

Automatic Classification of Soils and Vegetation with ERTS-1 Data

by
D. A. Landgrebe

The Laboratory for Applications of Remote Sensing
Purdue University
West Lafayette, Indiana

Automatic Classification of Soils and
Vegetation with ERTS-1 Data*

by

David Landgrebe

ERTS-A was successfully launched to become ERTS-1 one July 23, 1972. This was an historic event, though only a beginning, for it climaxed a long effort by many people to conceive, design, and bring into being such a data-gathering instrument. In parallel with this effort research had been under way to find methods to analyze data from such an instrument. One part of this latter effort was a search for ways to analyze such data by computing machine. These would be needed for problems requiring very rapid analysis or high data throughput, or both. It was only natural that such an analysis approach should be quickly tested on ERTS-1 data. In this way not only could the new analysis method be tested, but, more importantly, a concrete indication could quickly be obtained of the operating characteristics of ERTS-1 and the probable utility of its data for resource mapping.

It is the purpose of this brief report to present some of the results of this test. The analysis method used was a four-spectral-band, supervised, maximum likelihood, Gaussian classifier with training statistics derived through a combination of clustering and manual methods. This multispectral procedure is a multivariate analysis method which leads to the assignment of each resolution element of the data to one of a preselected set of discrete classes.

The data frame selected by NASA for this test was from an area over the Texas-Oklahoma border of the USA as shown in Figure 1. This data in the form of computer-compatible multispectral scanner tapes arrived at our analysis Laboratory

*Presented at the 23rd International Astronautical Congress, Vienna, Austria October 14, 1972. This paper is an abbreviated form of a more complete report of this experiment contained in LARS Information Notes 092972 and a presentation contained in "Preliminary Findings from Analysis of ERTS Observations", given at the NASA Goddard Space Flight Center, Greenbelt, Md., September 29, 1972. Work reported here was sponsored by the National Aeronautics and Space Administration under Grant NGL-15-005-112 and Contract NAS5-21773.

on the night of July 26, the day after it had been collected. The analysts had no ground observational information at that time and no familiarity with the area. Preliminary results were available 48 hours later. Following this a two-day ground observation mission of the area was conducted by light airplane and automobile. This was done to verify the results and to obtain information from which the results could be finalized. Revised results were available shortly thereafter and are summarized as follows.

The total experiment was carried out in several subprojects, the first of which was an analysis of the full frame into spectrally separable land use categories (Figure 2). Figure 2 shows a simulated color infrared photo of a large portion of the frame for orientation purposes. The Red River flows through the center of the frame and at one point has been dammed to form Lake Texoma. The Ouachita Mountains are in the northeast.

Figure 3 shows the analysis results for a portion of this area. Each resolution element of the data has been classified into one of 17 possible classes; the results are displayed here by associating a color with a class or group of classes. Eleven colors have been used for the 17 classes as follows:

•Yellow, tan, brown and light green	Rangeland and pastureland classes
•White and light gray	Sandy, bare soils and light vegetation and agriculture fields with sparse canopy
•Dark green	Forest and woodlots
•Aqua, blue-gray, dark blue and dark purple	Four classes of water

Figure 3 constitutes a preliminary form of a land use map. While viewing it one feels urged to draw boundaries around different areas and anotate them. Equally significant is the fact that at the same time this result was obtained, a computer tabulation was also obtained which showed that of the 3,425,300 hectares included in the full frame, 48.4% were classified into range and pasture land, 31.7% in cropland, 17.6% in forest, 1.4% in muddy water, 0.9% in clear water, etc.

Figure 3 gives a good overview of the type of information which can be derived in this manner. However, in order to better understand the detail possible with this data, two subframe areas were selected for study. Dr. Roger Hoffer,

in an earlier session of this Congress*, already reported on the results obtained in the Ouachita Mountain area in the northeast (Figure 4). He reported that it was possible to discriminate between five different categories of forest cover including one recently sprayed with a defoliant to kill non-productive trees in order to improve pasture productivity (Figure 5). Agricultural and pasture lands and water were also classified.

The second subframe (Figure 6) was selected in the Lake Texoma area. Special emphasis here was put on subcategories of water. Note the difference in appearance between Lake Tishomingo in the north and Lake Texoma. In this case it was possible to divide the data into 18 spectrally separable classes (Figure 7). They are grouped and displayed here in non colors as follows:

- Blue-gray, light blue, aqua, medium blue, and dark blue indicate subclasses of water;
- White indicates soil, light dry vegetation over soil, and crops with sparse canopy;
- Red shows crops with heavy canopy;
- Light green indicates forest with sparse canopy; and
- Dark green indicates forest with heavier canopy.

Notice that Lake Tishomingo, at the very top of the figure, was classified into the aqua, light blue and blue-gray classes. These are associated with shallow and muddy water. Lake Tishomingo is no more than one meter deep in most places and is muddy in appearance.

Lake Texoma was also classified into several water sub-categories. The west end (Figure 8) shows a hook-shaped delta where the Red River enters the lake. The aqua area adjacent to the delta is a portion of the delta barely covered with water. It is less than one meter deep in much of this area and is visibly muddy. The medium blue and dark blue areas represent relatively clear water with the dark blue tending to be in the deeper areas. This association is not always true, however, and may have been affected by surface water action. The line structure in the lake is due to the fact that the six sets

* "Agricultural and Forest Resource Surveys from Space", by R.M. Hoffer. Presented at the 23rd International Astronautical Congress, Vienna, Austria.

of detectors in the spacecraft scanner did not have exactly the same system gains.

Figure 9 shows a low-altitude air view of the delta area as seen from the north looking south. The hook shape of the dry portion can readily be seen. Gradations in the water as one moves out from the delta also appears to conform well to the classification.

Figure 10 shows a photo of the highway bridge taken from about one kilometer north of the north shore. The bridge is no more than 20 meters wide, well less than the resolution of the scanner system. A correct classification of the bridge was possible only because of its high reflectivity with the lake as background.

In summary, this study suggests that multispectral scanner data such as ERTS-1 produces, coupled with machine processing, shows promise for surveys of earth surface cover, although further evaluation of accuracy and the diversity of classes possible will be needed. Not only can maps be produced, but so can quantitative information about the amount of ground cover of a given type in a given area. This information is obtained for rectangular agricultural fields and for irregular forested areas and water bodies with equal ease.

But perhaps most important is the speed of this process. Processing speed is important not only for coping with the data volume, but because processing time is directly related to the cost of processing. If need be, the analysis of a full frame can presently be completed within 48 hours. The cost, for example, in computer time and manpower for producing the full frame land use classification is such that the total cost for this portion of a land use survey amounts to about .0075 cents/hectar (.003 cents/acre).

And finally I wish to congratulate the ERTS system designers, builders, and operators. The data it produces surpasses our expectations. Only by viewing the imagery and working with the data can one experience the real excitement that this satellite produces.



Figure 1.



Figure 2.

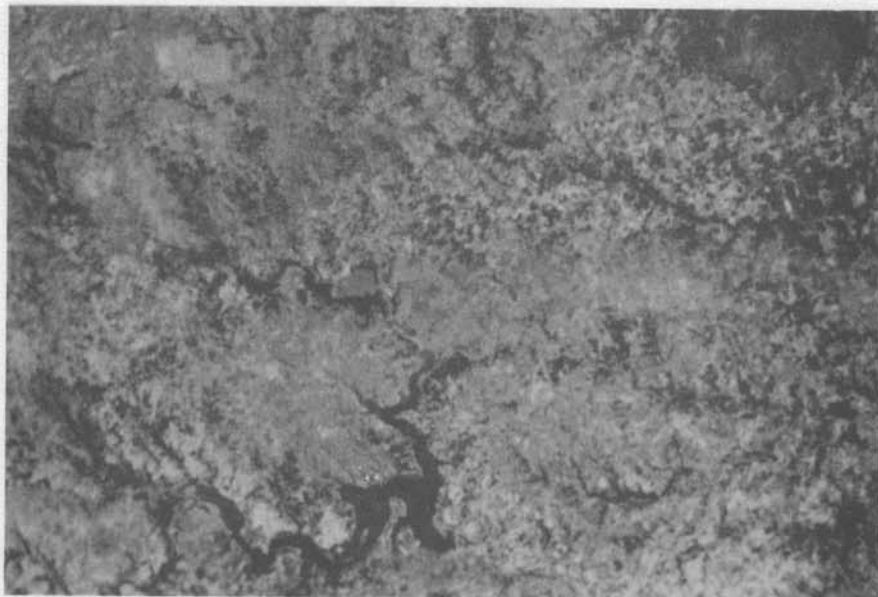


Figure 3.

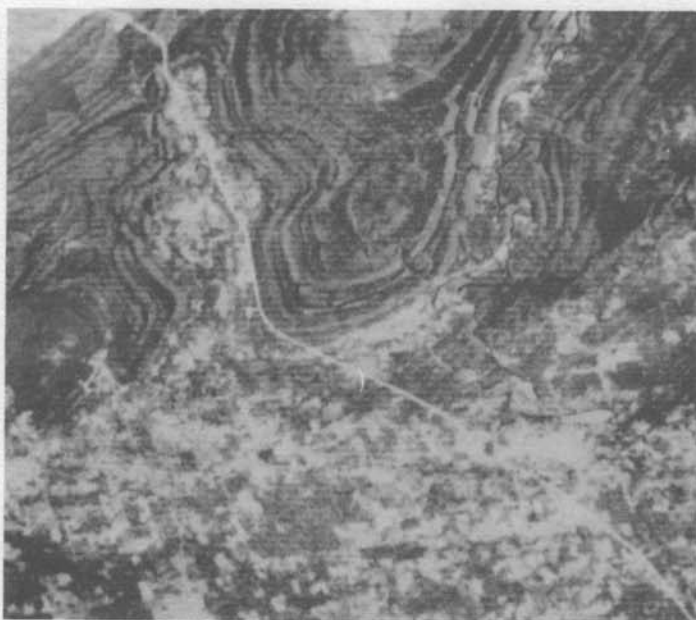


Figure 4.



Figure 5.



Figure 6.

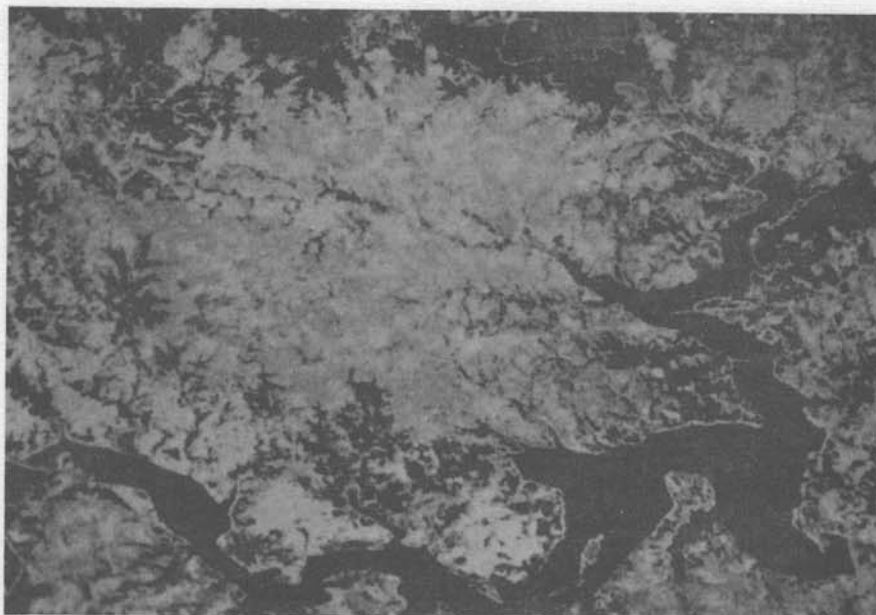


Figure 7.



Figure 8.



Figure 9.



Figure 10.