Reprinted from

Symposium on Machine Processing of Remotely Sensed Data

June 21 - 23, 1977

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana

IEEE Catalog No.
77CH1218-7 MPRSD

LANDSAT DIGITAL DATA APPLICATION TO FOREST VEGETATION AND LAND USE CLASSIFICATION IN MINNESOTA

ROY A MEAD AND MERLE P. MEYER
University of Minnesota

I. ABSTRACT

LANDSAT digital data were used to map eleven categories of land cover in north central Minnesota. The classification accuracy of these maps was found to be very low and they were not adequate for use by field level resource managers. A discussion of the advantages and disadvantages of various processing systems, different algorithms, and the problems in selecting training sets, is included.

II. INTRODUCTION

The purpose of this study was to determine whether information useful to field level resource managers is obtainable from LANDSAT data. Interested user-cooperators representing the U.S. Forest Service, Minnesota Department of Natural Resources, Blandin Paper Company and the Itasca County Land Commissioner's office assisted in ground truth collection (particularly in training set selection) and in the subsequent evaluation of classification accuracy. They also helped by suggesting meaningful categories for mapping and eventually determined the suitability of the maps produced in the study.

There were three major objectives:

1. To determine the categories of natural resources which are feasible to map from LANDSAT data in the study area - Itasca County, Minnesota.
2. To determine the best method(s), including the advantages and disadvantages of several data analysis systems, comparing various algorithms, and the relative value of interactive and batch processing, for analysis of LANDSAT data in north central Minnesota.
3. Produce vegetation maps for trial use by the field-cooperators.

III. STUDY AREA LOCATION, CHARACTERISTICS

The study area included all of Itasca County in north central Minnesota and the Chippewa National Forest (Figure 1). In this area, forests are the major land cover type - aspen/birch and maple/basswood being the most common deciduous forest cover types. In lowland areas, the predominant coniferous species are black spruce, balsam fir and tamarack - while red pine, and jack pine are commonly found on upland sites. Wetlands represent a considerable portion of the study area and are usually typified either as treeless sphagnum/leatherleaf bogs or as cattail/sedge meadows. In total, the study area can be described as a continuum of successional stages with many transitional zones and consisting of a mixture of the types.

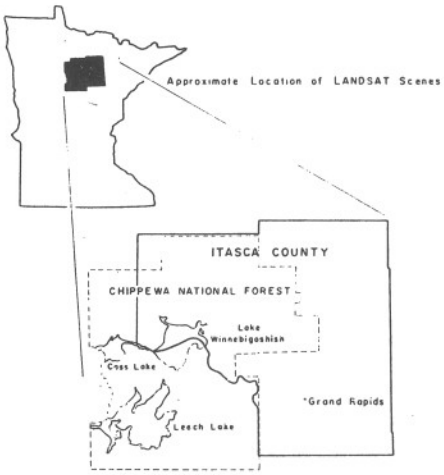

Figure 1. Map of Study Area.

IV. PROCEDURE

The ground truth collected for the entire study area included acquisition of the following:

1. Color infrared aerial photography flown in August, 1975, at a scale of 1:80,000 for all of Itasca County served as the primary data base. The selection of training sets was accomplished more readily through interpretation of these photos. The photography was also used extensively while working interactively on the computer in the classification process.
2. Forest cover type maps obtained from the field-cooperators and compiled on base maps identified vegetation types and classified items by stand density and size class. Knowledge of this nature was essential to insure that training sets were selected from stands representing a variety of conditions for each vegetation category. Also, this information helped the photo interpreter to differentiate specific species which were sometimes nearly impossible to identify on the small scale photography.
3. Intensive field checking over large areas comprised the third type of ground truth. It was necessary for the investigators to become familiar with the vegetation types in the study area, and gain an understanding of the variability within each type - as well as identify mixtures which occur. Field checking was also required to evaluate classification accuracy. The knowledge and advice of the field-cooperators who were very familiar with the study area, provided considerable help with respect to collection of the three types of ground truth.

Digital LANDSAT data recorded on May 29, 1973 and July 17, 1974, were selected for use in this study. These images were geometrically corrected, reoriented to a north-south base and rescaled to match the scale of U.S.G.S. maps at 1:24,000. All of the data from both dates were temporally registered, thereby permitting the use of data from both dates simultaneously.

Classifications were performed using the two pattern recognition routines most commonly used in analysis of LANDSAT data: maximum likelihood and parallelepiped. In addition to using both of these algorithms, an unsupervised clustering routine was also tried. The classifications were performed using the following: (1) an in-house batch processing system RECOG borrowed from Colorado State University, (2) the General Electric Image 100, located at the EROS Data center, (3) the Bendix Corporation M-DAS System, located in Ann Arbor, Michigan, and (4) the IDIMS System, manufactured by ESL Corporation, also located at the EROS Data Center. In this way, both interactive and batch processing systems were tested using the same data sets.

Each of the data analysis systems above required a specific input format and therefore, considerable preprocessing of the data was required. Due to the large preprocessing costs, work with the temporally-registered data was only done on the Image 100 and the IDIMS systems.

The output display of the classifications were in three forms:

1. Line printer character displays.
2. Color CRT displays.
3. Color film recorder hard copy prints.

The IDIMS and the Image 100 could display the classifications on color CRTs or as line printer character maps while, with RECOG, the line printer was the only format available. The M-DAS system, however, could display the data in all three formats, including hard copy prints generated with a color film recorder.

The hard copy format was considered the best for several reasons:

1. Hard copy afforded direct comparison with ground truth.
2. Hard copy was most practical for use by the field resource managers/cooperators.
3. Hard copy could be displayed at several scales which permitted determination of the optimum scale.

Considering the formats available with each system, it was decided to use the M-DAS system to produce the final classifications for an extensive and rigorous evaluation of the classification

Table 1. Land Cover Categories.

Water
Lowland conifer
Upland conifer
Mixed forest
Brush and shrub
Grassland and open
Agriculture
Mined land
Sedge meadow
Urban
Sphagnum/leatherleaf bog

accuracy. The results of the Image 100 and IDIMS systems were qualitatively evaluated but no evaluation of the line printer output from RECOG was made, although many of the original training sets and signatures were developed using the batch processing system.

Evaluation of the classification accuracy for the final hard copy land cover maps was accomplished by three methods described below. The first two techniques were employed because of their common use in other studies, and the third was used because it allowed for a more rigorous evaluation of the classification accuracy.

1. The areas designated as training set areas were classified by the computer. The results of these classifications were compared with the ground truth for each training set and the percent correct classification for each vegetation category was determined. Only errors of commission are possible when using this method.
2. Additional areas of known vegetative composition were selected for testing the classification accuracy. Three test areas representing "pure" blocks of high density were chosen for each vegetation category. These test areas were subjectively located in the center of large stands and varied in size. The machine solution was then compared pixel-by-pixel with the ground truth, by which means the errors of omission as well as commission could be estimated. Since these areas had not been used in the training process, it was felt that a more realistic estimate of the classification accuracy was obtained.
3. The third technique for evaluating the classification accuracy included a pixel-by-pixel comparison with known ground truth. A forest type map was prepared from the color infrared aerial photos using the same classification system. This map (in a transparent form) was laid directly over the color-coded LANDSAT classification at the same scale. This method gave the lowest and most realistic estimate of accuracy for classifying each category. In this case, large blocks which included transition zones and mixtures of types were evaluated rather than the relatively pure, atypical stands used in the first two methods.

V. RESULTS AND DISCUSSION

In accordance with the primary objective of this study, many categories of land cover were considered for classification. Specific land cover types of interest to field resource managers were considered and were combined until more general categories which seemed feasible to identify with the LANDSAT data were found. For example, red pine, white pine, and jack pine were all combined into the upland conifer category, and tamarack, black spruce, and white cedar were combined to form the lowland conifer category. In this manner, the final categories listed in Table 1 were determined.

The classification accuracies for these categories were determined by the three methods outlined above. A summary of the resulting classification accuracies found by the three methods are given in Table 2 and all of the omission and commission errors for methods I, II and III are shown in Tables 3, 4, and 5, respectively.

Considering the data in Table 2, it is very apparent that the classification accuracies were lower for method III and probably represent the most reliable estimate of accuracy. Four of the categories had accuracies of zero, and it is clear that only the water and sedge meadow had reasonably acceptable accuracies. There were no mined areas in the evaluation block and a test set for mined land was not selected. Therefore, no estimates of accuracy for this category are available except for the training sets.

The final color coded classifications were displayed at three scales: 1:250,000, 1:125,000, and 1:24,000. The largest scale was necessary for use in the accuracy evaluation since individual pixels are identifiable at this scale. However, the field cooperators felt that the scale was too large considering the poor site-specific information that was obtainable with the LANDSAT imagery. Therefore, it was concluded that the smaller scale imagery was more desirable since it forced the user to consider only the broad general patterns of land cover over the entire region with individual pixels practically unidentifiable. Also at the smaller scales, the users could not attempt to extract site-specific, detailed information from the imagery. Naturally, the classifications were no more accurate at the small scale than at the large scale but the scale limited the types of applications and thus put the data in a more appropriate form. It was concluded

Table 2. Summary of Classification Accuracy by Three Methods

Land Cover Category	Method I	Method II	Method III
Water	100	100	95
Lowland conifer	91	91	52
Upland conifer	89	88	64
Mixed forest	95	52	23
Brush & shrub	86	0	0
Grassland and open	95	-	0
Agriculture	98	0	0
Mined land	100	-	-
Sedge meadow	98	46	83
Urban	94	-	0
Sphagnum/leatherleaf	99	0	4

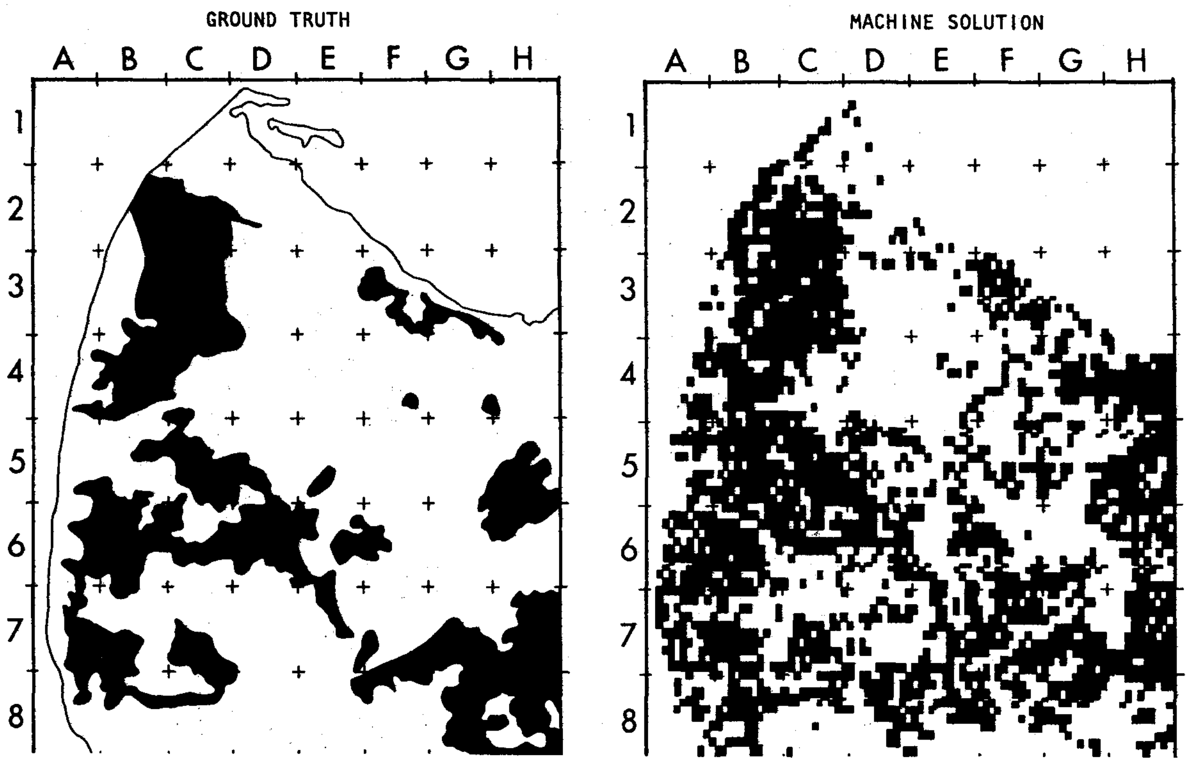

Figure 2. Upland Conifer: Interpreter's Solution Left, Machine Solution Right.

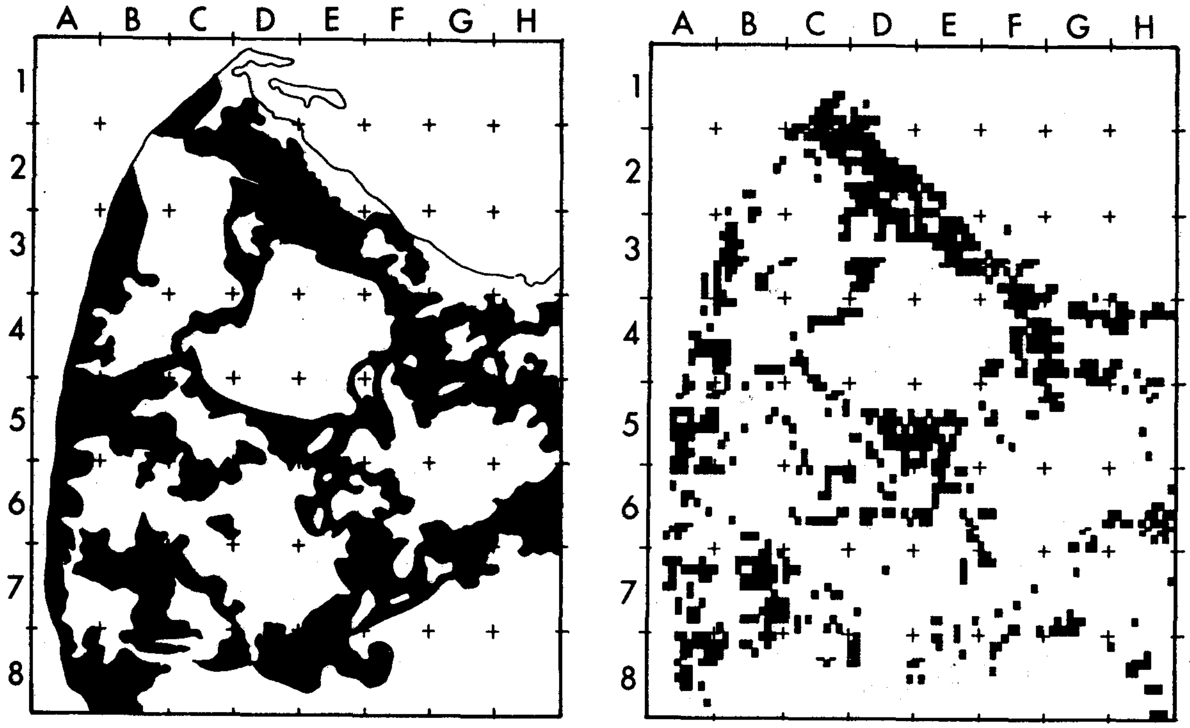

Figure 3. Mixed Forest: Interpreter's Solution Left, Machine Solution Right.

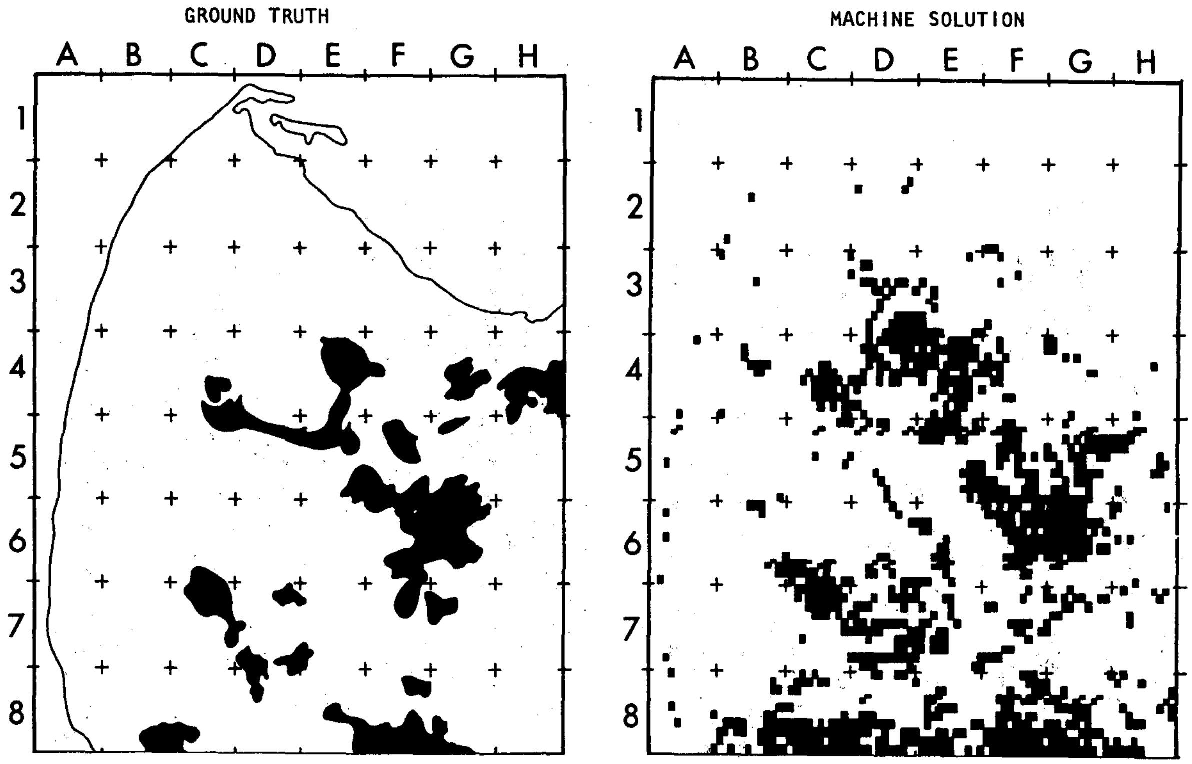

Figure 4. Lowland Conifer: Interpreter's Solution Left, Machine Solution Right.

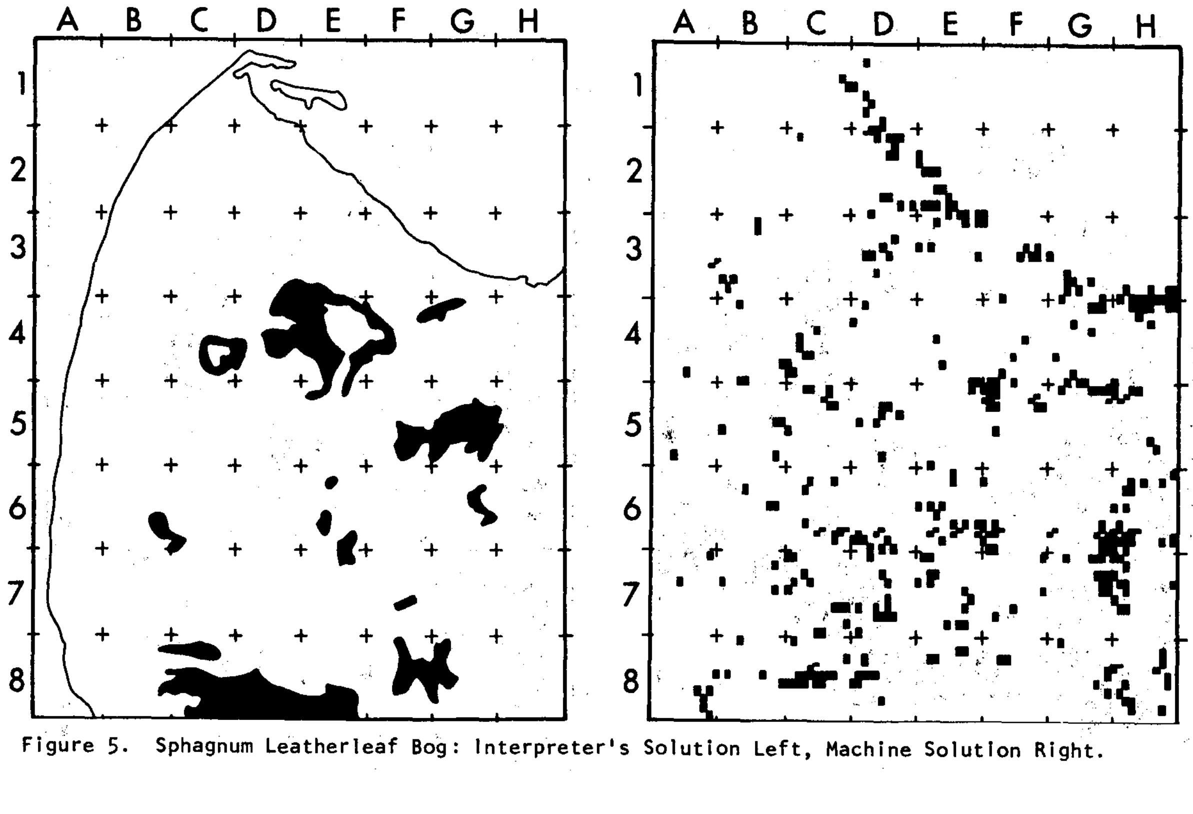

Figure 5. Sphagnum Leatherleaf Bog: Interpreter's Solution Left, Machine Solution Right.

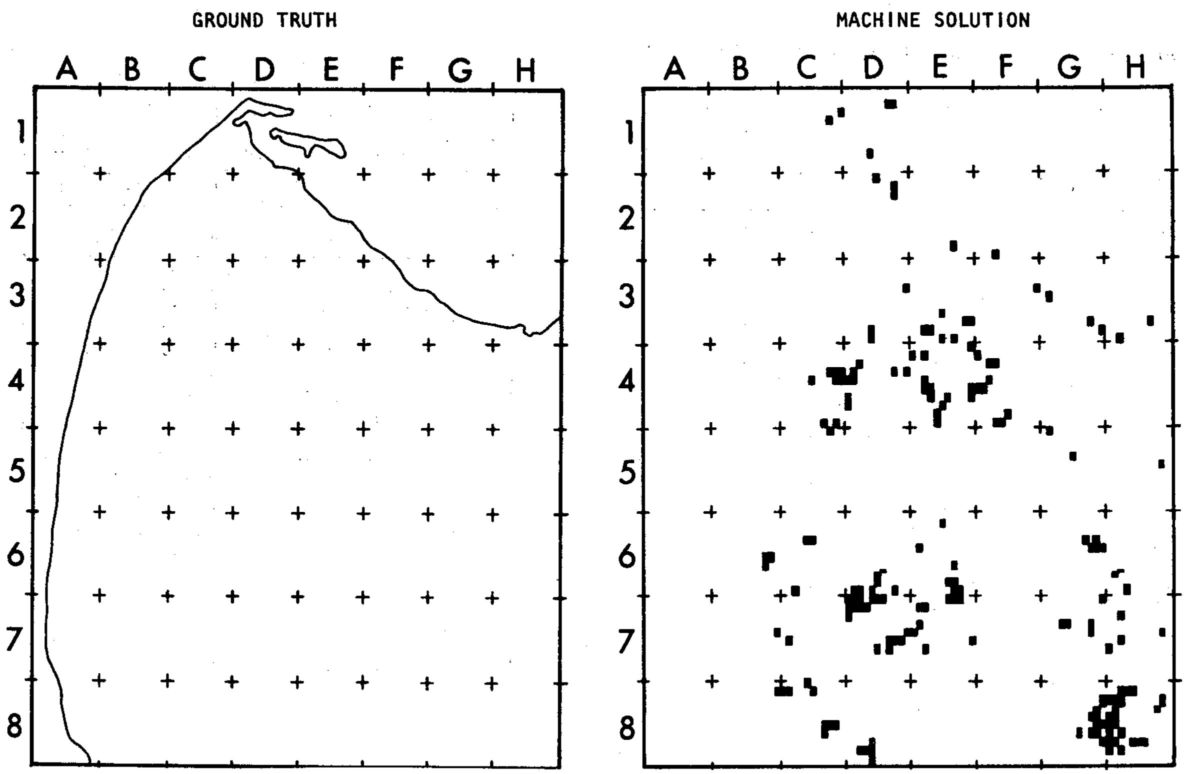

Figure 6. Unclassified: Interpreter's Solution Left, Machine Solution Right.

Table 3. Accuracy of Classification on Training Sets.

Ground Truth	Water	Lowland Conifer	Upland Conifer	Mixed Forest	Brush & Shrubs	Grassland & Open	Agriculture	Mined Land	Sedge Meadow	Sphagnum Leather-leaf	Urban	Total	Correct	Correct
Water	780											780	780	100
Lowland Conifer	9	103								1		113	103	91
Upland Conifer		9	173	12			1					195	173	89
Mixed Forest			5	110			1					116	110	95
Brush & Shrubs					32	4					1	37	32	86
Grassland & Open	1				2	82			1			86	82	95
Agriculture						1	156	3				160	156	98
Mined Land								132				132		100
Sedge Meadow	2								121			123	121	98
Sphagnum-Leather-leaf						1				100		101	100	99
Urban									1		16	17	16	94

Table 4. Accuracy of Classification Test Sets.

Ground Truth	Water	Lowland Conifer	Upland Conifer	Mixed Forest	Brush & Shrubs	Grassland & Open	Agriculture	Mined Land	Sedge Meadow	Sphagnum Leather-leaf	Urban	Uncategorized	Total	# Correct	% Correct
Water	238												238	238	100
Lowland Conifer		63	4							2			69	63	91
Upland Conifer		2	242	10						4		16	274	242	88
Mixed Forest			48	66			1			7		4	126	66	52
Brush & Shrubs		31	8	4						12	17		72	0	0
Grassland & Open	-	-	-	-	NOT IMPORTANT TYPE IN AREA				-	-	-	-	-	-	-
Agriculture		2			1				3	11	5	1	23	0	0
Mined Land	-	-	-	-	NOT IMPORTANT TYPE IN AREA				-	-	-	-	-	-	-
Sedge Meadow	1	5				2			16	9	2		35	16	46
Sphagnum-Leather-leaf		68	6						5		8		87	0	0
Urban	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 5. Accuracy of Classification on the Tamarack Point Evaluation Area.

Ground Truth	Water	Lowland Conifer	Upland Conifer	Mixed Forest	Brush & Shrubs	Grassland & Open	Agriculture	Mined Land	Sedge Meadow	Sphagnum Leather-leaf	Urban	Uncategorized	Total	# Correct	% Correct
Water	139	4	1						2				146	139	95
Lowland Conifer		71	47	6					7	2	2	2	137	71	52
Upland Conifer		39	225	51						7	8	21	351	225	64
Mixed Forest	1	65	229	115				1	3	51	11	30	506	115	23
Brush & Shrub		5	3	5						6			19	0	0
Grassland & Open											2		2	0	0
Agriculture		5	2				1		2	8	8		26	0	0
Mined Land													0	0	0
Sedge Meadow									5		1		6	5	83
Sphagnum-Leather-leaf	1	63	19						4	4	21		112	4	4
Urban													0	0	0
Total	141	252	526	177	0	0	1	1	23	78	53	53	1305		
# Correct	139	71	225	115	0	0	0	0	5	4	0				
% Correct	98.6	28.8	42.9	65.3	0	0	0	0	21.7	5.1					

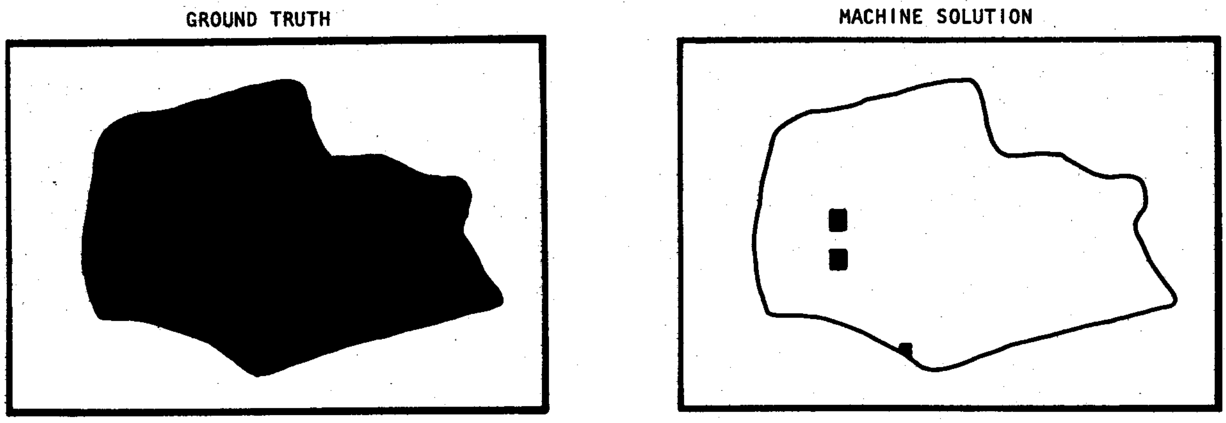

Figure 7. Red Pine Plantation: Interpreter's Solution Left, Machine Solution Right.

that the 1:125,000 scale was the largest reasonable scale that should be formed.

The cooperating field resource managers found the LANDSAT classifications "interesting", but indicated the information did not meet their needs. Providing the accuracy of the classification could substantially be improved, the imagery might possibly be useful for general planning purposes. Errors such as those shown in Figures 2, 3, 4, 5, and 6 for example, are quite unacceptable. Considering the time and effort put into training set selection, and the state-of-the-art classification techniques used, it is doubtful that useful forest land cover maps can be produced from the LANDSAT data. Further efforts might include consideration of scenes recorded on other dates, but achievement of satisfactory land cover maps does not appear promising.

Several data analysis systems were used in this study, including interactive and batch processing. The interactive systems permitted the analyst to work faster and more efficiently, but it became obvious that the interactive capability could only be of value if the analyst was quite familiar with the study area. In the analysis process, each training set had to be selected and the resulting classification (used individually or with other training sets) had to be evaluated on the spot - interactively. Unless the analyst was very familiar with the field conditions over a large portion of the study area, he could not objectively interact with the processing in the manner for which such machines were intended. If only one training set for each category is to be used, and no substitutions are to be evaluated, batch processing is certainly the best method. In many cases, interactive processing amounts to no more (or less) than an entertaining experience for the analyst in the development of themes and viewing attractive colors on the display screen.

It was found absolutely necessary to work as a team in the analysis process: (1) a machine operation specialist accustomed to working with a particular image processing system on a day-to-day basis, and who is familiar with all of the software options and how to access them; and (2) a qualified scientist who possesses a high degree of professional expertise and experience in the resource field pertinent to the application at hand and who is familiar with image processing techniques. It is impossible for the machine operator to fully understand the dynamics of signatures because he is unfamiliar with the canopy geometry, understory conditions, species and density variations, etc., that are found in the study area. By the same token, the resource management specialist cannot be familiar with all of the procedures and knob turnings involved in actual operation of the image processing system. This was particularly the case in this study where several different systems were used. To repeat -- there is no substitute for a team approach to the overall analysis process.

Two algorithms were used to classify the data into the eleven major land cover categories: (1) the parallelepiped approach and (2) a maximum likelihood classifier. The Bendix system, IDIMS system, and the batch processing software (RECOG) all utilize the Gausian Maximum Likelihood algorithm while the Image 100 system performs classification by the parallelepiped method. Although the same data were classified with the M-DAS and the Image 100 systems (using the same categories and nearly identical training sets), the resulting classifications were considerably different.

When the parallelepiped technique was used, less than half of the scene was classified and the remaining pixels did not fit into any of the areas defined in four-dimensional feature space, as established by the training sets. This implies that the training sets were "atypical" examples of the types they were selected to represent. Since all the training sets were validated in the field and carefully located on the graymaps, the results were at first not understandable. Further examination of the photography and more field checking, however, gave a possible explanation: i.e., pure "classic" stands of each type had been selected for use as training sets and were not representative of the types in general, since the natural vegetation patterns of the region are highly variable. Much of the total area is covered by mixes of the various types with complex transition zones and these patterns resulted in intermediate signatures not within the range of radiance values for any of the classes. Using the maximum likelihood classifer, however, permitted classification of more than 98% of the scene (Figure 6). With this algorithm more of the pixels of intermediate value were forced into one of the categories.

When using the parallelepiped method, the selection of training sets is very critical. That is, a training set must be selected which includes the full range of reflectance values for that particular land cover type. Therefore, a great deal of familiarity with the field conditions is needed, and more time is required since many training sets may be necessary. When themes are built by combining training sets, several dimensional units of feature space are required to give a complete signature of a land cover type. In this process, overlaps often occur, the pixels included in the overlap zone are left unclassified and consequently, even more areas are left unidentified. In this study, the overlaps were resolved by using a maximum likelihood method for putting each "overlapping" pixel into one of the prescribed land cover types. In summary, if little is known about the field conditions for a particular site except for a few verified training sets, a less-than-complete classification will often result - especially where signatures are highly variable.

The maximum likelihood classifier can be adjusted so that a larger percentage of the scene is classified. This is possible due to the option of varying the threshold parameter of each category. In essence, the confidence interval can be set on each category so that pixels are forced into one of the land cover types. This is accomplished by setting wider limits (i.e., 2 standard deviations rather than 1 standard deviation) as the surface of the n dimensional boundary in feature space, for each category. The categories can be uniformly or differentially varied according to the expected signature variability resulting from the field conditions of each land cover type. Thus, with the maximum likelihood method, various thresholds can be established without the need for training sets which include all possible field conditions and a large percentage of the scene can be classified. When the thresholds are set wide under the maximum likelihood method, a greater number of commission errors can be expected to result.

Selection of training sets was found to be the most critical portion of the analysis procedure and differed with the method of classification used. When working with parallelepiped, training sets representing the extremes of field conditions for each land cover type were needed. Therefore, very dense mature stands as well as low density and young stands of each vegetation type were needed. With the maximum likelihood method, training sets representing the central tendency or medium field condition for each land cover type were desired. This is not to say that only one training set and, therefore, only one n dimensional "cloud" of feature space was needed for each land cover category. In some cases, the signatures for the extremes of a given land cover type were sandwiched around the feature space for another category. However, it was felt that fewer training sets and less time-consuming training set selection were required with the maximum likelihood. The final major point concerning algorithms is that, in either case, familiarity with the land cover types and the variability of the field conditions for each type over the entire region is absolutely necessary if the optimum classification is to be obtained.

There are advantages and disadvantages to both classification techniques: (1) with the parallelepiped, it is nearly impossible to classify the entire scene, whereas (2) with maximum likelihood, pixels are apt to be classified incorrectly. There seems, however, to be a logical approach to the dilemma considering the tradeoffs. First, the scene should be classified using a parallelepiped technique - the results of which should give the best classification possible. The signatures classified with this technique would have to be quite similar to those for the "classic" training sets and should not include the mixtures and transition zones. Once these classifications are performed, they could be printed with various symbols all in one color - then the remaining pixels can be classified using the maximum likelihood algorithm. The results of this second classification could be printed with the same symbols used in the parallepiped technique but this time in a different color. Such a two-step procedure would permit mapping nearly the entire scene at two levels of confidence. The accuracy could be estimated for areas classified with each technique individually. The user of such a map would thus know for which areas photography and field checking are most needed.

The value of using temporally-registered data is difficult to determine. Two LANDSAT scenes for Itasca County (May 29, 1973 and July 17, 1974) were registered to one another by techniques used at the Jet Propulsion Laboratory, Pasadena, California. Training sets were defined and three classifications were performed using the Image 100 system. A portion of the scene was mapped first using only Bands 5 and 7 from the May imagery, then using the same bands on the July imagery. Finally, using Bands 5 and 7 from both dates simultaneously, a third classification was performed. The accuracies of these classifications were not quantified, however, the three classifications were quite different, although exactly the same training sets were used in each case. Thus, it can be concluded that the degree of crossover and types of crossover changes with the phenology of the vegetation.

The complicating factor in analysis of the temporal data is in the process of selecting the best training sets. Experience gained in this study suggests there should be optimum dates to image each individual land cover type. For example, were it possible to classify lowland conifer just with a May scene, then go on to successive categories using their optimum date for identification, the whole classification combined would give the best possible results. Such a procedure would, however, require extensive processing to resolve the areas which were classified more than once and would be very complex. Consequently, this study simply concluded that the temporally-registered data added too much complexity to the signatures and did not improve the accuracy appreciably.

VI. SUMMARY

Three methods of accuracy verification were applied to eleven land cover categories mapped from LANDSAT data. As a result, it was quite apparent that the accuracy of mapping land cover on large blocks of land, including transition zones and vegetation type mixtures, gave lower estimates of accuracy than was realized on either training sets or test sets.

Final vegetation maps were produced in color at scales of 1:24,000, 1:125,000 and 1:250,000. Application tests of these solutions scales led to the conclusion that, at such time as it is possible to obtain sufficiently accurate land cover solutions, the largest scale which should be used is 1:125,000. The 1:24,000 displays should only be used for classification accuracy verification. May imagery gave better classification than the July imagery; and temporally-registered (May + July) data did not appear to result in more reliable classifications than the May imagery used alone.

Interactive image processing systems were more efficient than the batch system under the following conditions: (1) considerable ground truth was available, and (2) a team consisting of a qualified programmer and an experienced resource applications scientist did the analysis together. There seem to be good reasons for using the maximum likelihood algorithm rather than parallelepiped - especially in areas with highly variable land cover. Also, training sets should be selected by criteria matched with the specific algorithm to be used.

Evaluation of the various map solutions by experienced field resource management cooperators resulted in the judgment that the classification accuracies were so low as to preclude practical use for their purpose at this time.

VII. ACKNOWLEDGEMENTS

This project was supported by NASA Grant NAS5-20985 (administered by the Minn. State Planning Agency), and the University of Minnesota's Agricultural Experiment Station and College of Forestry. A parallel forest aerial photography project funded by the McIntire-Stennis Cooperative Forestry Research Program provided the necessary aerial photography and ground truth. Soil Scientist Grant Goltz and Ranger Francis Voytas (U.S. Forest Service), Land Commissioner William J. Marshall (Itasca County), Forester L. Chris Peterson (Blandin Paper Co.), and certain of their staff members provided field support and evaluations. Fred C. Billingsley, Jet Propulsion Laboratory, did the LANDSAT data temporal registration; and James Wray, U.S. Geological Survey, arranged for geometric correction of the data. Greg Johnson, and other EROS Data Center personnel, provided invaluable assistance and permitted use of the Image 100 and IDIMS Systems. John Peine, Bureau of Outdoor Recreation, kindly permitted the use of their geometrically-corrected LANDSAT tapes and a number of their training sets. Virginia Carter, U.S. Geological Survey, most graciously arranged for the use of the EROS Data Center facilities and Dr. James A. Smith, Colorado State University, provided the RECOG programs and assisted in setting up the system locally; and Susan Moorlag, U.S. Geological Survey, assisted in arranging some of the image processing. James Marshall and Mary Hagen, IAFHE Remote Sensing Laboratory, assisted with the field work and forest type maps and with the pixel counting and drafting, respectively. This presentation is based on a portion of a Ph.D. thesis submitted to the University of Minnesota Graduate School

VII. REFERENCES

1. Marshall, J. and M. Meyer, 1977, A field evaluation of small-scale forest aerial photography: IAFHE RSL Res. Report 77-2, 18 pp.

2. Eller, R., M. Meyer and J. Ullman, 1973, ERTS-1 data applications to Minnesota forest land use classification: IAFHE RSL Res. Report 73-1, 35 pp.

ROY A. MEAD

B.S. (Botany) Northern Arizona University
M.S. (Earth Resources) Colorado State University
Presently graduate student and Instructor in the College of Forestry, University of Minnesota.

MERLE P. MEYER

BS (Forestry-Wildlife) Univ. of Minnesota
MF University of California
PhD (Forestry) Univ. of Minnesota
Presently Professor of Forestry and Director, IAFHE Remote Sensing Laboratory, University of Minnesota.