The New Camera Calibration System at the U. S. Geological Survey

Donald L. Light
U. S. Geological Survey, 511 National Center, Reston, VA 22092

ABSTRACT: Modern computerized photogrammetric instruments are capable of utilizing both radial and decentering camera calibration parameters which can increase plotting accuracy over that of older analog instrumentation technology from previous decades. Also, recent design improvements in aerial cameras have minimized distortions and increased the resolving power of camera systems, which should improve the performance of the overall photogrammetric process. In concert with these improvements, the Geological Survey has adopted the rigorous mathematical model for camera calibration developed by Duane Brown. An explanation of the Geological Survey's calibration facility and the additional calibration parameters now being provided in the USGS calibration certificate are reviewed.

INTRODUCTION

BEGINNING IN THE LATE 1950s, the U.S. Geological Survey (USGS) calibrated aerial cameras that were to be used by contractors on USGS projects (Bean, 1962). In recognition of this capability existing at the USGS, the responsibility for aerial camera calibration in the U.S. Government, except for the Department of Defense, was transferred from the National Bureau of Standards to the USGS on 1 April 1973 (Tayman, 1974). About this time, the USGS Optical Calibration Laboratory's multicollimator instrument was structurally upgraded and expanded to include a total of 53 collimators. This expansion would permit super-wide angle aerial cameras to be accommodated. The upgraded laboratory was located at Reston, Virginia, where it remains today. The laboratory team performs approximately 100 calibrations every year, and their records show that today's cameras have improved significantly in recent years.

Perhaps the National Aeronautics and Space Administration's Large Format Camera, built by Itek in the early 1980s, set a new standard for cameras with improved lenses and forward motion compensation. Following Itek's lead, the commercial aerial camera manufacturers—Wild Heerbrugg, Carl Zeiss Oberkochen, and Zeiss Jena—have made significant improvements in the performance of their cameras both in resolution and distortion. Table 1 indicates typical improvements over the years.

The techniques and procedures for calibrations at the USGS, including a users guide, have been well documented (Karen, 1968; Tayman, 1978; Tayman, 1984; Tayman and Ziemann, 1984; Tayman *et al.*, 1985). The aerial camera is the instrument that gathers the data necessary for subsequent photogrammetric processes and operations. It can be considered as a surveying instrument of great precision when properly calibrated. The metric characteristics and orientation of critical parts of the camera and their relations to one another must be determined by calibration before the photographic data can be used for precision work. Basically, these characteristics are (1) the focal length of the camera lens, (2) the radial and decentering distortions of the lens, (3) the resolution of the lens-film combination, (4) the position of the principal point with respect to the fiducial marks, and (5) the relative positions and distances between the fiducial marks.

Up to now, the USGS Calibration Laboratory has performed its calibration functions and compiled the report in the format given by Tayman (1984). The importance of quality and precision measurement has always prevailed in the lab, and relatively few changes have been made in the basic concept over the years. Now, in recognition of the improvements being introduced by the camera manufacturers, reduced distortion, increased resolution, forward motion compensation, and even stabilized mounts coupled with the photogrammetrist's capability, by means of computerized instrumentation to utilize the improvements, it is timely to apply a more rigorous theory in modeling the camera's parameters. A more rigorous model of the camera parameters will increase the accuracy obtainable in photogrammetric practice. The new analytical model is based on the projective equations outlined by Schmid (1953). This model for calibrating metric cameras has undergone continuous and significant development since Brown published his rigorous treatment of simultaneous determination of the exterior angular orientation, interior orientation, and symmetric radial lens distortion (Brown, 1956).

The original solution utilized a least-squares adjustment of the measured plate coordinates of stellar images taken with a ballistic camera. The star images were taken in support of the National Satellite Triangulation Program using stellar cameras to photograph the satellite against a background of stars (Case, 1955). Since then, significant advances were made by the introduction of a parameterized model of lens decentering distortion and Brown's (1966) refinement of the Conrady (1919) lens decentering distortion model. The culmination of the analytical camera calibration method was reached when DBA Systems issued a report titled *Advanced Methods for the Calibration of Metric Cameras* (Brown, 1968). Finally, a computer software program, *Simultaneous Multi-Camera Analytical Calibration* (SMAC), was produced under contract to the U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia, by Gyer *et al.* (1970). The software system to be utilized at the USGS and discussed in this paper is a modification of the SMAC Program adapted to the USGS multi-collimator instrument. The SMAC is being modified and installed on an IBM PS/2 Model 70 computer by DBA Systems, Inc. of Melbourne, Florida.

PRESENT SYSTEM'S DATA REDUCTION METHOD

The math model now in use is basically a least-squares version of the photogrammetric resection problem. The plate coordinates of the collimator images are the observations and the collimator directions are the known coordinates. Because the collimator image positions have been perturbed by lens distortions, the measured coordinates will be different from their true positions. In the reduction process the position of the interior perspective center and the orientation of the camera are allowed to adjust to minimize the differences between the observed and true image positions. The result is calibrated focal length, profile of the mean radial distortion, and the coordinates of the principal point. Thirty-three collimator images will appear on the

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING,
Vol. 58, No. 2, February 1992, pp. 185–188.

0099-1112/92/5802–185$03.00/0
©1992 American Society for Photogrammetry
and Remote Sensing

calibration plate for a standard 6-inch focal length (152.4 mm) aerial camera, giving a total of 66 observations.

Figure 1 illustrates the geometry of one bank of the USGS multi-collimator instrument. Figure 2 illustrates the distribution of the collimator positions that appear on the plate. Figure 3 illustrates the resolution target and the position reticle that is installed in each collimator.

TABLE 1. TYPICAL IMPROVEMENTS IN AERIAL CAMERAS

Camera Item	1960s	1990s
Resolution (AWAR)*	63 lp/mm	90 lp/mm
Radial Distortion	± 10 μm	± 3 μm

*AWAR is area weighted average resolution measured on a Kodak high resolution (400 lp/mm) micro-flat glass plate.

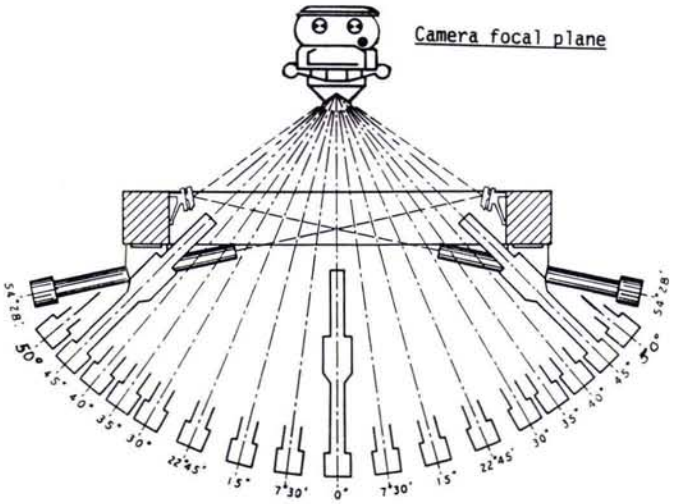

FIG. 1. Schematic diagram of one collimator bank.

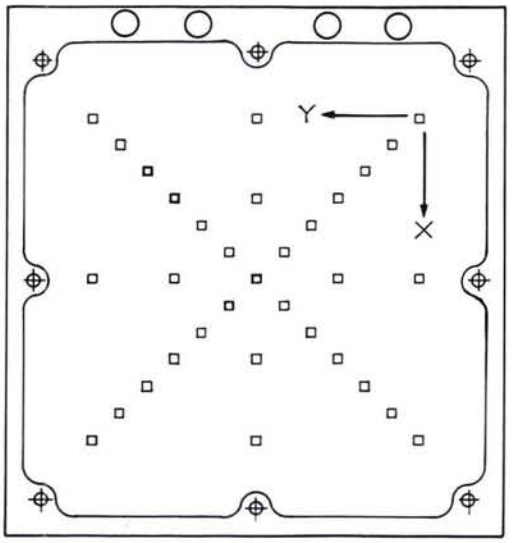

FIG. 2. Target positions that appear on glass plate.

RESECTION EQUATION AS A CALIBRATION MODEL

The well known projective equations of photogrammetry are employed in the present camera calibration system (Karen 1968). They are

$$x - x_p = f\left[\frac{A\lambda + B\mu + C\nu}{D\lambda + E\mu + F\nu}\right]$$
$$y - y_p = f\left[\frac{A'\lambda + B'\mu + C'\nu}{D\lambda + E\mu + F\nu}\right] \quad (1)$$

in which

x and y are the measured plate coordinates with respect to the photo coordinate system;

x_p, y_p, f are the coordinates of the principal point and focal length of the camera;

$\begin{bmatrix} A\ B\ C \\ A'B'C' \\ D\ E\ F \end{bmatrix}$ = orientation matrix elements which are functions of three independent angles α, ω, κ referred to an arbitrary X, Y, Z frame in object space; and

λ, μ, ν = X, Y, Z direction cosines of rays joining corresponding image and object points.

Summarizing the present method, it should suffice to point out that the data reduction program uses the linearized version of Equation 1 in the least-squares solution of the unknowns. The unknowns are

(x_p, y_p, f): the coordinates of the principal point and focal length of the camera, and

α, ω, κ: the three angles of exterior orientation. α and ω are approximately zero.

The residuals Δx and Δy on each point remaining after the least-squares adjustment are the basis for the radial distortion curve as reported in each USGS calibration report. In the actual reduction, the focal length is slightly adjusted to accommodate for the small shift in unknowns x_p, y_p away from the center of fiducials in the image plane.

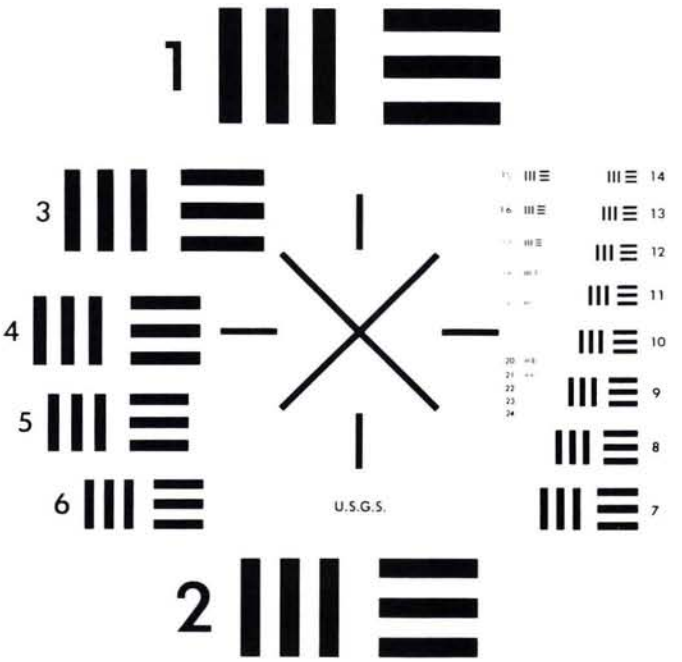

FIG. 3. Reticle/resolution target.

THE NEW ANALYTICAL MODEL WITH RADIAL AND DECENTERING LENS DISTORTION

The new model (Brown, 1966) incorporates Equation 1, the photogrammetric projective equations, and augments them with two parts: symmetric radial distortion and lens decentering distortion. Specifically, the analytical models for radial and decentering distortion are introduced directly into the projective equations. The parameters defining the distortion functions are then solved simultaneously with the projective parameters (α, ω, κ, x_p, y_p, f, λ, μ, ν) in a least-squares adjustment. There are 13 unknowns in the new model which rigorously define the camera parameters.

Models for Radial and Decentering Distortion

The distortion, δr, of a perfectly centered lens referred to the Gaussian focal length, f, can be expressed as

$$\delta r = K_1 r^3 + K_2 r^5 + K_3 r^7 + \ldots \tag{2}$$

in which K_1, K_2, K_3 are coefficients of radial distortion and r is the radial distance referred to the principal point.
More specifically,

$$r = [(x - x_p)^2 + (y - y_p)^2]^{1/2}$$

where x, y are the plate coordinates of the photographed collimator points and x_p, y_p are the coordinates of the principal point. The x-y components of radial distortion are given by

$$\begin{aligned} \delta x = \delta r \left(\frac{x}{r}\right) &= x(K_1 r^2 + K_2 r^4 + K_3 r^6 + \ldots) \text{ and} \\ \delta y = \delta r \left(\frac{y}{r}\right) &= y(K_1 r^2 + K_1 r^4 + K_3 r^6 + \ldots). \end{aligned} \tag{3}$$

The number of coefficients required to represent radial distortion with sufficient accuracy depends on the particular lens. Mapping lenses generally require two or three coefficients. Unusual lenses such as fisheye lenses may require several more, whereas a simple lens may require only a single coefficient.

After computing the appropriate coefficients, it is customary for practical production usage to transform the Gaussian distortion function (Equation 2) to a form in which the maximum positive and negative values of the distortion are equal. It was shown by Brown (1957) that a change, Δf, in the computed focal length results in a more useful distortion function given by

$$\begin{aligned} \delta r &= \frac{\Delta f}{f} + (K_1 + \frac{\Delta f}{f}) r^3 + (K_2 + \frac{\Delta f}{f}) r^5 + (K_3 + \frac{\Delta f}{f}) r^7 + \ldots \\ &= K'_0 + K'_1 r^3 + K'_2 r^5 + K'_3 r^7 + \ldots \end{aligned}$$

The K' values will be reported in the calibration report along with the associated focal length, f.

Decentering Distortion Model

Slight misalignments in lens assembly introduce what is termed decentering distortion. This distortion has both a radial and tangential component and can be described analytically by the expressions that follow (Brown, 1966):

$$\begin{aligned} \Delta x &= \left[1 + P_3^2 r^2\right]\left[P_1 (r^2 + 2x^2) + 2P_2 xy\right] \\ \Delta y &= \left[1 + P_3^2 r^2\right]\left[2P_1 xy + P_2 (r^2 + 2y^2)\right] \end{aligned} \tag{4}$$

where P_1, P_2, P_3 are unknown coefficients of decentering lens distortion. The higher order coefficient, P_3, only rarely proves to be significantly different from zero in today's aerial cameras.

The P_1, P_2, P_3 quantities in Equation 4 are actually parameterized terms of the actual decentering parameters Φ_0, J_1, J_2. These are introduced to simplify the computations (Livingston, 1980).

$$\begin{aligned} P_1 &= J_1 \sin \Phi_0 \\ P_2 &= J_1 \cos \Phi_0 \\ P_3 &= J_2/J_1 \end{aligned}$$

Reversal of these relations yields

$$\begin{aligned} \Phi_0 &= \arctan (P_1/P_2) \\ J_1 &= (P_1^2 + P_2^2)^{1/2} \\ J_2 &= (J_1)(P_3) \end{aligned}$$

where

Φ_0 = angle between the positive x axis and the axis of maximum tangential distortion, and
J_1, J_2 = coefficients of decentering distortion to be reported in the calibration report from the new system.

The Total SMAC Model

The model is composed of three parts: (1) projective equations of photogrammetry, Equation 1; (2) the radial distortion model, Equation 3; and (3) the decentering distortion model, Equation 4. Equation 5 represents the totally rigorous SMAC model. Table 2 shows a comparison of the current program output with the SMAC. Notice that the Ks and J_1, J_2, and Φ_0 are the additional values obtained by SMAC. Table 3 is a comparison of the calibrated values from an actual camera known to have large decentering distortion. Notice the 15 μm at 40° is reflected in the SMAC output. The current system offsets the x_p, y_p considerably as a result of the decentered lens, but it cannot directly account for the decentering as distortion. For the high quality cameras manufactured today, decentering distortion is generally expected to be near zero, but it will be defined. This may be a measure of the camera quality as well of precision in assembly of the lenses. However, in some state-of-the-art lenses decentering distortion may actually be significantly greater in magnitude than radial distortion. This may be largely due to computer lens designs where the designer may suppress radial distortion to a very low level, but the alignment and fabrication of the lens elements remains a tedious manufacturing step as in the past.

Table 2. Comparison of Program Outputs: Current vs New SMAC

Current Program	New SMAC Program	Definition
• x_p, y_p	x_p, y_p	Coordinates of Principal Point
• f	f	Calibrated focal length
• α, ω, κ	α, ω, κ	Orientation angles
• ▬	K'_0, K'_1, K'_2, K'_3	Coefficients of radial Distortion
• ▬	J_1, J_2, Φ_0	Parameters of Decentering Distortion
• Radial Dist. Table	Radial Dist. Table	Radial Distortion - Each Diagonal
• ▬	Decentering Distortion Table	Decentering Distortion - Each Diagonal

• All other parameters Resolution, Fiducial Coordinates, Shutter Efficiency, Stereomodel Flatness, etc., will remain as currently reported.

TABLE 3. CURRENT SYSTEM VS SMAC FOR A REAL CAMERA

• Angle	Current System Average Distortion	SMAC Radial Distortion	SMAC Decentering Distortion
7.50°	−8 μm	−5 μm	0 μm
15.00°	−8	−7	1
22.75°	−3	−4	2
30.00°	4	3	4
35.00°	3	7	6
40.00°	0	1	9
• Parameters			
x_p	21 μm	5 μm	–
y_p	−47	−21	–
f	152.599 mm	152.597 mm	–
K'_0	–	0.254×10^{-3}	–
K'_1	–	-0.553×10^{-7}	–
K'_2	–	0.241×10^{-11}	–
Φ_0	–	–	213°
J_1	–	–	0.558×10^{-6}

$K_3 = J_2 = 0$ (insignificantly different from zero by statistical test)

Simultaneous Multiframe Analytical Calibration - SMAC

• Calibration Math Model

$$x - x_p = f\left[\frac{A\lambda + B\mu + C\nu}{D\lambda + E\mu + F\nu}\right] + x\left[K_1r^2 + K_2r^4 + K_3r^6\right]$$

Projective Equation Radial Distortion

$$+\left[1 + P_3{}^2r^2\right]\left[P_1(r^2+2x^2)+2P_2xy\right] \quad (5)$$

Decentering distortion

$$y - y_p = f\left[\frac{A'\lambda + B'\mu + C'\nu}{D\lambda + E\mu + F\nu}\right] + y\left[K_1r^2 + K_2r^4 + K_3r^6\right]$$

$$+\left[1 + P_3{}^2r^2\right]\left[2P_1xy+P_2(r^2+2y^2)\right]$$

CONCLUDING REMARKS

In the past, many of the metric shortcomings of mapping cameras could be tolerated by virtue of the compensation provided by fairly dense networks of pre-established ground control. Establishing dense networks of control constitutes a major expense for mapping operations in both time and money. Mainly for this reason, coupled with advances in computer technology and photogrammetric instrumentation, the extension and densification of mapping control by means of block analytical aerotriangulation has gained widespread acceptance. Experiments by Brown (1966) have shown that the full promise of analytical methods depends in great measure on the precise calibration of the camera. This is because residual systematic errors propagate through analytical aerotriangulation in a most unfavorable manner.

It follows that the more comprehensive and more precise the calibration of the camera, the lower the requirements for absolute control in photogrammetric operations. The introduction of the more rigorous SMAC system at the USGS may be of fundamental importance to geodetic photogrammetry, to analytical photogrammetry, and to photogrammetric instrumentation nationwide.

The intent of this paper is to show the general equation form of the new calibration system and thereby to assist the users of the calibration report to better understand its value. The development team for the project is composed of Brad Johnson, George Schirmacher, Edward Cyran, and Donald L. Light of the USGS. Conversion of the original SMAC Code to the USGS system is being handled by Michael Babba and Thomas Riding of DBA Systems. Finally, testing and implementation of the system at the USGS is planned for completion by the end of 1991.

(Received 12 February 1991; accepted 18 April 1991)

REFERENCES

Bean, R. K., 1962. U. S. Geological Survey Camera Calibrator. Presented at the ACSM-ASP Annual Convention, Washington, D.C.

Brown, D. C., 1956. *The Simultaneous Determination of the Orientation and Lens Distortion of a Photogrammetric Camera,* AF Missile Test Center Technical Report No. 56–20, Patrick AFB, Florida.

———, 1957. *A Treatment of Analytical Photogrammetry with Emphasis on Ballistic Camera Applications,* AF Missile Test Center Technical Report No. 57–22, Patrick AFB, Florida.

———, 1966. Decentering Distortion and the Definitive Calibration of Metric Cameras, *Photogrammetric Engineering,* 42(5):xxx–xxx.

———, 1968. *Advanced Methods for the Calibration of Metric Cameras.* U.S. Army Engineer Topographic Laboratories Final Report under Contract DA-44-009-AMC-1457 (X), Fort Belvoir, Virginia.

Case, James B., 1955. Stellar Methods of Camera Calibration, Chapter 4 in *Manual of Photogrammetry,* 3rd Edition, American Society of Photogrammetry, Falls Church, Virginia, pp. 187–189.

Conrady, A., 1919. Decentered Lens Systems, *Monthly Notices of the Royal Astronomical Society,* 79:384–390.

Gyer, M. S., N. N. Haag, and S. K. Llewellyn, 1970. *Documentation for the Multi-Camera Stellar SMAC Computer Program.* U.S. Army Engineer Topographic Laboratories final report, DBA Systems.

Karen, R. J., 1968. Camera Calibration by the Multicollimator Method. *Photogrammetric Engineering,* 34(7):706–719.

Livingston, R. G. (ed.), 1980. Aerial Cameras, Chapter 4 in *Manual of Photogrammetry,* Fourth Edition (C. C Slama, editor). American Society of Photogrammetry, Falls Church, Virginia, pp. 261–263.

Schmid, H., 1953. *An Analytical Treatment of the Orientation of a Photogrammetric Camera,* Ballistic Research Laboratories Report No. 880. Aberdeen Proving Ground, Maryland.

Tayman, William P., 1974. Calibration of Lenses and Cameras at the USGS. *Photogrammetric Engineering,* 40(11):1331–1334.

———, 1978. Analytical Multicollimator Camera Calibration, *Photogrammetria,* 34:179–197.

———, 1984. Users Guide for the USGS Aerial Camera Report of Calibration. *Photogrammetric Engineering & Remote Sensing,* 50(5):577–584.

Tayman, W. P., and Hartmut Zieman, 1984. Photogrammetric Camera Calibration, *Photogrammetria,* 39:31–53.

Tayman, W. P., R. Vraga, and E. G. Schirmacher, 1985. *USGS Procedures Manual: Calibration of Aerial Mapping Cameras.* U. S. Geological Survey, Reston, Virginia.

Misclassification Bias in Areal Estimates

Raymond L. Czaplewski
USDA Forest Service, Rocky Mountain Forest and Range Experiment Station, 240 W. Prospect Street, Fort Collins, CO 80526–2098

ABSTRACT: In addition to thematic maps, remote sensing provides estimates of area in different thematic categories. Areal estimates are frequently used for resource inventories, management planning, and assessment analyses. Misclassification causes bias in these statistical areal estimates. For example, if a small percentage of a common cover type is misclassified as a rare cover type, then the area occupied by the rare type can be severely overestimated. Many categories are rare in detailed classification systems. I present an informal method to anticipate the approximate magnitude of this bias in statistical areal estimates, before a remote sensing study is conducted. If the anticipated magnitude is unacceptable, then statistical calibration methods should be used to produce unbiased areal estimates. I then discuss existing statistical methods that calibrate for misclassification bias with a sample of reference plots.

INTRODUCTION

REMOTELY SENSED AREAL ESTIMATES are typically treated as unbiased estimates of the true area for each cover type in a study area. However, Card (1982), Chrisman (1982), and Hay (1988) note that misclassification can bias areal estimates from remote sensing. My first objective is to demonstrate the cause of misclassification bias, and then to present an informal method to anticipate its magnitude before a remote sensing study is conducted. With a quantitative expectation of the approximate magnitude of this bias, the user of remotely sensed areal estimates can judge the practical importance of misclassification bias, given the unique requirements of each remote sensing study. If the anticipated magnitude of misclassification bias is unacceptable, then areal estimates should be calibrated with remotely sensed and reference classifications for a representative sample of reference plots from the study area. My second objective is to increase awareness of existing methods that can statistically calibrate for misclassification bias, and then to provide general guidance in the choice and application of an appropriate calibration method.

SOURCE OF MISCLASSIFICATION BIAS

Let the remotely sensed percentage of cover type A be denoted as the scalar Y, and the true percentage of cover type A be denoted as the scalar X. The true percentage of cover types other than A (labeled cover type B in the following) will equal the scalar $(100\% - X)$. Let scalar H_A represent the conditional probability that any pixel is interpreted as cover type A, given that the pixel is truly cover type A, where $0 \leq H_A \leq 1$; and let scalar $(1 - H_B)$ represent the conditional probability that any pixel is interpreted as cover type A, given it is truly type B, where $0 \leq H_B \leq 1$. H_A and H_B represent producer's accuracies, and are measures of omission error (Story and Congalton, 1986). The remotely sensed percentage (Y) of cover type A will be the following deterministic function of the true percentage (X) of cover type A and the true conditional probabilities of omission errors (H_A and H_B):

$$Y = [H_A X] + [(1 - H_B)(100\% - X)]. \quad (1)$$

Equation 1 shows that misclassification biases areal estimates from remote sensing; the remotely sensed percentage (Y) will not equal the true percentage (X) unless there are no omission errors, i.e., $H_A = H_B = 1$, or effects of omission errors exactly compensate, i.e., $(1 - H_B)(100\% - X) = (1 - H_A) X$. Either condition is rare in remote sensing.

Proportions or acreages of each cover type can be readily used in place of percentages in Equation 1. Instead of $(100\% - X)$, $(1 - X)$ would be used if X and Y are proportions, and $(T - X)$ would be used if X and Y are acreages, where T is the total acreage of the study area.

Assume classification accuracies are high for all cover types (e.g., $H_A = H_B = 0.95$). If cover type A truly occupies 90 percent of the study area (i.e., $X = 90$), then the remotely sensed percentage (Y) will equal 86 (see Equation 1). Similarly, Y equals 68 percent if X equals 70 percent, and Y equals 50 percent if X equals 50 percent. The bias in areal estimates for rare categories can be relatively high, even with such high classification accuracies. If cover type A truly occupies 10 percent of the study area, then the remotely sensed estimate will be 14 percent (see Equation 1). In this example, the remotely sensed percentage will be 40 percent larger than the true value. If a small percentage of a common cover type is misclassified as a rare cover type, then the area occupied by the rare type will be overestimated, unless there is a high rate of omission error in classifying the rare type. As the detail of a classification system increases, many categories will be rare. Figure 1 portrays the magnitude of misclassification bias for a wide range of classification accuracies.

MAGNITUDE OF MISCLASSIFICATION BIAS

Figure 1 or Equation 1 can be informally used to anticipate the approximate magnitude of misclassification bias for any cover type. However, preliminary expectations of classification accuracies and prevalence of various cover types must be used, rather than their true, but unknown, values. For example, assume you expect that classification accuracies for your study area will be similar to those given by Story and Congalton (1986), who used reference and remotely sensed classifications of 30 forested plots, 30 water plots, and 40 urban plots to construct an error matrix. From their error matrix, your preliminary estimate of producer's accuracy for the forest cover type in your study area is $H_A = 28/30 = 0.93$, and your preliminary estimate for non-forest accuracy is $H_B = (15+1+5+20)/(30+40) = 0.59$ (i.e., the water and urban types are pooled together). Assume your preliminary estimate of forest cover in your study area is 33 percent. Using Figure 1 with $H_A = 0.90$, $H_B = 0.60$, and $X = 33$, you anticipate misclassification bias will be approximately 25 percent; you can expect the remotely sensed areal estimate for forest will be $(33 + 25) = 58$ percent if forest cover is truly 33 percent in your study area. From the same error matrix, producer's accuracy for water cover is $H_A = 15/30 = 0.50$, and that for non-water (i.e., forest and urban) is $H_B = (28+1+15+20)/(30+40) = 0.91$. You can anticipate from Equation 1 that the remotely sensed areal estimate for water will be approximately $Y = 23$ percent if water truly covers $X = 33$ percent of your study area, i.e., misclassification bias of -10 percent. Finally, you can use Equation 1 and the same error matrix to anticipate that the remotely

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING,
Vol. 58, No. 2, February 1992, pp. 189-192.

0099-1112/92/5802–189$03.00/0
©1992 American Society for Photogrammetry
and Remote Sensing

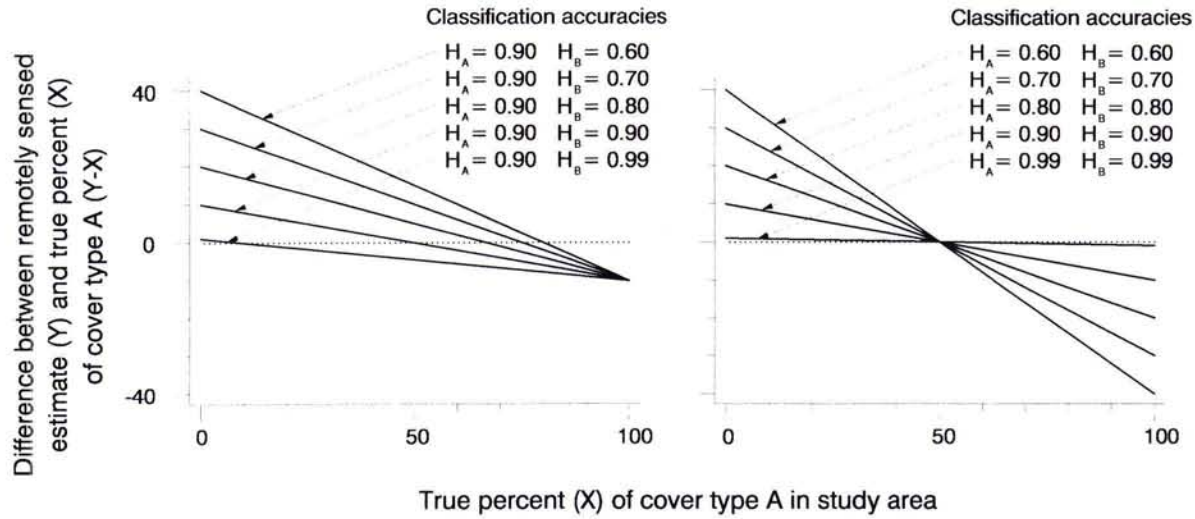

FIG. 1. Examples demonstrating magnitude of misclassification bias for various probabilities of omission errors. Misclassification bias is the difference between the remotely sensed estimate and the true percentage of a cover type. Remotely sensed estimate (*Y*) is a function (Equation 1) of prevalence of a cover type (*X*) and producer's classification accuracies (H_A and H_B), which are conditional probabilities of omission errors. H_A and H_B are sometimes equal in this example, but they are not necessarily equal in practice. This figure, or Equation 1, can be informally used to anticipate the magnitude of misclassification bias, given approximate expectations of classification accuracy and prevalence of cover types. If the anticipated magnitude is unacceptable, then more formal calibration methods should be considered, which are discussed in this paper.

sensed areal estimate for urban will be approximately 19 percent if urban areas truly cover 33 percent of your study area, i.e., misclassification bias of −14 percent.

CALIBRATION FOR MISCLASSIFICATION BIAS

If the anticipated magnitude of misclassification bias is unacceptable to the user, then more formal calibration techniques should be included in the remote sensing study. Calibration cannot identify misclassified pixels. Rather, calibration is a probabilistic technique; it uses proportions of imperfectly classified pixels in a reference sample to estimate conditional probabilities of various types of misclassification, and these estimated probabilities are then used to predict the true percentage of each cover type given the remotely sensed percentages. Proportions or acreages can be used in place of percentages. Statistical calibration requires accurate estimates of misclassification probabilities using reference plots from the study area, rather than preliminary expectations used above to anticipate the approximate magnitude of misclassification bias.

Two different calibration methods can be used if a reliable error matrix is acquired for a particular study. In remote sensing, Bauer *et al.* (1978), Maxim *et al.* (1981), Prisley and Smith (1987), and Hay (1988) demonstrate use of a classical multivariate calibration method, which was introduced into the statistical literature by Grassia and Sundberg (1982). Equation 1 is a univariate example of this method. To produce an unbiased areal estimate, Equation 1 is solved for the true percentage (*X*) given the remotely sensed estimate (*Y*), and accurate estimates of the probabilities of omission errors (i.e., H_A and H_B). The inverse calibration estimator of Tenenbein (1972) is an alternative to this classical estimator, and Card (1982) and Chrisman (1982) demonstrate use of this estimator in remote sensing. The inverse estimator uses probabilities of commission errors (i.e., user's accuracies), while the classical estimator uses probabilities of omission errors. Czaplewski and Catts (1990) give examples of these two calibration methods in remote sensing.

Based on unpublished Monte Carlo simulations, I found the inverse calibration estimator of Tenenbein is less biased, more precise, and less prone to numerical problems and infeasible solutions, especially for small sample sizes of reference plots. For example, the multivariate classical estimator requires a matrix inversion, and can produce negative areal estimates; the inverse estimator requires less complex algebra, and will always produce positive estimates.

These calibration techniques use misclassification probabilities from an error matrix that are estimated with a finite sample of reference plots from the study area. These estimated probabilities contain sampling errors, which are propagated into estimation errors for the true percentage of each category in the study area. As the sample size of reference plots increases, the sampling error deceases for estimates of misclassification probabilities, and accuracy of the calibrated areal estimate increases. Grassia and Sundberg (1982) and Tenenbein (1972) give approximate covariance matrices for these estimation errors; they assume a large sample of reference plots is available, and the reference plots are independent and homogeneous (i.e., each independent reference plot is classified into a single category with remote sensing, and a single category with the reference data). These covariance matrices are needed to construct confidence intervals, which describe the level of uncertainty in the calibrated areal estimate that is produced by uncertain estimates of misclassification probabilities.

An unstratified sample of homogeneous reference plots will include a small number of rare cover types. Stratification can provide more intensive sampling of rare types, which can improve accuracy of calibrated areal estimates for rare types. However, an inappropriate calibration technique can bias calibrated areal estimates from a stratified sample of reference plots. In general, the inverse estimator of Tenenbein should be used if stratification is based on the remotely sensed classifications. If the stratified sample is selected based on the reference classifications, then the classical estimator of Grassia and Sundberg should be used. This latter situation might exist if existing field plots are used for reference data, but the cost of accurate registration of existing field plots to the remotely sensed imagery limits the number of plots that can be registered. Bias from an inappropriate calibration technique can be eliminated with independent ancillary estimates of the true or remotely sensed

percentages in the study area, but more elaborate calibration methods are required.

These calibration methods require a large sample of representative reference plots to estimate misclassification probabilities. Representative reference plots are best selected with randomization methods. Training or labeling plots often have lower rates of classification error than are typical for the entire study area, and such plots will produce biased areal estimates if used for calibration. Brown (1982) discusses controlled calibration, which can use purposefully selected reference plots; however, this requires Bayesian estimation, which is vulnerable to subtle problems and undetected biases.

If the reference plots are not well registered to the remotely sensed imagery, then classification error will be confounded with registration error, and the misclassification probabilities will be poorly estimated. Large heterogeneous reference plots might be more successfully registered to remote sensing imagery than small homogeneous plots. If the heterogeneous plots are a simple random or systematic sample, and reference classifications are available for each pixel in these plots, then the calibration methods of Tenenbein (1972) and Grassia and Sundberg (1982) can be applied without modification. However, classification errors for adjacent pixels in the same reference plot are not independent, and different methods would be required to calculate the estimation error covariance matrix.

Different calibration methods are required if an error matrix cannot be constructed from the reference data. For example, reference data for agricultural surveys can be limited to areal estimates of different crop covers within large, heterogeneous plots; maps showing the location of each crop cover within the reference plots might not be available. Therefore, the reference classifications for each pixel within the reference plots are not available, and an error matrix cannot be constructed. Here, calibration can only use the remotely sensed and reference percentages of each cover type within each heterogeneous reference plot. Calibration estimators for this situation have been developed and evaluated by Chhikara *et al.* (1986), Fuller (1986), Heydorn and Takacs (1986), McKeon and Chhikara (1986), Hung and Fuller (1987), Battese *et al.* (1988), and Chhikara and Deng (1988). Similar situations arise when registration of pixels to reference plots is problematic, and reference classifications for individual pixels cannot be reliably obtained. Iverson *et al.* (1989) consider this situation for AVHRR data, where remotely sensed estimates for Landsat scenes serve as the reference data. In addition, Pech *et al.* (1986) describe a method to calibrate areal estimates from mixed pixels that cannot be classified into unique categories. All these methods are linear regression techniques rather than the probabilistic techniques of Tenenbein (1972) and Grassia and Sundberg (1982). Calibration based on regression methods can produce negative areal estimates. Lewis and Odell (1971), Liew (1976), and Shim (1983) propose quadratic programming techniques to avoid negative estimates, and van Roessel (USDA Forest Service, 1980) has applied this solution in remote sensing. Detection limits can affect misclassification bias in more complex ways. For example, an AVHRR pixel might require 30 percent deforestation before any deforestation can be detected. This can cause a nonlinear relationship between the remotely sensed and reference areal estimates for the reference plots, which might require nonlinear calibration estimators. Scheffé (1973) and Brown (1982) discuss the statistical aspects of nonlinear calibration, but this technique has not been applied in remote sensing.

All of these calibration methods correct areal estimates for misclassification error, despite the cause. Interpretation error is the most familiar cause. However, changes in land cover might occur between the dates that remotely sensed images and reference data were acquired, or there might be differences in definitions between the remote sensing and reference classification systems. Calibration treats the reference data as the standard, and calibrated areal estimates represent the acquisition dates definitions and protocol used for the reference data. For example, if the remotely sensed images were acquired in 1987 and the reference data in 1991, then the calibrated areal estimates are an unbiased estimate of the status in 1991. If users require areal estimates consistent with their existing definitions and protocol for field surveys, but other methods are used for the reference data in calibration (e.g., photointerpretation of large-scale imagery, or "windshield surveys"), then the calibrated areal estimates can be unacceptable to the user.

All of these calibration techniques are closely related to various multi-stage or multi-phase sampling designs, which can be more efficient than calibration if the sample size of reference plots is large. The remotely sensed data are analogous to the first level of a multi-level design, and the reference data are analogous to the second level. However, calibration methods have been developed that use areal estimates from all pixels in an image, and for multivariate and nonlinear situations; calibration might be more readily applied to these more complicated estimation problems than multi-level sampling designs.

CONCLUSIONS

Some users reject areal estimates from remote sensing because the magnitude of misclassification bias might be large. Some remote sensing specialists recommend that users ignore misclassification bias if classification accuracy is high. The most reasonable alternative might lay between these extremes. During the planning stage, remote sensing specialists should anticipate the approximate magnitude of misclassification bias. If the anticipated magnitude is unacceptable to the user of remotely sensed areal estimates, then the study plan should require statistical methods that will calibrate the final areal estimates. Reliable calibration requires an adequate, representative, and timely sample of accurately registered reference data from the study area.

REFERENCES

Battese, G. E., R. M. Harter, and W. A. Fuller, 1988. An error-component model for prediction of county crop areas using survey and satellite data. *Journal of the American Statistical Association*. 83:28–36.

Bauer, M. E., M. M. Hixson, B. J. Davis, and J. B. Etheridge, 1978. Area estimation of crops by digital analysis of Landsat data. *Photogrammetric Engineering & Remote Sensing*. 44:1033–1043.

Brown, P. J., 1982. Multivariate calibration. *Journal of the Royal Statistical Society, B*. 3:287–321.

Card, D. H., 1982. Using known map categorical marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering & Remote Sensing*. 48:431–439.

Chhikara, R. S., J. C. Lundgren, and A. A. Houston, 1986. Crop acreage estimation using a Landsat-based estimator as an auxiliary variable. *International Transactions in Geoscience and Remote Sensing*. 24:157–168.

Chhikara, R. S., and L. Y. Deng, 1988. Conditional inference in finite population sampling under a calibration model. *Communications in Statistics - Simulation and Computation*. 17:663–681.

Chrisman, N. R., 1982. Beyond accuracy assessment: correction of misclassification. *Proceedings of the 5th International Symposium on Computer-assisted Cartography*, Crystal City, Virginia. pp.123–132.

Czaplewski, R. L., and G. P. Catts, 1990. Calibrating area estimates for classification error using confusion matrices. *Proceedings of the 56th Annual Meeting of the American Society for Photogrammetry and Remote Sensing*, Denver, Colorado, Volume 4. pp. 431–440.

Fuller, W. S., 1986. Small area estimation as a measurement error problem. *Proceedings of the Conference on Survey Research Methods in Agriculture*, Leesburg, Virginia, 15–18 June.

Grassia, A., and R. Sundberg, 1982. Statistical precision in the calibra-

tion and use of sorting machines and other classifiers. *Technometrics*. 24:117–121.

Hay, A. M., 1988. The derivation of global estimates from a confusion matrix. *International Journal of Remote Sensing*. 9:1395–1398.

Heydorn, R. P., and H. C. Takacs, 1986. On the design of classifiers for crop inventories, *IEEE Transactions in Geoscience and Remote Sensing*. 24:150–155.

Hung, H. M., and W. A. Fuller, 1987. Regression estimation of crop acreages with transformed Landsat data as auxiliary variables. *Journal of Business and Economic Statistics*. 5: 475–482.

Lewis, T. O., and P. L. Odell, 1971. *Estimation in Linear Models*. Prentice-Hall, New Jersey, 193 p.

Liew, C. K., 1976. Inequality constrained least-squares estimation. *Journal of the American Statistical Association*. 71:746–751.

Iverson, L. R., E. A. Cook, and R. L. Graham, 1989. A technique for extrapolating and validating forest cover across large regions, calibrating AVHRR data with TM data. *International Journal of Remote Sensing*. 10:1805–1812.

Maxim, L. D., L. Harrington, and M. Kennedy, 1981. Alternative scale-up estimates for aerial surveys where both detection and classification error exist. *Photogrammetric Engineering & Remote Sensing*. 47:1227–1239.

McKeon, J. J., and R. S. Chhikara, 1986. Crop acreage estimation using satellite data as auxiliary information: multivariate case. *Proceedings of the Survey Research Methods Section, American Statistical Association*, University of Houston, Clear Lake, Texas.

Pech, R. P., A. W. Davis, R. R. Lamacraft, and R. D. Graetz, 1986. Calibration of Landsat data for sparsely vegetated semi-arid rangelands. *International Journal of Remote Sensing*. 7:1729–1750.

Prisley, S. P., and J. L. Smith, 1987. Using classification error matrices to improve the accuracy of weighted land-cover models. *Photogrammetric Engineering & Remote Sensing* 53:1259–1263.

Scheffé H., 1973. A statistical theory of calibration. *American Statistician*. 1:1–37.

Shim, J. K., 1983. A survey of quadratic programming applications to business and economics, *International Journal of Systems Science*. 14:105–115.

Story, M., and R. G. Congalton, 1986. Accuracy assessment: a user's perspective. *Photogrammetric Engineering & Remote Sensing*. 52:397–399.

Tenenbein, A., 1972. A double sampling scheme for estimating from misclassified multinomial data with applications to sampling inspection. *Technometrics* 14:187–202.

USDA Forest Service, 1980. *Evaluation of Multiresource Analysis and Information System (MAIS) Processing Components, Kershaw County South Carolina Feasibility Test*, Berkeley, California. 96 p.

(Received 27 June 1990; revised and accepted 19 March 1991)

Forthcoming Articles

Paul V. Bolstad and *T. M. Lillesand*, Rule-Based Classification Models: Flexible Integration of Satellite Imagery and Thematic Spatial Data.

John Crews, Overplotting Digital Geographic Data onto Existing Maps.

Claude R. Duguay and *Ellsworth F. LeDrew*, Estimating Surface Reflectance and Albedo from Landsat-5 Thematic Mapper over Rugged Terrain.

Peter F. Fisher, First Experiments in Viewshed Uncertainty: Simulating Fuzzy Viewsheds.

Steven E. Franklin and *Bradley A. Wilson*, A Three-Stage Classifier for Remote Sensing of Mountain Environments.

Clive S. Fraser, Photogrammetric Measurement to One Part in a Million.

Clive S. Fraser and *James A. Mallison*, Dimensional Characterization of a Large Aircraft Structure by Photogrammetry.

Clive S. Fraser and *Mark R. Shortis*, Variation of Distortion within the Photographic Field.

Peng Gong and *Philip J. Howarth*, Frequency-Based Contextual Classification and Gray-Level Vector Reduction for Land-Use Identification.

Christian Heipke, A Global Approach for Least-Squares Image Matching and Surface Reconstruction in Object Space.

David P. Lanter and *Howard Veregin*, A Research Paradigm for Propagating Error in Layer-Based GIS.

Richard G. Lathrop, Jr., Landsat Thematic Mapper Monitoring of Turbid Inland Water Quality.

Kurt Novak, Rectification of Digital Imagery.

J. Olaleye and *W. Faig*, Reducing the Registration Time for Photographs with Non-Intersecting Crossarm Fiducials on the Analytical Plotter.

Albert J. Peters, Bradley C. Reed, and *Donald C. Rundquist*, A Technique for Processing NOAA)AVHRR Data into a Geographically Referenced Image Map.

Kevin P. Price, David A. Pyke, and *Lloyd Mendes*, Shrub Dieback in a Semiarid Ecosystem: The Integration of Remote Sensing and Geographic Information Systems for Detecting Vegetation Change.

Benoit Rivard and *Raymond E. Arvidson*, Utility of Imaging Spectrometry for Lithologic Mapping in Greenland.

Kathryn Connors Sasowsky, Gary W. Petersen, and *Barry M. Evans*, Accuracy of SPOT Digital Elevation Model and Derivatives: Utility for Alaska's North Slope.

Omar H. Shemdin and *H. Minh Tran*, Measuring Short Surface Waves with Stereophotography.

Vittala K. Shettigara, A Generalized Component Substitution Technique for Spatial Enhancement of Multispectral Images Using a Higher Resolution Data Set.

Michael B. Smith and *Mitja Brilly*, Automated Grid Element Ordering for GIS-Based Overland Flow Modeling.

David M. Stoms, Frank W. Davis, and *Christopher B. Cogan*, Sensitivity of Wildlife Habitat Models to Uncertainties in GIS Data.

Khagendra Thapa and *John Bossler*, Accuracy of Spatial Data Used in Geographic Information Systems.

Thierry Toutin, Yves Carbonneau, and *Louiselle St-Laurent*, An Integrated Method to Rectify Airborne Radar Imagery Using DEM.

Paul M. Treitz, Philip J. Howarth, and *Peng Gong*, Application of Satellite and GIS Technologies for Land-Cover and Land-Use Mapping at the Rural-Urban Fringe: A Case Study.

William S. Warner and *Øystein Andersen*, Consequences of Enlarging Small-Format Imagery with a Color Copier.

Zhuoqiao Zeng and *Xibo Wang*, A General Solution of a Closed Form Space Resection.