Stan Aronoff*
TES Research and Consulting Ltd.
Calgary, Alberta T3A 2G6, Canada

The Map Accuracy Report: A User's View

The objective is to present the test data in a disaggregated form so that individual users can interpret the results for their specific application.

Introduction

A variety of map accuracy estimation methods and report formats have been developed to assess and compare the results of maps produced from remotely sensed data or other information. These tests have commonly addressed specific accuracy questions when the designer of the test is either the ultimate user or has designed it to provide a certain type of information for a hypothetical user. The objective of the report format outlined in this paper is to present the data of a map accuracy test in such a way that users with a variety of applications can each evaluate the suitability of the map for their objectives.

The report format is based on the theory of map accuracy estimation presented in Aronoff (1982). All values, except the minimum accuracy value, are obtainable from tables in Ginevan (1979) or by using the equations discussed below.

Abstract: *A reporting format for map accuracy test results is described and illustrated. The objective of the format is to present the test data in a disaggregated form so that individual users can interpret the results for their specific application. Consumer and producer risks, overall map accuracy, individual class accuracies, and minimum map accuracy values are also summarized.*

Map Accuracy Estimation

Map accuracy testing using the binomial sampling distribution is briefly outlined below. A more complete discussion can be found in Aronoff (1982).

To decide whether a map is of acceptable accuracy, a sample of map points is checked against ground data and a probabilistic statement is made about the true accuracy of the map. The statement generally claims some minimum level of accuracy with some high level of confidence, e.g., a minimum of 85 percent accurate at the 95 percent confidence level. The sampling problem is one of determining the optimal number (N) of map samples to be compared with ground data, and an allowable number (X) of misclassifications of these samples. After these values are determined, N map samples are selected and their classifications are compared against the true classification of the sample point (e.g., ground data). If X or fewer points were misclassified, then the map is accepted as accurate at the specified level of precision. (It is assumed that misclassification of a site can be unambiguously determined.)

Consumer's risk is the probability that a map of unacceptable accuracy will pass the accuracy test. It can be calculated from the binomial sampling distribution as follows:

$$\text{CONRISK} = \sum_{Y=0}^{X} \frac{N!}{Y!(N-Y)!} Q_L^{N-Y} (1 - Q_L)^Y \qquad (1)$$

where CONRISK = consumer's risk,
Q_L = the minimum accuracy required,
X = number of allowable misclassifications,
N = total number of points sampled, and
Y = number of misclassifications.

The producer's risk is the probability that a map of some acceptable accuracy Q_H will be rejected and is calculated as follows:

* Now at the Department of Forestry, University of California, Berkeley, CA 94720.

0099-1112/82/4808-1309$02.25/0

$$\text{PRODRISK} = \sum_{Y=X+1}^{N} \frac{N!}{Y!(N-Y)!} Q_H^{\,N-Y} (1 - Q_H)^Y \quad (2)$$

where PRODRISK = producer's risk and
Q_H = a selected high accuracy level.

Values of consumer and producer risks for specific sample designs have been tabulated in Ginevan (1979), some of which are shown in Table 1. The tables presented in the report cover sample numbers as high as 397, consumer risks of 5 and 1 percent; Q_L = 85 and 90 percent; and values of Q_H = 90, 95, and 99 percent.

The selection of values for consumer and producer's risks depend on the value of the information and cost of errors in a specific application. Though the values used to produce the "Summary Data" sections of the proposed accuracy report are not selected by the user, if they are not suitable the user can evaluate the data in the error matrices and calculate statistics of his choosing.

Method

Overall Map Accuracy

Consider the design of an accuracy assessement to test whether the overall map accuracy is at least 85 percent = Q_L with a consumer's risk of 5 percent. Suppose that after considering the ultimate use of the map and cost of misclassifications, the producer's risk is set at 30 percent for a map of true accuracy 90 percent. Referring to Table 1, the sample design found to meet these requirements requires 213 randomly selected points with 23 allowed misclassifications. The test results are tabulated in the Map Accuracy Test Error Matrix in Table 2. The Summary Data section for this error matrix states the parameters of the test and gives a minimum map accuracy value. The minimum map accuracy value is the highest map accuracy value which would pass the test given the observed sample results and the selected value of consumer risk, in this case 5 percent (i.e., it is the highest map accuracy for which the observed sample results would be a "pass"). It is calculated from Equation 1 by substituting the observed number of misclassifications equal to X (i.e., in this case X = 24), N = 213 as before, and solving for the highest value of Q_L such that the consumer's risk is less than or equal to that required, in this case CONRISK ≤ 5 percent. This can be done either by using tables of the cumulative binomial or with a simple computer program.

Instead of reporting only that a map passed or failed the test, a minimum map accuracy value indicates how well it passed or failed. In this example the map failed the test by having 24 misclassifications. However, by giving a minimum map accuracy value, in this case 84.5 percent, the user may decide that the map is still suitable for the application.

TABLE 1. OPTIMAL SAMPLE SIZES (N), THEIR ASSOCIATED ALLOWABLE ERRORS (X), AND PRODUCER'S RISK FOR TRUE ACCURACY VALUES 90%, 95%, AND 99%, GIVEN A REQUIRED CONSUMER'S RISK ≤5% [ADAPTED FROM GINEVAN (1979)]. FIRST DETERMINE THE DESIRED PRODUCER'S RISK FOR A TRUE MAP ACCURACY VALUE (Q) OF 90%, 95%, OR 99%. READ DOWN THE APPROPRIATE COLUMN UNTIL A PRODUCER'S RISK SMALLER THAN OR EQUAL TO THAT REQUIRED IS ENCOUNTERED. FOLLOW THIS ROW TO THE LEFT TO FIND THE REQUIRED N AND X VALUES.

	Producer Risks in % for Values of Q			
N	X	Q = 90%	95%	99%
19	0	87	62	17
30	1	82	45	4
40	2	78	32	1
50	3	75	24	0
59	4	72	17	0
68	5	69	12	0
76	6	65	9	0
85	7	62	6	0
93	8	59	4	0
102	9	58	3	0
110	10	55	2	0
118	11	52	2	0
126	12	50	1	0
134	13	47	1	0
142	14	45	1	0
150	15	43	0	0

Accuracy Assessement of Individual Classes

A further analysis of the map to assess the accuracy with which each class has been classified can be done using a method adapted from that of Hay (1979). Continuing the example, suppose that the minimum required accuracy Q_L is 85 percent and the consumer risk is set at 5 percent. It was also decided that cost considerations allowed only 50 points to be sampled from each class. From Table 1 the number of allowable errors is X = 3, and the producer's risk for a map of true accuracy 90 percent is 0.75 or 75 percent. Suppose that the 75 percent producer risk was considered tolerable except for class C which was of special interest, and that it was decided that 126 points would be sampled from this class with an allowable error of 12, thereby reducing the producer risk to 50 percent for a map of true accuracy = 90 percent (from Table 1).

Sampling is done by continuing to select random points after the initial 213 were selected for the previous test except that once sufficient points have been drawn for a given class, additional points for that class are discarded (i.e., this becomes a stratified random sample). No additional points are drawn for class D, for example, because 96 had already been selected. However, all the selected points are used. Referring again to Table 1, because N = 96 has not been tabulated, the next

TABLE 2. MAP ACCURACY TEST RESULTS

MAP ACCURACY ERROR MATRIX

		CLASSES (VERIFIED) A	B	C	D	E	Total	% Correct	% Commission
OBSERVED	A	26	1	0	0	1	28	93	7
	B	1	5	0	0	3	9	56	44
	C	2	0	43	1	2	48	90	10
	D	4	1	2	76	13	96	79	21
	E	0	0	2	1	29	32	91	9
	Total	33	7	47	78	48	213		
	% Omission	21	29	9	3	40			

SUMMARY DATA FOR MAP ACCURACY ERROR MATRIX

Points Sampled	213
Allowable Misclassifications	23
Observed Misclassifications	24
Consumer's Risk	5%
Producer's Risk (given true map accuracy of 90%)	30%
Minimum Map Accuracy	84.5%

lower value $N = 93$ is used. This gives an allowable error of 8 and a producer's risk of 59 percent for a map of true accuracy = 90 percent.

The results of this test are shown in the Class Accuracy Error Matrix in Table 3. The Summary Data section for this matrix is compiled as for the overall map accuracy test described previously, except that the values pertain to a single class. Thus, the minimum class accuracy value is the highest class accuracy value which would pass the test, given the observed results and the selected consumer's risk.

INTERPRETATION OF THE ACCURACY TEST

From the overall map accuracy results, it can be seen that the map narrowly failed by one misclassification compared with an allowable 23 misclassifications. The minimum map accuracy of 84.5 percent is quite close to the required 85 percent. The error matrix shows class B to be poorly classified. However, the class has a small sample size and conclusions drawn about the class from only nine sample points would be unreliable. Similarly, the other classes can be compared, recognizing that the larger the sample size the more reliable the data.

Individual classes can be more confidently compared using the class accuracy test results (Table 3). Class C is identified with relatively high accuracy and meets the minimum requirements with low commission and omission errors. Class D does not pass the test and has a 21 percent commission error but only a 3 percent omission error. Class E passes the test with a 6 percent commission error but a high omission error of 29 percent.

The values in the class accuracy error matrix may be more important to the potential user than knowing whether the class passed or not. Suppose, for example, that classes D and E represented unstable soils occuring as scattered small patches. The user needs the map to identify these areas because road construction costs over these areas are prohibitively high whereas detours around these areas cost relatively little. Then the cost of a commission error (needlessly avoiding the area) becomes relatively low, but the cost of an omission error is very high. Thus, the user may conclude that the map product is suitable for identifying class D even though it failed the test, because there was a low probability of an omission error. The user might also conclude that it was unsuitable for identifying areas of class E even though it passed the test, because the probability of a costly omission error is too high.

If, however, it was unnecessary to distinguish between classes D and E, then the two classes could be combined if the data of Table 2 are used. This would give a total sample size of $96 + 32 = 128$ observed points and $78 + 48 = 126$ verified points in the combined classes. The number of correctly identified points for the combined class would be $76 + 1 + 13 + 29 = 119$, giving an omission error of $(126 - 119)/126 = 0.06$ or 6 percent and a commission error of $(128 - 119)/128 = 0.07$ or 7 percent. The minimum accuracy for the combined class is 88 percent at a 5 percent consumer risk; thus, the combined class also passes the 85 percent accuracy test.

(Note that the values in Table 3 could not be used in this way because points for each class were selected separately, discarding points not from the class of interest, i.e., each class was a separate

TABLE 3. CLASS ACCURACY TEST RESULTS

CLASS ACCURACY ERROR MATRIX

		VERIFIED							
		A	B	C	D	E	Total	% Correct	% Commission
OBSERVED	A	47	2	0	0	1	50	94	6
	B	7	40	0	0	3	50	80	20
	C	7	0	116	1	2	126	92	8
	D	4	1	2	76	13	96	79	21
	E	0	0	2	1	47	50	94	6
	Total	65	43	120	78	66	372		
	% Omission	28	7	3	3	29			

CLASS ACCURACY SUMMARY DATA

	CLASSES				
	A	B	C	D	E
Points Sampled	50	50	126	96	50
Allowable Errors	3	3	12	8	3
Observed Errors	3	10	10	20	3
* Minimum Class Accuracy	85	68	87	71	85
** Producer's Risk	75	75	50	59	75

* Minimum class accuracy given in % and calculated for a consumer's risk of 5%.
** Producer's risk given in % and calculated for a map of 90% accuracy and using the allowable error values.

sample frame. The data in Table 2 were generated by randomly sampling points from the entire map. Thus, all points, regardless of class, were from a single sampling frame.)

In this case, by having the data from the accuracy test available, the user could do a detailed analysis of the map product for a specific application. Had only the minimum map accuracy value been available, the user might not have been able to afford the risk that the classes of interest were not mapped with sufficient accuracy.

Conclusion

The objective of a map accuracy report should be to present accuracy test data in a way that allows the user to extract the test results important for his specific application. This enables the user to evaluate the test data in a way other than that of the designer of the test or producer of the map product. A report format meeting these objectives was proposed which requires a minimum of calculations and is understandable to the user with a minimum of statistical training. At the same time the data are available for users wanting to calculate probabilities of specific types of error. In providing this flexibility of interpretation, the map accuracy report enables the map product to be more fully utilized because users can better predict its performance for their specific application.

Acknowledgments

The author gratefully acknowledges the critical reviews by Dr. R. N. Colwell and Dr. W. Libby of the Department of Forestry, and Mr. Andrew Benson of the Remote Sensing Research Program at the University of California, Berkeley. The author would like to thank the Department of Forestry of the University of California at Berkeley for supporting this research.

References

Aronoff, S., 1982. Classification Accuracy—A User Approach, *Photogram. Engin. and Remote Sensing* 48(8):1299-1307.

Ginevan, M. E., 1979. Testing Land-Use Map Accuracy: Another Look, *Photogram. Engin. and Remote Sensing* 45(10):1371-1377.

Hay, A. M., 1979. Sampling Designs to Test Land-Use Map Accuracy, *Photogram. Engin. and Remote Sensing* 45(4):529-533.

(Received 2 July 1981; revised and accepted 8 February 1982)