Reprinted from

Seventh International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Range, Forest and Wetlands Assessment

June 23 - 26, 1981

Proceedings

Purdue University
The Laboratory for Applications of Remote Sensing
West Lafayette, Indiana 47907 USA

TECHNIQUES TO UPDATE A LAND MANAGEMENT INFORMATION SYSTEM WITH LANDSAT

C.A. NELSON

Technicolor Graphic Serivces, Inc.
EROS Data Center, Sioux Falls, South Dakota

D.E. MEISNER

University of Minnesota
St. Paul, Minnesota

B. SMEKOFSKI

Minnesota State Planning Agency
St. Paul, Minnesota

Abstract

The Minnesota State Planning Agency has developed a geographically referenced Land Management Information System which is being used extensively for planning purposes. Land use categories in the system were originally coded from aerial photographs; this method is inefficient for updating the large-area data base. Landsat data and many computer-assisted techniques are available to analyze the classification system and to update the land use data base. The data derived from a Landsat analysis could be used to supplement the existing data base and to complement detailed interpretations of aerial photographs.

This study had as its primary objective an evaluation of computer manipulation, classification, and accuracy assessment techniques for use in updating land use data in the Land Management Information System. Four approaches to statistical computer manipulation (polygons selected from cathoderay tube displays, unsupervised clustering, polygons selected from aerial photographs and data extracted from the existing land use data base) were attempted. The resulting statistics were applied to the image data by three pattern-recognition algorithms: minimum distance to the mean, maximum likelihood, and canonical analysis with minimum distance to the mean. Twelve output images were compared to photointerpreted samples, ground-verified samples, and the current land use data base for accuracy assessment.

The results of this study indicate that for a reconnaissance inventory, statistical computer manipulation via polygons selected from aerial photographs applied with the canonical analysis and minimum distance algorithm is the most accurate and efficient approach. Cross-tabulation with the accuracy samples indicated classification accuracies between 20 to 40 percent. These accuracy levels could probably be increased with the availability of appropriate seasonal coverage and the collection of more timely multidate supporting data.

I. BACKGROUND

The Minnesota Land Management Information System (MLMIS) is a geographically referenced data base developed and maintained by the Minnesota State Planning Agency's Land Management Information Center (LMIC). The system was developed in an attempt to centralize storage and analysis of data on Minnesota's resources. MLMIS was developed cooperatively on the University of Minnesota computer by the University's Center for Urban and Regional Affairs and the State Planning Agency. MLMIS is both a depository of geo-graphically based information and a computer analysis system. It is used for applications such as siting of power plants, assessment of productivity for tax purposes, and location of combined resources.

The MLMIS data base includes natural resource variables such as soils, watersheds and climate, public ownership data, and local political unit boundaries. Information is stored in computer files by units of public land survey for every 40-acre parcel in the State. It is organized by region, county, and township and can be accessed for mapping or statistical analysis. Besides the 40-acre-parcel data base, data have often been collected for parcels of 2.5 acres for special small-area studies. For statewide studies, a file of data consisting of 5-kilometer cells has also been created.

Computer software has been developed to retrieve and manipulate the data and to produce information in tabular, statistical, map, or computer file form. The software is called the Environmental Planning and Programming Language (EPPL) and can be used on MLMIS data or any other data in a compatible format.

The land use data in MLMIS were produced using interpretation of high-altitude, black-and-white, 1:90,000-scale photographs acquired during the springs of 1968 and 1969. The nine classes of land use, given in table 1, were chosen so photointerpreters could determine the dominant land use within each 40-acre cell without much ancillary data.

The MLMIS data base has to be kept as current as possible so that the studies and decisions made by analyzing the data base are valid. Changes in land use, for example, from the pasture and open category to the urban residential category would make a difference in how an area is assessed for power-plant siting. Data resolution also should be improved because the 40-acre cell size is too large for use in many studies. These factors led to the recent LMIC acquisition of new polygon-digitization capabilities and to the decision to revise the geographic base of the MLMIS to a 100-meter grid referenced to Universal Transverse Mercator (UTM) coordinates. This new computer system also provides an added capability to use computer-assisted classification of digital data, such as Landsat, if it can be determined which classification algorithms are the most accurate and efficient. This study used existing software and the latest available Landsat data at the Earth Resources Observation System (EROS) Data Center to evaluate statistical training, classification, and accuracy assessment for accuracy and efficiency in updating the MLMIS data base.

II. ANALYSES PERFORMED

A diversity of training, classification, and accuracy assessment techniques were applied in the study for comparative purposes. Training sets were selected using four different selection procedures:

(1) Supervised polygon selection on a cathode-ray tube (CRT) display.
(2) Unsupervised clustering.
(3) Polygons digitized from interpretations of color-infrared aerial photography.
(4) Existing digital land use data (MLMIS data) as binary masks.

These four training sets were used as input to three classification algorithms:

(1) Minimum distance to the mean.
(2) Maximum likelihood.
(3) Minimum distance to the mean after canonical transformation.

Table 1.--Dominant land use for 40-acre parcels, determined from aerial photographs.

[Source - Land Management Information in northwest Minnesota, Report Number One, MLMIS study, Minneapolis: University of Minnesota, Center for Urban and Regional Affairs, 1972. Source Map Data - 1969. MLMIS Update - 1976. Interpretation by MLMIS staff.]

CLASS	TYPE OF LAND USE
1. Forested	Containing at least 10-percent crown cover of deciduous or coniferous trees.
2. Cultivated	Dominant type of land use appears to be recently tilled or harvested land.
3. Water	Permanent open water is the dominant surface feature.
4. Marsh	Dominant type of land use is nonforested, vegetated areas which are permanently wet.
5. Urban Residential	Containing five or more residential buildings and no commercial buildings.
6. Extractive	Dominant type of land use is the extraction of minerals and includes such features and facilities as mines, tailings, gravel pits, quarries, crusheries, and storage facilities.
7. Pasture and Open	Dominant type of land use is pasture land or land not used for any other identifiable purpose.
8. Urban and Nonresidential or Mixed Residential	Containing at least one commercial, industrial, or institutional development. Examples: schools, factories, hospitals, athletic fields, business districts, churches, warehouses, military installations, sewage disposal facilities junk yards.
9. Transportation	Dominant type of land use is facilities for the conveyance of people and/or materials. Examples: airports, railroad yards, highway interchanges, rights-of-way.

The three classifications were then applied to two data bases:

(1) System-corrected Landsat data (resampled to the Hotine Oblique Mercator map projection). These data were resampled by the nearest neighbor method to a 50-meter UTM grid following classification.
(2) System-corrected Landsat data resampled to a 50-meter UTM grid. These data were subjected to cubic convolution resampling (in effect, re-resampled data) prior to classification. This was needed for the binary-mask training selection because of the geographic referencing.

Finally, the results were assessed for accuracy by:

(1) Comparison to digitizer-encoded polygons which were photointerpreted from high-altitude color-infrared photography.
(2) Comparison to a random set of ground-verified sample plots.
(3) Comparison to the current (that is, 10-year-old) 40-acre-parcel MLMIS digital land use base.

III. STUDY AREA DESCRIPTION

Two criteria were used in selection of the study area. First, the system-corrected data had to be available in order to assess questions of advantages and/or disadvantages of geometric correction. This also tested the software for reading new-format[2] computer-compatible tapes (CCT's). Unfortunately, this criteria limited us to a single scene acquired on a non-optimum date (24 May 1979), making multidate analysis unfeasible. Figure 1 shows the Landsat band 5 scene which was used in the study.

Second, a study area located in an "urban fringe" region of the available scene was sought. This would provide data of the greatest utility to the planners in terms of updating their information. Urban fringe regions are also the most difficult areas for Landsat analysis because of the wide range of spectral characteristics of constituent cover types and considerable spatial complexity.

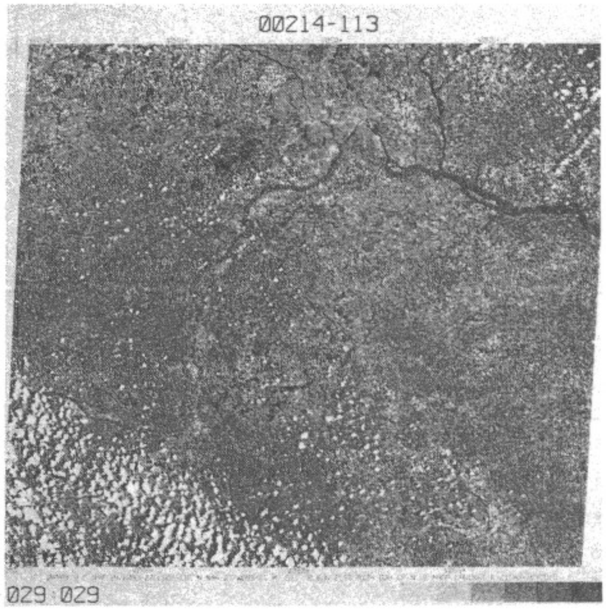

Figure 1. Landsat scene, band 5, shows the Twin Cities of Minnesota at the confluence of the Minnesota and Mississippi Rivers

The study area covered four U.S. Geological Survey 7.5-minute quadrangles (map names: Centerville, Hugo, White Bear East, and White Bear West) with White Bear Lake approximately in the center of the area. Distances between the various stages of urban development (high density, suburban, and exurban) and rural farmland are minimal in this northeast portion of the Twin Cities (Minneapolis and St. Paul). The area also contains many lakes, wetlands, and forested areas. Figure 2 shows the Landsat band 5 subscene that was selected for study, and the subscene illustrates the complexity of the region.

IV. DESCRIPTION OF ANALYSES

A. IMAGE CORRECTION

Radiometric correction by destriping.--The selected subscene was entered into the EROS Data Analysis Laboratory's interactive digital image analysis system and checked for missing data by displaying each of the four Landsat multispectral scanner (MSS) bands. The next step was to radiometrically correct the data for any residual striping caused by nonlinearity in the MSS detector responses. The algorithm used to do this is a procedure for "smoothing by a thresholded window." The program tests brightness values of pixels in a window of specified size (in this study, 7 lines of 7 samples each) against the brightness value of the center pixel. If the absolute difference between the center pixel and the 48 neighbors is less than a specified threshold (3 brightness levels in this study), the center pixel is replaced by the average of the 49 pixels. The size of the window and the threshold values had previously been determined empirically by studying water areas because these most clearly show the effects of the corrections.

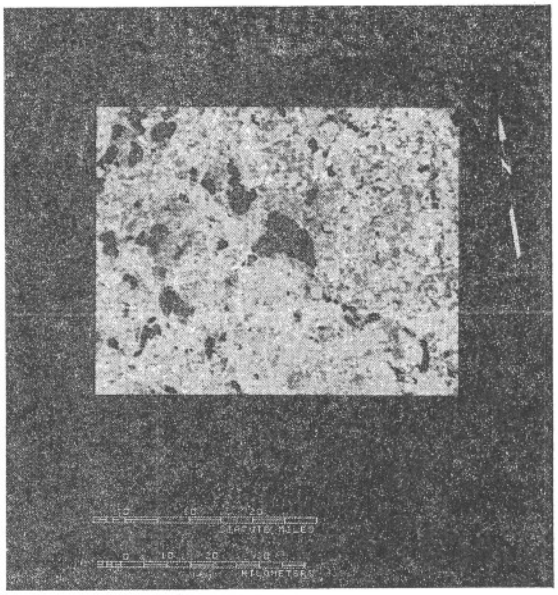

Figure 2. Landsat, band 5, subscene of the selected study area

Figure 3 is a side-by-side comparison of image data over White Bear Lake before and after "smoothing." The left-hand portion of the image illustrates the limitations of the radiometric correction applied to this system-corrected data set. The right-hand portion shows that the striping effect apparent in the left-hand side image has been reduced. The image in fig. 3 was enhanced to show the water area, but the algorithm has the same effect on all the data which meet the spatial and spectral criteria.

Geometric correction.--After the radiometric correction by smoothing was completed, geometric correction to a 50-meter UTM grid was done. This involved selecting 31 ground-control points located in the image and correlating their image coordinates to 7.5-minute map coordinates. (A CRT display and sonic digitizing table were used.) These points were used in a least-squares analysis to derive a second order polynomial transformation.

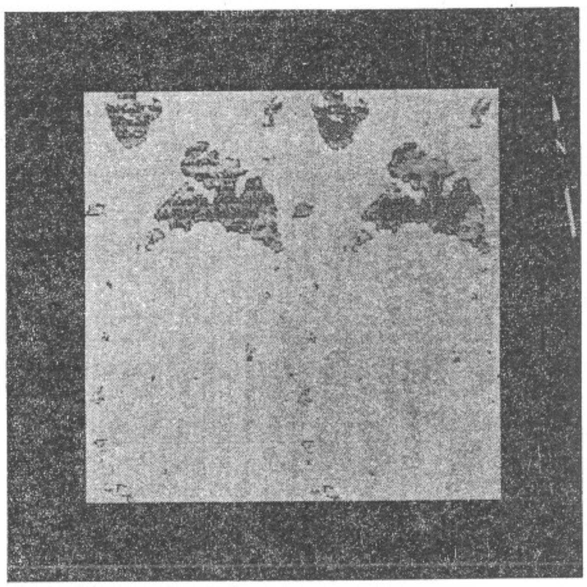

Figure 3. A comparison of before and after thresholded smoothing

The multispectral Landsat image was then registered, once before classification and once after classification, to a grid which allowed 50-meter pixel spacing and was aligned to the north in a UTM projection (fig. 4). Because it was desirable in the first registration that the least amount of spatial degradation occur, cubic convolution resampling was used. The selected study area registered to the 50-meter UTM grid is shown in fig. 4. In the second registration, performed after classification, nearest neighbor resampling was used. Registering all the data to the same grid allowed the comparison of the results by single sets of points in the accuracy assessment.

B. TRAINING SET SELECTION

Statistical training of the classification algorithms was done by four methods: (a) supervised polygon selection from CRT display, (b) unsupervised clustering, (c) airphoto-based supervised clustering, and (d) training from the existing MLMIS land cover data base.

Supervised polygon selection from CRT screen(A).--Polygons were selected on a color CRT on enlarged subscenes by an analyst familiar with remote sensing but not with the study area. Reference was made to 7.5-minute quadrangles to assist in the selection of areas falling within the

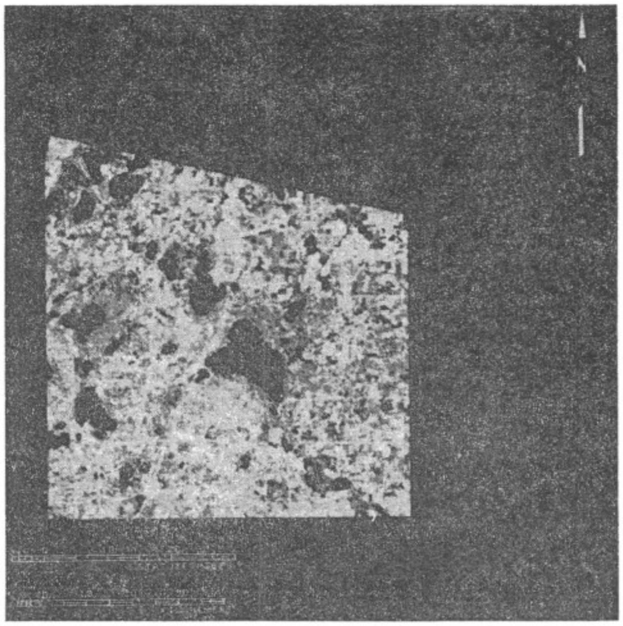

Figure 4. Landsat, band 5, subscene of the selected study

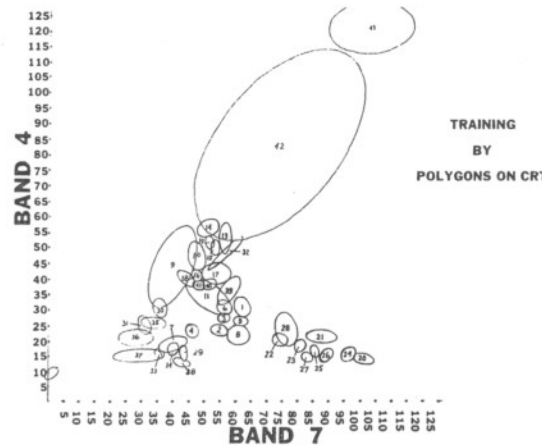

Figure 5. Spectral comparison plot of Landsat band 4 vs. Landsat band 7 training by polygons on CRT

desired land use classes. The polygons were not made so small as to be homogeneous training fields; they were large enough to contain approximately 5 to 10 spectral classes of land cover within the designated land use. Several polygons for each land use class were selected. A clustering algorithm[6] developed the mean vectors and covariance matrices for spectral clusters within each set of land use polygons.[1/] The statistics from these individual clusterings were consolidated into one file of 84 clusters and then were reduced to 43 by deleting clusters which overlapped. Overlap was determined by a separability algorithm[9] which uses the saturated transformed divergence as the test of separability of clusters. A divergence value of 1200 was used as the level below which clusters were deleted. Figure 5 is a spectral comparison plot showing the means in bands 4 and 7. The cumulative density function contour is 68 percent. This plot shows that in two dimensions many of the clusters overlap and that the contours vary greatly in size; there are also gaps in the spectral space definition by these clusters.

Unsupervised Clustering(B).--Unsupervised clustering was done of the entire area. If the area had been larger, a random sample could have been taken to reduce the computation time. Other studies[7,10] have shown that even 1- or 2-percent samples are representative.

The clustering algorithm, using the same parameters as used for the individual polygon clustering, grouped the data into 64 clusters which were then consolidated, again using the separability algorithm, to 60 clusters. Figure 6, another comparison plot with the same parameters as fig. 5, shows that there training data had little overlap and that with a few exceptions the 68-percent cumulative density function contour were of similar size.

Polygons Digitized from Interpretations of Color-Infrared Aerial Photography(C).--Polygons which had been interpreted by analysts who were familiar with the study area and who had done some field verification were digitized from 1:24,000-scale color-infrared aerial photographs. These polygons were composited in the analysis system by registering the photographs to maps, using ground reference points which were visible on both. The polygons were then used as masks to extract only those selected portions of the data. The selected portions were then clustered using the same separability parameters as before.

[1/] The parameters used in the program (known as ISOCLS) were: STDMAX = 2.0, DLMIN = 1.6, ISTOP = 10, NMIN = 10, MAXCLS = 64, CHNTHS = 1.6.

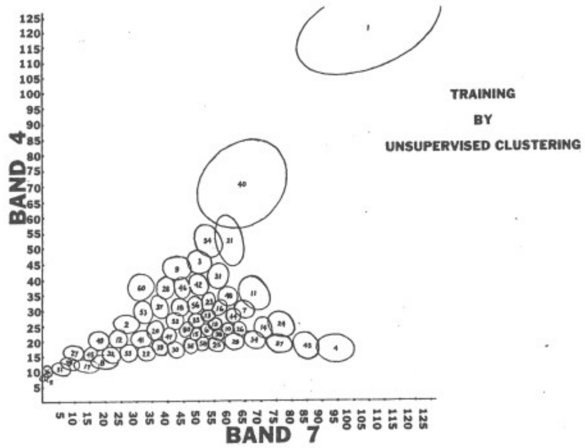

Figure 6. Spectral comparison plot of Landsat band 4 vs. Landsat band 7 training by unsupervised clustering

This technique was similar to technique A except for the source of the polygons. In this technique, 68 consolidated clusters were tested for separability and then reduced to 60 clusters. In fig. 7, the comparison plot shows that there was much spectral overlap between the two bands, possibly indicating classification confusion. Also, there were gaps in the spectral space definition of these clusters.

Existing MLMIS Digital Land Use Data as Binary Masks(D).--Data were extracted from the multispectral information by creating binary masks for each land use class from the existing MLMIS digital land use data. That is, marks for forested, cultivated, water, marsh, urban residential, extractive, pasture and open, urban and nonresidential, and transportation classes were created. These data had been resampled from an original 40-acre grid to the 50-meter cells used in this study. The resulting masked data were then clustered using the same parameters as in previous techniques.

This technique was similar to techniques A and C. The resulting 238 clusters were consolidated together and then were reduced to 58 clusters by the separability algorithm. Figure 8 illustrates the magnitude of spectral overlap in the two bands, possibly indicating classification confusion. As with the other polygon techniques, there were some spectral gaps, but not as many because of the large number of initial clusters.

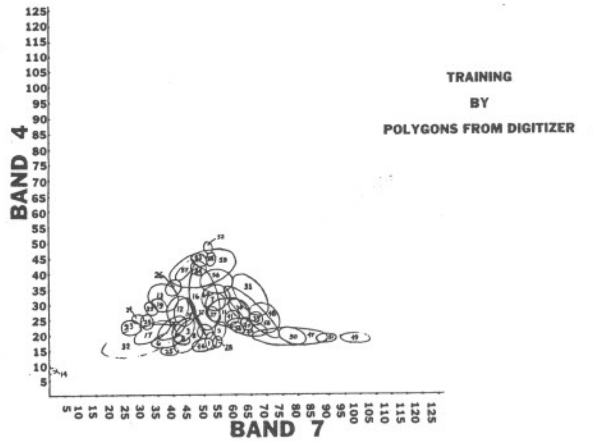

Figure 7. Spectral comparison plot of Landsat band 4 vs. Landsat band 7 training by airphoto digitizing

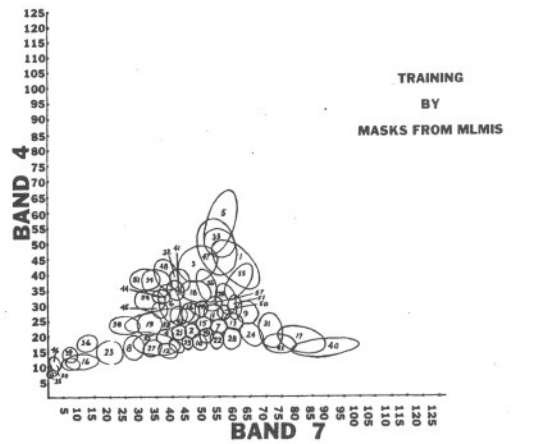

Figure 8. Spectral comparison plot of Landsat band 7 training by binary masks

C. PATTERN RECOGNITION (CLASSIFICATION) ALGORITHMS

The statistics developed by the training techniques described above were applied to the image data by three pattern recognition algorithms: minimum distance to the mean, maximum likelihood, and canonical transformation with minimum distance to the mean.

Minimum distance to the mean (city-block distance)(1).--This algorithm tests each multispectral vector of the data against the mean vectors in the statistics and assigns each pixel to that cluster which is closest in spectral space. A maximum-distance parameter is used to threshold all pixels which are not within a "reasonable" distance of any mean.

In this study, the same distance threshold (12.0 units) was used for all four iterations (one for each training technique) of the algorithm.

Maximum likelihood with threshold(2).--This technique computes a likelihood value for each pixel based on the mean and co-variance of each cluster and assigns the pixel to that cluster which has the maximum likelihood value.[1] The algorithm outputs two images. The first is the cluster assignment, and the second is the likelihood value transformed to values which are indicators of chi-squared values. The chi-image is used for thresholding those values which are not within a specified percentage.

A 5-percent threshold value was used for thresholding the classifications from all the training techniques.

Canonical transformation with minimum distance to mean(3).--This technique uses linear transformations of the data to make the classifier more accurate. Canonical analysis[5] transformations increase the separability of clusters while minimizing the differences occurring within each cluster. A linear transformation to uncorrelated variables is produced that has greatest amount of variance in the first variable and lesser amounts in the succeeding variables.

The image data that were transformed by the coefficients developed with this algorithm, and the transformed statistics for the training data, were input to the minimum-distance classifier.

D. ACCURACY ASSESSMENT

After classification by the three algorithms, the spectral cluster images were grouped according to the MLMIS land use classes. This grouping was done on the CRT by analyzing the spatial distribution of each cluster, using maps, aerial photographs, and knowledge of the area.

The grouped data (fig. 9 is an example) were then assessed for accuracy by three comparison techniques: comparison with photointerpreted reference data, comparison with the MLMIS data base, and comparison with ground-verification sample points.

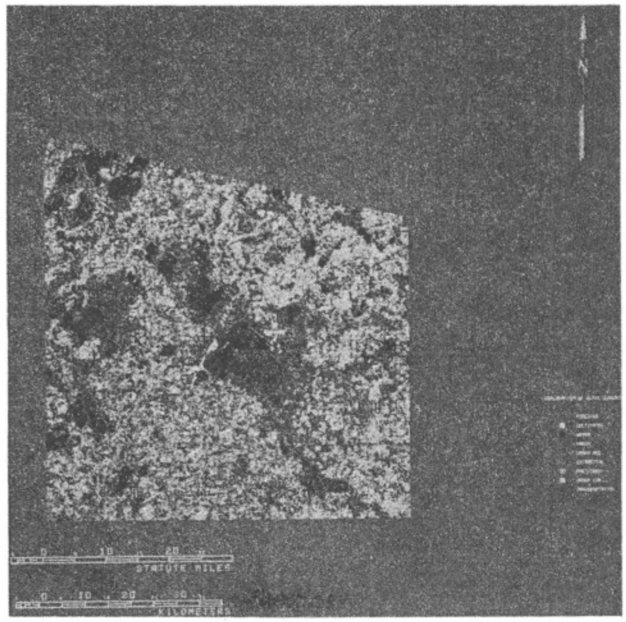

Figure 9. Color-coded image of training by unsupervised clustering with classification by canonical analysis-minimum distance to mean

Comparison with photointerpreted reference data.--The photography had been acquired in July and August of 1977 in support of wetlands mapping of the eastern metropolitan Twin Cities area. The coverage consisted of 9-by-9-in. color-infrared transparencies at a scale of 1:24,000. Six scenes covering approximately 30 percent of the study area were selected for analysis, and contact prints were generated. These prints were photointerpreted into 18 land cover classes, shown in table 2, by an analyst familiar with the area. About 12 hours of interpretation time were required. Because of a lack of overlap in the available prints, the interpretation was done monoscopically, but this did not present any problems in view of the scale and interpretation classes being used. A minimum mapping unit of one acre was used.

Table 2.--White Bear Lake Photointerpretation Classes

Land Use Class	Type of Cover
Urban	1. Commercial/Institutional
	2. Transportation
Residential	3. Residential-High Density
	4. -Medium Density, Little Forest
	5. -Med-Low Dens., Partly Forested
	6. -Low Density, Heavily Forested
Forest	7. Forest-Mature
	8. -Shrub Forest
Agriculture	9. Pasture-Open
	10. Cropland, Uncultivated
	11. Cropland, Cultivated
Wetland	12. Open Wetland
	13. Emergent Lake Vegetation
	14. Open Water
Miscellaneous	15. Golf Course
	16. Bare Land
	17. Extractive
	18. Forest, Plantation

Figure 10 is the interpretation overlay of one of the photographs. The complexity of the area and the irregular boundaries are well illustrated by this overlay.

The interpretations were input to the analysis system by digitizing the polygons from the interpretation overlay. After the digitizing was checked, an image having data values corresponding to the class numbers in table 2 was generated and registered to the Landsat subscene. For the verification, the 18 photo-interpreted classes were grouped into the eight MLMIS classes.

The accuracy of the training technique and classification combinations was then determined by a program which creates contingency tables of the two input images for comparison.

Comparison with the MLMIS data base.--Since the MLMIS data base and the training technique/classification images were already registered with one another, all that was required was to run the contingency table program to do the accuracy assessment.

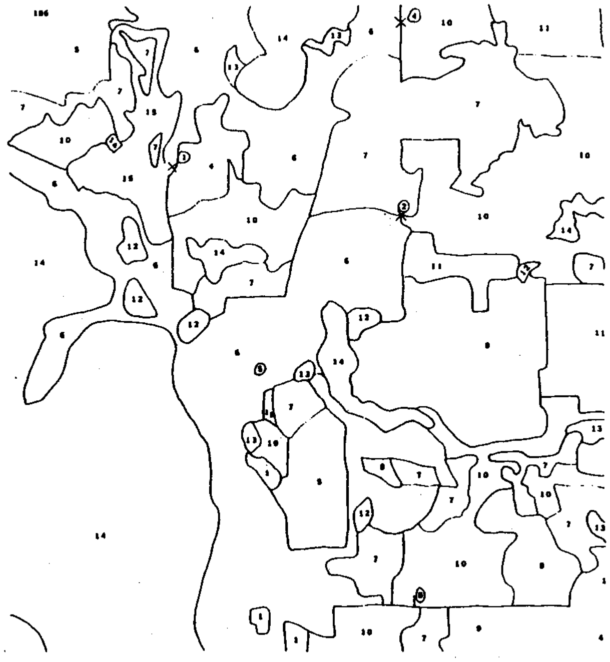

Figure 10. Interpretation overlay showing the complexity of the classes and the irregular boundaries of this study area

Comparison with ground verification sample points.--A set of 240 ground points were randomly selected, plotted on the 7.5-minute quadrangles, and then visited on the ground. The ground visits consisted of locating each point and noting the land use for that area. Test points were extracted from the training technique/classification images by using the digitizer and a point-extraction program. The ground information was then coded and input to the analysis system so that contingency tables could be processed with the sets of data.

V. RESULTS

This study indicates that there is no significant difference in the accuracy of geometric corrections done either before or after classification. The trade-off is in the computer time required for registration versus the efficiency and accuracy of training-set selection. Training by the exist-ing digital land use data base required that the registration be done before the training. The geometric transformation used for registration had residual mean errors of 23 meters in the X axis and 30 meters in the Y axis at the ground control points (root mean square error= 29 meters).

A comparison of the training techniques was made based on the amount of variance in each of the canonical analysis axes, shown in table 3. From this comparison, technique D appears to be best for capturing the variability of the image.

Table 3.--Comparison of the training techniques in each of the canonical analysis axes

Canonical Axis	Training Technique A	B	C	D
	Percent Variance			
1	83.86	90.57	90.62	92.47
2	15.44	9.37	8.83	7.37
3	0.51	0.04	0.37	0.12
4	0.19	0.02	0.18	0.03

The accuracy assessment techniques resulted in the rankings in tables 4, 5, and 6. These tables indicate that the most accurate means of training and classification is to digitize polygons from aerial photos and apply canonical analysis with a minimum-distance-to-mean algorithm.

Table 4.--Techniques ranked by comparison with photointerpreted data through contingency tables (values in percent).

Image	Rank	Percent Agreement
C3	1	37.6
A3	2	34.1
C1	3	33.6
B3	4	33.0
B1	4	33.0
B2	5	32.3
D3	5	32.3
C2	6	31.7
D1	7	31.3
D2	8	31.1
A1	9	28.9
A2	10	25.3

Table 5.--Techniques ranked by comparison with Minnesota Land Management Information System data through contingency tables (values in percent).

Image	Rank	Percent Agreement
C3	1	30.4
A3	2	29.7
B1	3	28.6
B3	4	28.0
C1	5	27.9
B2	5	27.9
A1	6	26.7
D1	7	26.6
D3	8	26.4
D2	8	26.4
C2	8	26.4
A2	9	23.0

Table 6.--Techniques ranked by comparison with randomly selected ground verification sites through contingency tables.

Image	Rank	Percent Agreement
A3	1	31.4
C3	2	31.2
A2	3	28.1
D1	4	28.0
A1	5	27.1
D2	6	26.9
D3	7	26.6
B3	8	26.0
B1	9	25.9
C1	10	24.8
B2	11	24.6
C2	12	18.8

After an evaluation of the contingency tables produced at EROS Data Center, and a simple visual assessment of the classified images, the LMIC staff agreed that the classification using digitized training sets processed with a minimum-distance-to-mean algorithm after canonical transformation was the best of the 12 methods used. The image that was classified in this way was subsequently transferred by tape from EDC to MLMIS, and an EPPL-compatible file was created to be used for comparison with existing MLMIS land use data.

The MLMIS land use data were resampled to a registered 50-meter grid that permitted cross-tabulation with the 50-meter Landsat classification.

It was also deemed desirable to aggregate the 50-meter Landsat data in 40-acre cells. This was done using the ZOOM program, a scale reduction and expansion routine for use with EPPL files in the MLMIS. In this application, the 50-meter data were reduced by a factor of eight to create 400-meter cells. Assignment of data values was done with a dominance rule. The 8-by-8-pixel grid window was determined by a grid of the specified factor superimposed on the data starting at the first row and column of data. The cross-tabulations showed that with 50-meter data there was an average correlation of 34.93 percent, whereas the 40-acre cells provided for an average correlation of 37.65 percent.

Besides the accuracy of the classification, the speed with which it is accomplished is very important. Table 7 shows the computer times required for training in the study. Training by polygons from a CRT is the most efficient. These training times, when combined with the computation times for the classification, shown in table 8, permit a comparison of the efficiency of individual factors affecting overall classification speed.

The combination of accuracy and timing factors show that training by polygons from aerial photos with the canonical analysis and minimum-distance-to-mean classification was the most accurate and efficient. However, this technique combination requires a skilled interpreter familiar with the area and classes desired.

VI. RECOMMENDATIONS

Many factors accounted for the low accuracies achieved in this study. In particular, the date of coverage of the Landsat image (selected primarily because it was the only cloud-free system-corrected image available at the time) was not optimum for some of the land use classes. This is certain to have affected classification accuracies and, as a result, may have influenced the technique comparison. This situation was compounded by the fact that none of the source data used in the study were acquired at the same times. The Landsat data were acquired in May of 1979, the aerial photography was acquired in 1977, and the ground verification data were collected in November and December of 1979. Such disparities seem likely to have been responsible for some of the errors in the grouping of the clusters into classes and the errors in verification.

Low accuracies not withstanding, the qualitative results of the comparison performed in this study appear valid.

If the project were to be repeated, significant improvements could be realized by increasing the lead time allowed for the collection of source data. This would ensure a greater probability of obtaining optional data in terms of both image content and temporal commonality.

Additional improvements could result from modifying the classification software, which could include:

(1). Implementation of a software test procedure in all classifiers so that masked areas (zero in all bands) would be placed in class 0 without any computation. This would make the use of masking a registered image by a geographic information system more feasible and productive. With this technique, the land cover within the land use could be determined. More ground verification and comparison with photo-interpreted data could improve the classes separated by the classification algorithms.

(2). Development of an output file of upper and lower bounds (95-percent confidence limits) from the canonical analysis, development of a parallelepiped classifier using these bounds, and development of a weighted minimum-distance classifier. These techniques would probably provide better results than the city-block, minimum-distance classifier currently used with the canonical analysis technique.

The intent of LMIC is to continue evaluation of the classified land cover data. Using the Landsat classified data with masks of existing MLMIS 40-acre data, LMIC plans to produce an updated land use/land cover map. The variables they will use in the masking process will include ownership, forest cover, the 1969 land use classification, and political boundry information. Although these data exist only at 40-acre-cell resolution, it is planned to produce a land use/land cover variable using 100-meter cells (approximately 2.5 acres) as the mapping unit.

REFERENCES

1. Fu, K. S., Landgrebe, D. A., and Phillips, T. L., 1969, Information processing of remotely sensed data; Proceedings of the Institute of Electrical and Electronic Engineers, v. 57, no. 4, April 1969, pp. 639-653.

2. Ginevan, M. E., 1979, Testing land use map accuracy: another look: Photogrammetric Engineering and Remote Sensing, v. 45, no. 10, pp. 1371-1377.

3. Holkenbrink, P. F., 1978, Manual on characteristics of Landsat computer com-

Table 7.--Results of study and computer times for training.

(All timings expressed are for an HP3000 CX computer; cpu is Central Processor unit time in seconds; cm is total connected time in minutes.)

Polygons from CRT used	531 cpu	185 cm	for selecting polygons
	468 cpu	106 cm	for clustering
	999 cpu	291 cm	
Unsupervised clustering used	38,479 cpu	690 cm	
Polygons from aerial photos used		120 minutes	for photo interpretation
	503 cpu	202 cm	for digitizing polygons
	544 cpu	128 cm	copying strata
	4,958 cpu	116 cm	clustering
	6,005 cpu	446 cm	
Binary masks from MLMIS used	965 cpu	56 cm	for mapping
	5,193 cpu	190 cm	for masking (multiplies)
	94,129 cpu	3,858 cm	for clustering
	100,287 cpu	4,104 cm	

Table 8.--Computation Time Comparison Between Training and Classification Techniques units: central processor units, seconds/connect time, minutes

Classification Technique	Training Technique: Mean	A CRT polygons	B Unsup.	C Photo polygons	D MLMIS mask
1. Minimum distance	4,032/76	4,473/76	3,603/64	4,452/85	3,599/79
2. Maximum likelihood	12,613/250	13,208/227	12,032/207	13,070/224	12,141/345
3. Canonical analysis	5,049/119	5,601/139	4,725/138	5,212/101	4,657/98

patible tapes produced by the Earth Resources Observation System Data Center Digital Image Processing System, U.S. Geological Survey, 1978, 78 pp.

4. Hord, R. M., and Brooner, W., 1976, Land use map accuracy criteria: Photogrammetric Engineering and Remote Sensing, v. 42, no. 5, p. 671-677.

5. Jenson, S. K., and Waltz, F. A., 1979, Principal components analysis and canonical analysis in remote sensing: American Society of Photogrammetry Annual Meeting, 45th, Washington, D.C., 1979, Proceedings: Falls Church, Va., American Society of Photogrammetry, v. I, p. 337-348.

6. Minter, R. T., 1972, ISOCLS-iterative self-organizing clustering program: Program C094, Johnson Space Center, Houston, Texas, 1972, 140 pp.

7. Rohde, W. G., Miller, W. A., Bonner, K. G., Hertz, E., and Engel, M. F., 1979, A stratified-cluster sampling procedure applied to a wildland vegetation inventory using remote sensing: XIII International Symposium on Remote Sensing of the Environment, Ann Arbor, Mich., 1979, Proceedings: Environmental Research Institute of Michigan (in press).

8. Swain, P. H., 1972, Pattern recognition: a basis for remote sensing data analysis: Laboratory for Applications of Remote Sensing Information Note 111572, Purdue University, West Lafayette, Ind., 1972, 40 pp.

9. Swain, P. H., Robertson, T. V., and Wacker, A. G., 1971, Comparison of the divergence and b-distance in feature selection: Laboratory for Applications Remote Sensing Information Note 020871, Purdue University, West Lafayette, Ind., 1971, 9 p.

10. Todd, W. J., and Gehring, D. G., 1979, Vegetation and terrain mapping of Lake Mead National Recreation Area using Landsat digital data (submitted publication of the Department of Geography and Geology, Indiana State University).

11. Van Genderen, J. L., and Lock, B. F., 1977, Testing land use map accuracy: Photogrammetric Engineering and Remote Sensing, v. 43, no. 9, p. 1135-1137.

Nelson, Charles A. Charles Nelson joined Technicolor Graphic Services, Inc., the prime contractor at the EROS Data Center, in September 1974. He is a Senior Data Analyst in the Data Analysis Laboratory where techniques in digital analysis of remotely sensed data are developed, implemented and taught. Prior to joining TGS/EROS, Mr. Nelson worked as a forestry scientist with Lockheed Electronic Corporation at Johnson Space Center. He worked with 24-band aircraft scanner and was in on the initial Landsat forest application experiments. Mr. Nelson received Master's and Bachelor's degrees in Forestry from the University of Minnesota.

Meisner, Douglas E. Douglas E. Meisner received his B.A. and B.Arch. degrees with honors from Syracuse University. He received his M.S. in Environmental and Resource Engineering from the SUNY College of Environmental Science and Forestry. He is a Research Specialist in the University of Minnesota Remote Sensing Laboratory, where he is responsible for the development of computer-aided analysis techniques to support the Lab's research programs. He also participates as a data analyst in many of these programs.

Smekofski, Bob Bob Smekofski received his B.A. (Major in Geography) from the University of Minnesota with the Land Management Information Center of the Minnesota State Planning Agency. Mr. Smekofski worked at the Center for Urban and Regional Affairs, University of Minnesota and studied toward a M.A. prior to joining MSPA.