

Conference Report

R. A. MEAD
School of Forestry and Wildlife Resources
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061
JOHN SZAJGIN
Technicolor Graphic Services, Inc.
EROS Data Center
Sioux Falls, SD 57198

Landsat Classification Accuracy Assessment Procedures*

An Account of a National Working Conference

INTRODUCTION

SINCE LANDSAT DATA first became available, many Landsat scenes have been digitally analyzed to classify land cover. These classifications are not without error and have been subject to close scrutiny by critics and potential users. However, methods for describing and quantifying classification errors have largely been developed on an ad hoc basis. Furthermore, the lack of standardized methods based on sound statistical theory has spurred many researchers to express concern. Thus, a conference addressing Landsat

statistical analysis techniques, (3) the associated costs and logistical problems, and (4) the value of the accuracy data to the remote sensing specialist and the resource manager. The conference was held at the Earth Resources Observation System (EROS) Data Center, Sioux Falls, South Dakota, on 12-14 November 1980. The conference focused on the following objectives:

- Determine the state-of-the-art of accuracy assessment procedures,
- Provide a forum for exchange of ideas,
- Identify research needs and recommend the ap-

ABSTRACT: *A working conference was held in Sioux Falls, South Dakota, 12-14 November, 1980 dealing with Landsat classification Accuracy Assessment Procedures. Thirteen formal presentations were made on three general topics: (1) sampling procedures, (2) statistical analysis techniques, and (3) examples of projects which included accuracy assessment and the associated costs, logistical problems, and value of the accuracy data to the remote sensing specialist and the resource manager. Nearly twenty conference attendees participated in two discussion sessions addressing various issues associated with accuracy assessment. This paper presents an account of the accomplishments of the conference.*

classification accuracy assessment procedures seemed appropriate.

Nearly 20 scientists from across the country, who had experience with Landsat classification accuracy assessment procedures, were invited to attend a three-day working conference consisting of formal presentations as well as small-group discussions.

Thirteen formal presentations were made on four general topics: (1) sampling procedures, (2)

* To be printed in the proceedings of the Symposium on Machine Processing of Remotely Sensed Data, held at Purdue University in June 1981.

proaches that should be taken to improve accuracy assessment procedures, and

- Publish a comprehensive proceedings of the conference and prepare a paper summarizing the discussions.

The first three objectives were accomplished during the conference. Preparation of the conference proceedings is currently underway, and this paper summarizes the major points of discussion.

It is difficult to summarize the full content of the discussions which took place. The intent is to highlight the issues and ideas which were repeatedly raised or which generated considerable

enthusiasm. A consensus was not necessarily reached on the items which follow. In some cases the point or issue is briefly identified, and in others, more complete explanations are given.

DISCUSSION TOPICS

- Topographic mapping procedures include routine evaluations for compliance with well-defined accuracy standards, and the accuracy attainable under specific conditions (terrain characteristics, mapping equipment used, and type of aerial photographs) are well known. However, national standards for reporting thematic map accuracy (such as those produced from digital classification of Landsat data) have not been established. Potential users of Landsat classifications often do not know the relative accuracies that are achievable in identifying various land-cover types. These relative accuracies have not been fully determined. Furthermore, no government agency is known to have published standards for expressing accuracy. Such standards should be established, and contractors should be required to utilize them. Standard methods for reporting accuracy will become more vital as these classifications become inputs for geobased information systems.

- There are two major types of accuracy assessment procedures: site-specific and non-site-specific. Non-site-specific accuracy is usually expressed as the similarity between the total number of acres in each land-cover type as determined by a Landsat classification compared to the corresponding acreage determined from measurements in the field or from photointerpretations. The non-site-specific method compares only total acreages without regard to location. Site-specific accuracy, however, considers the spatial nature of the data. That is, two spatially defined data sets (one being "ground truth") are registered and compared for the amount of agreement. Such comparisons can be made on a polygon, grid cell, or point basis. These comparisons result in a matrix showing the quantity of omission and commission errors. If properly conducted, the site-specific approach provides a more rigorous and more informative appraisal of a map product. This approach may not be warranted when spatial arrangements are not critical. For example, when only acreage proportions by type are of principal concern.

- Landsat classification accuracy assessments are often made with very inadequate reference data (that is, maps, photointerpretations, or actual visits to the field). These reference data should be distributed throughout the scene in such a way that all cover types, as well as zones of transition between the various cover types, are adequately represented. Furthermore, the time of reference data acquisition is an important consideration.

The use of training set data for accuracy assessment results in a biased and usually inflated estimate of accuracy. The amount of bias depends upon how well the training data represent the variability present in the scene. In some cases, this approach may be adequate for making intermediate estimates to aid in the classification process. However, final evaluation of classification accuracy should be accomplished using an independent sample.

The cost of an independent accuracy assessment can be minimized by collecting the necessary accuracy assessment data simultaneously with the training data. The data should be set aside during the classification process and used later to provide an independent estimate of accuracy. In this way, all necessary field data are collected during a single field effort.

When interpretations from aerial photographs are used as reference data in assessing classification accuracy, the photointerpretation may not be perfect. Therefore, ground data may be necessary to verify the adequacy of the photointerpretation data.

- When error matrices are developed between classification results and reference data, consideration must be given to the means for selecting the sample. Factors, such as the number of categories classified, the proportion of pixels assigned to each category, and the spatial diversity of the landscape, interact and affect decisions concerning sample size and method of allocation. Also, the cost of field data collection, the rigor of the accuracy evaluation desired, and the relative importance of each land-cover class impact the entire process.

- Numerous statistical techniques need to be evaluated for their utility in analysis of accuracy data. Those particularly well suited for this type of data should be identified, and their application to this work documented.

- One should not lose track of the difference between the usefulness of a specific product and its estimated accuracy. A quantitative accuracy assessment results in a numerical summary which may or may not represent the usefulness of the product or how well it compares with map products which were previously available.

- Further research is needed to determine the most appropriate sample designs for assessing the accuracy of classification results for landscapes of varying spatial diversity. In this regard, the advantages and disadvantages of cluster sampling should be investigated.

- A list of computer programs presently available for sampling classification results for assessing accuracy should be compiled. Development of additional computer programs may be needed to facilitate rapid accuracy assessments.

- There was a general consensus among those

participating in the conference that the costs and logistics required for conducting accuracy assessments are often prohibitive. Better estimates of these factors need to be published, and faster, less expensive methods that suit user requirements should be developed.

- Given the current level of knowledge, a general set of accuracy assessment guidelines should be written. These guidelines should be flexible because of the wide range of circumstances associated with the varying objectives of classification. For this reason, several authors would be needed to document adequately the many diverse aspects of assessing classification accuracy.

SUMMARY

Many issues were discussed and debated by conference participants. Topics for further research were identified, and major topics of discussion were summarized. A comprehensive report on the proceedings is being prepared in which state-of-the-art accuracy assessment procedures will be documented.

The participants recommended that a working group be established to write a manual or "guide book" on accuracy assessment procedures. This group could possibly be an ad hoc committee within the American Society of Photogrammetry.

New Sustaining Members

Orhan's Reproductions & Photomapping Ltd.

907-9th Avenue S.W., Calgary, Alberta T2P 1L3, Canada; (403) 265-7514

SINCE ITS FOUNDING in 1959-1960, Orhan's Reproductions & Photomapping Ltd., from a modest beginning has developed precision techniques and equipment geared to the fast pace of the industry.

The Company produces the following comprehensive services: Ground surveys, aerial surveys, conventional and computer assisted photogrammetry (digital mapping), analytical aerial triangulation, orthophoto mapping (color, black & white), terrestrial photogrammetry, conventional and automated cartography, U.T.M. metric base and well maps (from 40°00 to 84°00 latitude for North America at various scales for the oil and related industries), microfilming and microfiche, full reproductions, graphic arts, offset printing, general design and manufacturing, etc.

Our responsibilities are to provide to the client total in house capabilities or one stop service. Therefore, we have assembled the most sophisticated equipment and software in our new modern building (34,500 sq ft) facilities in Calgary, Alberta, which is the Oil capital of Canada.

SATLAB inc.

606 West 29th Place, Lawrence, KS 66044; (913) 749-0213

SATLAB INC. is a consulting firm specializing in the processing and interpretation of space- and air-borne remote sensing data for Earth resources applications. Our approach to image processing has been heavily influenced by our experience in various geology, vegetation, and land inventory projects. Our belief that visual interpretation techniques continue to play an important part in remote sensing applications is reflected in the high quality of our standard images and in unique, application-optimized image products which encourage the interpreter to utilize fully the skills which he or she has developed in the traditional field of photo interpretation. We have succeeded in translating multispectral, multi-temporal, and multisensor information into easily understandable formats by incorporating intensity, hue and saturation (IHS) transforms and stereoscopic presentations as additional tools for analysis.

As a small, independent company we maintain the flexibility to discuss, develop, and apply techniques to meet project-defined user requirements.

For additional information, contact Dr. Janét E. Bare, President, 606 West 29th Place, Lawrence, KS 66044; (913) 749-0213.