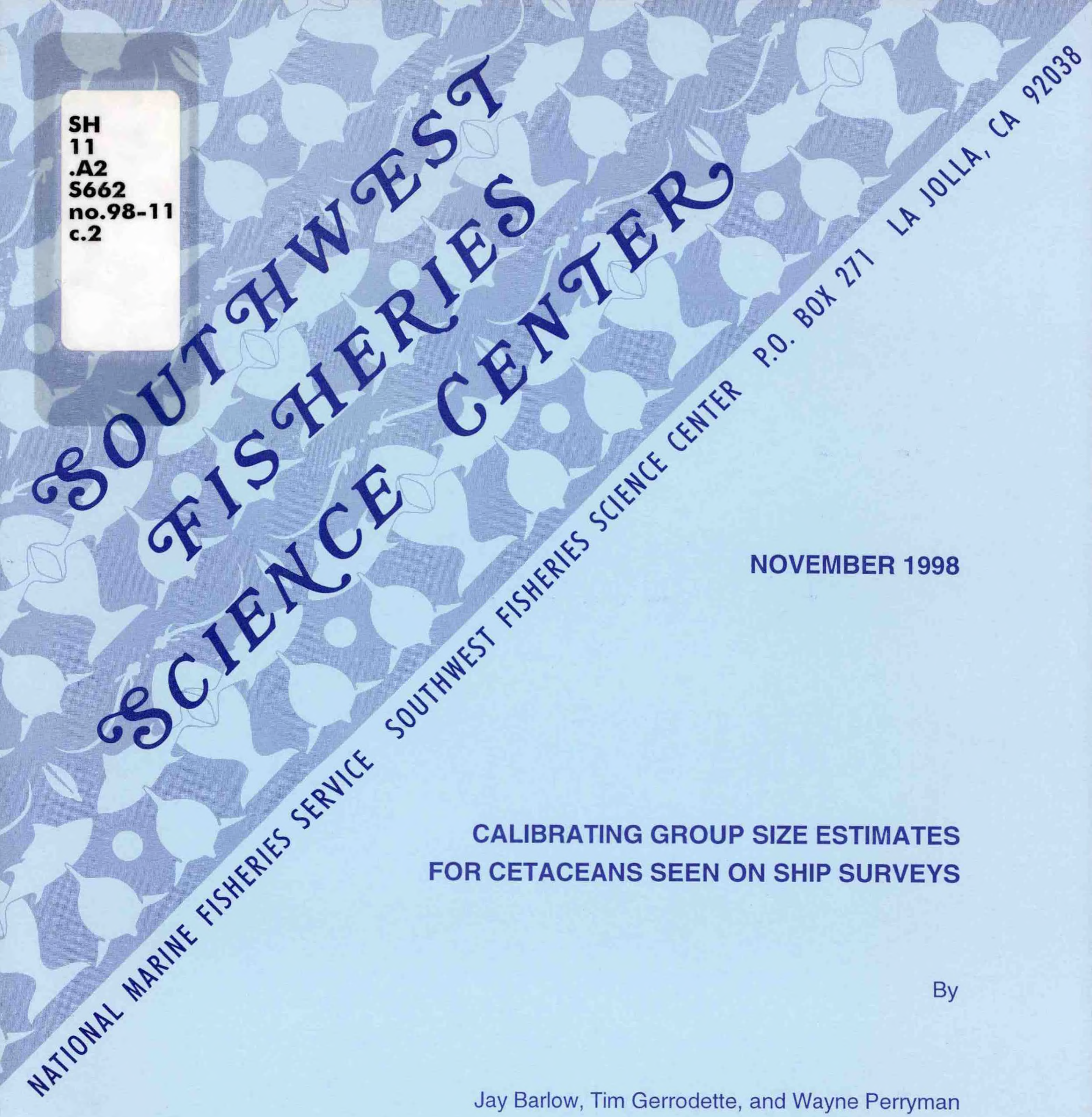


SOUTHWEST FISHERIES SCIENCE CENTER

NATIONAL MARINE FISHERIES SERVICE SOUTHWEST FISHERIES SCIENCE CENTER P.O. BOX 271 LA JOLLA, CA 92038

NOVEMBER 1998

CALIBRATING GROUP SIZE ESTIMATES FOR CETACEANS SEEN ON SHIP SURVEYS

By

Jay Barlow, Tim Gerrodette, and Wayne Perryman

ADMINISTRATIVE REPORT LJ-98-11

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
NOAA
U.S. DEPARTMENT OF COMMERCE

CALIBRATING GROUP SIZE ESTIMATES FOR CETACEANS SEEN ON SHIP SURVEYS

Jay Barlow, Tim Gerrodette, and Wayne Perryman

Southwest Fisheries Science Center
National Marine Fisheries Service, NOAA
8604 La Jolla Shores Drive
La Jolla, California 92037

NOVEMBER 1998

ADMINISTRATIVE REPORT LJ-98-11

CALIBRATING GROUP SIZE ESTIMATES
FOR CETACEANS SEEN ON SHIP SURVEYS

Jay Barlow
Tim Gerrodette
Wayne Perryman
Southwest Fisheries Science Center
P.O. Box 271, La Jolla CA 92038

ABSTRACT

In estimating cetacean abundance from ship line-transect surveys, group size is generally assumed to be estimated without error. To test this assumption, estimates of groups size were made from aerial photographs taken from a helicopter during cetacean ship surveys. We estimate calibration factors for ship-based observers by regression using these aerial counts as estimates of true group size. Methods differed from previous analyses by Gerrodette and Perrin (1991) in that (1) more regression models were considered for describing the relationship between estimated and true group sizes including one model based on regression through the origin, (2) covariates (such as year and sea state) were assumed to affect the slope of the log-regression of estimated group size on true group size (rather than the intercept), (3) data from the years of 1992 and 1993 were added to the analyses, and (4) bias correction factors were used where applicable. Methods were extended to more observers and species than were included in the earlier paper. Calibration factors were estimated for 34 individuals based on a cross-validation method. Calibrated estimates of group size were less biased and more precise than uncalibrated estimates. For most individuals, a simple model using regression through the origin was optimal. For others, the model fit was improved by adding an intercept, a Beaufort sea state term, or year-specific terms.

Aerial photographic estimates of group size are only available for a small subset of groups seen in fair weather conditions. To determine whether estimation methods are consistent for the majority of other groups, an indirect calibration method was used. Indirect calibration factors were estimated by regressing the uncalibrated estimates of each observer against the calibrated estimates of the other observers. A high correlation was found between direct and indirect calibration factors for each observer, indicating that observers are consistent in their tendencies to over or underestimate group size. Indirect calibration factors were estimated for surveys conducted in 1986, 1991, and 1996 which did not use a helicopter to photographically estimate group size.

INTRODUCTION

On ship surveys for cetaceans, group size is typically estimated by ship-board observers viewing the group through binoculars or by naked eyes. On Southwest Fisheries Science Center (SWFSC) surveys, the ship is typically diverted to the vicinity of the group to facilitate species determination and group size estimation. The amount of time spent estimating group size and the distance from which group size is estimated varies with the size and behavior of the group. At any one time, most of the individuals in a group are submerged, making it impossible to simply count the number of individuals in the group. Although group size estimation may involve counting, ship-board observers typically make a *gestalt* estimate based on many factors including the number of animals they actually saw, the behavior of the group, sighting conditions, and what they know from past experience.

Better estimates of group size can be obtained in some cases by photographing a dolphin group from an aircraft and counting the number of dolphins in the image (Scott et al. 1985; Gilpatrick 1993). Beginning in 1987, a helicopter was used to photograph a subset of dolphin groups seen on SWFSC cetacean surveys in the eastern tropical Pacific. Aerial photographic estimates of group size have been shown to be precise (CV ≈ 0.05, Gilpatrick 1993) and, if all animal in the group are within the image, are believed to be unbiased. Using these aerial photographic counts as estimates of true group size, Gerrodette and Perrin (1991) estimated individual calibration factors, β, for shipboard observers by $\log_{10}$ transformed least-squares regression:

$$\log_{10} N = \beta_0 \log_{10} S + \beta_1 + \beta_2 x_2 + \sum_{y=87}^{90} \beta_y x_y \qquad \textbf{(1)},$$

where N = observer' s estimate of group size,
S = true (aerial) estimate of group size, and
x_i = categorical variables for Beaufort sea state (x_2) and year ($x_{87\text{-}90}$).

They explored two cases of calibrated estimates: (1) with slope and intercept only (β_0, β_1) and (2) with slope, intercept, and covariate terms. The observer's estimate of group size, N, was calculated as a weighted sum of the observer's best, high, and low estimates (*B, H, L,* respectively) of group size

$$N = w_1 \cdot B + w_2 \cdot H + w_3 \cdot L \qquad \textbf{(2)},$$

where the optimal weights, w_i, were calculated by minimizing the residual sum of squares in a regression of $\log_{10} N$ against the logarithm of the true (aerial) group size, $\log_{10} S$. Weights were estimated interatively using the simplex algorithm (Press et al. 1988).

Gerrodette and Perrin (1991) found that estimates from many observers could be improved by using calibration factors. This study updates the analysis of Gerrodette and Perrin (1991) by including aerial group size estimates from two additional survey years (1992 and 1993), by exploring alternative regression models, and by incorporating bias correction factors where necessary.

Aerial estimates of group size are only available for a small subsample of cetacean groups, and this sample may be biased because group size can be estimated from aerial photographs only in calm conditions with clear water. To determine whether direct calibration factors from aerial photographs are generally applicable to the majority of groups, we examine whether the relative differences among observers are consistent over the entire sample. Using the sample of groups without aerial photographs, we estimate indirect calibration factors by regression using the mean of calibrated estimates from other observers as an estimate of true group size. This method of indirect calibration (Barlow 1995) is also used to estimate calibration factors for observers on surveys which did not have a helicopter (in the eastern tropical Pacific in 1986; off California in 1991, 1993, and 1996; and off Oregon and Washington in 1996).

METHODS

Field Methods

The SWFSC conducted ship line-transect surveys for cetaceans: (1) in the eastern topical Pacific in 1986-1990 (Gerrodette and Perrin 1991; Wade and Gerrodette 1993); (2) off California in 1991 (Hill and Barlow 1992; Barlow 1995); (3) off the west coast of Central America in 1992 (Mangel and Gerrodette 1994a); (4) off California, Baja California, and in the Gulf of California in 1993 (Mangel and Gerrodette 1994b); and (5) off California, Oregon, and Washington in 1996 (VonSaunder and Barlow, in prep.). The primary purpose of these surveys was to estimate the population size of dolphins and other cetacean species that are caught incidentally in commercial fisheries. A team of 3 observers searched simultaneously, two using 25X pedestal-mounted binoculars and one using naked eyes and a 7x binocular, and off-duty observers were rotated into this watch every two hours. Group size was typically estimated independently by all three on-duty observers if all obtained adequate views of the group. When the groups were being photographed by helicopter, off-duty observers were also called upon to make group size estimates. Observations by off-duty personnel might not have been as long or as complete as observations by on-duty observers, but previous analyses showed that duty status did not affect group size estimation (Gerrodette and Perrin 1991). Observers are asked to provide their "best" estimate of the number of individuals present in a group as well as a "high" and a "low" estimate corresponding to their maximum and minimum estimates of the number of animals present. Observers' estimates were recorded confidentially in personal notebooks and were transcribed into a master record by the cruise leader each evening; observers were strongly discouraged from discussing group size estimates among themselves. Prior to surveys, observers were given suggestions on how to best estimate group size, including the recommendation to count subgroups (e.g. in tens, fifties, or hundreds).

We photographed a sub-sample of tropical dolphin groups from a Hughes 500D helicopter that was carried aboard the NOAA Ship *David Starr Jordan*. The photographs were taken with a KA-62 military reconnaissance camera that we mounted vertically below the fuselage of the helicopter. This camera has a 76 mm lens and a motion compensation system which eliminates the blurring of photographs caused by the forward motion of the aircraft while the shutter is open. We photographed the dolphin groups from altitudes between 200 to 300 m with Kodak Aerial Ektachrome (2448) film. Cycle rate of the camera was adjusted to ensure 80% overlap between adjacent frames. The number of photographic passes over each group varied with the group configuration, size, and the behavior of the dolphins. Safety constraints limited helicopter operations to the lower end of the sea state spectrum (generally Beaufort 3 or less).

We reviewed the transparencies of each group through a dissecting scope mounted above a light table and selected the groups for which the entire aggregation of dolphins was captured within the sequential images (Gilpatrick 1993). For these groups we carefully reviewed the photographs from each pass and selected the pass in which the largest number of dolphins were clearly visible just below the surface. We counted the dolphins by attaching an acetate overlay on the image and marking the location of each dolphin with a fine tipped pen. After all the dolphins present in one image had been marked, the overlay was attached to the adjacent overlapping image, where the marks were checked and dolphins that were not detected in the previous frame were added. For each group, this procedure was completed independently by three individuals. Following the procedures of Gilpatrick (1993), if the coefficient of variation (CV) of the three aerial counts was greater than 0.10, the group was recounted. If the CV of the subsequent counts was greater than 0.15, the group was excluded from consideration here. We used the mean of the three independent counts as the photographic estimate of group size. Group size was determined for 46, 49, 39, 37, 48, and 58 groups, respectively, for the years 1987, 1988, 1989, 1990, 1992, and 1993.

Direct Calibration Methods

Like Gerrodette and Perrin's analysis, we estimated calibration factors directly by comparison with the "known" group size from aerial photographs. We used a different calibration model in which the covariates affect the slope rather than the intercept of a linear regression using natural logarithms:

$$\ln N = \left(\beta_0 + \beta_2 x_2 + \sum_{y=87}^{93} \beta_y x_y \right) \cdot \ln S + \beta_1 \qquad \textbf{(3).}$$

This model is more appropriate because the intercept might logically be close to zero (if there is only one animal, most observers would estimate one), but observers might still be estimating proportionately higher or lower in certain conditions or years. The parameters for this model were fit iteratively using the simplex algorithm (Press et al. 1988) to minimize the residual sum of squares. Eighteen permutations of this method were investigated for each observer by including or excluding various elements of the above formulae (Table 1). Regression through the origin was investigated by excluding β_1. The importance of Beaufort sea state effects were investigated by including β_2. Year effects were tested by including all year terms $\beta_{87\text{-}93}$ (and, for clarity of presentation, by setting $\beta_0=0$). We explored using only the observers' best estimates of group size *($N = B$)* and using Gerrodette and Perrin's approach of estimating a weighted sum of observers' best, high, and low estimates.

The logarithm of the expected ("calibrated") value of group size *i* was calculated as:

$$\ln\left(E(S_i)\right) = \frac{(\ln N_i - \beta_1)}{(\beta_0 + \beta_2 x_2 + \sum_{y=87}^{93} \beta_y x_y)} \qquad \textbf{(4).}$$

Cross validation was used to determine the most appropriate regression model (Gerrodette and Perrin 1991). Each group of cetaceans that was estimated by an observer was iteratively eliminated from the sample, calibration factors for that individual were calculated based on all other groups (Eq. 3), and the log-size of the eliminated group was estimated based on these calibration factors (Eq. 4). For each regression model and each observer, the average squared prediction error for their n estimates of log-group size was calculated as

$$ASPE = \frac{\sum_{i=1}^{n} (\ln S_i - \ln E(S_i))^2}{n} \qquad \textbf{(5),}$$

where $E(S_i)$ is estimated using a regression that excluded the i-th datum. For each observer, the regression model giving the lowest ASPE was chosen as the most appropriate model.

Indirect Calibration Methods

Calibration factors were estimated indirectly by comparison with the calibrated estimates of other observers. Because a simple model worked well for the direct calibration of most observers (see Results) and because indirect calibration is inherently less precise than direct calibration, we only used a simple log-regression model without an intercept or Beaufort covariate term:

$$\ln N = \beta_0 \ln \bar{S} \qquad \textbf{(6),}$$

where N = observer's "best"estimate of group size, and

$\bar{S}$ = mean of calibrated, bias-corrected estimates for all other calibrated observers.

To avoid circular inference, the groups used to estimate direct calibration factors were excluded when estimating indirect calibration factors. Groups were included only if group size was estimated by at least two "calibrated" observers. Outliers (29 out of 8,798 groups) were eliminated if any two observers' estimates differed by more than an order of magnitude. We estimated indirect calibration factors, β_0, separately for each observer and each year if the number of groups estimated jointly with other "calibrated observers" was greater than 10. If a calibration factor was not available for a "calibrated observer" in a given year (such as in 1986, 1991, or 1996 when no helicopter was used), calibrated estimates were made using the overall calibration factors (ie., excluding the year covariates).

Bias Correction

Regression methods commonly assume that measurement error does not exist for independent (x) variables. Measurement error in the independent variable reduces the slope and increases the intercept in linear regressions (Fuller 1987), and the bias in the slope can be estimated as

$$Bias = \frac{\sigma^2_x + \sigma^2_{x-err}}{\sigma^2_x} \qquad \textbf{(7)},$$

where σ^2_x = overall variance in x values, and
σ^2_{x-err} = variance in individual x values due to measurement error.

In our case, the CV of group size, S, from aerial photographs has been estimated from of replicate counts to be 5.4% (Gilpatrick 1993); therefore, the standard deviation of measurement error in ln$S$ is approximately 0.054, and given that photographic group sizes used in Eq. 3 are the average of 3 independent estimates, σ^2_{x-err} is approximately 0.001. The overall variance, σ^2_x, varies from 0.8 to 2.5 among observers depending on the range of group sizes that were estimated by that observer. Therefore, the multiplicative bias in slope due to measurement error in photographic group size is typically less than 1.001 and can be ignored. Measurement error in the independent variable has essentially no effect on regressions through the origin, so this bias can also be ignored for our indirect calibration method.

Log-transformations introduce a source of bias that cannot be ignored (Rothery 1988). If regression assumptions are met, the expected value of ln$S$ (Eq. 4) should be an unbiased estimate of the logarithm of the true group size. However, significant biases are introduced in back-transforming to obtain E(S) due to measurement error in estimating N. A bias-corrected formula can be used to estimate E(S):

$$E(S_i) = \exp\left[\frac{(\ln N_i - \beta_1)}{(\beta_0 + \beta_2 x_2 + \sum_{y=87}^{93} \beta_y x_y)} - F \right] \qquad \textbf{(8)},$$

where the bias correction factor F is typically estimated as

$$F = \frac{\sigma^2_{E(lnS)}}{2} \qquad \textbf{(9)}$$

(Rothery 1988). Because biases are expected to be multiplicative on a linear scale, multiplicative bias will be expressed as a ratio of observed to expected values.

RESULTS

Overall, the uncorrected "best" estimates of group size from shipboard observers are biased (-7%) relative to the mean aerial photographic estimates of group size (Table 2) and show considerable variability (Fig. 1). Directly calibrated estimates of group size are less variable (Fig. 2). The mean bias in the back-transformed, calibrated group size estimates is large (19%, Table 2), but this bias was reduced considerably (to 2%, Table 2) by using the bias correction formulae (Eq. 8 & 9 with ASPE as an estimate of $\sigma^2_{E(lnS)}$). A weighted average of calibrated group sizes is even less variable (Fig. 3) and has a small bias (4%, Table 2), and a weighted geometric mean of calibrated group sizes is the least variable estimate and has a small bias (-2%) (Table 2). The actual regression parameters show that some observers tend to overestimate group size and others tend to underestimate (Appendices 1 & 2).

For the 34 observers who were directly calibrated, the calibration procedures with the lowest prediction error were generally simple procedures (Table 1): either no correction at all (n=8), simple regression through the origin (n=11), or regression through the origin within each year (n=6). Using only these simple regression procedures, calibrated estimates of group size are unbiased and have only a slightly greater ASPE than the more complex models (Table 2, Method 4). An intercept term decreased the prediction error in only 6 cases (Table 1, procedures 3, 4, 8, and 9). A Beaufort term decreased the prediction error in only 5 cases (Table 1, procedures 4, 5, 6, and 9). Of the 25 observers who participated in multiple years, year-specific coefficients decreased the prediction error for only 8 (Table 1, procedures 6, 7, 8, and 9). For the majority of observers, "best" estimates of group size resulted in a lower ASPE than did a weighted sum of the "best", "high" and "low" estimates of group size (Table 1). Using the log-transformation, regression residuals for calibrated estimates (Fig. 2) and for the weighted mean of calibrated estimates (Fig. 3) are symmetrically distributed and show no signs of heteroscedasticity. Looking at only the within-year regressions through the origin for an observers "best" estimates of group size, we see a trend from an equal number of over and underestimators in the 1980's to an increasing number of underestimators in 1992-93 (Fig. 4).

Indirectly calibrated estimates of group size are in good agreement with the means of directly calibrated estimates (Fig. 5). Based on regression through the origin, there was a significant correlation between direct (Appendix 1) and indirect (Appendix 3) calibration factors estimated for a given observer in a given year (Fig. 6, $R^2 = 0.6$). Overall, squared regression residuals were smaller for the indirect calibration method (ASPE=0.262) than for the direct calibration method (ASPE=0.283).

DISCUSSION

Bias

Individual observers have consistent tendencies in estimating the number of cetaceans present in a group. Some observers are underestimators, some are overestimators, and some appear to be accurate without any calibration. The tendency to over or underestimate appears consistent within a broad range of group sizes (Fig. 1) and, for the majority of observers, does not appear to be affected by sighting conditions (Beaufort sea state) or species composition of groups ("target" vs. "non-target schools", see Gerrodette and Perrin 1991). The majority of observers also appear consistent in their estimation among years.

Indirect calibration factors provide further evidence of the consistency within individual observers. Aerial photographic estimates of group size could only be made for a minority of groups (approximately 10%) seen from one vessel only (R/V *David Starr Jordan*). These groups included mostly dolphin groups seen in good weather conditions. Observers knew when calibration photographs were being taken, and it is possible that they changed their methods of estimation for such groups. The other ship used in these surveys (R/V *McArthur*) surveyed in slightly different areas which had different species compositions (especially in 1993 when the *McArthur* surveyed primarily of California and the *Jordan* surveyed primarily off Mexico). The high correlation between directly and indirectly estimated calibration factors provides strong circumstantial evidence that the tendency for an individual observer to over or underestimate groups was also consistent for the majority of groups that could not be estimated photographically.

Although observers appear to be consistent in their biases, the size of their biases are large and should not be ignored. For some observers, direct calibration factors (β_0, based on a simple within-year regression through the origin) are as small as 0.77 or as large as 1.08 (which on a linear scale would result in a group of 500 being estimated, on average, as a group of 120 or as a group of 820, respectively). Such large biases could easily undermine attempts to detect changes in population size, especially, in cases where the biases change over time (Fig. 4). For the observers that were directly calibrated here, the change in bias towards an increasing number of underestimators appears to be caused more by the addition of new observers (with a tendency to underestimate) rather than a change in estimation of previous observers.

Our calibration procedures cannot be expected to eliminate or reduce bias unless aerial estimates of group size are themselves unbiased. When cetaceans are near the surface in clear, calm water, all individuals appear distinct in our images. Although some dolphins can be temporarily obscured by other dolphins or by white caps, the 80% frame overlap obtained with our cameras means that obscured dolphins are likely to be visible at some time. The greatest potential bias with aerial photographic counts would result if a portion of the group is not within the photographic series. This could occur if a portion of the group dives or is separate from the main group. Every effort was made to count only coherent groups that were traveling rapidly at the surface (and were therefore not likely to be diving). Nonetheless, some bias is possible in the aerial photographic estimates, and, because these estimates are based on counts, any bias is likely to result in an underestimate of group size. Overall, observers "best" estimates of group size were less than aerial photographic counts; therefore, our calibration and bias correction is justified. Our calibrated estimates of group size may underestimate true group size to the same degree that aerial photographic counts underestimate true group size.

Precision

In addition to the bias in some individuals, all observers have a high degree of imprecision. The average squared prediction errors for uncalibrated "best" estimates is 0.465 in log-space (Table 2) which corresponds to a coefficient of variation of 77% in untransformed-space. The use of calibration factors improved the average squared prediction errors for calibrated estimates to 0.283 in log-space (Table 2) which corresponds to a coefficient of variation of 57% in untransformed-space. However, for a single calibrated estimate of group size, the 95% confidence intervals would be from approximately 33% to 300% of an estimate. This uncertainty adds considerably to the overall uncertainty in cetacean density and abundance estimates. In line-transect analysis, the usual method to estimate the total variance in group size estimates implicitly includes the component of variance due to (unbiased) measurement error (Buckland et al. 1993). Improving the precision of group size estimation (such as by our calibration methods) should improve the precision of line-transect density and abundance estimates.

Another method to make group size estimates more precise is to average estimates from several observers. Typically on SWFSC cruises, group size is estimated independently by up to 3 different observers, which in theory could reduce the variance in group size estimates by 66% over having just one estimate. On average, 5.5 independent estimates were made for our photo-calibration groups; which, if all estimates were independent, should have resulted in a 82% reduction in variance. In practice, the measured reduction in ASPE for the mean of all observers (relative to a single estimate) was only 37%. This smaller-than-expected reduction in variance is probably because multiple estimates from the same platform are not truly independent. All observers on a ship are looking at a group from roughly the same perspective and are observing the same behaviors, all of which are likely to introduce correlations between estimates. This might also explain why the ASPE was lower for the indirect calibration method the for the direct calibration methods. In the former case, indirectly calibrated estimates were compared to directly calibrated estimates made from the same platform, whereas in the latter case, directly calibrated estimates were compared to estimates made from an entirely different perspective. Although improvement in precision can be expected by averaging group size estimates among observers, the expected improvement will typically be less than that predicted by assuming the estimates are independent.

Prior to this paper, the best method of obtaining an overall estimate of expected group size from multiple estimates had not been addressed. Calibrated estimates of group size appear to be unbiased and normally distributed on a log-log scale. If the logarithms of calibrated group size estimates are unbiased, logarithms of group size can be averaged without introducing bias which would favor using a geometric mean. In practice, we found that a simple arithmetic mean of the bias-corrected estimates performed almost as well as a geometric mean (Table 2, methods 6 & 7). In averaging estimates from different observers, it is obvious that some observers are more consistent than others; therefore, average group size should be computed as a weighted average with weights being the inverse of their estimation variance (ASPE).

Recommendations

Although the methods used here may have some application to similar studies, we are not optimistic that they will be widely applicable. Our direct calibration methods require an alternate source of unbiased group size estimates for at least a subset of groups. Our indirect method requires that a subset of observers be directly calibrated. Unbiased aerial estimates of group size are simply not feasible to collect on most surveys because of the expense and logistic difficulty of carrying a helicopter with the required photogrammetric cameras. Clearly, some less expensive alternative would be preferred for obtaining individual calibration factors. Because the tendency to over or underestimate group size appears to be fairly consistent for an observer, the same pattern may hold when an observer is asked to estimate group size from a still photograph or a video image. Future calibration efforts of this sort should also test whether the tendency to over or underestimate group size can be predicted by an observer's performance in an artificial, laboratory setting. Needless to say, the more closely the laboratory situation mimics the problem of estimating group size at sea (such as with a computer simulation), the better the chance that a calibration factor from the laboratory study will be appropriate for correcting survey estimates of group size.

ACKNOWLEDGMENTS

This project has been involved efforts by a large number of people. Photographic counts were made by J. Gilpatrick, M. Lynn, and R. Westlake. Shipboard estimates of cetacean group size were made by G. Aldana, C. Alvarez, E. Archer, W. Armstrong, L. Baraff, S. Beavers, S. Benson, C. Bisbee, P. Boveng, J. Carretta, P. Corkeron, J. Cotton, D. Everhart, T. Farley, K. Forney, C. Fried, G. Friedrichsen, D. Gendron, S. Hill, W. Irwin, D. Kinzey, S. Kruse, R. Leduc, H. Lira, M. Lowry, M. Lycan, R. Mellon, S. Miller, E. Mitchell, L. Morse, M. Newcomer, S. Norman, P. Olson, D. Palacios, J. Peterson, R. Pitman, S. Perry, T. Pusser, J. Quan, J. Raffeto, K. Raum-Suryan, K. Rittmaster, J. Rivers, R. Rowlett, S. Sinclair, D. Skordal, B. Smith, C. Speck, C. Stinchcomb, V. Thayer, J. Waite, and M. Webber. We also wish to thank all of the cruise coordinators, cruise leaders, officers, crew, and helicopter pilots, without whom this project would have never occurred. Portions of the computer programs used in this analysis were written by C. Show. Earlier versions of this manuscript were improved by helpful comments from J. Carretta, K. Forney, J. Gilpatrick, M. Scott, B. Taylor, and K. Van Waerebeek.

LITERATURE CITED

Barlow, J. 1995. The abundance of cetaceans in California waters. Part I: Ship surveys in summer and fall of 1991. Fish. Bull. 93:1-14.

Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 1993. Distance Sampling: Estimating abundance of biological populations. Chapman and Hall. London. 446pp.

Fuller, W. A. 1987. Measurement Error Models. John Wiley & Sons, New York. xxx pp.

Gerrodette, T. D. and C. Perrin. 1991. Calibration of shipboard estimates of dolphin school size from aerial photographs. Administrative Report LJ-91-36, available from Southwest Fisheries Science Center, P.O. Box 271, La Jolla, CA 92038. 73pp.

Gilpatrick, J. W., Jr. 1993. Method and precision in estimation of dolphin school size with vertical aerial photography. Fish. Bull. 91:641-648.

Hill, P. S. and J. Barlow. 1992. Report of a marine mammal survey of the California coast aboard the research vessel McARTHUR July 28-November 5, 1991. NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-169. NTIS #PB93-109908. 103pp.

Mangels, K. F. and T. Gerrodette. 1994a. Report of cetacean sightings during a marine mammal survey in the eastern tropical Pacific Ocean aboard the NOAA ships *McArthur* and *David Starr Jordan* July 28 - November 2, 1992. NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-200. 74pp.

Mangels, K. F. and T. Gerrodette. 1994b. Report of cetacean sightings during a marine mammal survey in the eastern Pacific Ocean and the Gulf of California aboard the NOAA ships *McArthur* and *David Starr Jordan* July 28 - November 6, 1993. NOAA Technical Memorandum NOAA-TM-NMFS-SWFSC-211. 86pp.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. 1988. Numerical Recipes: The Art of Scientific Computing. Cambridge Univ. Press. 818pp.

Rothery, P. 1988. A cautionary note on data transformation: bias in back-transformed means. Bird Study 35:219-222.

Scott, M. D., W. L. Perryman, and W. G. Clark. 1985. The use of aerial photographs for estimating school sizes of cetaceans. Bull. Inter-American Tropical Tuna Comm. 18:383-404.

VonSaunder, A. and J. Barlow. (In prep.). Report of the 1996 marine mammal survey off the coasts of California, Oregon, and Washington aboard the research vessels McARTHUR and JORDAN. NOAA Technical Memorandum.

Wade, P. R. and T. Gerrodette. 1993. Estimates of cetacean abundance and distribution in the eastern tropical Pacific. Rep. Int. Whal. Commn. 43:477-493.

Table 1. Frequency distribution for the best of eighteen calibration models (nine regression models and two estimation models) chosen for each of 34 observers based on the lowest ASPE.

Calibration Procedure	Estimated Parameters	Estimation Model "Best" Estimates	Estimation Model Weighted "Best-High-Low"	Total
1	none	4	4	8
2	β_0	6	5	11
3	β_0, β_1	2	2	4
4	β_0, β_1, β_2	1	0	1
5	β_0, β_2	2	0	2
6	β_2, $\beta_{87\text{-}93}$	1	0	1
7	$\beta_{87\text{-}93}$	3	3	6
8	β_1, $\beta_{87\text{-}93}$	0	0	0
9	β_1, β_2, $\beta_{87\text{-}93}$	0	1	1
Total		19	15	34

Table 2. Bias and precision of methods for directly calibrating group size estimates for individual observers (Methods 1-4) and for estimates averaged among observers (Methods 5-7). Calibrations for each observer are based on the procedure that gave the lowest ASPE or (for Method 4) the "simple" regression procedure that gave the lowest ASPE (simple regression models included Procedures 1, 2, and 7 (Table 1) that were based on regression through the origin using only the "best" estimates of group size). Multiplicative bias is estimated as the ratio of calibrated estimates to the mean aerial estimates, averaged over all observers' estimates (methods 1-4) or averaged over all groups (methods 5-7). ASPE is estimated as the squared difference between the logarithms of the calibrated and mean aerial estimates averaged over all observer estimates (procedures 1-4) or averaged over all groups (methods 5-7). ASPE is measured in log-space because errors are log-normally distributed; consequently, the bias correction is excluded when estimating ASPE.

Method	Multiplicative Bias	ASPE
1. Uncalibrated "Best"Estimates	0.933	0.465
2. Directly Calibrated Estimates (No Bias Correction)	1.192	0.283
3. Directly Calibrated Estimates (With Bias Correction)	1.025	0.283
4. "Simple" Directly Calibrated Estimates (With Bias Correction)	1.001	0.296
5. Mean Calibrated Estimates (With Bias Correction)	1.040	0.179
6. Weighted Mean Calibrated Estimates (With Bias Correction)	1.040	0.170
7. Weighted Geometric Mean Calibrated Estimates (With Bias Correction)	0.982	0.163

Figure 1. Observers' "best" estimates of group size versus the mean of three counts made from aerial photographs. Diagonal represents 1:1 line. Sample size (n=1,417) is the number of independent estimates made by observers.

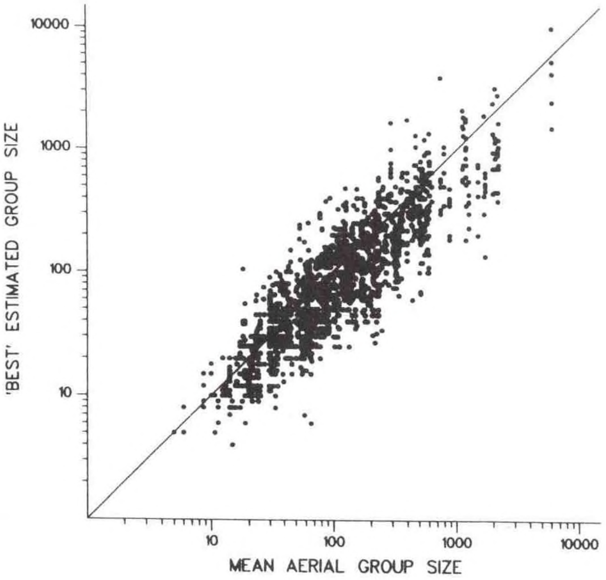


Figure 2. Observers' "calibrated" estimates of group size versus the mean of three counts made from aerial photographs. Calibrated estimates are based on the regression procedure that gave the lowest ASPE for a given observer and include bias correction. Diagonal represents 1:1 line. Sample size (n=1,417) is the number of independent estimates made by observers.

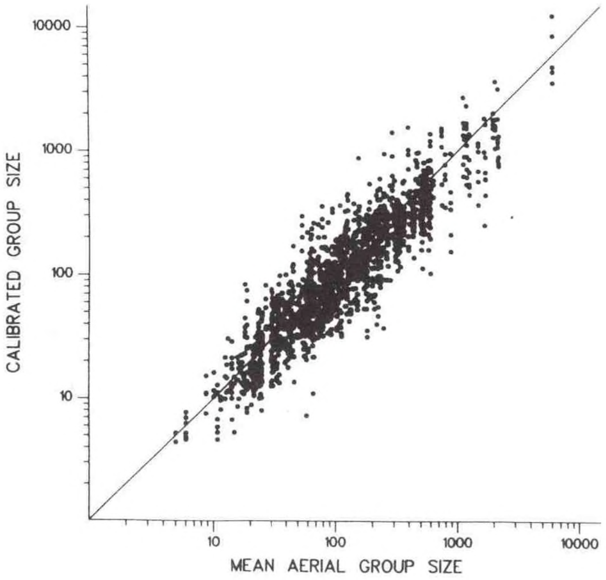


Figure 3. Weighted mean of directly calibrated estimates of group size versus the mean of three counts made from aerial photographs. Calibrated estimates are based on the regression procedure that gave the lowest ASPE for a given observer. Diagonal represents 1:1 line. Sample size (n=274) is the number of groups photographed.

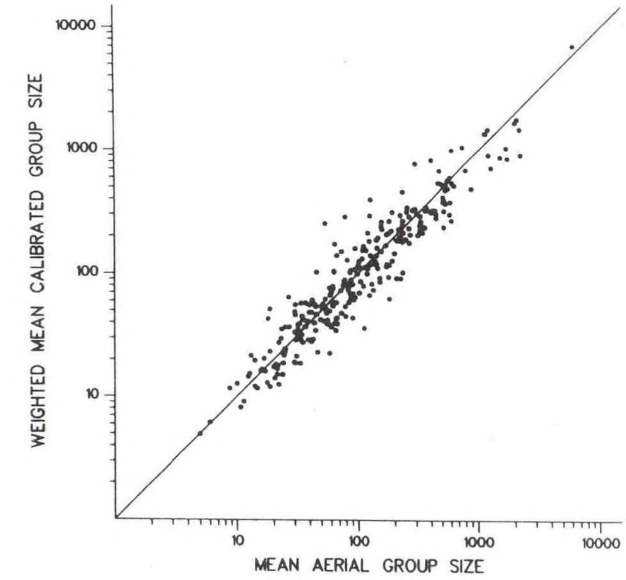


Figure 4. Direct calibration factors (β_0) based on a simple within-year regression through the origin for "best" estimates of group size for years 1987, 1988, 1989, 1990, 1992, and 1993. Lines connect coefficients estimated for the same observer in different years.

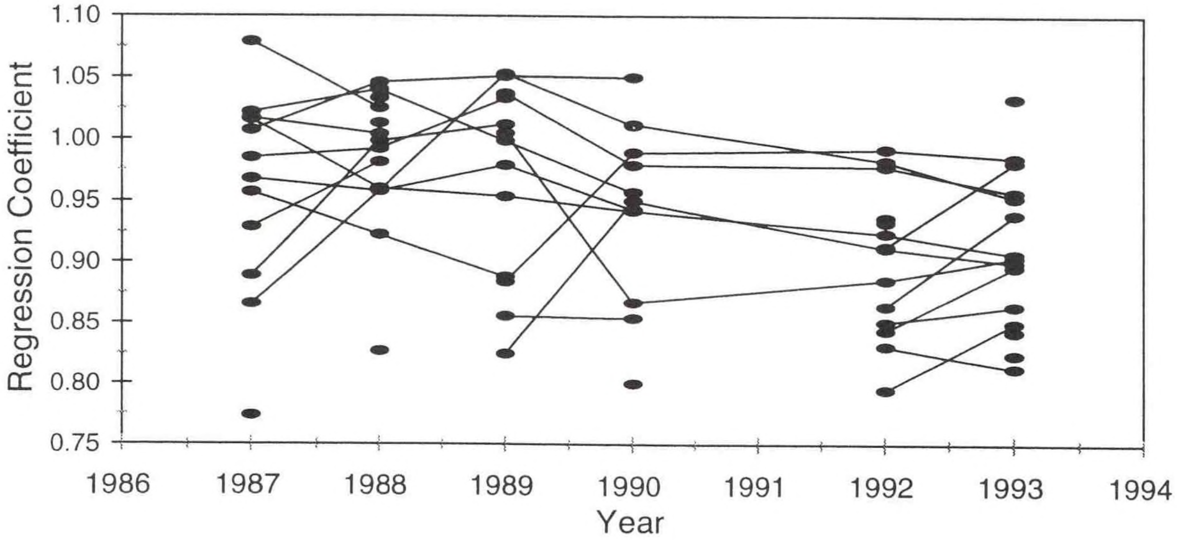


Figure 5. Indirectly calibrated estimates of group size versus the mean of two or more directly calibrated estimates. Directly calibrated estimates are based on the regression procedure that gave the lowest ASPE for a given observer. Indirectly calibrated estimates are based on a simple regression through the origin for all years pooled. Diagonal represents 1:1 line. Sample size (n=11,185) is the number of estimates for groups whose size was estimated by at least two directly calibrated observers and whose size was estimated by at least one other observer (excluding photographed groups used for direct calibration).

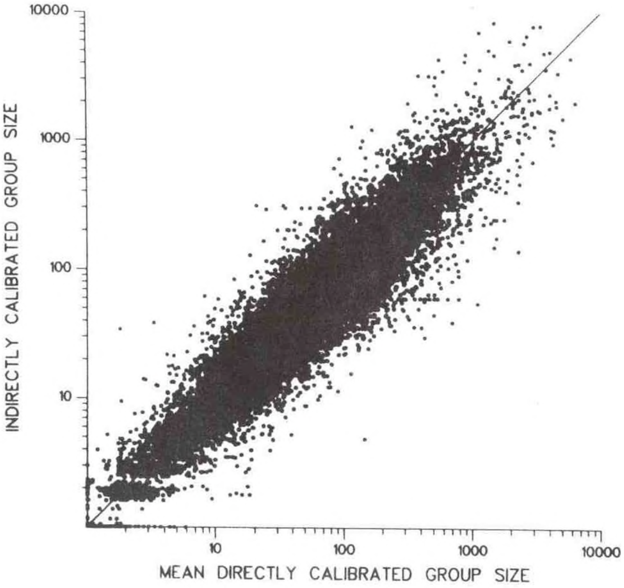
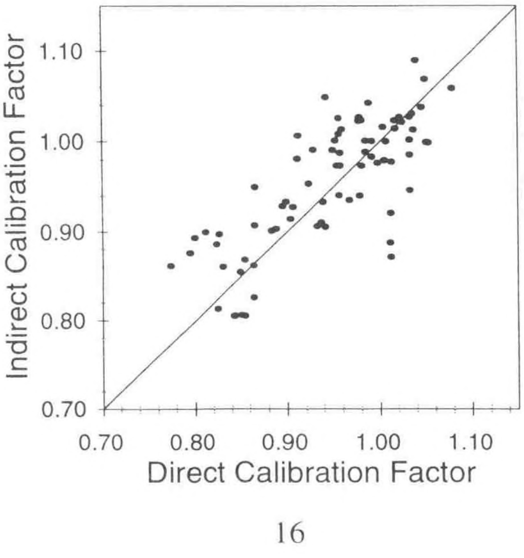


Figure 6. Comparison of direct and indirect calibration factors (β_0) for individual observers based on simple regression through the origin for the years 1987, 1988, 1989, 1990, 1992, and 1993. Direct calibration factors are based only "best" estimates made for "calibration schools". Indirect calibration factors are based on all other groups (excluding "calibration schools") that were estimated by at least two other "calibrated" observer. Diagonal line represents 1:1 line.



Appendix 1. Regression coefficients estimated for the direct calibration of group size based on a comparison of an individual observer's "best" estimates of group size with group size measured from aerial photographs. Coefficients were estimated for nine procedures which included different components of the regression model (Eq. 3). The lowest value (*) of the average squared prediction error (ASPE) indicates the best procedure for a given observer. The best overall procedure (comparing Appendices 1 and 2) is indicated with a "+" and is illustrated in Appendix 4. Sample size for all years is indicated by N.

Obs. Number	N	Procedure	Weights- w_i			Regression Coefficients									ASPE
			Best	High	Low	β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	
1	57	1	1.00	.00	.00										.2305
		2	1.00	.00	.00	.988									.2409
		3	1.00	.00	.00	.947	.201								.2693
		4	1.00	.00	.00	.904	.204	.068							.2480
		5	1.00	.00	.00	.945		.068							.2223*+
		6	1.00	.00	.00			.065	.000	.000	.990	.931	.941	.939	.2347
		7	1.00	.00	.00				.000	.000	1.037	.979	.978	.957	.2472
		8	1.00	.00	.00		.325		.000	.000	.976	.912	.914	.883	.2910
		9	1.00	.00	.00		.294	.064	.000	.000	.936	.871	.884	.872	.2724
2	82	1	1.00	.00	.00										.4502
		2	1.00	.00	.00	.964									.4612
		3	1.00	.00	.00	.943	.110								.4948
		4	1.00	.00	.00	.906	.205	.032							.5336
		5	1.00	.00	.00	.949		.026							.4769
		6	1.00	.00	.00			-.013	.000	.000	.887	.998	1.001	.992	.4597
		7	1.00	.00	.00				.000	.000	.884	.989	.992	.985	.4484*+
		8	1.00	.00	.00		.370		.000	.000	.805	.913	.926	.918	.5306
		9	1.00	.00	.00		.363	-.004	.000	.000	.807	.917	.930	.921	.5462
3	38	1	1.00	.00	.00										.6835
		2	1.00	.00	.00	.855									.3215*
		3	1.00	.00	.00	.848	.034								.3412
		4	1.00	.00	.00	.847	.036	.003							.3604
		5	1.00	.00	.00	.854		.002							.3391
		6	1.00	.00	.00			.002	.000	.000	.855	.853	.000	.000	.3564
		7	1.00	.00	.00				.000	.000	.856	.854	.000	.000	.3366
		8	1.00	.00	.00		.031		.000	.000	.849	.848	.000	.000	.3582
		9	1.00	.00	.00		.033	.003	.000	.000	.848	.846	.000	.000	.3800
4	29	1	1.00	.00	.00										.3514
		2	1.00	.00	.00	1.033									.3234*+
		3	1.00	.00	.00	.744	1.532								.5013
		4	1.00	.00	.00	.743	1.485	.017							.5320
		5	1.00	.00	.00	1.011		.039							.3348
		6	1.00	.00	.00			.039	.000	1.011	.000	.000	.000	.000	.3348
		7	1.00	.00	.00				.000	1.033	.000	.000	.000	.000	.3234
		8	1.00	.00	.00		1.532		.000	.744	.000	.000	.000	.000	.5010
		9	1.00	.00	.00		1.484	.017	.000	.743	.000	.000	.000	.000	.5318
5	49	1	1.00	.00	.00										.2638*
		2	1.00	.00	.00	1.007									.2708
		3	1.00	.00	.00	.909	.467								.3238
		4	1.00	.00	.00	.909	.466	.000							.3404
		5	1.00	.00	.00	1.003		.011							.2830
		6	1.00	.00	.00			.019	1.012	1.036	.000	.949	.000	.000	.2817
		7	1.00	.00	.00				1.022	1.040	.000	.957	.000	.000	.2709
		8	1.00	.00	.00		.388		.933	.960	.000	.882	.000	.000	.3208
		9	1.00	.00	.00		.370	.012	.931	.961	.000	.881	.000	.000	.3357

Appendix 1. (Continued).

Obs. Number	N	Procedure	Weights- w_i Best	High	Low	Regression Coefficients β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	ASPE
6	55	1	1.00	.00	.00										.3543*+
		2	1.00	.00	.00	.961									.3593
		3	1.00	.00	.00	.823	.700								.4686
		4	1.00	.00	.00	.825	.735	-.014							.4963
		5	1.00	.00	.00	.959		.004							.3776
		6	1.00	.00	.00			.005	.926	.978	.000	.000	.000	.000	.3765
		7	1.00	.00	.00				.929	.981	.000	.000	.000	.000	.3561
		8	1.00	.00	.00		.785		.767	.830	.000	.000	.000	.000	.5025
		9	1.00	.00	.00		.823	-.016	.768	.832	.000	.000	.000	.000	.5484
7	73	1	1.00	.00	.00										.2487
		2	1.00	.00	.00	.964									.2445*
		3	1.00	.00	.00	.924	.199								.2724
		4	1.00	.00	.00	.920	.184	.017							.2766
		5	1.00	.00	.00	.957		.019							.2495
		6	1.00	.00	.00			.019	.957	.953	.973	.936	.000	.000	.2659
		7	1.00	.00	.00				.968	.958	.979	.943	.000	.000	.2614
		8	1.00	.00	.00		.171		.933	.923	.943	.911	.000	.000	.2891
		9	1.00	.00	.00		.153	.017	.926	.922	.941	.907	.000	.000	.2924
8	35	1	1.00	.00	.00										.2714*
		2	1.00	.00	.00	1.010									.2808
		3	1.00	.00	.00	.971	.180								.3285
		4	1.00	.00	.00	.988	.014	.053							.3102
		5	1.00	.00	.00	.990		.054							.2836
		6	1.00	.00	.00			.056	.986	.993	.000	.000	.000	.000	.2993
		7	1.00	.00	.00				1.017	1.004	.000	.000	.000	.000	.2992
		8	1.00	.00	.00		.167		.979	.970	.000	.000	.000	.000	.3478
		9	1.00	.00	.00		.017	.056	.982	.990	.000	.000	.000	.000	.3290
9	10	1	1.00	.00	.00										1.0606
		2	1.00	.00	.00	.774									.2060*+
		3	1.00	.00	.00	.667	.484								.3335
		4	1.00	.00	.00	.665	.458	.018							.6549
		5	1.00	.00	.00	.762		.027							.2819
		6	1.00	.00	.00			.027	.762	.000	.000	.000	.000	.000	.2819
		7	1.00	.00	.00				.774	.000	.000	.000	.000	.000	.2060
		8	1.00	.00	.00		.484		.667	.000	.000	.000	.000	.000	.3335
		9	1.00	.00	.00		.458	.018	.665	.000	.000	.000	.000	.000	.6552
10	75	1	1.00	.00	.00										.2948
		2	1.00	.00	.00	1.039									.2491
		3	1.00	.00	.00	.810	1.086								.3134
		4	1.00	.00	.00	.807	1.052	.029							.3081
		5	1.00	.00	.00	1.024		.044							.2465*+
		6	1.00	.00	.00			.050	.984	1.036	1.035	1.030	.000	.000	.2571
		7	1.00	.00	.00				1.007	1.046	1.051	1.050	.000	.000	.2643
		8	1.00	.00	.00		1.188		.741	.800	.796	.821	.000	.000	.3343
		9	1.00	.00	.00		1.148	.036	.733	.801	.793	.814	.000	.000	.3245
11	120	1	1.00	.00	.00										.3468
		2	1.00	.00	.00	.944									.3144
		3	1.00	.00	.00	.856	.433								.3719
		4	1.00	.00	.00	.860	.458	-.020							.3821
		5	1.00	.00	.00	.950		-.013							.3201
		6	1.00	.00	.00			-.003	1.017	.961	.955	.943	.925	.909	.3207
		7	1.00	.00	.00				1.016	.960	.954	.942	.924	.907	.3140*
		8	1.00	.00	.00		.313		.946	.895	.887	.882	.862	.847	.3607
		9	1.00	.00	.00		.329	-.009	.947	.894	.887	.883	.864	.850	.3707

Appendix 1. (Continued).

Obs. Number	N	Procedure	Weights- w_i Best	High	Low	Regression Coefficients β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	ASPE
12	61	1	1.00	.00	.00										.4241
		2	1.00	.00	.00	.966									.4396
		3	1.00	.00	.00	1.054	-.451								.3666
		4	1.00	.00	.00	1.049	-.523	.030							.3705
		5	1.00	.00	.00	.954		.019							.4600
		6	1.00	.00	.00			.018	.878	.987	1.000	.000	.000	.000	.4189
		7	1.00	.00	.00				.889	.998	1.012	.000	.000	.000	.4017
		8	1.00	.00	.00		-.291		.949	1.054	1.067	.000	.000	.000	.3627*
		9	1.00	.00	.00		-.355	.026	.945	1.050	1.061	.000	.000	.000	.3670
13	13	1	1.00	.00	.00										1.0370
		2	1.00	.00	.00	.827									.3612*
		3	1.00	.00	.00	.668	.863								.5152
		4	1.00	.00	.00	.669	.692	.043							.7194
		5	1.00	.00	.00	.781		.066							.4174
		6	1.00	.00	.00			.066	.000	.781	.000	.000	.000	.000	.4174
		7	1.00	.00	.00				.000	.827	.000	.000	.000	.000	.3612
		8	1.00	.00	.00		.863		.000	.668	.000	.000	.000	.000	.5149
		9	1.00	.00	.00		.692	.043	.000	.669	.000	.000	.000	.000	.7190
14	7	1	1.00	.00	.00										.1022
		2	1.00	.00	.00	1.036									.0847*+
		3	1.00	.00	.00	.976	.306								.1526
		4	1.00	.00	.00	.999	.275	-.029							.2182
		5	1.00	.00	.00	1.054		-.030							.1011
		6	1.00	.00	.00			-.030	.000	.000	1.054	.000	.000	.000	.1012
		7	1.00	.00	.00				.000	.000	1.036	.000	.000	.000	.0847
		8	1.00	.00	.00		.306		.000	.000	.976	.000	.000	.000	.1525
		9	1.00	.00	.00		.274	-.029	.000	.000	.999	.000	.000	.000	.2182
15	16	1	1.00	.00	.00										.1783*+
		2	1.00	.00	.00	1.013									.1990
		3	1.00	.00	.00	1.044	-.159								.2108
		4	1.00	.00	.00	1.027	.055	-.049							.2452
		5	1.00	.00	.00	1.038		-.048							.2153
		6	1.00	.00	.00			-.048	.000	1.038	.000	.000	.000	.000	.2153
		7	1.00	.00	.00				.000	1.013	.000	.000	.000	.000	.1990
		8	1.00	.00	.00		-.160		.000	1.045	.000	.000	.000	.000	.2109
		9	1.00	.00	.00		.055	-.050	.000	1.027	.000	.000	.000	.000	.2452
16	21	1	1.00	.00	.00										.2046
		2	1.00	.00	.00	1.055									.1469
		3	1.00	.00	.00	1.071	-.071								.1548
		4	1.00	.00	.00	1.073	-.234	.089							.1268*+
		5	1.00	.00	.00	1.021		.084							.1309
		6	1.00	.00	.00			.076	1.041	1.004	.000	.000	.000	.000	.1412
		7	1.00	.00	.00				1.079	1.025	.000	.000	.000	.000	.1499
		8	1.00	.00	.00		-.207		1.128	1.070	.000	.000	.000	.000	.1487
		9	1.00	.00	.00		-.320	.081	1.115	1.071	.000	.000	.000	.000	.1300
17	38	1	1.00	.00	.00										.2974
		2	1.00	.00	.00	.961									.3052
		3	1.00	.00	.00	1.047	-.391								.2613
		4	1.00	.00	.00	1.030	-.483	.061							.2437
		5	1.00	.00	.00	.929		.052							.3015
		6	1.00	.00	.00			.052	.836	.925	1.017	.985	.000	.000	.3189
		7	1.00	.00	.00				.866	.958	1.053	1.011	.000	.000	.2905
		8	1.00	.00	.00		-.454		.971	1.052	1.156	1.121	.000	.000	.2262
		9	1.00	.00	.00		-.577	.065	.962	1.034	1.138	1.118	.000	.000	.2220*

Appendix 1. (Continued).

Obs.		Pro-	Weights- w_i			Regression Coefficients									
Number	N	cedure	Best	High	Low	β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	ASPE
18	80	1	1.00	.00	.00										.5021
		2	1.00	.00	.00	.896									.3031
		3	1.00	.00	.00	.902	-.029								.3060
		4	1.00	.00	.00	.857	.056	.047							.3123
		5	1.00	.00	.00	.869		.046							.2986
		6	1.00	.00	.00			.018	.000	.000	.819	.937	.899	.893	.2667
		7	1.00	.00	.00				.000	.000	.825	.950	.912	.900	.2609*+
		8	1.00	.00	.00		.081		.000	.000	.808	.933	.897	.886	.2773
		9	1.00	.00	.00		.111	.020	.000	.000	.795	.912	.877	.872	.2867
19	60	1	1.00	.00	.00										.2000*+
		2	1.00	.00	.00	.995									.2090
		3	1.00	.00	.00	.997	-.007								.2141
		4	1.00	.00	.00	.985	-.042	.031							.2156
		5	1.00	.00	.00	.977		.030							.2134
		6	1.00	.00	.00			.030	.966	.974	1.013	.000	.000	.000	.2255
		7	1.00	.00	.00				.985	.992	1.033	.000	.000	.000	.2201
		8	1.00	.00	.00		.017		.981	.988	1.029	.000	.000	.000	.2293
		9	1.00	.00	.00		-.018	.030	.970	.977	1.017	.000	.000	.000	.2318
20	28	1	1.00	.00	.00										.2619
		2	1.00	.00	.00	.944									.2331
		3	1.00	.00	.00	.874	.342								.2782
		4	1.00	.00	.00	.872	.323	.010							.2992
		5	1.00	.00	.00	.932		.019							.2508
		6	1.00	.00	.00			.016	.947	.000	.879	.000	.000	.000	.2338
		7	1.00	.00	.00				.957	.000	.888	.000	.000	.000	.2186*+
		8	1.00	.00	.00		.329		.889	.000	.822	.000	.000	.000	.2594
		9	1.00	.00	.00		.315	.007	.887	.000	.821	.000	.000	.000	.2783
21	53	1	1.00	.00	.00										.5114
		2	1.00	.00	.00	.912									.4190
		3	1.00	.00	.00	.883	.143								.4567
		4	1.00	.00	.00	.873	.145	.014							.4812
		5	1.00	.00	.00	.903		.014							.4415
		6	1.00	.00	.00			.014	.000	.000	.995	.857	.000	.894	.3822
		7	1.00	.00	.00				.000	.000	1.005	.867	.000	.904	.3607*
		8	1.00	.00	.00		.255		.000	.000	.956	.814	.000	.852	.4118
		9	1.00	.00	.00		.257	.014	.000	.000	.945	.803	.000	.842	.4354
22	34	1	1.00	.00	.00										.2158*+
		2	1.00	.00	.00	.996									.2298
		3	1.00	.00	.00	.976	.093								.2503
		4	1.00	.00	.00	.915	.198	.061							.2544
		5	1.00	.00	.00	.961		.057							.2221
		6	1.00	.00	.00			.042	.000	.000	.000	.981	.000	.941	.2253
		7	1.00	.00	.00				.000	.000	.000	1.012	.000	.953	.2244
		8	1.00	.00	.00		.221		.000	.000	.000	.966	.000	.900	.2591
		9	1.00	.00	.00		.261	.045	.000	.000	.000	.924	.000	.877	.2668
23	12	1	1.00	.00	.00										1.2228
		2	1.00	.00	.00	.800									.2978*
		3	1.00	.00	.00	.527	1.435								.3042
		4	1.00	.00	.00	.543	1.428	-.037							.3998
		5	1.00	.00	.00	.816		-.039							.3342
		6	1.00	.00	.00			-.039	.000	.000	.000	.816	.000	.000	.3342
		7	1.00	.00	.00				.000	.000	.000	.800	.000	.000	.2978
		8	1.00	.00	.00		1.434		.000	.000	.000	.527	.000	.000	.3043
		9	1.00	.00	.00		1.428	-.037	.000	.000	.000	.543	.000	.000	.3999

Appendix 1. (Continued).

Obs. Number	N	Procedure	Weights- w_i			Regression Coefficients									ASPE
			Best	High	Low	β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	
24	25	1	1.00	.00	.00										.4288
		2	1.00	.00	.00	.933									.3585*+
		3	1.00	.00	.00	.725	1.197								.4322
		4	1.00	.00	.00	.738	1.169	-.010							.4759
		5	1.00	.00	.00	.970		-.048							.3813
		6	1.00	.00	.00			-.048	.000	.000	.000	.000	.970	.000	.3813
		7	1.00	.00	.00				.000	.000	.000	.000	.933	.000	.3585
		8	1.00	.00	.00		1.197		.000	.000	.000	.000	.725	.000	.4323
		9	1.00	.00	.00		1.169	-.010	.000	.000	.000	.000	.738	.000	.4759
25	28	1	1.00	.00	.00										.5596
		2	1.00	.00	.00	.937									.5442
		3	1.00	.00	.00	.988	-.287								.5001*+
		4	1.00	.00	.00	1.030	-.376	-.033							.5159
		5	1.00	.00	.00	.950		-.017							.5856
		6	1.00	.00	.00			-.017	.000	.000	.000	.000	.950	.000	.5856
		7	1.00	.00	.00				.000	.000	.000	.000	.937	.000	.5442
		8	1.00	.00	.00		-.287		.000	.000	.000	.000	.988	.000	.5001
		9	1.00	.00	.00		-.375	-.033	.000	.000	.000	.000	1.030	.000	.5159
26	53	1	1.00	.00	.00										.7652
		2	1.00	.00	.00	.858									.2714*+
		3	1.00	.00	.00	.810	.260								.3050
		4	1.00	.00	.00	.833	.244	-.029							.3037
		5	1.00	.00	.00	.878		-.031							.2713
		6	1.00	.00	.00			-.029	.000	.000	.000	.000	.873	.881	.2782
		7	1.00	.00	.00				.000	.000	.000	.000	.851	.865	.2768
		8	1.00	.00	.00		.241		.000	.000	.000	.000	.809	.818	.3108
		9	1.00	.00	.00		.236	-.028	.000	.000	.000	.000	.832	.836	.3120
27	58	1	1.00	.00	.00										.9766
		2	1.00	.00	.00	.824									.2821
		3	1.00	.00	.00	.769	.288								.3234
		4	1.00	.00	.00	.796	.287	-.041							.3182
		5	1.00	.00	.00	.851		-.041							.2781
		6	1.00	.00	.00			-.032	.000	.000	.000	.000	.819	.869	.2631
		7	1.00	.00	.00				.000	.000	.000	.000	.795	.850	.2616*
		8	1.00	.00	.00		.244		.000	.000	.000	.000	.750	.803	.2940
		9	1.00	.00	.00		.247	-.032	.000	.000	.000	.000	.773	.821	.2963
30	33	1	1.00	.00	.00										.7109
		2	1.00	.00	.00	.870									.4122
		3	1.00	.00	.00	.670	1.018								.6743
		4	1.00	.00	.00	.673	.991	.006							.7229
		5	1.00	.00	.00	.853		.040							.4200
		6	1.00	.00	.00			.066	.000	.000	.000	.000	.804	.879	.3914*+
		7	1.00	.00	.00				.000	.000	.000	.000	.844	.896	.4153
		8	1.00	.00	.00		1.038		.000	.000	.000	.000	.639	.694	.6655
		9	1.00	.00	.00		.900	.032	.000	.000	.000	.000	.647	.713	.6420
31	47	1	1.00	.00	.00										.3598
		2	1.00	.00	.00	.957									.3591*
		3	1.00	.00	.00	.939	.093								.3826
		4	1.00	.00	.00	.939	.136	-.015							.4074
		5	1.00	.00	.00	.963		-.011							.3768
		6	1.00	.00	.00			-.009	.000	.000	.000	.000	.918	.986	.3853
		7	1.00	.00	.00				.000	.000	.000	.000	.913	.981	.3623
		8	1.00	.00	.00		.110		.000	.000	.000	.000	.891	.959	.3989
		9	1.00	.00	.00		.149	-.014	.000	.000	.000	.000	.891	.960	.4285

Appendix 1. (Continued).

Obs. Number	N	Procedure	Weights- w_i Best	High	Low	β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	ASPE
32	33	1	1.00	.00	.00										.6752
		2	1.00	.00	.00	.905									.5917
		3	1.00	.00	.00	.894	.054								.6218
		4	1.00	.00	.00	.879	.211	-.040							.7184
		5	1.00	.00	.00	.918		-.033							.6334
		6	1.00	.00	.00			-.012	.000	.000	.000	.000	.871	.942	.6476
		7	1.00	.00	.00				.000	.000	.000	.000	.865	.939	.5907*
		8	1.00	.00	.00		.017		.000	.000	.000	.000	.861	.935	.6256
		9	1.00	.00	.00		.079	-.015	.000	.000	.000	.000	.857	.927	.6961
33	36	1	1.00	.00	.00										1.0278
		2	1.00	.00	.00	.821									.4518
		3	1.00	.00	.00	.868	-.235								.4232*
		4	1.00	.00	.00	.871	-.283	.016							.4413
		5	1.00	.00	.00	.818		.007							.4799
		6	1.00	.00	.00			.003	.000	.000	.000	.000	.829	.812	.5123
		7	1.00	.00	.00				.000	.000	.000	.000	.831	.813	.4801
		8	1.00	.00	.00		-.222		.000	.000	.000	.000	.874	.859	.4534
		9	1.00	.00	.00		-.262	.013	.000	.000	.000	.000	.875	.862	.4769
35	19	1	1.00	.00	.00										.4895
		2	1.00	.00	.00	1.033									.5022
		3	1.00	.00	.00	1.248	-1.040								.3418*+
		4	1.00	.00	.00	1.249	-1.051	.004							.3993
		5	1.00	.00	.00	1.041		-.021							.5877
		6	1.00	.00	.00			-.021	.000	.000	.000	.000	.000	1.041	.5877
		7	1.00	.00	.00				.000	.000	.000	.000	.000	1.033	.5022
		8	1.00	.00	.00		-1.041		.000	.000	.000	.000	.000	1.249	.3417
		9	1.00	.00	.00		-1.050	.004	.000	.000	.000	.000	.000	1.249	.3993
36	9	1	1.00	.00	.00										1.3057
		2	1.00	.00	.00	.824									.7438*+
		3	1.00	.00	.00	.727	.536								2.1545
		4	1.00	.00	.00	.755	.347	.024							2.1395
		5	1.00	.00	.00	.815		.035							.9060
		6	1.00	.00	.00			.035	.000	.000	.000	.000	.000	.815	.9059
		7	1.00	.00	.00				.000	.000	.000	.000	.000	.824	.7438
		8	1.00	.00	.00		.537		.000	.000	.000	.000	.000	.727	2.1549
		9	1.00	.00	.00		.347	.024	.000	.000	.000	.000	.000	.755	2.1384
37	30	1	1.00	.00	.00										.8908
		2	1.00	.00	.00	.843									.4304*
		3	1.00	.00	.00	.757	.438								.5573
		4	1.00	.00	.00	.759	.443	-.003							.5937
		5	1.00	.00	.00	.840		.005							.4597
		6	1.00	.00	.00			.005	.000	.000	.000	.000	.000	.840	.4597
		7	1.00	.00	.00				.000	.000	.000	.000	.000	.843	.4304
		8	1.00	.00	.00		.437		.000	.000	.000	.000	.000	.758	.5574
		9	1.00	.00	.00		.443	-.003	.000	.000	.000	.000	.000	.759	.5937

Appendix 2. Regression coefficients estimated for the direct calibration of group size based on a comparison of an individual observer's weighted mean of "best", "high", and "low" estimates of group size with group size measured from aerial photographs. Coefficients were estimated for nine procedures which included different components of the regression model (Eq. 3). The lowest value (*) of the average squared prediction error (ASPE) indicates the best procedure for a given observer. The best overall procedure (comparing Appendices 1 and 2) is indicated with a "+" and are illustrated in Appendix 4. Sample size for all years is indicated by N.

Obs.		Pro-	Weights- w_i			Regression Coefficients									
Number	N	cedure	Best	High	Low	β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	ASPE
1	57	1	.56	.19	.25										.2440
		2	.56	.19	.25	.991									.2565
		3	.57	.19	.24	.947	.217								.2877
		4	.57	.19	.24	.904	.219	.068							.2654
		5	.56	.19	.25	.948		.068							.2370*
		6	.56	.19	.25			.066	.000	.000	.990	.931	.944	.945	.2469
		7	.57	.19	.24				.000	.000	1.038	.980	.982	.964	.2609
		8	.56	.19	.25		.337		.000	.000	.974	.911	.916	.887	.3083
		9	.56	.19	.25		.304	.065	.000	.000	.934	.869	.885	.876	.2875
2	82	1	.00	.04	.96										.5019
		2	.00	.04	.96	.942									.4794
		3	.00	.04	.96	.940	.013								.4974
		4	.00	.04	.96	.895	.129	.039							.5392
		5	.00	.04	.96	.922		.035							.4929
		6	.00	.04	.96			-.002	.000	.000	.864	.968	.972	.956	.4783
		7	.00	.04	.96				.000	.000	.864	.967	.971	.956	.4651*
		8	.00	.04	.96		.256		.000	.000	.809	.915	.926	.909	.5319
		9	.00	.04	.96		.267	.005	.000	.000	.806	.909	.920	.905	.5528
3	38	1	.00	.58	.42										.5714
		2	.00	.58	.42	.874									.3082*+
		3	.00	.58	.42	.846	.136								.3405
		4	.00	.58	.42	.837	.151	.015							.3604
		5	.00	.58	.42	.869		.014							.3232
		6	.00	.58	.42			.014	.000	.000	.871	.866	.000	.000	.3358
		7	.00	.58	.42				.000	.000	.875	.872	.000	.000	.3193
		8	.00	.58	.42		.138		.000	.000	.845	.846	.000	.000	.3520
		9	.00	.58	.42		.150	.015	.000	.000	.837	.837	.000	.000	.3735
4	29	1	.00	1.00	.00										.4866
		2	.00	1.00	.00	1.066									.3497*
		3	.00	1.00	.00	.680	2.039								.5637
		4	.00	1.00	.00	.679	1.985	.020							.5951
		5	.00	1.00	.00	1.037		.049							.3569
		6	.00	1.00	.00			.048	.000	1.037	.000	.000	.000	.000	.3568
		7	.00	1.00	.00				.000	1.066	.000	.000	.000	.000	.3497
		8	.00	1.00	.00		2.041		.000	.680	.000	.000	.000	.000	.5635
		9	.00	1.00	.00		1.984	.020	.000	.679	.000	.000	.000	.000	.5954
5	49	1	.00	.00	1.00										.2517*+
		2	.00	.00	1.00	.977									.2635
		3	.00	.00	1.00	.902	.355								.3066
		4	.00	.00	1.00	.902	.359	-.002							.3230
		5	.00	.00	1.00	.974		.006							.2764
		6	.00	.00	1.00			.014	.985	1.008	.000	.918	.000	.000	.2699
		7	.00	.00	1.00				.993	1.011	.000	.924	.000	.000	.2576
		8	.00	.00	1.00		.261		.933	.957	.000	.874	.000	.000	.2946
		9	.00	.00	1.00		.248	.009	.932	.957	.000	.873	.000	.000	.3096

Appendix 2. (Continued).

Obs. Number	N	Pro- cedure	Weights- w_i Best	High	Low	Regression Coefficients β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	ASPE
6	55	1	.76	.00	.24										.3785
		2	.75	.00	.25	.953									.3753
		3	.76	.00	.24	.823	.658								.4868
		4	.76	.00	.24	.825	.689	-.013							.5127
		5	.76	.00	.24	.951		.004							.3941
		6	.76	.00	.24			.005	.918	.970	.000	.000	.000	.000	.3934
		7	.75	.00	.25				.920	.972	.000	.000	.000	.000	.3725*
		8	.75	.00	.25		.737		.768	.830	.000	.000	.000	.000	.5194
		9	.76	.00	.24		.774	-.014	.770	.832	.000	.000	.000	.000	.5624
7	73	1	.00	.88	.12										.2301*+
		2	.00	.88	.12	.991									.2362
		3	.00	.88	.12	.908	.403								.2807
		4	.00	.88	.12	.904	.383	.022							.2824
		5	.00	.87	.13	.980		.026							.2387
		6	.00	.88	.12			.024	.987	.973	.995	.961	.000	.000	.2541
		7	.00	.88	.12				1.002	.977	1.003	.970	.000	.000	.2510
		8	.00	.88	.12		.383		.924	.899	.922	.897	.000	.000	.2970
		9	.00	.88	.12		.364	.019	.916	.899	.919	.893	.000	.000	.2991
8	35	1	.00	.00	1.00										.2523*+
		2	.00	.00	1.00	.980									.2707
		3	.00	.00	1.00	.952	.128								.3108
		4	.00	.00	1.00	.969	-.042	.054							.2911
		5	.00	.00	1.00	.960		.052							.2726
		6	.00	.00	1.00			.053	.959	.962	.000	.000	.000	.000	.2879
		7	.00	.00	1.00				.987	.972	.000	.000	.000	.000	.2880
		8	.00	.00	1.00		.107		.963	.950	.000	.000	.000	.000	.3278
		9	.00	.00	1.00		-.040	.055	.967	.970	.000	.000	.000	.000	.3088
9	10	1	.00	1.00	.00										.8299
		2	.00	1.00	.00	.810									.2257*
		3	.00	1.00	.00	.648	.733								.3599
		4	.00	1.00	.00	.645	.701	.022							.6903
		5	.00	1.00	.00	.793		.037							.2954
		6	.00	1.00	.00			.037	.793	.000	.000	.000	.000	.000	.2955
		7	.00	1.00	.00				.810	.000	.000	.000	.000	.000	.2257
		8	.00	1.00	.00		.733		.648	.000	.000	.000	.000	000	.3602
		9	.00	1.00	.00		.702	.022	.645	.000	.000	.000	.000	.000	.6904
10	75	1	.00	.76	.24										.3676
		2	.00	.76	.24	1.065									.2533
		3	.00	.76	.24	.816	1.179								.3147
		4	.00	.76	.24	.813	1.145	.030							.3085
		5	.00	.76	.24	1.049		.045							.2496*
		6	.00	.76	.24			.050	1.014	1.059	1.055	1.061	.000	.000	.2618
		7	.00	.76	.24				1.037	1.069	1.071	1.081	.000	.000	.2692
		8	.00	.76	.24		1.291		.748	.802	.794	.832	.000	.000	.3356
		9	.00	.76	.24		1.253	.034	.740	.803	.791	.825	.000	.000	.3259
11	120	1	.00	1.00	.00										.3006
		2	.00	1.00	.00	.970									.3027
		3	.00	1.00	.00	.848	.600								.3706
		4	.00	1.00	.00	.851	.623	-.017							.3806
		5	.00	1.00	.00	.973		-.009							.3085
		6	.00	1.00	.00			.002	1.045	.986	.977	.964	.947	.932	.3050
		7	.00	1.00	.00				1.046	.986	.978	.965	.948	.934	.2986*+
		8	.00	1.00	.00		.485		.939	.886	.874	.872	.852	.841	.3591
		9	.00	1.00	.00		.498	-.008	.939	.885	.874	.873	.854	.843	.3686

Appendix 2. (Continued).

Obs. Number	N	Pro- cedure	Weights- w_i			Regression Coefficients									ASPE
			Best	High	Low	β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	
12	61	1	.00	.00	1.00										.5761
		2	.00	.00	1.00	.901									.4221
		3	.00	.00	1.00	.995	-.484								.3406*+
		4	.00	.00	1.00	.993	-.513	.013							.3515
		5	.00	.00	1.00	.900		.001							.4424
		6	.00	.00	1.00			.000	.851	.918	.940	.000	.000	.000	.4452
		7	.00	.00	1.00				.852	.918	.940	.000	.000	.000	.4244
		8	.00	.00	1.00		-.382		.930	.991	1.012	.000	.000	.000	.3627
		9	.00	.00	1.00		-.407	.009	.929	.990	1.010	.000	.000	.000	.3766
13	13	1	.00	1.00	.00										.4280
		2	.00	1.00	.00	.909									.2811*+
		3	.00	1.00	.00	.687	1.196								.3581
		4	.00	1.00	.00	.689	.990	.052							.4539
		5	.00	1.00	.00	.850		.084							.2844
		6	.00	1.00	.00			.084	.000	.850	.000	.000	.000	.000	.2844
		7	.00	1.00	.00				.000	.909	.000	.000	.000	.000	.2811
		8	.00	1.00	.00		1.196		.000	.687	.000	.000	.000	.000	.3582
		9	.00	1.00	.00		.990	.052	.000	.689	.000	.000	.000	.000	.4537
14	7	1	1.00	.00	.00										.1037
		2	1.00	.00	.00	1.036									.0852*
		3	1.00	.00	.00	.976	.306								.1531
		4	1.00	.00	.00	.999	.275	-.029							.2187
		5	1.00	.00	.00	1.054		-.030							.1016
		6	1.00	.00	.00			-.030	.000	.000	1.054	.000	.000	.000	.1016
		7	1.00	.00	.00				.000	.000	1.036	.000	.000	.000	.0852
		8	1.00	.00	.00		.306		.000	.000	.976	.000	.000	.000	.1531
		9	1.00	.00	.00		.274	-.029	.000	.000	.999	.000	.000	.000	.2187
15	16	1	1.00	.00	.00										.1968*
		2	1.00	.00	.00	1.013									.1991
		3	1.00	.00	.00	1.044	-.159								.2246
		4	1.00	.00	.00	1.027	.055	-.049							.2597
		5	1.00	.00	.00	1.038		-.048							.2193
		6	1.00	.00	.00			-.048	.000	1.038	.000	.000	.000	.000	.2193
		7	1.00	.00	.00				.000	1.013	.000	.000	.000	.000	.1991
		8	1.00	.00	.00		-.160		.000	1.045	.000	.000	.000	.000	.2245
		9	1.00	.00	.00		.055	-.050	.000	1.027	.000	.000	.000	.000	.2596
16	21	1	.18	.00	.82										.1857
		2	.17	.00	.83	1.038									.1557
		3	.17	.00	.83	1.063	-.109								.1613
		4	.17	.00	.83	1.064	-.272	.090							.1316*
		5	.17	.00	.83	1.004		.085							.1380
		6	.17	.00	.83			.077	1.022	.988	.000	.000	.000	.000	.1490
		7	.17	.00	.83				1.060	1.010	.000	.000	.000	.000	.1601
		8	.17	.00	.83		-.237		1.117	1.060	.000	.000	.000	.000	.1560
		9	.18	.00	.82		-.355	.083	1.104	1.063	.000	.000	.000	.000	.1351
17	38	1	.00	.00	1.00										.4281
		2	.00	.00	1.00	.901									.3089
		3	.00	.00	1.00	1.073	-.782								.1919
		4	.00	.00	1.00	1.056	-.871	.060							.1787
		5	.00	.00	1.00	.875		.042							.3112
		6	.00	.00	1.00			.042	.794	.875	.944	.919	.000	.000	.3545
		7	.00	.00	1.00				.817	.902	.973	.939	.000	.000	.3124
		8	.00	.00	1.00		-.854		1.016	1.078	1.167	1.146	.000	.000	.1640
		9	.00	.00	1.00		-.975	.065	1.006	1.061	1.149	1.143	.000	.000	.1579*+

Appendix 2. (Continued).

Obs. Number	N	Procedure	Weights- w_i Best	High	Low	Regression Coefficients β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	ASPE
18	80	1	.00	.00	1.00										.7877
		2	.02	.00	.98	.847									.3049
		3	.02	.00	.98	.894	-.244								.2745*
		4	.02	.00	.98	.858	-.173	.039							.2813
		5	.02	.00	.98	.822		.042							.3014
		6	.00	.00	1.00			.025	.000	.000	.782	.842	.850	.858	.2945
		7	.02	.00	.98				.000	.000	.791	.863	.870	.868	.2895
		8	.00	.00	1.00		-.118		.000	.000	.815	.884	.890	.889	.2802
		9	.00	.00	1.00		-.082	.023	.000	.000	.800	.860	.866	.873	.2903
19	60	1	.57	.00	.43										.2170*
		2	.58	.00	.42	.981									.2265
		3	.57	.00	.43	1.001	-.102								.2205
		4	.58	.00	.42	.989	-.137	.031							.2215
		5	.58	.00	.42	.962		.030							.2309
		6	.57	.00	.43			.029	.952	.959	1.001	.000	.000	.000	.2427
		7	.57	.00	.43				.970	.976	1.020	.000	.000	.000	.2371
		8	.58	.00	.42		-.079		.986	.992	1.036	.000	.000	.000	.2348
		9	.57	.00	.43		-.115	.030	.975	.980	1.023	.000	.000	.000	.2367
20	28	1	1.00	.00	.00										.3152
		2	1.00	.00	.00	.944									.2860
		3	1.00	.00	.00	.874	.342								.3335
		4	1.00	.00	.00	.872	.323	.010							.3574
		5	1.00	.00	.00	.932		.019							.3096
		6	1.00	.00	.00			.016	.947	.000	.879	.000	.000	.000	.2894
		7	1.00	.00	.00				.957	.000	.888	.000	.000	.000	.2688*
		8	1.00	.00	.00		.329		.889	.000	.822	.000	.000	.000	.3116
		9	1.00	.00	.00		.315	.007	.887	.000	.821	.000	.000	.000	.3331
21	53	1	.00	.78	.22										.4181
		2	.00	.78	.22	.941									.3949
		3	.00	.79	.21	.878	.309								.4540
		4	.00	.79	.21	.864	.313	.020							.4760
		5	.00	.78	.22	.927		.019							.4147
		6	.00	.78	.22			.018	.000	.000	1.020	.884	.000	.917	.3580
		7	.00	.78	.22				.000	.000	1.033	.898	.000	.930	.3387*+
		8	.00	.78	.22		.421		.000	.000	.952	.811	.000	.844	.4028
		9	.00	.78	.22		.423	.019	.000	.000	.937	.796	.000	.831	.4238
22	34	1	.00	.00	1.00										.2437
		2	.00	.00	1.00	.952									.2312
		3	.00	.00	1.00	.964	-.056								.2360
		4	.00	.00	1.00	.906	.044	.059							.2401
		5	.00	.00	1.00	.917		.058							.2219*
		6	.00	.00	1.00			.045	.000	.000	.000	.934	.000	.899	.2282
		7	.00	.00	1.00				.000	.000	.000	.967	.000	.913	.2291
		8	.00	.00	1.00		.052		.000	.000	.000	.956	.000	.900	.2481
		9	.00	.00	1.00		.093	.046	.000	.000	.000	.913	.000	.877	.2539
23	12	1	.00	.00	1.00										1.2346
		2	.00	.00	1.00	.788									.2935*+
		3	.00	.00	1.00	.563	1.184								.4222
		4	.00	.00	1.00	.574	1.180	-.025							.4683
		5	.00	.00	1.00	.799		-.027							.3163
		6	.00	.00	1.00			-.027	.000	.000	.000	.799	.000	.000	.3163
		7	.00	.00	1.00				.000	.000	.000	.788	.000	.000	.2935
		8	.00	.00	1.00		1.184		.000	.000	.000	.563	.000	.000	.4222
		9	.00	.00	1.00		1.180	-.025	.000	.000	.000	.574	.000	.000	.4682

Appendix 2. (Continued).

Obs. Number	N	Procedure	Weights- w_i			Regression Coefficients									ASPE
			Best	High	Low	β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	
24	25	1	1.00	.00	.00										.4159
		2	1.00	.00	.00	.933									.3726*
		3	1.00	.00	.00	.725	1.197								.4579
		4	1.00	.00	.00	.738	1.169	-.010							.5099
		5	1.00	.00	.00	.970		-.048							.4069
		6	1.00	.00	.00			-.048	.000	.000	.000	.000	.970	.000	.4069
		7	1.00	.00	.00				.000	.000	.000	.000	.933	.000	.3726
		8	1.00	.00	.00		1.197		.000	.000	.000	.000	.725	.000	.4577
		9	1.00	.00	.00		1.169	-.010	.000	.000	.000	.000	.738	.000	.5099
25	28	1	.00	.00	1.00										.7510
		2	.00	.00	1.00	.902									.5894
		3	.00	.00	1.00	.973	-.397								.5123*
		4	.00	.00	1.00	1.010	-.474	-.029							.5299
		5	.00	.00	1.00	.909		-.009							.6252
		6	.00	.00	1.00			-.009	.000	.000	.000	.000	.909	.000	.6252
		7	.00	.00	1.00				.000	.000	.000	.000	.902	.000	.5894
		8	.00	.00	1.00		-.396		.000	.000	.000	.000	.973	.000	.5122
		9	.00	.00	1.00		-.474	-.029	.000	.000	.000	.000	1.010	.000	.5299
26	53	1	1.00	.00	.00										.7699
		2	1.00	.00	.00	.858									.2889*
		3	1.00	.00	.00	.810	.260								.3276
		4	1.00	.00	.00	.833	.244	-.029							.3236
		5	1.00	.00	.00	.878		-.031							.2866
		6	1.00	.00	.00			-.029	.000	.000	.000	.000	.873	.881	.2941
		7	1.00	.00	.00				.000	.000	.000	.000	.851	.865	.2947
		8	1.00	.00	.00		.241		.000	.000	.000	.000	.809	.818	.3351
		9	1.00	.00	.00		.236	-.028	.000	.000	.000	.000	.832	.836	.3334
27	58	1	.00	1.00	.00										.5440
		2	.00	1.00	.00	.883									.2516
		3	.00	1.00	.00	.735	.777								.2877
		4	.00	1.00	.00	.757	.776	-.032							.2867
		5	.00	1.00	.00	.905		-.033							.2530
		6	.00	1.00	.00			-.023	.000	.000	.000	.000	.866	.927	.2309
		7	.00	1.00	.00				.000	.000	.000	.000	.850	.913	.2264*+
		8	.00	1.00	.00		.731		.000	.000	.000	.000	.714	.771	.2475
		9	.00	1.00	.00		.732	-.023	.000	.000	.000	.000	.731	.785	.2529
30	33	1	.00	.21	.79										.7802
		2	.00	.21	.79	.859									.4191
		3	.00	.21	.79	.674	.943								.6819
		4	.00	.21	.79	.668	1.013	-.016							.7888
		5	.00	.21	.79	.851		.018							.4404
		6	.00	.21	.79			.047	.000	.000	.000	.000	.798	.881	.3982*
		7	.00	.21	.79				.000	.000	.000	.000	.826	.893	.4020
		8	.00	.21	.79		.965		.000	.000	.000	.000	.635	.705	.6427
		9	.00	.21	.79		.912	.013	.000	.000	.000	.000	.638	.712	.6742
31	47	1	.75	.25	.00										.3460*+
		2	.75	.25	.00	.986									.3643
		3	.75	.25	.00	.929	.290								.4105
		4	.75	.25	.00	.930	.334	-.016							.4364
		5	.75	.25	.00	.989		-.005							.3810
		6	.75	.25	.00			-.003	.000	.000	.000	.000	.943	1.012	.3842
		7	.75	.25	.00				.000	.000	.000	.000	.942	1.011	.3635
		8	.75	.25	.00		.308		.000	.000	.000	.000	.881	.951	.4230
		9	.75	.25	.00		.348	-.014	.000	.000	.000	.000	.881	.951	.4539

Appendix 2. (Continued).

Obs. Number	N	Procedure	Weights- w_i Best	High	Low	β_0	β_1	β_2	β_{87}	β_{88}	β_{89}	β_{90}	β_{92}	β_{93}	ASPE
32	33	1	.00	.63	.37										.5301
		2	.00	.63	.37	.942									.5122*+
		3	.00	.63	.37	.878	.328								.6073
		4	.00	.63	.37	.863	.481	-.039							.6962
		5	.00	.63	.37	.951		-.022							.5525
		6	.00	.63	.37			-.009	.000	.000	.000	.000	.922	.966	.5954
		7	.00	.63	.37				.000	.000	.000	.000	.916	.964	.5370
		8	.00	.63	.37		.306		.000	.000	.000	.000	.857	.903	.6460
		9	.00	.63	.37		.409	-.025	.000	.000	.000	.000	.851	.889	.7309
33	36	1	.00	.22	.78										1.0431
		2	.00	.22	.78	.818									.4515
		3	.00	.22	.78	.879	-.301								.4032*+
		4	.00	.22	.78	.881	-.349	.016							.4217
		5	.00	.22	.78	.816		.005							.4811
		6	.00	.22	.78			.005	.000	.000	.000	.000	.817	.815	.5128
		7	.00	.22	.78				.000	.000	.000	.000	.819	.817	.4787
		8	.00	.22	.78		-.301		.000	.000	.000	.000	.879	.879	.4278
		9	.00	.22	.78		-.356	.017	.000	.000	.000	.000	.880	.884	.4475
35	19	1	.77	.23	.00										.5259
		2	.77	.23	.00	1.043									.5526
		3	.78	.22	.00	1.243	-.968								.3837*
		4	.77	.23	.00	1.243	-.973	.003							.4404
		5	.77	.23	.00	1.050		-.020							.6407
		6	.78	.22	.00			-.020	.000	.000	.000	.000	.000	1.050	.6406
		7	.78	.22	.00				.000	.000	.000	.000	.000	1.042	.5526
		8	.77	.23	.00		-.967		.000	.000	.000	.000	.000	1.243	.3836
		9	.78	.22	.00		-.974	.003	.000	.000	.000	.000	.000	1.243	.4404
36	9	1	.41	.59	.00										1.0691
		2	.41	.59	.00	.862									.8200*
		3	.41	.59	.00	.681	.999								2.3047
		4	.41	.59	.00	.714	.777	.028							2.2945
		5	.41	.59	.00	.848		.053							.9618
		6	.41	.59	.00			.053	.000	.000	.000	.000	.000	.848	.9616
		7	.41	.59	.00				.000	.000	.000	.000	.000	.862	.8200
		8	.41	.59	.00		1.001		.000	.000	.000	.000	.000	.680	2.3084
		9	.41	.59	.00		.776	.028	.000	.000	.000	.000	.000	.714	2.2942
37	30	1	.00	1.00	.00										.6325
		2	.00	1.00	.00	.875									.3486*+
		3	.00	1.00	.00	.783	.471								.4481
		4	.00	1.00	.00	.787	.494	-.014							.4774
		5	.00	1.00	.00	.878		-.004							.3714
		6	.00	1.00	.00			-.004	.000	.000	.000	.000	.000	.878	.3714
		7	.00	1.00	.00				.000	.000	.000	.000	.000	.875	.3486
		8	.00	1.00	.00		.471		.000	.000	.000	.000	.000	.783	.4478
		9	.00	1.00	.00		.494	-.014	.000	.000	.000	.000	.000	.788	.4774

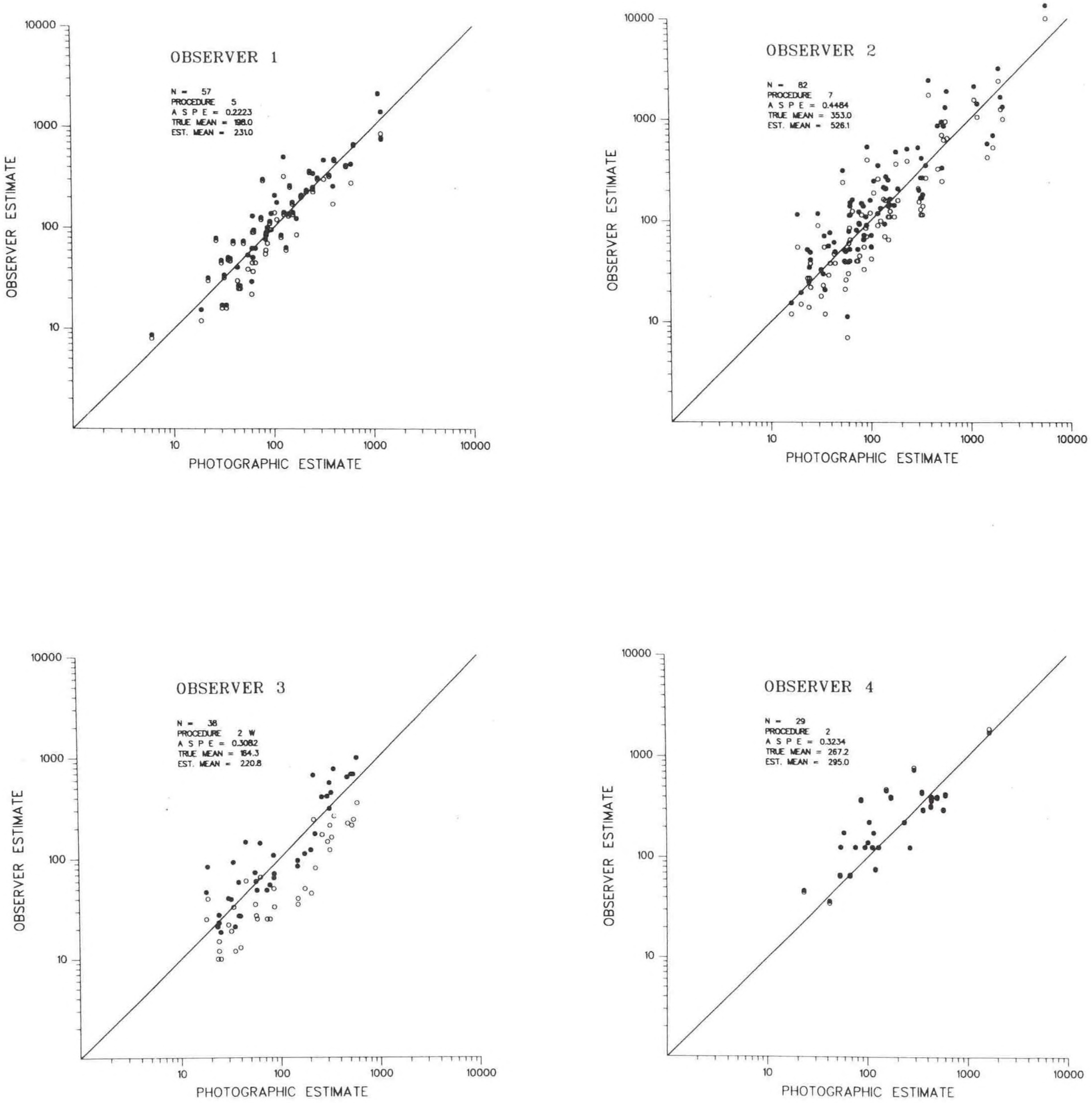
Appendix 3. Regression coefficients, β_0, estimated for the indirect calibration of group size based on a comparison of an individual observer's "best" estimates of group size with the mean calibrated group size estimated from two or more other "calibrated" observers for a given year. ASPE indicates the average squared predition error using this regression coefficient. Sample size for all years is indicated by N.

Year Observer Number	N	β_0	ASPE
1986			
8	85	.891	.4216
9	53	.911	.2283
10	121	.997	.2209
11	80	1.015	.2757
15	96	.958	.2122
16	124	1.023	.1649
17	82	.907	.3586
19	70	.985	.2518
20	138	1.053	.1643
38	11	.966	.7026
39	26	.999	.1190
40	33	.932	.2580
41	42	.870	.2938
42	86	1.040	.2381
43	116	.999	.2195
1987			
4	33	.985	.3280
5	151	1.027	.1725
6	145	.991	.3063
7	89	.936	.2408
8	137	1.015	.1490
9	129	.862	.1507
10	141	1.000	.1134
11	128	1.023	.1769
12	133	.904	.2270
16	167	1.059	.1295
17	95	.908	.3223
19	92	.989	.1606
20	157	1.026	.1974
1988			
4	100	1.027	.2671
5	112	1.090	.1399
6	141	1.024	.2703
7	104	.987	.1651
8	117	1.016	.1856
10	105	1.038	.1617
11	118	1.014	.2112
12	115	.976	.2528
13	36	.898	.1362
15	87	.977	.1930
16	117	1.022	.1608
17	91	.974	.2448
19	94	1.001	.2276
1989			
1	85	1.013	.3857
2	146	.902	.2161
3	130	.869	.2262
7	108	1.027	.2152
10	106	.999	.1767
11	124	.973	.1774
12	87	.921	.2573
14	50	1.031	.1776
17	53	.999	.3316
18	128	.814	.3213
19	54	1.002	.2261
21	87	.979	.4071
1990			
1	141	.941	.3541
2	61	1.043	.2035
3	104	.806	.2970
5	107	.941	.2633
7	106	1.049	.1500
10	108	1.069	.1631
11	105	.906	.3386
17	95	.888	.3336
18	106	.990	.2833
21	144	.950	.2899
22	155	.872	.3537
23	98	.894	.5404
44	54	.856	.4263
1991			
11	71	.957	.1859
18	134	.922	.1494
22	137	.987	.1157
23	63	.983	.1016
24	169	.882	.2210
45	20	.890	.1324
46	31	.914	.2673
47	32	.851	.3765
48	52	.848	.2025
49	170	.763	.3374
50	15	.957	.0803

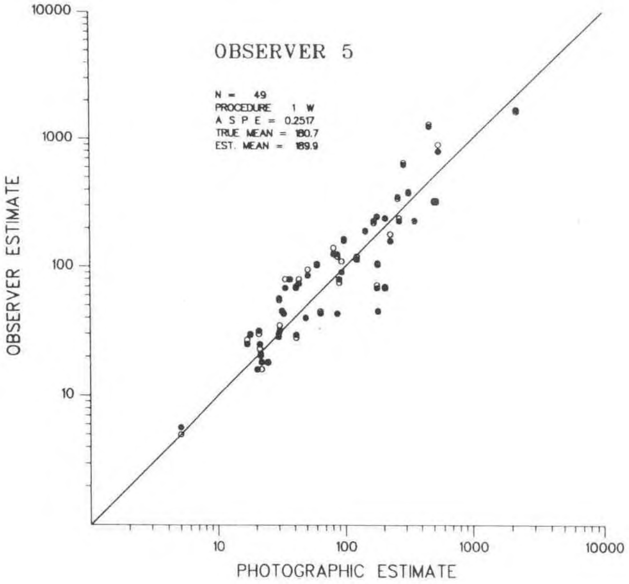
Appendix 3. (Continued)

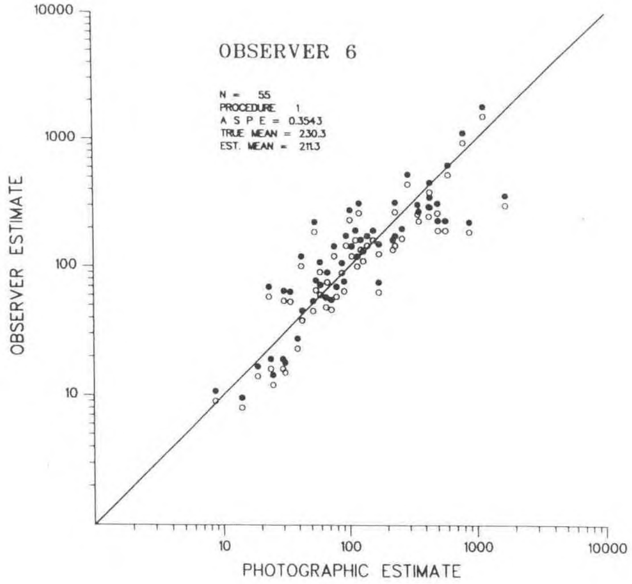
Year Observer Number	N	β_0	ASPE
1992			
1	117	1.023	.2128
2	129	.983	.3064
11	130	.953	.3053
18	108	.981	.2322
24	100	.906	.1582
25	109	.910	.2119
26	115	.806	.2737
27	106	.876	.1890
30	130	.805	.2961
31	128	1.007	.3662
32	116	.863	.3167
33	122	.861	.2524
63	16	.889	.2649
1993			
1	144	1.008	.1990
2	53	1.001	.2007
11	202	.928	.1925
18	43	.934	.2800
21	92	.915	.2029
22	86	1.001	.2290
26	208	.827	.2320
27	192	.855	.1735
30	112	.929	.2611
31	188	.973	.3119
32	112	.933	.2556
33	121	.900	.1794
35	120	.946	.5699
36	32	.886	.1835
37	210	.805	.3411
64	149	.895	.2446
1996			
30	10	.878	.1850
52	12	1.008	.2701
53	15	.924	.0812
58	33	.895	.2261
59	19	.839	.1341
60	20	.916	.1556
All Years Pooled			
1	494	0.992	.2948
2	396	0.956	.3028
3	234	0.843	.2794
4	133	1.017	.2845
5	370	1.019	.2390
6	286	1.008	.2911
7	407	1.002	.2202
8	339	0.977	.2795
9	182	0.879	.1840
10	581	1.019	.1771
11	958	0.971	.2527
12	335	0.934	.2636
13	36	0.898	.1362
14	55	1.033	.1679
15	183	0.967	.2040
16	408	1.036	.1527
17	416	0.933	.3461
18	519	0.911	.3504
19	310	0.994	.2120
20	300	1.033	.1952
21	323	0.952	.3060
22	378	0.930	.2920
23	161	0.917	.3813
24	269	0.893	.1974
25	109	0.910	.2119
26	323	0.818	.2485
27	298	0.863	.1804
28	17	0.907	.3476
29	14	0.869	.1142
30	252	0.861	.3506
31	316	0.988	.3367
32	228	0.898	.3075
33	252	0.880	.2290
35	120	0.946	.5699
36	32	0.886	.1835
37	210	0.805	.3411
38	11	0.966	.7026
39	26	0.999	.1190
40	33	0.932	.2580
41	42	0.870	.2938
42	86	1.040	.2381
43	116	0.999	.2195
44	54	0.856	.4263
45	21	0.892	.1264
46	33	0.926	.2623
47	32	0.851	.3765
48	54	0.850	.1957
49	170	0.763	.3374
50	15	0.957	.0803
52	12	1.008	.2701
53	15	0.924	.0812
58	33	0.895	.2261
59	19	0.839	.1341
60	20	0.916	.1556
64	149	0.895	.2446

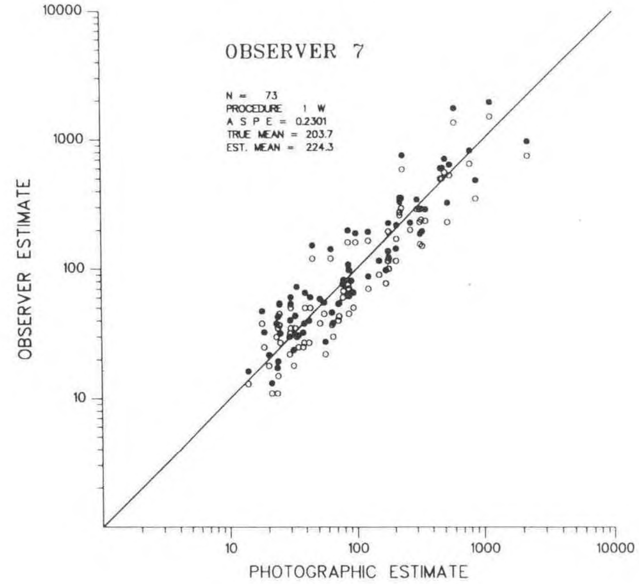
Appendix 4. Directly calibrated estimates of group size (●) plotted against group size from aerial photographs. Uncalibrated "best" estimates of group size are also illustrated (○) when they differ from calibrated estimates. Only the procedure resulting in the lowest average squared prediction error (ASPE) is shown. "W" following a procedure number indicates that the best procedure used a weighted average of "best", "high" and "low" group size estimates. "N" indicates the sample size of calibration groups.

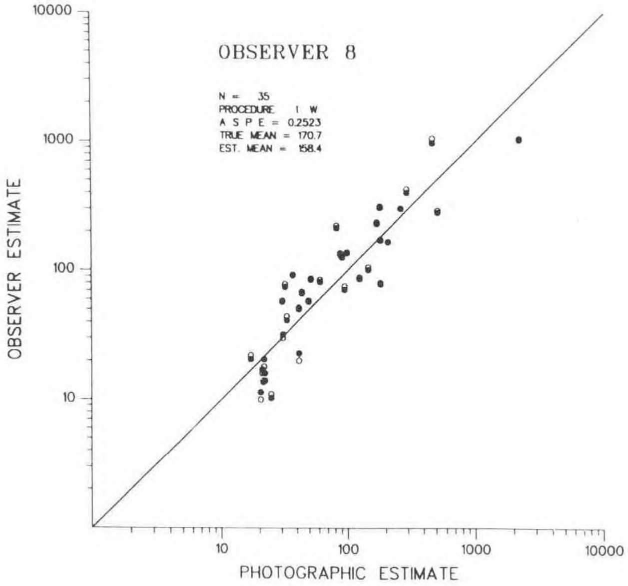


Appendix 4. (Continued).

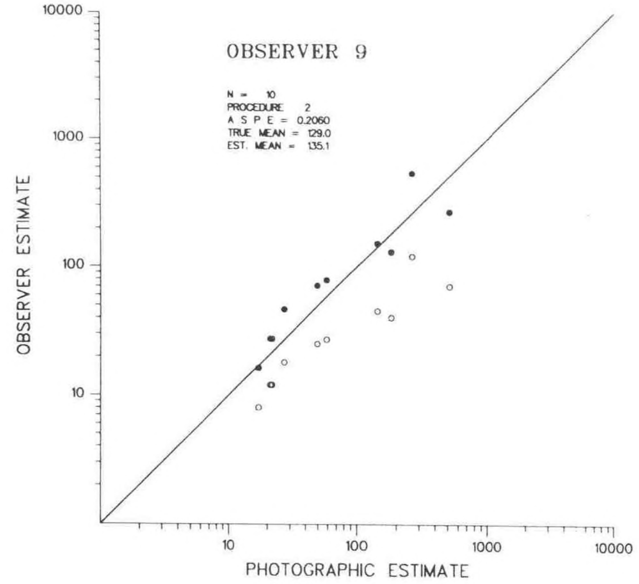


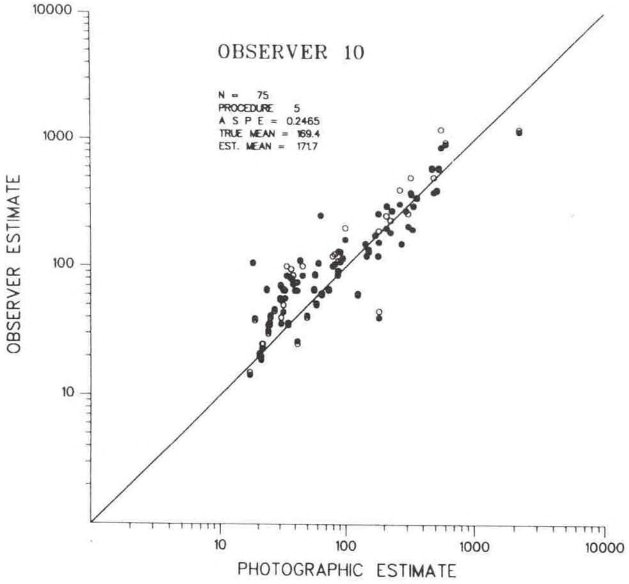


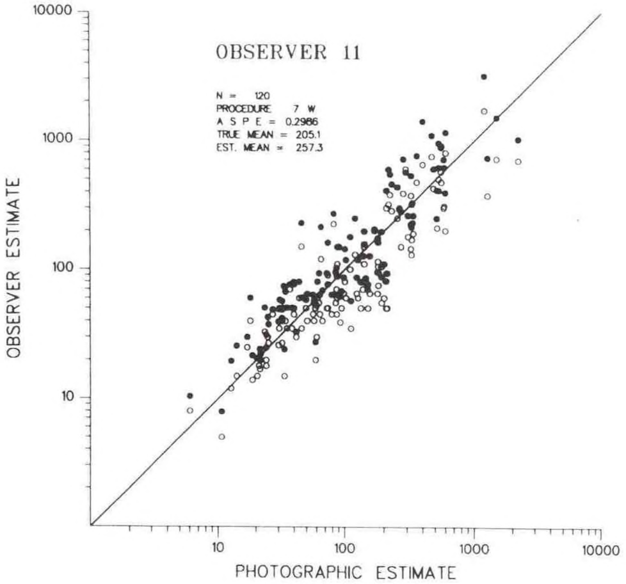


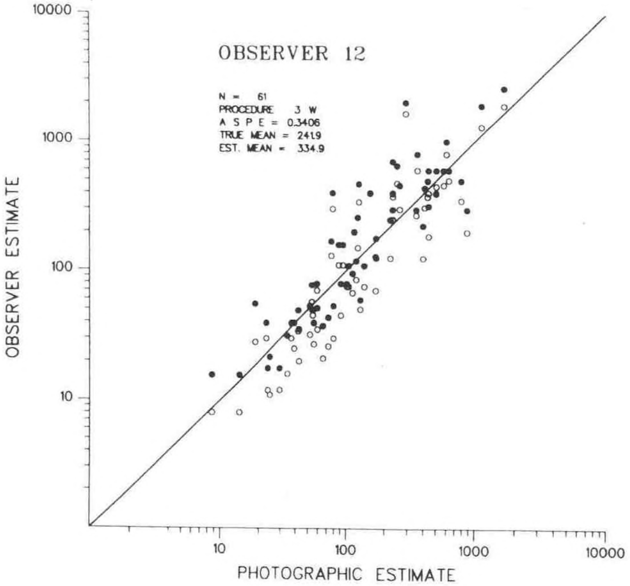


Appendix 4. (Continued).

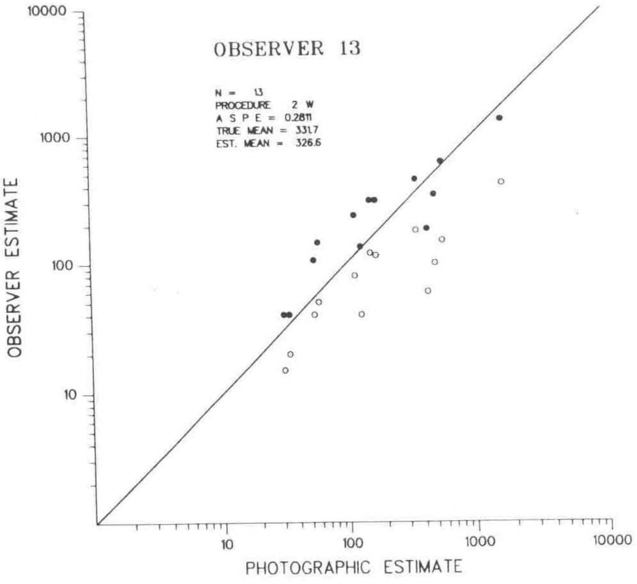


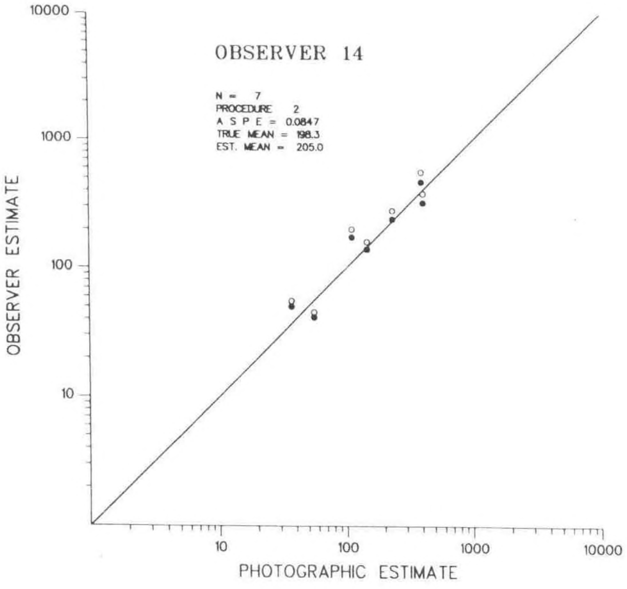


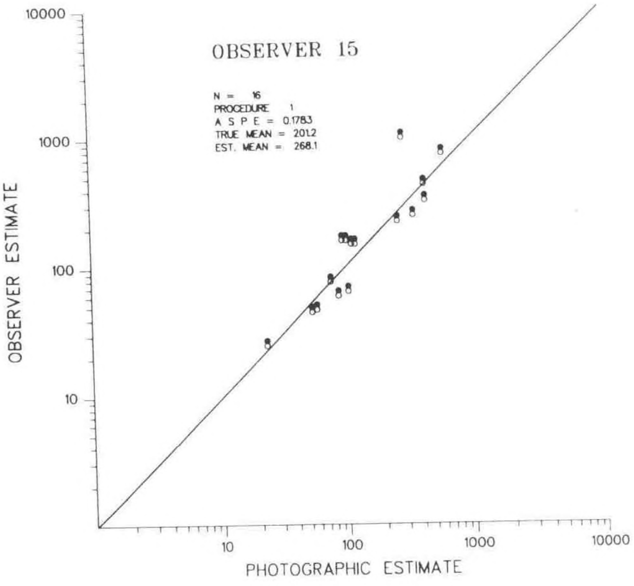


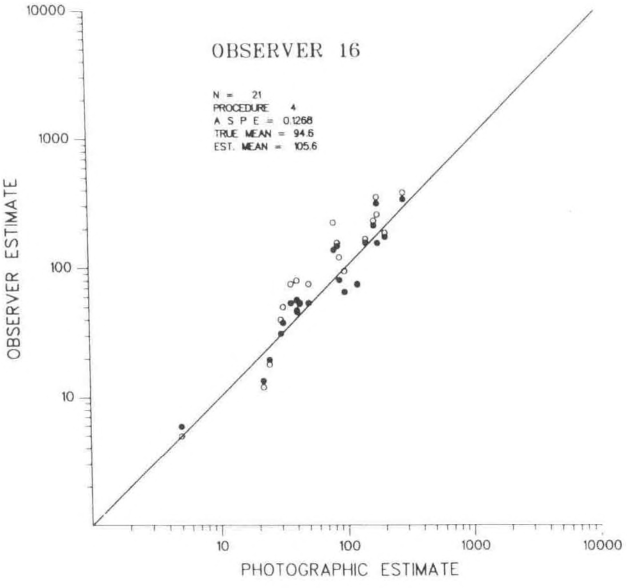


Appendix 4. (Continued).

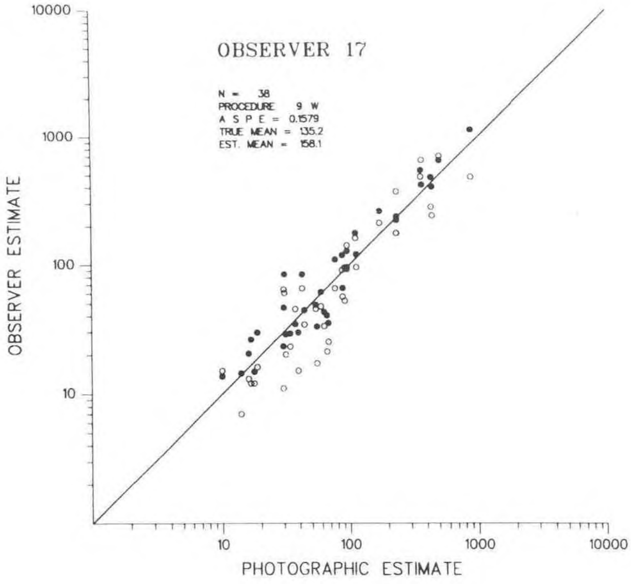


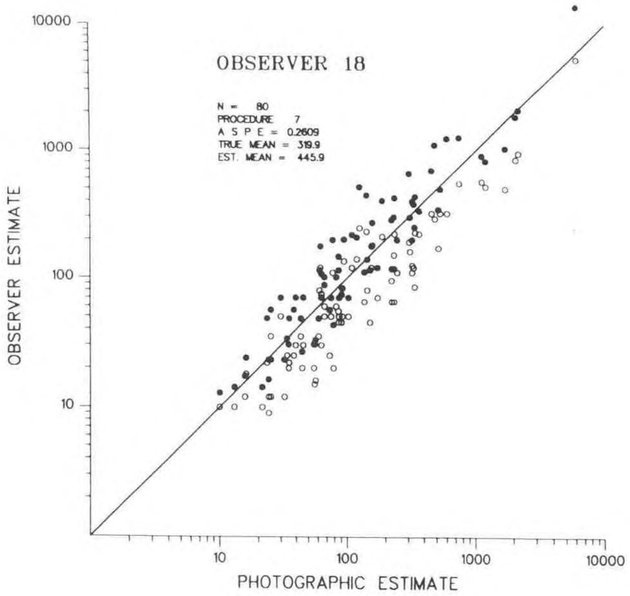


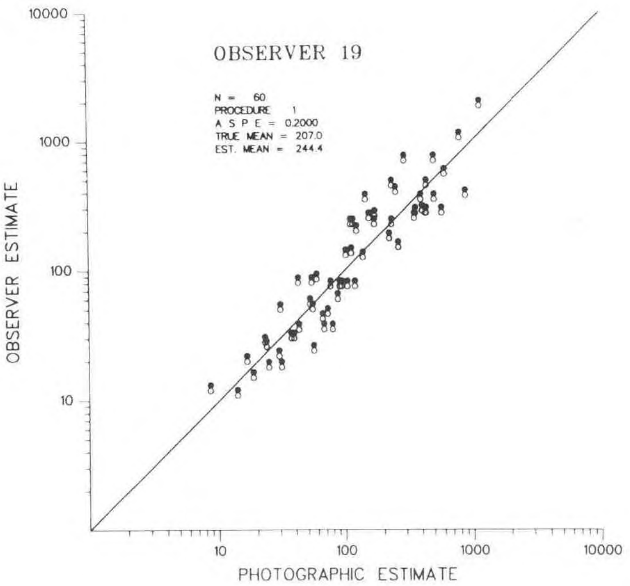


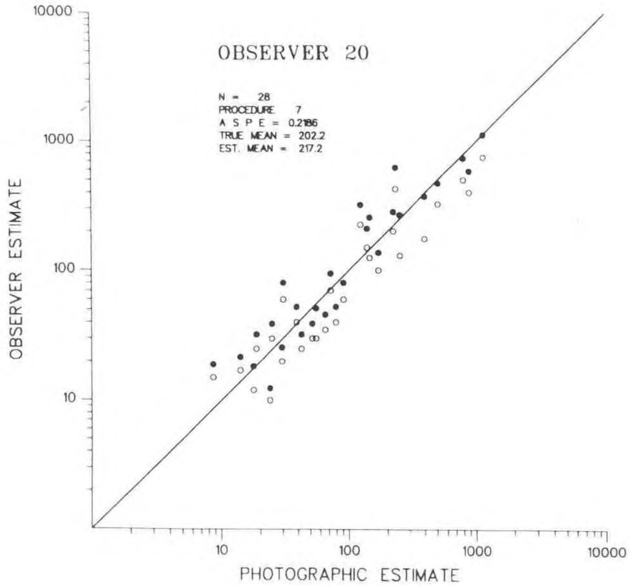


Appendix 4. (Continued).

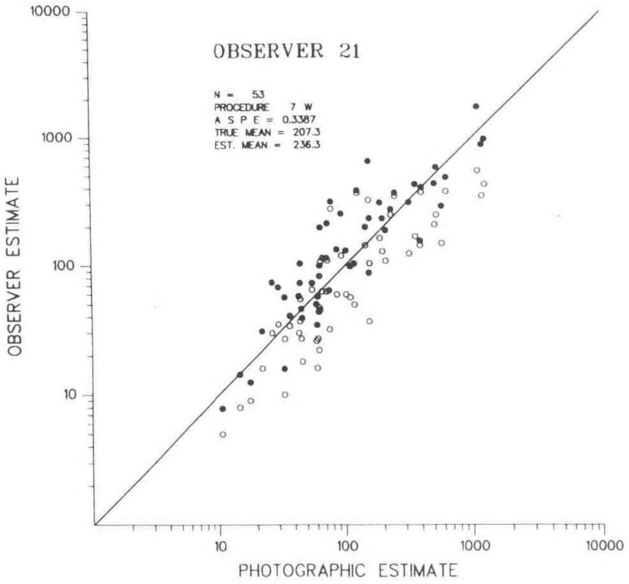


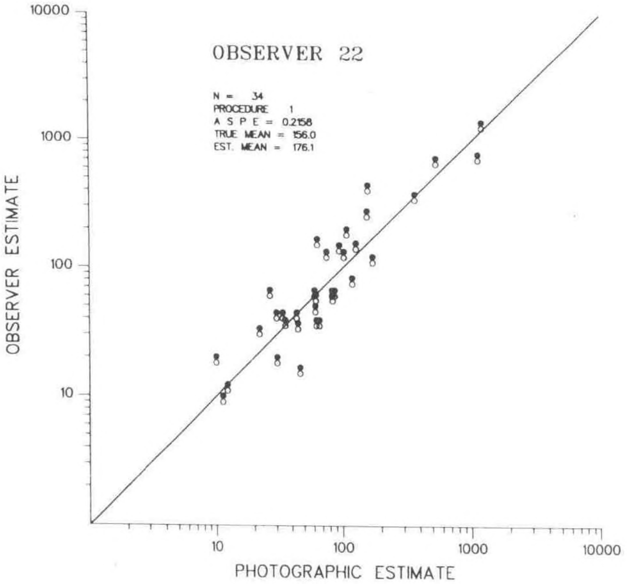


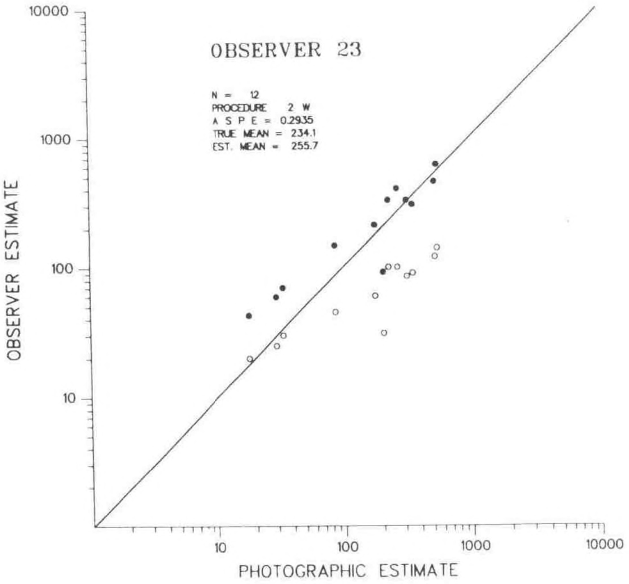


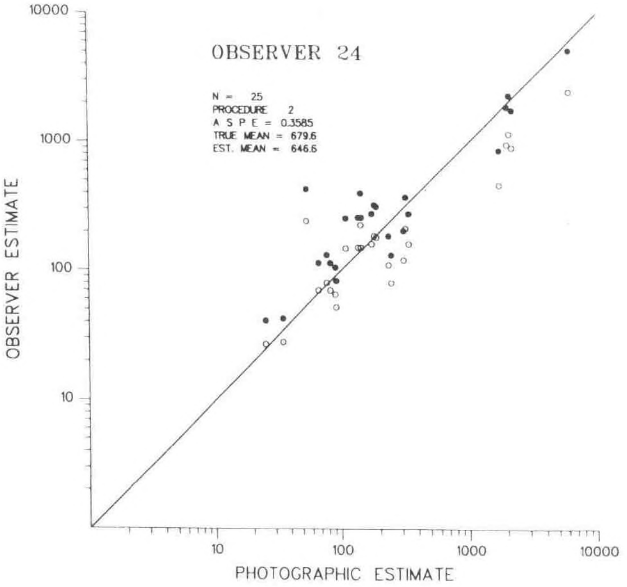


Appendix 4. (Continued).

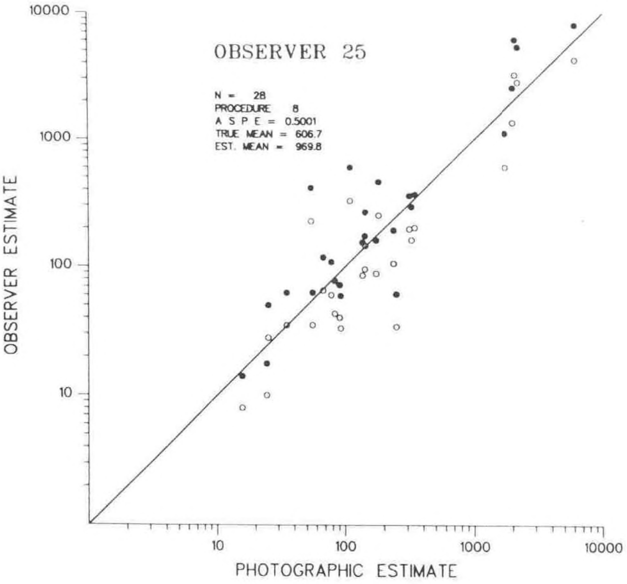


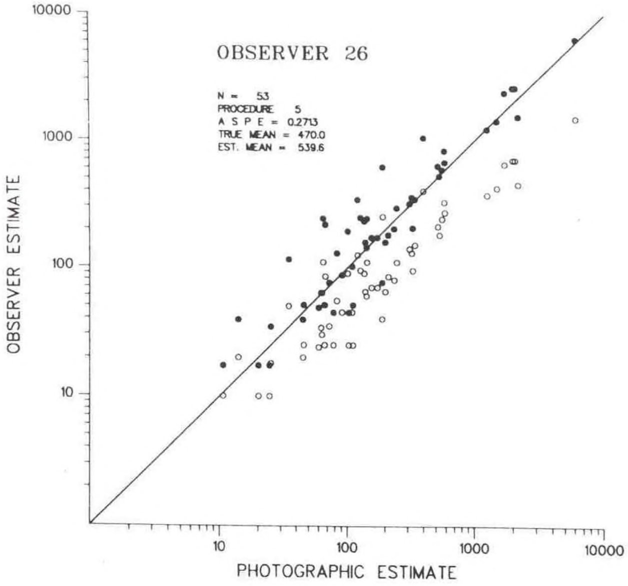


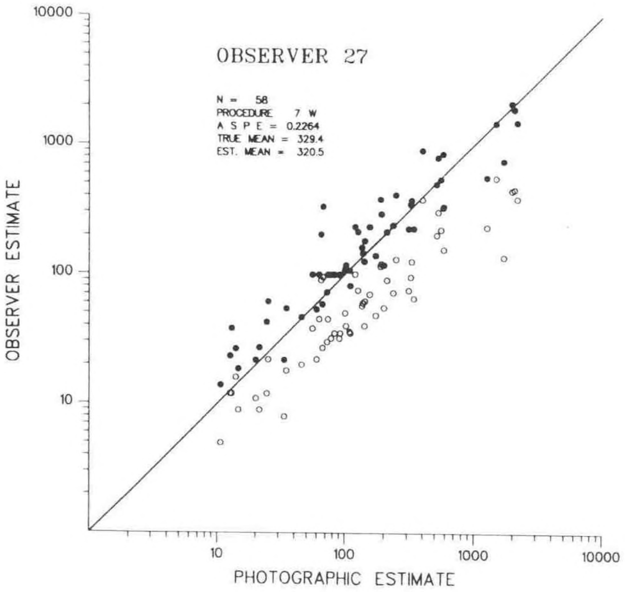


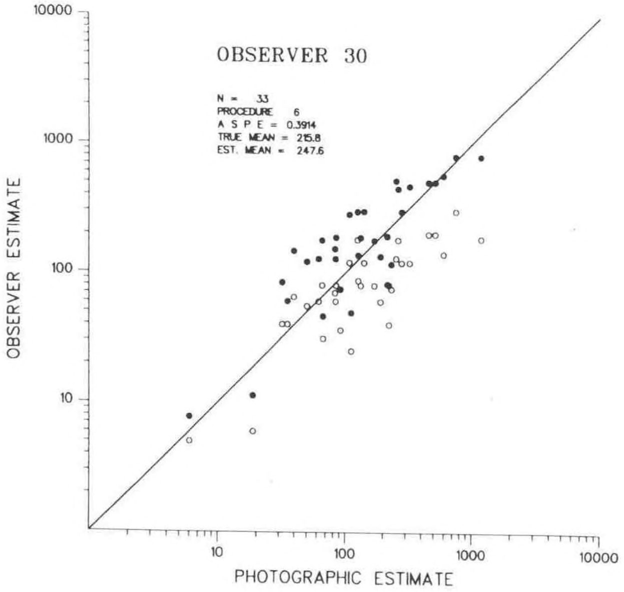


Appendix 4. (Continued).

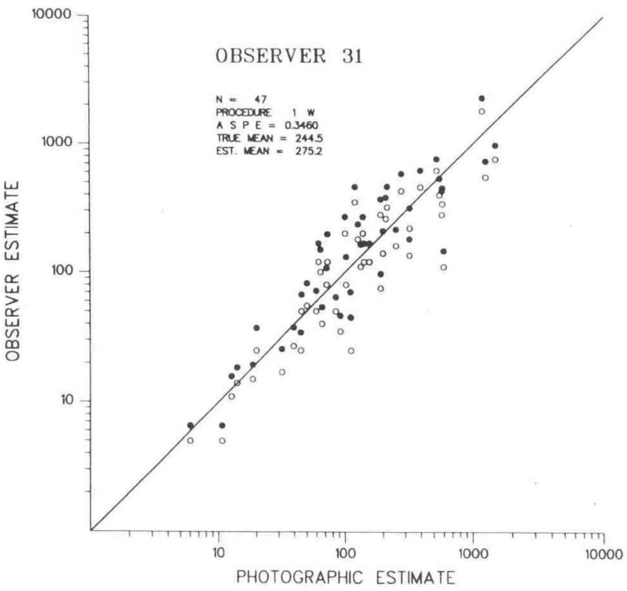


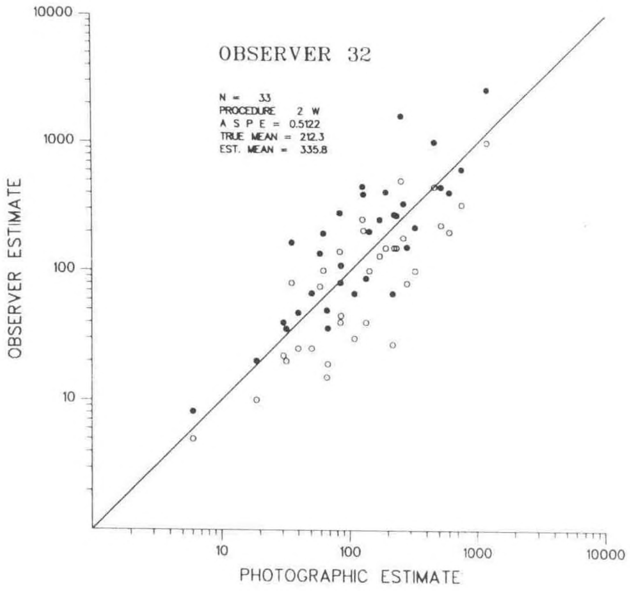


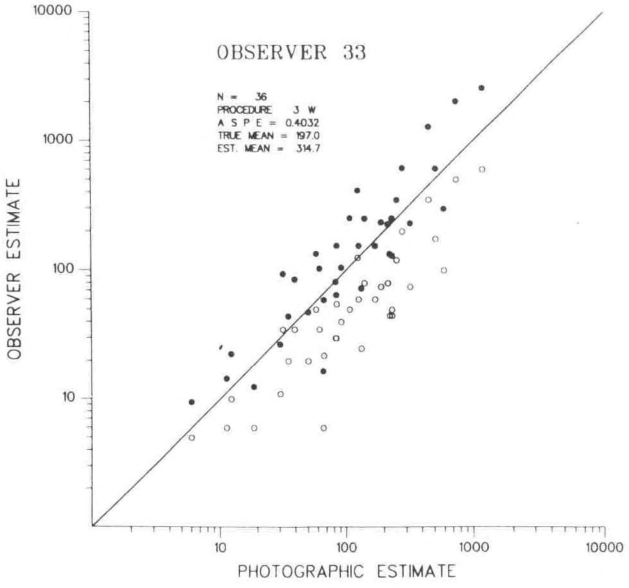


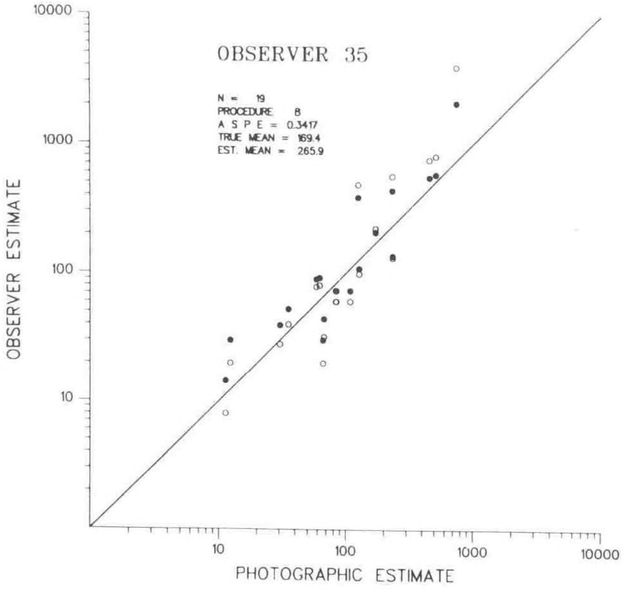


Appendix 4. (Continued).









Appendix 4. (Continued).