*Quantitative Evaluation of Photo Interpretation Mapping**

HAROLD E. YOUNG[1] *and*
ERNEST G. STOECKELER[2]

ABSTRACT: *The accuracy of maps prepared by photo interpretation techniques has often been questioned by those who make the maps as well as by many users. In the past five years a number of American workers in the field of photo interpretation have emphasized the need for quantitative evaluation of photo interpretation in* PHOTOGRAMMETRIC ENGINEERING.

To illustrate the type of information required and the analytical procedures necessary to make a quantitative evaluation, an engineering soils map of a 210 square miles area (544 square kilometers) in the State of Maine was used as an illustration. Data were presented comparing photo interpretation classification versus actual field classification of soils. By applying recognized statistical procedures, the authors prepared a table presenting photo interpretation errors based on field sampling at 19:1 odds. These data were further refined to show the error at 19:1 odds within soil types. In addition, a hypothetical illustration was presented to show the method of determining the accuracy of boundary line location between adjacent soil types.

In the authors' opinion, most mapping studies of the earth's surface and vegetation should be evaluated quantitatively so that the users can employ the map information more intelligently. With an accumulation of results of quantitative evaluation of a number of photo interpretation studies, it will be possible to compare photo interpretation accuracy resulting from such controllable factors as scale, season, film and filter as well as differences between photo interpreters.

INTRODUCTION

FOREST type maps, engineering soils maps, maps of geologic studies, geographic studies and terrain analysis maps for military purposes of hundreds of thousands of square miles of territory have been prepared with the aid of photo interpretation techniques. It is inevitable that many people employing the maps will ask the question—how accurate is the photo interpretation?

In most instances, it is not feasible to field check more than a small portion of the total area, and in some cases, it is not possible to do any field checking at all. Thus the concern over map reliability is very real. In the past five years Colwell (1), Lundahl (4), Sammi (8), Katz (2), Rogers (7), Pomerening and Cline (5) and Young (10) have emphatically expressed the need for quantitative evaluation of photogrammetry and photo interpretation.

The papers and comments cited above should have aroused some interest among those who are doing photo interpretation of the earth's surface and vegetation. Krumbein (3) and Potter and Siever (6) have recently employed

* A Contribution of Commission VII, Photographic Interpretation, International Society of Photogrammetry.

[1] Assistant Forester, Maine Agricultural Experiment Station.

[2] Research Assistant, Soils Mechanics Laboratory, University of Maine.

statistical methods in evaluating geologic maps. It is the purpose of this paper to present the general procedures for evaluating photo interpretation in terms of the accuracy of the designation of a map unit, and the accuracy of the location of the boundary lines between map units.

Methods

Descriptions of the map symbols, type of photography, soil type photo patterns and field checking methods have been omitted purposely to avoid complicating the presentation of procedures involved in the quantitative evaluation of photo interpretation and map accuracy. Principles and procedures will vary for different types of photo interpretation studies. The five general phases in the preparation of an engineering soils map* of the Bangor Quadrangle, along with the collection of data necessary for the evaluation, are as follows:

A. The making of a field reconnaissance with photos in hand to determine the significance of various soil patterns and photo elements in order to develop a set of diagnostic features.
B. The delineation of soil type boundaries by photo interpretation in the office.
C. The field checking of the soil type designations and the accuracy of the location of the boundary lines.
D. The transference of photo detail to a base map.
E. The spot checking of the map in the field to determine the accuracy of the final product.

While in the field (Step C above) a record was maintained of the photo interpretation classification and the field classification of 449 sampling locations (Table II), and interpretation errors in designation and location of boundary lines were rectified. For this study, the number of areas of each soil type was counted and the acreage of each type was determined by use of a dot grid with 16 dots to the inch. Data for the Bangor Quadrangle, Maine, an area of 134,500 acres or approximately 210 square miles, are summarized in Table I. Map unit symbols represent soil types recognized in the Maine Reconnaissance Engineering Soils Classification System. Minimum area mapped for this study was five acres.

Table I

Engineering Soils Information, Bangor Quadrangle, Maine

Map Unit	Total Acreage	Number of Areas	Av. Acreage/Area
R	14,050	493	28.5
BG	14,770	131	112.7
B	48,200	152	317.1
G	3,280	135	24.3
F	34,200	96	356.3
S	7,970	294	27.1
P	5,400	43	125.6
W‡	6,630	11	602.7
	134,500	1,355	

‡ Signifies water which does not appear in other tables and in the discussion.

* The Maine Reconnaissance Engineering Soils Classification System was developed by E. G. Stoeckler as a part of a photo interpretation project financed by the Maine State Highway Commission in cooperation with the U. S. Bureau of Public Roads.

TABLE II

COMPARISON OF ENGINEERING SOILS TYPES AS CLASSIFIED BY PHOTO INTERPRETATION AND FIELD CLASSIFICATION, BANGOR QUADRANGLE, MAINE

CLASSIFICATION BY PHOTO INTERPRETATION

Soil Type	R	%	BG	%	B	%	G	%	F	%	S	%	P	%
R	*58*	*79*	24	32	9	10			2	2				
BG	11	15	*47*	*62*	3	3								
B	3	4	3	4	*64*	*72*	1	4	15	16	2	4		
G	1	1	2	3	4	4	*20*	*77*	4	4				
F					9	10	5	19	*72*	*76*	1	2		
S									2	2	*40*	*74*	5	14
F											11	20	*31*	*86*
Total	73	99	76	101	89	99	26	100	95	100	54	100	36	100

Italicized values indicate correct photo interpretations as confirmed by field checking. For example, of the 73 areas classified by photo interpretation as *R*, 58, or 79 per cent, were found to be correct. Of the remainder, 11, or 15 per cent, were found to be erroneously classified as *BG;* 3, or 4 per cent, were called *B;* and 1, or 1 per cent, was called *G*.

ACCURACY OF MAP UNIT DESIGNATION

An examination of the engineering soils map indicated that the existing road grid traversed the map units to such a degree that it would be convenient to do the sampling in the vicinity of the roads. Although this is a questionable departure from the concept of random sampling, it was accepted because of the limitations in time and funds.

If the field check is accepted as the final authority, then the photo interpretation is either correct or incorrect. This makes the data amenable to analysis employing the binomial theorem. Snedecor (9) discusses this in detail and provides a table on page 4 of *Statistical Methods*. Within this table the confidence interval at 19:1 odds or 99:1 odds can be determined without any computations at all. In chapter 16, Snedecor provides the background information to make computations leading to a greater degree of refinement than is possible by using the table.

ACCURACY OF BOUNDARY LINE LOCATION

The accuracy increment, or degree of precision, in map checking is dependent on map scale, type of interpretation, whether the study is reconnaissance or detailed in nature, and other factors peculiar to any given mapping project. The reconnaissance engineering soils map of the Bangor Quadrangle is at a scale of 1:31,680. At this scale, the actual width of soil type lines on the final map represents approximately 50 feet on the ground. For this study the minimum error, or accuracy increment, to be measured was established as 100 feet. For other types of mapping, the accuracy unit may be a foot or a hundred yards depending on the scope and purpose of the particular map.

A problem in checking the boundary line accuracy is locating the boundary in the field. In some instances the boundary is well defined, but in others there is a transitional band of considerable width between map units. Unless these bands are extensively sampled, different observers could easily locate the boundary in different places. For reconnaissance mapping, however, time-consuming and expensive sampling is not justified. In soils engineering, forestry and geology there are situations in which boundary delineations by photo interpretation are more accurate than boundary lines located by orthodox field checking methods. In this particular study on map accuracy there were only two instances out of 77 field checks where the actual soil boundary was more than 100 feet removed from the location on the map. It is well to point out that field checking the finished map is a combined check on the photo interpretation, the transfer of detail to the base map, and on any possible drafting errors. Thus an error of 100 feet might actually be attributable to transfer of detail by the sketchmaster or to drafting error rather than to photo interpretation.

Had there been more errors between the photo interpretation and the field check of the boundary line location, then the errors would have been sorted

TABLE III

HYPOTHETICAL DATA TO ILLUSTRATE ANALYTICAL METHODS FOR DETERMINATION OF BOUNDARY LINE ACCURACY BETWEEN MAP UNITS

Accuracy Increments of 100 feet	Adjoining Types						
	R/BG	R/B	R/G	R/F	R/S	R/P	total
0	10	7	3	2	2	1	25
1	3	2	1	0	0	0	6
2	2	0	0	0	0	0	2
3	1	0	0	0	0	0	1
Total	16	9	4	2	2	1	34

ANALYSIS OF VARIANCE

Due to	D.F.	Sum of Squares	Mean Square	F. Ratio
Between accuracy Units	3	72.47	24.16	7.09*
Within accuracy Units	22	75.00	3.41	
Total	25	147.47		
Correction factor	1	38.53		
Grand Total	26	186.00		

* Significant at 99:1 odds. Of the 3 degrees of freedom the independent comparison between (0) and (1, 2, and 3) accuracy units accounts for 69.0 of the 72.47 in the sum of squares. The other two independent comparisons are not significant.

The interpretation of the above hypothetical analysis is that errors of 100 feet or more in delineation of boundary units will occur less frequently than zero errors.

and analyzed as in the hypothetical illustration in Table III. Snedecor (9) discusses the analysis of variance technique in chapters 10 and 11 appropriate to such an analysis. No tables similar to those for the binomial theorem are available thereby making necessary the computations for each set of data.

RESULTS

A comparison of Tables I and II shows that the number of areas in the field check was not proportional to the number of map unit areas because some of the larger map unit areas were sampled in several places. Tables IV and V present the results of the statistical analysis of each soil type or map unit.

The *R* soil type is used to illustrate proper interpretation of the tables mentioned above. Seventy-three soil type areas, out of a total of 493, classified as

TABLE IV

PHOTO INTERPRETATION ERRORS BASED ON FIELD SAMPLING AND EXPECTED RANGE OF ERROR AT 19:1 ODDS, BANGOR QUADRANGLE, MAINE

Soil Type	Number of Field Samples	Per Cent Error (Actual)	Range of Error at 19:1 Odds in Per Cent
R	73	21	14–34
BG	76	38	31–55
B	89	28	19–38
G	26	23	10–46
F	95	24	15–32
S	54	26	15–41
P	36	14	6–32
	449		

R by photo interpretation, were checked in the field. Fifteen of these areas, or 21 per cent, were found to be something other than *R* in the field check. The statistical analysis indicates that, based on this sample, the expected range in error for the entire quadrangle sheet for the *R* soil type, at 19:1 odds, ranges from a minimum of 14 per cent to a maximum of 34 per cent. In other words, on the basis of photo interpretation alone, without field checking and accompanying map corrections, of the 493 areas classified as *R* it can be expected that error in designation will range from a minimum of 69 to a maximum of 168.

To analyze the data further, Table V indicates the nature of the errors within each soil type. Errors in *R* turned out to be mainly *BG* with some *B* and a little *G*. The range of these errors are also calculated by use of the binomial theorem. Although 15 per cent of the areas designated as *R* by photo interpretation actually were *BG* by field sampling, in the photo interpretation of the entire quadrangle, designations of *R* can be expected to be *BG* ranging from a minimum of 8 per cent to a maximum of 27 per cent of the time.

The data represent only a comparison of office study designation of soil types by photo interpretation with actual field sampling. It should be stated here that Tables II, IV and V do not apply to the finished map of the Bangor Quadrangle as all errors observed in field work, and extending beyond the range of these data, were eliminated.

In the field checking of the boundary lines of the finished map, only two

TABLE V

PHOTO INTERPRETATION ERRORS BASED ON FIELD SAMPLING AND EXPECTED RANGE OF ERROR AT 19:1 ODDS WITHIN SOIL TYPES, BANGOR QUADRANGLE, MAINE

Per Cent Misclassified within a Soil Type and Expected Range of Error at 19:1 Odds

Soil Type*	R %	Error %	BG %	Error %	B %	Error %	G %	Error %	F %	Error %	S %	Error %	P %	Error %
R			32	25–48	10	5–19			2	0– 7				
BG	15	8–27			3	1–10								
B	4	1–12	4	1–12			4	0–21	16	9–24	4	0–14		
G	1	0– 8	3	0–10	4	1–12			4	1–10				
F					10	5–11	19	7–41			2	0–11		
S									2	0– 7			14	6–32
P											20	12–36		

* Classification by field sampling.

instances out of 77 indicated differences of as much as 100 feet. Thus it appears that boundary lines can be located with more precision than the map units can be designated by means of photo interpretation. This certainly indicates the need for more field checking of the latter than the former.

Discussion

It is the opinion of the authors that most studies of the earth's surface and vegetation should be evaluated quantitatively by the general procedure outlined above. Photo interpretation has not reached, and for this type of problem never will reach, the point where we can expect perfect results from the interpreters. Field sampling provides a check on the accuracy of the photo interpretation, both type designation and boundary line location. If the field checker would take the time, less than a minute per sampling location, to maintain a record of photo interpretation classification/field classification, he can, with very little effort, collect the data necessary to make a quantitative evaluation such as is outlined in this paper.

Once quantitative methods are established, it should become normal procedure to include estimate of error in planning so that the standards of photo interpretation can be established with sufficient field checking to yield results within established limits. It is understood that the standards will vary with the purpose for which the map is made, e.g., the precision of a reconnaissance survey will be of a lower level than that of a strip survey for highway location.

Quantitative evaluation provides each photo interpreter with a sound basis for measuring his own ability, his rate of progress and a means of comparing his work with that of others working in the same field.

The accumulation of the results of a number of photo interpretation studies that have been quantitatively evaluated will make possible comparing variations in interpretation due to such controllable variables as scale and date of photography. For example, other things being equal, if the over-all accuracy of interpretation with photos at a scale of 1:10,000 is 75 per cent, and if interpretation has an accuracy of 70 per cent at a scale of 1:20,000, then perhaps the additional gain of 5 per cent in accuracy does not justify the additional expense in procuring large scale coverage.

The authors have heard experienced photo interpreters make the following statements: (a) "1:20,000 *is* the optimum scale for studies in my field" (b) "Winter photography for my purposes *is* useless as I must have photos taken in the spring before the leaves are out." (c) "Photos taken in April *are* much better than those taken in June for soils interpretation." (*d*) "I *can* get sufficient accuracy from two inch to the mile photos." (e) Infrared photos taken in mid summer at a scale of 1:15,840 *are* the very best for my purposes." In practically all instances those who make such statements are voicing opinions which *may* or *may not* be valid. It is not a difficult matter to field check most photo interpretation studies so that data will be available for quantitative evaluation to determine the validity of such opinions.

Summary and Conclusions

The widespread use of photo interpretation as the primary means of preparing forestry, soils engineering, geology, geography and military terrain analysis maps generates the requirement that the finished product be quantitatively evaluated since perfection cannot be expected. Methods of quantitatively evaluating the accuracy of map unit designations, and the accuracy of boundary line location between map units as determined by photo interpretation, are presented. An engineering soils map prepared by photo interpretation provided data to illustrate these methods. It is the opinion of the authors that quantitative methods should be used in the preparation of maps involving photo interpretation as a means of establishing and maintaining satisfactory levels of precision, and permitting intelligent evaluation by map users.

ACKNOWLEDGEMENT

The authors express thanks to Professor F. X. Schumacher for assistance in the analytical procedures, and Mr. Gerald Hale for making the counts on the Bangor Quadrangle.

BIBLIOGRAPHY

1. Colwell, Robert, "The P. I. Picture in 1955," PHOTOGRAMMETRIC ENGINEERING, 21: 720–724.
2. Katz, Amrom H., "Photogrammetry Needs Statistics," PHOTOGRAMMETRIC ENGINEERING, 18: 536–542.
3. Krumbein, W. C., "Statistical Analysis of Facies Maps," *Jour. of Geology*, 63: 452–470.
4. Lundahl, Arthur, Remarks at close of Photo Interpretation Panel, PHOTOGRAMMETRIC ENGINEERING, 20: 462.
5. Pomerening, James A. and Marlin G. Cline, "The Accuracy of Soil Maps Prepared by Various Methods that Use Aerial Photograph Interpretation," PHOTOGRAMMETRIC ENGINEERING, 19: 809–918.
6. Potter, Paul E. and Raymond Siever, "A Comparative Study of Upper Chester and Lower Pennsylvania Stratigraphic Variability," *Jour. of Geology*, 63: 429–451.
7. Rogers, Earl J., "A Plan for Research in Fields of Aerial Photo Interpretation," PHOTOGRAMMETRIC ENGINEERING, 19: 801–805.
8. Sammi, John C., "The Application of Statistics to Photogrammetry," PHOTOGRAMMETRIC ENGINEERING, 16: 681–685.
9. Snedecor, George W., "Statistical Methods," 4th ed., Iowa State College Press, Ames, Iowa, 1946.
10. Young, Harold E., "The Need for Quantitative Evaluation of the Photo Interpretation System," PHOTOGRAMMETRIC ENGINEERING, 21: 712–714.