. The same occurrence was noted within the lacustrine area, where spectral responses of low responding soils were checked. The poorly drained soils were representative of the lowest reflective soil, but both graphs also revealed the next lowest responding soil as higher reflecting in the third channel. If vegetation was masking the response of soil, then perhaps bare soils within the same response group are classified to the nearest group responding as a bare soil with no vegetative influence (Figures 23, 24).

Initial success at identifying soils was based on the ability to differentiate drainage profiles; thus it was surprising to note two poorly drained soils within the same parent material. Spectral characterization of soil parameters is needed to define the extent to which each parameter contributes to overall spectral response.

Unique Characteristics of the Data

Scattered vegetation was evident across all classifications and contributed to interference with the homogeneity of all map units. These scattered vegetation-soil complexes were at first considered to not be a valid delineation. Further inspection of response values found that these

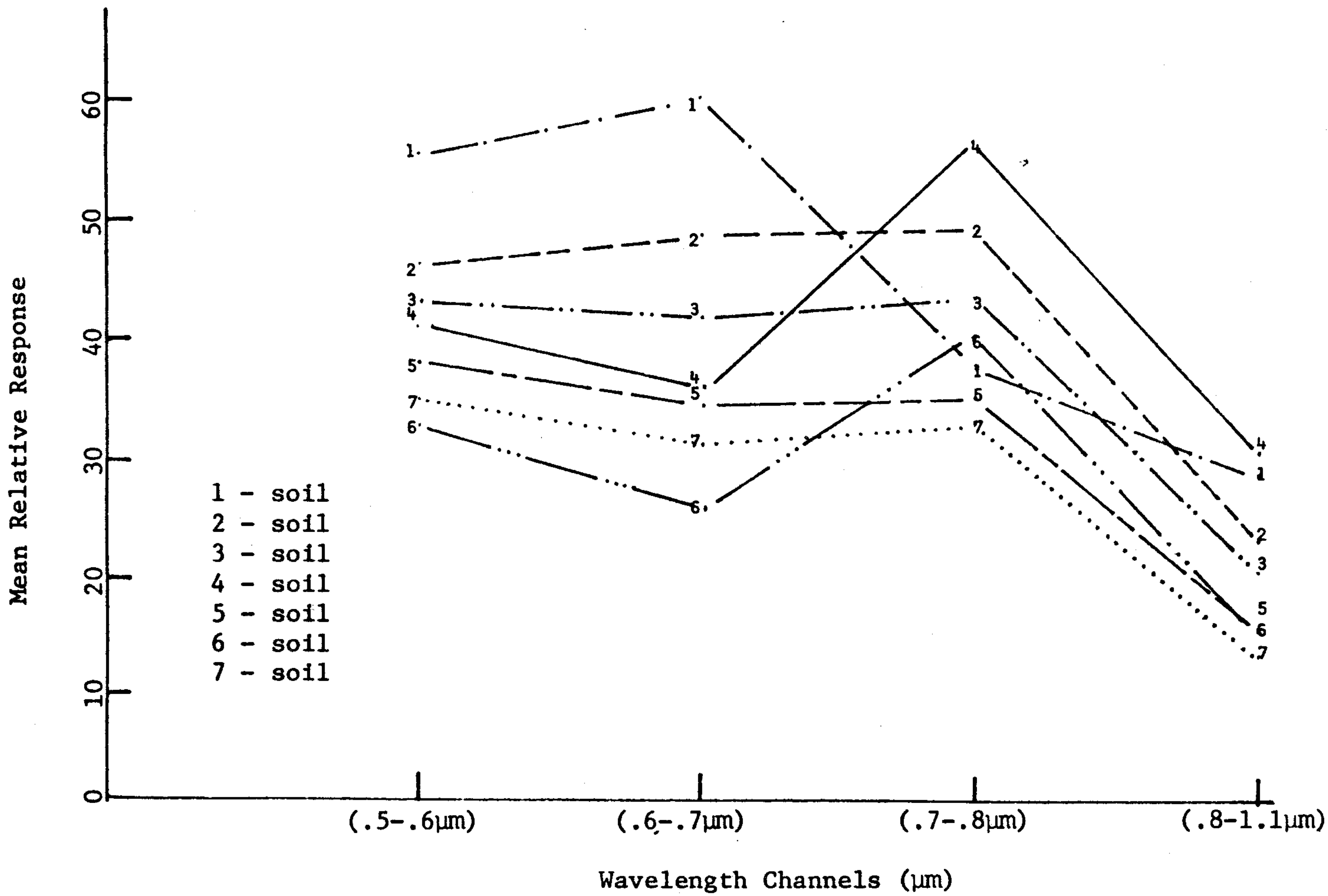


Figure 23. Classification four lacustrine soil spectral responses.

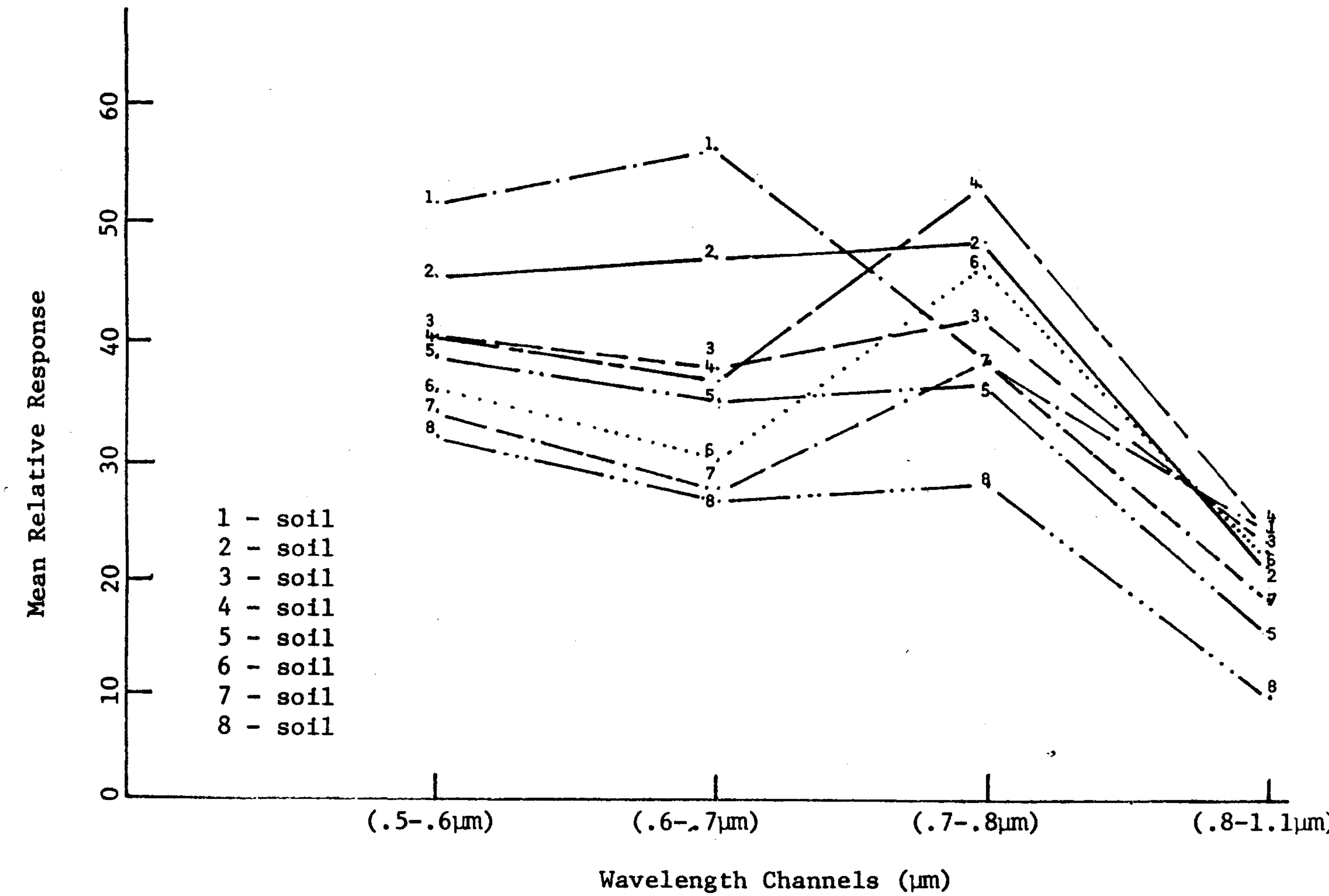


Figure 24. Soil spectral responses from outwash area, classification four.

data points were indeed a combination of vegetation and soil responses. Crop information from June 1973 found ninety percent of the corn and sixty percent of the soybeans were planted by early June. This gives explanation to why so many scattered data points were found in evaluations. These data should have been gathered before decisions were made regarding the date to be used in analyzing the remotely sensed data. At this time it is not known exactly how vegetation influences a soil response, that is, whether it gives an overall high response or low response or only influences the response in channel two and three. Information then cannot be extrapolated from these combination pixels as to the type of soil the vegetation is occurring on. If the combination points appear in a map unit, those points cannot be assumed to be a part of that map unit because of the possibility of inclusions occurring within the unit. The easiest way to eliminate this problem is to choose a date that is known to be relatively free of interfering ground cover. The next step would be to analyze responses to predict the soil from the combination response.

Aerial photographs were not rectified which contributed to error in matching map images. The resulting rectified halftone transparencies were used for reregistration which should produce a more accurate map representative of the county. Map quality photos should be essential for creation of registration data and map quality output.

Evaluation of Quarter Sections

Evaluation of quarter sections was, in general, a subjective approach with map units and spectral classifications being overlaid for comparison. One analyst did use a numerical approach by counting data points within each map unit and calculating the percent soil each spectral class represented. The purest map units or those spectral classes that represented the largest portion of any single map unit were found in the last classification. All analysts agreed that the last two classifications were the most representative of the four classification techniques.

A more quantifiable evaluative technique is necessary to provide an objective approach in selecting classifications. Bias was also integrated in the analysis techniques by the same individuals mapping the quarter sections and evaluating the classifications. By varying the individuals that mapped the quarter sections and evaluated the quarter sections, a more objective evaluation would result.

A statistical evaluation was attempted to test the validity of separating the parent materials. Both analyses (all highest responding classes and all lowest responding classes across parent materials) proved highly significant at the .01 level; therefore, the hypothesis of the homogeneity of distributions was rejected. These values may have been overly inflated due to the large number of points used in compilation of the distributions. Calculation of degrees of freedom is based on the total number of points used in the set of distributions; therefore, the large number of points contributed to the significant values. The problem was further complicated because at least six classes were needed for testing so no classes could be eliminated to reduce point size. A test more sensitive to relationships of distributions and less sensitive to point quantities is needed.

Delineations Made by the Classifier

Favorable correlations with the classification map were found when field observations were made. As in the past, drainage patterns and organic matter differences were found to be highly correlated to reflectance. Organic differences were evidenced by the separable Histic inclusions in the north and southeast. Minor differences in texture also were evident especially in the outwash area. Again, it is not certain how much contribution each of these soil parameters make to the overall soil reflectances.

Areas of moderate to severe erosion located in the till region were found to correlate almost 100% with one spectral class. Two separate areas were checked and both gave evidence to good correlation. The second area showed large areas of erosion running east to west that when field checked were not that extensive. This could be caused by east-west bias that occurs in clustering. Clustering samples point left to right across a line; therefore, the probability of points lying next to one another being in the same class is slightly higher than for points lying to the north or south. Brighter points of erosion could also have influenced neighboring pixels which would have averaged the area around the areas making it look larger than its actual extension.

The eroded class, in both areas, was not the highest responsive class. In general, the highest responding class tended to have the largest variance because it is an all inclusive class of points above a certain response. Erosion representation, since it is not the highest responding class, has definite limiters on its response range which contributes to a better defined distribution with smaller variance.

The sand ridges, in the northern part of the county, were defined by the vegetative response of the scrub oak, that occurred on the ridges. This provides the ability to map Plainfield sand, the predominant soil of the sand ridges, by delineated scrub oak areas. Areas of native vegetation or in this case scrub oak could be used to identify underlying soils if certain soils supported unique vegetation types.

A vegetation map of the county was also available through the use of the tree design processor which was used to delineate the Jasper-Pulaski Fish and Wildlife Refuge, the location of rivers and creeks, drainageways, and pastures and/or wheat fields. County roads and Interstate 65 were also visible on the final classification. Boundaries of parent materials could also be obtained and individual parent material classifications could be printed because of the nature of the layered processor (Figures 25-30).

Map products of the soils classification can be all or any part of the county at any scale. The products can be on acetate or computer printout with grey scale values, alphanumeric or symbol sets. This map quality product gives a synoptic view of Jasper County that has not been available without Landsat except through mosaicing aerial photographs.

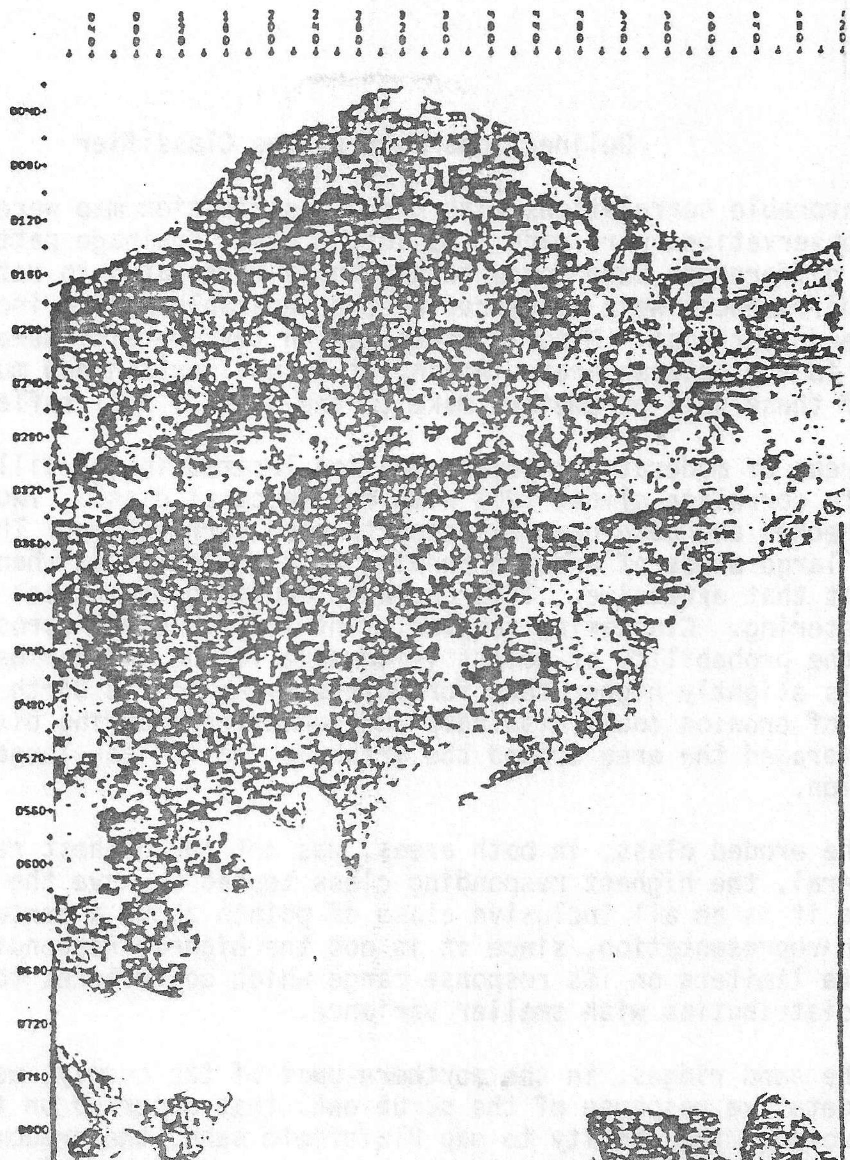


Figure 25.

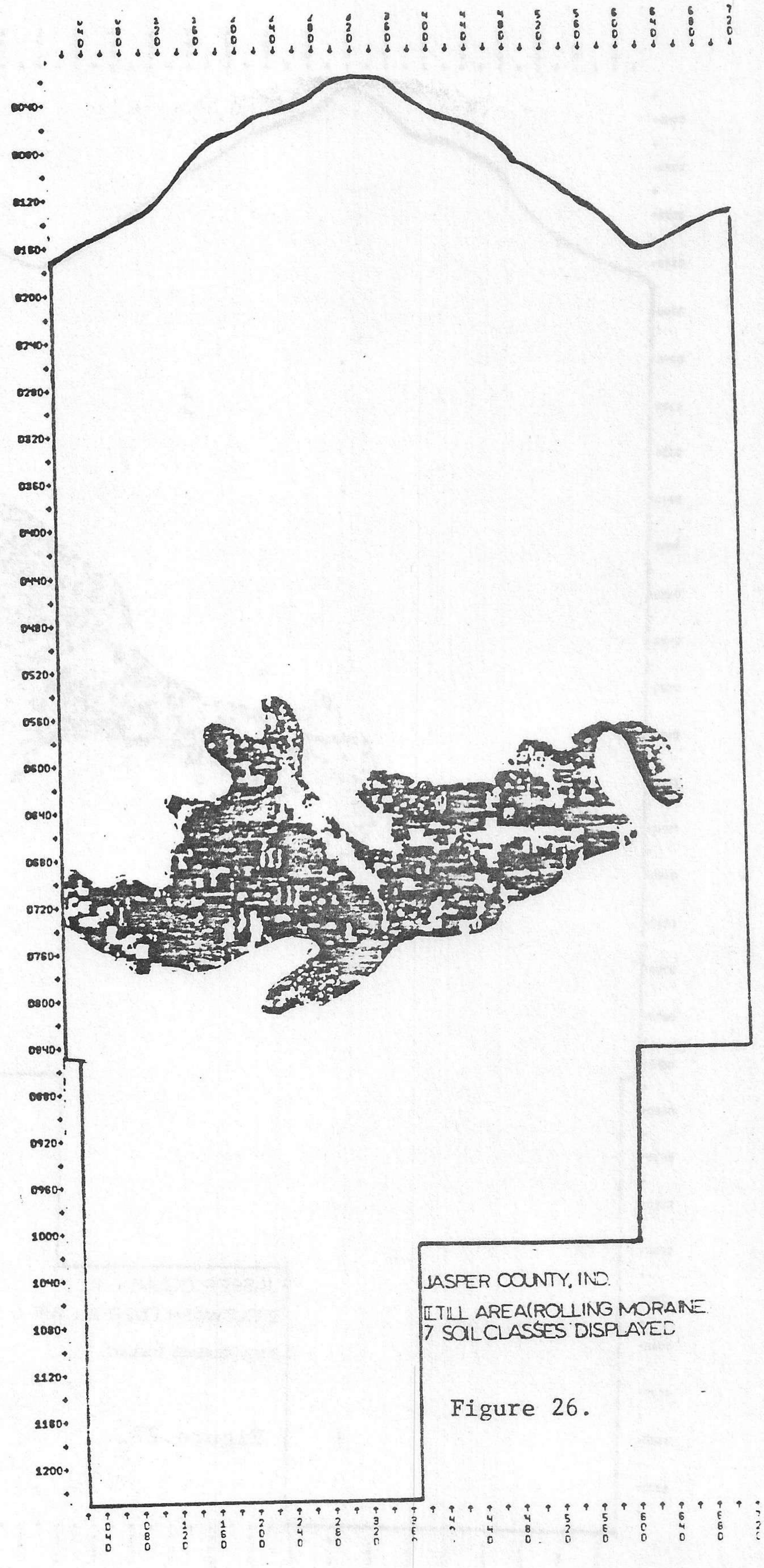

JASPER COUNTY, IND.
II TILL AREA(ROLLING MORAINE
7 SOIL CLASSES DISPLAYED

Figure 26.

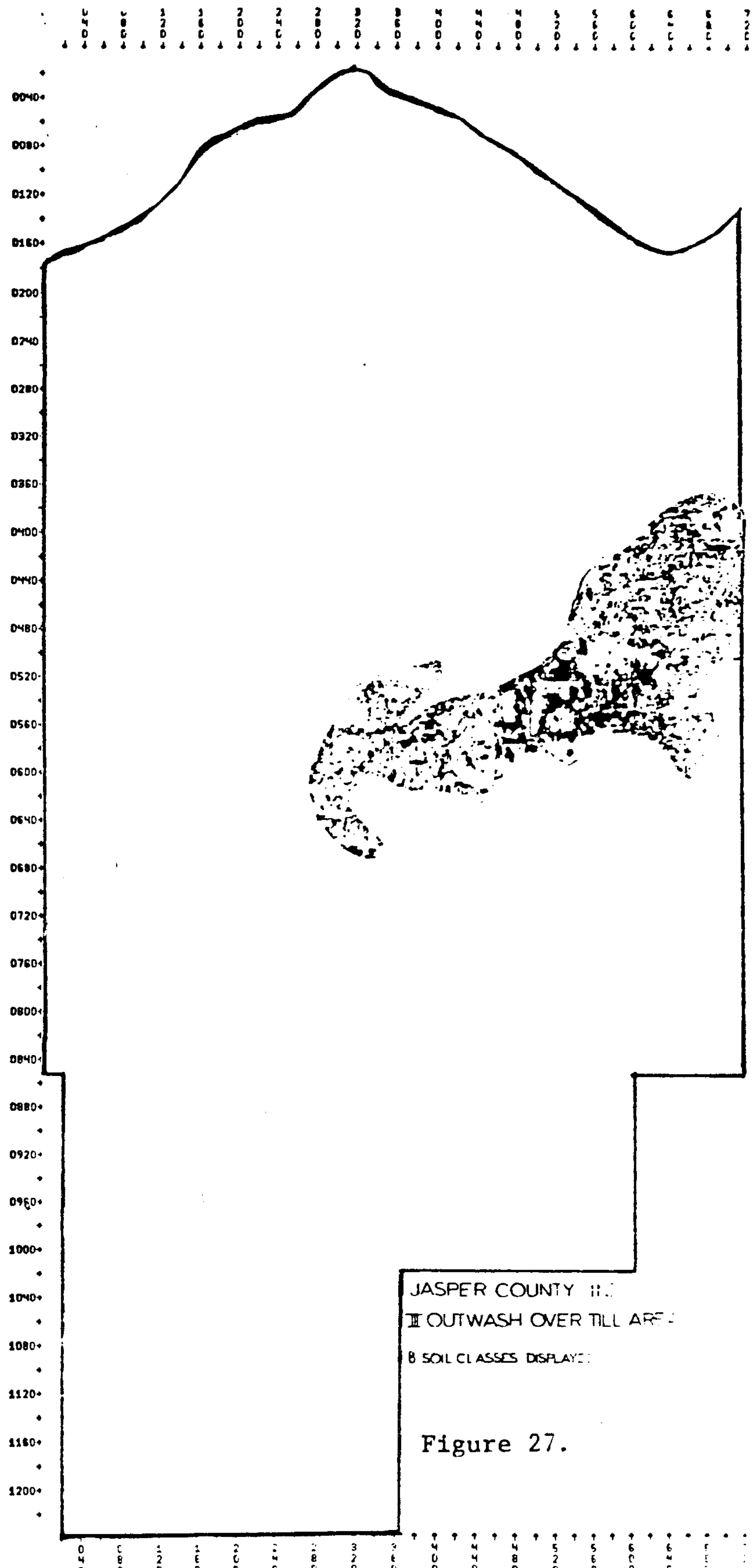

0040
0080
0120
0160
0200
0240
0280
0320
0360
0400
0440
0480
0520
0560
0600
0640
0680
0720
0760
0800
0840
0880
0920
0960
1000
1040
1080
1120
1160
1200
JASPER COUNTY
III OUTWASH OVER TILL
8 SOIL CLASSES

Figure 27.

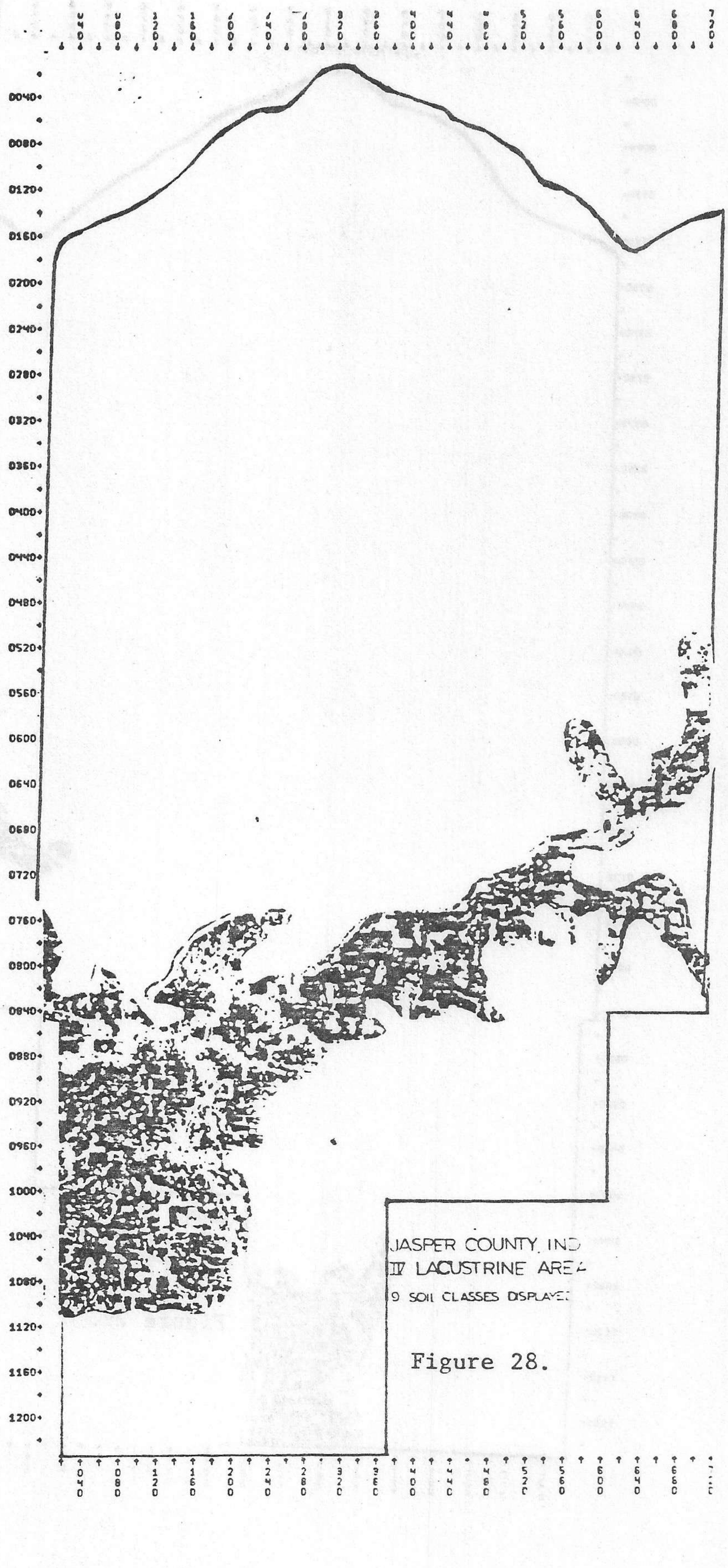

JASPER COUNTY, IND
IV LACUSTRINE AREA
9 SOIL CLASSES DISPLAYED

Figure 28.

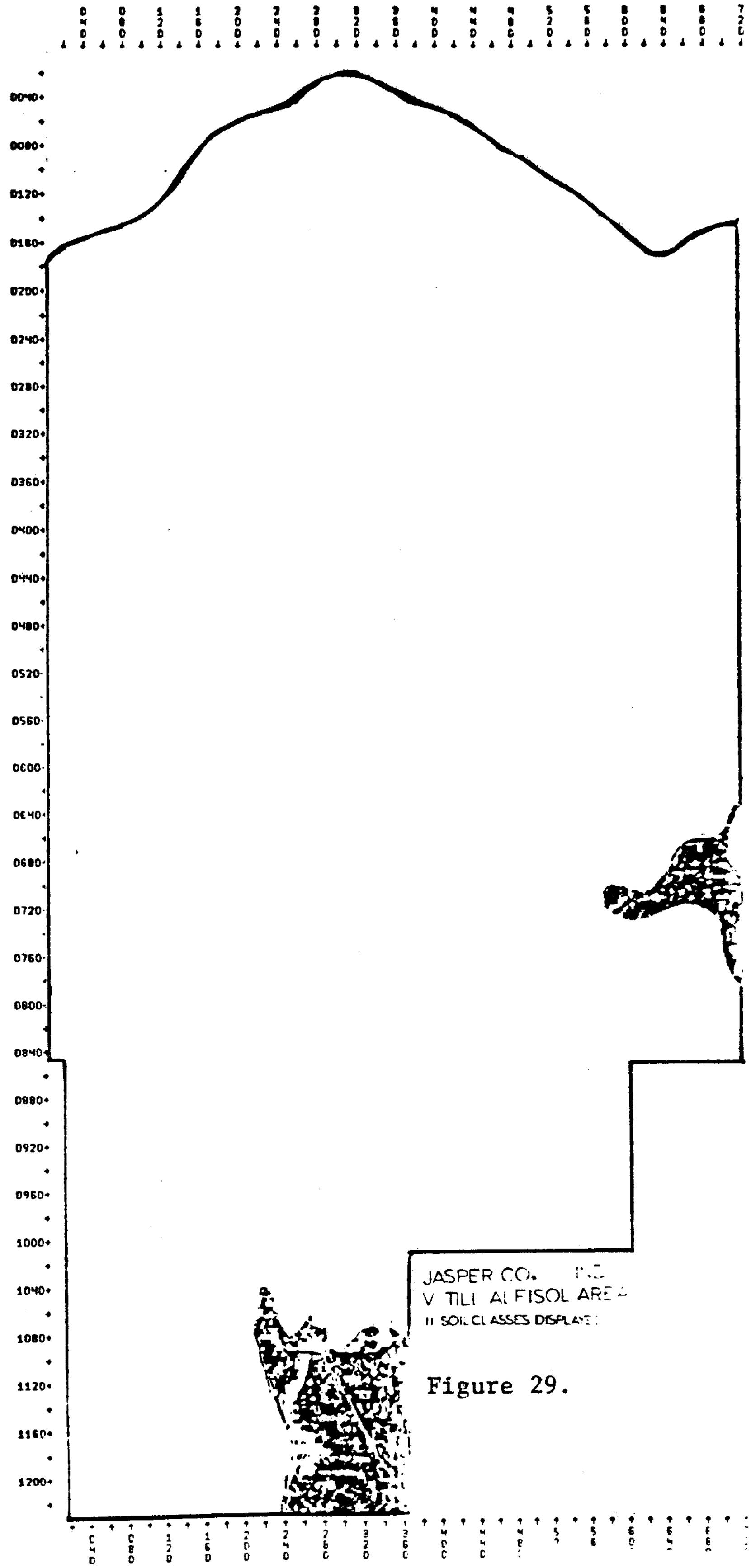

0040
0080
0120
0160
0200
0240
0280
0320
0360
0400
0440
0480
0520
0560
0600
0640
0680
0720
0760
0800
0840
0880
0920
0960
1000
1040
1080
1120
1160
1200
JASPER CO.
V TILL ALFISOL
SOIL CLASSES

Figure 29.

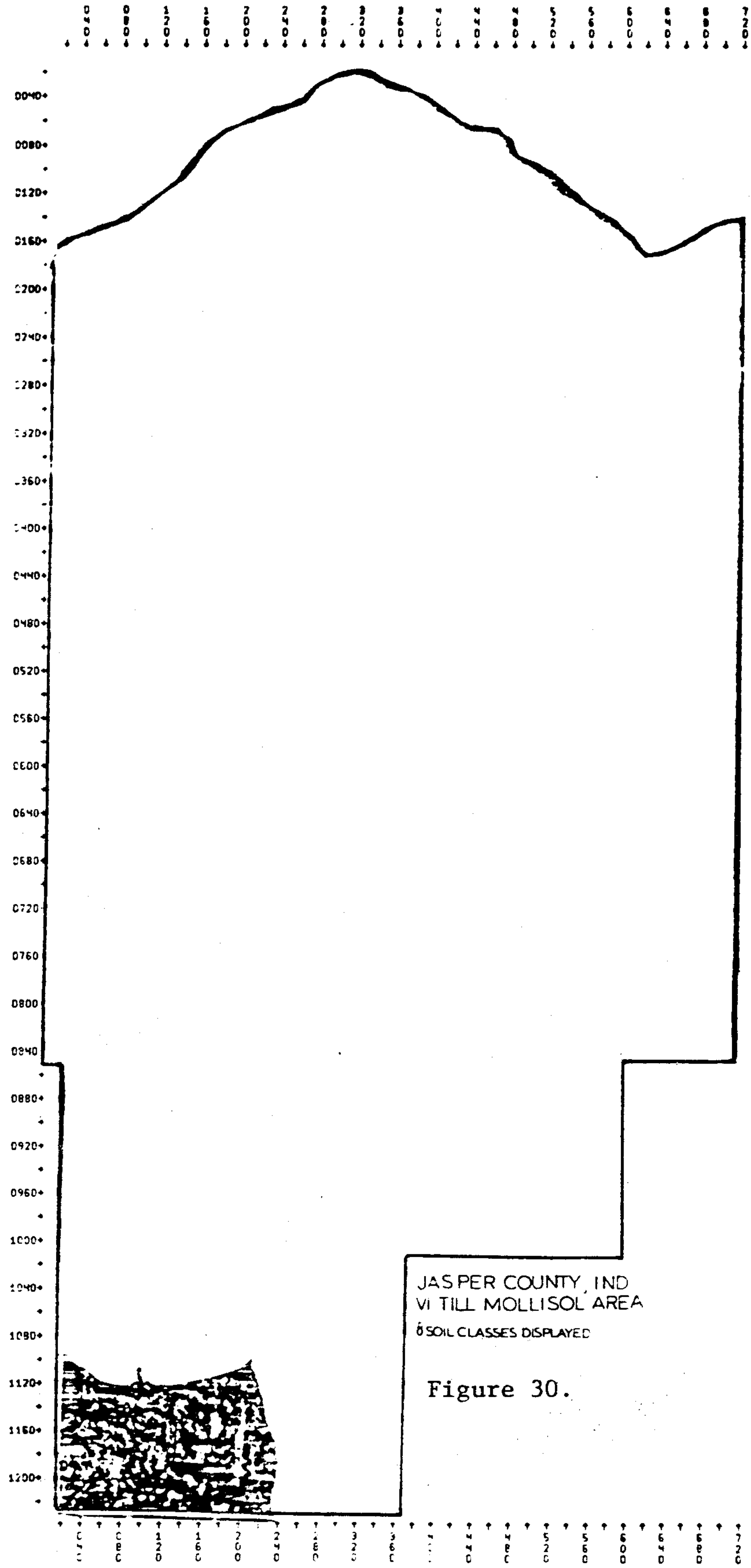

JASPER COUNTY, IND
VI TILL MOLLISOL AREA
6 SOIL CLASSES DISPLAYED

Figure 30.

An Augmented Procedure for the County Soil Survey

If this type of analysis has the potential to be used in soil surveying, where does it fit into the county plan for a survey? The decision to use remotely sensed data could be made at the same time the designation to initiate the soil survey is made. Data preparation and imagery analysis would then be instituted at the same time preliminary investigation was to take place. If photography were used as a base map for registration, then it must be taken in advance. If 7½ minute topographic maps are to be used for registration, photographs need not be available until the usual time. Digital analysis, evaluation, refinement, and creation of map products can also be done before soil mapping begins. Map products could then aid in beginning the soil mapping by locating spectrally similar soils, identifying inclusions, providing information to areas not readily accessible, identifying drainage profiles, locating possible areas of erosion, and identifying textural and organic differences. If a parent material or soil association map were created, this could aid in developing soil interpretations and establishing soil series within the county. Finally, the remotely sensed data could be used as a quality control for map units by identifying the percent inclusions, their extent and location. Figure 31 shows a possible augmented soil survey procedure.

Limitations and Difficulties

The greatest limitation was the interference of vegetation with soil response. Consideration of planting dates should influence the date when the remotely sensed data are chosen. Future remotely sensed data systems may not have the same difficulty as the Landsat MSS data, but now this is extremely important.

Registration is important if close correlation to resolution size elements is to be made. Aerial photography should be of map quality if good correlations are desired. The photographic imagery and remotely sensed data should be collected at approximately the same time.

Compilation of statistical distributions is of extreme importance for successful classification. In the systematic sampling of data points distributions were more uniform unlike the subjective sampling of data points, although in the layered processor large distributions of vegetation weighted classification of vegetation. Since the classifier has been altered to accept more than 60 classes of data, the problem should be alleviated. Large variances and platokurtic distributions should be avoided when smaller leptokurtic distributions are part of the same set of statistical distributions. The reasoning for this follows that the probability of points being classified in the larger variant distribution is greater than for the smaller variant distribution.

Still more research should be done to identify the parameters that affect soil reflectance. It is not known which parameters contribute the most to overall soil reflectance. Investigation should be made as to how they affect the response, whether higher or lower response is made because of their presence.

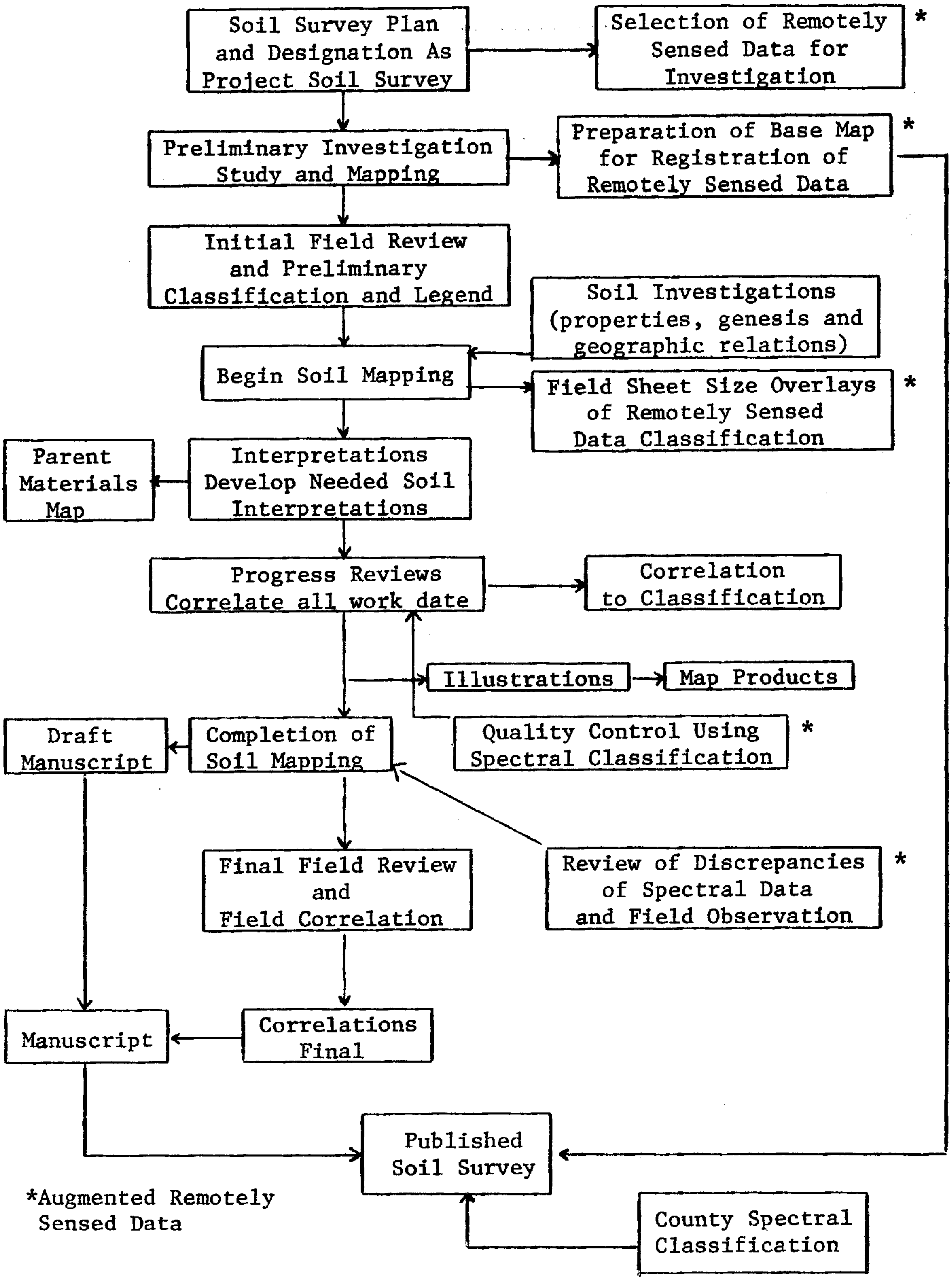


Figure 31. Steps in a soil survey that can be augmented by remotely sensed data.

SUMMARIES AND CONCLUSIONS

Preparation of a spectral map representative of soils within Jasper County resulted in relatively inexpensive quality map products that could be used in the future county soil survey. Development of a methodology from acquisition of data to creation of usable map products will aid future attempts at augmenting the traditional soil survey techniques. Heretofore, costs in acquiring map products other than panchromatic black and white photography have been prohibitive, but Landsat data analysis expense was estimated to be a reasonable expense for county use if the duration of a county survey could be shortened by increasing the daily mapping capacity.

Remotely sensed data should be closely associated in time to any ancillary data that would be used for registration purposes and/or for correlation (field checking). Prior knowledge of the amount of ground cover type and growth stage of corn, soybeans or other crops that contribute to interference with soil response would be of importance in selecting dates of data acquisition. Difficulties encountered in class confusion in the Jasper County spectral map were largely due to scattered vegetation masking soil responses.

Image interpretation of the Landsat image and field checking of the image boundaries resulted in the creation of a county parent materials map that made an obvious improvement when used as ancillary data in the statistical classification of the county.

All classifications provided a means of identifying map units that could be quantified. Soil series could be identified with the aid of the ancillary data (parent material boundaries) along with the ability to specify drainage characteristics. Organic matter differences were easily identifiable throughout the county from the muck soils in the north and the well drained sandy soils in the east with little organic matter content. Erosion was strikingly separable within the till area. (The other areas have not been checked for an erosion class.)

Difficulties in delineating closely associated soils in some areas were encountered. Somewhat poorly drained soils were confused with moderately well and well drained soils which is not surprising when the closeness of their drainage characteristics are considered. Some classes of somewhat poorly drained soils are so minutely different from better drained soils that discussion as to their delineation can be controversial even upon field inspections. These difficulties must be considered when criticism arises against MSS remotely sensed data being used because of inability to make certain soil delineations.

Evaluation of these classifications indicated the classification involving a systematic data point sampling technique for compilation of training statistics within unique parent material areas to be the most representative. Other classifications that used training samples across the entire county resulted in statistical distributions that were too broad for a fine delineation of spectral responses. Establishing a statistical representation across such a large area as Jasper County created distributions that diminished subtle differences in responses.

The subjective nature of the evaluative techniques was not adequate to evaluate classification performance quantitatively. A homogeneity test was used to determine the necessity of parent material delineation but this also proved inadequate. A more objective approach for determining classification performance and a test less sensitive to point quantities and more sensitive to relationships of distributions are needed.

Random quarter section evaluation was a sufficient means of sampling the county soils, but not all parent materials were sampled. Therefore, questions about the outwash over till area on the last classification remained unanswered. Future classification evaluation should include a larger sampling over a more extensive area.

Initially it was thought that the soil parameter most affecting Landsat spectral responses was drainage characteristics. Results in the outwash area produced spectrally separable soils of the same drainage characteristics indicating that either minor textural or organic matter differences might also significantly affect soil spectral response. Although a successful classification has been produced that will greatly aid Jasper County in their soil survey, more research is needed to determine the soil parameters that make spectral separations possible and to what extent each of the parameters contribute to overall soil response.

Final map products are available that delineate parent materials, vegetation across the entire county, specific sections or any area of the county at any map scale. These map products can be printed on acetate or paper with soil and vegetation classes represented by alphanumeric characters, symbols or varying grey scale values.

Products from this study are to be available along with rectified halftone transparent aerial photographs to be used in mapping the soils of Jasper County. The two images printed at the same scale (1:15840) were specifically designed to be a readily usable tool for field investigations. These products will provide information in areas not readily accessible and can provide the opportunity of extending the mapping time during the summer months when covered crop canopies make it extremely difficult to map.

In conclusion, this research has investigated a number of capabilities using remotely sensed data. Specifically, the research resulted in the following:

1) Designing a methodology for using remotely sensed data from the initiation of a county soil survey to evaluation of the map units;

2) Successfully creating a parent materials map through image interpretation of Landsat data;

3) Analyzing four statistical methods of classifying data points and recommending the most representative of the four to be used in county soil mapping;

4) Finding drainage characteristics, textural and organic matter differences, erosion, and scattered vegetation to be significant contributors to soil responses;

5) Map units that were easily characterized as to their homogeneity, and drainage characteristics in relation to other soils;

6) Readily available single feature maps such as vegetation maps;

7) Definable parent material areas that contribute to a more representative statistical classification of a county soil map;

8) Finding that selection of data acquisition dates is extremely important;

9) Vegetation affecting soil responses across the Landsat channels; however, it was not known how much and to what extent the response was affected;

10) Creating statistical distributions for classification of an area to be of extreme importance if an accurate classification is desirable;

11) Landsat providing a synoptic view of Jasper County that has not been available for other counties unless aerial photographs were mosaiced together;

12) Map products designed to be readily used in the research of county soils.

BIBLIOGRAPHY

1. Anuta, P. E. 1973. Geometric correction of ERTS-1 digital multispectral scanner data. Technical Report #103073. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN.

2. Anuta, P. E., S. J. Kristof, D. W. Levandowski, R. B. MacDonald and T. L. Phillips. 1971. Crop, soil and geological mapping from digitized multispectral satellite photography. Proc. 7th Int. Symp. on Remote Sensing 3:1983-2016. University of Michigan, Ann Arbor, MI.

3. Anuta, P. E. and R. B. MacDonald. 1971. Crop surveys from multiband satellite photography using digital techniques. Remote Sensing of the Environment 2:53-67.

4. Ball, G. H. and D. J. Hall. 1965. ISODATA: a novel method of data analysis and pattern classification. Stanford Research Institute, Menlo Park, CA.

5. Bauer, M. E., P. H. Swain, R. P. Mroczynski, P. E. Anuta, and R. B. MacDonald. 1971. Detection of southern corn leaf blight by remote sensing techniques. Proc. 7th Int. Symp. on Remote Sensing of the Environment 3:693-704. University of Michigan, Ann Arbor, MI.

6. Bauer, S. A. and R. J. Hanks. 1965. Reflectance of radiant energy from soil. Soil Sci. Soc. Am. J. 100:130-138.

7. Baumgardner, M. F., S. J. Kristof, C. J. Johannsen and A. L. Zachary. 1970. The effects of organic matter on multispectral properties of soils. Proc. Ind. Acad. Sci. 79:413-422.

8. Beck, R. H. 1975. Spectral characteristics of soils related to the interaction of soil moisture, organic carbon and clay content. M.S. Thesis, Purdue University, West Lafayette, IN.

9. Beckett, P. H. and S. W. Bie. 1976. Reconnaissance for soil survey II: pre-survey estimates of the intricacy of the soil pattern. Soil Sci. Soc. Am. J. 27:101-110.

10. Bochenek, Z. T. and W. A. Waldemar. 1977. Land use classification of the Warsaw, Poland area by digital analysis of Landsat multispectral data. Technical Report #090777. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN.

11. Bushnell, T. M. 1929. Aerial photographs of Jennings County. Proc. Ind. Acad. Sci. 39:229-230.

12. Colwell, R. N., W. Brewer, G. Landis, P. Langley, J. Morgan, J. Kenker, J. Robinson and A. Sorem. 1963. Basic matter and energy relationships involved in remote reconnaissance. Photogrammetric Engineering 29:761-799.

13. Cipra, J. E., M. F. Baumgardner, E. R. Stoner and R. B. MacDonald. 1971. Measuring radiance characteristics of soil with a field spectroradiometer. Soil Sci. Soc. Am. Proc. 35(6):1014-1017.

14. Dominiquez, O. A. 1960. A comparative analysis of black and white aerial photographs as aids in mapping of soils in wild land areas. p. 398-402. In Manual of Photographic Interpretation. American Society of Photogrammetry, Washington, D.C.

15. ERTS Data Users' Handbook. 1972. Document #715-4249. Goddard Space Flight Center, NASA, Greenbelt, Maryland.

16. Frey, T. H. 1967. Agricultural applications of remote sensing-the potential from space platforms. USDA Agric. Info. Bull. 328:3.

17. Fu, K. S., D. A. Landgrebe and T. L. Phillips. 1969. Information processing of remotely sensed agricultural data. Proc. IEEE 57(4): 639-653.

18. Hoffer, R. M. 1967. Interpretation of remote multispectral imagery of agricultural crops. Purdue Agric. Exp. Stn. Bull. 831.

19. Hoffer, R. M., P. E. Anuta and T. L. Phillips. 1972. ADP, multiband and multi-emulsion digitized photos. Photogrammetric Engineering 38(10): 989-1001.

20. Horvath, E. H. and M. F. Baumgardner. 1971. Multispectral remote sensing of soils. II. Optimum spectral wavelength for computer-implemented mapping of soil organic matter. Agronomy Abstracts (Division S-5):99.

21. Hovis, W. W. 1966. Infrared spectral reflectance of some common minerals. Applied Optics 5:245-248.

22. Kirschner, F. R., S. A. Kaminsky, R. A. Weismiller, H. R. Sinclair and E. J. Hinzel. 1977. Soil map unit composition assessment by digital analysis of Landsat data. Technical Report #092377. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN.

23. Kristof, S. J. and A. L. Zachary. 1974. Mapping soil features from multispectral scanner data. Photogrammetric Engineering 40(12):1427-1434.

24. Longlois, K. H., L. C. Osterholz and F. R. Kirschner. 1976. Use of Landsat imagery as a base for making a general soil map. Proc. Ind. Acad. Sci. 85:126.

25. Marill, T. and D. M. Green. 1963. On the effectiveness of receptors in recognition systems. IEEE Trans. on Information Theory IT-9:11-17.

26. Mathews, R. L., J. E. Cunningham, J. E. Cipra and T. R. West. 1971. Application of multispectral remote sensing to soil survey research in southeastern Pennsylvania. Agronomy Abstracts (Division S-5):103.

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

28. Nilsson, N. J. 1965. Learning Machines. McGraw-Hill Book Co., New York.

29. Parry, J. T., W. R. Cowan and J. A. Hegenbottom. 1969. Soil studies using color photos. Photogrammetric Engineering 35(2):44-56.

30. Peterson, J. B. 1975. Quantitative inventorying of soil and land use differences. Technical Report #063075. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN. (Invitational Paper, Annual Meeting of Soil Conservation Society of America, San Antonio, Texas.)

31. Peterson, J. B., F. E. Goodrick and W. N. Melhorn. 1975. Delineation of the boundaries of a buried pre-glacial valley with Landsat-1 data. Technical Report #100175. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN. (Invitational Paper, Earth Resources Survey Symposium, Lyndon B. Johnson Spacecraft Center, Houston, Texas.)

32. Phillips, T. L. (ed.). 1973. LARSYS Users Manual, I and II. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN.

33. Powell, J. C. and M. E. Springer. 1965. Composition and precision of classification of several mapping units of the Appling, Cecil and Lloyd Series in Walton County, Georgia. Soil Sci. Soc. Am. Proc. 29(4):454-458.

34. Reeves, R. G., A. Anson and P. Landon (editors). 1975. Manual of Remote Sensing. American Society of Photogrammetry, Falls Church, VA.

35. Shockley, W. G., S. J. Knight and E. B. Lipscomb. 1962. Identifying soil parameters with an infrared spectrometer. Proc. 2nd Symp. on Remote Sensing of the Environ. University of Michigan Press, Ann Arbor, MI.

36. Simonett, D. S. 1965. Pedologic and ecologic applications of orbital radar imagery for studies of tropical environments in use of imaging radar and two frequency altimeter scatterometer on manned spacecraft to investigate lunar and terrestrial surfaces. Tech. Rep. to National Aeronautics and Space Administration. University of Kansas, Lawrence, KS. p. 154.

37. Singh, A. N., S. J. Kristof and M. F. Baumgardner. 1977. Delineating salt-affected soils in the Ganges Plain, India, by digital analysis of Landsat data. Technical Report #111477. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN.

38. Soil Survey of Pulaski County, Indiana. 1968. U.S. Dept. of Agric. U.S. Government Printing Office, Washington, D.C.

39. Soil Survey Staff, Bureau of Plant Industry, Soils and Agricultural Engineering. 1951. Soil Survey Manual. USDA Handbook No. 18, U.S. Government Printing Office, Washington, D.C.

40. Sontirat, V. 1977. Land use classification of the Bangkok, Thailand area by digital analysis of Landsat multispectral data. M.S. Thesis, Purdue University, West Lafayette, IN.

41. Steinhardt, G. C., D. P. Franzmeier and J. E. Cipra. 1975. Indiana soil associations compared to Earth Resources Technology Satellite imagery. Proc. Ind. Acad. Sci. 84:463-468.

42. Stockton, J. G., A. L. Zachary and M. F. Baumgardner. 1973. The influence of internal drainage on multispectral reflectance from surface soils. Technical Report #030873. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN

43. Stoner, E. R. and E. H. Horvath. 1971. The effect of cultural practices on multispectral response from surface soil. Proc. 7th Int. Symp. on Remote Sensing 3:2109-2113.

44. Swain, P. H. 1972. Pattern recognition: a basis for remote sensing data analysis. Technical Report #111572. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, IN.

45. Swain, P. H. and H. Hauska. 1977. The decision tree classifier: a design and potential. IEEE Trans. on Geoscience Electronics GE-15(3): 142-147.

46. Timm, N. H. 1975. Multivariate Analysis with Applications in Education and Psychology. Brooks-Cole Publishing Co., Monterey, CA.

47. USDA Agric. Info. Bull. #352. 1971.

48. Weismiller, R. A., J. L. Ahlrichs and J. L. White. 1967. Infrared studies of hydroxy-aluminum interlayer material. Soil Sci. Soc. Am. Proc. 31(4):459-463.

49. Weismiller, R. A., I. D. Persinger and O. L. Montgomery. 1977. Soil inventory from digital analysis of satellite scanner and topographic data. Soil Sci. Soc. Am. J. 41(6):1166-1170.

50. Westin, F. C. 1973. ERTS-1 imagery: a tool for identifying soil associations. Proc. to Earth Survey Problems through the Use of Space Techniques, General Assembly of Committee on Space Research. Ronstanz, West Germany.

51. Westin, F. C. and C. J. Frazee. 1976. Landsat data, its use in a soil survey program. Soil Sci. Soc. Am. J. 40(1):81-89.