Assessing the effects of measurement errors in line transect sampling

Russell Alpízar-Jara, L. A. Stefanski,
*K. H. Pollock, and J.L. Laake**
Department of Statistics
North Carolina State University
Box 8203, Raleigh, NC 27695-8203

National Marine Mammal Laboratory*
Alaska Fisheries Science Center
NMFS, 7600 Sand Point Way N.E.
Seattle, WA 98115

Institute of Statistics
Mimeograph Series No. 2508
North Carolina State University
Raleigh, North Carolina

June, 1998

Abstract

We evaluate the effects of measurement error in population parameter estimation from a line transect sampling model, characterize the error distribution, and illustrate using data from a controlled field experiment (Stakes data set, Laake 1978). We describe a methodology to estimate the measurement error variance from a replicated experiment assuming an additive measurement error model, and correct for systematic measurement error bias in the population estimates using calibration (Carroll et al. 1995). A simulation based method of inference for parametric measurement error models is suggested to correct for random effects measurement error biases (Cook and Stefanski 1994). A simulation study is conducted to show the potential effects of measurement error in a line transect studies. Two main sources of measurement error biases were found in the stake data set, systematic and random effects. In this particular study, the systematic measurement error bias causes overestimation, and the random effects measurement error biases cause underestimation of population size and density. The random effects biases were not as severe as the systematic bias.

KEY WORDS: Calibration, line transect, SIMEX, systematic and random measurement error, wildlife density estimation.

1 Introduction

One of the fundamental assumptions of line transect sampling is that distances from the transect line are measured without error (Buckland *et al.*, 1993). Enough evidence exists in the literature to show that measurement errors have an important effect on the estimates of population size, N, and density D. However, sound approaches to quantify and account for this effect in population estimates have not received enough attention.

In distance sampling, measurement errors are possibly due to improper determination of the centerline. Often systematic errors occur in measurements of perpendicular distances or sighting angles and distances. The accuracy of measurements near the transect line is crucial for estimation of D. If distances near the transect line were rounded to zero distance, the estimate of population size and density are very poor. Errors away from the center line are undesirable but have far less effect on the estimate of density (Schweder 1977). Burnham et al. (1980, p. 105) emphasize the importance of examining the data by means of histograms, order statistics, or other techniques before attempting an analysis. The authors illustrate the problem of measurement errors using some data from Robinette et al. (1974). Their analysis suggests that considerable measurement and rounding errors occurred near the transect line.

Buckland et al. (1993, p. 317) include a section on measurements and emphasize the importance of collecting accurate measurements in distance sampling, and some implications of errors in the measurements. The authors mention several reasons to support the use of perpendicular rather than radial distance models, and discuss possible solutions to cope with measurement errors in distance sampling data. Buckland and Anganuzzi (1988) proposed an *ad hoc* method of estimating the smearing parameters, and compared their method with other existing methods. "Smearing" is a method that has been often used to analyzed cetacean shipboard surveys where sighting distance and angle data are more likely to be collected. The concept of smearing was first introduced by Butterworth (1982). For a description of this method see Buckland et al. (1993, p. 319).

Borchers and Haw (1990) pointed out the need for more accurate estimation methods when analyzing data based on radial distances (sighting distances and angles) rather than perpendicular distances. They conducted experiments on Antarctic sighting surveys and identified several sources of biases and imprecision in the measurements. Schweder et al. (1991) attempted to estimate error and bias in radial measurements based on triangulated data from the parallel ship experiments. Øien and Schweder (1992) obtained estimates of bias and variability in visual distance measurements made by observers.

Recently, a more formal treatment to the problem of measurement error in line transect surveys has been addressed by Chen (1997). Chen proposes a method of moments estimator to correct for measurement error induced bias, and assumes an exponential power series detection function. He also assumes up to a fourth moment of the error distribution, and his method requires knowledge of the side of the transect on which an observation has been collected (i.e. signed perpendicular distances).

In this paper, systematic and random effects of measurement error in line transect data are evaluated using the recent developments in general measurement error theory proposed by Carroll et al. (1995). The next section introduces the basic line transect model and some notation. Section three introduces key concepts of measurement error models as they apply to a line transect sampling study with measurement error. Section four describes a controlled line transect study which is used to illustrate the methods suggested in this paper (Stake data set, Laake 1978). Section five characterizes the distribution for the stake data set using some exploratory data analysis techniques. An additive measurement error model and measurement error variance estimation is discussed in section six. Bias correction using calibration and SIMEX is discussed in section seven. Finally, a discussion and future research directions in section eight concludes.

2 Basic line transect model

A brief description of this model is given here. See Buckland et al. (1993) for a good detailed exposition on theory and applications.

Line transect sampling consists of defining straight lines across an area of known boundaries. Observers travel along the lines (walking, riding on horseback, on a ship, on an airplane, etc.) detecting a sample of target objects, and recording at the moment of detection the perpendicular distance from the line to a detected object, or the sighting distance from the observer and the angle between the line of travel and the line of sight to the object. If possible perpendicular distances are measured, but if not then the sighting distance and angle are usually converted to perpendicular distance by a trigonometric relation (i.e. perpendicular distance equals sighting distance times sine of the sighting angle) (Buckland et al. 1993).

2.1 Assumptions

The basic assumptions underlying the line transect sampling model are:

(*i*) Objects directly on the line have probability one of being detected.

(*ii*) Objects do not move before being detected and no object is counted more than once.

(iii) The measurements of distances and angles are exact (i.e. there are no measurement errors or rounding errors).

(iv) The sighting of one animal is independent of the sighting of another.

These assumptions can easily be violated according to the particular practical conditions of each study. The validity of the model will depend on there being little violation of the assumptions; therefore, the design of the study should attempt to minimize all possible violations. This paper focus on issues related to assumption (iii). This assumption is violated if measurements are only roughly estimated. Extensive work have been done to relax the other assumptions, but little have been done on measurement errors. See Buckland et al. (1993) for a detailed discussion of assumption violations.

2.2 Notation

N_w is the population size in the surveyed area.

L is the total line length in a line transect survey.

w is the half-width of a strip transect of L length.

$A_w = 2wL$ is the surveyed area.

$i = 1, ..., N_w$ is used to index the observations (i.e. objects, animals).

$j = 1, ..., K$ is used to index the replicates (i.e observers in the stakes study).

k_i is the number of replicates in which observation i is detected.

n_j is the number of observations (sample size) detected in replicate j.

X_i are the exact perpendicular sighting distances.

W_{ij} are the observed perpendicular sighting distances measured with error.

U_{ij} are the errors associated to the measurements.

$\underline{\theta}$ is the vector parameter describing the detection function.

$g(\cdot|\underline{\theta})$ is the detection function.

$f(\cdot|\underline{\theta})$ is the probability density function of perpendicular distances.

Although all objects are potentially detectable from the transect line, not all of them need to be detected. Objects that are far away from the transect center line usually have lower probability of being detected than those closer to the transect center line. Modeling this drop in detection probability is central to line transect sampling. This drop is modeled with a decreasing function relating the probability of detecting an animal and its perpendicular distance from the line. This function is known as the detection or sighting function, $g(X_i)$, which properly scaled can be used as a probability density function, $f(X_i)$. Thus, population and density estimates are directly dependent on accurate measurement of the distances because $g(X_i)$ needs to be estimated from the available distance information. If distances are measured inaccurately, the estimate of the detection function will be biased. It can be shown that if the above assumptions, in particular assumption (i), are satisfied then the estimators of density and population size are given by

$$\hat{D} = \frac{\hat{N}_w}{A_w} = \frac{n\hat{f}(0)}{2L} \tag{1}$$

where $\hat{N}_w$ is an estimator of the total population in the area covered, $A_w = 2wL$. L is the total length of the transect strips of width w. It is extremely important to randomly locate lines in the area; otherwise, is necessary to assume that objects are randomly and independently distributed over the population area which in many cases it is an unreasonable assumption. Also, it is recommended that several random lines be used so that robust variances of estimates can easily be obtained.

In equation (1), note that the key parameter to estimate is $f(0)$, the probability density function evaluated at distance zero. This parameter is directly proportional to the population size estimator $\hat{N}_w = nw\hat{f}(0)$, and if distances near the line are measured with error, the estimators of N and D are biased.

3 Measurement error models

In recent years, measurement error models have been widely documented and publicized. Two comprehensive references are Fuller (1987), which gives a detailed exposition of measurement error in linear models, and more recently Carroll *et al.* (1995) on measurement error in nonlinear models. In the remainder of this paper, some of the key concepts of measurement error models as they apply to a line transect sampling study with measurement errors are introduced. The effect of measurement errors on the estimators of density and population size is assessed.

The discussion is based on the existence of an exact predictor X and measurement error models which provide information about this predictor. Our exact predictor is the perpendicular distance measured without error, sometimes also referred as "true" or "exact" distance.

In line transect surveys, we are interested in modeling the probability density function, $f(X|\underline{\theta})$, as a function of the exact perpendicular distance, the predictor X. In practice, X can not be observed exactly for all individuals in the study. However, we can observe a variable W (i.e. observed perpendicular distance) which is related to X. The parametric model $f(X|\underline{\theta})$ can not be estimated directly by fitting $f(\cdot|\underline{\theta})$ to X. Substituting W for X, but making no adjustments in the usual fitting methods (i.e. maximum likelihood or nonlinear least squares) for this substitution, leads to estimates that are biased, sometimes seriously (Carroll et al. 1995). The goal of using measurement error modeling is to obtain nearly unbiased estimates of $\underline{\theta}$.

In assessing measurement errors, careful attention needs to be given to the type and nature of error, and the sources of data which allows modeling of this error. Measurement errors can be due to systematic or random effects. A systematic error is caused by the tendency of consistently underestimating or overestimating the true distances, whereas the random error component changes from one observation to the next, sometimes being positive, sometimes negative, but tending to zero on average. To ascertain the distributional properties of measurement error, one requires a replicated experiment. In this paper, replicates consist of measurements taken by different independent observers.

4 A controlled field study: the stake data set

Laake (1978) conducted an investigation of line transect sampling by placing a known number of wooden stakes in a rectangular area of sagebrush meadow east of Logan, Utah. The stakes were randomly placed and uniformly distributed as a function of distance from the line. Eleven independent observers walked along a transect line of length $L = 1000$ meters and a fixed half- width strip of $w = 20$ meters.

The stake data is a very well known data set that has been analyzed on several occasions. Burnham *et al.* (1980) analyzed this data set and found that the number of observed stakes varied substantially among observers, and that the underlying detection functions $g(X)$ differed greatly among observers. The data provided an estimate of the average density of 31.6 stakes per ha., but the true density was 37.5. Burnham *et al.* (1981) considered that this negative bias is at least partially due to the failure of the assumption that all stakes on the centerline were detected. Buckland *et al.* (1993) have done some additional analyses and found that some of the negative biases in the population estimates are due to measurement error for stakes near the line. Measurement errors in the stake data could also be related to heaping and rounding effects when observers estimate distances from the centerline.

For the stake data set, the exact location of each stake was known, and also the estimated distances for the stakes seen by each of the eleven observers are available. It is possible to identify which stakes were seen by each observer, and an estimate of the measurement error variance can be obtained as we shall show. Another advantage of using the stake data set is that the population size is known ($N = 150$ stakes), hence our population estimates can be compared to the true population parameter. Also, it is an immobile population, and thus we do not have to worry about violation of the assumption that objects do not move before being detected, (*ii*), which also introduces an effect in the measurement of the distances.

A data set like this is difficult to obtain in practice, but it certainly allows illustration of general measurement error methods and emphasizes the importance of collecting some replicate measurements in all studies if possible. Replicate measurements allow the estimation of the measurement error variance. In this study, each observer's estimated distance to a detected stake was considered to be a replicate measurement of the distance to that stake.

5 Characterization of the error distribution

To evaluate the effect of measurement errors in the stake data set, an exploratory data analysis was carried out using the distances measured by the eleven observers. Often, measurement errors are assumed to be normally distributed to simplify analysis and the construction of confidence intervals for population parameters. The error distribution is first characterized instead of blindly assuming normality.

Figure 1 displays the observers errors vs. the "true" distances. Important features of this line transect study can be obtained from examining this plot. Notice for instance that there is a concentration of points near to zero distance, and sometimes a few points far away from the line. There is a clear drop in the sightings as distance increases. It is also possible to appreciate the magnitude and direction of the error. Sometimes errors were as large as one meter or more, and there is a slight tendency to underestimate the true distance. Table 1 summarizes the distribution of the errors by distance categories.

[Table 1 and Figure 1 near here]

A set of exploratory data analysis plots was made for each distance category (i.e. 0-5, 5-10, 10-15, and 15-20 meters). Figure 2a shows a histogram, a boxplot, a density plot, and a normal qq-plot of these data sets. The resulting plots for each distance group indicate that the errors come from a nearly normal distribution. The shape of the histograms, and the linearity of the qq-plots indicate that the error distribution is approximately Gaussian. There may be little departure from normality for a specific distance category, but overall we can assume that the distribution of the errors is close to normal. Figure 2b shows a similar plot, but considering the deviations of the observed distances from their average rather than from the true distance.

[Figures 2a,b near here]

6 The effects of measurement error and estimation of measurement error variance

A simple univariate additive error model is considered, where conditional on X, the errors have mean zero and constant variance, σ_u^2. For a single observation, assume the model $W_{ij} = \beta_0 + \beta_x X_i + U_{ij}$ with $U_{ij} \sim (0, \sigma_u^2)$.

For each stake i that was detected by at least one observer, we compute the average of the observed distances given that a stake was located at true distance X_i (fixed by design). It is reasonable to assume that if there is no effect of measurement errors, the plot of the average of the observed distances ($\overline{W}_{i\cdot}$) versus the true distances (X_i) should fall on a straight line with slope one and intercept zero. We then hypothesize the following model for the average of observed distances,

$$\overline{W}_{i\cdot} = \beta_0 + \beta_x X_i + \overline{U}_{i\cdot} \quad \textit{with} \quad \overline{U}_{i\cdot} \sim N\left(0, \frac{\sigma_u^2}{k_i}\right) \tag{2}$$

where

$$\overline{W}_{i\cdot} = \frac{\sum_{j=1}^{k_i} W_{ij}}{k_i}, 1 \leq k_i \leq K, i = 1, \cdots, N_w.$$

Note also that $E(\overline{W}_{i\cdot}|X_i) = \beta_0 + \beta_x X_i$, and $Var(\overline{W}_{i\cdot}|X_i) = \frac{\sigma_u^2}{k_i}$, where σ_u^2 is the measurement error variance for a single observation.

To find out if there is any systematic measurement error bias the following hypotheses need to be tested,

$$H_0 : \beta_x = 1 \quad \textit{and} \quad H_0 : \beta_0 = 0.$$

We expect that if $E(\overline{W}_{i\cdot}|X_i) = X_i$ then there are not major systematic biases. If $H_0 : \beta_x = 1$ is rejected, then we conclude that there is some effect of measurement error biases in the distances. Determining the magnitude and direction of this effect is also of interest. If $H_0 : \beta_0 = 0$ is rejected, then this indicates that measurement errors have a strong effect on distance measurements close to line (at zero distance).

Since an observation depends on whether a stake was detected or not, the number of replicates for each stake is not fixed. Observations do not contribute equally to the fit. We use weighted least squares (Draper and Smith, 1981) to estimate the parameters of this model. After fitting the model, the data can be calibrated by adjusting the observed averages to obtain the additive error model

$$\overline{W}^*_{i\cdot} = X_i + \overline{U}^*_{i\cdot} \tag{3}$$

where

$$\overline{W}^*_{i\cdot} = \frac{\overline{W}_{i\cdot} - \hat{\beta}_0}{\hat{\beta}_x}, \quad \overline{U}^*_{i\cdot} = \frac{\overline{U}_{i\cdot}}{\hat{\beta}_x} \quad and \quad \overline{U}^*_{i\cdot} \sim N\left(0, \frac{\sigma^2_u}{\hat{\beta}^2_x}\right). \tag{4}$$

Under these model assumptions, an estimate of the measurement error variance after calibration is given by

$$\hat{\sigma}^2_{\overline{u}^*} = \frac{\hat{\sigma}^2_u}{\hat{\beta}^2_x}, \tag{5}$$

which is estimated dividing the mean squared error of the fitted weighted regression model by the the estimate of the slope.

To determine whether there is some degree of heteroscedasticity in the measurement error variance given the "true" distance, one could hypothesize that the variance of the observed distances would increase for observations further away from the transect line. The following could be used as a baseline model,

$$S^2_i = \gamma_0 + \gamma_x X_i + \epsilon_i \quad with \quad \epsilon_i \sim (0, \sigma^2_\epsilon). \tag{6}$$

In this case, for each stake i detected by at least two observers we compute

$$S_i^2 = \frac{1}{k_i - 1}\sum_{j=1}^{k_i}(W_{ij} - \overline{W}_{i\cdot})^2,$$

and then fit the model using least squares weighting by the number of observations in which each S_i^2 is based (i.e. $k_i - 1$). Note that $E(S_i^2|X_i) = \gamma_0 + \gamma_x X_i$, and $Var(S_i^2|X_i) = \frac{\sigma_\epsilon^2}{k_i - 1}$. We are interested in testing the following hypothesis,

$$H_0 : \gamma_x = 0$$

If $H_0 : \gamma_x = 0$ is rejected then the variance of the observed measurements given the true distance increases linearly as true distance increases. If $\gamma_x = 0$, then constant variance should be assumed and an additive model is appropriate. Also, an estimate of the measurement error variance for a single observation is given by the estimate of the intercept ($\hat{\gamma}_0$) since $E(S_i^2|X_i) = \gamma_0$.

Results from the fitted models, $\widehat{\overline{W}}_{i\cdot} = \hat{\beta}_0 + \hat{\beta}_x X_i$ and $\widehat{S}_i^2 = \hat{\gamma}_0 + \hat{\gamma}_x X_i$, using the stake data set are given in Table 2. We used SAS procedure REG (SAS Users Guide) for the analysis. Plots corresponding to these fits are given in Figure 3.

[Table 2 and Figures 3a,b near here]

The null hypothesis $H_0 : \beta_x = 1$ is rejected in favor of the alternative $H_0 : \beta_x < 1$. $\hat{\beta}_x = 0.989$, with standard error 0.004. This result indicates that there is a systematic effect of the measurement errors in the measures of perpendicular distances. Observers tend to underestimate the true perpendicular distance. This result is also confirmed by the exploratory data analysis form the previous section (Figure 1), in which generally more than half of the observations detected by each observer are below the zero line. An estimate of the measurement error variance of the calibrated measurements is given by

$$\frac{\hat{\sigma}_u^2}{\hat{\beta}_x^2} = \frac{0.2826}{(0.9892)^2} = 0.2888.$$

Although we reject the null hypothesis $H_0 : \gamma_x = 0$, ($p = 0.0154$), the model does not give a good fit to the data ($r^2 = 0.067$) (Figure 3). Also, the exploratory data analysis does not suggest strong evidence of heteroscedasticity for this model (Figure 2).

In assessing the measurement error effects we also fitted several parametric models to the data collected by each observer. Models were fitted to both, "true" and observed perpendicular distances. Program DISTANCE (Laake et al. 1993) was used for model fitting using maximum likelihood estimation, and model selection was carried out using the Akaike Information Criteria (Akaike 1973). Models that "best" fitted the data with the corresponding population estimates are shown in Table 3. Note that in most cases a half-normal model was selected as the "best" fit, and in some occasions a uniform detection function with up to three cosine adjustment terms (see models for Observer 6).

[Table 3 near here]

These results suggest that there may be some effects of measurement errors in model selection. For instance, for Observer 1 a uniform model with one adjustment cosine term was selected when fitting the model with the true distances, but a half-normal model with no adjustment terms was selected when using the observed distances. Also, note that in most cases the true population size $N = 150$ is underestimated although the true value is generally contained in the confidence intervals. Possibly, this negative bias is largely due to violation of the assumption that stakes on the centerline have probability one of being observed rather than to the effect of measurement errors.

The other important outcome from these analyses is that for the stake data there is some indication of systematic measurement errors to overestimate population size when models were fitted with observed distances rather than true distances. This positive bias of the population size under the presence of measurement error can be severe. For example for Observer 11, overestimation is about 24.2% of the error-free population parameter estimate. Overestimation in population size in this case is due to the fact that in this data set true distances were consistently underestimated. To correct for this systematic bias, calibration of the data is required.

7 Bias correction using calibration and SIMEX

In this section we use some stake data and simulations to illustrate bias correction using calibration to account for the systematic measurement error effects, and to demonstrate the random measurement error effects on population size. We also describe and suggest the SIMEX algorithm (Cook and Stefanski, 1994) to correct for bias due to the random component.

For the sake of illustration the data collected by Observer 4 has been chosen, and fitted a half-normal model with no adjustment terms for the true and observed distances. The maximum likelihood estimates of population size and their standard errors obtained from these fits are $\hat{N}_w = 122(19.3)$ and $\hat{N}_w = 125(19.6)$ for the true and observed distances respectively. As noticed from the data analyses for all the observers in the preceding section, overestimation is caused by the tendency of each observer to underestimate the true perpendicular distances. Observer 4 was not an exception. We calibrate the observed data as described earlier, (see equation 4 in previous section), to reduce this systematic bias. After calibration of the observed distances, we fitted the same model and obtained an estimate of population size of $\hat{N}_w = 123(19.3)$. Note that this population estimate is now closer to the error-free population estimate and we have corrected for some of this systematic effect.

The simulations to be described are based on the SIMEX algorithm. SIMEX is a simulation-based method of inference for parametric measurement error models to adjust for the random effects of measurement error bias (Cook and Stefanski 1994, and Stefanski and Cook 1995). The authors show that the magnitude of this random effect bias in the estimates depends on the size of the measurement error variance. The simulation step of SIMEX is used to examine the effect of the random component. The method requires that the measurement error variance is known or at least well estimated. Since in the stake data set the random measurement error effects are not large enough, we do not use the extrapolation step of the SIMEX algorithm to correct for this effect. If the effects of random measurement errors were large, we suggest using the full SIMEX algorithm to correct for bias due to the random component.

Cook and Stefanski (1994) developed a simulation-based estimation procedure for measurement error models known as SIMEX (Simulation-Extrapolation). The SIMEX method has been shown to produce less biased parameter estimates. The main idea using SIMEX is to experimentally determine the random effect of measurement errors on an estimator via simulation. If an estimator is influenced by random measurement errors, simulation experiments can be considered in which the level of the measurement error, (i.e. its variance), is intentionally varied (Carroll et al. 1995, p.80). SIMEX estimates are obtained by adding additional measurement error-induced bias versus the variance of the added measurement error, and extrapolating this trend back to the case of no measurement error. In synthesis, the SIMEX algorithm consists of four main steps: (*i*) Additional measurement error is added in known increments to the observed data. (*ii*) For each increment of added measurement error, parameter estimates are computed from the further-contaminated data. (*iii*) A trend is established between the parameter estimate and the added measurement error. (*iv*) The trend is extrapolated back to the case where there is no measurement error. Details on the theoretical support of the SIMEX estimate can be found in Stefanski and Cook (1995). SIMEX is best suited to problems with additive measurement error models, but additivity of errors is not crucial and the method can be extended to more general models (Carroll et al. 1995). Although the focus in this paper has been an additive measurement error model, additional analyses have been done assuming a multiplicative error model. For reasons of space, these analyses are not presented here.

In this section, only the simulation step of the SIMEX method is illustrated using the stake data set. A computer program was written in MAPLE (Char et al. 1991) to carry out the simulations, and to obtain maximum likelihood estimates for the scale parameter of the half-normal detection probability density function, $f(\cdot|\theta)$. In the simulation step we create additional data sets of increasingly larger measurement error $(1+\lambda)\sigma_u^2$, $\lambda = \{0.0, 0.25, 0.50, ..., 2.0\}$. The measurement error variance estimate obtained in the previous section, given by (5) was used as an estimate of σ_u^2. For any $\lambda \geq 0$, we define

$$W_{b,i} = W_i + \sqrt{\lambda} U_{b,i}, i = 1, ..., n, b = 1, ..., B, \tag{7}$$

where the computer-generated pseudo-errors, $\{U_{b,i}\}_{i=1}^{n}$, are mutually independent, independent of all the observed data, and identically distributed normal random variables with mean zero and variance σ_u^2. Once the new predictors are created, the average of the estimates obtained from a large number of experiments, ($B = 500$ in our study), with the same amount of measurement error are computed and plotted versus the amount of error added (λ). The estimated value at $\lambda = 0$ (i.e. no measurement error added to the contaminated data) is known as the the NAIVE estimator. The extrapolation step consists of modeling these averages as a function of λ, for $\lambda > 0$, and extrapolating the fitted models back to $\lambda = -1$. The extrapolated value at $\lambda = -1$ corresponds to the SIMEX estimator.

A simulation study was carried out in which known amount of measurement error can be added to error-free line transect data (i.e. true perpendicular distances) to determine the effects of random measurement error in the parameter estimates. If we were to do the extrapolation step, it would not make sense to extrapolate for the true distance data because the SIMEX and the NAIVE estimators are equivalent since the data are known to be error-free. The SIMEX simulation step was used to compare population estimates obtained from fitting the true, observed and calibrated perpendicular distances while adding known amounts of random measurement errors.

When fitting the calibrated perpendicular distances, we account for heteroscedasticity of the measurement error variance by modeling it as a function of distance as described in the previous section (see equation 6). Pseudo-data sets are generated using an estimate of the measurement error variance given by $\hat{\sigma}_u^2 = \dfrac{\hat{\gamma}_0 + \hat{\gamma}_x X_i}{\hat{\beta}_x}$.

Results from these simulations are summarized in Figure 4. The solid line corresponds to fits of a quadratic extrapolant function. Clearly, adding more random measurement error to distance data causes underestimation of the population size estimates. The same declining trend is observed for analyses of the observed and true perpendicular distance data. Note however, that the random effect of measurement errors in the population estimates is minimal since the lines are almost flat curves (Figure 4a). Also note that the calibrated data produce estimates that are closer to those estimates obtained from the true distances even after adding large amounts of random measurement error (Figure 4b). This result suggests that a combination of calibration and SIMEX analysis removes of most of the measurement error.

[Figures 4a,b near here]

8 Discussion and future research

In this paper, the stake data set was used to obtain an estimate of the measurement error variance. The stake data set also allows characterization of the error distribution. Several advantages of using the stake data set are that the true distances at which the stakes were placed are known, and there were eleven observers walking along the line which allows reasonable estimation of the measurement error variance. The true population size is also known. Two main sources of measurement error biases were identified and quantified for this study, systematic and random measurement errors. Calibration was used to correct for systematic measurement error bias, and the simulation step of the SIMEX algorithm to examine the effects of random measurement error bias. For simplicity, a univariate model for the detection function (half-normal) is considered to illustrate the estimation methods. In synthesis, this paper suggests a general but simple methodology to assess the effect of measurement errors in line transect surveys, and to correct for measurement error bias.

Future research could focus on examining mixed effects measurement error models to incorporate the variance component of the observer's effect. The effect of measurement errors in model selection using Akaike information criteria and maximum likelihood ratio tests also needs to be assessed. The robustness of estimation methods (i.e. maximum likelihood estimation, nonlinear least squares, kernel density estimation, etc.) to measurement errors can be evaluated. Violation of the assumption of constant variance and the use of an alternative multiplicative model need to be investigated since the measurement error variance may be a nonlinear function of the distances (i.e. quadratic or exponential function). Some analyses in this direction have been done, but because strong evidence of heteroscedasticity was not found, and for reasons of space they are not included here.

In this particular study, the effects of measurement error bias are overestimation due to systematic error, and very little underestimation of the population size due to the random component. The bias effect due to measurement error variance is small compared to the amount of systematic bias introduced by underestimating the distances. However, we would like to emphasize that in other situations systematic errors could cause either underestimation or overestimation of population size depending on the direction of the systematic bias when recording the perpendicular distances, and the assumed measurement error model. The effect of random measurement error would be more severe for a mobile wildlife population. For the stake data set, measurement errors are likely smaller than in other studies because it is a controlled experiment. In most wildlife populations the conditions for an study will be obviously different. This analysis suggests a lower bound on the effects of measurement errors in line transects studies. Our simulation study clearly demonstrates the potential measurement error effects in population parameters when using line transect surveys. It also emphasizes the importance of bias correction when evidence of measurement errors exists.

References

Akaike, H. (1973) Information theory and an extension of the maximum principle. In *International Symposium on Information Theory*, 2nd edn (eds B. N. Petran and F. Csàaki), Akadèemiai Kiadi, Budapest, Hungary, pp. 267-81.

Borchers, D. L. and M. D. Haw. (1990). Determination of minke whale response to a transiting survey vessel from visual tracking of sightings. *Report of the International Whaling Commission*, **40**, 257-69.

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.

Buckland, S. T. and A. A. Anganuzzi. (1988). Comparison of smearing methods in the analysis of Minke sighting data from IWC/IDCR Antarctic cruises. *Report of the International Whaling Commission*, **38**, 257-63.

Burnham, K.P., D.R. Anderson, and J.L. Laake. (1980). Estimation of density from line transect sampling of biological populations. *Wildlife Monographs*, **72**, 202 pp.

Carroll R.J. and L.A. Stefanski. (1990). Approximate quasi-likelihood estimation in models with surrogate predictors. *Journal of the American Statistical Association.*, Theory and Methods. **85(411)**, 652-663.

Carroll R.J., D. Ruppert, and L.A. Stefanski. (1995). Measurement errors in nonlinear models. Monographs on statistics and applied probability 63. Chapman and Hall, London. 305pp.

Char, B.W *et al.* (1991) Maple V Library reference manual. Waterloo Maple Software. Springer-Verlag, New York.

Chen, S. X. (1997). Measurement errors in line transect surveys. Technical paper submitted for publication to Biometrics.

Cook, J.R. and L.A. Stefanski. (1994). Simulation-extrapolation estimation in parametric measurement error models. *Journal of the American Statistical Association.*, Theory and Methods. **89(428)**, 1314-1328.

Draper, N. R. and H. Smith (1981). Applied regression analysis. John Wiley and Sons, New York, New York, U.S.A.

Laake, J. L. (1978) Line transect estimators robust to animal movement, MS Thesis, Utah State University, Logan, UT, USA, 55pp.

Laake, J. L., Buckland, S. T., Anderson, D.R. and Burnham, K. P. (1993) DISTANCE User's Guide. Colorado Cooperative Fish and Wildlife Research Unit, Colorado State University, Fort Collins, CO 80523, USA.

MathSoft, Inc. (1995). S-PLUS Guide to Statistical and Mathematical Analysis. Version 3.3 for Windows and Unix. Seattle, WA. U.S.A.

Øien, N. and T. Schweder. (1992). Estimates of bias and variability in visual distance measurements made by observers during shipboard surveys of northeastern atlantic minke whales. *Report of the International Whaling Commission*, **42**, 407-412.

Robinette, W. L., Loveless, C. M. and Jones, D. A. (1974). Field tests of strip census methods. *Journal of Wildlife Management*, 38, 81-96.

SAS Institute Inc., (1990). SAS/STAT User's Guide, Version 6, Fourth Edition, Volume 2. Cary, North Carolina, 846 p.

Schweder, T. (1977). Point process models for line transect experiments, in *Recent Developments in Statistics* (eds). J.R. Barba, F. Brodeau, G. Romier and B. Van Cutsem), North-Holland Publishing Company, New York, USA, pp. 221-42.

Schweder, T., G. Høst, and N. Øien. (1991). Estimates of the detection probability for shipboard surveys of Northeastern Atlantic Minke whales, based on a parallel ship experiment. *Report of the International Whaling Commission*, **41**, 417-432.

Stefanski, L. A. (1985). The effects of measurement error on parameter estimation. *Biometrika* 72(3):583-92.

Stefanski, L. A. and Cook, J. R. (1995). Simulation-Extrapolation: The measurement error jackknife. *Journal of the American Statistical Association*, **90**, 1247-1256.

Table 1: Distribution of the errors by distance categories

Errors	0 – 5 m	5 – 10m	10 – 15 m	15 – 20m	Total
< –1.00	1	0	0	0	1
–1.00 – –0.75	1	0	1	2	4
–0.75 – –0.50	6	2	7	3	18
–0.50 – –0.25	25	18	6	7	56
–0.25 – 0.00	125	69	30	23	247
0.00 – 0.25	114	69	35	22	240
0.25 – 0.50	19	13	15	14	61
0.50 – 0.75	7	2	1	1	11
0.75 – 1.00	2	0	1	0	3
> 1.00	1	0	0	0	1
Total	301	173	96	72	642

TABLE 2: ANOVA tables for models described in section 6

MODEL1 AVERAGE = Bo+Bx*TRUE+Ui.
Dependent Variable: AVERAGE

Source	DF	SS	MS	F-value	Prob>F
Model	1	17182.9	17182.9	60805.8	0.0001
Error	99	28.0	0.28259		
C-Total	100	17210.8			
Root	MSE	0.53159		R-square	0.9984
Dep	Mean	6.49621		Adj-Rsq	0.9984
C.V.	8.18304				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for Ho: Param=0	Prob>|T|
NTERCEP	1	-0.0543	0.0339	-1.604	0.1120
TRUE	1	0.9892	0.0040	246.588	0.0001

Test for Ho: Slope= 1

			df		
SLOPE1	Numerator:	2.0586	1	F-value:	7.2850
	Denominator:	0.282586	99	Prob>F:	0.0082

MODEL2 VARS = Go+Gx*TRUE+Ei.
Dependent Variable: VARIANCE

Source	DF	SS	MS	F-value	Prob>F
Model	1	0.1720	0.1720	6.120	0.0154
Error	85	2.3886	0.0281		
C-Total	86	2.5606			
Root	MSE	0.16763		R-square	0.0672
Dep	Mean	0.07178		Adj-Rsq	0.0562
C.V.	233.53414				

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for Ho: Param=0	Prob>|T|
INTERCEP	1	0.0498	0.0114	4.359	0.0001
TRUE	1	0.0035	0.0014	2.474	0.0154

Test for Ho: Slope= 0

			df		
SLOPE2	Numerator:	0.1720	1	F-value:	6.1205
	Denominator:	0.028101	85	Prob>F:	0.0154

Table 3: Estimation of population size and density using the model selection procedure in program DISTANCE

	Estimate	%CV	df	95%-CI	
OBSERVER 1					
TRUE-DISTANCE					
Uniform/Cosine (1)*					
**D	0.0030	14.16	71	0.0022	0.0039
N	119	14.16	71	90	158
OBSERVED-DISTANCE					
Half-normal					
D	0.0030	15.15	71	0.0023	0.0041
N	122	15.15	71	90	164
OBSERVER 2					
TRUE-DISTANCE					
Half-normal					
D	0.0028	17.67	47	0.0020	0.0040
N	113	17.67	47	80	161
OBSERVED-DISTANCE					
Half-normal					
D	0.0029	17.60	47	0.0020	0.0041
N	116	17.60	47	81	165
OBSERVER 3					
TRUE-DISTANCE					
Half-normal					
D	0.0030	15.64	73	0.0022	0.0041
N	120	15.64	73	88	163
OBSERVED-DISTANCE					
Half-normal					
D	0.0030	15.63	73	0.0022	0.0041
N	121	15.63	73	89	165
OBSERVER 4					
TRUE-DISTANCE					
Half-normal/Cosine (1)					
D	0.0037	19.90	58	0.0025	0.0055
N	148	19.90	58	100	220
OBSERVED-DISTANCE					
Half-normal/Cosine (1)					
D	0.0039	19.47	58	0.0027	0.0058
N	157	19.47	58	106	230
OBSERVER 5					
TRUE-DISTANCE					
Half-normal/Cosine (1)					
D	0.0034	20.26	58	0.0023	0.0051
N	136	20.26	58	91	203
OBSERVED-DISTANCE					
Half-normal/Cosine (1)					
D	0.0035	20.22	58	0.0023	0.0052
N	138	20.22	58	93	206
OBSERVER 6					
TRUE-DISTANCE					
Uniform/Cosine (3)					
D	0.0034	17.06	69	0.0025	0.0048
N	138	17.06	69	98	193
OBSERVED-DISTANCE					
Uniform/Cosine (3)					
D	0.0035	14.88	69	0.0026	0.0047
N	139	14.88	69	103	186
OBSERVER 7					
TRUE-DISTANCE					
Half-normal/Cosine (1)					
D	0.0034	21.57	53	0.0022	0.0052
N	136	21.57	53	89	208
OBSERVED-DISTANCE					
Half-normal/Cosine (1)					
D	0.0035	21.21	53	0.0023	0.0053
N	138	21.21	53	91	210
OBSERVER 8					
TRUE-DISTANCE					
Half-normal/Cosine (2)					
D	0.0032	20.75	58	0.0021	0.0048
N	127	20.75	58	84	192
OBSERVED-DISTANCE					
Half-normal/Cosine (2)					
D	0.0032	20.67	58	0.0021	0.0048
N	129	20.67	58	85	194
OBSERVER 9					
TRUE-DISTANCE					
Half-normal					
D	0.0021	18.62	45	0.0015	0.0031
N	85	18.62	45	59	124
OBSERVED-DISTANCE					
Half-normal					
D	0.0021	18.56	45	0.0015	0.0031
N	85	18.56	45	59	124
OBSERVER 10					
TRUE-DISTANCE					
Half-normal/Cosine (1)					
D	0.0031	21.85	39	0.0020	0.0048
N	125	21.85	39	81	194
OBSERVED-DISTANCE					
Half-normal/Cosine (1)					
D	0.0032	21.85	39	0.0021	0.0050
N	129	21.85	39	83	199
OBSERVER 11					
TRUE-DISTANCE					
Half-normal					
D	0.0030	16.51	53	0.0022	0.0042
N	120	16.51	53	86	167
OBSERVED-DISTANCE					
Half-normal/Cosine (1)					
D	0.0037	19.95	52	0.0025	0.0055
N	149	19.95	52	100	222

* Number of adjustment terms ** Density per square meter

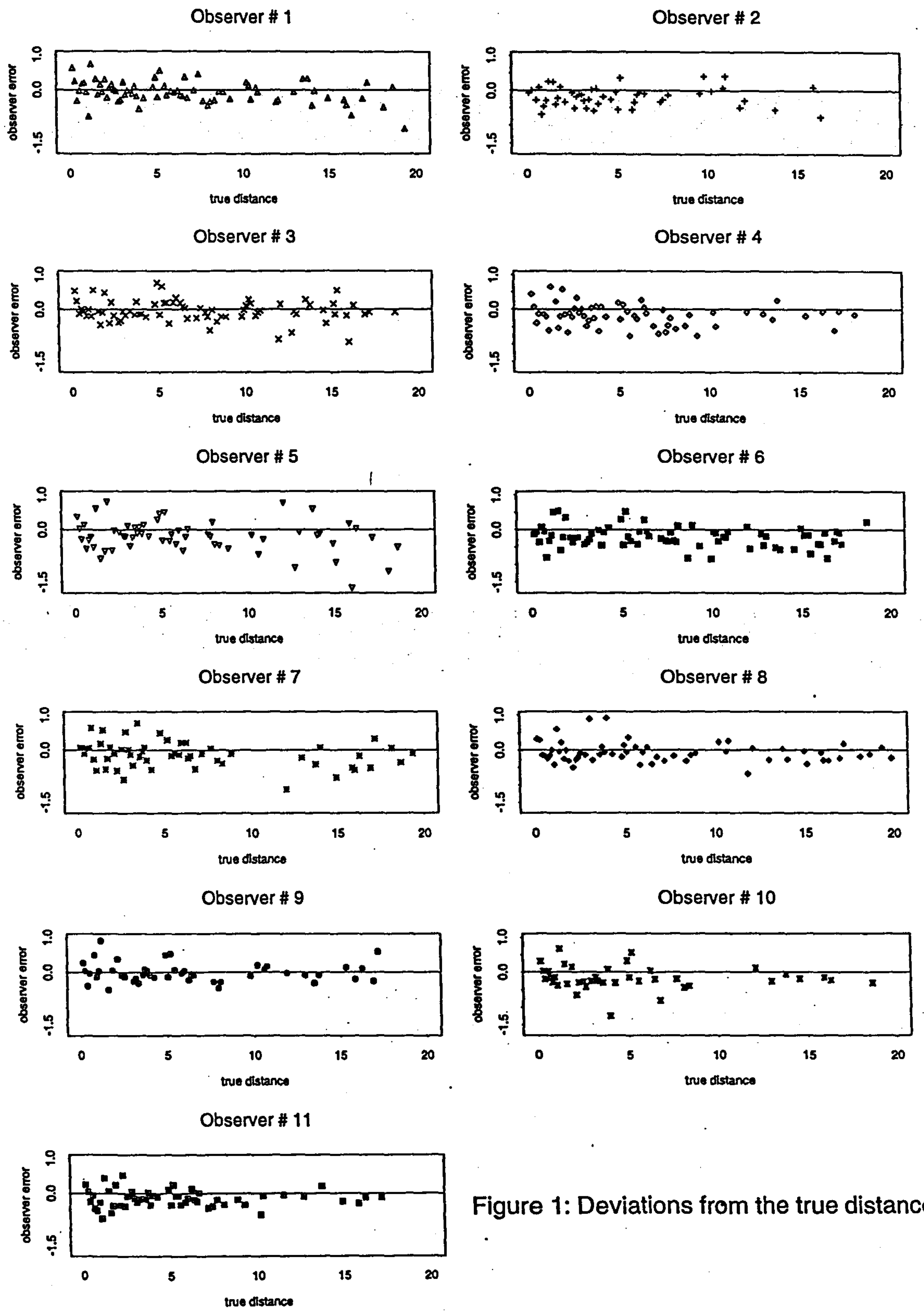

Figure 1: Deviations from the true distance

Figure 2a. Exploratory data analysis for measurement errors based on true distance.

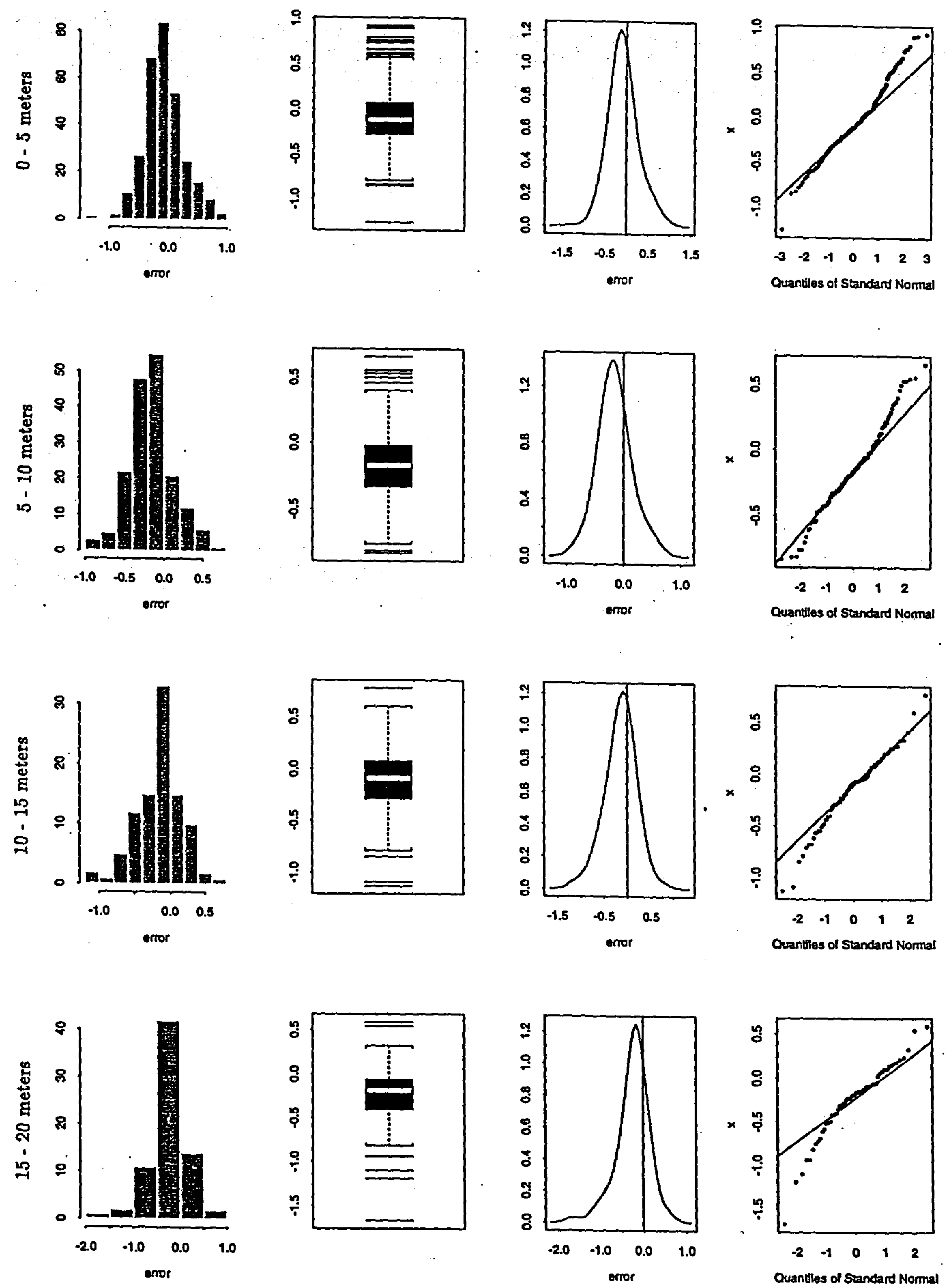

Figure 2b. Exploratory data analysis for measurement errors based on observed average of the distances.

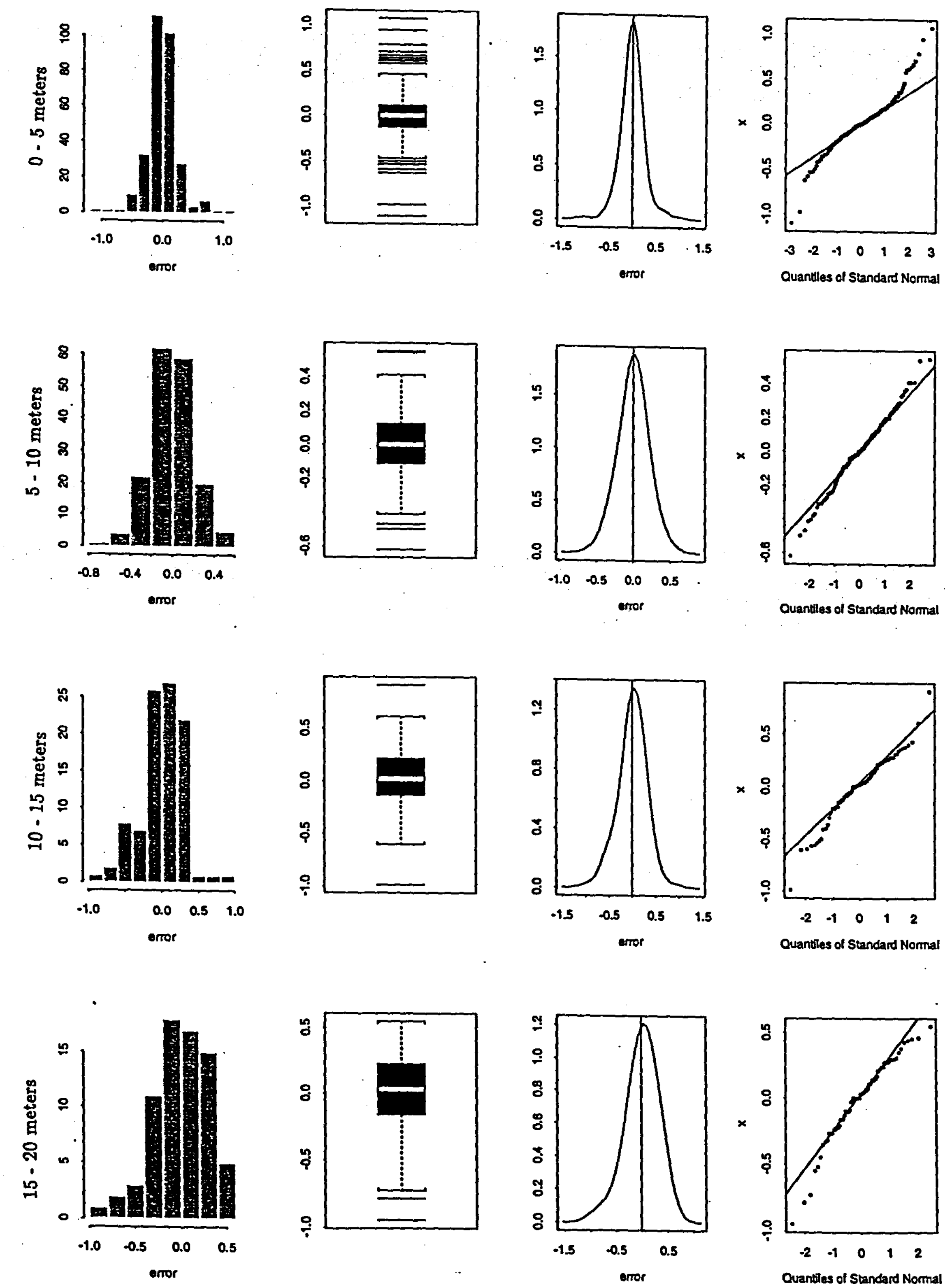

Figure 3a: Observed average vs. true distance and fitted line

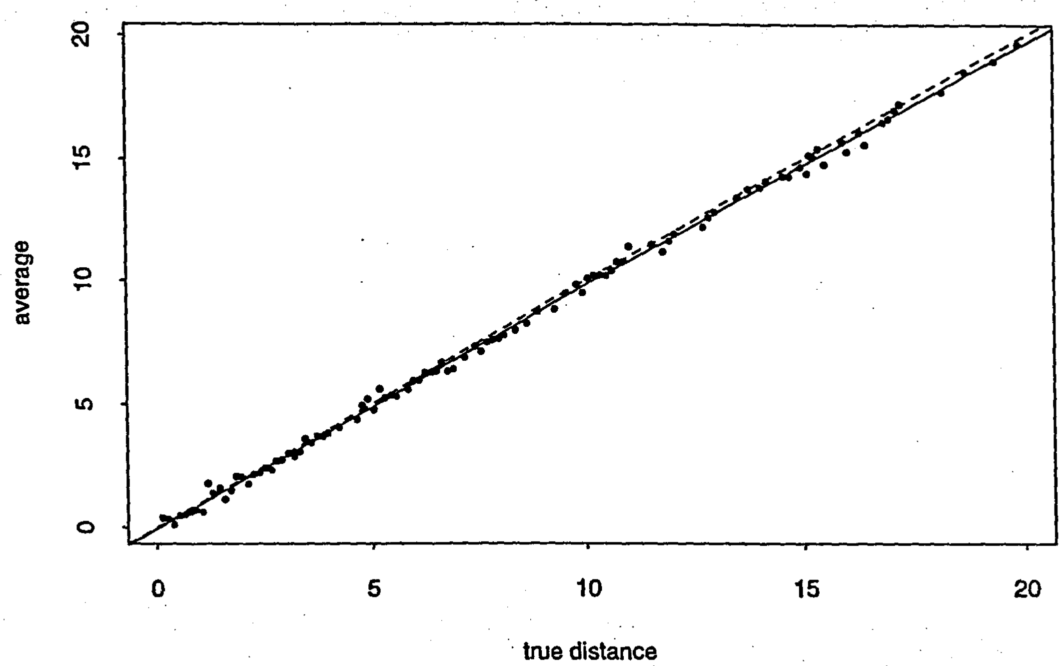

Figure 3b: Observed variance vs. true distance and fitted line

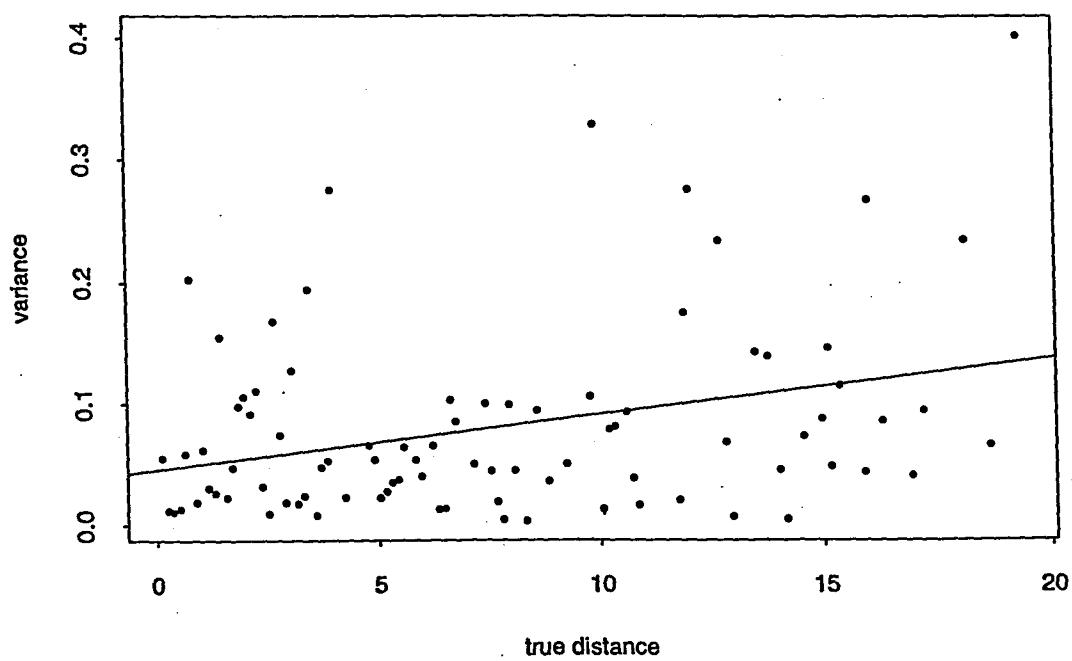

Figure 4a: Effect of adding random measurement error to distance data

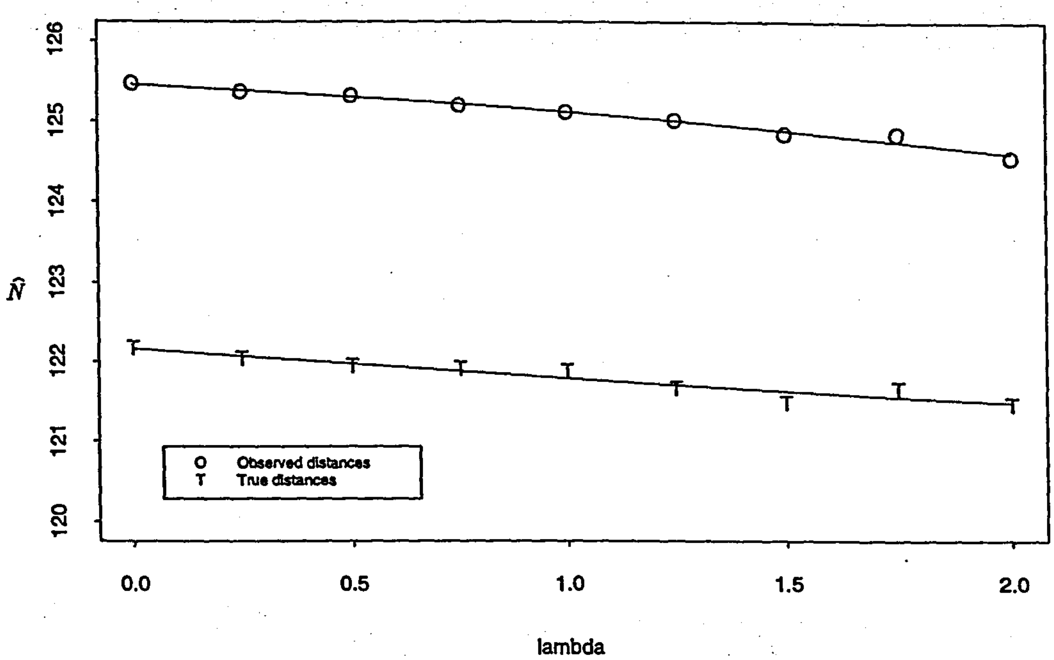

Figure 4b: Using calibration to correct for systematic effects