Reprinted from

Seventh International Symposium

Machine Processing of

Remotely Sensed Data

with special emphasis on

Range, Forest and Wetlands Assessment

June 23 - 26, 1981

Proceedings

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

A CASE FOR STANDARDIZED TEST SITES

CHARLES E. OLSON, JR.

School of Natural Resources
University of Michigan
Ann Arbor, Michigan

I. ABSTRACT

Continued progress in remote sensing can be fostered by more standardized test procedures. Standardized test sites, including sites with totally enumerated populations, have particular advantages in testing sampling and modeling methods. Due to the multitude of purposes for using remote sensor data, it is unlikely that any single test site can meet needs of all users.

II. INTRODUCTION

Remote sensing is coming of age. Having passed through periods of widespread skepticism and unwarranted optimism, remote sensing is finding a reasonably realistic middle-ground based on proven capabilities and proven limitations. The maturation has been rapid, and continues at a rapid pace.

Within the USA, relatively free access to remote sensor data has been at least as important a stimulus as the amount of federal dollars invested. All with an interest in trying, have had an opportunity to do their thing. Many departed from accepted methodology, most because they didn't know any better. Of these, many failed, or had to learn the hard way what others had learned before them. A few succeeded, however, proving previously accepted methodology lacking, if not actually wrong.

If progress in remote sensing has been stimulated by easy access to data and the opportunity for all of us to make our own mistakes, why is there continuing interest in standardization? Could standardization stifle progress? It is my belief that standardization is a natural evolution stemming from the rapid growth of the remote sensing field. Properly handled, standardization can foster progress. If I am correct, there must be something lacking in our present test sites and test site procedures.

III. OBJECTIVES

The primary objective of this paper is to identify some of the reasons why standardized test sites seem to be needed. The primary objective of standardization test sites, as I see them, is to provide a basis for realistic evaluation of the utility of various techniques. In short, how do we determine which techniques actually work, or work best?

There is seldom a simple answer to the question: "Does it work?" Sometimes that question means: Can results be duplicated by others? At other times it means: Are the results accurate? or, Is the method cost-effective? With all due respect to the statisticians and modelers whose contributions to remote sensing progress have been immense, statistics and models are based on assumptions and these assumptions should be tested when feasible to do so. Testing of assumptions, as well as verification of methods and models, can be a function of standardized test sites.

IV. WHAT IS A TEST SITE?

Many investigators have examined remote sensor data of specific areas, made interpretations based on those data, and some have gone further and actually field-checked the site to determine whether or not their interpretations were correct. Any location used in this way can be considered a test site, whether or not actually visited by the investigator or the investigator's agents. One form of remote sensor data may be used to access the accuracy of interpretations from another form of remote sensor data, as when conventional aerial photographs are used as the data source for determining if thematic maps from LANDSAT data adequately represent the terrain mapped. Increasing use of such surrogates for field-checking have prompted some to suggest replacing the term "ground truth" with "reference data." Thus, any site for which reference data exists may be a test site. Of the multiplicity of test sites fitting this definition, few are documented well enough to permit use by others. Sites for which appropriate reference data are available to any user are of particular value in some kinds of testing. With the rapid evolution of different software packages for processing digital remote sensor data, replication of experiments at a single test site with different software would help determine which provide most acceptable results.

A. STANDARDIZED TEST SITES

The term "standardized test site" is used in this paper to denote a test site for which sufficient reference data are available to all users to permit multiple, and comparative, tests at the same site. The term has no magic, and no attempt to coin a new bit of jargon is intended. Many such sites seem to exist. Some of these were established during the extensive period of investigations in the 1960s and 1970s which preceded the launching of LANDSAT-1. Others have come into being more recently. Yet, I know of no investigators who have published complete descriptions of test sites in a form usable by others. I assume, however, that there may be some; but, have any of these site records been maintained so as to provide opportunity for continued tests at the same site? Wouldn't it be nice to have a site,or a series of sites, where new techniques could be tested against an established yardstick?

B. STANDARDIZING TEST SITES

If standardized test sites are accepted as a desirable concept, what is to be standardized, and how it is to be done, must be determined. Because data requirements vary from user to user, it is unlikely that any single set of specifications can be developed to meet all needs at reasonable cost. Few investigators can afford to gather all of the data needed by others. This does not mean that standardized test sites must be centralized and institutionalized, although some well documented test sites might be worth widespread support. Rather, the need to keep standardization within reasonable cost points to the need for integrating local and regional/national sites into a single network. Such integration could be started on a voluntary basis now, providing those in charge of existing test sites are willing to share their baseline data. This will require an investment, a cost. Is this cost justified, and by whom should it be borne?

In many cases data available for existing test sites are so fragmentary and time dependent, that the cost of making the limited baseline data available cannot be justified. In others, available data are not in an appropriate form to permit ready dissemination. Few agricultural experiment stations have a centralized record of all of the treatments applied by all of the investigators using any experimental field, and the records of such stations may be the best available. Even in those small number of cases where baseline data do exist in appropriate form, the size of the area for which documentation exists is usually miniscule when viewed in LANDSAT terms. Yet, these data points can be important if a way can be found to make them more generally available, for all remote sensing is not based on LANDSAT. I offer no solution to the obvious problem, but believe a solution will foster continued rapid development of remote sensing applications. Almost certainly, a central clearing house of test site information will be needed. Possibly a new subject descriptor within existing abstracting services can meet this need.

C. LARGE-SIZE TEST SITES

Test sites to be meaningful in LANDSAT terms must be large. Because they must be large they present special problems in standardization. Is it practical to enumerate all of the pixels of a LANDSAT scene? If such an enumeration were practical could it be maintained and updated often enough to be of continued utility? If the answers to both of these questions are affirmative, what might be gained from 100 percent enumeration of a LANDSAT scene?

Many large area inventories seem to require sampling, actual measurement of a small sub-set of the total population, and inferring population parameters from this sample. Statistical techniques for handling the mathematics involved are well developed, and development continues as new needs are encountered. When working with samples and inferring population parameters from these samples, we never really know the actual value of the population parameter. To overcome this, we claim a value within certain error bounds. Such claims are based upon assumptions concerning the mathematical distribution of the parameter measured, the detectability of the thing to be measured, the independence of the observations making up the sample, the representativeness of the sample to the population, and others. Repeatability of measurements, or estimates, is no guarantee those measurements or estimates are correct. Biases may exist which provide repeatable results that are wrong. The only way to be sure a sampling system provides a good estimate of a population parameter is to know the population parameter. Especially when designing new sampling schemes, an accurate indication of the efficiency of the population estimator is desirable. Work at The University of Michigan with totally enumerated populations of forest data for a three-county area have shown that some statistical techniques are superior to others, and that the most cost-effective technique is not always the most obvious. We hope to expand this effort to include enumeration of all of the pixels of a LANDSAT scene to provide a site for testing statistical sampling schemes for world-wide inventories. Some of the data are already in hand, and availability of the Michigan state-wide, 1:24,000 color-infrared aerial photography of 1978-79 makes preparation of such a baseline data set feasible.

If we complete this undertaking, would such a test site be useful to others? Preliminary indications are positive, even if the data base is not updated. Such a test site would provide a real population with the natural variability of real terrains altered by man. With population parameters known, a wide variety of performance testing would be possible and a yardstick would be available against which different techniques could be evaluated.

At the present time, we expect to record land cover/use data to Level IV of the Michigan Land Cover/Use Classification System whenever possible, but may be forced to stop at Level II except in forest areas.

Still to be determined is the format in which the baseline data will be stored. A pixel-by-pixel format has some unique advantages for work with a single scene, but certain liabilities should we elect to correlate subsequent scenes for change detection.

A large-sized, enumerated test site, such as described above, would certainly have potential utility to some outside of Michigan. Should such a test site be "standardized" and maintained for use by others? Once established, should such a test site be maintained? Can it be maintained? We think the answer to these questions is, yes. The proposed LANDSAT scene includes three sites of continuing research interest on other projects. These form a base from which the additional monitoring work can be conducted, at least on a periodic basis. Only this active work on other prjects makes it possible for us to consider maintaining such a data set on a long term basis.

This proposed test site would not meet all needs of all users. Possibly, similar test sites should be developed in other areas to meet other user needs. This test site concept has been identified, here, as a means of stimulating additional thoughts from those present at this meeting. If I have only begun a discussion that will continue beyond this meeting, I will have accomplished what I set out to do.

V. CONCLUDING REMARKS

In this brief paper, some of the reasons for continuing interest in standardized test sites for use with remote sensor data have been identified. These reasons stem from the fact that remote sensing, as a field, is coming of age. Part of the coming-of-age process involves calibrating the work we do against some meaningful yardstick. Standardized test sites provide one way to accomplish the needed calibration.

Charles E. Olson, Jr. Professor Olson received his B.S.F. and Ph.D. degrees from The University of Michigan in 1952 and 1969, respectively. In between, he received his M.F. degree from the University of Minnesota in 1953, served three years of active duty with the U.S. Navy as an image interpreter (1953-56), and spent six years with the University of Illinois as a Research Forester, Assistant Extension Forester, and Assistant Professor of Forestry. He joined The University of Michigan faculty in 1963, was named Associate Professor in 1969, and Professor of Natural Resources in 1973. He has served as Graduate Chairman (1972-74) and Dean (1974-75) of the School of Natural Resources at The University of Michigan and been the Director of the School of Natural Resources Remote Sensing Laboratory since 1972.