Dr. J. L. van Genderen
Fairey Surveys Limited
Maidenhead, Berkshire, England
B. F. Lock
Salisbury College of Advanced Education
Adelaide, Australia

Testing Land-Use Map Accuracy

A simple statistical sampling procedure for determining the accuracy of remote sensing-derived land-use maps is described.

Abstract: *In providing land-use maps to meet prescribed specification accuracy standards from remote sensing data, some form of reliable statistical testing is necessary. Cost considerations in operational remote sensing land-use surveys necessitate the use of a sampling strategy for field checking. Such a strategy should require that only the minimum number of field areas necessary need be checked, yet ensure that results are still statistically valid. Such a procedure is explained and recommended for inclusion in all land-use survey job specifications issued by international, government and other user agencies.*

Introduction

In recent years the development of techniques for collecting and processing remotely sensed data has progressed very rapidly, but many problems still persist in the reliable utilization of the information. In order to achieve wider acceptance among users of land-use mapping from remote sensing data, the interpreter must be able to specify the accuracy of his product. Due to time and cost constraints, it would be virtually impossible, in the practical sense, to check completely each land-use parcel throughout a region. Therefore, a valid sampling procedure is required to estimate classification accuracy. Stratified random techniques have been accepted as the most appropriate method of sampling in land-use studies using remote imagery, so that smaller areas can be satisfactorily represented (Rudd, 1971; Zonneveld, 1972). But the problem remains concerning the selection of the best (i.e., minimum) sample size for each category.

Although several alternative methods have been used in other projects, few provide sufficient statistical justification for the allocation of sample points in each category of land use using remote sensing imagery. Papers by Stobbs (1968) and Hord and Brooner (1976) are some of the few published reports where the mathematical bases for determining the number of sample points are adequately detailed. However, their design parameters do not permit extensive utilization of their system for adoption for use with most forms of remote sensing imagery including orbital data. The following paragraphs describe a simple yet reliable method for determining the most appropriate (i.e., minimum) sample size acceptable in order that valid statistical testing of remote sensing land-use accuracy may be carried out.

The function of the ground truth survey in an operational remote sensing land-use survey is to utilize a sound statistical sampling design which will test the correctness of the attribution, by interpretation, of specific sites to classes in the classification. That is, for any sample point, it should be shown

TABLE 1. MATRIX SHOWING HYPOTHETICAL NUMBERS OF SITES IN ACTUAL AND INTERPRETED LAND-USE CATEGORIES.

		Land Use (on the ground)			
		A	B	C	Sum
Land Use (interpreted from imagery)	A	12	1	4	17
	B	2	19	0	21
	C	1	0	12	13
	Sum	15	20	16	51

whether the remote sensing attribution to a class within the classification is correct or in error.

Some of the main aspects that need to be considered in such a remote sensing sampling design are—

- The frequency that any one land-use type (on the ground) is erroneously attributed to another class, for example, in Table 1 3/17 of A is erroneously attributed to other classes;
- The frequency that the wrong land use (as observed on the ground) is erroneously included in any one class, for example, 5/17 of A attributions are erroneously interpreted;
- The proportion of all land (as determined in the field) that is mistakenly attributed by the interpreter, for example, 8/51 of all attributions are incorrect; and
- The determination of whether the mistakes are random (so that overall proportions are approximately correct) or subject to a persistent bias, for example, there is a significant tendency to mis-attribute land use C (on the ground) to category A, i.e., 4/16.

Thus the successful design of a sampling and statistical testing procedure will allow an approximate answer to each of these aspects.

SAMPLE SIZE

In order to determine the optimum sample size (defined as the minimum number of points that need to be checked in the field yet still meet a specification requirement of 'q' accuracy) for a stratified random sample of a region which has been mapped by remote sensing techniques, it is necessary to consider, primarily, one land-use type or category (stratum) which has been identified from remote sensing imagery. A sample of x points in that land-use type can then be selected and the number of errors (f) checked in the field. If such a procedure adopts a very small sample (e.g., $x = 10$), the number of errors would normally also be small (e.g., $f = 0, 1, 2, \ldots$). However, the achievement of perfect results (i.e., $f = 0$) in such a small sample does not imply that the method is error-free, because the result may occur by chance in a situation where a substantial portion of the land-use classification was in fact erroneous. This fact is seldom appreciated by many image interpreters when checking the accuracy of the results of their remote sensing land-use survey. The proportion of the interpretation which is in error would be identified in a very lengthy study, and is normally called p per cent (or p as a decimal fraction). The probability of making no interpretation errors when taking a sample of x from a remote sensing based classification, with real errors having a probability, p, is given by the binomial expansion $(p + q)^x$, in which $q = 1 - p$. In the case of no errors in the interpretation, the last term of the expansion is the only one of interest (i.e., q^x).

Table 2 shows the probability of scoring no interpretation errors in samples of varying sizes taken from a population with a range of real error proportions p. This table indicates that no-error sample results can quite easily arise in small samples when the true error rate is high. Taking the conventional probability level of 0.95 / 0.05 (95 percent / 5 percent), the table can be divided into two parts by a 'stepped' line. Above and

TABLE 2. PROBABILITY OF SCORING NO ERRORS IN SAMPLES OF VARYING SIZES FROM A POPULATION WITH A RANGE OF REAL ERROR PROPORTIONS q

		Sample Size										
	q \ x	5	10	15	20	25	30	35	40	45	50	60
specified interpretation accuracy	0.99											0.5472
	0.95						0.2146	0.1661	0.1285	0.0994	0.0769	0.0461
	0.90			0.2059	0.1216	0.0718	0.0424	0.0250	0.0148	0.0087	0.0052	0.0461
	0.85			0.0874	0.0388	0.0172						
	0.80		0.1074	0.0352								
	0.70	0.1681	0.0282									
	0.60	0.0778										
	0.50	0.0313										

——— stepped line indicates approximate 0.05 level of probability.

to the left of the line, the probabilities of obtaining error-free sample results are low, even when true errors are present in appreciable numbers. Below and to the right of the line, it is possible to identify the high probability that these error-free results could have been obtained only from a method which was relatively error free.

Thus, if the permissible error rate in the image interpretation is predetermined, for example, 85-90 percent as suggested by USGS circular 671 (Anderson *et al.*, 1972) or as required in an operational job specification, the sample size for each land-use category (stratum) necessary for 85 percent interpretation accuracy should be at least 20, for 90 percent accuracy at least 30, and so on. Therefore, by using Table 2, the minimum sample size required for checking any interpretation accuracy level can be determined. It is a minimum as for any smaller sample size, even a perfect (i.e., error-free ground check) result signifies very little.

Sampling Strategy

In order to locate the required number of points (e.g., 30 points for a 90 percent interpretation accuracy level), random point sampling within a land-use category (or stratum) can be performed by sampling using random spatial coordinates. Further details of the sampling strategy, sampling design, and analysis of results may be found in van Genderen and Lock (1976).

Conclusion

The concept developed and described incorporates the probability of making incorrect interpretations at particular prescribed accuracy levels, for a certain number of errors and for a particular sample size. This contrasts with the usual practice of expressing the interpretation errors as a percentage of a subjectively derived number of sample sites. Consequently, it is considered that the approach presented here offers a more meaningful explanation of the interpretation accuracy level for an entire remote sensing land use survey and also within each category.

The sampling strategy presented has an added advantage in that it can be adapted easily for use with most forms of remote sensing imagery, including orbital data. It provides a reliable framework for testing the accuracy of any remote sensing image interpretation-based land-use classification using the minimum number of sample points possible, thereby saving time and money, especially if it is employed in operational surveys where high specification accuracy levels need to be guaranteed.

Acknowledgment

The authors wish to thank Dr. A. Hay, University of Sheffield, for his assistance with the development of the statistical sampling procedure.

References

Genderen, J. L. van, and B. F. Lock, 1976. *A Methodology for Producing Small Scale Rural Land Use Maps in Semi-Arid Developing Countries Using Orbital M.S.S. Imagery,* Final Contractor's Report—NASA—CR-151173 September 1976, 270 pp.

Hord, R. M., and W. Brooner, 1976. "Land-Use Map Accuracy Criteria," *Photogrammetric Engineering and Remote Sensing,* Vol. 42, No. 5. May 1976, pp. 671-677.

Rudd, R. D., 1971. "Macro Land-Use Mapping With Simulated Space Photographs," *Photogrammetric Engineering,* Vol. 37, pp. 365-372.

Stobbs, A. R., 1968. "Some Problems of Measuring Land Use in Under Developed Countries: the Land Use Survey of Malawi," *Cartographic Journal,* Vol. 5, pp. 107-110.

Zonneveld, I. S., 1974. "Aerial Photography, Remote Sensing and Ecology," *ITC Journal,* 1974, Part 4, pp. 553-560.