Information Note 010772

Engineering Soils Mapping from Multispectral Imagery Using Automatic Classification Techniques

by
T.R. West

The Laboratory for Applications of Remote Sensing

Purdue University, West Lafayette, Indiana

1972

LARS Information Note #010772

ENGINEERING SOILS MAPPING FROM MULTISPECTRAL IMAGERY USING AUTOMATIC CLASSIFICATION TECHNIQUES

by

Terry R. West

Purdue University

Paper Presented at the 51st Annual Meeting of the Highway Research Board

January, 1972

ENGINEERING SOILS MAPPING FROM MULTISPECTRAL IMAGERY USING AUTOMATIC CLASSIFICATION TECHNIQUES*

T. R. West

Purdue University

ABSTRACT

Multispectral imagery collected over southeastern Pennsylvania by Willow Run Laboratories, University of Michigan, was analyzed at the Laboratory for Applications of Remote Sensing (LARS) using its current capabilities for automatic classification. The project involved the evaluation of imagery in thirteen discrete bands of the spectrum as a source of data on engineering soils. Detailed computer classifications of two 4-mile segments of the Pennsylvania flightline, each containing predominantly one parent soil material, were obtained. Classification accuracy as measured by training field and test field performance was in excess of 90%. Using these two segments as a basis, a computer-implemented map was obtained for 12 miles of the flightline.

Soil pattern recognition using the LARS approach predates this study but most previous work differed from this in its objectives and analysis procedures. Agricultural soil studies have involved detailed mapping of relatively small areas (about 100 acres) located in the glaciated areas of the midwestern corn belt where topographic relief is low. A previous engineering soils study (Tanguay and Miles), in which the LARS and other remote sensing techniques were considered, was made over the same general terrain as the agricultural studies using the LARS techniques available at that time (1969). The Pennsylvania flightline, over a region of bedrock-derived residual soils having moderate relief, presented a new challenge to the continually evolving LARS techniques.

*Research reported in this paper was supported by National Aeronautics and Space Administration (NASA) Grant NGR 15-005-112 and U. S. Department of Transportation Contract FH-11-7565.

INTRODUCTION

In June 1970, the research project entitled "Analyze Multisensor Data Collected over Various Test Sites by Digital Computer" was initiated by the Laboratory for Applications of Remote Sensing (LARS), Purdue University. Sponsored by the Federal Highway Administration (FHWA), U. S. Department of Transportation, in cooperation with NASA (the primary funding agency of the LARS facility), this research involves the analysis of remote sensing data using the LARS digital computer and automatic multispectral data analysis techniques. Previous work applying the LARS techniques to engineering soils has been reported by Rib and Miles (4) and Tanguay and Miles (5).

Test sites in Pennsylvania, Kansas City, Kansas, and rural Kansas and Virginia were flown and the multispectral data tapes forwarded to LARS for analysis. In all, tapes covering more than 100 miles of flightline, with repeat flights for some portions, were received. Analysis of these data constitutes a continuing project. The work done on one of these flightlines, the Pennsylvania Test Site, is discussed in this paper.

DESCRIPTION OF TEST SITE

The Pennsylvania site is located in southeastern Pennsylvania about 38 miles east of Harrisburg. It is approximately 47 miles long, originating near the town of White Horse on the south and extending north-west to terminate at Tremont (and Interstate 81) on the north. From south to north, the flightline crosses Lancaster, Lebanon and Berks Counties (see Figure 1). The southern end is located in the Piedmont Plateau physiographic province, but the flightline soon enters the Valley and Ridge province and ends there about 40 miles to the north.

The test site location was selected by Federal Highway Administration personnel because of its accessibility from their offices in Washington, D. C., and because of the favorable geologic features of the area. The flightline lies at right angles to the prominent structural trend of the Appalachian Mountain system (or perpendicular to its strike) in this portion of Pennsylvania. This orientation maximizes the number of different rock units intercepted by the flightline. Also, the extensive cultivation in the area maximized the percentage of bare soil present in the spring when the flight was made.

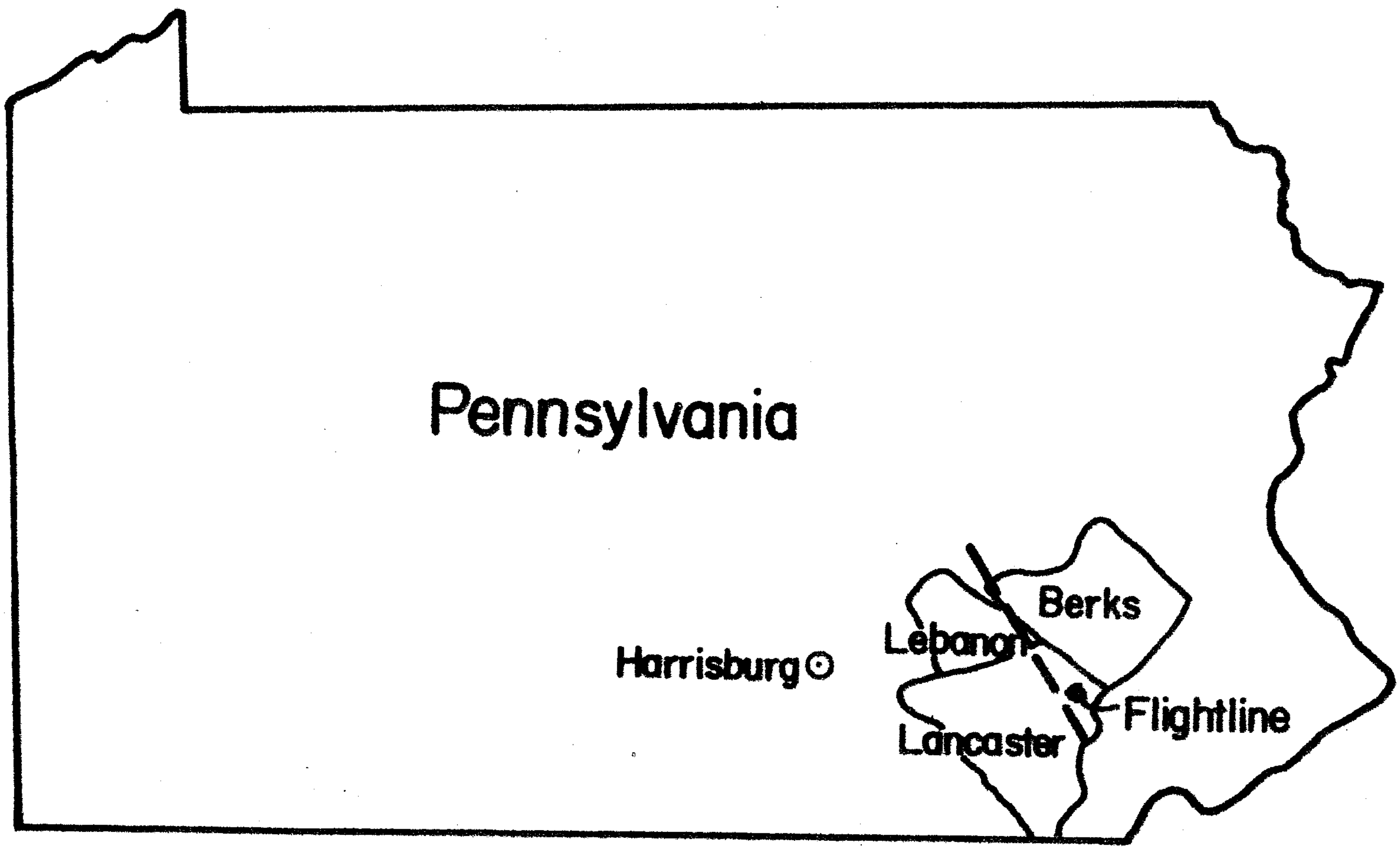

Figure 1. Locations Map for Pennsylvania Test Site

The geology of the Pennsylvania flightline area can be described in general as consisting of simply-to-complexly folded, steeply dipping sedimentary rock strata cut by several fault systems and igneous intrusions of Triassic age. The sedimentary rocks are primarily sandstone, shale, limestone and dolomite sequences of Paleozoic age, but they also contain some Triassic sandstone, shale and conglomerate. Southeastern Pennsylvania has not been glaciated and the soils consist of residual material formed by surface weathering of the parent bedrock. Local relief in the area ranges from 100 to 500 feet and the elevation from 350 to 1000 feet above sea level. Annual precipitation is approximately 41 inches per year and dense forests develop in areas not under active cultivation.

Because of the complex structure of the sedimentary rocks, portions of the flightline were selected which contained predominantly one parent bedrock material. These specific "pure soil" areas were designated along the flightline with the intention that analysis would first be accomplished for these and subsequently extended to the more diverse areas between them. Table 1 lists the pure soil areas for the flightline, giving the location of each and its assigned number, running consecutively from south to north.

TABLE 1. PURE SOIL AREAS PENNSYLVANIA TEST SITE

Pure Soil Area No.	Designation	Multispectral Scanner Tape	Parent Material	Geologic Age
1	Blue Ball	1	Limestone	Cambrian
2	Blue Ball	1	Limestone	Cambrian
3	Independence School	1	Sandstone & Shale	Triassic
4	Reamstown to Denver	1	Limestone	Ordovician
5	Union House	2	Sandstone & Shale	Triassic
6	Sheridan	2	Limestone	Ordovician
7	Mt. Aetna	2	Sandstone & Shale	Ordovician
8	Bethel	2	Sandstone & Shale	Ordovician

REMOTE SENSING DATA COLLECTED FOR THE PENNSYLVANIA TEST SITE

Remotely sensed data can be collected by ground-based, airborne and (ultimately) satellite-borne instruments, for example, from an elevated platform just above the field, from an aircraft aloft, and from a satellite in earth orbit. In this study, data were obtained from an airborne sensor and the following discussion is limited to that aspect of remote sensing data acquisition.

Currently, multispectral imagery of the terrain is obtained for LARS by the Willow Run Laboratories, University of Michigan, using a multispectral optical-mechanical scanner mounted on their DC-3 aircraft. Details on the airborne data acquisition are available in papers by Landgrebe and Phillips (2) and Hoffer and Goodrick (1).

The specific multispectral bands recorded on aircraft storage tapes vary according to the needs of the researcher. In this study, a total of fifteen channels of data were obtained, thirteen in the visible and reflective infrared portion of the spectrum (between 0.4μm and 2.6μm) and two in the thermal infrared (4.5-5.5μm and 8.0-14.0μm).

The flightline was flown May 15, 1969, at 3000 feet above the average terrain elevation. The data were recorded on two tapes, Tape One for the southern half and Tape Two for the northern half of the flightline.

Each segment was flown both during the day and at night. For the daytime flights, visible, reflective infrared and thermal infrared data were obtained. At night measurements were limited to the thermal infrared.

In addition to the aircraft data tapes, other information made available to LARS included: black and white 9" x 9" aerial photos plus an aerial photographic mosaic for the flightline; topographic maps for the area and agricultural soils maps for Berks and Lancaster Counties; cathode-ray tube imagery of three wavelength bands (one each in the visible, reflective infrared and thermal infrared); 70mm color photography; and 9" x 9" color and color infrared photography, available on a loan basis.

LARS TECHNIQUES FOR MULTISPECTRAL DATA ANALYSIS

The following section is intended to familiarize the reader with the current LARS analysis techniques for multispectral remote sensing data. Owing to space limitations this discussion has been abbreviated; for previous descriptions of the techniques see papers by Tanguay and Miles (5); Tanguay, Miles and Hoffer (6); Hoffer and Goodrick (1); and West (7).

Remote sensing involves the identification and classification of surfaces through analysis of data from sensing devices not in direct contact with those surfaces. At LARS these data are analyzed on a digital computer using multivariate pattern recognition techniques. Currently, imagery collected in the visible through thermal infrared portions of the spectrum are included in pattern recognition studies.

Following the multispectral remote sensing flights, the aircraft scanner data tapes are forwarded to LARS. By means of the LARS data-handling system, these analog data tapes are converted to digital form. Each analog scan is sampled, normally at a sampling rate which yields 220 data points for an 80° field of view across the flightline. For later reference, each digitized data point is assigned a unique address in a two-dimensional coordinate system based on scan line numbers ("line numbers") and samples within the line ("column numbers").

A final preprocessing function, data overlaying, must precede analysis of imagery from multiple-aperture scanners. In 1969, when these data were obtained, the University of Michigan scanner system collected data from three apertures (visible, reflective infrared, and thermal infrared). In such cases, data overlaying is performed to align the data so that pattern recognition can be based on all channels. Check points, such as field corners and highway intersections, are located on gray scale printouts (described below) for channels obtained through different apertures. After an array of such check points is established throughout the flightline, the computer aligns these points and forces the data between them to line up as well as possible. Typically, the discrepancy remaining after data overlaying is at most two or three resolution elements. A resolution element or remote sensing unit (RSU) is the area on the ground represented by a single reflectance symbol or letter on a gray scale printout. At the 3000-foot altitude of the Pennsylvania flight, this represents about a 20-foot square on the ground.

The computer programs used at LARS to analyze multispectral scanner data are outlined in Figure 2. Indicated are the sequential and alternative steps involved in the analysis procedures. This figure constitutes an updating of a similar diagram presented in the paper by Tanguay and Miles (5).

Typically the initial step in the analysis is to obtain gray scale printouts of the data for several channels ($PIC in Figure 2). Gray scale printouts are digital displays of the spectral response of the terrain but limited to one band or channel of the scanner data per display. They resemble low resolution photographs. Alphanumeric symbols (letters of the alphabet, typing symbols and single-digit numbers) are assigned to the radiance levels so that high-intensity or light areas are printed with symbols which cover a low proportion of the paper (example: ---), and dark areas are printed with symbols with a high proportion of cover (example: MMM). Because the data are not equally distributed between the highest and lowest radiance values, the symbols are not simply assigned so that each represents an equal range of radiance. Instead, to achieve a maximum contrast display, the data are first histogrammed and the symbols are assigned so that each will occur with approximately the same frequency.

Areas of known materials are located on the gray scale printouts by the researcher and their addresses recorded. This information, referred to as "ground truth," is used to train the computer to recognize similar material. In agricultural studies this training information may consist of fields containing corn, wheat, oats, soybeans, etc.: in forestry studies, conifers and deciduous trees or individual tree species; in geology and highway engineering, specific bedrock and soil types. While some of the ground truth areas are used to train the computer to recognize the classes of interest, the remaining areas are reserved for testing the acuuracy of the computer classification after it is completed. Addresses of training areas are provided to the computer on punched cards; these addresses are boundary corners indicated by column and line number, and the areas are limited to rectangular shapes (commonly referred to as "fields").

After January 1, 1971, an alternative to the use of gray scale printouts became available at LARS. This tool, known as the digital display unit (FIELDSEL) provides a television-like image of the scanner data with each digital value represented by a different brightness level. This yields an image on the screen having greater detail than is possible with gray scale symbols on computer paper. Images for each channel can be displayed, and fields of interest can be outlined on the screen using a light pen with their addresses automatically punched on cards.

Another method for obtaining training field sites is sometimes used in conjunction with the above procedure of manually selecting

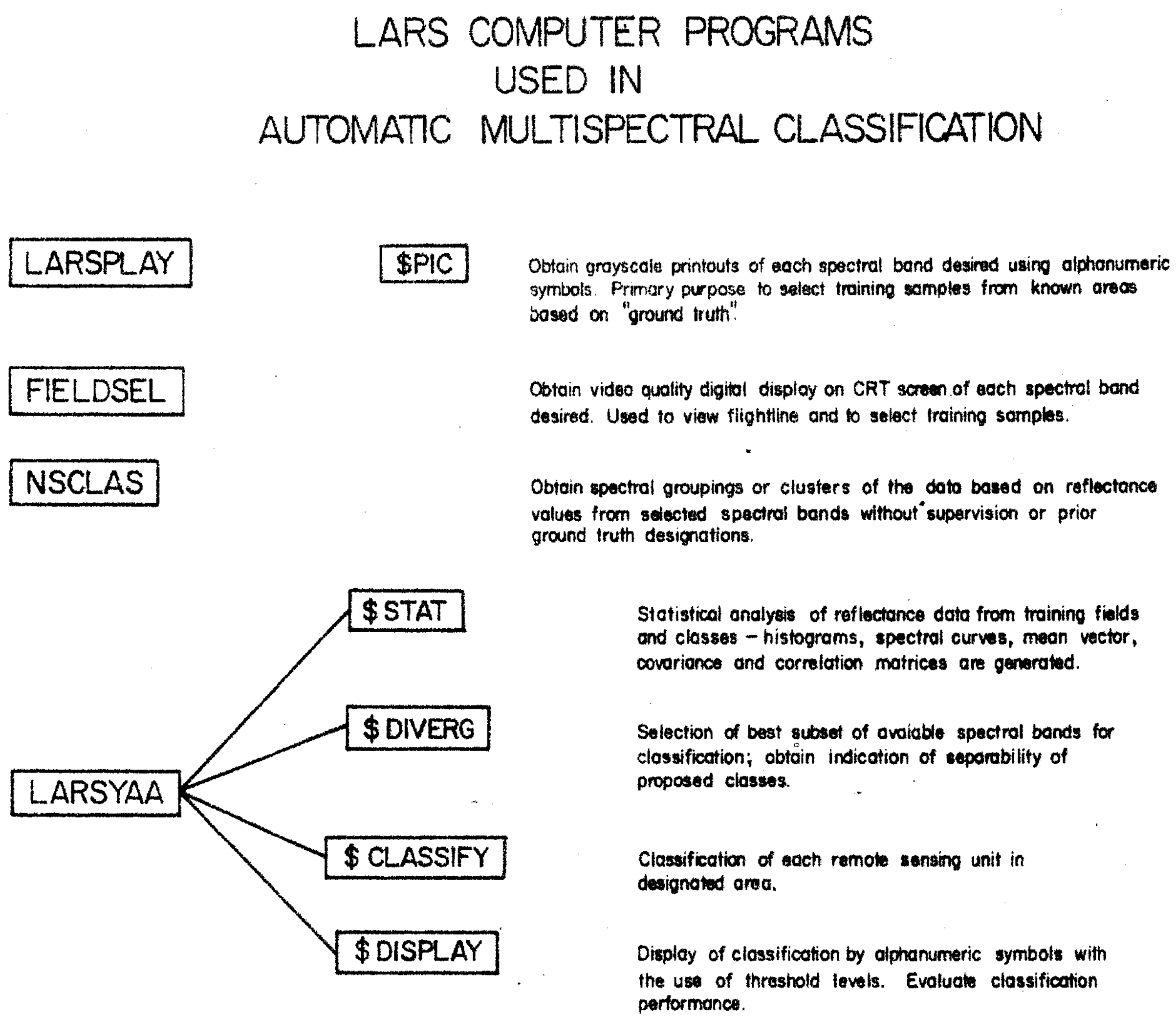

Figure 2. Computer programs for classification of multispectral scanner data.

these sites. Accomplished by using the clustering program NSCLAS (non-supervised classifier, see Figure 2), this relatively new technique divides the scanner data into groups or clusters based on similarity of spectral response within clusters. Typically four to six channels of imagery are analyzed simultaneously and the program is requested to obtain ten clusters. A map of the results is printed for each area analyzed in this way and the researcher can observe the patterns of spectrally differentiable material occurring in the data. The training fields may then be selected from these maps. The resulting clusters may not in fact be spectrally distinct, but the researcher must decide this to his own satisfaction based on separability information for the clusters which is printed by the program. He then has the option of repeating the analysis using a different number of clusters to increase their separability. An important point is that this method of selecting training fields takes into detailed account the multispectral response data as well as the ground truth information.

In the clustering approach, the actual differences in spectral response are displayed, but a major problem exists in determining what each of the clusters represents. Some may be vegetation types, some water bodies, man-made features or tonal aspects of bare soil and rock. Aerial photography is helpful in affixing names to the patterns observed. Despite this difficulty, clustering is a powerful tool in obtaining workable training fields. The clustering technique, which was under development during the analysis phase of the work reported here, was not applied to this data. It was subsequently used for detailed soils mapping of the Pennsylvania data (Mathews et al., 3) and in a brief study in Indiana (West, 8).

The next step in the analysis is to obtain histograms for each class of material identified in the training fields ($STAT). An example of classes for a geologic study area might be alluvium, limestone soil, shale soil, trees, mixed crops, and water. The histograms for each class show the distribution of reflectance intensity for each spectral channel.

Unimodal or single-peaked distributions in the histograms for a class suggest that the proposed class is spectrally an individual group. Bimodal or trimodal distributions must be subdivided manually into unimodal classes by the researcher; histograms for individual fields can be obtained to help in locating the multimodal contribution within a class.

The next operation involves the application of a divergence (statistical separability) analysis (known as $DIVERG) to determine the best channels to use for classification. Only the best four to six channels are used for classification in order to save computer time; in general, the accuracy of classification is not meaningfully increased when more channels are added. In addition to indicating the preferred channels, the divergence analysis indicates the separability of the designated classes. If separability between significant materials is poor, some of the preceding steps may be repeated in an attempt to improve this separability and hence classification accuracy.

Next the training field statistics are used to classify the designated portion of the flightline ($CLASSIFY). The computer classifies each data point (RSU) based on a maximum likelihood criterion. On request, the computer calculates how accurately it classified the areas used for training by comparing the classification of each point in the training fields with the initial ground truth designation. A high level of agreement means there is minimal confusion within the training field statistics of the various materials and that the classes are being separated properly.

The final steps are to print out a computer classification map for the whole area ($DISPLAY) and to determine how well the test fields were classified. The test fields are those areas of known material from ground truth studies that were not previously used for training purposes. If test fields and training fields show a high degree of accuracy and the test fields are representative of the entire area, the classification is a good one. If the accuracy is low, some reworking of the classes should be done. The researcher may also have to conclude that the classes involved are not spectrally separable.

RESULTS OF STUDY FOR THE BLUE BALL AND REAMSTOWN-DENVER AREAS

Initial work during the contract period was performed on data from the Blue Ball, Pennsylvania area. Prior to this, a brief study had been made on a 12-mile section immediately to the south of Blue Ball near Welsh Mountain, a prominent quartzite, tree-covered ridge. This preliminary study, set up to determine the potential of the LARS system for engineering soils was most rewarding in that it suggested an improved approach for analyzing the total flightline.

The Welsh Mountain study was complicated by the diverse geologic and topographic nature of the area. It is a complex region of the Piedmont Plateau Province displaying up to 500 feet of relief due to differential erosion of rocks with contrasting resistance to weathering. Soils derived from limestone, schist, quartzite, and a mixture of these, talus or colluvium, mantle the area.

Because of these complexities, very accurate ground truth information was necessary for training the computer to distinguish between the various materials. Such detailed information, as later determined, was unavailable. In areas of high relief composed of resistant rocks, it is common for pieces of durable rock to be moved down slope by gravity, accumulating as a mixture of materials at a lower level. This colluvium displays some of the spectral characteristics of all the parent materials involved. As the extent of colluvium exposed at the surface was not known, it was omitted from the training samples.

In the computer-generated classification of the Welsh Mountain area, vegetation and water were accurately delineated; however, classification of the soil types did not agree well with the known geological conditions of the area. Strips of rock types (or soils derived from them) were designated in a manner that paralleled the man-made field boundaries. This is not the actual situation at the site; in reality, individual rock types appear as broad bands several miles wide rather than as narrow, repetitive strips measured in tens of feet. Because of these inaccuracies in this complex area it was decided to concentrate on pure soil areas where the complications are minimized. This supplied the impetus to use the pure soils area approach for the Pennsylvania flightline.

The first pure soil area studied, located near Blue Ball, is 3.6 miles long and approximately 5000 feet wide. It lies near the southern end of the flightline in Lancaster County and is designated as pure soil areas 1 and 2 in Table 1. The following discussion, a step-by-step description of the manner in which the final Blue Ball classification was accomplished, is presented as an example of the details involved in obtaining an accurate classification.

This flightline segment consists primarily of residual limestone soil with some river alluvium located along a small creek. It was eventually determined that three conditions exist in this residual limestone soil: 1) eroded areas where the subsoil is exposed, 2) non-eroded areas, 3) places where local alluvium derived from erosion has accumulated. These combined conditions (which can be observed on the color photos and to a lesser extent on the gray scale printouts of the multispectral scanner data) made analysis of the Blue Ball site difficult. Visible and reflective infrared channels were included in the following analysis.

The first classification of the Blue Ball area was based on ten fields outlined on airphotos that were suggested for training by the FHWA photointerpreter. Training fields in the river alluvium along the creek, which had not been represented previously, were added, and the Lancaster County soils map provided ground truth information. Early results suggested that more spectral classes of limestone soil were present than were represented by the training fields. In the subsequent refinement, many additional limestone training fields were added from the entire area in an attempt to represent all possibilities. Test fields were selected at the same time.

For this next classification, results still indicated difficulty in differentiating river alluvium, local alluvium and limestone soil. Field checks were made by staff from the Department of Agronomy, Pennsylvania State University, at LARS' request. This field study indicated that the agricultural soil map was not entirely accurate, and this new information was used to discard some training fields, regroup others, and add new ones, specifically categories for eroded limestone and local alluvial accumulation. A series of classifications was made, deleting fields that proved troublesome and adding several new factors such as training fields from a limestone quarry. Approximately 75 limestone training fields were used in the final classification. The training fields were evaluated after the classification was completed, and these results (over 97% correct) are presented in Table 2. The final classification of the area appears as Figure 3.

Following the classification, 58 test fields in the limestone soil were used to determine the accuracy of the classification. For these fields, which comprised 3340 data points or RSU's, an accuracy of 90.6% was obtained.

In the process of improving the classification, several class divergence analyses were made. As a result, the 75 limestone fields were eventually combined into 14 separable classes. Several of the final classes were relatively similar, but collectively were sufficiently different so that if combined, would yield a combined class too broad for useful classification. Hence the 14 classes were maintained as separate entities.

The divergence analysis identified the four best channels for classification. As vegetation types are not an important aspect in this study, the vegetation classes were omitted in the divergence analysis. This meant that only separability between different soil groups would be a factor in determining which channels to use. Maximum significance was assigned to distinguishing between limestone, local alluvium, river alluvium, and eroded soil. This resulted in the selection of channels 2, 7, 12 and 13 as the four best channels for classification (.44-.46, .58-.62, 1.00-1.40, and 2.00-2.60μm respectively).

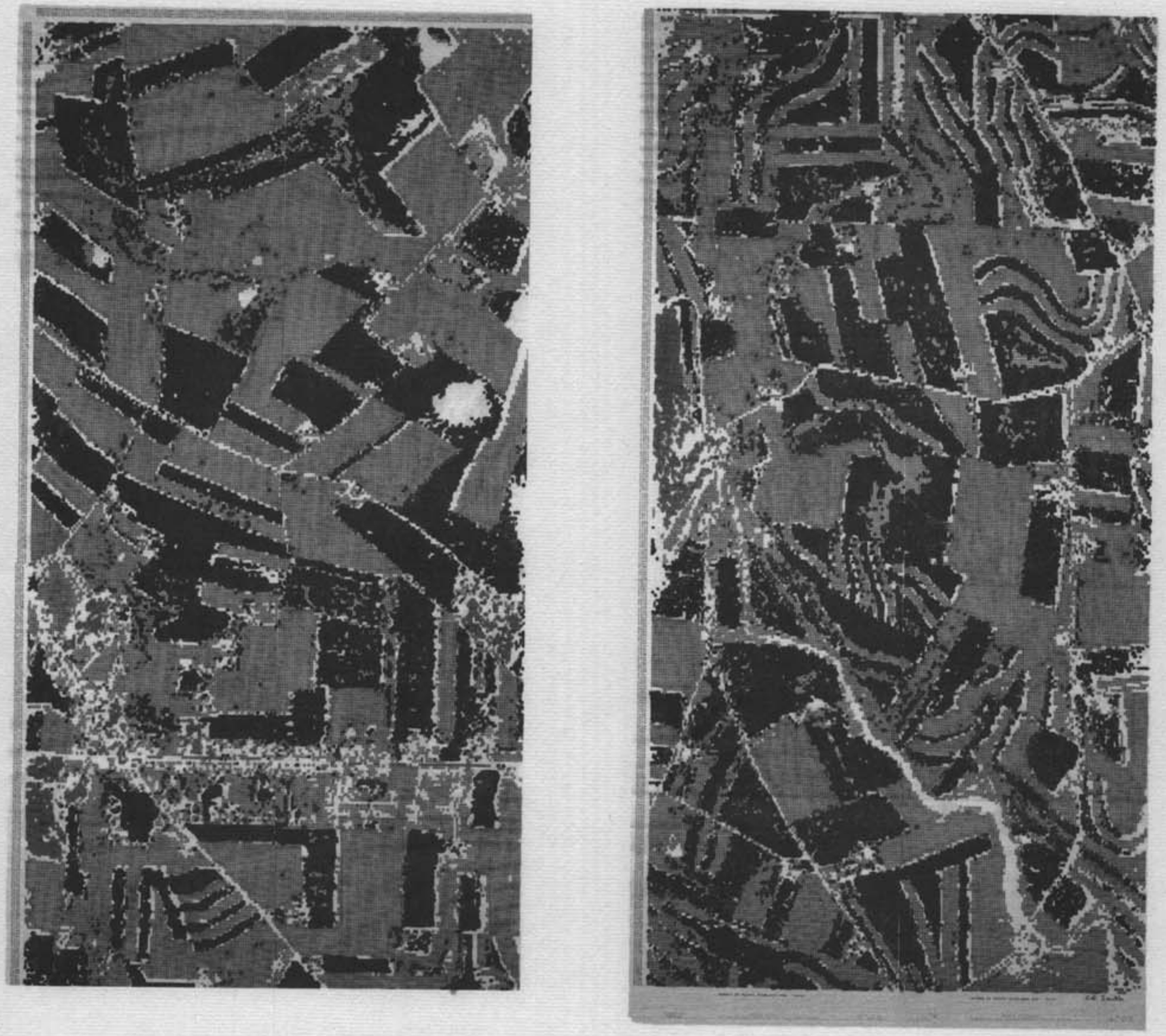

Figure 3. Classification Printout for Blue Ball, Pa.

TABLE 2. TRAINING FIELD PERFORMANCE BLUE BALL, PENNSYLVANIA

	Group	No. of Samples	Pct. Correct	Number of Samples Classified Into: Veg	Lime	Alv	Loclv	Rival	Erod	Threshold*
1.	Vegetation	519	100.0	519	0	0	0	0	0	0
2.	Limestone Soil	3586	97.4	16	3491	3	62	1	0	13
3.	Alluvium	43	97.7	0	0	42	0	1	0	0
4.	Local Alluvium	86	96.5	0	3	0	83	0	0	0
5.	River Alluvium	149	98.7	0	1	1	0	147	0	0
6.	Eroded Limestone Soil	32	96.9	1	0	0	0	0	31	0
		4415		536	3495	46	145	149	31	13

Overall Performance 4313/4415 = 97.7%

Average Performance by classes 587.1/6 = 97.8%

*No classification made (data very unlike any training class)

A third divergence analysis, identical to the preceding except for the deletion of channels 12 and 13, was then performed in order to determine if channels 12 and 13 contributed enough toward class separability to warrant their inclusion. Since these channels are obtained from a scanner aperture different from channels 1 through 11, these upper channels are significantly involved in the overlay problem, with the misalignment ranging up to three resolution units despite overlay adjustments. Results indicated a marked reduction in separability of some key classes when channels 12 and 13 were omitted, signifying that they were needed for proper separation. Thus channels 2, 7, 12 and 13 appeared best for the classification.

The Reamstown-Denver area, designated in Table 1 as pure soil area 4, was the second pure soil area studied. Reamstown is located 7 1/2 miles northwest of Blue Ball, and Denver lies 2 miles beyond Reamstown in the same direction. The Reamstown-Denver pure soil study area is about 4 1/2 miles long and 5000 feet wide.

This study area contains several soils which are derived from very different parent materials (bedrock). Much of the area is underlain by Ordovician age limestone which yields a limestone soil similar to that found in the Blue Ball area. The limestone soil conditions in Reamstown-Denver are simpler, however, since the severely eroded areas are absent.

Present also in the Reamstown-Denver area are 1) soils derived from dark-gray Ordovician age shales and 2) soils derived from Triassic age red-colored shales, sandstones and conglomerates. In addition, river alluvium and river terrace materials are present. The water and vegetation classes are similar to those in the Blue Ball area.

Training samples were selected from the different soil types in the Reamstown-Denver site, combined into classes, and used to classify the existing surface materials. Twenty-three classes consisting of 76 fields were used. The same channels used in the Blue Ball classification, 2, 7, 12 and 13 (.44-.46, .58-.62, 1.00-1.40 and 2.00-2.60μm respectively), were used. Several classifications were required since adjustments for the troublesome training fields were needed before an acceptable classification was obtained for the area. The training fields were evaluated for the Reamstown-Denver classification and these results (98% accurate) are presented in detail in Table 3.

Following the completion of the Reamstown-Denver classification, the training classes for both the Blue Ball and the Reamstown-Denver areas were combined yielding 43 classes and approximately 150

TABLE 3. TRAINING FIELD PERFORMANCE REAMSTOWN-DENVER, PENNSYLVANIA

Group	No. of Samples	Pct. Correct	Number of Samples Classified Into: Quarry	Turn-Pike	Roof	Water	Allu-vium	Terrace	Lime	Shale	Triassic
1. Quarry	142	97.2	138	1	3	0	0	0	0	0	0
2. Turnpike	27	100.0	0	27	0	0	0	0	0	0	0
3. Roof	53	98.1	0	0	52	0	0	0	0	1	0
4. Water	104	100.0	0	0	0	104	0	0	0	0	0
5. Alluvium	70	92.9	0	0	0	0	65	0	0	1	4
6. Terrace	194	98.5	0	0	0	0	2	191	1	0	0
7. Duffield (Limestone Soil)	534	99.3	0	0	0	0	1	1	530	0	2
8. Berks (Shale Soil)	224	99.6	0	0	1	0	0	0	0	223	0
9. Penn (from Triassic red beds)	165	94.5	0	0	0	0	5	0	4	0	156
	1513		138	28	56	104	73	192	535	225	162

Overall Performance 1486/1513 - 98.2%

Average Performance by Classes 880.0/9 = 97.8%

training fields. The area from Blue Ball continuously through to Denver, a distance of 11.6 miles, was classified using the resulting statistics. Included in this 11.6-mile segment is pure soil area 3, Independence School. This area of Triassic age shale, sandstone and conglomerate was mapped in this manner. The results obtained for area 3 are realistic and quite good considering that no training samples were taken directly from that area.

For the 11.6-mile segment the training field performance averaged 95.6% correct. This high level is somewhat misleading in regard to the actual accuracy as no test fields were evaluated and no training fields from the Independence School area were evaluated. A problem that can be observed is the confusion between shale-derived soils in Reamstown-Denver and the limestone soils in Blue Ball. This misclassification is related to the presence of transition local alluvium and limestone soil. Despite these difficulties, the extended classification is a significant step toward the objective of engineering soils mapping for large areas of a flightline using the LARS computer techniques.

CONCLUSIONS AND FUTURE WORK

Results from the two pure soils areas, Blue Ball and Reamstown-Denver, indicate that engineering soil materials can be mapped at the Pennsylvania test site in considerable detail when adequate ground truth is available. Vegetation type, water, roads, roof tops and quarries can also be discerned. Vast amounts of detail are available from multispectral analyses, much of which is due directly to soil properties and conditions. For example, locations of wet soil and areas where soil cover over bedrock is thin may be outlined. Hence a subsequent accomplishment may be the determination of engineering soils properties within a soil unit after these units have been delineated by a less detailed analysis.

Classification of the 11.6-mile segment from Blue Ball to Reamstown-Denver marks the first step toward classifying longer portions of a flightline by extending classification from smaller segments. This could possibly be extended to achieve a classification of the southern half of the flightline if additional training fields from the unrepresented soil types were obtained. The achievement of a combined flightline map was an objective of the study with the pure soils area concept initiated in hopes of accomplishing it.

Two problems arise when mapping a large area which contains numerous spectral classes of material. First, there is a physical limit on the number of classes that can be processed by the computer programs, this being primarily a function of computer memory limitations. The current limit at LARS is sixty classes, and the 11.6-mile classification with 43 classes approached this limit. Second, accurate classification becomes more difficult when numerous material types are considered. Simply stated, the probability of having overlapping spectral characteristics of materials is increased when greater numbers of similar materials are included.

A new approach now being explored is to include in the first classification attempt only the soil types that are anticipated in the extended area. For example, soils derived from metamorphic rocks such as schist, quartzite and slate would not be included as training classes for an area derived from sedimentary bedrock. An initial subdivision into broad soils groups along the flightline would be made first. The cluster analysis would be applied to delineate these general groups, removing the human bias in training field designation. Detailed analysis of each general group using typical classes of material for that area would follow.

ACKNOWLEDGEMENTS

This research project was sponsored by the Federal Highway Administration in cooperation with LARS and its primary funding agency, the National Aeronautics and Space Administration (NASA). The on-going project is a portion of the Cooperative Remote-Sensing Research Program of FHWA monitored by Dr. H. T. Rib, whose efforts in coordinating the research are hereby acknowledged. Special thanks are afforded to Dr. R. L. Cunningham and Mr. H. L. Mathews, Department of Agronomy, Pennsylvania State University, who aided in the field identification of surface materials, and to Dr. P. H. Swain, of LARS, who critically reviewed the manuscript.

REFERENCES CITED

1. Hoffer, R. M. and F. E. Goodrick . 1971. "Geographic Considerations in Automatic Cover Type Identification," Proceedings of the Indiana Academy of Science, Vol. 80.

2. Landgrebe, D. A. and T. L. Phillips. 1967. "A Multichannel Image Data Handling System for Agricultural Remote Sensing," Computerized Imaging Techniques Seminar, Washington, D. C.

3. Mathews, H. L., R. L. Cunningham, T. R. West and Jan E. Cipra. 1971. "Application of Multispectral Remote Sensing to Soil Survey Research in Southeastern Pennsylvania," paper presented at American Society of Agronomy, New York (in preparation).

4. Rib, H. T. and R. D. Miles. 1968. "Multisensor Analysis for Soils Mapping," Highway Research Board Spec. Report 102.

5. Tanguay, M. G. and R. D. Miles. 1970. "Multispectral Data Interpretation for Engineering Soils Mapping," Highway Research Record 319, Highway Research Board.

6. Tanguay, M. G., R. D. Miles and R. M. Hoffer. 1969. "Multispectral Imagery and Automatic Classification of Spectral Response for Detailed Engineering Soils Mapping," Sixth Symposium on Remote Sensing, Ann Arbor, Michigan.

7. West, T. R. 1970. "Applications of Remote Sensing to Highway Locations," Proceedings of the 21st Highway Geology Symposium, Lawrence, Kansas.

8. West, T. R. 1971. "Engineering Soils Mapping in Indiana by Computer from Remote Sensing Data," Indiana Academy of Science, Vol. 81 (in press).

ADDITIONAL REFERENCES: LARS MULTISPECTRAL DATA ANALYSIS TECHNIQUES

1. Anuta, P. E., S. J. Kristof, D. W. Levandowski, R. B. MacDonald and T. L. Phillips, 1971. "Crop, Soil and Geologic Mapping from Digitized Multispectral Satellite Photography." Seventh International Symposium on Remote Sensing of Environment Ann Arbor, Michigan.

2. Atwell, R. B., R. B. MacDonald and L. A. Bartolucci. 1971. "Thermal Mapping of Streams from Airborne Radiometric Scanning," Water Resources Bulletin 7(2):228-242.

3. 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," Proceedings of the 7th International Symposium on Remote Sensing of Environment, University of Michigan, Ann Arbor, Michigan.

4. Baumgardner, M. F., L. F. Juggins, R. B. MacDonald and R. M. Peart. 1967. "Applicability of Remote Sensing and Systems Analysis Techniques to River Basin Planning," Proceedings of the Third Annual American Water Resources Conference, San Francisco, California.

5. Baumgardner, M. F., S. J. Kristof, C. J. Johannsen and A. Zachary. 1970. "Effects of Organic Matter on Multispectral Properties of Soils," Proceedings of the Indiana Academy of Science, 79:413-422.

6. Fu, K. S., D. A. Landgrebe and T. L. Phillips. 1969. "Automatic Reduction of Remotely Sensed Earth Resources Data," Remote Environmental Sensing Issue of the Proceedings of the Institute of Electrical and Electronic Engineers.

7. Hoffer, R. M., R. A. Holmes and J. R. Shay. 1966. "Vegetative, Soil, and Photographic Factors Affecting Tone in Agricultural Remote Multispectral Sensing," Proceedings of the Fourth Symposium on Remote Sensing of Environment, Ann Arbor, Michigan.

8. Hoffer, R. M. and C. J. Johannsen. 1969. "Ecological Potentials in Spectral Signature Analysis," Remote Sensing in Ecology, P. Johnson, ed., University of Georgia Press, Athens, Georgia. pp. 1-16.

9. Hoffer, R. M. and F. E. Goodrick. 1971. "Variables in Automatic Classification over Extended Remote Sensing Test Sites," Proceedings of the 7th International Symposium on Remote Sensing of Environment, University of Michigan, Ann Arbor, Michigan.

10. Holmes, R. A. 1968. "Data Requirements and Data Processing in Earth Resources Surveys," Thirteenth Annual Technical Symposium of Society of Photo-Optical Instrumentation Engineers, Washington, D. C.

11. Holmes, R. A. and R. B. MacDonald. 1969. "The Physical Basis of System Design for Remote Sensing in Agriculture," Remote Environmental Sensing Issue of the Proceedings of the Institute of Electrical and Electronic Engineers.

12. Kristof, S. J. 1968. "Spectral Soil Mapping," Proceedings of Indiana Academy of Science of 1967.

13. Kristof, S. J. and A. L. Zachary. 1971. "Mapping Soil Types from Multispectral Scanner Data," Proceedings of the 7th International Symposium on Remote Sensing of Environment, University of Michigan, Ann Arbor, Michigan.

14. Landgrebe, D. A., P. J. Min, P. H. Swain and K. S. Fu. 1968. "The Application of Pattern Recognition Techniques to a Remote Sensing Problem," Seventh Symposium on Adaptive Processes, Los Angeles, California.

15. MacDonald, R. B. 1969. "Space Age Sensing for Agriculture," Farm Technology.

16. MacDonald, R. B. and D. A. Landgrebe. 1967. "Remote Sensing for Agriculture and Natural Reosurces from Space," Proceedings of National Symposium of American Astronautical Society, Huntsville, Alabama.

17. Shay, J. R. 1966. "Some Needs for Expanding Agricultural Remote Sensing Research," Proceedings of Fourth Symposium on Remote Sensing of Environment, Ann Arbor, Michigan.

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

19. Swain, P. H. and D. A. Germann. 1968. "On the Application of Man-machine Computing Systems to Problems in Remote Sensing," Eleventh Midwest Symposium on Circuit Theory, Notre Dame, Indiana.