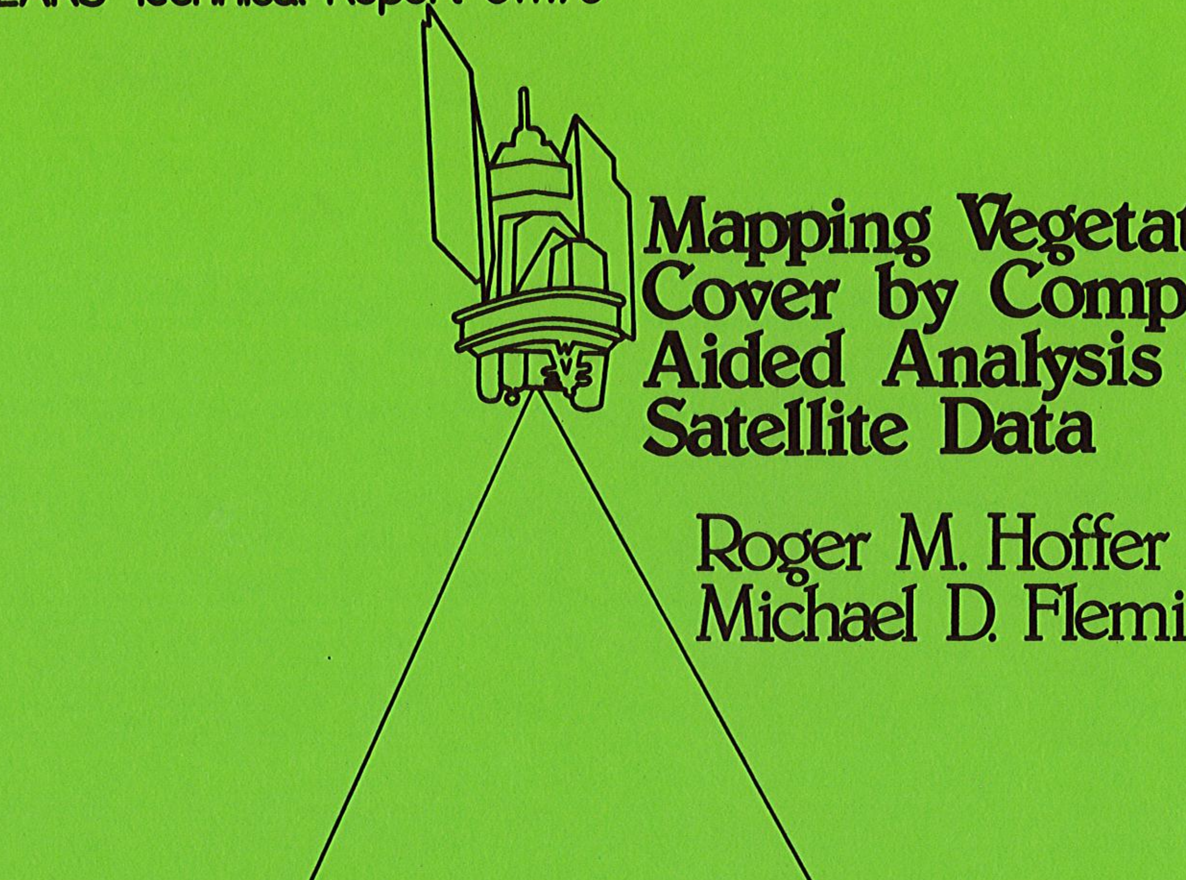
LARS Technical Report 011178

Mapping Vegetative Cover by Computer-Aided Analysis of Satellite Data

Roger M. Hoffer
Michael D. Fleming



Department of Forestry and Natural Resources
and
Laboratory for Applications of Remote Sensing
Purdue University West Lafayette, Indiana 47906 USA
1978

This paper was published in : Lund, H. Gyde, Vernon J. LaBau, Peter F. Ffolliott, and David Robinson, technical coordinators. 1978. Integrated Inventories of Renewable Natural Resources: Proceedings of the Workshop. (Tucson, Arizona, Jan. 8-12, 1978). General Technical Report RM-55, Rocky Mountain Forest and Range Experiment Station, Forest Service, U.S. Department of Agriculture, Fort Collins, Colorado. 482 p.

Mapping Vegetative Cover by Computer-Aided Analysis of Satellite Data[1]

Roger M. Hoffer and Michael D. Fleming[2]

Several techniques involved in digital analysis of data from satellite scanner systems are discussed. Major cover types for a mountainous test site of approximately one million hectares were mapped with an accuracy of over 85% using both Landsat and Skylab data. Acreage estimates based on computer analysis of satellite data were compared to photo interpretation estimates, resulting in correlation coefficients ranging from 0.93 to 0.97. Topographic data (elevation, slope, and aspect) were digitally overlayed onto the satellite data, creating a data base that enabled various map products to be produced for resource management purposes.

INTRODUCTION

In many disciplines, the need for reliable and timely resource information is critical. Because of the extensive and dynamic nature of the world's agricultural and forest resources, the synoptic type of data that can be obtained at regular intervals from spacecraft altitudes has many potentials. The launch of Landsat-1 (previously ERTS-1) in 1972 has clearly shown the capability to obtain high quality data in a quantitative format. As the potential value for this type of data becomes more apparent, many questions are raised concerning techniques that can most effectively handle and analyze such large quantities of data. One approach, which was first attempted in 1966 at the Laboratory for Applications of Remote Sensing (LARS), Purdue University, utilizes pattern recognition theory applied to measurements obtained from multispectral scanner (MSS) systems. This approach was initially developed for agricultural situations and involved data obtained from aircraft altitudes. In the last few years, modification and refinement of the basis procedures, followed by extensive testing with Landsat and Skylab data has proven that computer-aided analysis techniques can successfully map and tabulate various earth resources using data obtained from spacecraft altitudes.

It is the purpose of this paper to briefly discuss the basic steps involved in computer-aided analysis of multispectral scanner data, and then to describe the results of some of our work using these techniques applied to both Landsat and Skylab data for purposes of mapping forest and other major cover types.

TECHNIQUES FOR COMPUTER-AIDED ANALYSIS OF SATELLITE MSS DATA

The digital processing and analysis of data from Landsat or other multispectral scanner systems normally involves several major steps, including:

- Data Reformatting and Preprocessing
- Development of Training Statistics
- Classification of the MSS Data
- Evaluation of the Classification Results
- Display and Tabulation of Information (for Resource Management or Other Purposes)

The data reformatting and preprocessing can include several types of activities, such as reformatting Landsat data to allow a full frame to be contained on a single 9-track, 1600 b.p.i. data tape; geometric correction

---

[1] Invited paper presented at the Society of American Foresters National Workshop on Integrated Inventories of Renewable Natural Resources, Tuscon, Arizona, Jan. 8-12, 1978. Purdue Agricultural Experiment Station Journal No. 7017.

[2] Professor of Forestry, and Research Assistant, respectively, Department of Forestry and Natural Resources, in cooperation with the Laboratory for Applications of Remote Sensing (LARS), Purdue University, West Lafayette, Indiana.

and scaling the data to a specified map base; overlaying multiple sets of digital data; etc. Due to distortions in "bulk" or "system corrected" Landsat data tapes, the geometric correction process developed by Anuta (1973) to rotate, deskew, and rescale Landsat data without causing any changes in the radiometric values of the data has been of great value. The usual output of this process is a geometrically corrected data tape which, if every resolution element or "pixel" is displayed on a standard computer line printer, results in a 1:24,000 scale alphanumeric printout, oriented with north at the top. This scale allows the analyst to superimpose the Landsat data directly on 7½ minute U.S.G.S. topographic map (or other 1:24,000 scale maps or images). Such a capability has proven very beneficial to the analyst for accurately locating various features in the satellite data printouts.

The concept behind most computer-aided analysis techniques involves a man/machine interaction, whereby the man "trains" the computer to recognize specific combinations of numbers that represent reflectance measurements in each of several wavelength bands, for the cover types of interest. This training process involves fairly limited areas for which accurate information exists concerning the type and condition of the ground cover. After a representative set of training statistics have been developed, the computer is programmed to classify the reflectance values for each resolution element (or a statistical subsample thereof) in the entire data set. In this way, the speed of the computer is used to advantage and a large geographic area can be mapped and acreage tabulations obtained at a much faster rate than would be possible using normal image interpretation techniques.

The development of training statistics and the actual classification of the data are often considered to be a single task. Much attention has been given to the various algorithms that can be utilized to classify the data, but relatively little emphasis has been given to the procedures used to train the classifier. However, it is our belief that the process involved in developing the training statistics is very critical, and indeed is the key to effective use of the computer for mapping vegetative cover using satellite MSS data.

The most common approach used for developing training statistics is the so-called "supervised training fields" technique, whereby the analyst designates to the computer the X-Y coordinates of "training fields" of the various cover types which have informational value or are of interest. For example, at a certain X-Y location in the data is a stand of ponderosa pine; a stand of aspen is at another location; other areas contain Douglas fir, grassland, water, etc. This supervised technique has been used quite effectively for agricultural mapping (Bauer, 1973), and several forestry application studies have utilized this technique, but the varying degrees of success (Bryant and Dodge, 1976; Williams and Haver, 1976; Mead and Meyer, 1977). Our own experience, involving analysis of Landsat data in nine different states and a wide variety of conditions and cover types, has shown that for wildland areas, where the cover types of interest are often not spectrally homogeneous, use of this supervised technique often does not yield acceptable accuracy or reliability. The primary reason for this is the difficulty for the analyst to define locations in the data that represent <u>all</u> significant variations in spectral response for every cover type of interest.

Another approach to developing training statistics is the so-called "clustering" technique (sometimes referred to as the "nonsupervised" technique). With this approach, the analyst simply designates the area to be classified and the number of spectrally distinct classes into which the data should be divided. The computer is programmed to classify the data into the designated number of spectral classes and prints out a map indicating which resolution elements in the data belong to which spectral classes. The analyst then relates this classification output map to aerial photos or surface observation data, and determines which resources are represented by each of the spectral classes (e.g. Spectral Class 1 is aspen, Class 2 is ponderosa pine, etc.) Experience has shown that this technique effectively overcomes the primary limitation of the "supervised" approach, but when working with large areas, the amount of computer time involved in the interative clustering sequence makes this technique very expensive. In addition, the number of spectral classes defined is often very large, since a single cover type of interest is usually represented by several spectral classes. In areas where the vegetative cover is complex (e.g. small stands, variations in stand density, species mixtures, etc.) it is often difficult to reliably relate each of the spectral classes defined by the computer to the vegetative cover type.

A recently completed research project has shown that several methods can be utilized to combine various aspects of the "supervised" and the "clustering" techniques, and that the resulting hybrid techniques can cause significant differences in the amount of analyst time required, computer time required, and classification accuracy achieved (Fleming and Hoffer, 1977). The most effective procedure defined to date is described as the "Multi-Cluster Blocks" technique for developing training statistics. In this method, several small blocks of data are located, each of which contains several cover types and spectral classes (Figure 1). Each data block is individually clustered and then the spectral classes for all cluster areas are combined to form a single data deck which statistically describes the spectral characteristics of all cover types in the entire study site. In most situations, each cover type of interest is represented by several spectral classes in the MSS data. In essence, this Multi-Cluster Blocks technique entails discovering the natural spectral groupings present in the scanner data, and then correlating the resultant spectral classes with the desired informational classes (cover types, vegetative conditions, etc.). In most cases, less than one percent of the data involved in the final classification is utilized in this training phase. However, it has been shown that the classification results can have significantly higher accuracy (e.g. 14%) when the Multi-Cluster Blocks technique is utilized rather than the more common Supervised technique (Fleming and Hoffer, 1977). Most studies to date which have been directed at mapping forest cover by computer-aided analysis techniques using Landsat data have utilized the Supervised technique, but often the results have been less than satisfactory (e.g. Mead and Meyer, 1977). Therefore, we believe that the development of a more effective technique to develop training statistics offers a significant improvement in the potential utility of these computer-aided analysis techniques and satellite data sources.

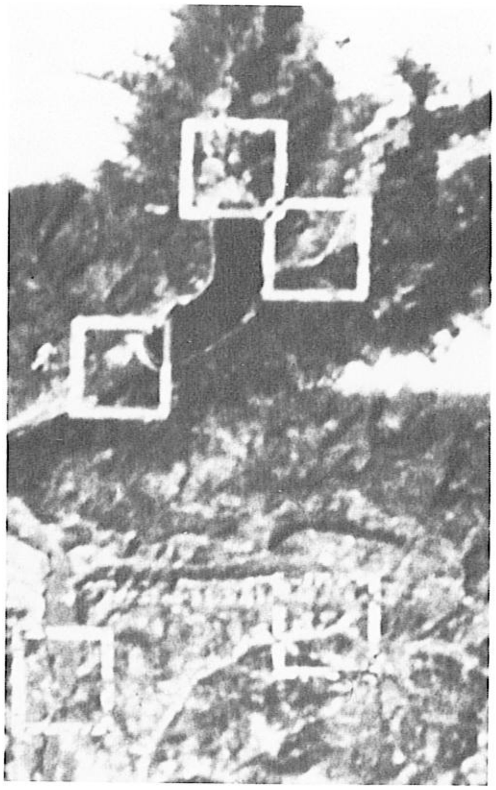


Figure 1 -- Digital Display Example of Landsat MSS Data with Several Training Cluster Blocks Defined.

After the training statistics have been defined the next step in the analysis process involves the actual classification of the data. In the classification, the measurement values of each resolution element "sensed" by the scanner system are assigned to one of the spectral classes defined by the training statistics. One of several different algorithms can be utilized to classify the data. The "Maximum Likelihood" algorithm based upon an assumption of a Gaussian distribution of the data is one of the most powerful and widely utilized algorithms, and is the one used to obtain most of the results reported later in this paper. Other algorithms such as the Minimum Distance of the Means, Table Look-up, Parallelepiped, etc. can also be used, but these have generally been found to be less accurate than the Maximum Likelihood algorithm, although they are faster.

Most classifications of satellite MSS data involve an independent classification decision for each resolution element in the data. This results in a very detailed classification map. Indeed, the amount of detail is sometimes more than is needed or desired by the user agency, since the classification maps may have a "salt and pepper" appearance in areas where there are many small (i.e. 1-3 acres) areas of different cover types. A new classification technique called ECHO Extraction and Classification of Homogeneous Objects) enables the computer to define areas of the same general cover type and then classifies the entire area in a single classification decision. With this technique, individual resolution elements that are spectrally different from the surrounding data are grouped into the surrounding cover type. Therefore, the resultant classification map looks much like a normal forest cover type map obtained by standard techniques. Depending on the particular application, some users express preference for the detailed per-point classification cover type maps, while other users prefer the more generalized ECHO map products. Figure 2 shows an example of each, and indicates the potential for using different techniques to produce various levels of detail in the map products.

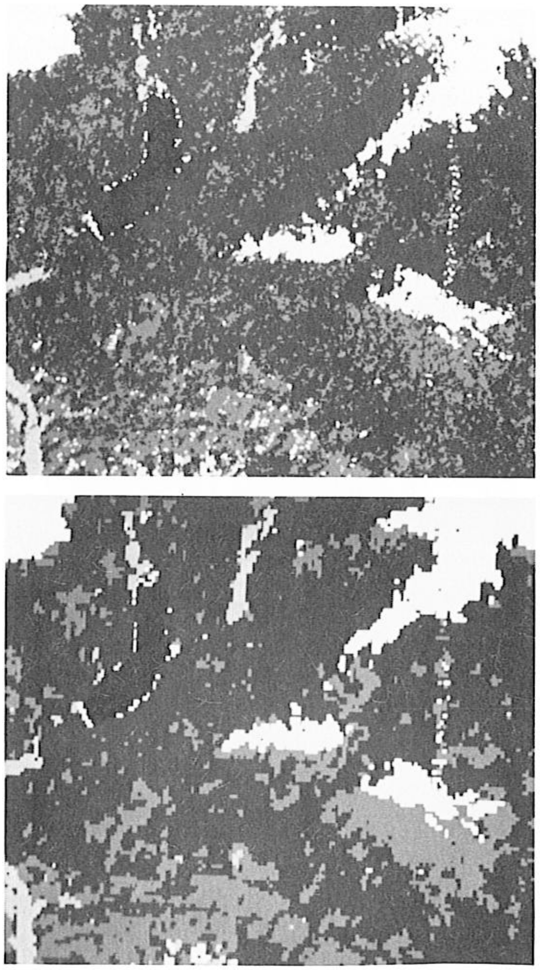


Figure 2 -- Comparison of Detailed Per-Point Classification Map (TOP) and more Generalized "ECHO" Classification (Bottom) Using Skylab MSS Data. White=Snow; Off-White=Exposed Rock & Soil; Light Grey=Pasture; Medium Grey=Deciduoud Forest; Dark Grey=Coniferous Forest; Black=Water.

After the data is classified, the results should be evaluated to determine if they are reasonably accurate. It is not difficult to classify or map a large geographic area very rapidly using computer analysis techniques, but unless the resultant classification is accurate and/or reliable enough to meet the users' requirements, the results are of little practical value. Several techniques have been used to evaluate the classification results, including both qualitative and quantitative approaches. The most common qualitative approach is to visually compare a classification map to an aerial photo or cover type map of the same area. A more quantitative approach involves designating a large number of "test areas" on which the actual cover type is known (Figure 3).

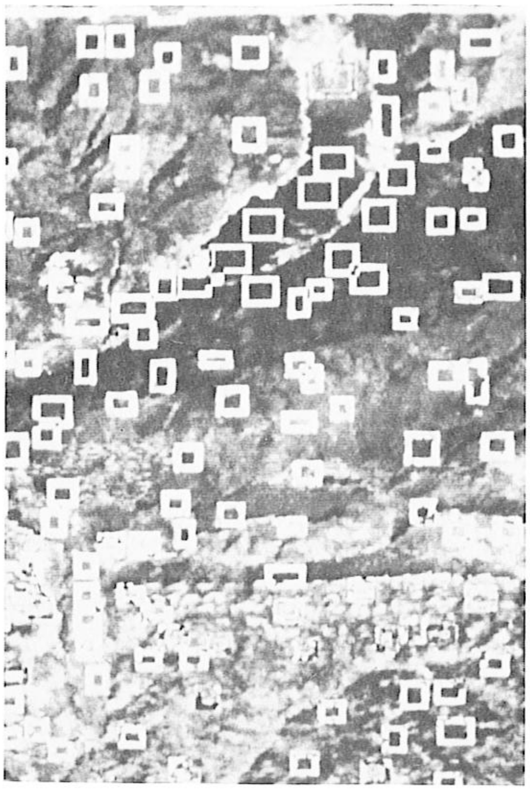


Figure 3 -- Landsat-1 MSS Band 5 (0.60-0.70μm) image of the Vallecito Reservoir study area showing the location of the test fields used to evaluate the classification accuracy.

The actual computer classification is then compared to the "correct" classification and a result is expressed in terms of classification performance or percent of the data points that were correctly classified. Tables showing both inclusive and exclusive errors can also be easily produced and studied to determine causes of classification errors (see Table 1). If acreage estimates for the various cover types are available or if they can be obtained from aerial photos, one can also compare these to acreage estimates from the computer classification results for the same areas (e.g. townships, sections, quadrangles, watersheds, etc.).

If the evaluation of the classification results indicates that they are reasonably accurate, the final step in the use of computer-aided analysis techniques is to display and tabulate the data in a format specified by the user. Different users require different scales of map outputs or may have specific requirements for tabular results. If the results of analysis of remote sensor data are to be truly useful, the results must be provided in a format that is suitable to meet the particular needs of the various user agencies and individuals. Figure 2 showed examples of maps that can be obtained in black and white or in color from various types of output devices. Another common output format (particularly in research programs) is the computer line-printer map, in which different alphanumeric symbols are used to represent the various cover types. An example of this type of computer-generated map is shown in Figure 4a. Figure 4b shows the same classification results, but in this case, the final output map was generated on a Calcomp plotter, thereby producing a map format that is preferred by some users.

RESULTS OF MAPPING FOREST COVER BY COMPUTER-AIDED ANALYSIS OF SATELLITE DATA

Several studies involving use of computer-aided analysis techniques on Landsat and Skylab data have been conducted for different areas in the Rocky Mountains of Colorado. These studies have involved the classification and mapping of forest cover and other major cover types to determine the capabilities and limitations of computer-aided analysis techniques in areas of significant topographic relief. Many researchers have carried out studies involving flat-land areas where the topographic effects on spectral response are negligible. However, because a significant proportion of the forest resources of the world are located in mountainous regions, it is our belief that

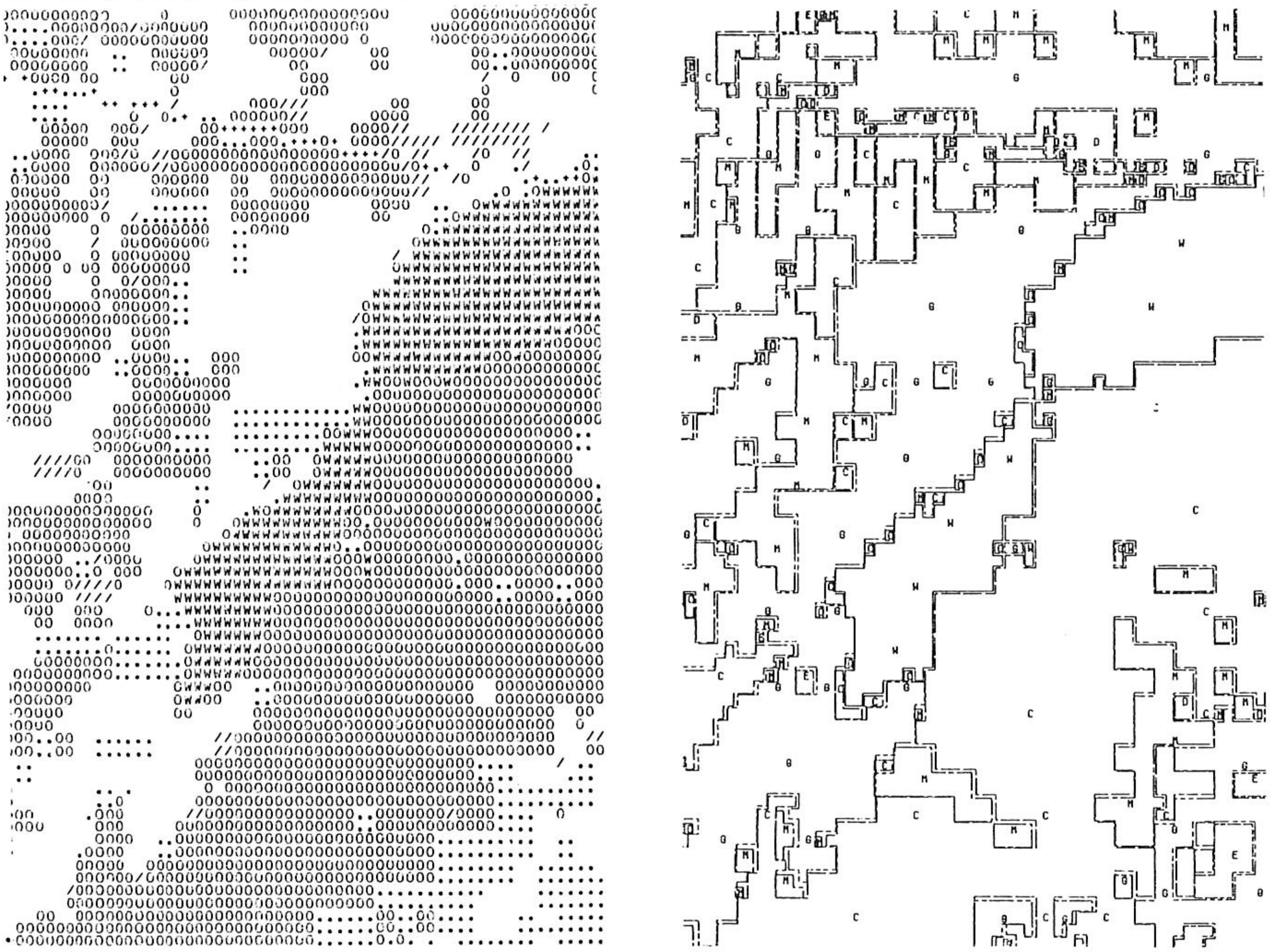


Figure 4a (left) and b (right). Two Different Types of Computer-Generated Classification Maps. A standard line-printer was used to obtain the map on the left whereas the map on the right was obtained by a Calcomp plotter. Both maps were generated using the same classification results data tape. 0=Coniferous Forest, /=Deciduous Forest, W=Water, •=Exposed Rock & Soil, Blank=Grassland.

computer-aided analysis techniques must be tested in mountainous as well as flatland areas, if we are to determine the true practical capabilities and limitations of these techniques.

Most of our work has concentrated on the San Juan Mountain area of Colorado which contains a complex mixture of forest types, rangeland, alpine tunda, agricultural areas, water bodies, geological features and various manmade features. The topography of this test area is rugged, ranging in elevation from less than 2000 meters to over 4200 meters. Within this range of elevation, there is a distinct distribution of cover types according to altitude. Much of the area is dominated by ponderosa pine (Pinus ponderosa) forest, but Douglas fir (Pseudotsuga menziesii var. glauca), Engelmann spruce (Picea engelmanii), and subalpine fir (Abies lasiocarpa) are found at the higher elevations and on steep north slopes. There are also many stands of quaking aspen (Populus tremuloides), primarily on sites that have been disturbed by fire or avalanches. At lower elevations, the drier steep southern slopes are dominated by Gambel oak (Quercus gambelii), and the valley bottoms are occupied by agricultural land (mostly hayfields). Timberline in the region is at approximately 3600 meters, and extensive areas of tundra are found above this elevation.

One of the major studies in the San Juan Mountains utilized Landsat data to map and tabulate forest and other major cover types over a relatively large area of 993,800 hectares (2,456,000 acres). The Landsat data used had been obtained on Sept. 21, 1973. Sixteen training blocks were defined, each of which contained four to six different cover types. The location of these training blocks was based in part upon the availability of aerial photography for those particular areas, as well as the spectral heterogeneity of these blocks. The Multi-Cluster Blocks technique was utilized and each training area was clustered into 12 to 18 spectral classes. Spectrally similar classes were then combined, resulting in 14 distinct, separable spectral classes. Each of these spectral classes were identified using existing aerial photography and type maps of the area. It was determined that the spectral classes present could be grouped into five "Major Cover Types" or "Land Use" Categories[1], including Coniferous Forest, Deciduous Forest, Grassland, Water, and Barren (exposed rock, outcrops, soil, and sparsely vegetated tundra). The grassland category included both cultivated pasture and rangeland areas because they could not be reliably separated on the basis of spectral response.

Test areas which included a total of 16,170 resolution elements were then obtained from quadrangles in which no training blocks were located. Aerial photos and subsequent field checks were used to accurately identify the cover type and characteristics (e.g. stand density) of each test area. After the entire area was classified, a qualitative evaluation of the resultant maps indicated that the classification appeared to be reasonably accurate. To obtain a quantitative evaluation, the computer classification of the designated test areas was tabulated, the results of which are shown in Table 1.

This type of tabulation allows an effective method of evaluate both the inclusive and exclusive errors present in the classification, and to determine performance for individual cover types as well as for the overall classification. Table 1 indicates that a relatively accurate classification had been obtained for these cover types, using computer-aided analysis techniques and Landsat satellite data. This result was believed to be particularly significant in view of the topographic and vegetative complexity of the area, and the size of the test site involved.

To evaluate the classification using acreage comparisons, the number of resolution elements classified into each cover type within each of the 63 quadrangles (U.S.G.S. 7½ minute quadrangles) in the entire test were tabulated, and area estimates based upon the computer classification were obtained. A separate team of people utilized planimeters and dot grids to determine the area of the various cover types according to type maps which had been developed from aerial photos using standard photo interpretation techniques. A random sample of seven quadrangles (totalling about 112,400 hectares) were utilized for the photo interpretation acreage estimates. Comparison of these two data sets resulted in a correlation coefficient (r) of 0.97. Such a correlation coefficient would indicate that the area estimates obtained by computer analysis of Landsat data are in close agreement with the estimates obtained from aerial photos, thereby providing additional confidence in the classification accuracy.

A cost evaluation of this analysis indicated that the total cost (including computer, personnel salaries, etc.) for pre-processing, developing the training statistics, classifying the data, and evaluating the results was

---

[1] A combination of Level I and Level II Land Use Categories as defined by U.S. Geological Survey Circular 964 (Anderson et al, 1976).

Table 1 -- Classification Performance of Major Cover Types in the San Juan Mountain Test Site (993,800 hectares).

Cover Type	No. of Samples[1]	No. of Samples Classified as: Coniferous	Deciduous	Grassland	Barren	Water	Shadow[2]	Percent Correct
Coniferous	9,634	9,110	22	53	21	96	332	94.6
Deciduous	1,475	113	1,286	76	0	0	0	87.2
Grassland	3,677	49	129	2,988	510	0	1	81.2
Barren	35	0	0	1	34	0	0	97.1
Water	1,349	6	0	0	0	1,334	9	98.9
Totals	16,170	9,278	1,437	3,118	565	1,430	342	

Overall Performance = (9,110 + 1,286 + 2,988 + 34 + 1,334)/16,170 = 91.2%

---

[1] Each "sample" is a Landsat resolution element. The column labelled "No. of Samples" indicates the total number of resolution elements of the various cover types actually present in the test areas (assuming that each test area contains only a single cover type).

[2] One of the 14 spectral classes that had been defined involved areas of topographic shadows, but since this was not an actual cover type, any resolution elements belonging to this spectral class were considered as errors in the classification.

approximately $0.0025 per hectare (0.1¢ per acre). Since this analysis was done on a medium-speed digital computer programmed for research types of activities, it would appear that in the future, such analyses could be conducted on special purpose, high speed digital computers in a relatively cost-effective manner.

Table 2 -- Classification Performance of Forest Cover Types for the Vallecito Reservoir Study Site.

Cover Type[1]	No. of Samples	No. of Samples Classified as: Pine	Spruce/Fir	Oak	Aspen	Grassland	Water	Barren	Percent Correct
Pine	1111	904	169	5	9	3	1	20	81.4
Spruce/Fir	747	254	485	2	6	0	0	0	64.9
Oak	481	8	0	297	95	81	0	0	61.7
Aspen	204	5	0	33	160	6	0	0	78.4
Grassland	242	2	0	6	0	232	0	2	95.9
Water	240	0	0	0	0	0	240	0	100
Barren	98	0	0	0	0	6	0	92	93.9
Totals	3123	1173	654	343	270	328	241	114	

Overall Performance = (904 + 485 + 297 + 160 + 232 + 240 + 92)/3123 = 77.2%

---

[1] Pine = Ponderosa Pine; Spruce/fir = Engelmann spruce, Douglas-fir and subalpine fir; Oak = Gambel oak, Aspen = Quaking aspen.

A second phase in the computer-aided analysis of the satellite data from this area involved mapping forest cover types over a more limited test site, referred to as the Vallecito Reservoir Study Site, which covered an area of about 23,000 hectares. The analysis procedures previously described were again utilized in this classification but since a more detailed level of classification was involved, a larger number of spectral classes had to be defined and utilized. In this case, a total of 24 spectral classes were obtained for the analysis. The computer classification of forest cover types resulted in a map of this intensive study site that qualitatively looked fairly good, but it was not as accurate as the map of major cover types. In several areas, individual forest cover types appeared to have been misclassified within the general categories of coniferous or deciduous. This was substantiated by a quantitative evaluation using test areas which had been carefully field checked. These results are shown in Table 2.

The results shown in Tables 1 and 2 indicated that a higher level of classification performance can be achieved for major cover types than for individual forest cover types. This is to be expected, since fewer and more easily separated spectral classes (i.e. major cover types) normally can be classified more accurately. In situations where there are larger numbers of spectral classes, (all of which are in the general category of "green vegetation"), the individual classes usually are not as spectrally separable and the resultant classification accuracy is generally lower. In this case, the 60-80% accuracy for the individual forest cover types probably is not accurate enough to be useful from a practical standpoint. However, preliminary results indicate that techniques which allow the use of topographic data in addition to the spectral data may enable the accuracy to be improved (Hoffer, 1975a). Additional work is needed to determine the level of detail that can be achieved and the conditions under which these techniques can produce satisfactory results.

In evaluating the specific reasons for classification errors within individual forest cover types, preliminary tests indicated that aspect, slope, elevation, and stand density all cause a statistically significant impact on spectral response (Hoffer, 1975b). A later more detailed analysis on the spruce/fir cover type has shown that decreasing stand density can be correlated quite well with increasing spectral response. Figure 5 shows this relationship for crown closures ranging from 0% to 100%, in 20% increments. The actual response values were slightly higher in the infrared than the predicted values which had been based on the assumption that the grass/forest mixture would produce a linear relationship between 0% grass & 100% spruce/fir and 100% grass & 0% spruce/fir. A similar relationship between spectral response and crown closure has been observed for three groupings of stand densities in southern pine by Williams and Haver (1976).

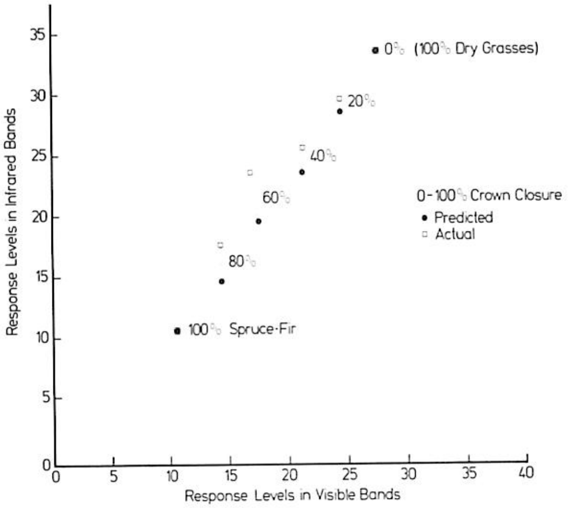


Figure 5 -- Spectral Response in Relation to Percent Crown Closure for the Spruce/Fir Forest Cover Type.

Analysis of Skylab MSS data produced results that are generally comparable to those obtained with the Landsat data. The data were obtained on June 5, 1973, and there was some snow present in the test site. Therefore, the classes defined and mapped included coniferous forest, deciduous forest, grassland, water, and snow. Even though the quality of the data was rather poor (i.e. low signal to noise ratio) the improved spectral resolution and spectral range of the Skylab MSS data (as compared to the Landsat data) enabled results to be obtained that were only slightly less accurate than those that had been obtained with the Landsat data. A classification performance of 85.1% was found for the major cover types, based upon a set of test data that included a total of 2400 pixels. The classification for individual forest cover types was 71.0%, again indicating the difficulty of obtaining highly accurate classifications for forest cover types in this area of complex topography.

Acreage estimates of major cover types obtained by computer-aided analysis of the

Skylab data were again compared to acreage estimates obtained by standard photo interpretation techniques, on a quadrangle-by-quadrangle basis for the five 7½ min. U.S.G.S. quadrangles in the test site. The photo interpretation was done by researchers at the University of Colorado, using 1:120,000 scale color infrared photos obtained the day after the Skylab MSS data were obtained. The result of this acreage comparison is shown in Figure 6.

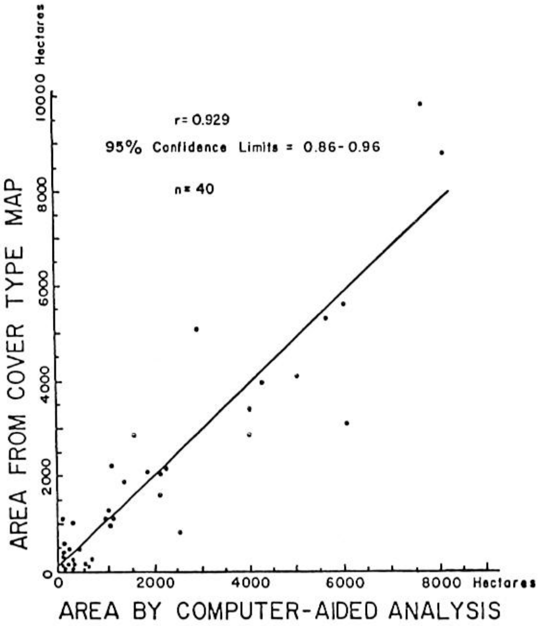


Figure 6 -- Acreage Estimates of Cover Types Obtained by Computer-Aided Analysis of Skylab MSS Data, Compared to Aerial Photo Interpretation Estimates.

As was the case with the Landsat data, a high degree of correlation was again found between these two very different methods of obtaining acreage estimates (e.g. computer analysis of satellite data and human interpretation of aerial photos). Because of the statistical basis for the computer classification of the MSS data, a particular individual resolution element will have a certain probability of being mis-classified. However, over a fairly large geographic area such as a 7½ minute U.S.G.S. quadrangle, some inclusive and exclusive classification errors tend to balance out, thereby resulting in acreage estimates that are somewhat better than one might anticipate, given the classification accuracy achieved.

One of the most significant results of our work in this Colorado test site has involved the overlay of topographic data with the satellite data. The Defense Mapping Agency has utilized 1:250,000 scale U.S.G.S. topographic maps to digitize the elevation data, using approximately a 50 meter grid. A data tape containing this digital elevation data for the San Juan Mountains test site was obtained from DMA, and techniques were developed at LARS to digitally overlay this data onto the Landsat and Skylab MSS data. An interpolation procedure was then developed to define the aspect and slope of each resolution element.

The combination of cover type classifications plus elevation, slope, and aspect data in a computer-compatible format for this large area in the San Juan Mountains has enabled a very flexible and useful type of map output to be produced for resource managers. We have worked with U.S. Forest Service personnel involved in management and land use planning activities for this area, producing a variety of map combinations for their evaluation and use. In one case, for example, wildlife management personnel needed "big game winter range" maps of this area, on a quadrangle-by-quadrangle basis. In this case, big game winter range was defined as "grassland" cover (obtained from the computer classification of satellite data), for elevations between 7500 and 9000 feet, at slopes of less than 30^0, and only for S.-SE. to S.-S.W. aspects. This combination of cover type and topographic characteristics was quickly and inexpensively displayed in map and tabular format and supplied to the Forest Service personnel in the form of 1:24,000 scale maps.

We believe that the speed and flexibility of being able to combine various cover type and topographic features has tremendous potential for the resource manager. Overlays of additional types of data, such as land ownership, political boundaries, soils data, etc. would allow data bases to be developed that would enable planners, resource managers, and others to obtain maps and tabulations of a wide variety of combinations of data. Such combinations could be defined by the user so as to meet a particular information need for a specific area. We believe that it is toward this type of system for data manipulation that we should be striving. Such systems could make use of computer-derived classifications of cover type (obtained from satellite data) as simply one part of the entire data set. The use of digital computer systems in such a manner would provide the resource manager with a flexible tool to meet his specific requirements for many types of information.

SUMMARY AND CONCLUSIONS

The application of computer-aided analysis techniques to satellite MSS data has shown that major cover types (such as coniferous forest, deciduous forest, grassland, exposed rock and soil, snow, and water) can be identified and mapped with a reasonable degree of accuracy (85-90%), even in areas of rugged mountainous terrain and spectrally complex vegetative cover types. Analysis of Landsat MSS data on a large (993,800 hectares) test site resulted in a classification performance for major cover types of 91.2%. (This figure is based upon the tabulation of 16,170 resolution elements from test areas that were carefully field checked and that had been designated in quadrangles other than those used for developing the training statistics.) Classification of Skylab MSS data resulted in a classification performance of 85.0% for the major cover types, which was somewhat less accurate than the Landsat classification, probably due to the poor quality of the Skylab data. The Landsat MSS data was also used to map individual forest cover types. The resultant 77.2% overall classification performance tends to indicate that results at this level of detail would probably not be accurate enough to provide useful information for most users. Further testing should be carried out in large test sites to determine if a satisfactory level of classification performance can be achieved for individual forest cover types in areas that are not as topographically or spectrally complex.

Acreage estimates of major cover types obtained by computer-aided analysis of Landsat and Skylab MSS data were compared to acreage estimates obtained by manual interpretation of 1:120,000 scale aerial photos. Five different tests were conducted, resulting in correlation coefficients ranging from 0.93 to 0.98. These results indicate that for fairly large areas, the analysis of satellite data by computer can provide acreage estimates that are very similar to those obtained by conventional techniques.

The development of the training statistics is a particularly critical aspect in the analysis process. A recently developed procedure for developing representative training statistics, called the "Multi-Cluster Blocks" technique has enabled significant improvements (e.g. 14%) in the final classification accuracy as compared to the "Supervised Training Fields" technique utilized by many analysts. Various types of output map formats can be produced from the computer classification tapes, including black and white or color-coded outputs, alphanumeric line-printer printouts, Calcomp plotter maps and others. These offer considerable flexibility in obtaining a format most suitable for meeting the particular needs of a user agency. The type of classification algorithm used can also enable different formats of classification maps to be produced. A newly developed "ECHO" algorithm appears very promising for forestry applications.

A particularly significant result of these studies involved the successful digital overlay of Landsat and Skylab MSS data with topographic data (elevation, slope, and aspect). Maps were generated on a quadrangle-by-quadrangle basis for various combinations of cover type, elevation, slope, and aspects, as specified by the resource managers and planners. Evaluation of these maps by the user agency (i.e., U.S. Forest Service) indicated that the results were adequate to meet many of their existing needs. Such a capability to combine the various features quickly and cost-effectively has tremendous potential. Overlay of additional data (e.g. soils, land ownership, political boundaries, etc.) would allow even more useful and flexible data bases to be developed. Such data bases, in combination with both existing and developing computer processing capabilities, will enable resource managers and planners to obtain map and acreage data on a custom-order basis, specific to their particular information needs. It is toward this goal that we must strive.

LITERATURE CITED

Anderson, J.R., E.E. Hardy, J.T. Roach, and R.E. Witmer. 1976. "A Land Use and Land Cover Classification System for Use With Remote Sensor Data". U.S. Geological Survey Professional Paper 964, Washington, D.C. 28 pp.

Anuta, P.E. 1973. "Geometric Correction of ERTS-1 Digital Multispectral Scanner Data." LARS Information Note 103073. Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana. 23 pp.

Bauer, M.E., and J.E. Cipra. 1973. Identification of Agricultural Crops by Computer Processing of ERTS MSS Data. Proceedings of the Symposium on Significant Results from ERTS-1, NASA SP-327, Washington, D.C. Pp. 205-212.

Dodge, A.G., and E.S. Bryant. 1976. "Forest Type Mapping With Satellite Data." Jour. Forestry 74(8):526-531.

Hoffer, R.M. and Staff. 1975a. "Natural Resource Mapping in Mountainous Terrain by Computer Analysis of ERTS-1 Satellite Data." Agricultural Experiment Station Research

Bulletin 919, and LARS Information Note 061575, Purdue University, West Lafayette, Indiana. 124 pp.

Hoffer, R.M. and Staff. 1975b. "Computer-Aided Analysis of Skylab Multi-spectral Scanner Data in Mountainous Terrain for Land Use, Forestry, Water Resource and Geologic Applications." (Final Report on Contract No. NAS9-13380, Skylab EREP Project 398.) LARS Information Note 121275. 381 pp.

Mead, R., and M. Meyer. 1977. "Landsat Digital Data Application to Forest Vegetation and Land-Use Classification in Minnesota". IAFHE Remote Sensing Laboratory Research Report 77-6, University of Minnesota, St. Paul, Minnesota. 57 pp.

Williams, D.L., and G.F. Haver. 1976. "Forest Land Management by Satellite: Landsat-Derived Information as Input to a Forest Inventory System." Interlab Project #75-1, Earth Resources Brancy, NASA/Goddard Space Flight Center, Greenbelt, MD. 36 pp.

Acknowledgments:

This work was done under NASA Contracts NAS 5-20948 and NAS 9-14970, Task 1. Acknowledgment is made of the work of Dr. Paula Krebs and her associates at the University of Colorado for their photo interpretation, field work, and coordination efforts, and to Messrs. Paul Anuta and Dennis Adams, Laboratory for Applications of Remote Sensing, Purdue University, for their work in developing the computer software for overlaying and processing the D.M.A. topographic data with the Landsat and Skylab data.