D,

N81-12506

SECTION A

LANDSAT APPLICATIONS TO WETLANDS CLASSIFICATION IN THE UPPER MISSISSIPPI RIVER VALLEY

(Final Report)

Dr. T. M. Lillesand and Lee F. Werth
Remote Sensing Laboratory
Institute of Agriculture
University of Minnesota
St. Paul, Minnesota

INDEX

5

Even with increased recognition since the 1960's of wetland importance, growing world population and food demand have resulted in extensive wetland drainage. Since World War II nearly 25% of the wetlands in the Prairie Pothole Region of the Dakotas and Minnesota has been drained. Sound federal, state, and local management decisions on acquisition, preservation or drainage of wetlands must be based on sound inventory data. Currently, the U.S. Fish and Wildlife Service is involved in a National Wetland Inventory to satisfy these data needs at the federal level. Likewise, various state and local agencies are undertaking, or have completed weltand inventories for a range of purposes. For example, in Minnesota wetland inventories serve as baseline data to evaluate site modification permits, non-point pollution sources, and generally to execute the public water law.

In short, accurate classification and mapping of wetland cover types is considered essential to the implementation of a number of federal, state, and local legislative mandates and resource management programs. At the same time, the task of mapping wetlands on a statewide basis is challenging in terms of cost, time and manpower. On a national level, the task is monumental.

The general objective of the federal wetland inventory is to classify and locate the nation's wetlands as accurately as possible on 1:100,000 (or in some cases 1:24,000) scale maps (Cowardin et al., 1977). To accomplish this, two possible remote sensing data bases suggest themselves: (a) high altitude, small-scale aerial photography and/or (b) LANDSAT. Of the two, aerial photography appears to be the favored choice from the standpoint of spatial resolution - but photography is not always available and, when available, it not always adequate for the purpose.

Neither technique, LANDSAT in particular, had been sufficiently tested in areas representative of Minnesota prior to this study. Thus, this study was proposed to test the applicability of automated processing of LANDSAT multispectral scanner (MSS) in the Minnesota regional context.

STUDY OBJECTIVES

Initially, the sole objective of this study was to test, under local conditions, the capability of Landsat data analysis techniques to position adequately and classify wetlands in support of the National Wetlands Inventory. This involved a comparison of single-date vs. double-date sets as well as a comparison of data sets preprocessed with a coarse vs. a precision geometric correction. The test area chosen for these comparisons consisted of five contiguous 7.5-minute quadrangles located within the western part of the Twin City 7-County Metropolitan Area (Figure 1). The entire Metropolitan Area has been mapped previously through airphoto interpretation by the Remote Sensing Laboratory under contract with a consortium of funding agencies.[1] (In addition, in-house funds had already been used to perform the single-date Landsat analysis of the study area.)[2] The large amount of existing data on this area provided a good test base for the study. The test site contains one of the most concentrated areas of wetland diversity in the west half of the Metropolitan Area. It is a

---

[1] The Minnesota Department of Natural Resources, the U.S. Army Corps of Engineers, the U.S. Fish and Wildlife Service, the USDA Soil Conservation Service, and the Twin Cities Metropolitan Council. (Werth et al., 1977; Owens and Meyer, 1978).

[2] The University of Minnesota Agricultural Experiment Station and the College of Forestry.

OF POOR QUALITY

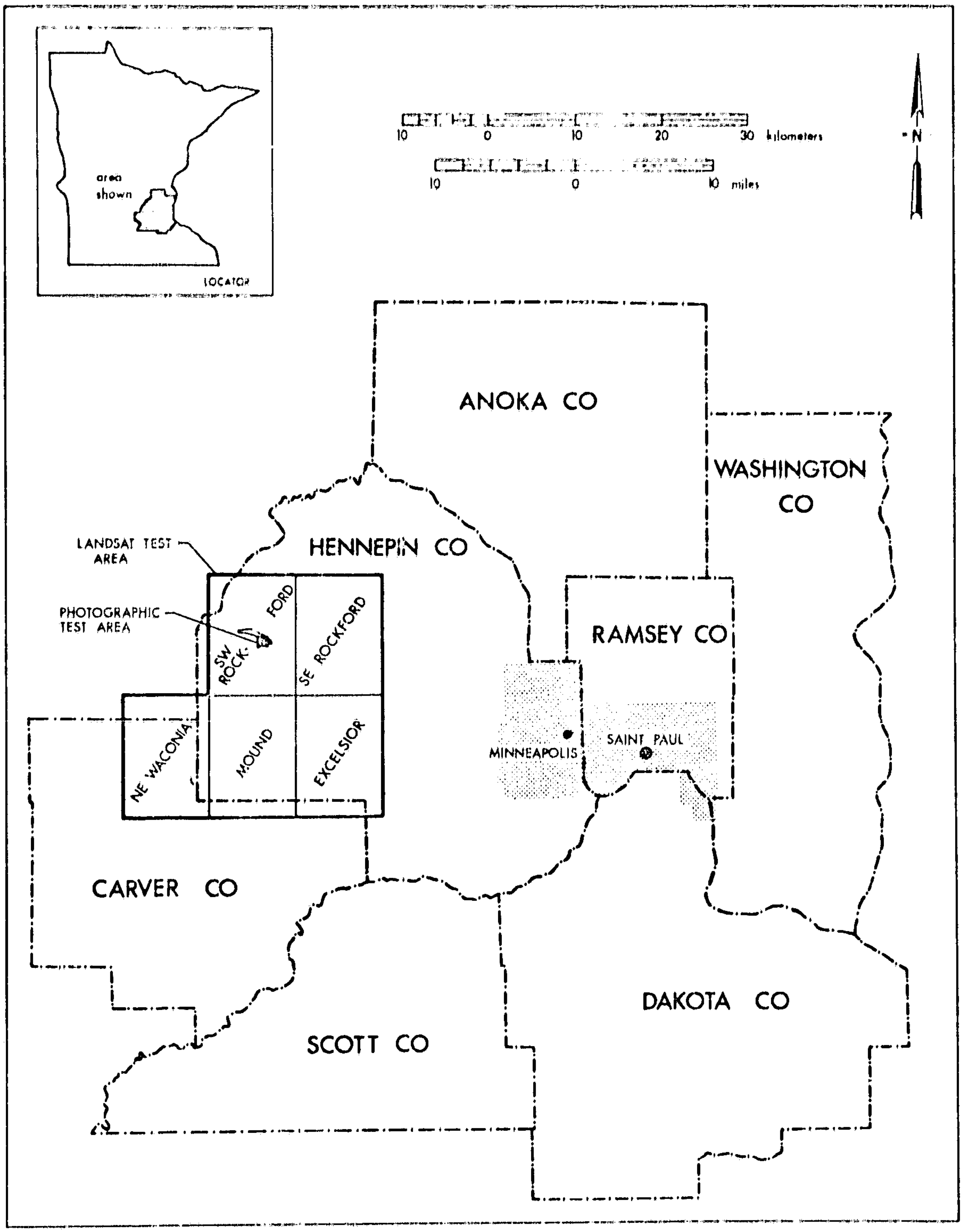


Figure 1. Study Site (located in the western part of Hennepin County, Minnesota).

complex, heterogeneous area of undulating glacial topography representatime of a transition zone between tall grass prairie and northern hardwood forest. The landscape is interspersed with small farms, villages, cities, subdivisions, parks and golf courses.

As the tests of the various LANDSAT data analysis techniques were being applied to the 5 quadrangle test area, a parallel test of digitized color infrared photographic data was undertaken in a limited sub-area (see Figure 1). The digitized photographic data virtually removed the spatial resolution restrictions inherent in the LANDSAT data. These restrictions were found to be acute due to the spatial complexity of land cover in this area. In this way, the photographic analysis permitted a more basic test of the spectral characteristics and separability of the wetland types. This was deemed important for judging the potential applicability of Landsat data analysis procedures in areas which are less complex spatially than those in the Metropolitan study area. It also provided some insight into the potential advantages to be realized with the improved spatial resolution of the Thematic Mapper on LANDSAT-D. In addition, the photographic research permitted an initial investigation into the potential utility of using digitized color infrared photography per se as a data source for wetland mapping in spatially complex areas.

The methods and results of the LANDSAT test and the test of the digitized photography comprise, respectively, the next two sections of this report.

LANDSAT DATA ANALYSIS

As previously indicated, both single-date and double-date LANDSAT data sets were analyzed in this study. Both data sets were processed at the Purdue University Laboratory for Applications of Remote Sensing (LARS). The single-date analysis was performed on data acquired on July 6, 1976. A coarse geometric correction (i.e., not using ground control points) was applied to this data set. The correction resulted in an average pixel positioning error of approximately 2 pixels N-S and 4.5 pixels E-W when image data were compared to a 1:24,000 map base. This correction uses a transformation model based on nominal satellite parameters and the geographic position of the test site, to rescale, deskew and rotate the original data. Following the geometric correction, classification was performed using the LARS hybrid multi-block cluster classification procedure (Hoffer and Fleming, 1978). This procedure entails selecting blocks of training areas containing heterogeneous cover types, clustering these blocks individually (unsupervised), identifying the informational identity of cluster classes, pooling training statistics from similar classes in various blocks, and using these statistics to perform a maximum likelihood classification of the entire area of interest. Following this procedure, thirty-five spectral classes were extracted from the single-date data to represent the following information classes: Non-wetland, Forest, Shrub, Emergent, Submergent and Water. Overall classification accuracy as measured from 2555 randomly selected pixels of verified identity was 72%. (That is, the total number of pixels classified correctly, ÷ 2555 = 0.72). However, the individual classes had highly variable classification accuracies. For example, 96% of the Water test pixels were classified correctly, whereas only 10% of pixels known to be of the Shrub class were classified correctly. The average of all of the class accuracies was 41%, which was deemed inadequate for the application at hand.

The generally poor classification performance and geometric correction results of the single-date analysis were improved upon in the double-date analysis process. First, application of the LARS geometric correction, using ground control points, resulted in average additional pixel positional corrections of approximately 0.5 pixels N-S and 0.75 pixels E-W. The geometrically corrected data for two dates were registered and the multi-block cluster technique was used for classification. (In this case the July 5, 1976 data were merged with those of August 7, 1975). The LARSYS feature selection process was employed to judge which of the eight available bands of data was optimum for subsequent classification. Using this process, bands 5 and 6 from both dates were deemed best suited for the classification. Given the poor performance of the Shrub class in the single-date analysis, this class was eliminated from the double-date classification. Also, it was hypothesized that separation between the Non-wetland class (basically corn) and Emergents (cattail) would improve with the introduction of the August data, since the corn had tasseled by this date.

The test field accuracy of the double-date classification was evaluated in a manner similar to the single-date effort. This involved evaluation of some 450 randomly selected test fields that varied in size from one to nine homogeneous pixels. The cover type present in each field was verified by ground site visits. Table 1 summarizes the results of comparing the LANDSAT and ground data. As can be seen from Table 1, the

Table 1. Test Field Accuracy Assessment of Double-Date LANDSAT Classification†

Known Cover Types		% Classified As Indicated Cover Type			
Cover Type	No. of Test Pixels	Non-wetland	Forest	Emergent	Water
Nonwetland	1,706	75	8	17	
Forest	120	34	56	10	
Emergent	121	33	30	35	2
Water	523			2	98

Overall Classification Accuracy: 77%

Average Accuracy Per Class: 66%

---

†A detailed discussion of the classification and mapping accuracy of the single and double-date LANDSAT analyses is given in Werth (1980).

overall classification accuracy for the double-date analysis was 77%. The accuracy for the various information classes ranged from a low of 35% for the Emergent class to 98% for Water. The per class average accuracy was 66%. The comparative levels of errors of omission and commission for each class are also shown in Table 1.

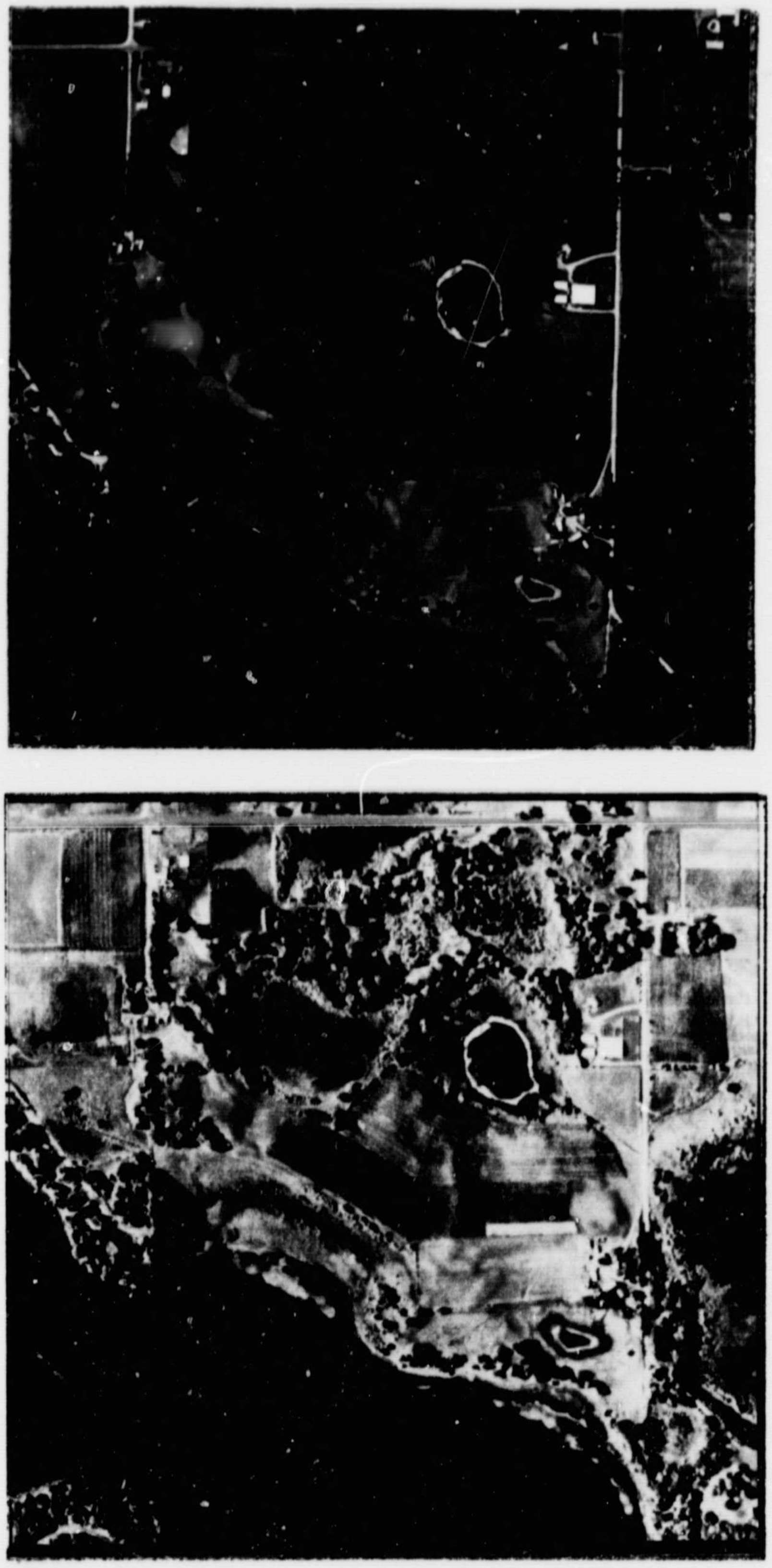
In short, the two-date results indicated an increase in overall classification accuracy of approximately 5% (from 72 to 77%) and an improvement in per class classification accuracy of about 25% (from 41 to 66%). Additional classification accuracy appears to be precluded by the spatial and spectral resolution limits of the LANDSAT MSS relative to the complexity of the study area. This is not to say that LANDSAT has limited application to wetland mapping in general. Rather, we have simply concluded that one should expect poor success in single-date analyses and only moderate success in double-date analyses of wetland areas which are as complex as those in our study area.

ANALYSIS OF DIGITIZED PHOTOGRAPHY

As mentioned previously, digitized color infrared aerial photography was also analyzed in this study as an extension of the LANDSAT analysis effort. Again, this source of digital data virtually removes the spatial resolution restrictions limiting the LANDSAT automated classification process, thereby permitting more thorough study of the spectral properties of the wetland cover types under investigation. Constraints of time and funding limited the photographic analysis to coverage of a single study site located within the LANDSAT test area (Figure 1). This site contains a number of small wetland areas which generally characterize those of the surrounding region.

Figure 2a is a (1.7x) enlargement of the original 70mm position transparency used in the photographic analysis. The original image was a 1:50,000 photograph (Kodak type 2443) taken with a Hasselblad 500 EL (f=50mm) camera on July 21, 1977, in support of previous wetland research efforts conducted by the Remote Sensing Laboratory. The image was digitized using a modified P-1700 Optronics drum scanner made available by the University of Wisconsin Environmental Monitoring and Data Acquisition Group (UW/EMDAG). Density readings from 0-3D were digitized into 256 levels and recorded on computer compatible tape. Broad band separation filters were used to obtain the approximate image density record for each of the three film layers. A measurement spot size of 50μm was employed, so that each pixel represented a ground area approximately 2.5m square. The spectral and spatial quality of the resulting digital data can be seen in the color rendering of the digitized data shown in Figure 2b. This figure was generated with the University of Minnesota D-47 Dicomed Color Image Recorder. The digital data from each image band were contrast enhanced by a histogram equilization algorithm prior to being recorded. This algorithm alters the distribution of original image values so that each display level comprises approximately an equal area portion of the display image in each band of data. Figure 2b represents a digitally generated color composite of the three original bands enhanced in this manner.

The photographic analysis effort took on two forms: 1) a full-scene classification, the intent of which was to classify all pixels and cover types present in the scene, and 2) a sub-scene classification,



(a)

(b)

Figure 2. Original color infrared aerial photograph of test area (a) and film recorder output of digital microdensitometer data (b).

to classify wetland areas and cover types only. In both cases, a supervised classification approach was employed using the "Eppler Optimization" form of the maximum likelihood classifier that is integrated into the University of Minnesota Image Processing Software (UMIPS). A description of the two classification procedures follows.

Full-Scene Classification

Training for the photographic classifications was accomplished by photo-interpreting training areas on prints produced from the enhanced Dicomed image. The boundary of each training area was then digitized in terms of x and y image coordinates. In turn, these coordinates were converted to column and row addresses in the scanning microdensitometer data using a polynomial two-dimensional coordinate transformation. This transformation took the form

$$\text{column no.} = a_1 + a_2 x + a_3 y + a_5 x^2 + a_6 y^2$$

$$\text{row no.} = a_7 + a_8 x + a_9 y + a_{10} xy + a_{11} x^2 + a_{12} y^2$$

A least squares observation equation solution for coefficients a_1, $a_2 \ldots a_{12}$ was performed by measuring both the image and line printer output coordinates of readily identifiable ground control points in the scene. Using 10 points for this purpose, the root mean square error (rms) of the transformation was found to be 0.89 pixels in the x and 0.74 pixels in the y image direction.

Training statistics were developed for the 13 major cover types appearing in the scene: Water, Duckweed (Lemna minor), Water Lily (Nymphaea spp.- Nuphar spp.), Cattail (Typha spp.), Reed Canary Grass (Phalaris arundinacea), Willow (Salix spp.), Forest, Cropland, Pasture, Bare

Soil, Road, Roof, and Shadow. Multiple spectral classes (25) were derived from 35 training areas to adequately train for the 13 cover types.

Table 2 is a confusion matrix indicating the training field classification accuracy obtained in the full-scene analysis. Note that for some types the accuracy of classificaion in the training areas was excellent (e.g., Pasture = 97%). Other types, most notably Willow, were very poorly classified (25%). Figure 3 is a color-coded Dicomed image showing the results of applying the full-scene classification. As can be seen from both Table 2 and Figure 3, similar to the case with the LANDSAT data, many of the cover types present in the scene could not be spectrally separated. However, of particular note is the fact that virtually all wetland types were mutually separable. That is, errors of omission and commission in each wetland type were primarily attributable to the spectral confusion between wetland and non-wetland cover types, rather than between wetland cover types themselves.

Sub-Scene Classification

Because of the spectral confusion between the wetland and non-wetland cover types, it was hypothesized that if the training and classification procedures were limited to wetland areas only, the classification accuracy would improve greatly. Hence, the training process was repeated in wetland areas only, using the following classes: Duckweed, Water, Water Lily, Cattail, Reed Canary Grass, and Willow. The boundaries of the wetland areas were delineated again using conventional image interpretation. This process requires much less labor than discriminating all cover type boundaries occurring within the wetlands. The purpose here was simply to separate wetland and non-wetland image areas. Spectral pattern

Table 2. Classification Performance for Training Fields Used in Full-Scene Analysis

Known Cover Type		% Classified As Indicated Cover Type												
Cover Type	No. of Train. Pixels	Duckweed	Water	Water Lily	Cattail	Reed Canary	Willow	Forest	Cropland	Pasture	Bare Soil	Road	Roof	Shadow
Duckweed	67	87		2					4		1	6		
Water	1,700		81								1			18
Water Lily	240			89					10	1				
Cattail	119				91		9							
Reed Canary	664					74		10	16					
Willow	255				17	3	25		55					
Forest	48			2		4		73	17	4				
Cropland	3,392			4		2	6	1	86		1			
Pasture	815			1					2	97				
Bare Soil	232	7							11		81	1		
Road	400	2							1	3	1	93		
Roof	56												100	
Shadow	86		7											93

Overall Classification Accuracy: 84%

Average Accuracy Per Class: 82%

Color Key

Duckweed - magenta

Water - blue

Water Lily - magenta

Cattail - orange

Reed Canary - red

Willow - yellow

Forest - black

Cropland - tan

Pasture - green

Bare Soil - brown

Road/Roof - white

Shadow - black

Figure 3. Full-scene classification of digital photographic data.

Table 3. Classification Performance for Training Fields Used in

Known Cover Type		% Classified As Indicated Cover Type					
Cover Type	No. of Train. Pixels	Duckweed	Water	Water Lily	Cattail	Reed Canary	Willow
Duckweed	67	88		4	8		
Water	1,700		100				
Water Lily	240			100			
Cattail	119	1			89		10
Reed Canary	664					97	3
Willow	255				40	21	39

Overall Classification Accuracy: 94%

Average Accuracy Per Class: 85.5%

Table 4. Classification Performance in Eight Wetland Test Areas

Known Cover Type		% Classified As Indicated Cover Type				
Cover Type	No. of Test Pixels	Duckweed	Water	Water Lily	Cattail	Reed Canary
Duckweed	984	95		3	2	
Water	6,173	2	97			1
Water Lily	2,240	11		72	1	16
Cattail	13,586				81	19
Reed Canary	9,920	1		5	2	92

Overall Classification Accuracy: 87%

Average Accuracy Per Class: 87%

recognition was then used to provide the more subtle delineation of wetland species, a more time consuming process for the photo-interpreter. Table 3 indicates the classification accuracy obtained in the "wetland only" training areas. Note the improvement in training field classification accuracy which results when non-wetland classes are eliminated from the analysis. Figure 4 shows the results of the sub-scene classification. As can be seen from Table 3 and Figure 4 the Willow type introduced the largest errors in the sub-scene classification.

As a final classification effort, the Willow type was dropped and the resulting statistics were again applied to the "wetland only" subset (with predominantly willow areas excluded) of the original image. The accuracy of this classification was evaluated in eight test areas located in the scene. The result of the classification in these areas is shown in Figure 5. The classification accuracy in each of the eight test areas was obtained by comparing the computer-derived classification to that obtained by photointerpretation. This was done for all pixels in the eight areas (approximately 33,000). In those cases where the photointerpretation was ambiguous, a field visit to the site was made to be certain of the correct identification of the cover type present.

The results of the pixel-by-pixel accuracy assessment are given in Table 4. Errors in the classification of Duckweed and Water are believed to be caused by edge effects, in which a given pixel actually overlays two cover types near a boundary between the two types. To some extent, this effect enters into the accuracy problems for Water Lily as well. It is felt that the classification accuracy of the Water Lily and Reed Canary classes could be improved somewhat with additional training. Most of the Water Lily/Reed Canary confusion occurred in an

Color Key

Duckweed - magenta

Water - blue

Water Lily - magenta

Cattail - orange

Reed Canary - red

Willow - yellow

Figure 4. Sub-scene classification of photgraphic data in wetland areas only.

Color Key

Duckweed - magenta

Water - blue

Water Lily - magenta

Cattail - orange

Reed Canary - red

Figure 5. Sub-scene classification excluding the Willow cover type.

area of the scene which was not used for training. The errors in classification of Cattail are believed to be due to the presence of mixed Cattail/Willow in many areas. A sparse scattering of Willow occurred in many areas of both the Cattail and the Reed Canary cover types. Wherever present, Willow tended to cause misclassification due to the highly variable spectral characteristics of this cover type. These difficulties may be resolvable by the inclusion of a "texture" variable in future analysis procedures.

CONCLUSIONS

Based on the results of this investigation, the following general conclusions have been reached:

1. If, within the accuracy levels it is capable of providing, LANDSAT data are judged to be acceptable for wetlands classification, a precision geometric correction should be applied and multi-date data sets should be used. In this study a 25% improvement in average classification accuracy was realized by processing double-date vs. single-date data (from 41% to 66%). Under the spectrally and spatially complex site conditions characterizing the geographical area used in this study, further improvement in wetland classification accuracy is apparently precluded by the spectral and spatial resolution restrictions of the LANDSAT MSS.
2. Full-scene analysis of scanning densitometer data extracted from small scale color infrared photography failed to permit discrimination of many wetland and non-wetland cover types. This was due essentially to the same spectral confusion between cover types that

was realized in the LANDSAT analysis. When classification of photographic data was limited to wetland areas only, much more detailed and accurate classification could be made. The final subset classification made in this study yielded an average classification accuracy of 87% with a 2.5m spatial resolution.

3. The integration of conventional image interpretation (to simply delineate wetland boundaries) and machine-assisted classification (to discriminate among cover types present within the wetland areas) appears to warrant further research. Though both the delineation of wetland boundaries and the development of valid training area statistics are labor-intensive tasks, the results from the combined approach appear to be much more accurate and economical than those obtained in conventional full-scene digital analysis. Additional research is needed to study the feasibility and cost of extending this methodology over a large area using LANDSAT and/or small scale photography.

ACKNOWLEDGEMENTS

The LANDSAT data analysis performed in this study represents partial fulfillment, by Lee F. Werth, of the requirements for the Ph. D. degree at the University of Minnesota. Many aspects of this study were supported directly or indirectly by funding from the University of Minnesota College of Forestry and Agricultural Experiment Station (Proj. MIN-42-003 and 037). The cooperation of the Purdue University Laboratory for Applications of Remote Sensing (LARS) in the LANDSAT analysis is recognized, as is that of the University of Wisconsin Environmental

Monitoring and Data Acquisition Group (EMDAG) in the photographic analysis. Douglas E. Meisner developed the University of Minnesota Image Processing Software used in the photographic analysis and Mark Columbo, Carl Markon, John Minor and Mark A. Springan assisted in the LANDSAT accuracy assessment. William L. Johnson and Katherine A. Knutson assisted in the production of this report. Finally, the University of Minnesota Space Science Center is acknowledged for administering the NASA Grant funding this research.

LITERATURE CITED

Cowardin, L. M., V. Carter, F. C. Golet, and E. T. LaRoe, 1977. Classification of Wetlands and Deep-Water Habitats of the United States (an operational draft). USDI Fish and Wildlife Services. 100pp.

Hoffer, R. M. and M. D. Fleming, 1978. Mapping Vegetation Cover by Computer-Aided Analysis of Satellite Data, LARS Technical Report 011178. 12pp.

Owens, T. and M. Meyer, 1978. A Wetlands Survey of the Twin Cities 7-County Metropolitan Area - East Half. IAFHE RSL Res. Report 78-2, St. Paul, MN. 19pp.

Werth, L. 1980. A Comparison of Remote Sensing Technique Applications to Wetland Classification in Minnesota's 7-County Metropolitan Area, Ph. D. Dissertation, University of Minnesota, (in preparation).

Werth, L., M. Meyer and K. Brooks, 1977. A Wetlands Survey of the Twin Cities 7-County Metropolitan Area - West Half. IAFHE RSL Res. Report 77-10, St. Paul, MN. 18pp.

D₂

N81-12507

SECTION B

DEVELOPMENT OF ALTERNATIVE DATA ANALYSIS TECHNIQUES FOR IMPROVING THE ACCURACY AND SPECIFICITY OF NATURAL RESOURCE INVENTORIES MADE WITH DIGITAL REMOTE SENSING DATA

(Progress Report)

(July 1, 1979 - December 31, 1979)

Dr. T. M. Lillesand and Douglas E. Meisner
Remote Sensing Laboratory
Institute of Agriculture
University of Minnesota
St. Paul, Minnesota

INDEX

3 0