Reprinted from

Eighth International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Crop Inventory and Monitoring

July 7-9, 1982

Proceedings

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

REGIONAL AQUIFER SYSTEM ASSESSMENT THROUGH LANDSAT DIGITAL IMAGE ANALYSIS

H.N. ANDERSON, M.W. GROSS, S.M. MANNING

Idaho Image Analysis Facilty
Boise, Idaho

I. ABSTRACT

The Idaho Department of Water Resources in conjunction with the U.S. Geological Survey was engaged in an analysis of the regional aquifer system of the Snake River Plain. As a major part of the study, an irrigated cropland inventory was performed.

To map irrigated and nonirrigated lands in the 15,600-square-mile area being analyzed, a combination of Landsat data, aerial photography, ground truth and irrigation district information was used to develop two spatially organized data sets. An irrigation water source data set, produced from irrigation district maps, was combined with the irrigated cropland distribution derived from Landsat and ground data to establish a land-cover classification of irrigated lands by water source.

All digital data were machine-processed using VICAR-IBIS software and an interactive image display system. Some 200 sample units, 2.5-km^2, were chosen using a stratified random sampling design to insure representative data and to quantify results. Training statistics for land-cover classification were developed using a maximum likelihood classifier in a modified clustering approach. A simple linear regression was performed to determine the relationship of Landsat data to ground data. The accuracy of Landsat crop-type discrimination was demonstrated in a contingency table analysis. The coefficient of determination between classified Landsat data and irrigated acreage was 0.89.

II. INTRODUCTION

The Water Resources Division of the U.S. Geological Survey is assessing the hydrology of the Snake River Plain regional aquifer system underlying most of the southern portion of the state of Idaho. An integral part of the assessment is an irrigated cropland inventory conducted by the Idaho Image Analysis Facility of the Idaho Department of Water Resources. The area inventoried includes approximately three-quarters of all irrigated lands in Idaho; in all, the aquifer underlies an area spanning 15,600 square miles from the Oregon border on the west to the Wyoming border on the east.

In addition to an irrigated lands inventory, it was the objective of this project to establish the source of irrigation water--from surface, ground or a combination--throughout the entire area. To produce a land-cover classification by water source, graphic information contained on county irrigation district maps was interfaced with the Landsat classification ground data to achieve the desired result: a map of the Snake River Plain designating irrigated lands by cover type and irrigation source.

III. STRATIFICATION

For this project, data was stratified for two primary reasons. First, it was the objective to insure a representative sample population by employing a stratified random sampling design; stratification provides a post-classification means for reducing confusion between land-cover types in irrigated and nonirrigated areas. Secondly, it aids in resolving conflicts due to variations in cropping practices, crop distribution, etc.

The initial stratification was prepared using six 1978 Landsat scenes. On the basis of visual interpretation, each scene was stratified on a mylar overlay to designate regions of obviously similar ground cover and agricultural practices. The mylar overlays were then digitized online creating files containing polygon labels and vector information in coordinate form. Corner points from each scene were registered to corresponding pixel row-and-column coordinates, allowing the vector data to be converted to raster-image form. Pixels within each polygon were homogeneously coded with a unique value and polygons were linked with their appropriate strata using polygon centroids. The result correlates on a one-to-one pixel basis directly with the Landsat scene from which it was constructed. For resolving class conflicts across strata boundaries, pixel values were logically altered.

IV. SAMPLING

Sample units were selected using a 2.5-km grid scaled to the stratification overlays (1:250,000). Grid cells were identified by row-column location and by stratum. Using a random number table for selection without replacement, over 200 sample units were chosen. The sample size was based on individual sub-basin analysis to estimate irrigated acreage to within plus-or-minus 5 percent at a 95 percent confidence level for each of the three sub-basins within the study area. Although sub-basin analysis was not mandatory, the sample size insured an adequate sample population to allow for possible conflicts in spectral responses between sub-basins (i.e. scenes) that could have necessitated individual analysis. The following equation was used to determine sample size per sub-basin:

$$n = \frac{t^2 \, CV^2 \, (1-r^2)}{AE^2}$$

where

- n = the number of sample units
- t = value from "student's t" distribution
- CV = coefficient of variation, or the standard deviation divided by the mean and multiplied by 100
- AE = allowable error, expressed as a percentage value
- r^2 = coefficient of determination between the prior information (Landsat classification) and criteria being measured (photo/ground observation)

Estimates of the expected coefficient of determination were based upon a similar study conducted in 1977 for the Pacific Northwest Regional Commission. The study resulted in a CV of 56 and an r^2 of 0.94. To compensate for sample size inadequacy experienced due to variation in data in the 1977 study, a r^2 of 0.85 was used for this project.

$$n = \frac{2^2 (56)^2 \; (1 - .85)}{5^2} = 76 \text{ sample units}$$

Multiplied by three and rounded, the figure yielded a sample size of approximately 200.

Once all sample units were located using the grid, their locations were plotted on USGS 7.5' quadrangle maps. Aerial photography was used to map field boundaries within the sample units. Field visits were then made to half of the units (100) to record data on method of irrigation, water source and crop type. From among those visited, 50 were digitized in a manner analagous to irrigation districts. The recorded field data was matched with the appropriate polygons and the information was processed by VICAR-IBIS to produce tabular summaries and digital maps of the sample units.

V. CLASSIFICATION OF LANDSAT DATA

A multispectral classification of the Landsat scenes was accomplished through the established procedures of spectral pattern recognition. With the aid of a computer, the analyst quantitatively evaluates the various spectral responses of each resolution element (pixel) in each of four wavelength bands. A raster image of 2983 scan lines and 3548 samples per line for each scene comprised a multispectral data set of over 40 million data points.

To analyze the spectral responses of various land-cover types, data was plotted numerically in four dimensions representing the four wavelength bands. Because the spectral response patterns for different land-cover types are not mutually exclusive, data points form elliptical clusters in certain portions of feature space in two dimensions. Through statistical analysis of the spatial relationships between data points in different dimensions and programmatic parameter control, clusters were developed to reflect specific land-cover types.

A modified clustering approach was used because it offered the necessary accuracy in the classification, given the limited availability of ground data. The basic premise of any clustering criterion is to minimize the distance between points within a cluster while maximizing the separation between clusters. The separability between clusters indicates the quality of data point assignments to clusters.

From among the 200 sample units, training sites were chosen to develop training statistics. Within each Landsat scene, approximately eight agricultural and seven wildland training sites were selected. To account for spectral variation due to factors such as crop maturity, soil type and topography, user-specified parameters controlled the maximum number of clusters generated by the clustering algorithm. For sites consisting of a large number of fields with high variability, the maximum number of clusters was defaulted to 35. Sites consisting of more homogeneous cover types with low variability generally produced fewer clusters. The cluster separability (Delta) value was set to 0.72.

After training site data were clustered, sample units were classified using a Gaussian maximum likelihood classifier to form spectral classes (land-cover types). Probability density functions were developed using the mean vector and covariance matrices for each cluster. These functions calculate the probability of a pixel belonging to a specific spectral class and assign it to that class for which the highest probability exists.

Once all clusters were assigned to their respective spectral classes, the number of training clusters within a scene was reduced from 250-300 to 50-60 clusters representing all cover types within that scene for inclusion in the master statistics file. Clusters were then inspected to insure that within-class separability was at least 0.45 and that the separability between classes was at least 0.60. These separabilities verified the inclusion of distinctly separable clusters within unique spectral classes.

All pixels within a scene were classified using a master statistics file. The resulting classified data set was geometrically rectified and recorded on magnetic tape. The final image was verified on the video image display device. Once errors were resolved, the final output was created.

VI. RESULTS

Two analytical steps were taken to test the reliability of the data once it was classified. First, a contingency table analysis demonstrated the crop-type discrepancies between ground data and Landsat data by indicating the number of pixels that were classified in each of four crop types: row crops, small grains, alfalfa-hay and pasture. Secondly, a simple linear regression was performed to compare actual irrigated acreage on the ground with the irrigated acreage recorded by Landsat. The overall coefficient of determination (r^2) for irrigated cropland of the entire study area was 0.89. Strata consisting of small areas of widely divergent land-cover types that were not clearly differentiated in the analysis contributed to lowering the overall accuracy of the land-cover classification. Nonetheless, results satisfied project goals and provided cost-effective, baseline data on the Snake River Plain to the state and participating agencies.

Mr. Anderson is currently serving as Director of the Idaho Image Analysis Facility (IIAF) within the Idaho Department of Water Resources. IIAF is the designated leader for promotion and applications of remotely sensed data for the State of Idaho. Mr. Anderson has more than six years' experience in the use of Remote Sensing for resource inventory and monitoring. He has a Bachelor Degree and a Master of Science Degree in Forestry Resources, with specialties in Remote Sensing, from the University of Idaho.

Mark Gross is a Senior Water Resource Analyst for the Idaho Department of Water Resources. His responsibilities include user training and applications with existing digital image processing software (VICAR/IBIS and System 511), software maintenance and support, and participation in the development of additional image analysis and geographic information system capabilities. Prior to joining the Department of Water Resources, Mr. Gross worked as a geologist in geothermal exploration and applications. Mr. Gross received his B.S. in Geology from Boise State University.

Stephen Manning is a Water Resource Analyst with the Idaho Department of Water Resources. His responsibilities with the department include the applications and development of digital image processing and spatial information systems capabilities for the Idaho Image Analysis Facility within the department. Prior to joining the Department of Water Resources, Mr. Manning worked for more than five years in the development of quantum detectors and electro-optical scanners. Mr. Manning received his B.S. in an interdisciplinary program of remote sensing applications in geology and geography from the University of Idaho.