E82-10109

AgRISTARS

RRI-L1-04143
JSC-17418
NASA-CR-167447

A Joint Program for Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing

Renewable Resources Inventory

September 1981

AN EVALUATION OF ISOCLS AND CLASSY CLUSTERING ALGORITHMS FOR FOREST CLASSIFICATION IN NORTHERN IDAHO

Lee F. Werth

(E82-10109) AN EVALUATION OF ISOCLS AND CLASSY CLUSTERING ALGORITHMS FOR FOREST CLASSIFICATION IN NORTHERN IDAHO (Lockheed Engineering and Management) 21 p HC A02/MF A01 CSCL 02F G3/43

N82-21650

Unclas 00109

Lockheed Engineering and Management Services Company, Inc.
1830 NASA Road 1, Houston, Texas 77058

Lyndon B. Johnson Space Center
Houston, Texas 77058

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
RRI-L1-04143, JSC-17418		

4. Title and Subtitle	5. Report Date
An Evaluation of ISOCLS and CLASSY Clustering Algorithms for Forest Classification in Northern Idaho	September 1981
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Lee F. Werth	LEMSCO-17154
	10. Work Unit No.
9. Performing Organization Name and Address	11. Contract or Grant No.
Lockheed Engineering & Management Services Company, Inc. 1830 NASA Road 1 Houston, Texas 77058	NAS 9-15800
	13. Type of Report and Period Covered
12. Sponsoring Agency Name and Address	Final - August 1981
National Aeronautics and Space Administration Lyndon B. Johnson Space Center Houston, Texas 77058 Tech. Monitor: R. O. Hill	14. Sponsoring Agency Code

15. Supplementary Notes

The Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing is a joint program of the U.S. Department of Agriculture, the National Aeronautics and Space Administration, the National Oceanic and Atmospheric Administration (U.S. Department of Commerce), the Agency for International Development (U.S. Department of State), and the U.S. Department of the Interior.

16. Abstract

This paper represents one of many studies funded under the Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing program to determine how Landsat can contribute to the U.S. Department of Agriculture Forest Service's Renewable Resource Inventory effort as mandated by the National Forest Management Act of 1976. The specific objective of this study was to test clustering algorithms used at Johnson Space Center - ISOCLS and CLASSY. The test site was the Elk River, 1:24,000 quadrangle in northern Idaho.

17. Key Words (Suggested by Author(s))	18. Distribution Statement
Classification Clustering algorithms Forestry Idaho Landsat Remote sensing	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified		

*For sale by the National Technical Information Service, Springfield, Virginia 22161

JSC Form 1424 (Rev Nov 75) NASA — JSC

RR-L1-04143
JSC-17419

FINAL REPORT

AN EVALUATION OF ISOCLS AND CLASSY CLUSTERING ALGORITHMS FOR FOREST CLASSIFICATION IN NORTHERN IDAHO

Job Order 72-542

This report describes activities of the Renewable Resources Inventory project of the AgRISTARS program.

PREPARED BY

L. F. Werth

APPROVED BY

A. V. Mazade, Supervisor
Forestry Applications Section

LOCKHEED ENGINEERING AND MANAGEMENT SERVICES COMPANY, INC.
1830 NASA Road 1
Houston, Texas 77058

Under Contract NAS 9-15800

For

Earth Resources Applications Division

Space and Life Sciences Directorate

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
LYNDON B. JOHNSON SPACE CENTER
HOUSTON, TEXAS

September 1981

LEMSCO-17154

PREFACE

The Agriculture and Resources Inventory Surveys through Aerospace remote Sensing is a multiyear program of research, development, evaluation, and application of aerospace remote sensing for agricultural resources, which began in fiscal year 1980. This program is a cooperative effort of the U.S. Department of Agriculture, the National Aeronautics and Space Administration, the National Oceanic and Atmospheric Administration (U.S. Department of Commerce), the Agency for International Development (U.S. Department of State), and the U.S. Department of the Interior.

The work which is the subject of this document was performed by the Earth Resources Applications Division, Space and Life Sciences Directorate, Lyndon B. Johnson Space Center, National Aeronautics and Space Administration, and Lockheed Engineering and Management Services Company, Inc. The tasks performed by Lockheed Engineering and Management Services Company, Inc., were accomplished under Contract NAS 9-15800.

The following scientists and other personnel contributed to this work:

- Kent Lennington of Lockheed Engineering and Management Services Company, Inc., for assistance with the CLASSY program.
- P. J. Aucoin and Susanne O'Brien of Lockheed Engineering and Management Services Company, Inc., for software development and testing.
- Rigdon Joosten of the National Aeronautics and Space Administration, for technical support and comments.

The author gratefully acknowledges the work of the following persons in support of this task.

- Frederick P. Weber, Program Manager, Nationwide Forestry Applications Program for technical support and comments.
- Irv Johnson and Robert Spencer of the U.S. Department of Agriculture Forest Service, Clearwater National Forest, for field assistance and reference data.

Contents

ACRONYMS

AgRISTARS	Agriculture and Resources Inventory Surveys Through Aerospace Remote Sensing
CIR	color infrared
CCT	computer-compatible tape
CMS	Conversational Monitoring System
CPU	Central Processing Unit
EOD	Earth Observations Division
EOD/LARSYS	Earth Observations Division version of Laboratory for Applications of Remote Sensing System
GSFC	Goddard Space Flight Center
ISOCLS	Iterative Self-Organizing Clustering System
JSC	Lyndon B. Johnson Space Center
LACIE	Large Area Crop Inventory Experiment
LARS	Laboratory for Applications of Remote Sensing
LARSYS	Laboratory for Applications of Remote Sensing System
MSS	multispectral scanner
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
OBC	optical bar camera
pixel	picture element
RIDS	Resource Information Display System
RRI	Renewable Resources Inventory
USDA	U. S. Department of Agriculture
USGS	U.S. Geological Survey

1. INTRODUCTION

One objective of the Renewable Resources Inventory (RRI) project in the Agriculture and Resources Inventory Through Aerospace Remote Sensing (AgRISTARS) program is to investigate the feasibility of using remote sensing analysis techniques for monitoring nationwide forest and range renewable resources. The Current Technology Assessment task, involving research, development, testing, and evaluation, is part of this project.

Two automated data processing systems have been examined to date:

- The classification systems developed by the Large Area Crop Inventory Experiment (LACIE), a joint endeavor of the U.S. Department of Agriculture, (USDA) National Aeronautics and Space Administration (NASA), and the National Oceanic and Atmospheric Administration (NOAA) of the U.S. Department of Commerce.
- Portions of the Resource Information Display System (RIDS), which was developed by the Geometronics Development Group in the Washington, D.C., office of the USDA Forest Service for application in all forest regions.

The purposes of this task are to evaluate the system developed by the Earth Observations Division (EOD) at the NASA Johnson Space Center (JSC) and the Laboratory for Applications of Remote Sensing (LARS) at Purdue University, and to determine its utility in supporting RRI project objectives. The following are specific objectives of the task:

- To evaluate the Iterative Self-Organizing Clustering System (ISOCLS) and CLASSY clustering algorithms for forest land classification.
- To determine the utility of selected clustering algorithms for analysis of forestry and range study sites and subsequently to incorporate the classification products of these sites as one of the overlays to be used in the RIDS.

1.1 BACKGROUND

Forest and range managers require accurate and timely vegetation inventory information, such as species composition, density, productivity, and condition, to make sound management decisions. Data gathered by satellite multispectral scanner (MSS) and analyzed using various classification schemes may help achieve the inventory goals of the AgRISTARS program. The EOD version of the LARS System (EOD-LARSYS) has several classifier options that can be evaluated to find the best one for forest classification.

1.2 SCOPE

The scope of this task is to determine how accurately forest and rangeland classes (U.S. Geological Survey - Anderson) can be identified using ISOCLS and CLASSY. If one of the algorithms can accurately classify forest land, it will be a useful classification tool for forest and rangeland inventories.

1.3 APPROACH

To address this task, the two algorithms will be applied to forest and nonforest classes for one 1:24,000 quadrangle map of northern Idaho. The classification and mapping accuracies of the classes for ISOCLS and CLASSY, the two algorithms to be assessed, were evaluated with 1:30,000 color infrared (CIR) aerial photography. Confusion matrices for the two clustering algorithms were generated and evaluated to determine which one is most applicable to forest and rangeland inventories on future projects.

1.4 OBJECTIVES

The objectives of this study are twofold: (1) to evaluate ISOCLS and CLASSY for land cover classification, and (2) to determine the classifier to use in other AgRISTARS Program test sites.

2. DESCRIPTION

The study area selected for the ISOCLS-CLASSY comparison is in the Clearwater National Forest (Region 1) of northern Idaho (figure 2-1). Upon the request of AgRISTARS for an area to test various remote sensing systems, the Clearwater National Forest staff selected the Elk River 7-1/2-minute quadrangle (1:24,000). This study area, located approximately 83 kilometers northeast of Moscow, Idaho, represents a complex heterogeneous site of the northern Rocky Mountain coniferous forest. The climate, topography (elevation, aspect, and slope), and soils combine to produce coniferous cover types of mixed conifer, western white pine (*Pinus monticola*), western larch (*Larix occidentalis* , Douglas-fir (*Pseudotsuga menziesii*), sub-alpine fir-spruce (*Abies lasiocarpa* - *Picea engelmannii*), western red cedar (*Thuja plicata*), and mountain hemlock (*Larix mertensiana*). The most abundant is the mixed conifer type, in which a single species is not dominant and the amount of each species in the mixture varies by location. Some of the species that may occur in the mixed type are western larch, Douglas-fir, grand fir (*Abies grandis*), western red cedar, western white pine, ponderosa pine (*Pinus ponderosa*), Englemann spruce (*Picea engelmannii*), lodgepole pine (*Pinus contorta*), and sub-alpine fir. Associated species of the larch type may include western red cedar, grand fir, Douglas-fir, ponderosa pine, and western white pine. Grand fir occurs with the western red cedar type, and Douglas-fir and mountain hemlock occur with the sub-alpine fir-spruce type. Associated species of the mountain hemlock type include lodgepole pine and whitebark pine (*Pinus albicaulis*).

Other land cover/land use classes in the quadrangle area include clearcuts, riparian area, urban area (town of Elk River), meadows, and water (Elk Creek Reservoir). The clearcuts, which vary in size and composition, occur throughout the quad, but are most common on the west half. Most clearcuts have shrubs as the dominant vegetation, but may also have rock outcrops, large cedar stumps or overtopping regeneration of Dourglas-fir, sub-alpine fir, lodgepole pine, and Englemann spruce. The most dominant shrub in some clearcuts is shinyleaf (*Ceanothus velutinus*) or redstem ceanothus (*Ceanothus sanguineaus*). Other shrubs that may be dominant or occur with the ceanothus

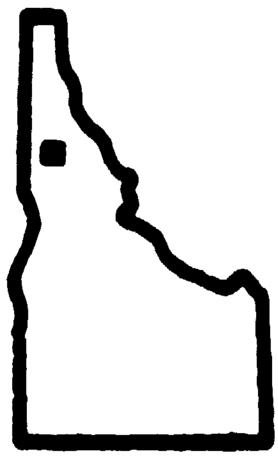

Figure 2-1.- Elk River test site in northern Idaho.

types are elderberry (*Sambucas caerulea*), willow (*Salix* spp), Rocky Mountain maple (*Acer glabrum*), alder (*Alnus sinuata*), and ninebark (*Physocarpus malvaceus*). The riparian area is a cottonwood-willow type (*Populus trichocarpa-Salix* spp), and the meadows contain grasses, sedges, forbs, and shrubs. Some species occurring in the meadows are Idaho fescue (*Festuca idahoensis*), bluebunch wheatgrass (*Agropyron spicatum*), sedges (*Carex* spp), yarrow (*Achillea melletolium*), goatweed (*Hypericum perforatum*), wild strawberry (*Fragaria vesca*), and snowberry (*Symphoricarpos albus*). There are wetland species of cattail (*Typha latifolia*), bulrush (*Scripus* spp), duckweed/waterlily (*Lemma* spp/*Nypha* spp), and willow, where Elk Creek empties into Elk Creek Reservoir and along the fringe of the reservoir.

All five of the management units for the Elk River Planning Unit (Palouse Ranger District) are represented in the study quad. The management units are Elk River foothills, granitic uplands, Elk Creek breaks, intermediate mountain-slope lands, and high-ridge lands. The elevation in the study area varies from 0 to 10 percent in the valleys to 60 percent plus on the mountain faces.

3. METHOD

The objective of this study was to compare two unsupervised techniques, or clustering algorithms, on one 7-1/2-minute quadrangle in northern Idaho. The data sources were satellite digital radiometric values, stored on computer-compatible tape (CCT).

3.1 DATA SELECTION

Digital tapes of the Landsat imagery (ID-293317281) of Idaho were ordered from Goddard Space Flight Center (GSFC) under the AgRISTARS program. Landsat digital tapes from August 12, 1977, were selected because they represent the peak of forest development and are high quality, cloud-free tapes. The color infrared (CIR) optical bar camera (OBC) aerial photographs that will be used to evaluate classification accuracy were taken August 29, 1978. An orthophoto quad (made with aerial photographs taken September 24, 1975) of the Elk River test site will also be used to supplement the CIR photographs.

3.2 PREPROCESSING

The Elk River data set represents a 132-square-kilometer area that is recorded on 28,731 data points or samples. Each data point represents approximately 0.45 hectare on the ground. Digital tapes from GSFC were sent to LARS to be reformatted into a Universal format and to perform a geometric correction of the Elk River data set. Pixel-radiance values were resampled, using the nearest-neighbor rule.

3.3 PROCESSING

All digital processing was done on LARS remote terminals located in JSC Building 17. The LARS host computer at Purdue University, West Lafayette, Indiana, is an IBM 3031 (formerly an IBM 370). The software to run the hardware is called LARSYS, and the version used at JSC is EOD-LARSYS.

3.3.1 CLUSTERING

The two clustering algorithms selected for the study were ISOCLS and CLASSY. Both algorithms have been used at JSC for LACIE. Basically, a clustering algorithm searches for the inherent separability or structure of the data without any prior knowledge or training; this process is opposed to a supervised technique, which does require training to separate the data.

ISOCLS is a clustering algorithm that is similar to ISODATA developed by Ball and Hall (ref. 1). If seeded, or starting, vectors for the desired information classes are not used, then the algorithm initializes its own spectral class mean and, according to the specified parameters, tries to partition the data set into spectral class groups. Unknown samples are compared to determine the group to which they belong (to which spectral class mean vector they are closest, using the city block, or L1, distance); if necessary, a new group is created, to which the sample is then assigned. After one iteration through the data, the mean vectors of the spectral classes are recomputed. The first iteration through the data may terminate with only two or three cluster classes, but subsequent iterations will probably produce more cluster classes through a sequence of splitting, combining, and chaining operations. The main difference between ISODATA and ISOCLS is that ISODATA does not have the chaining operation. For a more in-depth description of ISOCLS see Kan (ref. 2).

The number of cluster classes that is produced for a given data set depends, of course, on the data set complexity and selection of parameters. With ISOCLS many parameters can be varied to produce different results. For example, changing STDMAX from 4.5 to 3.0 will result in more cluster classes, and setting DLMIN (the distance between cluster centers) from 3.2 to 2.0 will also increase the number of cluster classes produced. Other parameters that can be varied include ISTOP (number of iterations); NMIN (minimum number of samples in a spectral class on the first and next-to-last iteration); PMIN (minimum number of samples in a spectral class at the last iteration); maximum number of clusters; and percent N [the percentage of stabilized clusters with standard deviations less than the threshold parameter or STDMAX (maximum standard deviation) in the initial split iteration sequence]. Since all of the parameters are interrelated, changing any one of them will alter the results; thus, finding an optimum set of parameters for a given data set is difficult.

CLASSY is a more sophisiticated algorithm which alternates maximum likelihood procedures with splitting, joining, and eliminating operations (refs. 3,4,5). CLASSY starts with a model and assumes the data is normally distributed. First, the data set is scrambled so that samples may be randomly selected. The algorithm begins with parent clusters and then determines if the distribution of the parent should be broken down into subclusters or be maintained. If the likelihood ratio is higher for the parent than the subclusters, the parent cluster is maintained. The CLASSY algorithm looks at four moments of the mixture density (histogram) to see if the cluster distribution assumes the normal bell shape. The moments looked at are mean vector, covariance matrix, skewness (the measure of symmetry of the tails of the curve), and kurtosis (the measure of the height of the peak of the curve or the flatness of the curve).

Even though CLASSY is not on EOD-LARSYS, it can be run using the LARS version of the IBM Conversational Monitoring System (CMS) 370. The only parameters the analyst can vary with CLASSY are the number of iterations, the size of the smallest cluster (based on a percent of the data set), and the maximum amount of time the program is allowed to run. CLASSY requires far more subroutines than ISOCLS.

3.3.2 LABELING

After final cluster maps were produced for each algorithm, the next step was to use 1:30,000 CIR OBC aerial photographs, stand maps, and the orthophoto quad for assigning information classes to the spectral classes (or labeling).

3.3.3 CLASSIFICATION

The spectral class or classes that represented an information class were then assigned an alphanumeric symbol. The resultant map with information classes becomes a classification map.

3.4 GROUND TRUTH OR REFERENCE BASE

The reference base for evaluating classification accuracy was primarily the 1:30,000 CIR OBC imagery. The OBC is a panoramic camera that produces exposures on a film strip that covers 3.7 by 59.5 kilometers (at nadir). Since there is so much inherent image distortion in the imagery away from the nadir (optical center of the film), a variety of equal-area grids have been developed by the USDA Forest Service Geometronics group in Washington, D.C., for every other exposure of the overlapping stereo pairs. The grid selected for this study contained 1.01-hectare tick marks.

After a preliminary manual photointerpretation of the 1:30,000 imagery, any cover types that could not be identified were examined and verified during a field trip.

3.5 EVALUATION PROCEDURES

The accuracy of the ISOCLS and CLASSY classification maps were assessed with stratified random sampling. The strata such as grass, cut-over, and forest were delineated on an overlay registered with the orthophoto quad. The number of samples, or one-pixel test fields, in each stratum was determined according to the areal extent (in percent) that each stratum covered on the orthophoto quad. For example, if stratum 1 comprised 15 percent of the quad, then 15 percent of the samples were taken in stratum 1. The locations of the test fields were marked on the classification maps, which could then be viewed on a light table after the orthophoto quad (positive transparency) was superimposed on them. The test fields in each stratum were evaluated according to the Landsat and the orthophoto quad - OBC solution.

Confusion matrices for each classifier were generated, and errors of omission and commission were calculated. Overall and class-classification accuracies were also calculated. In addition, overall and class-mapping accuracies were calculated according to the method used by Kalensky and Schrek (ref. 6) While classification accuracy includes only the omission error, mapping accuracy differs, in that it includes both the errors of omission and commission.

4. RESULTS AND DISCUSSION

4.1 PREPROCESSING

The general geometric correction produced by LARS resulted in line printer maps that had a systematic error of two to three pixels east to west and one to two pixels north to south. The greatest error was in the southwest corner of the quadrangle. An attempt was made to improve the positional error with a precision registration, but the error was not improved because a sufficient number of well-distributed points could not be found. The Elk River data set is in a relatively remote area that does not have well-defined roads, and stream intersections were not reliable because many of the streams were low at the time of the Landsat overpass (August).

4.2 PROCESSING

4.2.1 ISOCLS

Many runs of ISOCLS on the Elk River quadrangle were made with different parameters, but it became readily apparent that the interdependence of the parameters would prevent the search for an optimum set of parameters. Therefore, it was decided that PMIN and NMIN would be set at -4 and 0, respectively, so that even a one-pixel cluster would be retained if it was unique. The maximum number of clusters allowed was set at 60, and N was set at 80 percent, based on LACIE studies (ref. 2). The DLMIN parameter was set at 3.2 after it was observed that values below this produced spectral-class pairs that were separated by a DLMIN of less than the amount specified.

The criteria for determining the best value of STDMAX were, first, how well the water pixels were separated from slope shadows and, second, whether the assortments of spectral classes were too numerous to be identified and labeled. Also, the ease of identifying the meadows and the town of Elk River was another factor in judging the algorithm's performance. In essence, the analyst evaluated the cluster maps according to the data set's recognizable land cover features. The optimum number of spectral classes was found to be between 21 and 24, since any lower number resulted in water pixels all over the mountain slopes and since cluster maps with more than 24 spectral classes did not have well-defined meadows.

An attempt was made to separate out the forest cover types, but it became obvious that there was not a unique spectral class (or classes) for any of the forest types. Also, it would be beneficial if the various cut-over areas could have been discriminated because of their variability; but, again, this task was impossible to do in this study. Most of the cut-over areas were clearcuts, but there were also selection and shelterwood cuts. Therefore, for the Elk River data set, the 24 spectral classes were labeled into the information classes of Coniferous Forest, Cut-over, Grass (meadow), and Water.

The overall classification accuracy using the ISOCLS clustering algorithms was 81 percent, and the overall mapping accuracy was 60 percent (table 1). The classification and mapping accuracies for Coniferous Forest were 78 percent and 78 percent, respectively. Forest class was confused with Cut-over, Grass and Water. Most of the omission error was with Cut-over. It should be noted, however, that some of the spectral classes that have been labeled Cut-over could indeed be low-density forests. The classification and mapping accuracies of the Cut-over class were 72 percent and 55 percent, respectively. The omission error was equally divided between Coniferous Forest and Grass. This result is not surprising, since the area receives high annual precipitation (114.3 millimeters and greater), and the cut-over areas (especially the clearcuts) are quickly filled in with shrubs and grasses. Also, standing trees left on the cut-over areas contribute to the confusion with the Coniferous Forest class. The classification and mapping accuracies of the Grass class were 100 percent and 50 percent, respectively. Both Coniferous Forest and Cut-over contributed to the commission error. The classification accuracy of the Water class was 100 percent, but since some of the Coniferous Forest was classified as water, the mapping accuracy was only 80 percent. Since this error occurred along the edge of the reservoir, it could have been caused by pixel-positional accuracy or shadows near the water's edge.

4.2.2 CLASSY

CLASSY was developed to be used for clustering agriculture crop areas, not wildlands. Since the largest data set it could previously cluster was a LACIE segment (9 by 11 kilometers), software changes had to be made to do a quadrangle-size data set. Because this was the first time CLASSY had been used in forest classification, the experience gained in agricultural work was used initially to set the algorithm parameters. It was recommended that the number of iterations be between three and seven and the total run time be set at 150 minutes (ref. 7). Three iterations were usually used for LACIE work, but since the data set was very heterogeneous, initial trials were run with four and five iterations. When six iterations were used, the algorithm converged better, so most of the runs were made with six iterations. Compared to ISOCLS, the smallest size cluster class that can be maintained is 1 or 2 percent of the data set. Using 28,731 samples, for example, the smallest cluster class at the 1-percent level would be 287 pixels, or samples. Theoretically, any cluster class smaller than this would not be retained at the end of the final iteration. Because of LACIE problems of running CLASSY with 1 percent of the scene, it was decided to use 2 percent of the scene as the smallest cluster class. Later, the 2-percent cluster size was changed to 1 percent.

At the 2-percent threshold, only six cluster classes were produced. The Elk Creek Reservoir is between 50 and 100 pixels in area, so it was too small to show on the cluster maps. With the 1-percent threshold the number of spectral classes decreased from six to five, and the Water class was still too small to be separated. The resulting five spectral classes were labeled into the information classes of Coniferous Forest, Cut-over, and Grass (meadow).

The overall classification and mapping accuracies using the CLASSY clustering algorithm were 77 percent and 67 percent, respectively (table 2). Though the Coniferous Forest classification accuracy was 12 percent higher than it was using ISOCLS, the mapping accuracy was 1 percent lower. Both Cut-over and Grass were confused with Coniferous Forest, but there was more omission with Grass than was the case with ISOCLS. Grass also contributed to a commission error for the Coniferous Forest class. The classification accuracy was 31 percent lower. Most of the Cut-over samples were actually Coniferous Forest. Pixel-positional accuracy and some of the problems already mentioned with ISOCLS contributed to the errors. The classification accuracy of Grass was 7 percent lower than ISOCLS, yet a higher commission error resulted in a mapping accuracy that was 19 percent lower. Spectral response overlap between grass, shrub, and trees is believed to result in the commission error for both ISOCLS and CLASSY. Since the Water class was too small to be discriminated by CLASSY, no water comparison can be made between CLASSY and ISOCLS results.

Both ISOCLS and CLASSY indicated that there was a grass meadow in the north-west portion of the quadrangle, but aerial photos showed that the area was a cut-over area which had grass between the remaining standing trees. The area in question is not like the grass meadows in the southern part of the quadrangle, which are frost-pocket areas (created by deforestation) that can no longer support tree seedlings.

The Cut-over class is very important to forest managers, and any accuracy lower than 70 percent probably would not be acceptable to them. Recent cut-over areas have a better chance of being detected, especially the clean clear-cuts. However, as cut-over areas are filled in with grass, shrubs, or seedlings, the chances for discrimination are decreased. More work needs to be done in this area since change detection and updating of geographic information systems will greatly assist land-management planning in the future.

If terrain data had been used, some of the individual forest-cover types could have been separated, particularly the mountain hemlock and the Douglas-fir types.

The classification and mapping accuracy of ISOCLS and CLASSY differ little, except for some of the classes. CLASSY requires more Central Processing Unit (CPU) time per run, than does ISOCLS; however, ISOCLS requires more trial and error runs, since it has so many parameters that can be varied. ISOCLS in the unseeded mode requires 10 to 20 runs to separate land cover/use classes compared to only two to three runs with CLASSY for a given data set. If ISOCLS is used, it is recommended that starting (seeded) vectors be used.

TABLE 4-1.- ACCURACY EVALUATION OF LANDSAT FOREST CLASSIFICATION USING ISOCLS CLUSTERING ALGORITHM

Class	Landsat					Omissions		Mapping accuracy
	A	B	C	D	Total	No.	%	
Coniferous Forest (A)	100	13	9	6	128	22	22	78
Cut-over (B)	6	31	6		43	12	28	55
Grass (C)			15		15	0	0	50
Water (D)				24	24	0	0	80
Total indicated	106	44	30	30	210			
Total committed	6	13	15	6				
Percent commission	6	30	50	20				
Overall classification accuracy	81%							
Overall mapping accuracy	69%							

TABLE 4-2.- ACCURACY EVALUATION OF LANDSAT FOREST CLASSIFICATION USING CLASSY CLUSTERING ALGORITHM

Class	Landsat				Omissions		Mapping accuracy
	A	B	C	Total	No.	%	
Coniferous Forest (A)	135	5	10	150	15	10	77
Cut-over (B)	24	12	9	45	33	73	24
Grass (C)	1		14	15	1	7	41
Total indicated	160	17	33	210			
Total committed	25	5	19				
Percent commission	16	29	58				
Overall classification accuracy	77%						
Overall mapping accuracy	67%						

5. CONCLUSION

ISOCLS in a pure, unsupervised mode is an ad hoc algorithm that requires many trial-and-error runs to find the proper parameters, such as STDMAX, to separate desired information classes. On the other hand, CLASSY is a more refined algorithm, which in a single run tells the analyst more concerning the classes that can be separated. One major drawback to CLASSY is that important forest and range classes that are smaller than a minimum cluster size will be combined with other classes; futhermore, the algorithm requires so much computer storage that only data sets as small as a quadrangle can be done at one time. However, CLASSY appears to show more promise for forest stratification than ISOCLS and shows more promise for consistency. This study is not conclusive: more research needs to be done comparing the two algorithms in different areas and using any new improvements to either ISOCLS or CLASSY.

ORIGINAL PAGE IS OF POOR QUALITY

6. REFERENCES

Ball, G. H.; and Hall, D. J.: ISODATA, an Iterative Method of Multivariate Analysis of Pattern Classification. Proceedings of the International Communication Conference (Philadelphia), June 1966.

⟨an, E. P. F.: The JSC Clustering Program ISOCLS and Its Applications. ockheed Electronics Company, Inc., (Houston), Technical Report LEC-0483, uly 1973.

ennington, R. K.; and Malek, H.: The CLASSY Clustering Algorithm escription, Evaluation, and Comparison with the Iterative Self-Organizing ustering System (ISOCLS). Lockheed Electronics Company, Inc. (Houston), chnical Memorandum LEC-11289, March 1978.

nnington, R. K.; and Rassbach, M. E.: CLASSY - An Adaptive Maximum elihood Clustering Algorithm. Proceedings of the Ninth Annual Meeting the Classification Society (North American Branch), (Clemson, S C.), 1978.

ington, R. K.; and Rassbach, M. E.: Mathematical Description and Pro- Documentation for CLASSY, An Adaptive Maximum Likelihood Clustering od. Lockheed Electronics Company, Inc., (Houston), Technical randum LEC-12177, April 1979.

sky, Z.; and Scherk, L. R.: Accuracy of Forest Mapping from Landsat ter Compatible Tapes. Proceedings of the Tenth International Sympo- on Remote Sensing of Environment, (Ann Arbor, Mich.), Oct. 1975.

gton, K.: Personal Communication. Lockheed Engineering and Manage- ervices Company, (Houston), June 26, 1979.

NASA-JSC