79-10089
NASA-CR-157981

SDSU-RSI-78-14

REMOTE SENSING APPLICATIONS TO

RESOURCE PROBLEMS IN SOUTH DAKOTA

(E79-10089) REMOTE SENSING APPLICATIONS TO RESOURCE PROBLEMS IN SOUTH DAKOTA Annual Progress Report, 1 Jul. 1977 - Jul. 1978 (South Dakota State Univ.) 127 p HC A07/MF A01 CSCL 08F G3/43

N79-15367

Unclas 00089

For

National Aeronautics and Space Administration
Office of University Affairs
Washington, D.C.

July 1, 1977 - July 1, 1978

Annual Progress Report
Grant No. NGL 42-003-007

Remote Sensing Institute
South Dakota State University
Brookings, South Dakota 57007

Abstract

Applications of remote sensing techniques are being utilized in projects of current interest in South Dakota. The cooperative projects involve South Dakota agencies in an ongoing, technologic conveyance of operational remote sensing applications. Each of five projects is abstracted in the following paragraphs.

Habitat Change

Change in the vegetative structure is taking place in the Black Hills. Temporal analysis of the areal extent of open meadows was accomplished using black and white (1961) and color infrared (1976) aerial photography. A reduction of nearly 1100 hectares of open meadows was determined using photo-interpretive methods. However, the extent of individual meadow change was determined to be non-significant using a nonparametric statistic. In addition, temporal analysis of road changes was determined to be nonreliable due to the scale of photography. Powerline corridors were found to increase by 45 miles from 1961 to 1976. The use of color infrared imagery is shown to be useful in wildlife and wildlife habitat assessment in preliminary timber sale planning.

Meandered Lakes

The South Dakota Dept. of Game, Fish and Parks is currently implementing a management program for meandered lakes in South Dakota. The objective of this project was to develop and document application of remote sensing to aid in the management of meandered lakes. Techniques that were developed include: use of Landsat imagery for continuous monitoring, classification of hydrophytes on low altitude CIR imagery, and planning and evaluation of improvements and multiple uses on aerial photography and photo mosaics. Changes in water regime, vegetation structure, agrarian and municipal use and burning were interpreted on Landsat imagery. Hydrophyte classes were manually interpreted on low altitude infrared photography, but required interpretation experience and knowledge of hydrophyte characteristics. An alternate approach which employs the use of photographic separations and automatic density "slicing" was developed to minimize the need for interpretation experience. Using this technique the overall classification accuracy (K) and the overall positional mapping accuracy (M) were 88.0% and 79.8%, respectively. A photo mosaic of low altitude CIR was used to aid in the planning of an extensive level ditching scheme that was constructed in a large meandered lake. High and low altitude aerial photography was collected to evaluate the short and long term effects of different management practices. The techniques that were developed represent an operational use of remote sensing.

Agricultural Censusing

Landsat data has been investigated extensively to determine crop types and acreage. However, confounding site factors often reduce overall accuracy. Soils data in southeast South Dakota were used to

stratify Landsat data in a small contiguous area (intensive site) and across the entire Landsat scene (general area). Landsat data (05 June and 29 July) were analyzed statistically from both study areas to determine the effect of soils stratifications of corn signatures. Band 5 early season and Band 7 late season recorded the strongest evidence of the influence of soils on corn signatures. Significant strata were determined by a multiple range test. More reliable results were attributed to the intensive area because of reasonably uniform plant stage, soil texture, and drainage; slope and parent material differences were evident in the derived strata. The general study area because of variable plant stages across the scene resulted in strata that were less predictable than the intensive site strata.

Aspen Mapping

The South Dakota Department of Game, Fish and Parks is currently involved in the study of the management potentials of aspen (*Populus tremuloides*) within the Black Hills of South Dakota. Aspen has known benefits as a wildlife habitat and potential benefits as a livestock food component.

To assist in the development of a management plan for aspen, an inventory of aspen and other forest types has been produced using computer analysis of Landsat CCT data. Classification procedures are reviewed, accuracy results are discussed and sample output products are provided.

Results indicate an acceptable inventory has been produced. The inventory provides an accurate up-to-date inventory of forest cover types.

Cell Size

Computerized resource analysis systems, which manage, analyze and map spatially referenced data sets on a cellular basis, can only be effectively employed in conjunction with a systematic or scientific procedure for selection of the cell size. This study looks at system performance in terms of inventory accuracy and mapping accuracy as a function of changing cell size. Furthermore, a data set (map) characteristic is sought which also relates to changing cell size. The objective is to directly relate a map characteristic(s) to performance of the information system with varying cell size. The existance of an estimable characteristic provides a route to scientific selection of cell size via sampling of the candidate input data.

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

LIST OF FIGURES

LIST OF FIGURES - Continued

LIST OF TABLES

LIST OF TABLES - Continued

ACKNOWLEDGEMENTS

The National Aeronautics and Space Administration (NASA) and the State of South Dakota provided funding for these applications projects. Gratitude is extended to Mr. Joe Vitale, NASA Technical Officer, for support of these and past projects.

Contribution of Mr. J.T. Streckfuss' Game, Fish, and Parks (GF&P) is appreciated in the meandered lake study. Mr. Nicholas Grkovic (GF&P) contributed much assistance in the habitat change study.

REMOTE SENSING APPLICATIONS TO RESOURCE PROBLEMS IN SOUTH DAKOTA

The major objective of the NASA Project is to provide a means of applying existing remote sensing technology in current programs of state, local, and other South Dakota agencies. In this manner, these agencies are introduced to a means of gathering, compiling, and incorporating remotely sensed data into decision-making and management procedures. The ensuing annual report is composed of five individual projects.

The projects may be briefly described as follows: 1) assessing wildlife habitat change in the Black Hills (1961-1976); 2) classifying meandering lakes in eastern South Dakota; 3) improving classification of agricultural crops in southeast South Dakota; 4) mapping aspen in the Black Hills by classification of Landsat tapes; and 5) determining cell size criteria to ensure pre-determined accuracy limits and/or cost constraints. Each project report is presented as a separate section.

Application of Remote Sensing to Detect Changes of Habitat in the Black Hills (1961 to 1976)

by
K.J. Dalsted

Introduction

Photographs taken during Custer's 1874 expedition through the Black Hills provide a historical background of forest structure before the implementation forest management techniques. A comparison of photographs taken then (1874) and now (1974) shows the large increase in numbers of ponderosa pine (Pinus ponderosa) in the last 100 years in the Black Hills (Progulske, 1974). It is the decrease in the quality and extent of wildlife habitat (i.e. enclosure of open meadows by pine, increased density of pine stands or "doghair", and decrease in aspen acreage) that has concerned the Department of Game, Fish and Parks (GFP) personnel. Measurement of natural resource changes as is made possible by sequential remote sensing products is an essential component of resource management (Brothers and Fish, 1978). Quantitative data are needed to enumerate temporal changes in forest structure of the Black Hills in order to effect integrated programs of wildlife habitat preservation.

Addition of roads and power line corridors also have an effect on the quality of habitat and must be evaluated on a temporal basis. Remote sensing techniques provide a relatively quick and inexpensive means whereby a reconnaissance survey can temporally quantify vegetative changes. The area or areas where change is found to be most extensive, as determined using remote sensing technology, can then be studied more intensively.

Remote sensing techniques offer other advantages to the resource manager. The implementation of good management plans is dependent on an accurate assessment of germane ground parameters. The use of color infrared imagery (CIR) has been long recognized as an effective tool in vegetation analyses (Knipling, 1969). An integrated program of CIR and ground investigation can lead to an accurate evaluation of variables required to establish good management plans (Appendix A). The added perspective gained through the use of remote sensing technology, in addition to its aforementioned attributes, can lead to: 1) an overall improved assessment of habitat (e.g. GF&P environmental assessment report (EAR), a recommendation for wildlife habitat management required before a timber sale in the Black Hills); 2) a reduction in excess, time-consuming field investigations and a determination of field investigation emphasis that is necessary in preplanning.

This report will address two specific goals: 1) completion and comparison of areal and linear measurements of open meadows, roads, and power line corridors from 1961 to 1976; and 2) the role of remote sensing technology in habitat evaluation as input into pre-timber sale planning.

Methods and Materials[1/]

Black and white 1961 panchromatic photography (USDA Forest Service 9/14/61) in the form of photo index sheets at a scale of 1:63,360 was photo-interpreted for open meadow-like areas, roads and power line corridors in the Black Hills. Mylar and translucent drafting paper overlays were used to record the data. Spatial Data Equipment was used to measure the areal extent of open meadows; the average of two measurements gave the final figure. The Wold-Wolfowitz runs test (Winkler and Hays, 1975) was used to determine if the 1961 mean open area measurement was significantly different from mean open area in 1976. Using equations 1, 2, and 3 the critical Z value was calculated.

$$(1)\quad E(R) = \frac{2n_1\, n_2}{n_1 + n_2} + 1$$

$$(2)\quad \sigma^2_R = \frac{2n_1\, n_2\, (2n_1\, n_2 - n_1 - n_2)}{(n_1 + n_2)^2\, (n_1 + n_2 - 1)}$$

$$(3)\quad Z = \frac{R - E(R)}{\sigma_R}$$

[1/] Note - Methods and materials and results and discussion of open areas measurements are more fully described in FY 1978 NASA Semi-Annual Report (Myers et al., 1977).

where: E = expected value

R = runs. After ranking and collating 1961 and 1976 open areas by size, it is the relative degree of dispersion as determined by the number of times the 1961 and 1976 open areas change.

n_1 = observations 1961

n_2 = observations 1976

σ^2_R = variance associated with R

σ_R = standard deviation with R

Z = standard normal variable: mean = 0, $\sigma^2 = 1$

This test did not make any inference to the total open area differences, 1961 vs 1976. It tested only the individual open areas on a paired-like comparison.

Road types were recorded as paved, all weather, dirt, or primitive; a Forest Visitor's Map (U.S. Forest Service, compiled in 1961 and partly revised in 1972) was used to aid in identifying and locating the road types. However, only the roads which were visible on the imagery were recorded. Linear determinations of edge and roads were determined from the average of two measurements using an electronic planimeter (Numonics Model 1224).

RB-57 color infrared imagery was photo-interpreted to provide the 1976 data. Data derived from the 9 September 1976 imagery was extracted and recorded in the same manner as the 1961 data. A slightly smaller scale (approximately 1:100,000) of the CIR was used; the cost factor involved in enlarging the original color transparencies to 1:63,360 was prohibitive. It is felt that the superior quality of the CIR more than offset the scale differential.

However, to ensure that scale difference did not bias the results, the 1961 open area data was systematically checked against the 1976 open area data to make sure that all corresponding open areas were accounted for. Field investigations were undertaken to confirm many of the 1976 delineations, to investigate any questionable areas, and to establish criteria for a photo-interpretive key.

The field investigations also served to familiarize GF&P personnel with the CIR representations of Black Hills' vegetation. Data collected during these field examinations was used to develop a photo-interpretive key for GF&P personnel. The key at a 1:24,000 scale includes identification of the following classes: 1) ponderosa pine (a minor component may be Black Hills spruce) and three divisions of crown or canopy density; 2) aspen (may include other deciduous trees); 3) open meadows and differentiae based on moisture regimes; 4) pine infestation by beetles; 5) geologic formations which are apparent, e.g. taconite deposits.

The overlays and photo-interpretive key will be made available by GF&P personnel for review and accuracy checks. The data supplied may then be used to aid in field evaluations and ultimately improve management plans (see Appendix A).

Results and Discussion

This study endeavored to compare and measure structural habitat changes in the Black Hills from 1961 to 1976 and to establish the usefulness of color infrared imagery for current wildlife habitat evaluations.

Imagery availability limited this project to a fifteen year interval. Although new seedlings will have started in this interval, their growth to maturity will not have taken place. The "signature" of immature pine and that of mature pine is considerably different (see Appendix B, photo-interpretive key). However, the increase in crown density, the encroachment of pine on open meadows, and the shading out of aspen will have taken place by the further growth of existing immature pine (since 1961).

Table 1 shows the results of the open area measurements. The results show a noticeable decrease in meadow-like areas from 1961 to 1976. Further discussion and examples of change are in the Semi-Annual Report to NASA (Myers et al., 1977).

Table 1. Areal and edge measurements in the Black Hills. The 1961 data were derived from black and white photo index sheets (U.S. Forest Service) and the 1976 data were derived from color infrared RB-57 imagery (Myers et al., 1977).

	Open Meadow-like Area (hectares)	Edge of Open Area (kilometers)	Edge/Unit Area (km/ha)
1961	54,083	4,423	.082
1976	52,899	3,846	.073

The results of open area means were statistically compared. Mean meadow sizes from the two years were compared by means of the Wold-Wolfowitz nonparametric runs test. A Z value of 0.53 was calculated. The Z value, under the conditions and assumptions associated with the Z distribution, is an indicator used to test the hypothesis that the 1961 open area mean and 1976 open area mean are not significantly different (H_o: $M_{1961} = M_{1976}$ vs H_a: M_{1961} M_{1976}). In this case the null hypothesis (H_o) was accepted because the critical Z value (.05 level of significance), which was 1.96, was not exceeded. It must be emphasized that this test in no way made any allusions to the total differences (Table 1); those differences are a total not a mean value and as such are the population not an estimator of the population (an undetermined error value is associated with the results in Table 1).

Although on an individual basis open area changes are not significant, a consistent decrease in open area adds up. Severson and Boldt (1977) point out that from understory and overstory production viewpoint the maximum production was attained under a square feet of basal area per acre or GSL (growing stock level) of 60 to 100 (medium density). It follows that forest thinning and other applicable programs (e.g. prescribed burning) which control forest encroachment can lead to optimal lumber production and quality of habitat.

Results of roads and power line corridor measurement are shown in Table 2. A comparison of all the categories indicates similarity

Table 2. Roads and power line corridor measurements, 1961 and 1976.

	Road Types (Miles)				
	Paved	All Weather	Dirt	Primitive	Power line
1961	295	250	380	865	215
1976	300	240	375	740	260

in all but the primitive road and power line corridor categories. Firstly, color infrared imagery is probably not as well suited to road delineation, especially nonpaved roads, as conventional photography because of its response to vegetative reflection of infrared radiation. This fact may have biased results somewhat giving an indication of little road addition or even some decrease between 1961 and 1976. A closing of pine canopy since 1961 near some of the more primitive roads may have biased mileage determination in 1976. Overall, the results are not conclusive evidence that road additions since 1961 are inconsequential. A more effective approach to road determination would require lower level imagery. Power line corridors (see photo-interpretation key) are quite apparent on the imagery. The increase in mileage of power line corridors is likely to be a more accurate indication of change than roads measurements.

Remote sensing products have been used with considerable success by GF&P personnel in habitat assessment programs (Appendix A). As a result of this program a photo-interpretation key (Appendix B was prepared for use by Rapid City Regional GF&P personnel. The actual use of CIR was to aid in providing wildlife habitat recommendations for environmental assessment reports, preliminary to timber harvest activities, in order to make allowances for habitat change.

References

Brothers, G.L., and E.B. Fish. 1978. Image enhancement for vegetative pattern change analysis. Photogram. Eng. and Remote Sensing. Vol. 44 (5):607-616.

Knipling, E.G. 1969. Leaf reflectance and image formation on color infrared film. p. 17-29. In P.L. Johnson (ed.) Remote sensing in ecology. Univ. of Georgia Press, Athens.

Leedy, D.L. 1960. Photo interpretation in wildlife management. In R.N. Colwell (ed.) Manual of Photo Interpretation. Amer. Soc. of Photogram., 1515 Massachusetts Ave., N.W., Washington, D.C. pp. 521-530.

Myers, V.I., R.G. Best, K.J. Dalsted, J.C. Eidenshink, F.A. Schmer, and M.E. Wehde. 1977. Application of remote sensing technology in South Dakota to assess wildlife habitat change, describe meandering lakes, improve agricultural censusing, map aspen, and quantify cell selection criteria for spatial data. Semi-Annual Progress Report for NASA. SDSU-RSI-77-17. Remote Sens. Inst., SDSU, Brookings.

Progulske, D.R. 1974. Yellow ore, yellow hair, yellow pine. Ag. Exper. Sta. South Dak. State Univ., Brookings.

Severson, K.E., and C.E. Boldt. 1977. Options for Black Hills Forest owners: timber, forage, or both. Range. J. 4(1):13-15.

Smith, D.R. 1977. A report on the wildlife and habitat parameter workshop. Rocky Mt. Forest and Range Exper. Sta., Colorado State Univ., Fort Collins.

Winkler, R.L., and W.L. Hays. 1975. Statistics: probability, inference, and decision. Holt, Rhinehart, Winston, New York, N.Y.

Appendix A

Example of the use of remote sensing data in wildlife and wildlife habitat planning before a timber sale in the Black Hills.

SOUTH DAKOTA
Department of Game, Fish and Parks
Pierre, South Dakota 57501 Phone 224-3387

Division of Administration

3305 West South St.
Rapid City, SD 57701

March 22, 1978

Denny Miller, Ranger
U.S. Forest Service
Pactola Ranger District
Rapid City, SD 57701

Dear Mr. Miller:

On several occasions, during the month of February and in company with Dennis Lowry, I made a reconnaisance of the Commissary, Tomaha and Rapid Creek timber sale area. In addition, Rapid Creek was reconned from the Pactola Lake spillway through McGee Siding. The purpose was to make a meaningful wildlife and wildlife habitat input into the Environmental Assessment report you are preparing for the three sales.

The following is objectively submitted considering time spent in the field, research information and historical data available. Two enclosures are submitted. One is a habitat recon form; the second is a tracing made from infra red photos (scale 1:24000). Both containing recommendations.

The value of remote sensing imagery is contained in better area perspective, vegetation differentiation and identification plus canopy closure differentiation, open area and mineral outcrop differentiation. All provide for better management understanding and treatment. The imagery can be made available in stereo pairs.

Seven areas as provided by format concerning USFS projects are addressed as follows:

1. Turkey, whitetailed deer, mule deer and ruffed grouse are the major species to be considered with raptors, furbearers, small game and passerine species considered as secondary species in that order.

   Rapid Creek is of blue ribbon trout stream classification, and should be given individual attention. DEP classification is cold water permanent. Other streams in the area are classified marginal. Extensive natural reproduction in Rapid Creek in the timber sale area has been documented through sampling by electro-fishing with a wild brown trout population (over 20cm in length) calculated as 1500 per

SOUTH DAKOTA

mile (R. C. Ford personal communication). This is considered a low figure. With strategic road closures, Rapid Creek in the are of the sales, could be managed as a blue ribbon walk-in fishery. Road closure location recommendations are: 1) the end of the road at Placerville and, 2) the crossing at McGee Siding. Since McGee Siding is in private ownership, it is recommended this parcel be considered in the land exchange program. This area is also extensively grazed by cattle.

2. On site population objective for whitetailed deer is to provide for a proportionate deer increase in the overall Black Hills deer population increase of 6,000 animals as stated in the Comprehensive Wildlife Management Plan for National Forest System Lands in South Dakota. This area is a part of Management Unit III (of 4) into which this agency has the Black Hills segmented for the on-going Pittman-Robertson Project W-95-R-11 (Deer Harvest Regulation Study). This study provides deer population trends through pellet count data analysis assisted by browse utilization transect data. Also, this area provides for a portion of Unit III winter range which winters a sizeable number of whitetailed deer. Pellet count data analysis established the parameter of 25-50 deer per section for the Black Hills as a whole and with a recent population trend downward. Browse utilization percentages (leader use) for transects in the timber sale areas show an irregular pattern with cattle-deer competition. Locations of pellet group and browse transects are available on request. Therefore, management practices which will increase winter range in the timber sale area are warranted and desireable from a wildlife resource-multiple use standpoint.

   This area supports an undetermined portion of the Black Hills turkey population. Turkey habitat is good and many man-days of turkey hunting are provided with reasonable success. The area can support a proportionate increase in the objective of 1,000 bird increase as stated in the Comprehensive Wildlife Management Plan. In my opinion, the objective of a 1,000 bird increase for the Black Hills is a low figure. Cattle-turkey competition (oak and hawthorn regeneration) has been noted, especially with hawthorn shoots browsed by cattle to the extent that, in my estimation, existing hawthorn patches are decadent. Continuing the trend will eventually cause the species to extirpate. Fencing with rest-rotation grazing can alleviate the condition.

   Some habitat management recommendations for other species are contained in the literature which is reflected in the enclosures (snags, unevenage aspen management, water production, road closures, riparian habitat management, etc.).

3. Priority ranking is as follows: fisheries management, turkey and whitetailed deer habitat management, raptor habitat considerations along Rapid Creek and especially in Dark Canyon. Raptor habitat is classified excellent in Dark Canyon (J. Sharps, personal communication). Ruffed grouse habitat considerations are dependent on unevenage aspen management and aspen clone preservation along drainage bottoms.

Furbearers, dependent on riparian habitat along Rapid Creek, will benefit by reduced competition with cattle along the creek. Cavity nesters and other perching bird habitat requirements are contained in the literature.

4. Detailed habitat requirements for the species listed above are well documented in the literature and need not be reiterated. It is again meaningful to emphasize the need for deer winter range acreage increase, the need for reduced deer-turkey-cattle competition on certain forage species and the fencing requirement.

5. Suggested techniques for habitat improvement for species listed above include: prescribed burning for browse and grass stimulation and rapid nutrient turnover, fencing for rest rotation and deferred grazing programs, road closures, leaving snags, nest and roost trees, slash for turkey nesting sites near water, and preservation of natural openings by reducing pine encroachment.

6. A travel management plan for this area might include through access from east to west by means of USFS roads 159 and 164 which could be classified as a system road. For wildlife considerations, other roads needed for wood fiber removal would be temporary and should be closed upon completion of logging operations and also should be limited in number. Feeder roads should be obliterated and reseeded. Seclusion, solitude and escape cover for wildlife species may be limiting habitat factors in some instances.

7. Specific wildlife programs for this area were not identified during the briefing on 6 December 1977. Recommendations enumerated above can be included in future budgetary planning considerations.

If I can clarify questions concerning any recommendations submitted, do not hesitate to call on me.

Sincerely,

Nicholas Grkovic
Resource Management Specialist

NG/i
cc: Black Hills Nat'l Forest(Staff Wildl. Specialist)
Dennis Lowry
Lloyd Thompson
Director, Game Management Division

Appendix B

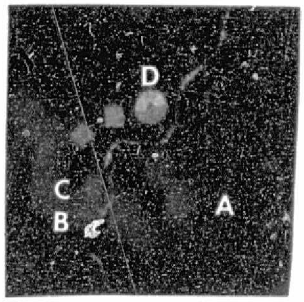
Photo-interpretative key for use in the
Black Hills by G,F,&P personnel.

Prepared by RSI at a scale of 1:24,000
from 1976 RB-57 high altitude color infrared imagery.

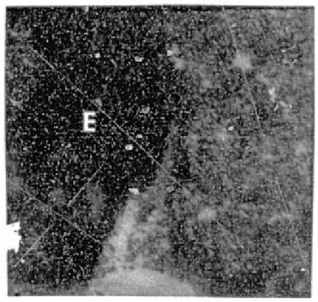
FOREST

PONDEROSA PINE (MAY INCLUDE SOME SPRUCE)

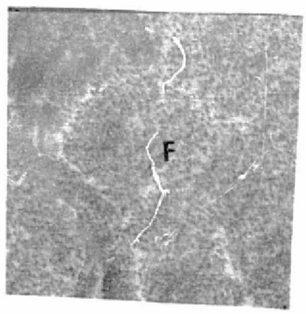
EXPERIMENTAL FOREST: THREE LEVELS OF CROWN DENSITY A) HIGH, B) MEDIUM, C) LOW
D) BARE SOIL



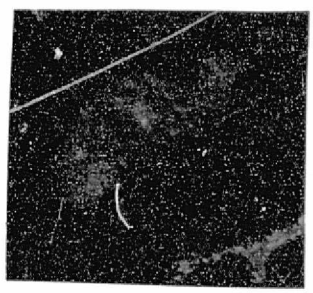
HIGH CROWN DENSITY (E)



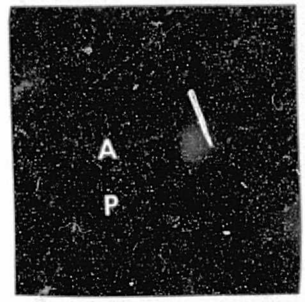
MEDIUM CROWN DENSITY (F)



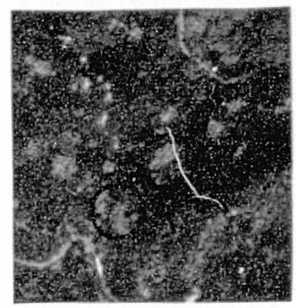
LOW CROWN DENSITY (G)



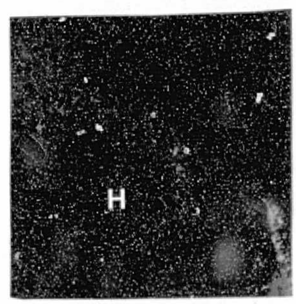
PINE - ASPEN



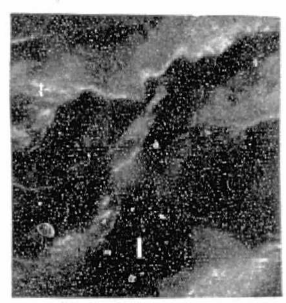
BEETLE INFECTED PINE



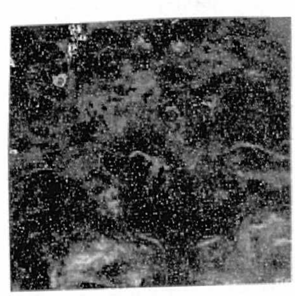
ASPEN (H)



YOUNG PINE (15 TO 25 YEARS)



YOUNG ASPEN

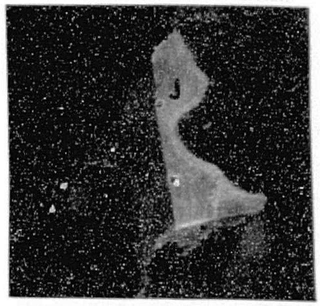


CHOKECHERRY AND OTHER BROWSE

NON-FOREST

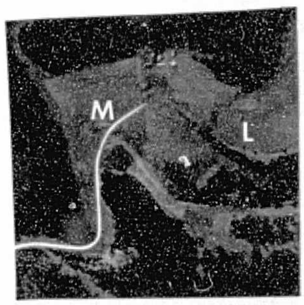
CROPLAND

BARE (J)
TAME GRASS (K)

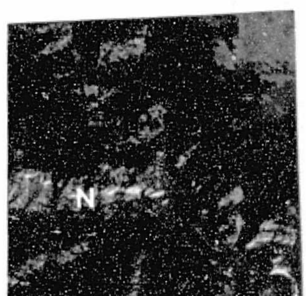


MEADOW (L)

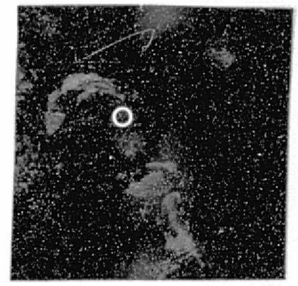
RIPARIAN VEGETATION (M)



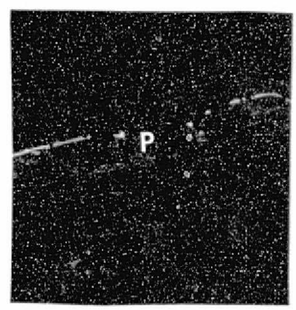
ROCK OUTCROP (N)



TACONITE DEPOSIT (O)



WATER (P)

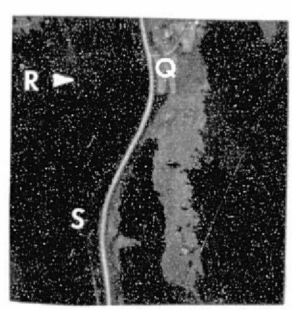


GENERAL

ROAD (Q)

POWERLINE CORRIDOR (R)

RAILROAD (S)



APPLICATIONS OF REMOTE SENSING IN THE MANAGEMENT OF MEANDERED LAKES IN SOUTH DAKOTA

By R.G. Best and J.T. Streckfuss[1]

Introduction

In 1957 the South Dakota Game, Fish and Parks Department (GFP) was given authority and responsibility by the State Legislature to administer meandered lake or trust lands (SDCL 41-2-18). Meandered lake lands are those areas that were omitted from the original government survey. The Federal Government issued no titles to these lakes and lake beds but dedicated them to the People of South Dakota. These areas range in size from 20 to over 7,850 acres and they vary from closed stands of hydrophytes to open water prairie lakes. These 250 meandered lakes represent a primary source of outdoor recreation for the people of South Dakota such as hunting, fishing, boating, etc.

Until recently, there has been only limited management of the meandered lakes. Game, Fish, and Parks personnel are now implementing a management program to utilize the optimum recreational potential for these areas. An inventory of these areas based on original survey plots has been completed (Dvorak, 1973) but descriptions in the original survey of the meander line are vague and often incomplete. Maps and inventories of vegetation, use of adjacent lands, etc. needed to formulate management plans are not available and would require

[1]Associate Wildlife Specialist, Remote Sensing Institute, South Dakota State University, and Assistant Regional Supervisor, South Dakota Dept. of Game, Fish and Parks, Webster, SD.

considerable manpower to complete in a ground level survey program. Remotely-sensed data can be used effectively in conjunction with existing information in the planning of improvements for the maximum use of meandered lakes.

Game, Fish and Parks personnel will develop "cooperative management agreements" to prevent permanent damage or further encroachment of adjacent agricultural land use within the meandered areas. The overall objective of the cooperative management agreements is to enhance the watershed and associated wildlife while increasing their recreational use by the public and allow legal agrarian uses of these trust lands by the farming community. These agreements may include: level ditching schemes, opening of closed stands of emergent vegetation by controlled grazing and prescribed haying and burning, planting of agricultural crops during dry periods and providing public access to the areas. Parameters required to formulate these management plans include: normal high water line, species of hydrophytes present, area of open water, and land use within and adjacent to the high water line.

Data on the distribution and type of aquatic macrophytes is required for management planning in meandered lakes. Present plans call for management of the area bounded by the normal high water line, which in most cases, differs from the meander line because of natural lake succession and changes in peripheral land use. The normal high water line is defined as the limit of hydrophytes. Some species of hydrophytes can also be utilized as feed for livestock (Boyd, 1974). Interpretations of the distribution of species suitable for livestock feeding are needed to plan prescribed grazing and haying within meandered lakes. Furthermore, an accurate classification of hydrophytes can improve the accuracy of wetland classifications from aerial photography.

The objective of this project was to develop remote sensing applications in the management of meandered lakes. The following techniques were developed to achieve this objective: use of Landsat imagery for continuous monitoring, classification of hydrophytes on low altitude CIR imagery, and planning and evaluation of improvements and multiple uses on aerial photography and photo mosaics. Remotely-sensed data provide a perspective of the meandered lakes that is not available in ground level observations. Interpretations of satellite and aircraft imagery provide the data required for management planning in a timely fashion and without the large manpower requirements of ground based data collection. Results of this project indicate an operational use of remote sensing.

A documentation of the techniques that were developed in this project was provided to Game, Fish and Parks personnel. Aerial photography, photo mosaics, and Landsat imagery were provided for continuing applications.

METHODS

Kingsbury County in east-central South Dakota was selected by the Land Management Specialist from GFP as the test site for the analysis of remotely-sensed data. The area has 17 meandered lakes with a wide variety of characteristics which provide a major source of outdoor recreation. Numerous cloud-free, high quality Landsat scenes are available for the county. In addition, high-altitude color and color-infrared RB-57 imagery are available for portions of the county. Low-altitude color infrared imagery of all the meandered lakes was collected with the RSI aircraft on 8/10/77 and 8/17/77. Color-infrared and black and white (wratten 25A filter) low altitude photography was also collected on 6/5/78 for meandered lakes in Kingsbury County and selected lakes in six northeastern counties. High altitude color and color infrared imagery of 4 northeastern counties where numerous meandered lakes are concentrated was collected by NASA RB-57 on 9/14/77. Color and color infrared 35 mm obliques of the meandered lakes within the four counties were exposed from low altitude aircraft on 9/19/77 to supplement ground truth data and aid in the interpretation of RB-57 imagery.

Landsat Monitoring of Meandered Lakes

The advantages of temporal analyses were illustrated by selecting cloud-free Landsat scenes at approximately one month intervals from May 1976 to November 1976. Scenes from October 1974-1976 were chosen to illustrate yearly Landsat monitoring of meander lakes. A scene from January 1974 was also selected to illustrate the advantages of snow cover for vegetation analysis. Black and white enlargement prints of MSS 5 and MSS 7 were used for interpretation. A county road map was superimposed onto each during printing to aid in location of areas of interest. Changes in water regime, vegetation, encroachment of farming, municipal use and in one case uncontrolled burning from date to date were recorded.

Classification of Hydrophytes

An area within a large meandered lake (Lake Preston) in Kingsbury Co. was selected to develop and test interpretation techniques. Low altitude (10,000 ft. AGL) color infrared (KODAK 2443) imagery was collected on 8/17/77. The area was selected because relatively large pure dense stands of hydrophytes common to most meandered lakes were present. Nine classes representing 6 different stands of aquatic macrophytes, 1 of upland grasses, 1 open water and a non-typical spoil pile class were interpreted and delineated on enlargement prints. The interpretations were verified and adjusted by ground verification.

A second technique which employs photographic separation of the original color film and a color encoded density analysis was developed and evaluated. Black and white separations using red (wratten 25), green (wratten 74) and blue (wratten 98) filters were made in an attempt to partition the effects of infrared, red, and green reflectance on the original film. Separations are made on panchromatic film which is equally sensitive to all 3 wavelengths and processed to $\delta = 1$ to maintain the relative reflectance effects of the original film. Pin registered combinations of positive and negative separations were used to produce a ratioing effect as described by Lockwood (1975).

Each image and ratio sandwich was "sliced" in 4 equal density increments and color encoded on a CRT monitor of Spatial Data model 703. Enlargement prints were made from 35 mm slides taken of the monitor. The classification of hydrophytes was made by identifying the relative density of each class in each separation. Interpretations were rectified to remove monitor distortion by registering to enlargement prints on a zoom transfer scope.

Density sliced interpretations and the ground verified photo interpretation of vegetation classifications were encoded into a computerized spatially-oriented data base in order to illustrate the capabilities of the computerized storage and retrieval system and to calculate the accuracy of the density sliced interpretation. A grid of known cell size (1 cell = .017 hectares) was superimposed on the interpretations (see figure 1). The process of digitization of the interpretation data involves coding on a row-by-row basis those cells where the dominance of one class changes to another. Output products include computer plotted maps (see Figure 2) at any scale, areal tabulations and statistical comparisons.

Confusion matrices, statistical accuracy and mapping accuracy were calculated using methodology developed by Kalensky and Scherk (1975). Each cell in the data was treated as an individual pixel.

Planning and Evaluation of Improvements

A photo mosaic of low altitude color infrared imagery collected 8/17/78 was used to plan an extensive level ditching scheme in Lake Thompson, Kingsbury County. The criteria for selection of the site were: location within but adjacent to the meandered line, availability of public access and water levels suitable for construction.

On 9/14/77 high altitude color-infrared imagery was collected by NASA RB-57 for a 4 county area (Marshall, Day, Grant and Roberts) in which a majority of the state's meandered lake occur. Low altitude, color infrared and Black and White (wratten 25a filter) imagery was collected on 6/5/78 for most of the meandered lakes in the 4 county area.

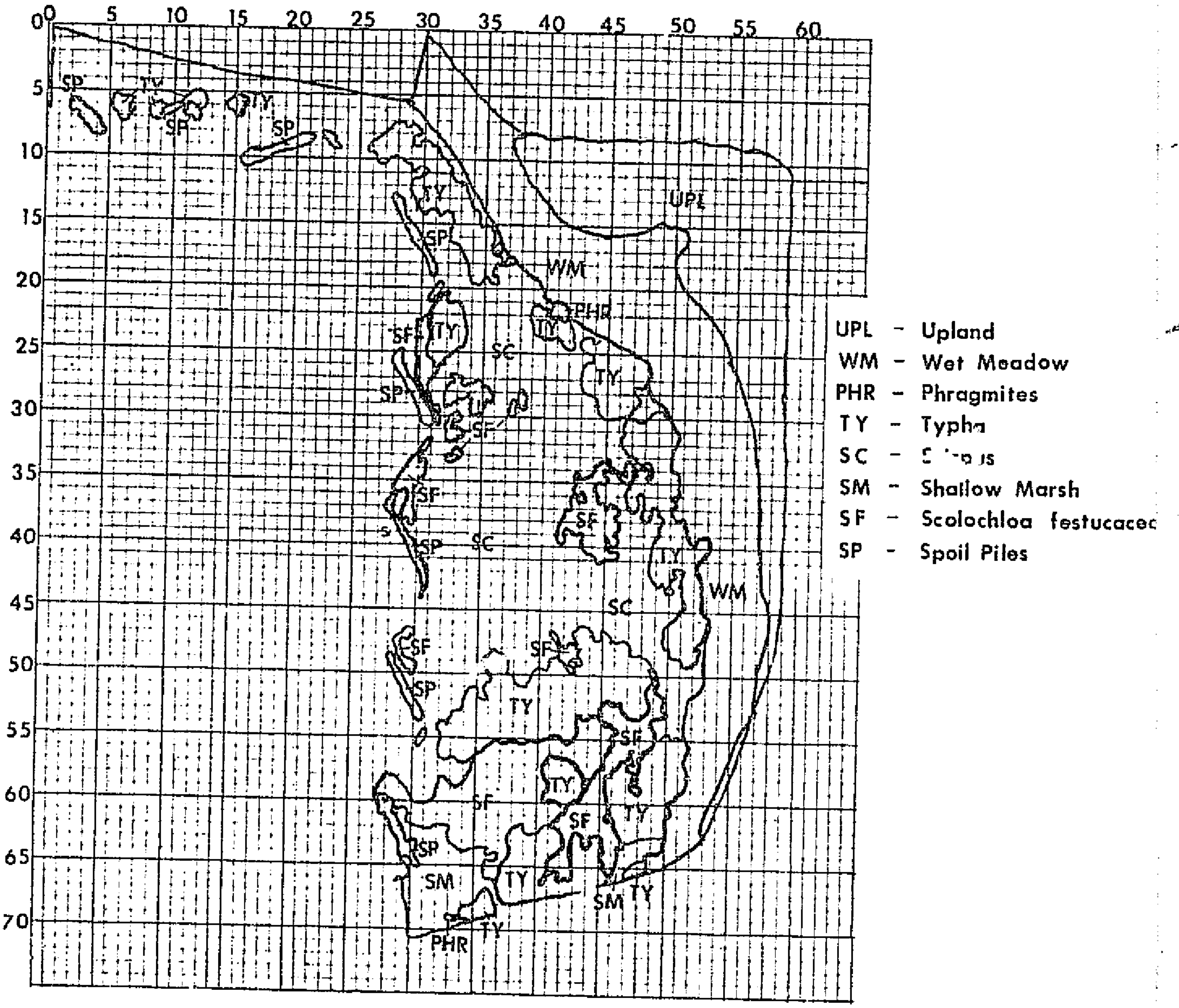


Figure 1. Illustration of cellular matrix used for encoding hydrophyte classification interpretation. 1 cell = .017 hectares.

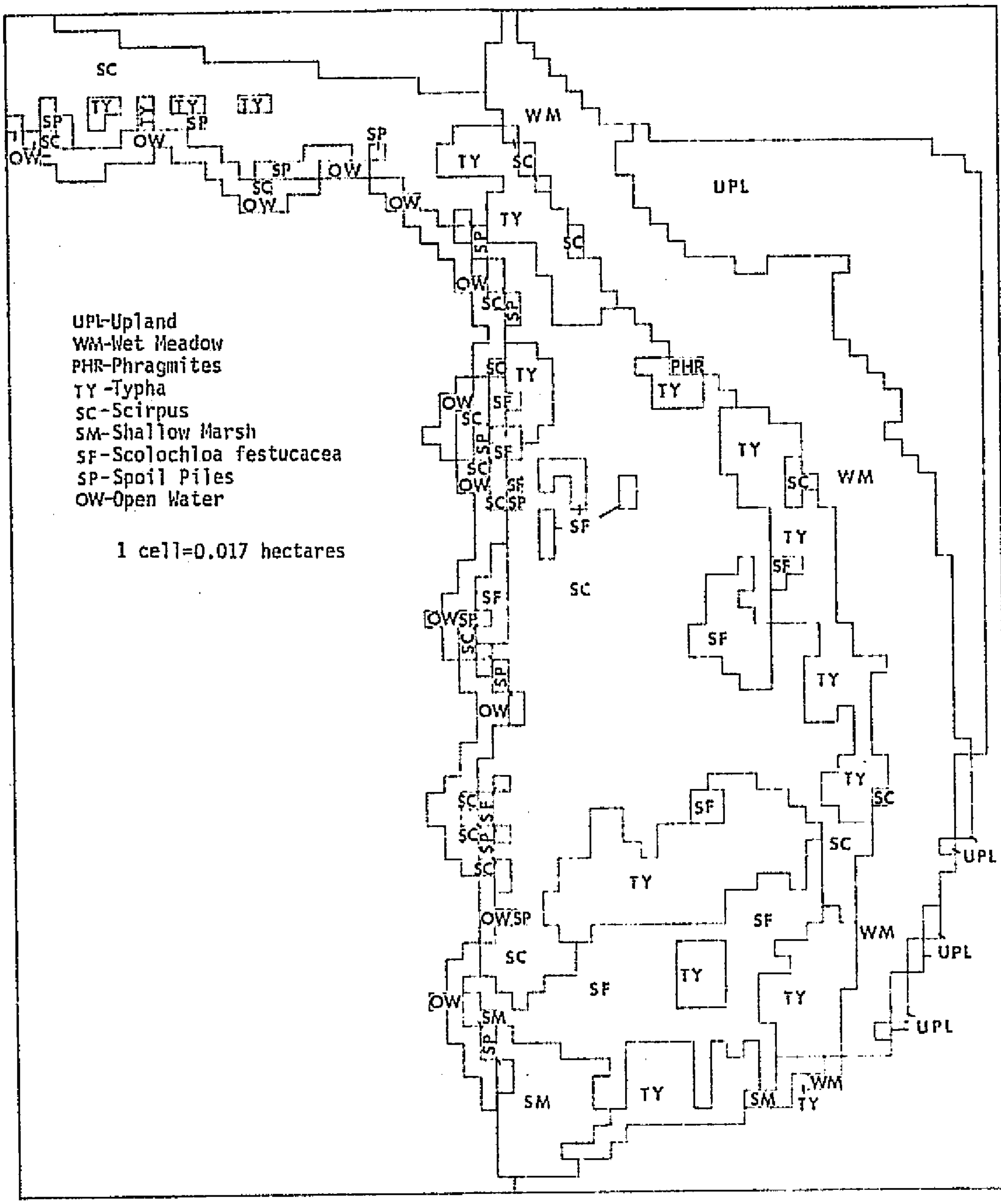


Figure 2. Example computerized plot of hydrophyte classification from spatially oriented data base.

Enlargement prints of these data were provided to Game, Fish and Parks personnel to evaluate the short and long term effects of management practices which were implemented during and previous to 1977.

Results and Discussion

Numerous characteristics of meandered lakes can be interpreted on Landsat imagery. Repetitive coverage allows monthly or yearly monitoring when cloud-free scenes are available. A display with monthly scenes from May to November, 1976, scenes for October 1974-1976 and an example snow-cover scene was presented to the South Dakota GFP Commission at their September 1977 meeting. Examples of these were included with the semi-annual report (Best et al., 1978). Changes in water regime are the most obvious characteristic on the imagery and perhaps the most important. As the water level drops numerous changes, including vegetation structure, water quality and encroachment of agriculture within the meander line, occur in the lakes. During 1976 many lakes dried up because of drought conditions and portions were plowed and planted to small grains which could be interpreted on the Landsat imagery. The amount of agricultural use within meander line must be monitored to prevent excessive damage to wetland habitat. An uncontrolled burn, which destroyed a large area of wetland vegetation which would otherwise provide valuable winter cover, was easily identifiable by the light tones on MSS 7. Landsat imagery of June, 1977 showed that the burn area was completely revegetated and snow melt and rainfall have partially filled the lake which indicated that the burn had little or no permanent effect on the lake. Resolution and the length of time required to receive Landsat imagery are factors which limit the use of Landsat in an operational program. The low cost of temporal Landsat imagery may outweigh both limitations.

Vegetation classes can be manually interpreted by tonal and textural differences on enlargement prints of low altitude color infrared aerial photography. Two classes, Typha sp. and Scirpus sp. are genus' in which species could not be separated. Only two species of hydrophytes, Scholochloa festucacea and Phragmites communis could be consistently interpreted when they occurred in pure dense stands. A shallow marsh class consisting of a mixed stand Scirpus sp., Eleocharis sp., Sparganum sp., Polygonum sp. and grasses could be interpreted by unique textural difference. The wet meadow zone, a mixed stand of grasses and sedges which delimits the extent of the basin, could be interpreted by uniform tones and its peripheral position. Considerable interpretation experience and knowledge of aquatic plant communities is required to make an accurate manual interpretation of the above classes.

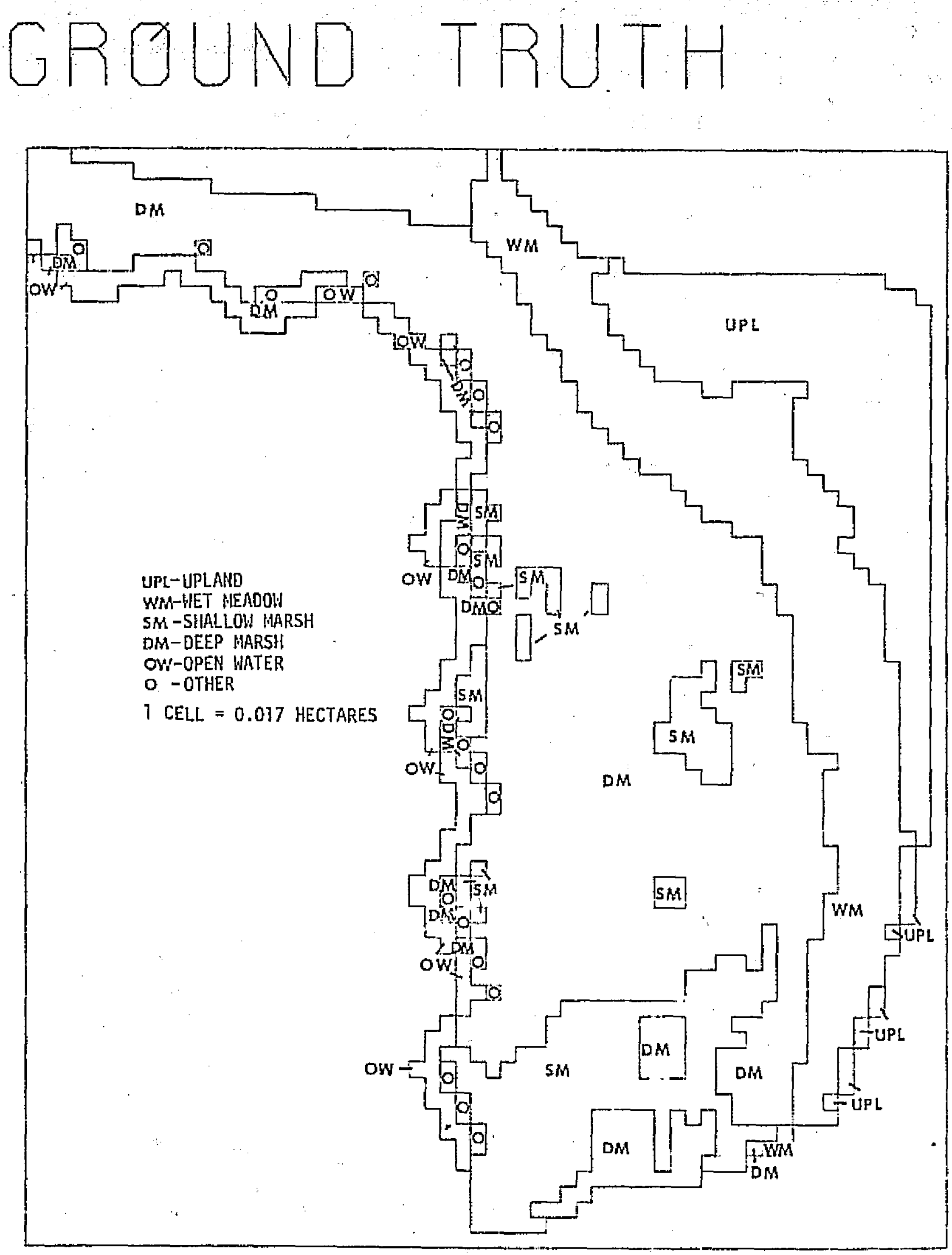
The photographic separation and automated density slicing technique was developed to eliminate the need for experienced photo interpreters. The best separation of hydrophyte classes occurred in the infrared separation. Ratios of separations provided no additional information. This may be attributed to the fact that the near infrared sensitive film layer is approximately 1 1/2 stops less sensitive with respect to day light illumination than the other two layers to visible red and green. Table 3 is a confusion matrix comparing the density sliced interpretation to the photo interpretation and ground verification which is considered as accurate. The table shows an overall classification accuracy of 77.9% and a positional mapping accuracy of 67.3%. The overall mapping accuracy (M) is weighted to reflect differences in the sample size of each category. The mapping accuracy (M_I) of each class is presented in the far right column of the table. The sample size for Phragmites is too small to be compared to the relative accuracy of the other classes. The low accuracy for delineating spoil piles can be expected because the vegetation present is not representative of a wetland ecosystem.

Figures 3a and b are examples of computer plots of the classification data which illustrate the composite capabilities of the data base. They are also presented to show the sources of error presented in Table 4. Table 4 is a confusion matrix in which classes of hydrophytes with similar characteristics have been combined. Upland, wet meadow and open water classes remain the same. The Scolochloa festucacea was combined with the shallow marsh class. The Phragmites, Typha and Scirpus classes were grouped into the deep marsh class and the spoil piles are termed other. The overall classification accuracy and positional mapping accuracy have increased to 88.0% and 79.8%, respectively. This indicated that some of the errors were misclassification of hydrophytes with similar characteristics. The errors in misclassification are evident in Figure 3a and b. The primary source of positional error can be attributed to the non-linearity of the CRT monitor and errors in rectification.

Table 3. Confusion matrix and mapping accuracies of photographic separation and automatic density slicing interpretation of hydrophyte classes.

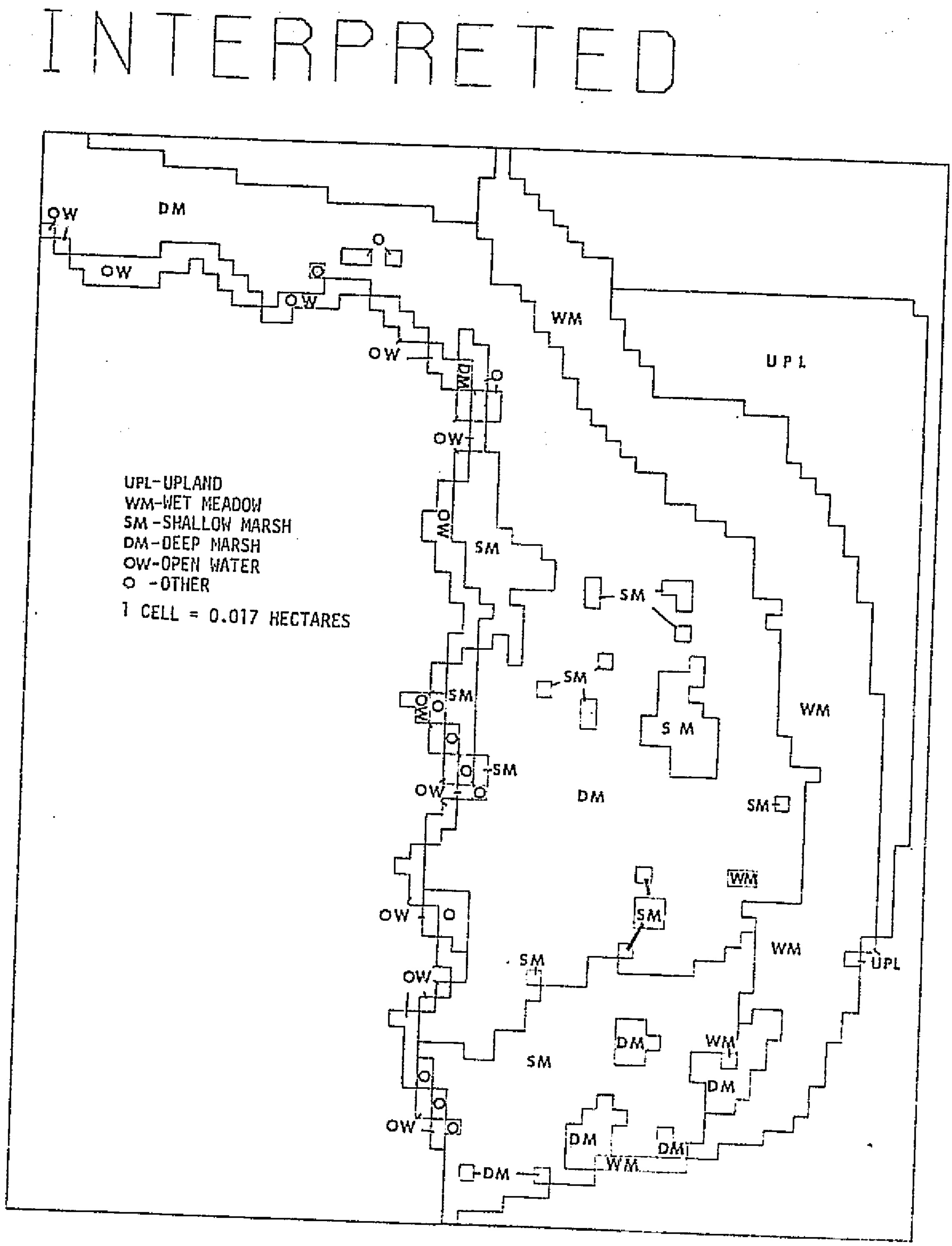
CLASS	UPL	WM	PHR	TY	SC	SM	SF	SP	OW	TOTAL	Omissions No.	Omissions %	M_I%
Upland (UPL)	251	18	0	0	0	0	0	0	0	269	18	6.7	88.7
Wet Meadow (WM)	14	353	0	25	20	6	0	0	0	418	65	15.5	77.1
Phragmites (PHR)	0	0	2	0	0	0	0	0	0	2	0	0	100
Typha Sp. (TY)	0	1	0	212	86	2	9	1	1	312	100	32.0	54.5
Scirpus Sp. (SC)	0	1	0	51	508	0	9	17	2	588	80	13.6	68.2
Shallow Marsh (SM)	0	0	0	0	0	37	5	2	0	45	8	17.8	62.7
Scolochloa festucacea (SF)	0	0	0	49	31	6	160	7	1	254	94	37.0	57.1
Spoil Piles (SP)	0	0	0	1	11	0	1	15	3	31	16	51.6	25.4
Open Water (OW)	0	0	0	0	9	0	1	1	90	101	11	10.9	89.8
TOTAL	265	373	2	338	665	51	186	43	97	2020			
Commissions No.	14	20	0	77	157	14	26	28	7				
Commissions %	5.3	5.4	0	22.8	23.6	37.8	14.0	65.1	7.2				

Overall Classification Accuracy K	77.9%
Overall Mapping Accuracy M	67.3%



3a. Ground truth

Figure 3. Example computerized plots of hydrophyte classification data illustrating the composite capabilities of the data base. Note the sources of error in the density sliced interpretation.



3b. Density sliced interpretation

Figure 3. Continued.

Table 4. Confusion matrix and mapping accuracies of composited density sliced interpretation of hydrophyte classes.

CLASS	UPL	WM	SM	DM	OW	Other	TOTAL	Omissions No.	Omissions %	M_I%
Upland (UPL)	251	18	0	0	0	0	269	18	6.7	88.7
Wet Meadow (WM)	14	353	6	45	0	0	418	65	15.5	77.1
Shallow Marsh (SM)	0	0	209	80	1	9	299	90	30.1	63.9
Deep Marsh (DM)	0	2	20	859	3	18	902	43	4.8	82.0
Open Water (OW)	0	0	1	9	90	1	101	11	10.9	89.8
Other	0	0	1	12	3	15	31	16	51.6	25.4
TOTAL	265	373	237	1005	97	43	2020			
Commissions No.	14	20	28	146	7	28				
Commissions %	5.3	5.4	13.4	14.5	7.2	65.1				

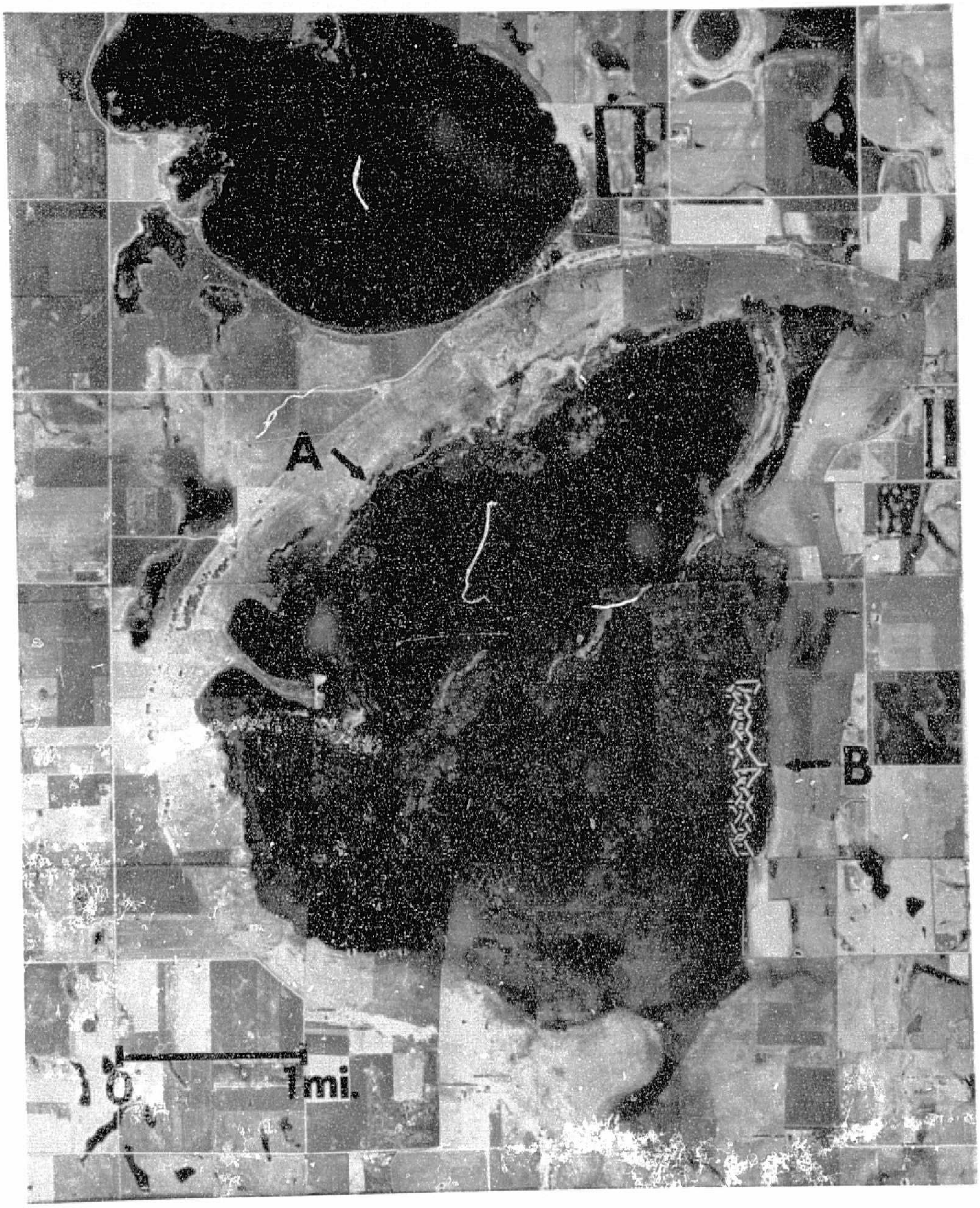
Overall Classification Accuracy K	88.0%
Overall Mapping Accuracy M	79.8%

A level ditching scheme for Lake Thompson in Kingsbury Co. was planned with the aid of a photo mosaic of low altitude color infrared aerial photographs collected 9/17/77 (see Figure 4). The level ditches are excavated in a "zig-zag" pattern to an average depth of 5 feet with side slopes of 3 to 1 or greater with a drag line (see Figures 4b and 5). The berm is a minimum of 3 feet wide with the spoil deposited on alternate sides of the ditch as construction moves from one leg to the next. The completed level ditches enhance the marsh for waterfowl and fur bearer production and big game use. In addition to the benefits for game, waterfowl hunters are able to utilize the ditches as access routes in pursuit of their recreation. The project was constructed by the Game, Fish and Parks in the fall of 1977 at a cost of $60,000 with funding provided by federal drought disaster funds. The level ditching scheme is located adjacent to the meander-line with a public access (see Figure 4b at "B"). The level ditching scheme at "A" in Figure 4a and b was completed several years earlier and is located well within the meanderline and is not readily accessible to the public.

Remotely sensed imagery can also be used effectively to monitor the short and long term effects of different management practices on the meandered lakes. Enlargements of low altitude color-infrared and black and white photography collected on June 5, 1978 and NASA RB-57 imagery collected September 14, 1977 were provided to Game, Fish and Parks to determine the short term effects of

4a. 8/17/77

Figure 4. Photomosaics of Lake Thompson in Kingsbury County used for planning of level ditching project "B". Level ditching scheme at "A" was completed several years prior to data collection.



4b. 6/5/78

Figure 4. Continued.

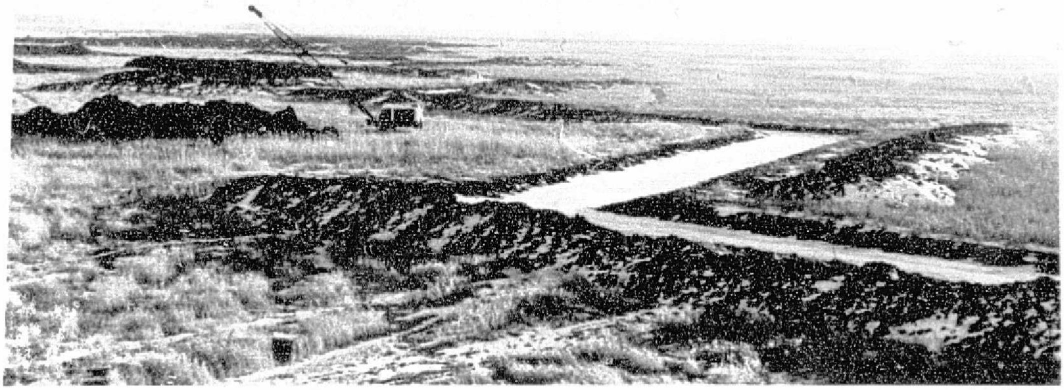


Figure 5. Photograph illustrating level ditch construction.

management plans implemented during 1977. These data will also provide a historic record for comparisons with future data in order to determine the long term effects.

Conclusions

A potential operational application of remote sensing techniques for use in the management of meandered lakes has been demonstrated (see attached letter). Several types of imagery including Landsat and high and low altitude aerial photography were used effectively to reduce the man power requirements of conventional ground survey techniques.

Landsat enlargements can be used to monitor changes in; water regime, vegetation structure, extent of burning and agrarian use within the meander line. Aerial photographs and photo mosaics can be used for the planning of improvements and for the evaluation of different management practices. Hydrophytes can be classified on CIR imagery using either manual interpretation or photographic separation and automated density slicing techniques. The need for interpretation experience and knowledge of hydrophytes that is required for manual interpretation is eliminated in the automatic density slicing technique which provides sufficient accuracy for use. However, the photo-interpreter has the advantage of making interpretations based on texture, patterns, and position in addition to tone.

LITERATURE CITED

Best, R.G., K.J. Dalsted, J.C. Eidenshink, M.E. Wehde. "Application of Remote Sensing Technology in S.D. to Assess Wildlife Habitat Change, Describe Meandering Lakes, Improve Agricultural Censusing, Map Aspen, and Quantify Cell Selection Criteria for Spatial Data." Semi-Annual Report, July 1, 1977 - December 31, 1977. (Submitted to NASA under contract no. NGL 42-003-007).

Boyd, C.E. 1974. Utilization of Aquatic Plants in Aquatic Vegetation and Its Use and Control. D.S. Mitchell ed. ISBN 92-3-101082-4 UNESCO.

Dvorak, D. 1973. Listing of Meandered Lakes. State of South Dakota Department of Game, Fish and Parks, Pierre, South Dakota. 9 pp.

Kalensky, Z. and L.R. Scherk. 1975. Accuracy of Forest Mapping from Landsat Computer Compatible Tapes in Proceedings of the 10th International Symposium on Remote Sensing of the Environment. ERIM Ann Arbor, Michigan. pp. 1159-1167.

Lockwood, H.E. 1975. Photographic Image Enhancement and Processing. The Fifth Biennial Workshop on Color Aerial Photography in Plant Sciences and Related Fields. American Society of Photogrammetry, Falls Church, Virginia. pp. 29-44.

SOUTH Dakota

Department of Game, Fish and Parks

Division of Game and Fish

Sigurd Anderson Building
Pierre, South Dakota 57501

JUN 1978 REC'D Remote Sensing Institute

June 27, 1978

Mr. Victor I. Meyers
Director, Remote Sensing Institute
South Dakota State University
Brookings, South Dakota 57006

Dear Mr. Meyers:

This is to advise that the data gathered and interpreted employing the use of Remote Sensing techniques during the current Fiscal period has been extremely beneficial in the development of Meandered Lake plans, particularly in the planning of the Level Ditching project in Lake Thompson completed this past winter.

In addition, it appears that Remote Sensing has an excellent potential for the continuous monitoring of Meandered Lake beds and in providing base data for formulating management plans.

Along with the Meandered Lake data, the monitoring of changes on other public lands will aid our Department personnel in the application of management practices.

Sincerely,

Jerry Lounsberry

Jerry Lounsberry, Director
Division of Wildlife

JL/JTS/d

SOUTH DAKOTA

Stratification of Landsat Data by Soil Associations For Partitioning Corn Spectral Signatures

by
K.J. Dalsted and M.E. DeVries

INTRODUCTION

Since its inception, Landsat has been supplying data that has numerous applications throughout the diverse and important field of agriculture. One such use is the estimation of crop acreage; the classification of Landsat Computer Compatible Tapes (CCT) in crop acreage estimation studies has produced various degrees of success (Bizzel et al., 1976; Stockton et al., 1975; Heilman, 1974). The levels of accuracy in these and other studies have been influenced greatly by numerous site factors, including physiographic region, agricultural diversity, climate, and soil influences. The degree of these influences fluctuates from region to region. Wiegand et al. (1977) noted the need for "tuning" of crop classification studies to the crops and soils of physiographic regions.

Other researchers have also noted the effect of soils on Landsat-recorded vegetative signatures (Richardson et al., 1975; Myers et al., 1977a and 1977b; Tucker and Miller, 1977) and of vegetation on Landsat-recorded soils discrimination (Siegal and Goetz, 1977; Westin and Lemme, 1978). These studies point out the need for a more thorough understanding of the heterogeneity of edaphic, agronomic, and climatologic factors which influence crop signatures.

The Statistical Reporting Service, now called the Economic Statistics and Cooperatives Service (ESCS) has conducted numerous and extensive investigations using Landsat data to determine crop classification accuracies as well as augmenting operational ground survey methods (Sigman et al., 1977; Von Steen and Wigton, 1976). Their classification results have been mixed. ESCS investigators concluded from an Idaho study that temporal analysis may improve the results of crop acreage estimates (Von Steen and Wigton, 1976). They also noted that crop variety, soil type, weather conditions, and state of crop maturity caused unaccounted for variation in crop signatures over a large area. The ESCS uses a stratification system which is based on percentage or intensity of cultivation. However, this system divides regions into only three to four strata of different intensity of cultivation. Soil information is accounted for only in its effect on the general land use patterns of an area.

Since the use of Landsat data (together with salient satellite data) appears to offer the only viable, real-time means of assessing world-wide crop acreage and subsequent production, every means of improving classification accuracy should be investigated. Vegetation or crop production is respondant to its growth environment. It follows that any attempt to classify crop acreage and its ensuing production should exploit the existing resource information, within reasonable bounds, that will improve classification of Landsat data.

This study will explore the efficacy of incorporating existing soil association data into crop area estimates from Landsat data. The attributes of soil interpretive data will also be examined.

BACKGROUND

Before any validity can be attached to the results of this study, certain assumptions must be shown to hold true. The following discussion will list and elaborate upon the key concepts which underlie this investigation.

By definition a soil association is "...two or more defined taxonomic units occurring together in a characteristic pattern ..." (Soil Science Society of America, 1978). This definition does not preclude the multiple occurrence of a soil unit in more than one soil association; however, the significance of this definition is in "characteristic pattern." A soil association will usually have from one to three major soil components although minor soils or inclusions may occur in all but the most homogeneous units. Soil association maps are designed to show general soil-landscape relationships. This is contrasted with detailed soil surveys which are made for more intensive applications.

Boundary accuracy in soil association maps, as in detailed soil maps, is often qualitative, being based on the judgment of the individual mapper. The gradual continuum of change and the complexity of soils on the landscape, including transitional soils, limits the accuracy of line maps which discretely divide soil units. Although these boundary effects do occur to a degree in the study area of this project, this investigation will assume the boundaries to be accurate, discrete delineations. The only other option is the use of detailed soil information which is at a resolution which is too small to consistently use in conjunction with Landsat data.

In order to establish the utility of soil association data in Landsat stratification procedures, it must first be established that soil associations are visible on Landsat imagery. Secondly, the influence of soil properties within soil associations that cause variation in crop growth or phenology and the succedent crop signature must be evaluated. Thus the influence of soils on crop signatures may be direct (background reflectance) or indirect (crop phenology).

The use of Landsat data to identify soil association units has been well documented (Frazee et al., 1972; Westin and Frazee, 1976; Weismiller et al., 1976). The visibility (i.e. characteristic reflectance) of soil associations on Landsat must be established because until the crop canopy closes completely, soil background reflectance will be part of the integrated crop signature. Stratification by soil association or accounting for variable soil background reflectance may allow acquisition of data of the earliest possible optimum date for classification of a Landsat CCT - a date which may have considerable soil reflectance in the composite crop signature. This is crucial since turn-around time is an important factor in establishing real-time crop inventory estimations.

The next assumption to validate is the influence of soil properties on crop signatures. Kristof and Baumgardner (1975) concluded that certain soil properties influence the reflectance of vegetative cover. They also noted the integrated effect of soil background on maize signatures using a June 30 CCT. However, their study employed data obtained from a low level aircraft scanner. Landsat data may not have the resolution capability to discriminate differences that are of a small magnitude or occur in a small area. Westin and Lemme (1978) point out that different soil drainage and soil profile characteristics influence the planting date and the rate of growth of the respective crops. Thus, canopy development and crop phenology are variables which may influence the results of Landsat classification on dates after the soil has been masked by the canopy. Their study endeavored to recognize soil associations through crop canopy differences. However, the results indicated that soil associations in their study area could not be consistently separated within one vegetative type.

Conceptually, a number of soil properties could cause variability within crop signatures. Moisture and fertility are two strong influences on crop development. Certain soil properties such as inherent fertility, texture, organic matter content, depth of solum, slope, and parent material are important in providing the crop with adequate moisture and fertility. Consequently, this would lead one to believe that, under similar management, a crop growing on a highly productive soil will have a different signature than the same crop growing on a less productive soil at the same time of the growing season. (This is not intended to suggest that Landsat data can easily be used to estimate crop yield since biomass production in many crops isn't always directly correlated to yield.) Obviously if soil reflectance is not masked by canopy at the time of classification, soil color will be a major factor, but the preceding discussion was aimed at a situation with a closed canopy. If the theory of soil productivity is to be tested, then a soil interpretive grouping which provides an integrated representation of the soil properties which influence productivity is needed.

Soil productivity groupings have been developed through the Soil Conservation Service (SCS). The SCS groupings are based on a relative scaling of soil series in a particular area (usually a county). The soils are rated as to production of the various crops grown in a particular region. This interpretation will be used to aid in analysis of the crop per soil signatures.

This study will investigate: 1) the influence of soil background reflectance on Landsat-derived corn signatures at two dates; and 2) the influence of soil properties on Landsat-derived corn signatures.

STUDY AREA

The area of study (Fig. 6) is located in a warm, moist prairie region (Udic Ustoll soils). Soils in this area have formed in an interlace of topography, surface texture, and parent material (Fig. 7).

Two intensities of data collection were undertaken. An intensive study site was located in southern Lincoln County (Fig. 7). In this area, corn is a major crop and contrasts in parent material and slope are quite evident on the landscape. The general or less intensive study area encompassed the rest of the region. Data collected in this area were sufficient for overall analyses but not as spatially concentrated as in the intensive area.

METHODS AND MATERIALS

The approach to data collection, extraction, and analysis is very similar to recent efforts in South Dakota (Myers et al., 1977a

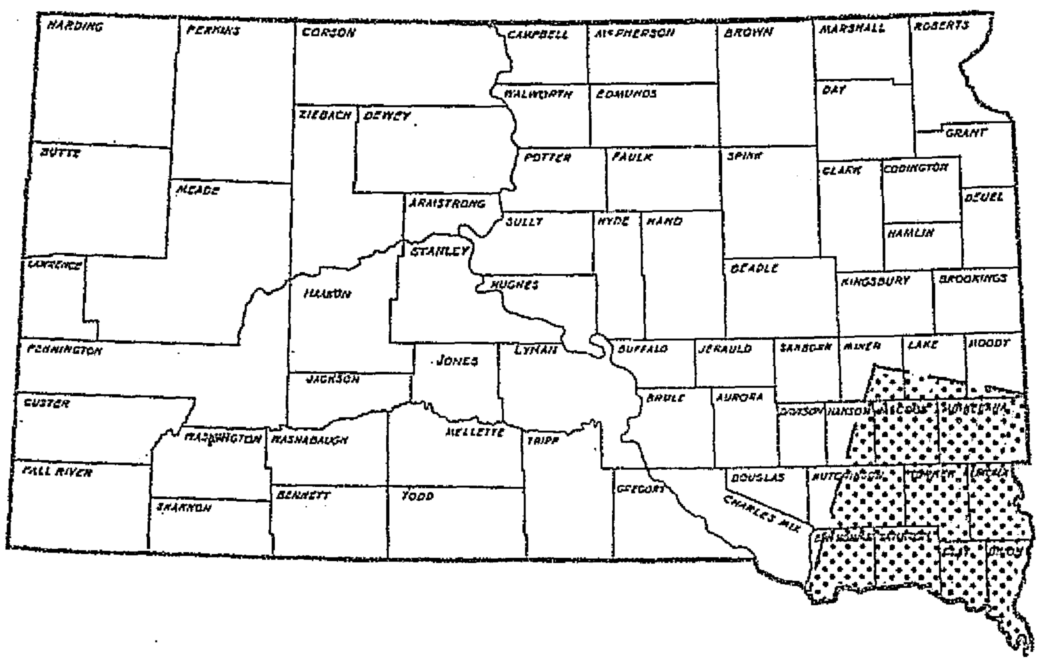


Fig. 6. Location of study area in agricultural censusing.

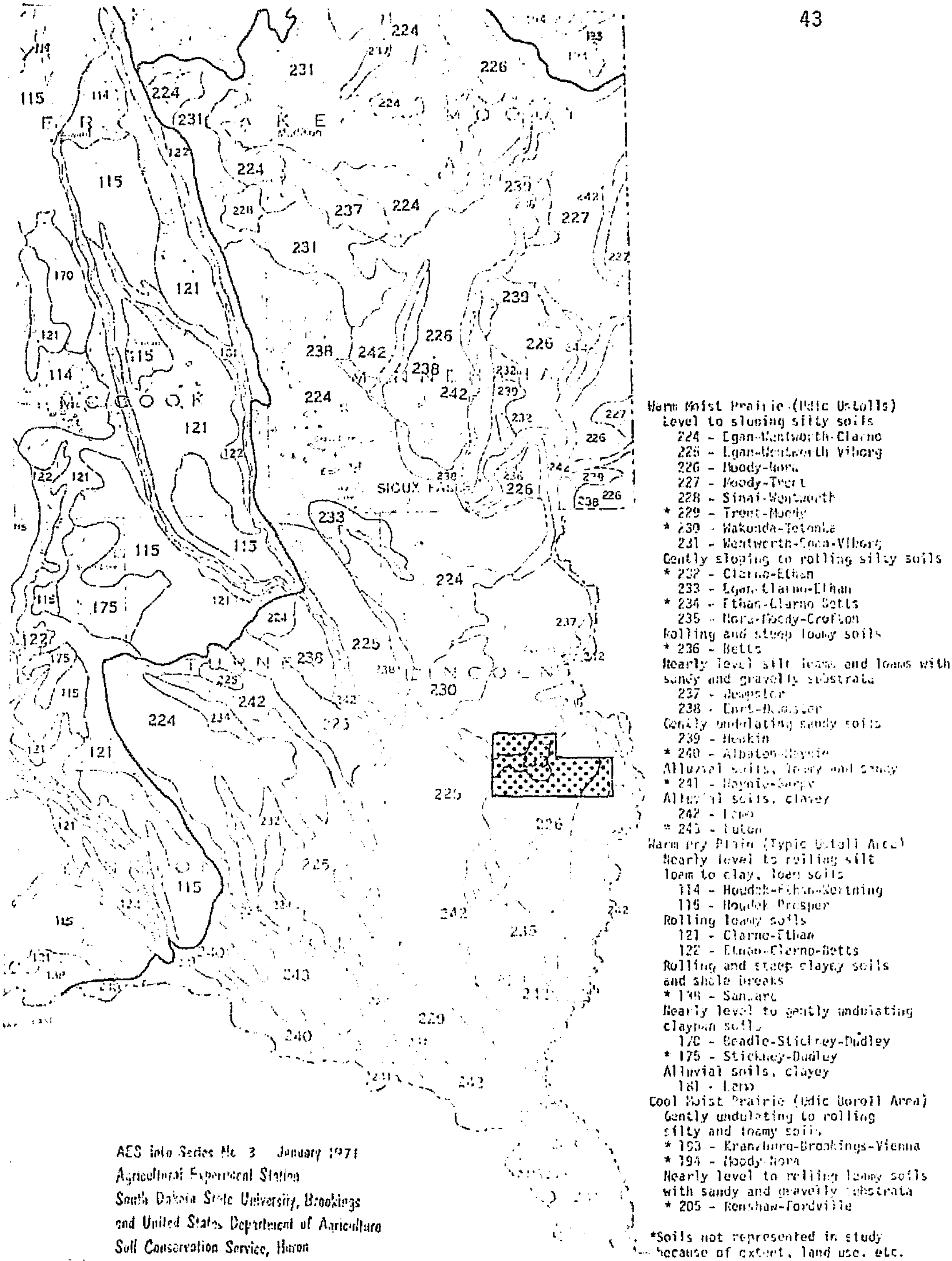


Fig. 7. Soil associations of southeast South Dakota. The blocked area is the location of the intensive study site. The dark line running through the length of the map differentiates the Warm, Moist Prairie from the Warm, Dry Prairie.

and 1977b). A review of these reports and the proposed Fiscal Year 1979 project discloses the evolution of methods reported in this study.

Field data were collected by strata on a soil association basis to establish the contribution of soil to row crop signatures. However, several soil associations in the study area were not represented in this study (Fig. 7). The determination of the inclusive soil associations was based on areal extent within the study area, land use (i.e. cropped or not), and shape (e.g. narrow floodplain soils). In this manner only the major cropped soil associations in the study area were sampled.

The collection of ground data was undertaken during the summer of 1977. The data collected included various crops but the major crop was corn. The collection of corn and other crop data was dictated by the locations of major soil associations within the study area. Ground observations were made at road intersections and crop locations were recorded on county highway maps (scale approx. 1:140,000). Prior to field work, soil association boundaries were transferred to county highway maps via a reflecting projector. Thus, a running tally could be kept to insure that sufficient ground data (minimum 10 fields per association) were collected within each major soil association. The road intersections also provided an efficient means of registering or locating the field data on Landsat data. In addition, Skylab coverage of the study area provided a valuable means of recognizing and confirming the shape and configuration of the fields as they were notated on the highway map.

Two Landsat CCT (Computer Compatible Tapes) were obtained for data extraction, 5 June 1977 I.D. 2865-16123 and 29 July 1977 I.D. 2919-16095. Some fields in the general study area had to be deleted from the 29 July analyses because the July scene coverage had shifted to the south and some cloud coverage also occurred.

Shadeprints, or symbol overstrike line printouts, were generated with six levels chosen to maximize the contrast over the area of interest. Obvious features, such as roads, railroads and lakes, were used as reference points for locating the cornfields on the shadeprints. A rectangular area within each field was delineated such that this rectangular area contained no border pixels. Coordinates and rectangle size were then recorded for each field. This information was first used to verify the field locations and then to extract the pixel data for each field. The program which was used to extract the field data also computes the mean and standard deviation for all MSS bands of each field. Those fields with suspect means and/or high standard deviation were checked for possible shadeprint positional errors.

After the fields had been satisfactorily located and extracted, statistical analysis procedures were initiated. Three statistical procedures from the Statistical Analysis System (SAS) package (Barr et al., 1976) provided the bulk of the analyses. These procedures were MEANS, ANOVA and DUNCAN. The MEANS procedure produces simple univariate descriptive statistics for data set numeric values, including the mean, standard deviation, coefficient of variation, and minimum and maximum values. ANOVA is an analysis of variance procedure used for testing for significant differences among mean values. The DUNCAN procedure performs Duncan's multiple range test to separate the data into significantly different (or significant) groups.

The intensive study data were grouped several ways for analysis. One data set contained all Lincoln County corn data grouped by soil association. A second data set contained all the data sorted by detailed soil mapping units. A third data set had an even sample size, that is, a 175 randomly selected pixels per soil association; this data set was used almost exclusively in intensive area analysis and interpretation.

A fourth data set of Lincoln County data was generated when the data were transformed using the Kauth and Thomas Transform [KTT] (Kauth and Thomas, 1976). All of the ground truth area (all pixels - not just corn fields) was also extracted from the June CCT and transformed using the KTT transform. The four resulting channels of information (soil, green stuff, yellow and non-such) were output to an Optronics black-and-white film recorder for further photo-interpretive analysis.

The SAS package was used to generate the same statistics for the general study area as were used for the intensive study site. Detailed soil maps were not available for much of the general area - hence only soil association data were used. Since the scope of the general study area was broader and more spatially spread than the intensive area, the fewer ground data that were collected resulted in even sample sizes of 75 pixels per soil association.

Table 5 shows the various data sets for both the Lincoln County and general study areas and the statistical analyses performed each.

Table 5. Statistical tests and procedures used to evaluate 1977 CCT-extracted crop inventory data.

STUDY AREA	Means 05 June	Means 29 July	Analysis of Variance (ANOVA)[†] 05 June	Analysis of Variance (ANOVA)[†] 29 July	Duncan's New Multiple Range Test 05 June	Duncan's New Multiple Range Test 29 July	Difference Data 05 June vs 29 July
Intensive							
Soil Association Data							
All data	x	x					means
Equal sample size (175 pixels)	x	x	x	x	x	x	means, ANOVA, DUNCAN
GIN transforms[§]	x	x					
Detailed Soils							
All data	x	x					means
General							
Soil Association Data							
All data	x	x					
75 pixels	x	x	x	x	x	x	

[†] Where checked (x), ANOVA tested differences among soil association or detailed soils. Also, in the intensive study area - 175 pixel category - field differences within each association were tested.

[§] Green index transformations, GIN, (Kauth and Thomas, 1976) completed for 05 June and 29 July CCT.

Results and Discussion

The results and discussion of this study are partioned into the two areas of investigation - the intensive and the general study area. Both study areas have been stratified by soil association data compiled by Westin et al., 1971. Initially, the intensive study in Lincoln County was also stratified by detailed soils information.

At the onset of this study two intensities of study were established to: 1) evaluate contrasting soils within a close proximity of one another as they affect crop signatures, hence the intensive area; and 2) evaluate crop signatures across a Landsat scene for a regional picture, hence general study area. Landsat data encompassing the intensive study area, because of its reasonably small size, was transformed using the Kauth & Thomas transformation. Also, more detailed information is incorporated in the intensive study data set.

Intensive Study Area

This area, which is located in southeast Lincoln County, has a wide range of soil properties. The major soil properties which influence soil background reflectance and crop growth in this area are: 1) slope; 2) parent material, loess derived soils being lighter colored than glacial drift derived soils in this area; and 3) horizon thickness, an indicator of erosion or accretion, landscape position, and inherent fertility (Table 6). Soil productivity ratings, which

Table 6. Important soil properties and productivity rating of soil associations in intensive study area.

Soil Association[†]		Slope	Parent Material[§]	Thickness Of A Horizon	Solum (A+B Horizons)	Production rating. Corn (bu) under common management under dry land[¶]
225	Egan	Nearly level (0-3%)	Drift and till	8 in	30 in	40-49
	Wentworth			7 in	34 in	50-52
	Viborg			17 in	34 in	50-53
226	Moody	Nearly level to sloping (0-9%)	Loess	6 in	38 in	50-53
	Nora			7 in	30 in	35-50
233	Egan	Gently sloping to rolling (3-15%)	Drift and till	8 in	30 in	30-47
	Clarno			8 in	30 in	30-40
	Ethan			6 in	24 in	25-30
235	Nora	Gently sloping to strongly sloping (3-15%)	Loess	7 in	30 in	35-50
	Moody			6 in	38 in	35-50
	Crofton			6 in	6 in	20-30

[†] Association soils are listed in order of dominance

[§] Most soil textures are silty, inclusions are loamy

[¶] No B horizon present

are used to help analyze the ensuing results, are assumed to represent the integrated effect of all the soil properties which affect yield. The indication of Table 6 is that soil 225 and 226 are the most productive and soil 233 is the least productive.

Detailed soil information (Driessen et al., 1976) further divides the soil association data into another level of resolution. Table 7 lists the mapping units and the frequency of occurrence within the soil associations. It is obvious that detailed soils information does not always adhere to soil association boundaries. This is to be expected since soil association data are general and represent only a characteristic pattern of soil occurrence (see Background section). Only the mapping units which occurred more than 8 percent in the field ground data are listed.

Soil property and productivity ratings for these soils are listed in Table 8. Productivity ratings are somewhat similar for soils of the same slope. The shallower soils are usually found on the steeper slopes, are generally lighter colored, have less available moisture, and are lower in fertility than the surrounding, deeper soils. It follows that in a normal year shallow soils will have shorter, perhaps moisture- and fertility-stressed plants and that soil influence may be more apparent in CCT crop signatures. Because of variable occurrence, i.e. more than one mapping unit per field, analysis and accompanying statistics of Landsat data was not reported on the detailed soil mapping units.

Table 7. Percentage of mapping unit occurrence within soil association data

Mapping Unit		Frequency Within Soil Association (%) 225	226	233	235
WhA[†]	Wentworth-Chancellor si cl[§]	35	4	39	--
Ca	Chancellor-Tetonka si cl	31	--	17.5	--
EcB	Egan-Chancellor si cl	15	--	17.5	--
AcA	Alcester si cl	--	6	--	22
MoB	Moody si cl	--	32	--	22
MpC2	Moody-Nora si cl, eroded	--	28	--	30
CpD2	Crofton-Nora sil[¶]	--	6	--	7
MoA	Moody si cl	--	8	--	--
EaB	Egan si cl	19	2	26	--
Ah	Alcester and Lamo si cl	--	8	--	4
MpB	Moody-Nora si cl	--	6	--	15

[†] A,B,C,D refer to slope 0-3,3-6,6-9,9-15 percent respectively and § indicates an eroded soil phase

[§] Si cl - silty clay loam

[¶] Sil - silt loam

Table 8. Important soil properties and productivity ratings of soils in the intensive study area - detailed mapping units.

Mapping Unit		Slope	Parent Material§	Thickness of A Horizon (inches)§	(Solum) A+B Horizon (inches)§	Productivity of corn dryland under common management (bu)†
Wha	Wentworth- Chancellor	0-2	Drift, alluvium	7 18	34 36	60
Ca	Chancellor- Tetonka	0-2	Alluvium	18 19	36 41	50
EcB	Egan- Chancellor	2-4	Drift, alluvium	8 18	30 36	58
AcA	Alcester	0-2	Alluvium	18	42	74
MoB	Moody	2-6	Loess	6	38	68
MpC2	Moody- eroded Nora,	6-10	Loess	6 7	38 30	48
CpD2	Crofton- Nora	9-17	Loess	6 7	6 30	--¶
MoA	Moody	0-2	Loess	6	38	72
EaB	Egan	3-6	Loess	8	30	61
Ah	Alcester and Lamo	0-2	Alluvium	18 19	42 34	74 65
MpB	Moody- Nora	2-6	Loess	6 7	38 30	68

† Rating is based on the mapping unit as a whole not the component soils

§ Values obtained from SCS soil series descriptions

¶ Not suitable for cultivation, consequently a crop estimate would be very low

With the aid of detailed and general soil information, the soil association data appear to break into two apparent groups, (1) 225 and 233 and (2) 226 and 235. This is an obvious break because of parent material differences. Thus, a prior partitioning of the influence of these soils on crop signatures occurred before the examination of Landsat CCT data.

The means of Landsat CCT data and Kauth and Thomas Transformations (KTT) are listed by soil association in Table 9. The ratioed data on both CCT dates has very little range. Band 5 and 6 have the wider range of means on the early date; while Band 6 and 7 have the wider spread on the later date. KTT 1, which is transformed to enhance soils, loses much of its spread in mean values as the canopy closes from June to July. The change in other KTT transformed data do not have as a dramatic change. Similarity of mean values is noted for the loess soils and the drift soils.

ANOVA-derived F values on the fields within each soil association are listed in Table 10. This test provides a method of analyzing variability within each soil association. Soil 235, which has the most soil variability, especially slope, also has the greatest variability of reflection among fields. This is expected because of the range in soil colors within this association. The high values in Bands 6 and 7 for soil 235 may be explained in part by anomalous data, such as grassed waterways or tree clumps which were not observed in the field during data acquisition.

Table 9. Mean Digital Count and KTT transformations means (175 pixels per soil association) of soils information in the intensive study area.

Soil†	Ratio(5/7)		Band 4		Band 5		Band 6		Band 7		KTT 1 (Soil brightness)		KTT 2 (Vegetation)		KTT 3 (Yellow stuff)		KTT 4 (Nonsuch)	
	1977																	
	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul		
225	1.61	0.64	23.70	15.95	28.52	15.40	41.03	45.47	18.46	24.60	80.37	52.81	114.52	102.67	125.01	93.96	99.14	114.77
226	1.74	0.57	24.63	15.88	31.22	15.03	41.78	48.31	18.32	26.60	87.91	56.74	110.30	108.52	127.93	94.54	98.02	114.93
233	1.78	0.60	21.78	15.93	26.13	14.92	34.64	46.11	15.11	25.16	59.26	53.25	103.78	104.54	124.73	93.37	100.35	115.04
235	1.72	0.59	26.24	16.13	33.52	15.36	45.64	48.63	20.24	26.65	102.78	57.78	114.79	108.42	128.38	94.39	97.40	114.71

† Refer to Table 1 for a listing of association soil components

Table 10. ANOVA results for fields within soil association.

F Values
All data are significant at 0.001 level of significance

Soil	Num. d.f. (fields)	Denom. d.f.	Band 4		Band 5		Band 6		Band 7		Ratio	
			05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul
225	18	172	21.58	3.56	17.72	3.69	25.42	18.99	18.85	12.72	12.74	4.17
235	15	171	35.23	8.22	45.83	18.42	44.58	43.18	41.78	36.53	25.83	21.39
226	43	416	19.68	4.06	19.74	5.42	38.32	20.24	27.61	18.85	7.04	7.53
233	26	296	23.91	3.11	24.35	5.67	29.96	11.13	27.94	15.64	8.27	8.87

ANOVA tests on differences among soil associations (75 pixels per unit) are shown by F values in Table 11. Again, changes from early season to later season are apparent. Relative significance on these results can be estimated in general by ranking the F values. One such ranking method is the DMRT. The results of the Duncan Multiple Range Test (Table 12) provide a statistical way of determining which means are significantly different from each other (5 percent level of significance). Early season differences are most apparent on Bands 4 and 5, while the later date differences are revealed in the infrared bands, Bands 6 and 7.

An obvious conclusion from the foregoing results is that early season signature differences are most apparent on the visible bands while later season differences which are due to crop differences are best noted by IR reflection on Bands 6 and 7. Thus, before the crop canopy closes, the visible bands recorded the varying soil background reflectance as part of the signature. After canopy closure the soil productivity ratings provided an interpretation which helped partition Landsat crop signatures. Although the differences at the later season do not appear to be very great, they are significant as determined by statistical tests.

General Study Area

The general study area encompassed one Landsat scene. The cropped soils in this area do not have the abrupt contrast which is found in the intensive study site. Type of soil parent material

Table 11. ANOVA results of soil associations within Landsat bands in the intensive study area.

	F Value†		Coefficient of Variation (%)	
	05 June 1977	29 July 1977	05 June 1977	29 July 1977
Ratio (5/7)	8.49	7.16	20.13	25.95
Band 4	48.13	1.69§	15.26	9.23
Band 5	49.21	2.84¶	21.13	16.18
Band 6	63.61	19.43	19.29	10.45
Band 7	47.56	24.03	23.65	11.63
KTT 1 Soil brightness	72.98	21.00	35.42	13.67
KTT 2 Green index	15.12	16.53	16.58	9.50
KTT 3 Yellow stuff	18.39	1.57¶	4.85	3.94
KTT 4 Nonsuch	20.38	0.08§	4.09	3.00

† All F values significant at .001 level of significance except those noted

§ Not significant

¶ Significant at .05 level of significance

Table 12. Significant separations of mean digital counts of soil associations in the intensive study areas as determined by Duncan's Multiple Range Test. Level of significance is .05.

Ratio		Band 4		Band 5		Band 6		Band 7	
05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul
233	225	235	235	235					
226					225	235	226	235	226
235		226	225	226	235		235		235
	233		226		233	226		225	
225	235	225	233	225	226	225		226	
	226						233		233
		233		233		233	225	233	225

is relatively uniform throughout the scene. A listing of pertinent soil properties and productivity ratings for this area is presented in Table 13.

Landsat-derived corn digital means are listed on a per soil association basis in Table 14. A comparison of Tables 13 and 14 reveals a general relationship between corn yield and Landsat derived reflectance. Low reflectance of the June Band 5 data is related to high July Band 7 reflectance. A striking exception is the 237 soil association unit; its 5 June Band 5 reflectance is low and its 29 July Band 7 reflectance is also low. This indicates, if the CCT data has been extracted correctly, that the soil property of sand and gravel at 20-40 inches below the soil surface has had some influence on the reduced amount of biomass due to moisture stress. Another exception is the 226 unit which behaves in a reverse manner, i.e. high June Band 5 and high July Band 7. This is caused by the light colored loess parent material.

The F values and coefficient of variation results are listed in Table 15. These results indicate that the difference among soils per Landsat band are significant. In general, a qualitative way of looking at the data involves ranking the F values. Again, the differences between dates entail a Band 5 decrease and a Band 7 increase in F values. This data indicates only that one or more of the soil associations is significantly different from the others.

Table 13. Important soil properties and productivity ratings of the soil associations in the general study area.

Soil Association		Slope	Parent Material	Thickness of A horizon (inches)	A+B horizon (inches)	Productivity rating, corn on dryland under common management (bu)
224	Egan	Level to	Drift and till	8	30	42-63[†]
	Wentworth	sloping		7	34	51-63
	Clarno			8	30	66-67
226	Moody	Level to	Loess	7	38	38-64[†]
	Nora	sloping		7	30	27-55
227	Moody	Level to		7	38	38-64
	Trent	sloping	Loess	8	37	72
228	Sinai	Level to	Drift and till	7	32	35-54
	Wentworth	sloping		7	34	51-63
231	Wentworth	Level to	Drift and till	7	34	51-63[‡]
	Egan	sloping		8	30	42-63[†]
	Viborg			17	34	66-67
237	Dempster	Nearly level	Outwash terrace[¶]	9	32	40-53
238	Enet	Nearly	Outwash terrace[¶]	38	28	40-41
	Dempster	level		9	32	40-53
239	Henkin	Gently Undulating	Outwash	6	32	33-40[§]
242	Lamo	Nearly level	Alluvium	28	28	30-65
114	Houdek	Nearly	Drift and till	6	28	33-53
	Ethan	level to		6	24	35-39[§]
	Worthing	rolling				44[§]

Table 13. Continued

Soil Association	Slope	Parent Material	Thickness of A horizon (inches)	A+B horizon (inches)	Productivity rating, corn on dryland under common management (bu)
115 Houdek	Nearly level to rolling	Drift and till	6	28	33-53
Prosper			9	30	54-57
121 Clarno	Rolling	Drift and till	8	30	31-54
Ethan			6	24	35-39
122 Ethan	Rolling	Drift and till	6	24	35-39†
Clarno			8	30	31-54#
Betts			4	8	26-30#
170 Beadle	Nearly level to undulating	Drift and till	7	30	38-45
Stickney			10	34	37-45
Dudley			9	30	24-26

† Lowest yield on eroded slopes

§ Highest yield on drained field

¶ Silt loam and loam over sandy and gravelly strata

Yield under drained conditions

Table 14. Mean digital count of soil associations in general study area.

Soil	Ratio (5/7)		Band 4		Band 5		Band 6		Band 7	
	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul
224	1.58	0.65	21.61	17.12	26.13	16.05	37.38	46.25	17.13	25.04
226	1.74	0.85	26.96	17.20	35.48	18.56	46.29	43.73	20.25	22.94
227	1.68	--[†]	21.94	--	26.56	--	36.18	--	16.10	--
228	1.73	0.72	24.52	17.20	28.17	16.44	38.06	44.72	16.48	23.33
231	1.72	0.65	23.44	16.38	28.60	16.68	37.72	44.97	16.72	24.18
237	1.64	0.99	21.69	17.56	25.84	18.70	36.37	40.20	16.74	20.02
238	1.68	0.63	24.05	16.17	30.60	15.14	40.52	46.25	18.33	24.77
239	1.57	--[†]	22.81	--	26.65	--	37.93	--	17.82	--
242	1.70	0.66[††]	22.41	17.84	26.24	17.12	35.42	50.81	15.66	27.18
114	1.75	0.76	23.88	17.84	29.12	17.85	38.13	45.09	16.96	23.62
115	1.58	0.88	24.33	18.62	29.08	19.98	40.64	44.54	18.77	22.94
121	1.66	0.82	25.32	18.30	31.17	19.09	42.04	44.18	19.12	23.30
122	1.65	0.79	25.17	18.21	30.85	19.01	42.34	45.77	19.00	23.93
170	1.70	0.70	25.45	18.05	31.20	17.84	39.98	48.68	18.45	26.06

[†] July CCT data not available due to scene shift or cloud coverage

[††] 33 pixels in this sample

Table 15. ANOVA results of soil associations with Landsat bands in the general study area.

	F Value†		Coefficient of Variation (%)	
	05 June 1977	29 July 1977	05 June 1977	29 July 1977
Ratio (5/7)	3.60	16.56	16.61	30.95
Band 4	20.87	11.60	12.76	11.40
Band 5	25.06	11.90	16.19	22.72
Band 6	16.17	12.98	16.22	11.07
Band 7	10.32	19.54	20.39	12.76

† All data are significant at .001 level of significance.

The Duncan's Test on the general study area data is shown in Table 16. The significant groupings as derived from Duncan's reveals a similar although weaker trend as was observed in the intensive study site. That is, the lower reflectance on Band 5 leads to a high reflectance on Band 7 in the groupings. There are several factors which confuse the results in this study area. Planting date varies from south to north because of soil temperature. Variable rainfall amounts inevitably occur. In the general study area on June 5, the corn was approximately 10 inches tall; however, the corn which was taller may have caused differences in the Band 7 July readings.

The general study area, because of its size and the factors previously enumerated, did not produce results which were as reliable as those noted in the intensive study site. The range in production capability of the various soils also reduced the worth of that interpretation in the analysis. The soils examined in this study vary in slope and the amount of erosion which has taken place; this results in variable reflections and yield potentials.

Summary and Conclusions

Stratification of Landsat CCT by soil associations produced strata of significantly different corn spectral signatures. The influence of soil properties was more apparent in the June 5 CCT than the July 29 CCT. Data from both dates were significant,

Table 16. Duncan's Multiple Range Test (.05 level of significance) in the general study area data. Observations per soil association are 75 pixels in all but the denoted soils.

Ratio		Band 4		Band 5		Band 6		Band 7	
05 Jun	29 Jul§	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul	05 Jun	29 Jul
114	237	226	115	226	115	226	170	226	170
226	115	170	121	170	121	122	238	121	224
228	226	121	122	121	122	121	224	122	238
231	121	122	170	122	237	115	122	115	231
170	122	228	114	238	226	228	114	170	122
227	114	115	237	114	114	170	231	238	114
238	228	238	226	115	170	114	228	239	228
121	170	114	228	231	228	228	115	224	121
122	231	231	224	228	224	239	121	114	115
237	224	239	231	239	231	231	226	237	226
115	238	227	238	227	238	224	237	231	237
224		237		224		237		228	
239		224		237		227		227	

§ Soil 227 and 239 are not in the July data set due to cloud coverage and Landsat. Soil 242 had only 33 pixels.

however, the July 29 data was less obvious and more subtle than the June 5 CCT.

Data from the intensive study site produced results that were more reliable or predictable from a soils viewpoint than those of the general study area. The general study area data because of its variability in growth stage had less predictable results, although the general trends or responses of the soil association data was still close to that which was expected.

A general relationship between Landsat Bands 5 and 7 was apparent during the course of the study. Soil properties and productivity ratings aided greatly in interpreting the response of various soil association - corn data. The higher producing soils on the less sloping terrain had a low 5 June Band 5 reflectance (exception is loess derived soils) and a high Band 7 July 29 reflectance. Results which didn't follow this general trend could be explained by particular causitive soil properties.

Results from this study lead to recommendations for further investigation in different physiographic regions and using a computer classifier to test improvement of soil stratification.

References

Barr, A.J., J.H. Goodnight, Q.P. Sall and Q.T. Helwig. 1976. A user's guide to SAS 76. SAS Institute, Inc. Raleigh, N.C.

Bizzel, R.M., F.G. Hall, A.H. Feiverson, M.E. Bauer, B.J. Davis, W.A. Malila, and D.P. Rice. 1975. Results from the crop identification technology assessment for remote sensing (CITARS) projects. Proc. 10th International Sympos. of Remote Sensing of Environ., Ann Arbor, Michigan.

Driessen, J.L., G.E. Murphy and W.D. Wiesner. 1971. Soil survey of Lincoln County, South Dakota. SCS-USDA and S. Dak. Exper. Sta., U.S. Gov. Print. Off., Washington, D.C.

Frazee, C.J., V.I. Myers, and F.C. Westin. 1972. Density slicing techniques for soil survey. Soil Sci. Soc. Amer. Proc. 36:693-695.

Frazee, C.J., P.H. Rahn, F.C. Westin, and V.I. Myers. 1974. Use of ERTS-1 imagery for land evaluation in Pennington County, South Dakota. p. 549-568. Proc. ninth intl. symp. on remote sens. of envir. Ann Arbor, Michigan.

Heilman, J.L. 1974. The application of ERTS-1 multispectral imagery for crop resource evaluation and inventory. M.S. Thesis, S. Dak. State Univ., Brookings.

Kauth, R.J., and G.S. Thomas. 1976. The tasselled cap - a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. Proc. Symp. Machine Processing of Remote Sensing Data, 4B:41-51. LARS, Purdue Univ., West Lafayette, Ind.

Kristof, S.J., and M.F. Baumgardner. 1975. Changes of multispectral soils patterns with increasing crop canopy. Agronomy J. 67:317-321.

Myers, V.I., Best, R.G., K.J. Dalsted, J.C. Eidenshink, M.E. Wehde. 1977a. Application of remote sensing technology in South Dakota to assess wildlife habitat change, describe meandering lakes, improve agricultural censusing, map aspen, and quantify cell selection criteria for spatial data. Semi-Annual Report to NASA, SDSU-RSI-77-17. Brookings.

Myers, V.I., K.J. Dalsted, R.G. Best, J.R. Smith, J.C. Eidenshink, F.A. Schmer, A.S. Andrawis, and P.H. Rahn. 1976. Application of remote sensing in South Dakota to provide accurate inventories of agricultural crops, enhance contrast in photographic products, monitor rangeland habitat loss, map aspen, and prepare hydrogeologic surveys. Annual Progress Report to NASA. SDSU-RSI-77-08, Brookings.

Richardson, A.J., C.L. Wiegand, H.W. Gausman, J.S. Cuellar, and A.H. Gerberman. 1975. Plant, soil, and shadow reflectance components of row crops. Photogram. Eng. and Remote Sens. Vol. 41(11):1401-1407.

Siegal, B.S., and A.F.H. Goetz. 1977. Effect of vegetation on rock and soil type discrimination. Photogram. Eng. and Remote Sens. 43(2):191-196.

Sigman, R., C.P. Gleason, G.A. Hanuschak, and R.R. Starbuck. 1977. Stratified acreage estimates in the Illinois crop-acreage experiment. Machine Processing of Remotely Sensed Data Sympos. Purdue Univ., West Lafayette, Ind.

Soil Science Society of America. 1978. Glossary of soil science terms. 677 South Segoe Road, Madison, Wi.

Stockton, J.G., M.E. Bauer, B.O. Blair, and M.F. Baumgardner. 1975. The use of ERTS-1 multispectral imagery for crop identification in a semi-arid climate. LARS Info. Note 040775. The Laboratory for Application of Remote Sensing. Purdue University, West Lafayette, Ind.

Tucker, C.J., and L.D. Miller. 1977. Soil spectra contributions to grass canopy spectral reflectance. Photogram. Eng. and Remote Sens. Vol. 43(6):721-726.

Von Steen, D.H., and W.H. Wigton. 1976. Crop identification and acreage measurement utilizing Landsat imagery. Statistical Reporting Service, USDA, Washington, D.C.

Weismiller, R.A., I.D. Persinger, and O.L. Montgomery. 1977. Soil inventory from digital analysis of satellite scanner and opographic data. Soil Sci. Soc. Am. J., 41:1166-1170.

Westin, F.C., et al. 1971. Soil associations of South Dakota. Ag. Exper. Sta. Info. Series No. 3. So. Dak. State Univ., Brookings.

Westin, F.C., and C.J. Frazee. 1976. Landsat data, its use in a soil survey program. Soil Sci. Soc. Amer. Proc. 40(1):81-89.

Westin, F.C., and G.D. Lemme. 1978. Landsat spectral signatures: studies with soil associations and vegetation. Photogram. Eng. and Remote Sens. Vol. 44(3):315-325.

Wiegand, C.L., H.W. Gausman, R.W. Leamer, A.J. Richardson, J.H. Everitt, and A.H. Gerbermann. 1977. Soil, water, and vegetation conditions in south Texas. Type III Final Report for Goddard Space Flight Center, Greenbelt, Md.

THE USE OF REMOTE SENSING TECHNOLOGY TO MAP ASPEN

by
J.C. Eidenshink and F.A. Schmer

INTRODUCTION

The South Dakota Department of Game, Fish and Parks is currently involved in the study of the management potentials of aspen (*Populus tremuloides*) within the Black Hills of South Dakota. The study is focused on the management of aspen for improved wildlife habitat, especially for ruffed grouse. Game, Fish and Parks personnel have also been involved in a study utilizing processed aspen as a feed component for livestock. Tests have shown that food rations containing processed aspen were superior to an alfalfa control ration in average daily gain and feed required per pound of gain (Kamstra et al.).

Knowledge of the location and size of aspen stands is absolutely necessary for the improved management of wildlife habitat and potential commercial harvesting of aspen as a livestock food source. Comparison of recent high altitude aerial photography and U.S. Forest Service type maps showed gross differences in the size and location of the aspen stands. It was very evident the U.S. Forest Service type maps were inaccurate (Schmer et al.).

Realizing the need for a better inventory of aspen the Department of Game, Fish and Parks and the Remote Sensing Institute initiated a cooperative project funded by NASA to utilize remote sensing techniques to produce an inventory of aspen and other forest cover types. Based on previous research funded by NASA it was determined that digital analysis of Landsat CCT data was the most effective and economic method for developing such a map.

STUDY AREA

The Black Hills encompass approximately 1.2 million hectares (3 million acres) located in western South Dakota and eastern Wyoming along the South Dakota-Wyoming border. The Black Hills are composed of deciduous and coniferous forest with intermingled grass meadows. The major conifer type is ponderosa pine (Pinus ponderosa) and the major deciduous type is aspen (Populus tremuloides).

Over the last 75 years the forest has become over stocked with conifer. This is due primarily to the prevention of forest fires and the limited logging activity. The conifers have overwhelmed many of the aspen stands and meadows. Consequently, there has been a severe loss of wildlife habitat.

REVIEW OF CLASSIFICATION PROCEDURE

The Landsat CCT data for the Black Hills were classified using a maximum likelihood decision rule classifier. Training sample data representing each of the land cover types were input to train the classifier. The land cover classes were pine, aspen, pine mixed, aspen mixed, grass and water. All land cover classes were based on a percentage of cover type. In order for a cell to be assigned to the pine or aspen category the dominant tree type had to be in excess of 80 percent. For the pine mixed and aspen mixed category the dominant type had to be in excess of 50 percent but less than 80 percent. A cell was considered to be grass if there was greater than 50 percent grass.

An attempt was made to classify water but some problems were encountered. Because of the topography in the Black Hills several shadows are prevelant in the data. Even after several attempts it was impossible to separate the water category from the areas of shadow. Consequently water was not identified from the CCT data. Instead, the water was interpreted from high altitude photography and inserted manually into the classified data. Areas of shadow not interpreted into the water category were arbitrarily placed in the pine category.

The performance of the classifier was determined by classifying and field checking three test areas of approximately 30,000 pixels (13,300 hectares) each. The verification procedures and results are discussed later. Final classification of the entire Black Hills was performed on approximately 2.5 million pixels (1.1 million hectares). The classified data were geometrically corrected using Digital Image Rectification System (DIRS) software developed by NASA. DIRS uses a system of ground control points (GCP's) to geometrically correct the data. Once corrected any point within the data can be converted to a Universal Transverse Mercator (UTM) coordinate point. This capability allows for easy location and extraction of specific areas of interest. The capability will be used extensively to extract quadrangle areas during final product production.

CLASSIFICATION ACCURACY

Knowledge of the accuracy is necessary for acceptance and use of the classified data. Therefore considerable time and effort was spent determining the accuracy of the data. Three methods were used to determine the accuracy. Each method involved the comparison of the 0.45 hectare (1.1 acre) data aggregated to 4.05 hectare (10 acres) cells with some form of ground truth. The format of the output map, which will be discussed more completely later, was a black and white photographic print scaled to 1:31,680. The scale is identical to the U.S. Forest Service type maps. Six grey levels were used on the output map to represent the classes. In addition the road network was superimposed onto the map to aid in locating specific areas of interest. The road network was obtained from USGS 7.5 minute quadrangles.

The first method used to test accuracy had a twofold purpose. First and foremost was to determine the accuracy of the data and secondly to determine the usefulness of the output map for locating specific areas of interest in the field. The procedure involved selecting three test sites each corresponding to the precise area encompassed on a USGS 7.5 minute guadrangle. The represented quadrangles were CROOKS TOWER, MT. RUSHMORE, and CUSTER. The areas are in the northern, eastern, and southern portions of the Black

Hills, respectively. These areas were selected to determine how well the classifier performed in different portions of the hills and where the predominant land cover type was different.

Ground truth data were collected for each of the three test sites by driving through the areas and coding the land cover type that was present along the road onto the output maps. The superimposed road network and other features were used to determine approximate location with respect to the output maps. Since the data were aggregated to 4.05 hectare (10 acre) cells it was attempted to determine the predominant land use for each cell adjacent to the road. The land cover types were determined by Game, Fish and Parks personnel based on the criteria used to define the categories. The ground truth information was coded directly onto the output maps.

Two problems were encountered using this procedure. First it was very difficult to determine actually where the boundaries or extents of a 4.05 hectare (10 acre) cells were from a position on the road. In turn it was somewhat difficult determining the predominant land cover type for the cell because of the terrain made if difficult to visualize an entire 4.05 hectare (10 acre) unit. The second problem was that because the ground truth data were coded onto the output maps there was obvious bias introduced in areas where it was difficult to determine the predominant land cover type. Whenever there was any indecision the end result was to code the ground truth to correspond to the output map unless of course the output map was obviously in error. Table 17 is a confusion matrix showing the statistical and mapping accuracy results.

The accuracy figures are extremely high in comparison to similar studies. The accuracy figures were equally high for both the CUSTER and MT. RUSHMORE quads. No data were obtained for the CROOKS TOWER quad because of heavy snow accumulation making all roads in the area impassible. The information acquired for the CUSTER and MT. RUSHMORE quads indicated that the classifier was performing equally in both areas. These areas are predominantly pine and grass with very little aspen. Therefore, it was necessary to collect ground truth data for the CROOKS TOWER quad, where aspen is dominant, at a later time when snow was not a problem.

The second ground truth data collection procedure was used to verify the accuracy of the data for the CROOKS TOWER quadrangle. In this procedure the data were collected without using the output map. Ground truth data were obtained by driving along roads and recording cover type which was determined by Game, Fish and Parks personnel. The ground truth data were coded onto a 4.05 hectare (10 acre) cell grid that was overlayed with the output map to determine the accuracy. Table 18 is a confusion matrix showing the results of the accuracy verification.

As can be seen the accuracy figures are significantly less than those obtained in the first method. There are two probably reasons

Table 17. Accuracy figures from field verification method #1.

CLASS	Grass	Pine	Pine Mixed	Aspen Mixed	Aspen	Total	Omissions No.	Omissions %	Mapping Accuracy %
Grass	*232	1	3	0	0	236	4	2	97
Pine	0	373	14	0	0	387	14	4	94
Pine Mixed	2	7	162	2	0	173	11	6	85
Aspen Mixed	2	0	0	40	0	42	2	5	89
Aspen	0	0	0	1	17	18	1	6	94
Total	236	381	179	43	17	856			
Commissions No.	4	8	17	3	0				
Commissions %	2	2	9	7	0				
Overall Classification Accuracy %	96								
Overall Mapping Accuracy %	93								

*Values represent number of 4.05 hectare (10 acre) cells

Table 18. Accuracy figures from field verification method #2

CLASS	Grass	Pine	Pine Mixed	Aspen Mixed	Aspen	Total	Omissions No.	Omissions %	Mapping Accuracy %
Grass	* 75	14	18	6	2	115	40	35	39
Pine	14	140	62	8	2	226	86	38	52
Pine mixed	28	23	40	17	6	114	74	65	20
Aspen mixed	22	6	9	9	7	53	44	83	10
Aspen	11	0	0	2	3	16	13	81	9
Total	150	183	129	42	20	524			
Commissions No.	75	43	89	33	17				
Commissions %	50	23	69	79	85				
Overall Classification Accuracy %	51								
Overall Mapping Accuracy %	26								

*Values represent number of 4.05 hectare (10 acre) cells

for this. First, the area is much more diverse which magnifies the errors that may be associated with not being able to locate precisely the area encompassed by a 4.05 hectare (10 acre) unit. Secondly, an output map wasn't present to bias any decisions made in difficult areas.

After comparison of the two previous accuracy verifications with results obtained in similar investigations it was determined that neither approach was acceptable. Therefore, a third more classic verification procedure was used. The procedure involved comparison of the classified data to a visual interpretation of high altitude color infrared aerial photography. The aerial photography was flown 8 days later than the Landsat overpass used in the computer analysis.

Photographic enlargements (scale = 1:24,000) of an area corresponding to the CROOKS TOWER quadrangle were interpreted by an experienced photo interpreter. Criteria for determining the individual land cover types were the same as those used in the machine classification. Polyginal boundaries were delineated for all cover types within the area. The photo interpretation was digitized into 4.05 hectare (10 acre) cells. Each cell was assigned to the category which was dominant. The gridded photo interpretation and machine classification were compared to determine the accuracy of the classification. The results are reported in Table 19.

Table 19. Accuracy figures from field verification method #3.

CLASS	Grass	Pine	Pine Mixed	Aspen Mixed	Aspen	Total	Omissions No.	Omissions %	Mapping Accuracy %
Grass	*303	16	27	28	5	379	76	20	56
Pine	57	1247	301	23	7	1635	388	24	65
Pine Mixed	62	254	471	106	13	906	435	48	36
Aspen Mixed	36	19	88	114	74	331	217	65	23
Aspen	7	1	3	8	34	53	19	36	22
Total	465	1537	890	279	133	3304			
Commissions No.	162	290	419	165	99				
Commissions %	35	19	47	59	74				
Overall Classification Accuracy %	66								
Overall Mapping Accuracy %	40								

*Values represent number of 4.05 hectare (10 acre) cells

C-2

As can be seen, the accuracy values are quite different than either of the first two methods. Comparison shows that the accuracy figures are more in line with results achieved by other investigators (Kalensky and Scherk, Heller et al.). This is understandable since the third verification procedure used is the widely accepted procedure.

Another representation of the accuracy of the data can be seen in Table 20. The aspen and aspen mixed category and the pine and pine mixed categories have been combined to show the accuracy of the data based on only three categories: pine, grass and aspen. The mapping and statistical accuracy increases significantly when similar categories are combined. This emphasizes the fact that much of the classification error is among similar categories.

OUTPUT PRODUCTS

Production of suitable output products is very important to overall acceptance and use of the data. Several alternatives were investigated to determine the most effective and efficient method available (Best et al.). The method that was selected involved utilizing a film recorder device.

The process involved aggregating the 0.45 hectare (1.1 acre) classified data into 4.05 hectare (10 acre) cells. The data were aggregated for two purposes. First, to reduce the amount of data handling. Second, because for management purposes 4.05 hectare (10 acres) was the smallest unit of interest.

Table 20. Accuracy figures for category combinations from field verification method #3.

CLASS	Grass	Pine	Aspen	Total	Omissions No.	Omissions %	Mapping Accuracy %
Grass	* 303	43	33	379	76	20	56
Pine	119	2273	149	2541	268	11	84
Aspen	43	111	230	384	154	40	41
Total	465	2427	412	3304			
Commissions No.	16	2	154	182			
Commissions %	35		6	44			
Overall Classification Accuracy	85						
Overall Mapping Accuracy	60						

*Values represent number of 4.05 hectare (10 acre) cells

Aggregation of the data to 4.05 hectare (10 acre) cells caused the loss of some information. In particular a narrow strip of any cover type may be completely eliminated. This is actually no problem because those areas generally have no management potential because they are in fact too small.

Film negatives of the data were produced using a film recorder device and black and white photographic prints made. Each map was enlarged to the scale 1:31,680 as requested by Game, Fish and Parks. Six grey levels were used to represent the land cover types. Existing road network, section lines, section number, quadrangle name and township and range numbers were superimposed onto the map using photographic processes. The information was obtained from the USGS 7.5 minute quadrangle maps. Figure 8 is an example output map.

Production of the maps for the entire Black Hills area has not been completed as yet. The procedure will involve producing a film negative of the 4.05 hectare (10 acre) data for the entire area. The areas corresponding to the quadrangles will be extracted from the negative using another film negative which serves as a mask. Each area will be enlarged to 1:31,680 and will include the map information included on the test area maps.

The film negative of the entire area will also be used to produce a small scale (1:500,000) map of aspen for inclusion in documents as well as for general information about the area.

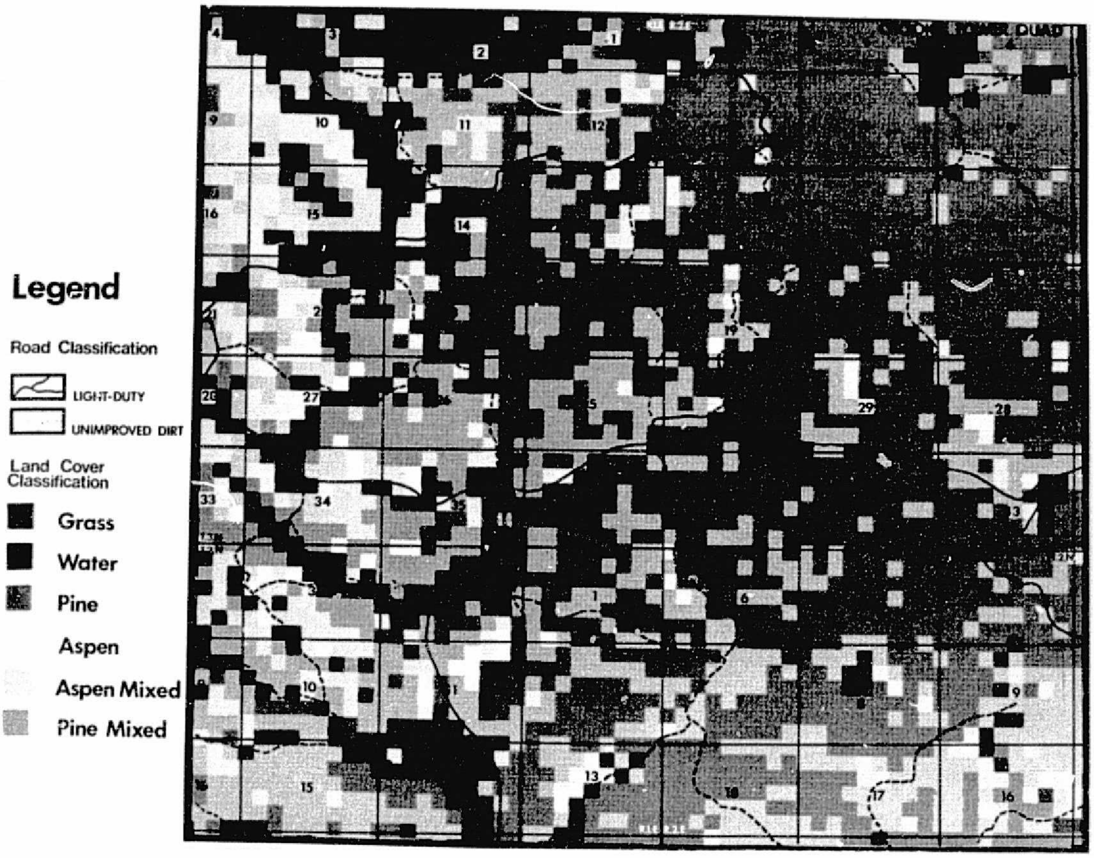


Figure 8. Sample black and white map produced using a film recorder system.

Statistical information will be produced for all areas of interest. This includes statistical information for each quadrangle area or any other governmental or physical area of interest.

RESULTS

An inventory of forest land cover types has been produced for the Black Hills of South Dakota. Although the location and size of aspen stands is of primary importance the inventory also provides information on the other forest cover types such as grass and pine. The statistical and mapping accuracy of the individual classes as well as overall accuracy are at acceptable levels for use in development of a management plan. Statistical and map information will be produced for the total area and for areas corresponding to the USGS 7.5 minute guadrangle series.

CONCLUSION

At the completion of this project the Department of Game, Fish and Parks will have in hand an up-to-date inventory of forest cover types within the Black Hills. The inventory will be used to assist in development of a management plan for aspen in the Black Hills. Through the use of statistical and map information areas will be designated for prescribed cutting of decadent aspen stands. It is hopeful that the rejuvenation of decadent stands will improve the quantity and quality of wildlife habitat especially for ruffed grouse.

The inventory will enable an assessment of the potential for commercial harvesting of aspen for livestock feed. The inventory will help determine the capital outlay necessary to begin commercial harvesting and in what portion of the Hills to begin harvesting.

The inventory provides base line data for the monitoring of changes over time. It will be possible to monitor the effects of aspen management, encroachment of pine on unmanaged areas of aspen and the encroachment of forest into the grasslands.

In general, the inventory will increase the overall understanding of the resources of the Black Hills from both an environmental and economic standpoint.

REFERENCES

Best, R.G., K.J. Dalsted, J.C. Eidenshink, F.A. Schmer, M.E. Wehde. Application of Remote Sensing Technology in S.D. to Assess Wildlife Habitat Change, Describe Meadnering Lakes, Improve Agricultural Censusing, Map Aspen, and Quantify Cell Selection Criteria for Spatial Data. Semi-Annual Report, July 1, 1977-December 31, 1977. (Submitted to NASA under contract no. NGL 42-003-007).

Heller, Robert C. 1975. Evaluation of ERTS-1 Data for Forest and Rangeland Surveys. Pacific Southwest Forest and Range Experiment Station, USDA Forest Service, Research Paper PSW-112/1975. Berkeley, California.

Kalensky, Z. and L.R. Scherk. 1975. Accuracy of Forest Mapping from Landsat Computer Compatible Tapes in Proceedings of the 10th International Symposium on Remote Sensing of the Environment. ERIM Ann Arbor, Michigan. pp. 1159-1167.

Kamstra, L.D., L.B. Embry, and M. Singh. 1975. Utilization of Aspen Tree Material as a Reminant Feed Component. South Dakota State University Experiment Station. Brookings, South Dakota.

Schmer, F.A., R.E. Isakson, and J.C. Eidenshink. 1977. Investigation of Remote Sensing Techniques as Input to Operational Resource Management. Type III Final Report. SDSU-RSI-77-12. Remote Sensing Institute, Brookings, South Dakota.

Quantification of Cell Selection Criteria for Spatial Data

by

M.E. Wehde

Introduction

The technological growth and development of digital hardware and systems software in the last decade has resulted in a broadened spectrum of data processing applications in spatial analysis. Digital processing offers analytical power and processing performance not achievable with alternative techniques, i.e. the storage, manipulation, and retrieval of geographically-based resource information.

The motivation for digital processing of map-oriented resource information is three-fold:

(1) Analytical power is available; multiple themes and dates of spatial inventory information can be analyzed for combinatorical, temporal and/or suitability analyses.

(2) Display of information is flexible; line printers, plotters, cathode-ray tube devices, and film recorders are typical display mediums. Furthermore, the scale of the output can be customized to the needs of the user.

(3) Inventory tabulations in spatial units can be made within themes and geographical division.

Although several approaches to geographical information systems exist - parcel, point, network, area, and grid - only the latter two, which have become known as polygon and cell systems respectively, actually maintain boundary information suitable for generation of spatial displays. {1,2}

The Remote Sensing Institute has chosen to develop AREAS, Area REsource Analysis System, as an economy-oriented cellular system. The cellular system was chosen because many forms of remote sensing data already exist in cellular form. Furthermore the implementation of processing functions is often more economical via the cellular rather than the polygon approach.

In AREAS digital image matrices are converted directly to the compact, sequential records. When reference maps or auxiliary maps are used as input data, a computer-generated grid on transparent mylar is used as an overlay guide in the hand coding process. The columns are coded for each row where data classification changes. These entries are followed by the identification of the new data class. The data storage program edits the coded data for inconsistencies and coding format errors as it converts to a byte compacted, variable length, blocked record data set for disk storage and retrieval. An error print out allows correction of the input before the next step. A compositing processor allows multiple themes to be overlaid by interleaving the change points and reporting the theme combinations which occur in the process. An interpretation routine allows assignment of integer interpretation numbers to any one or more of the codes in the data set. This routine also generates tabulations of spatial inventory for the raw input and interpreted output data sets. Finally, the primary output products are Calcomp drum plotter maps in one of several optional formats at a user specified scale.

The Remote Sensing Institute deals with state, national, and international resource managers, scientists, and government personnel have expressed interest in computerized resource data banks for geographical areas and themes of particular interest to them. In some cases the technology is transferred to these resource managers, while in others the Remote Sensing Institute provides assistance in generating the data base and producing the final products. In either situation, however, the initial decision is choice of cell size. Typically in either case, advice is sought from Remote Sensing Institute personnel on recommendations for cell size. Clearly, a system for coding, storing, processing, and displaying cellular resource information is not a complete package until systematic guidelines for cell size selection are included.

Objectives

The primary objective of the study is to inter-relate three factors (1) data set characteristics, (2) processing performance, and (3) application costs, in search of guidelines for cell size selection.

Intuitively decreasing cell size increases performance (mapping and inventory accuracies) but also increases operating costs. Derivation of performance versus cell-size relationships will be emphasized in this report; AREAS operating costs which are relevant to AREAS users only will not be documented here.

Procedure

The interaction of data set characteristics and processing performance is of general interest to all cell system users. To facilitate the study of the inter-relationship a sample map segment of 2 miles by 2 miles from a detailed soil survey was selected for intensive study. An area which had a moderate boundary density and a mixture of various sizes and shapes was selected. The intent was to avoid bias that might enter artificially created data or that might arise from data with regular shapes and/or sizes. The data map segment is shown in Figure 9.

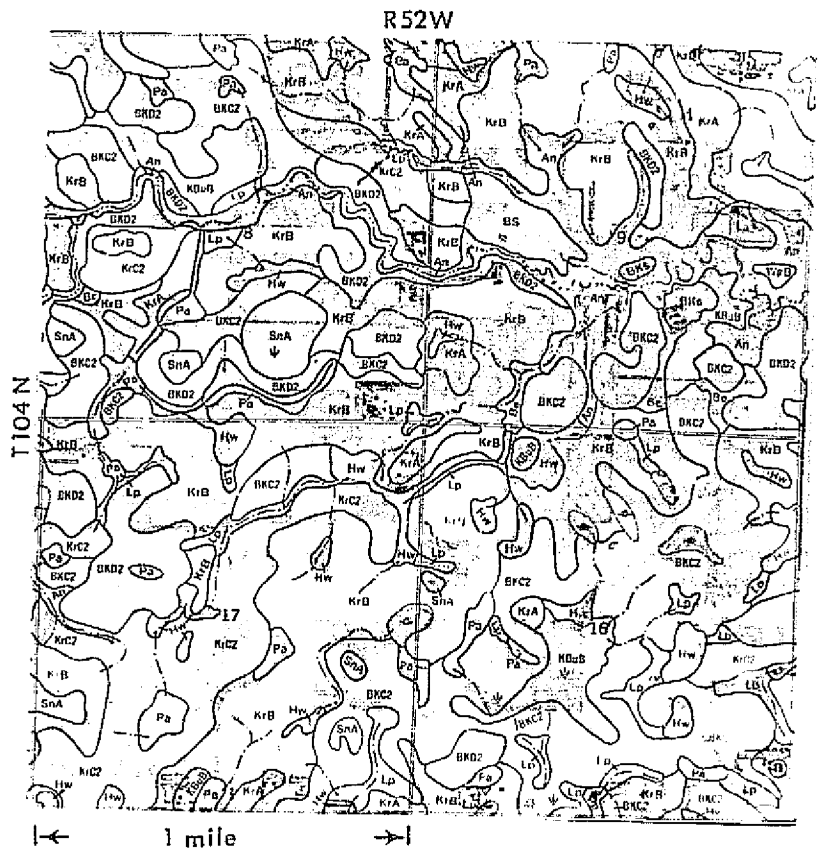


Figure 9. The original soils map of the intensive study data set. The four sections of soil association data are located in Minnehaha County, South Dakota. Mapping descriptions are located in the soil survey (3).

A very fine grid network or cell size was necessary in the study of cell size influence since the smallest cell size was considered optimum and used to evaluate the performance of larger cell sizes and aggregates. An initial size of cells smaller than the smallest separation of adjacent boundary lines was considered. In addition a constraint was imposed that the cell size be an even integer divisor of the commonly used 1.008, 4.032, and 16.128 ha (2.5, 10, 40 acre) cells. Under this constraint, the study of cellular resolutions would pass through cell sizes which users have employed and for which users have developed a conceptual feeling.

The smallest cell size selected was 0.007 ha (0.0174 acre) resulting in a map grid of 384 elements square. The data set produced from this cell size was considered the reference or "true" map for all analyses (Figure 10). The data set was generated by enlarging the detailed soil map, preparing a computer-drawn grid (0.007 ha), and manually encoding or digitizing the cell contents.

In order to analyze various cell size results, data sets of the same map at larger and larger cell sizes were required. Rather than attempt to manually digitize the soils map for each cell size a computer aggregation was used because of its superior time, cost, and dependability aspects.

An AREAS processing program was created to aggregate cells into larger and larger sizes. An integer is specified, for example four, and the change points of the data set are altered to align on column and row boundaries divisible by four. The algorithm sequentially

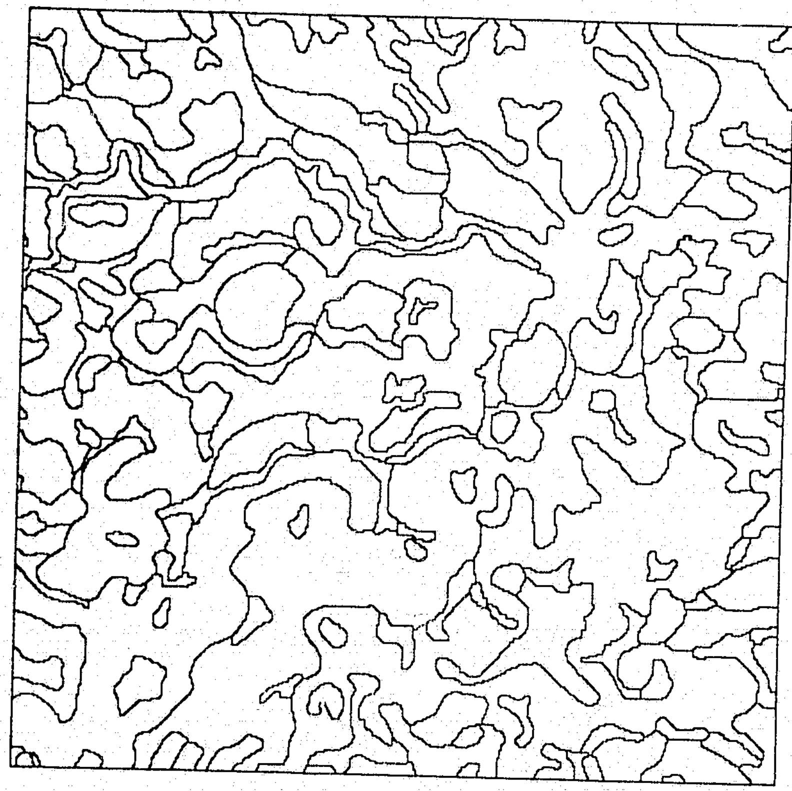


Figure 10. The computer map of the "true" data set at 0.007 ha (0.0174 acre) cellular grid. (Compare to Figure 9).

examines four by four arrays of cells to determine the dominant theme. Simple plurality applies and m-way ties are broken by an m sided coin toss -- a random number process. The output data set extents, rows, columns and codes, are maintained in agreement with the input data set.

Where a specified aggregation does not evenly divide the input rows and columns, the program handles code assignments based on whatever pixels remain at the end of the rows or the end of the data set. Continuing the example above, a fourteen by fourteen input data set would be aggregated by four with two pixel by four row coverage at the row ends, a two row by four pixel coverage at the data set end, and a two pixel by two row coverage at the ends of the last two rows of data. This approach allows any integer aggregation to be processed up to a program limit of sixty four.

However, for uneven aggregation division of the row-column dimensions of the input data set, the pixels processing at the right and bottom edge introduces artificial structure into any analysis of boundary spacing. Therefore aggregation was only applied for even divisors of the 384 by 384 base data set. The aggregations were 2,3,4,6,8,12,16, 24,32,48, and 64, with corresponding cell sizes of 0.028, 0.063, 0.112, 0.252, 0.448, 1.008, 1.792, 4.032, 7.168, 16.128, and 28.672 ha (0.069, 0.156, 0.278, 0.625, 1.111, 2.500, 4.444, 10.000, 17.778, 40.000, and 71.111 acres).

The eleven aggregations were also processed with an AREAS option which generates data sets with reduced row and column dimensions. Rows are selectively omitted and column designations are divided by the aggregation factor to achieve a reduction in the number of cells. This reduction agrees with the enlargement of the individual cell size. These reductions are useful for more rapid and economical plotting as well as for boundary structure analysis. The original data set, the eleven aggregated data sets and the eleven aggregated-reduced data sets comprise the twenty-three data sets of the data base for intensive study.

Performance from the product viewpoint was analyzed as mapping accuracy and inventory accuracy. The variable under study was cell size. The source data set with the finest resolution cell was the "true" map reference for mapping accuracy. An inventory tabulation of the spatial quantities of each map unit in the original data set became the "true" inventory reference for the study of inventory accuracy. The processing applied to each of the eleven data sets with larger cell size is diagrammed in Figure 11. The inventory and mapping performance was observed by map units and summarized for the entire map. Data characteristics were analyzed from an interboundary distance distribution standpoint.

Results

The apperance of the maps corresponding to the twelve resolutions in the data base can be compared in Figure 12. Note that the basic data set (upper left in Figure 12) and the aggregation by 2 to the immediate right are of such a fine cell size for the output scale being used that the cellular nature is not even apparent. Cellular representations can be as cosmetically pleasing as polygon approaches if the processing of many cells can be afforded. Corresponding mapping errors, the mismatch of areas when comparing maps of larger cell sizes to the reference, are displayed spatially in Figure 13.

The data flow process (Fig. 11) yields mapping error and tabulation error for each aggregate cell size hereafter termed the "resolution number" of the experiment. Figure 13 graphically shows the mapping results. The error rates are plotted in Figure 14 together with corresponding inventory errors. Omission and commission errors, on a

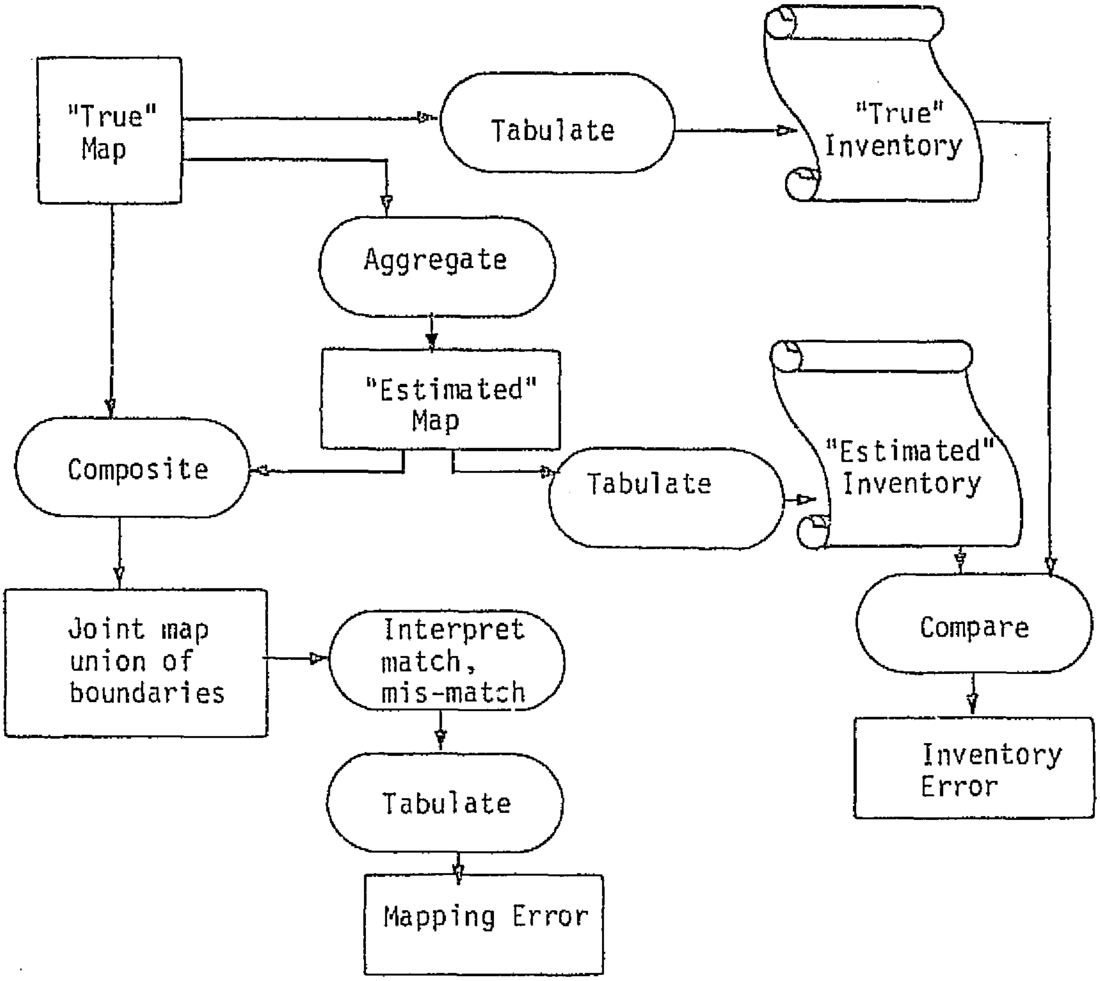


Figure 11. Performance evaluation processing diagram for an increased cell size.

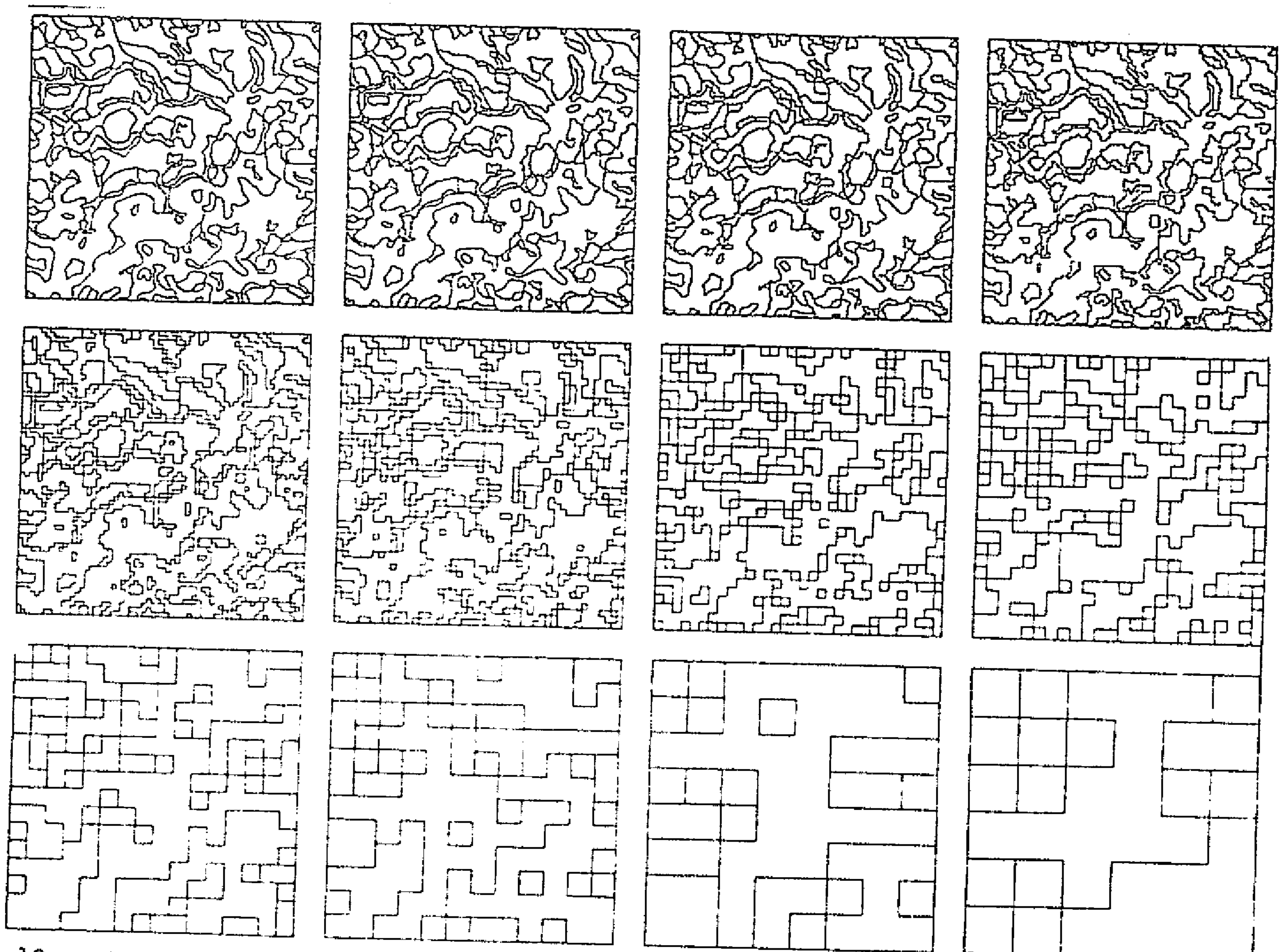


Figure 12. The twelve maps compared for mapping accuracy versus cell size. Cell sizes from top left to lower right are 0.007, 0.028, 0.063, 0.112, 0.252, 0.448, 1.008, 1.792, 4.032, 7.168, 16.128, and 28.672 hectares.

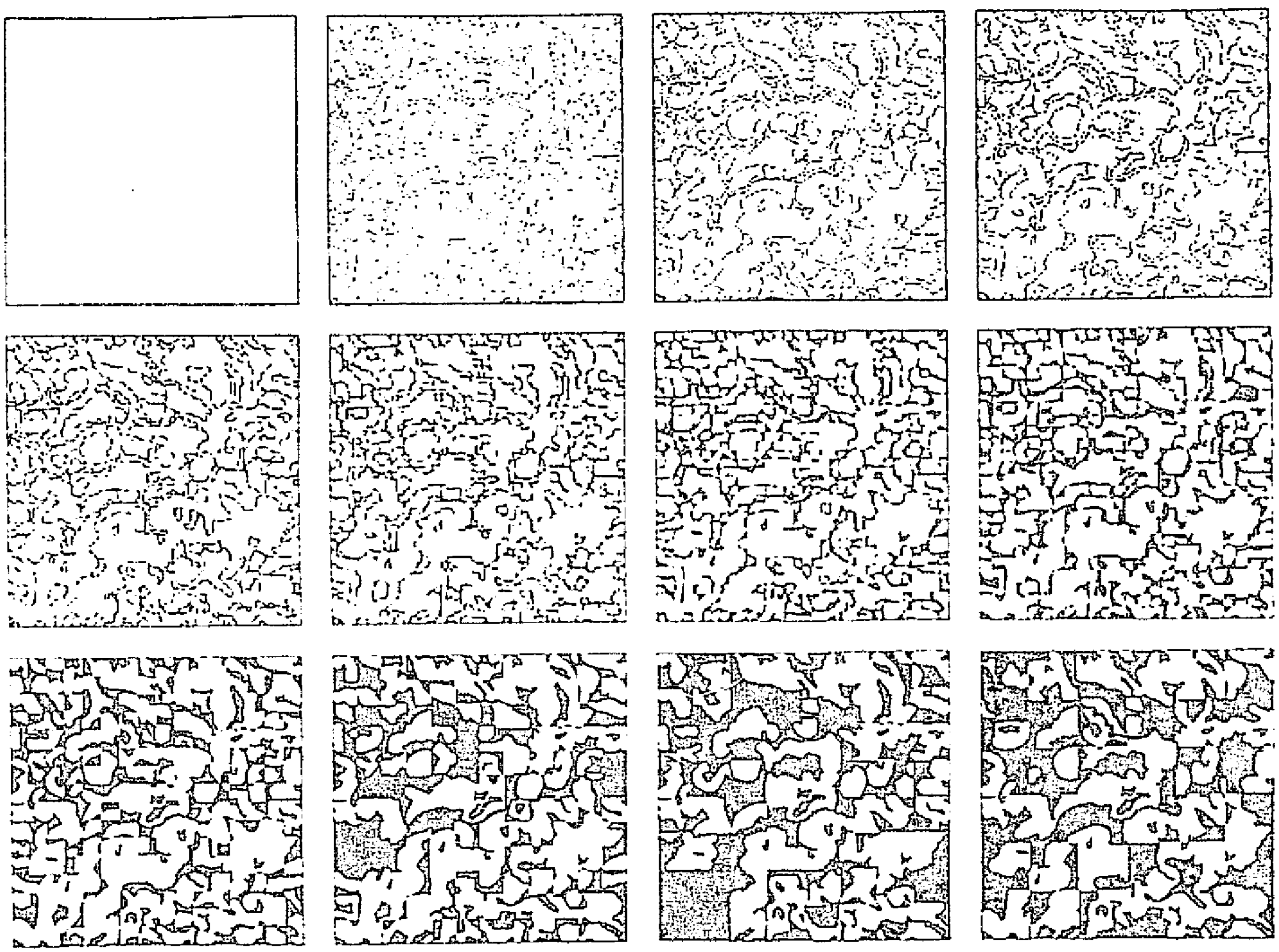


Figure 13. Representations of mapping error. Maps correspond to Figure 12.

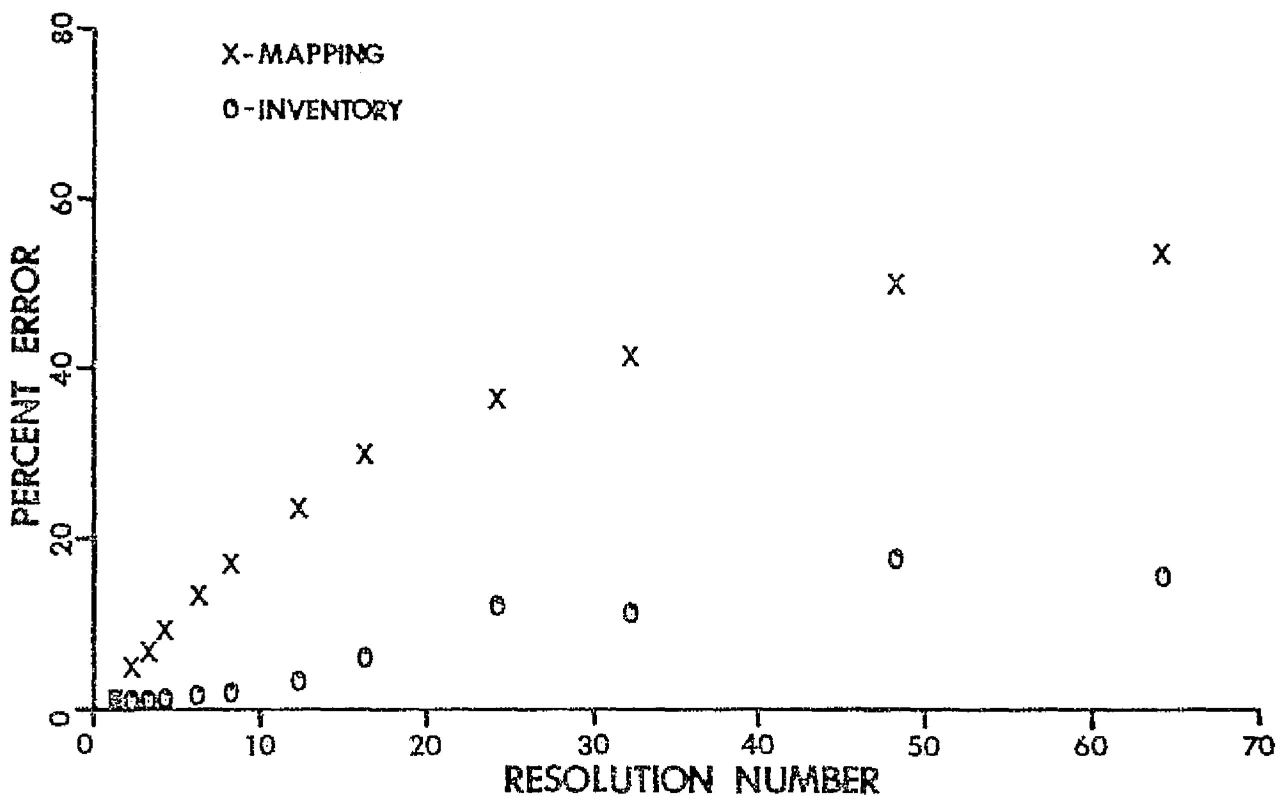


Figure 14. Mapping and inventory errors versus cell size expressed as a resolution number.

category basis within the map, partially cancel each other resulting in lower error rates in the inventory results. Summing over all categories in the map would yield a total inventory error of zero; therefore a root sum square was utilized to gauge inventory performance of the total map.

The test area contained moderate boundary density and a mixture of region shapes and sizes. It was selected with the intent of minimizing error variance. Regularly shaped regions of a size less than or equal to a given cell size can yield a wide range of errors depending on the spatial relationship of the cell to the region. These effects are random spatial relationships and should be averaged to yield a net interrelationship. Figure 14 shows that the averaging sought was achieved in as much as the relationship of error to changing cell size is non-random.

The pursuit of the defined objective not only requires evaluation of the system performance under various cell sizes as shown in Figure 14 but also an investigation of the characteristics of the data set at each resolution.

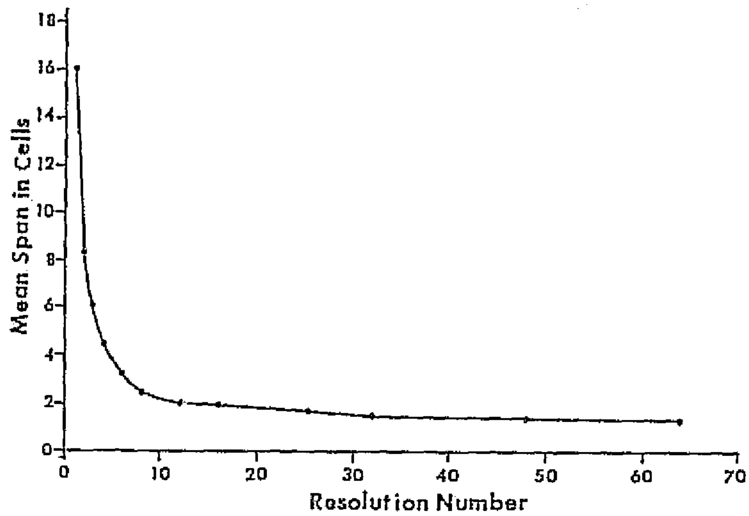
The "character" of the map data is the distribution of distances between boundaries. These distances may be called spans and the actual span distribution for original input maps is a continuous random variable. Cellularization quantizes continuous span into discrete levels. The quantization takes place both horizontally (along rows) and vertically (between rows).

A processing program within AREAS allows generation of horizontal, vertical, and total span distributions; it then calculates the mean for each distribution and performs a Kolmogorov-Smirnov test of equality (4) between the directional span distributions. Applying this program to each of the maps of Figure 12 yields a mean span for each distribution (i.e. each resolution of the map). The mean span in cells times the cell size involved (resolution number) is a mean interboundary distance in the units of the reference map cell.

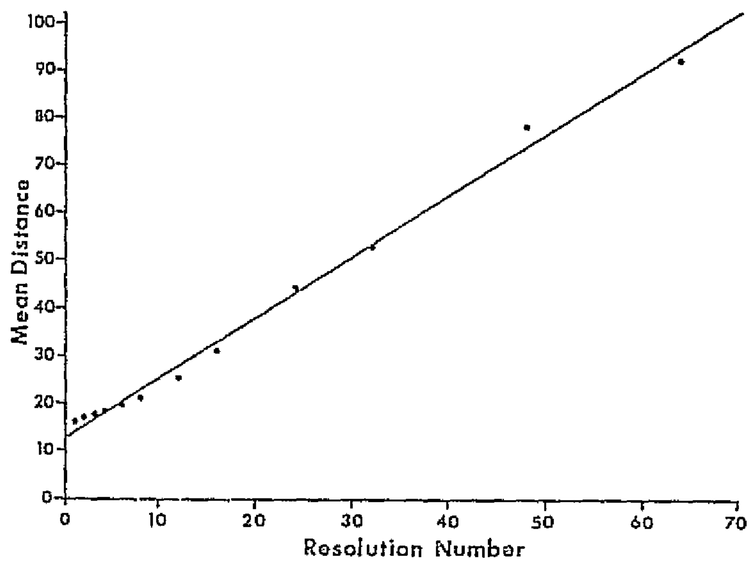
Figure 15a shows the mean span behavior versus corresponding resolutions and Figure YYb shows mean distance in terms of reference map cells versus changing resolutions. In 15a, an increasing cell size causes fewer cells between boundaries. If the cell size grows to equal the map size, the asymptotic limit on the mean span at that cell size would be one. Conversely at map resolution equal to one (the reference map) the mean span is the best estimate of the actual mean interboundary distance of the map.

In Figure 15b it is easier to guage the departure from an optimum. If a resolution shift does not vary the inter-relationship of boundaries, the mean span will decrease by the resolution factor as the cell size grows by that same factor. The product, i.e. the mean distance, would remain constant over all resolutions - a horizontal line with a zero slope. Although optimum this is an impossible occurrence which would correspond to 100% accuracy in representation of the actual map by any resolution cell. Whether linear or non-linear, the departure of data points from a horizontal line (Figure 15b) is an indication that the alteration of the boundary structure occurs in the process of changing cell size.

Since the relationship of mean distance to map resolution is linear, the map resolution axis of Figure 14 can be directly rescaled to mean distance. This establishes the direct relationship of a span distribution parameter - the mean - to the mapping and inventory errors. Mapping



a. The mean span for the distribution corresponding to the cellular resolution.



b. The mean distance between boundaries in the map corresponding to the cellular resolution.

Figure 15. Inter-boundary distance distribution is estimated by the span distribution. The behavior of the mean relates to changing cell size.

performance is a more rigid measure than inventory performance since correctness is required spatially on a cell by cell basis. Prediction of mapping performance will be discussed throughout the remainder of this report.

A mathematical relationship was not sought by standard statistical and/or curve fitting techniques since it was realized that such a relationship - even though clearly existent - would not likely hold for other map data sets. For example, maps with multimodal span distributions would have mean distances that change in an altogether different pattern with changing resolution. This would quite possibly make the equivalent of Figure 15b non-linear and completely alter the relationship of errors to the span distribution mean.

Since the span distribution characterizes the composition of map interboundary distances and the spans correspond directly to the compact, sequential geocoding process, estimation of mapping error at a certain cell size should be possible from the relative frequencies of the span distribution rather than the mean span. Various distributions including multimodal possibilities may arise from different map structures and still produce identical mean spans. The span distributions, however, will represent these differences in map structure.

The philosophical key to the relationship between mapping error and span distribution is as follows: The cumulative span distribution will provide information on the proportions of the spans which are less than or equal to a candidate cell size and the span distribution itself yields the proportion of spans which are not wholely divisible by a candidate cell size. These situations give rise to errors of omission of map spans and rearrangement of map boundaries, respectively. Furthermore these two mapping error sources effectively ignore the interplay between the sequentially adjacent spans. Such interplay at a candidate cell size can result in mapping error even within spans which independently would be evenly divided and representable by that cell size.

Experiments revealed that the cumulative span distribution alone relates quite well to mapping errors. Accounting for the relationships between cells and adjacent spans for spans greater than the cell size has not yet significantly improved prediction of mapping error.

Beginning with a simple example, first consider a closed region isolated in the spatial framework such that none of the surrounding space is of mapping interest. Circles, squares, ellipses, rectangles etc. are examples and one of each of these figures is created on a reference cellular base map for study. Running a sequence of cellular aggregations, all with respect to a single origin, soon reveals a widely varying error rate as the cells size approaches the region size. It was very apparent that chance spatial relationship of the aggregate cells to the region is a significant factor.

In a general map with a variety of region placements and orientations there is an average effect of changing cell size which is conceptually no different than averaging all possible orientations of an aggregate cell size with respect to a single region. This becomes the experiment. Aggregation of cells in 2 x 2, 3 x 3 etc. patterns were made and mapping error noted. For the two by two aggregation there were two horizontal positions and two vertical positions or four ways to aggregate. For an N by N aggregation there are N^2 ways. These were all observed in order to acquire data for an aggregation average. Cumulative span distributions were calculated for comparison. Results were as follows:

1. The average mapping error over aggregations for simple isolated regions followed a linear model from zero error at the reference resolution of one to a 100% error when the aggregate cell area became more than twice the region size. This is a very appealing and logical result.
2. The cumulative distribution approximated the linear model very well.

Consider next the effects of shape complexity. Concave-convex figures or donut-shaped regions will have mapping errors of higher magnitude than a linear model based simply on region area. The span distribution however should still adequately represent the nature of the data.

Finally consider the effect of mixed regions. In any one region, boundary mapping errors must necessarily affect the adjacent region. The modeling of mapping errors from a span distribution has this very viewpoint -- analyzing every span separate from the next adjacent span. One might anticipate the over estimation of error by a factor of two.

This was indeed the case and the inclusion of a weighting factor of one half on the cumulative span distribution for the data set of Figure 10 resulted in the predicted mapping error as compared to actual experimentally observed mapping error as shown in Figure 16.

Without attempting any distributional derivations, the direct relationship of the cumulative span distribution and mapping error is established in a general sense which is applicable to any data set. An implied assumption is made that the diversity of region locations in the map or the diversity of region shapes (or both) is effectively equivalent to a single region undergoing an "average positional cellularization." The meaning of equivalence, a measure of degree, and how critical the requirement may be for application of the model are unknowns requiring further study.

The significance of the existence of this relationship, however, is that resource maps may be sampled in an extremely fine grid or even continuous measure to yield an estimate of the span distribution from which mapping error versus cell size can be predicted. Thus the selection of cell size can be assisted with estimates of error rates prior to digitization of the map.

Conclusions and Recommendations

Conclusions of the study are as follows:

1. In a map of diversely shaped, sized and oriented regions, there is a well-behaved net relationship between cell size and mapping error.

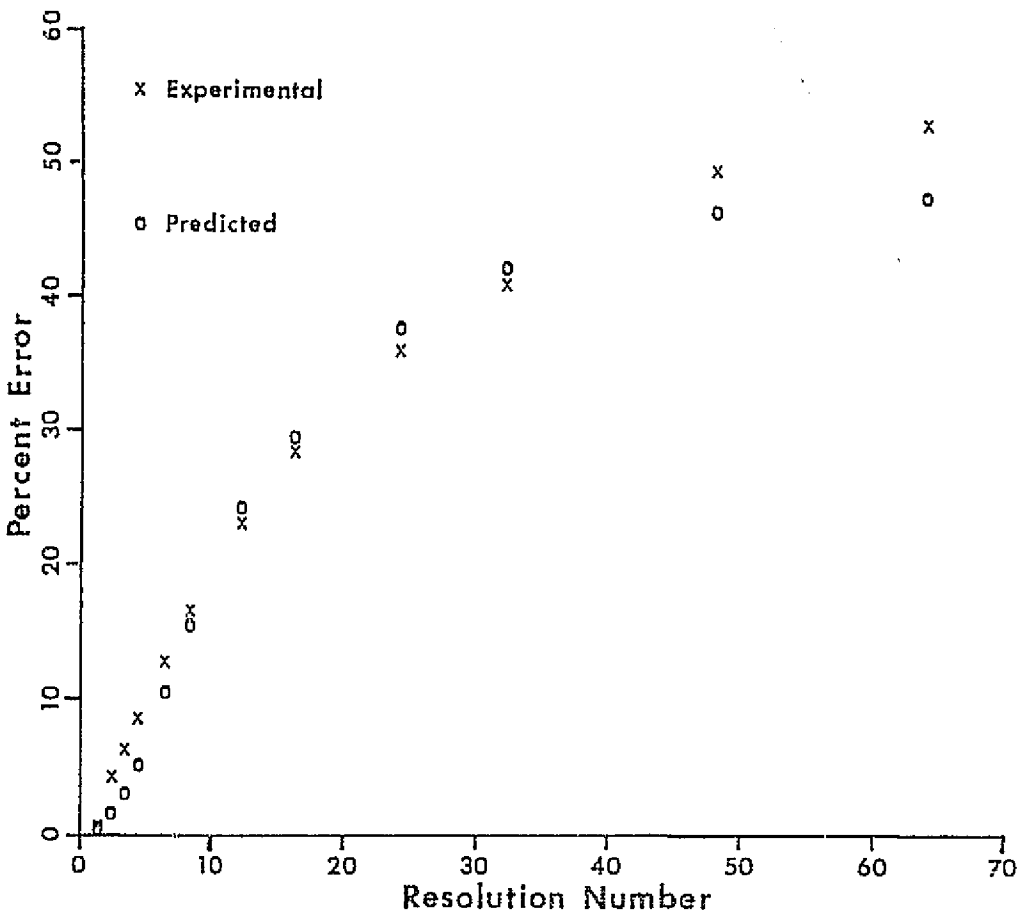


Figure 16. The mapping error experimentally observed as compared to the mapping error predicted by one-half the cumulative span distribution values for the finest resolution data set available.(closest estimate of actual inter-boundary distance distribution).

2. A continuous inter-boundary distance distribution characterizes a resource map.
3. A discrete span distribution in its cumulative form serves to estimate mapping error under various cellularizations of a resource map.
4. A scientific approach to selection of cell size is feasible.

Recommendations for further study and development are:

1. Define the sampling process to be used to analyze resource maps for an adequate estimate of the inter-boundary distance distribution.
2. Pursue an error prediction for composited maps based on the inter-boundary distance distributions of the input maps.
3. Analyze further the influence of the assumption that average aggregation positional mapping error is equivalent to the affect of aggregation over diverse region locations and shapes.
4. Analyze additional data sets for possible prediction of inventory errors from mapping errors or from a separate relationship to the interboundary distribution.

REFERENCES

{1} "IRIS - ILLINOIS RESOURCES INFORMATION SYSTEM" University of Illinois Center for Advanced Computations Feasibility Study - Final Report, Urbana, Illinois, 1972.

{2} Phillip A. McDonald and Jerry D. Lent, "MAPIT - A Computer Based Data Storage, Retrieval and Update System for the Wildland Manager" in Proceedings of the 38th Annual Meeting of American Society of Photogrammetry, Washington, D.C. March 12-17, pp. 370-397, 1972.

{3} Soil Survey Minnehaha County, South Dakota USDA-SCS in cooperation with South Dakota Agricultural Experiment Station, Map Sheet 1, 1964.

{4} Robert Winkler and William Hays STATISTICS PROBABILITY, INFERENCE AND DECISION, Second edition, Hoet, Reinhart and Winston, Chicago, p. 845-848, 1975.