Aerial Verification of Polygonal Resource Maps: A Low-Cost Approach to Accuracy Assessment

Thomas H. George
Geophysical Institute, University of Alaska, Fairbanks, AK 99775-0800

ABSTRACT: Map verification requires that additional data be collected independently of information used to generate a resource map. This often overlooked part of map production provides a measure of quality control for map makers and users.

A method was devised to collect verification data by making observations from a light aircraft. These data were compared with photo-interpreted range maps to make an estimate of map accuracy.

The verification procedure was performed on a range survey area on the Seward Peninsula, Alaska. Results were used by map makers to identify areas requiring additional work to improve the quality of these preliminary maps.

This procedure offers a low-cost approach to verification of range maps which cover large geographic areas. With further refinement, it should be adaptable for use with other polygonal resource maps in different geographic settings.

INTRODUCTION

MAP VERIFICATION requires that additional data be acquired independently of the information used to generate a resource map. This often overlooked part of map production is important in providing a measure of quality control for (1) map makers to insure that their product meets the intended specifications, and (2) map users, who need to know how much to trust a specific map they may be using for a resource planning or management activity.

In Alaska, a great deal of mapping is performed at reconnaissance levels, ranging in scale from 1:30,000 to 1:250,000. Typical project areas vary from one to several million hectares in size. These large project areas, combined with the short field season and lack of surface transportation systems, severely restrict the collection of field data. With the difficulty and high cost of collecting field data, it is not too suprising that map verification is sometimes overlooked or given minimal attention.

Dozier and Strahler (1983) provide an excellent discussion of verification and accuracy assessment. Normally, map verification is performed to determine the level of accuracy of a map, or how well it meets its design specifications. As a matter of economy, a sample of point observations is collected and examined. Methods to determine the "correct" classification of a sample point vary from detailed ground inspections to interpretation of aerial photography. These reference data are compared to the map in question to establish its accuracy.

Two problems plague the conventional verification approach when addressing large-area resource mapping projects. First, a large number of samples is required to obtain reasonable estimates of map accuracy. Hay (1979) recommends a minimum of 50 samples per map category, using a stratified sample design. More would be required for a simple random sample. The cost of helicopter-supported field crews to gather large numbers of ground observations is often prohibitively expensive.

The second problem pertains to the physical size of the sample "point" relative to the ground size of a map delineation. Sample points are typically plots that range from a single metre-square area, to a transect of smaller plots, or in some cases, an area up to half an hectare in size. In contrast, a map delineation 2 centimetres on a side at a map scale of 1:30,000 represents 36 hectares on the ground. The same size delineation on a 1:250,000 map covers 2500 hectares. A ground sample between one square metre and half a hectare does indeed become a "point" in comparison.

Landscapes normally contain enough diversity that one cannot expect to find map-unit-sized areas that belong totally to a single mapping classification. To account for this diversity, it is customary to allow a delineation to contain areas of a foreign type of land cover or soils, called an inclusion. This limit of al-

PHOTOGRAMMETRIC ENGINEERING AND REMOTE SENSING,
Vol. 52, No. 6, June 1986, pp. 839–846.

0099-1112/86/5206-839$02.25/0

lowable inclusions are specified as a design parameter of the mapping project, and may occupy up to 15 or 20 percent of a mapping unit. In other words, up to 15 or 20 percent of the time, a "point" sample may fall into some type other than that labeled to a specific delineation. When this occurs, the reference data will then indicate the map unit is wrong when, in fact, the label was correct.

This paper presents a technique for collecting aerial observations as a reference data set to assess map reliability. This procedure was developed in the course of an operational range inventory and mapping project conducted in northwest Alaska.

BACKGROUND OF THE MAPPING PROJECT

The USDA Soil Conservation Service, assisted by the University of Alaska, has conducted a five-phase inventory and mapping project of the reindeer ranges on the Seward Peninsula (Swanson *et al.*, 1983).

The project area is located in northwest Alaska (see Figure 1). Tundra vegetation is the predominant cover of this landscape, ranging from low wet coastal plains, through moist rolling hills to dry alpine sites at higher elevations (see Figures 2 to 5). Soils tend to be poorly drained silts, which are almost entirely underlain by permafrost. Situated just below the Arctic Circle, winters tend to be long, cold, and dark, while summers are short and highly productive in terms of plant productivity.

Range maps have been derived by photo-interpretation of high altitude, color-infrared aerial photography on 1:60,000-scale photographic prints. Minimum map units are 65 hectares, although delineations typically range from 100 to 500 hectares in size. Mapping categories are vegetation and soil types called range sites (Shiflet, 1973), which are listed in Table 1. Map units usually contain an individual range site; however, a maximum of two sites may occur as a complex. Allowable inclusions of an unlabeled range site are limited to 20 percent of a map delineation. Desired overall mapping accuracy is 85 percent.

The survey has inventoried and mapped approximately 6.5 million hectares over a six-year period. This area presently supports 17 to 20 thousand reindeer. These maps represent the focus of this verification effort.

VERIFICATION METHOD

The process of recording aerial observations is somewhat analogous to photo-interpretation. An observer views the Earth's surface and examines the composition of objects that cover a landscape unit. Tone, texture, shape, size, and position all provide visual clues to the identity of cover types (in this case, range sites). The observer must be familiar with the categories to be evaluated and the terrain characteristics associated with each. Aerial observations differ in that the observer has only one opportunity to examine the landscape, where a photograph may be studied at length. Consequently, aerial observation conditions must be carefully planned and executed.

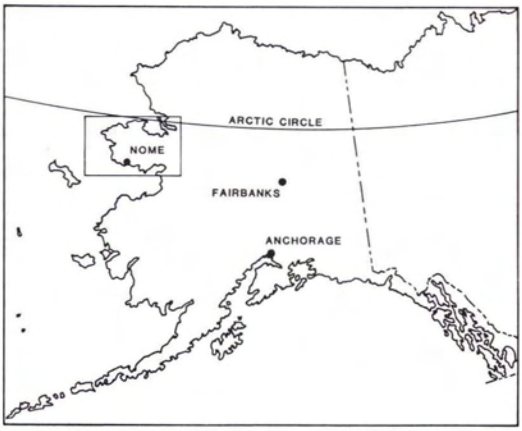

FIG. 1. The Seward Peninsula is the location of the Soil Conservation Service range inventory program.

The use of an aircraft as an observation platform influences data collection. Airplanes are most stable when flown straight and level. Excessive maneuvering can lead to disorientation and air sickness, both of which make it difficult to record observations. Even under ideal conditions, fatigue becomes a factor which influences the observation quality. To minimize these problems, straight line transects were selected for data collection. Transects have the added advantage of being easy to navigate, which helps when it comes to matching the location of observations with the appropriate map unit.

The information we wish to acquire is a measure of how well map units, as defined by mapping category descriptions, match the landscape. To collect suitable data to address this question, mapping delineations were used as the units to be sampled. To accomplish this, the observer is asked to look at the delineation defined by the map and determine what range site(s) fall within it. Using this approach, the observer is biased by knowing the boundary of the units, but not the label assigned by the map maker. Using this strategy, a four step data collection and analysis procedure was devised including (1) project planning, (2) observer training, (3) data collection, and (4) data analysis.

PROJECT PLANNING

The maps to be verified are studied to determine the overall project area, size of map units, and nature of mapping categories (see Figure 6). Transects are established across the project area to sample the variability of the terrain. Along each transect, a corridor is established within which aerial observations are recorded. Individual map units serve as the sample unit for data collection. Map unit boundaries that fall within the corridor are transferred, without labels, to an aerial photograph or topographic base map for verification (see Figure 7).

FIG . 2. Moist Tundra. Low rolling hills cover the majority of the Seward Peninsula. Tussocks of cotton grass (Eriophorum) are a common feature in these sites.

FIG . 3. Alpine Tundra. In mountainous areas, lichens and very low growth forms of shrubs and herbs are dominant.

FIG . 4. Wet Tundra. The coastal plains and other lowland areas are covered by wet tundra plant communities. Soil is normally saturated throughout the summer months.

FIG . 5. Shrublands. In areas that are well drained, shrubs manage to grow to heights of one or two metres.

The width of the observation corridor is determined by a combination of factors including map unit size, type and diversity of land cover, aircraft speed, and flying height above the ground. The observer must be able to see the entire corridor, have time to examine each map unit, and record the observation.

Once flight lines are established, the duration of observation periods is considered. Individual flight lines should be kept short (less than 20 minutes each) to avoid observer fatigue. Because it is generally not feasible to land every half hour, the use of two observers, who record alternate transects, should be considered.

Factors relating to the aircraft also must be incorporated. Aircraft range, the availability of fuel, and location of airstrips are all integrated into the planning process. After a verification mission plan is developed, the transects are plotted on air navigation charts.

TRAINING

The success of collecting usable verification data from an aircraft relies heavily on the advance preparation of the data collection team. The quality of data is largely dependent on the observer's skill and experience. Basic proficiency in map reading and cross-country navigation is required. The observers must be acquainted with the range sites (vegetation types) in the survey area, and be able to identify them both on the ground and from the air. Map symbols must be memorized so that observations can be recorded efficiently.

Care must be taken to organize all the material before the flight. Aerial photographs are arranged in sequential order and oriented in the same direction. Material for each flight line should be grouped separately, and extra pens kept within reach in flight. Items mislaid or out of order contribute to confusion, stress, and disorientation, which reduce the observers ability to accurately identify sample units.

DATA COLLECTION

Normally, a high-wing, single engine aircraft is used for data collection. The nominal crew consists of a pilot, navigator, and two observers.

The pilot is responsible for aircraft operation and overall flight safety. He should be briefed concerning the mission objectives and requirements. It is noteworthy that this is a specialized use of an aircraft, quite different from point-to-point transportation.

TABLE 1. SEWARD PENINSULA GENERAL MAPPING SITE LEGEND

Physiognomy	Code	Mapping Unit	Physiognomy	Code	Mapping Unit
Water: (1-9)	1	Lakes larger than 40 acres but smaller than 160 acres	Herbaceous	50	Dunes (beach)
	2	Lakes larger than 160 acres but smaller than 640 acres		51	Marsh (tidal)
	3	Lakes larger than 640 acres		52	Sedge (wet meadow)
	4	Lagoon		54	Sedge (drainageway)
	5	Ocean		55	Cottongrass-water sedge (low center polygons)
Trees: (10-19)	10	Mixed forest (floodplain)		56	Breached lake bed
	11	Black spruce		57	Sedge (wet lake bed)
	12	White spruce (upland)	Herbaceous: (60-69) (mat)	60	Lichen (tussock tundra)
	13	Spruce-lichen (upland)		61	Lichen meadow (mountain)
	14	Paper birch (upland)		63	Lichen-sedge (coastal tundra)
	15	Spruce-lichen (palsa)		64	Lichen-sedge meadow (upland)
Tall Shrubs: (20-29)	20	Tall shrub (floodplain)		65	Lichen slope (upland)
	21	Tall shrub (drainageway)		66	Lichen mat (lowland tundra)
	22	Tall shrub (hillside)	Mat and cushion	70	Lichen granitic slope (alpine)
Low Shrubs: (30-39)	32	Mixed shrub (tundra)		71	Dryas limestone slope
	34	Low shrub (floodplain)		72	Bald limestone slope
	35	Low shrub (hillside)		74	Dryas-lichen (alpine ridges)
Low Shrubs: (40-49) (herbaceous)	41	Shrub meadow (mountain)	Miscellaneous: (80-89)	80	Lava bed
	42	Tussock (tundra)		81	Barren
	42a	Wet phase		82	Riverwash
	42b	Mesic phase	Burn: (90-99)	90	Burned forest
	43	Meadow (alpine)		91	Burned tundra
	44	Shrub-lichen (upland)			
	45	Water sedge-muskeg (bog-fen)			

From SCS Alaska State Office

The pilot is an intergral part of the crew and must be able to work as a member of the team.

The navigator directs overall operation of the flight. Knowing the lines to be flown and observer requirements, he directs the pilot to the study area and along flight lines. The navigator keeps track of the aircraft position to insure that the observers have an unobstructed view of the corridor at all times. The navigator works closely with the pilot to evaluate inflight weather conditions and other logistical constraints, and selects alternate plans as required. In general, the navigator supervises the flight operation, leaving the observers free to concentrate solely on data collection.

The observers are the backbone of the project. Their tasks are to (1) locate the sample unit delineated on the aerial photograph or map, (2) examine the landscape below, and (3) record the appropriate symbol for the range site observed. As in photo-interpretation, they must consider decision rules concerning the presence of more than one range site, and the amount of allowable inclusions. Intense concentration is required to perform these tasks for extended periods of time. As previously mentioned, it is desirable to use two observers who take turns recording alternate flight lines.

Flight lines are flown at a height above terrain which allows the observer to view the entire corridor and see sufficient detail to identify the range site. Factors influencing flying height include the size of map units, difficulty of range site identification, and local weather. Typically, 300 meters above terrain seemed adequate for the conditions encountered in this project.

DATA ANALYSIS

After data collection, observations recorded from the air are compared with preliminary mapping results. The aerial observations are physically overlaid on the maps and tabulated on a unit-by-unit basis. A contingency table is constructed to show the agreement between the map and reference data set (see Table 2). By summing the units labeled correctly (diagonal of the matrix) and dividing by the total number of samples, an estimate of overall map accuracy is derived.

An estimate of map accuracy is more realistic if

FIG . 6. Preliminary photo-interpretive mapping product to be verified. Map unit labels correspond to the range sites listed in Table 1.

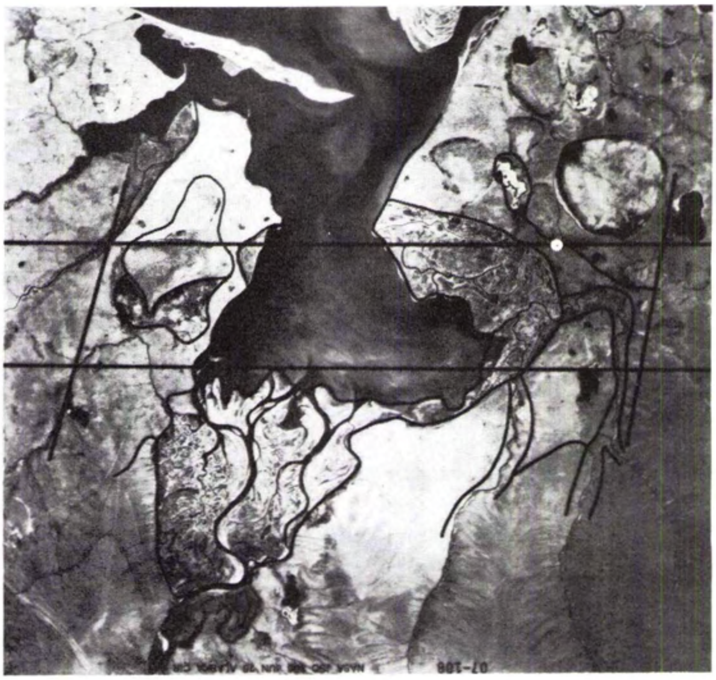

FIG . 7. The corridor to be verified is approximately 2.5-km wide. Map unit boundaries have been transferred from the preliminary map. An observer in the aircraft labels each unit with the symbol for the range site observed within.

TABLE 2. 1983 VERIFICATION DATA
Error Matrix

MAP UNIT CATEGORIES

AERIAL OBSERVATIONS	20	21	22	32	34	35	41	42	43	44	51	52	54	55	56	57	60	61	70	71	72	81	82	91
20	7	1																						
21	2	44	1	3		2		1																
22		2	3																					
32		2		14		2	4	3										1						
34		1			2																			
35		4	2			4		2	2	1		1												
41				2			2	1	1									3						
42				3			7	54	2			6												
43							2		1			1						3						
44																								
51	1										4				1									
52							3					2							1					
54													1	1										
55														9										
56					1										8	1								
57					1											1								
60			1	2			2	4	1	3		1					9	9						
61									1								1	7	4	1				
70																			4					
71																								
72																				1	5			
81																								
82																								
91																								7
Total	10	54	7	24	4	8	20	65	8	4	4	11	1	10	9	2	10	23	9	2	5	0	0	7
Percent	70	81	43	58	50	50	10	83	13	0	100	18	100	90	89	50	90	30	44	0	100	–	–	100

Total Observations = 297
Correct Observations = 188
Overall Agreement = 0.63 ± 0.05 (0.95 confidence level)

confidence limits are established. If the assumption is made that the sample is randomly drawn, 0.95 confidence limits for the binomial proportion can be calculated using the formulas:

$$\text{Upper confidence limit} = p + 1.96\sqrt{(pq/n)}$$
$$\text{Lower confidence limit} = p - 1.96\sqrt{(pq/n)}$$

where p is the estimate of percent correct, $q = (1-p)$, and n is the sample size (Snedecor and Cochran, 1980).

Finally, information contained in the off-diagonal elements of the contingency table may provide insight concerning the nature of classification errors. If a particular mapping category is consistently confused with another, it may be relatively easy to pinpoint the category requiring attention. Confusion between a wide range of mapping categories may result from factors such as poor interpretation, lack of adequate image detail, or mapping categories that are too detailed.

1983 VERIFICATION

In July 1983, the verification method described above was implemented on a portion of the Seward Peninsula. The objective was to examine some preliminary mapping covering a 1.6 million hectare area on the northern side of the peninsula. Figure 6 is a sample of the preliminary mapping which was photointerpreted the previous winter without the benefit of field reconnaissance. This verification was intended to test the mapping and indicate problem areas which could be addressed subsequently in the field season.

Project planning involved establishing approximately 500 kilometres of flight lines. Corridors 2.5-kilometres wide were devised, and plotted on aerial photographs. Map units were transferred from the preliminary maps without map labels.

The verification team spent two days visiting selected sites by helicopter to become familiar with range site characteristics. Data collection was conducted in one day using a Cessna 208 (high wing) aircraft with a nominal crew of pilot, navigator, and two observers. Observers took turns recording alternate flight lines. Due to low overcast conditions and rain showers, data collection was limited to two thirds of the planned flight lines. In total, 297 observations were collected.

Aerial observations were compared with the preliminary mapping. A contingency table was prepared, as previously described (see Table 2). Overall agreement between the reference data set and the preliminary mapping was only 0.63 ± 0.05 (0.95 confidence limit).

Inspection of the off-diagonal values in Table 2 indicate the problems with this mapping. Range site 60 (Lichen Tussock Tundra) and range site 61 (Lichen Mountain Meadow) exhibit a wide range of confusion with low shrub tundra classes. Subsequent examination of the aerial photography revealed that the lichen component of this range site was strongly influenced by image quality, which varied considerably over the survey area. This error will be addressed by additional work in the field to separate lichen from shrub tundra units.

The remainder of the classification errors are distributed around the diagonal of the error matrix. Because these errors are primarily among similar classes, the interpreter may have had difficulty distinguishing between closely related range sites. This suggests that the level of classification detail desired from this scale of photography was too great. To test this possibility, similar range sites were grouped and the verification data retabulated. Table 3 shows the results, and increase in overall agreement to 0.75 ± 0.05 (0.95 confidence level). A good deal of the confusion is cleaned up by this grouping of mapping categories.

Because higher accuracy was desired at the more detailed level, additional field work was performed by the SCS to revise the preliminary mapping. Presently, the maps are being edited for final presentation. Additional verification will be required to establish the accuracy of the final map.

TABLE 3.
1983 VERIFICATION DATA
Grouped Range Site Error Matrix

	Grouped Range Sites	
20	tall shrub	20, 21, 22
30	low shrub	32, 34, 35
40	low shrub	41, 42, 43, 44
51	tidal marsh	51
50	herbacious	52, 54, 55, 56, 57
60	lichen tundra	60, 61
70	mat α cushion	79, 71, 72
91	burned tundra	91

	MAP Unit Categories							
	20	30	40	51	50	60	70	91
20	60	5	1					
30	9	22	12		1	1		
40		5	70		7	6		
51	1			4	1			
50		2	3		23		1	
60	1	2	11		1	26	5	
70							10	
91								7
Totals	71	36	97	4	33	33	16	7
Percent	85	61	72	100	70	79	63	100

Total Observations = 297
Correct Observations = 222
Overall Agreement = 0.75 ± 0.05 (0.95 confidence level)

DISCUSSION AND CONCLUSIONS

The verification procedure described here has evolved through several iterations of an operational range mapping project (George, 1985). This approach offers a rapid, low cost means to evaluate map accuracy over large geographic areas. Two features are attractive in comparison to verification methods using point data required by ground visits. First, a relatively large number of samples can be acquired in a short time (roughly 300 per day). Second, the sample unit is a map unit, rather than a point. Because a map unit is a spatial entity, which by definition has smaller spatial impurities (inclusions) contained within, it seems more appropriate to compare larger size areas directly.

The cost of this procedure, in comparison to the entire mapping project, is relatively low. Table 4 provides the breakdown of labor and aircraft time required to collect and evaluate the data presented here. To put these figures in perspective, a typical field season is six calendar weeks and requires 24 work-weeks labor and over 100 hours of helicopter time.

TABLE 4. 1983 VERIFICATION COSTS

Task	Labor	Aircraft Time
preparation	1 workday	
observer training	2 workdays	5 hrs helicopter
data collection	3 workdays	5 hrs fixed-wing
data reduction	2 workdays	
Total	8 workdays	10 hrs aircraft

This method of map verification deals primarily with classification error, that is, determining if a map unit is correctly labeled. Another source of error is from improper placement of boundaries between map units. Boundary line error is noticeable when sharply contrasting map categories are adjacent to one another. More often, the transition from one map category to another is gradual, and line placement is largely a matter of interpretation. The procedure presented here only recognizes boundary problems when they become so severe that they cause the identification of the map unit to change and, thus, become a classification error.

Much of the effort to develop this verification technique was focused on identifying and controlling factors that influence data collection from the air. The use of two observers causes a potential problem by biasing data collection if they don't record observations in a consistent fashion. In this study, the mean percent agreement between observers was compared and found not to vary significantly (George, 1985). More work needs to be done to explore observer reliability. Experiments should be performed to compare aerial observations to ground transects.

The sampling design of this study was driven by characteristics of the aircraft and observation technique. It does not conform exactly to any of the commonly used sampling schemes described by Berry and Baker (1968). In the data analysis, a random sample was assumed in order to calculate confidence limits. Depending on the orientation of the transects relative to the variability in the terrain, this assumption is questionable. The use of an unaligned systematic sample, suggested by Aronoff (1985) may eliminate bias from the sample selection. Further study of the sample design may suggest ways to improve sample selection and allow more rigorous and detailed data analysis.

Maps are made to satisfy definite informational needs, usually related to a specific resource problem. Map verification is a tool to help evaluate how well a map meets its design requirements. This tool may be applied by map makers as part of the mapping process, or it can be applied by map users to evaluate maps of unknown accuracy.

ACKNOWLEDGMENTS

This project was supported by the USDA Soil Conservation Service and the Vice Chancellor for Research and Advanced Studies, University of Alaska, Fairbanks. The author would like to thank Sharon L. Wall for her assistance in reviewing the manuscript.

REFERENCES

Aronoff, S., 1985. The minimum accuracy values as an index of classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 51(1):99–111.

Berry, B. J. L, and A. M. Baker, 1968. Geographic Sampling, in B.J.L. Berry and D.R. Marble (eds.), *Spatial Analysis—a reader in statistical geography*, Prentice Hall, pp. 91–100.

Dozier, J., and A. J. Strahler, 1983. Ground Investigations in Support of Remote Sensing. in R. Colwell (ed), *Manual of Remote Sensing*. American Society of Photogrammetry, Falls Church, Virginia, pp. 981–984.

George, T. H., 1985. *Verification of polygonal resource maps using aerial observations*. Masters Thesis, University of Alaska, Fairbanks.

Hay, A. M., 1979. Sampling designs to test land-use map accuracy. *Photogrammetric Engineering and Remote Sensing*, 45(4):529–533.

Shiflet, T. N., 1973. Range sites and soils in the United States. *Arid Shrublands—Proceedings of the Third Workshop of the United States/Australia Rangeland Panel*, Tucson, Arizona. pp. 26–33.

Snedecor, G.W., and W.G. Cochran, 1980. *Statistical Methods*, Seventh Edition. Iowa State University Press, p. 121.

Swanson, J. D., M. P. Kinney, and P. C. Scorup, 1983, Reindeer Range Vegetation Inventory Procedures. *Acta. Zool. Fennica* 175:39–41.

(Received and accepted 3 March 1986)