Measurement Errors in Line Transect Surveys
Author(s): Song Xi Chen
Source: *Biometrics*, Vol. 54, No. 3, (Sep., 1998), pp. 899-908
Published by: International Biometric Society
Stable URL: http://www.jstor.org/stable/2533844
Accessed: 22/04/2008 23:23

BIOMETRICS 54, 899–908
September 1998

Measurement Errors in Line Transect Surveys

Song Xi Chen

School of Statistical Science, La Trobe University,
Victoria 3083, Australia

SUMMARY

An important assumption used in the analysis of line transect data for animal abundance estimation is that the perpendicular sighting distances of detected animals to transect lines are measured accurately. However, in many line transect surveys, the sighting distances are likely to be subject to measurement errors. In this paper, we investigate the effect of measurement errors on animal abundance estimation and show that measurement errors cause systematic bias, which cannot be reduced by simply increasing the survey effort. Estimators that correct for measurement errors using the method of moments by assuming an exponential power series detection function and knowledge of only a few moments of the error distribution are proposed.

1. Introduction

Line transect sampling as a method of estimating wildlife abundance has been used for many decades. It has been a practical and inexpensive approach for wildlife management. The population density D of a biological population is defined as $D = N/A$, where N is the unknown population size and A is the area occupied by the population. To estimate D, an observer traverses a distance L along randomly generated nonoverlapping transect lines within the area. Each object sighted from the transect lines is counted, and its perpendicular distance to the transect lines is measured.

An important assumption in line transect estimation is that the perpendicular sighting distances are measured accurately (Buckland et al., 1993, p. 18). However, the sighting distances are likely to be subject to measurement errors. In many line transect surveys, the perpendicular distances are derived by measuring both the sighting angles and the radial distances from the observer to the sighted animals by using protractors and range finders, respectively. It is well known that the sighting angles are subject to rounding errors, as they are often rounded to the nearest 5 or 10^o. The radial distances are also subject to random errors. In some aerial surveys, protractors and range finders are operated onboard a small airplane. Measurement errors are increased by air turbulence. For surveying animals such as birds and marine mammals, measurement errors arise as the position where the animal is first detected is very likely to be different from the position where the measurement is taken.

Measurement errors will be present in line transect surveys even though the latest modern measuring techniques have been employed. In an aerial line transect survey for the commercially valuable Southern bluefin tuna, as reported in Chen (1996), a satellite based Global Positioning System (GPS) is used to obtain the perpendicular sighting distances. The perpendicular sighting distances are derived from the GPS recordings in latitude and longitude of positions where the tuna schools are first detected, where the airplane breaks the transects to observe closely the tuna schools and the exact positions of the tuna schools. The use of GPS has largely improved the accuracy of the distance measurements compared to those obtained by using protractors and range finders. However, there are still two sources of measurement errors. One is that the satellite signals used by the GPS have been deliberately corrupted due to security concerns. The other is that the exact positions of the tuna schools may not be taken accurately by the pilot. Although measurement

---

Email address: Song.Chen@latrobe.edu.au

Key words: Abundance estimation; Double exponential power series; Method of moments.

errors are quite wide spread in line transect surveys, there has been little research on this problem. Gates et al. (1985) considered data heaping. Schweder (1996) and Schweder et al. (1996) studied the size of measurement errors and proposed some correction methods.

In this paper, we investigate the effect of measurement errors and show that measurement errors cause systematic bias in the estimation of animal abundance and that this bias cannot be reduced by increasing the survey effort. Estimators that correct for measurement errors using the method of moments by assuming an exponential power series detection function and knowledge of only a few moments of the error distribution are proposed.

The effect of measurement errors on the estimation of D is studied in Section 2. Method of moments estimators for removing measurement errors are proposed in Section 3. Section 4 presents a numerical example. Some simulation results are reported in Section 5.

2. Effect of Measurement Errors

Suppose n objects are detected independently from the transect lines. Let $X_1, X_2, \ldots, X_n$ be the exact perpendicular sighting distances, which are independent and identically distributed with a probability density function f_x. Due to measurement errors, what we actually measure are

$$Y_i = X_i + Z_i \qquad i = 1, \ldots, n,$$

where Z_i are measurement errors. It is reasonable to assume that $Z_1, \ldots, Z_n$ are independent and identically distributed with a symmetric probability density function q and that measurement errors Z_i are independent of the exact sighting distances X_i. As pointed out by a referee, measurement errors might be subject to bias in some situations. If the structure of the bias is known, the bias can be removed and only random errors remain. The referee also pointed out that, in situations where perpendicular sighting distances are derived from radial distance and sighting angles, the assumption that X_i and Z_i are independent may require strict conditions regarding sighting angles, sighting distance, and measurement errors. However, the primary aim of line transect estimation is to estimate $f_x(0)$, which basically depends largely on small sighting distances where the independence between X_i and Z_i is reasonable.

In this paper, the sighting distances X_i and Y_i can take negative values, as detection is made on either side of the transect lines. These signed sighting distances are very natural in line transect surveys. A negative (positive) distance indicates that the animal was detected on the left (right) of the transect line. Traditionally, the absolute perpendicular sighting distances have been used in the analysis of line transect surveys, with no record of the side of the transect on which a detection is made. When the distances can be measured accurately, using absolute distances is convenient for analysis. However, in the presence of measurement errors, it is awkward to use the absolute distances, as $|Y_i| \neq |X_i| + |Z_i|$ and the mathematics involved is very complicated. The advantage of using the signed distances in the presence of measurement errors is that the probability density f_y of Y_i is a convolution of f_x and q.

Let $g(x)$ be the conditional probability of detecting an object given the object is at a perpendicular distance x from the transect line; g is commonly called the detection function. Assuming $g(0) = 1$ and the objects are not responsive to the observers, the abundance density

$$D = \frac{\mathrm{E}(n) f_x(0)}{L}, \tag{1}$$

which is just a simple modification of that for the absolute distances. A general estimator for D is

$$\hat{D} = \frac{n \hat{f}_x(0)}{L}, \tag{2}$$

where $\hat{f}_x(0)$ is an estimator for $f_x(0)$.

In the presence of measurement errors, $\hat{f}_x(0)$ cannot be estimated directly from the observed distances $Y_1, \ldots, Y_n$. If measurement errors are ignored, the density D is instead estimated by

$$\hat{D}^\star = \frac{n \hat{f}_y(0)}{L}.$$

In the rest of this section, we show that there is a systematic bias associated with $\hat{D}^\star$ as an estimator for D and that the bias cannot be reduced by increasing the sample size as a result of increased survey effort L.

As almost any reasonable estimator $\hat{f}_y(0)$, parametric or nonparametric, for $f_y(0)$ based on the observed distances $Y_1, \ldots, Y_n$, is at least asymptotically unbiased, $\hat{D}^\star$ is at least asymptotically unbiased for

$$D^\star = \frac{\mathrm{E}(n) f_y(0)}{L}.$$

In other words,

$$\mathrm{E}(\hat{D}^\star) = D^\star + o(1). \tag{3}$$

Here we denote by $o(1)$ a term that converges to zero when L, the total length of the transect line, approaches infinity.

However, what we want to estimate is D. To see the effect of measurement errors on line transect estimation, let us assume $g(0) = 1$ and $P(g(X) < 1) > 0$, where X is a random variable representing the sighting distance. These two assumptions mean that $f_x(0) = \max f_x(x)$ as $f_x(x) = g(x)/\int g(u)du$. Since f_y is a convolution of f_x and q,

$$f_y(t) = \int_{-\infty}^{\infty} f_x(u) q(t-u) du.$$

As the error distribution is symmetric about zero,

$$\begin{aligned} f_y(0) &= \int_{-\infty}^{\infty} f_x(u) q(-u) du = f_x(0) \int_{-\infty}^{\infty} \{f_x(0)\}^{-1} f_x(u) q(u) du \\ &< f_x(0) \int_{-\infty}^{\infty} q(u) du = f_x(0). \end{aligned} \tag{4}$$

So measurement errors lower the probability density value at zero and, as a consequence,

$$D^\star < D.$$

This and (3) mean that $\hat{D}^\star$ underestimates D.

To give a more detailed quantification for the bias of $\hat{D}^\star$ as an estimator for D, let us assume that f_x belongs to the following double exponential power series family:

$$f_x(t) = \frac{1}{2\lambda\Gamma(1+1/p)} \exp\left\{-\left(\frac{|t|}{\lambda}\right)^p\right\} \qquad t \in (-\infty, \infty), \lambda > 0, \text{ and } p > 0, \tag{5}$$

where λ and p are the scale and the shape parameters, respectively. The density given in (5) is a symmetric reflection of the exponential power series family for absolute sighting distances proposed by Pollock (1978). Then

$$f_y(t) = \frac{1}{2\lambda\Gamma(1+1/p)} \int_{-\infty}^{\infty} \exp\left\{-\left(\frac{|u|}{\lambda}\right)^p\right\} q(t-u) du \tag{6}$$

and

$$f_y(0) = \frac{1}{\lambda\Gamma(1+1/p)} \int_0^{\infty} \exp\left\{-\left(\frac{u}{\lambda}\right)^p\right\} q(u) du = C f_x(0), \tag{7}$$

where

$$C = 2\int_0^{\infty} \exp\left\{-\left(\frac{u}{\lambda}\right)^p\right\} q(u) du,$$

which is always less than one for any p and λ.

The bias of $\hat{D}^\star$ as an estimator for D is

$$\begin{aligned} B = \mathrm{E}(\hat{D}^\star) - D &= \frac{\mathrm{E}(n)\{f_y(0) - f_x(0)\}}{L} + o(1) = \frac{\mathrm{E}(n) f_x(0)(C-1)}{L} + o(1) \\ &= D(C-1) + o(1) \qquad \text{as } L \to \infty. \end{aligned}$$

Clearly, the bias cannot be reduced by increasing the survey effort L, as C does not depend on the sample size and is always less than one. This means that measurement errors cause underestimation of the abundance. The exact amount of bias depends on the value of C, which depends on the parameters p and λ, and the error distribution.

As an example, we consider the case where

$$q(u) = \frac{p}{2\sigma_z^\alpha \Gamma(\alpha/p)} |u|^{\alpha-1} \exp\left(-|u/\sigma_z|^p\right). \tag{8}$$

So the error distribution is a symmetric generalized gamma distribution with a scale parameter σ_z and has the same shape parameter p as f_x given in (5). Simple algebra shows that

$$C = 2\int_0^\infty \exp\left\{-\left(\frac{u}{\lambda}\right)^p\right\} q(u)du = \left\{1 + (\sigma_z/\lambda)^p\right\}^{-\alpha/p}. \tag{9}$$

Obviously, $C < 1$. When $\alpha = 1$ and $p = 2$, which corresponds to a normal detection function and $\mathrm{N}(0, \sigma_z^2)$ measurement errors, then

$$C = \sqrt{\frac{\sigma_x^2}{\sigma_x^2 + \sigma_z^2}},$$

where $\sigma_x^2 = \lambda^2/2$. This is just the reliability ratio in measurement error models considered in Fuller (1987) and Carrol, Ruppert, and Stefanski (1995). So the larger the variance of measurement errors, the further away the C value is from 1 and the larger (in absolute value) is the systematic bias.

3. Correcting for Measurement Errors

To correct for measurement errors, we have to assume a parametric form for f_x. This is because there is no effective nonparametric method at the moment for density estimation in the presence of measurement errors. Using the deconvolution method to develop estimators for f_x has been studied by Carroll and Hall (1988), Stefanski and Carroll (1990), and others. Even if the method assumes a nonparametric form for f_x, the error distribution has to be completely known. When the error is normally distributed and f_x has k bounded derivatives, the fastest rate of convergence of any density estimators, including both the kernel and the Fourier series estimators, is only $\log(n)^{-k/2}$. The slow convergence rate suggests that, to use the deconvolution estimators, the sample size has to be very large. This can be difficult to achieve for a line transect survey.

In this paper, we take a semi-parametric approach to correct measurement errors. We assume that f_x is in the double exponential power series family as given in (5). We also assume knowledge of a few moments of the measurement error distribution rather than complete knowledge of this distribution.

The maximum likelihood approach is difficult to apply for parameter estimation in the current situation. There are two reasons. First, the error distribution may not be known. Second, even if we know the error distribution, the density function f_y is still difficult to obtain except in some isolated cases. To appreciate the second reason, let us assume normal measurement errors with variance σ_z^2. Substitute $q(t) = \phi_{\sigma_z^2}(t)$, the p.d.f. for $\mathrm{N}(0, \sigma_z^2)$, into (6) and we have

$$f_y(t) = \frac{1}{\lambda\Gamma(1+1/p)\sqrt{2\pi}\sigma_z}\int_0^\infty \exp\left\{-\left(\frac{u}{\lambda}\right)^p\right\}\left[\exp\left\{-\frac{(t-u)^2}{2\sigma_z^2}\right\} + \exp\left\{-\frac{(t+u)^2}{2\sigma_z^2}\right\}\right] du.$$

This density is not tractable except when $p = 2$.

To avoid the difficulties of maximum likelihood, we propose using Karl Pearson's method of moments for parameter estimation. The method of moments requires only knowledge of $\sigma_z^2 = \mathrm{E}(Z_i^2)$ and $m_{z4} = \mathrm{E}(Z_i^4)$, which is much less than a complete knowledge of the error distribution required by both the maximum likelihood and the deconvolution methods. However, knowledge of these moments may still be regarded as demanding. Generally speaking, to correct for measurement errors, we need to have some knowledge about them. Sometimes σ_z^2 may be obtained from the resolution specification given by the manufacturer of the measuring devices. Otherwise, calibrations for the measurement have to be carried out. This is the price paid for not being able to measure the distances accurately.

In this section, estimators derived using the method of moments are proposed for two situations: (1) when the shape parameter p is unknown and (2) when p is known.

3.1 *p is Unknown*

If the shape parameter p in the double exponential power series model (5) is unknown, the method of moments requires knowledge of both σ_z^2 and m_{z4}. As X_i and Z_i are symmetrically distributed, the first and third moments of X_i and Z_i are zero. From Johnson, Kotz, and Balakrishnan (1995, p. 388),

$$\mathrm{E}(X_i^2) = \lambda^2\Gamma(3/p)/\Gamma(1/p) \quad \text{and} \quad \mathrm{E}(X_i^4) = \lambda^4\Gamma(5/p)/\Gamma(1/p).$$

As X_i and Z_i are independent,

$$\mathrm{E}(Y_i^2) = \mathrm{E}(X_i^2) + \mathrm{E}(Z_i^2) = \lambda^2\Gamma(3/p)/\Gamma(1/p) + \sigma_z^2$$

and

$$\begin{aligned}\mathrm{E}(Y_i^4) &= \mathrm{E}(X_i^4) + 6\mathrm{E}(X_i^2)\mathrm{E}(Z_i^2) + \mathrm{E}(Z_i^4)\\ &= \lambda^4\Gamma(5/p)/\Gamma(1/p) + 6\sigma_z^2\lambda^2\Gamma(3/p)/\Gamma(1/p) + m_{z4}.\end{aligned}$$

Replacing the population moments by their sample counterparts, we have the following two estimating equations:

$$\hat{m}_2 = \lambda^2\Gamma(3/p)/\Gamma(1/p) + \sigma_z^2 \tag{10}$$

and

$$\hat{m}_4 = \lambda^4\Gamma(5/p)/\Gamma(1/p) + 6\sigma_z^2\lambda^2\Gamma(3/p)/\Gamma(1/p) + m_{z4}, \tag{11}$$

where $\hat{m}_2 = n^{-1}\,\Sigma\, Y_i^2$ and $\hat{m}_4 = n^{-1}\,\Sigma\, Y_i^4$. After some algebra, an estimating function for p is

$$T(p) = Q(\hat{m}_2, \hat{m}_4), \tag{12}$$

where

$$Q(\hat{m}_2, \hat{m}_4) = \frac{\hat{m}_4 - m_{z4} - 6(\hat{m}_2 - \sigma_z^2)\sigma_z^2}{(\hat{m}_2 - \sigma_z^2)^2}$$

and

$$T(p) = \frac{\Gamma(5/p)\Gamma(1/p)}{\Gamma^2(3/p)}.$$

LEMMA. *$T(p)$ is a monotonic decreasing function of p and $\lim_{p\to\infty} T(p) = 1.8$.*

We defer the proof of the lemma to the Appendix. As $Q(\hat{m}_2, \hat{m}_4)$ is determined by the values of $\hat{m}_2$, $\hat{m}_4$, and σ_z^2, an estimate for p, say $\hat{p}$, can be found by solving equation (12). Then substituting $\hat{p}$ in (10), an estimate for λ is

$$\hat{\lambda} = \sqrt{(\hat{m}_2 - \sigma_z^2)\Gamma(1/\hat{p})/\Gamma(3/\hat{p})}. \tag{13}$$

It may be shown that, with probability one, $Q(\hat{m}_2, \hat{m}_4)$ should be larger than the lower bound 1.8 when $L \to \infty$, which ensures a unique finite $\hat{p}$ value. However, sometimes $Q(\hat{m}_2, \hat{m}_4) \leq 1.8$ for finite samples. This happens when data exhibit a large shoulder near $x = 0$. In this case, the estimate for p is infinity, which implies that X has a uniform distribution on $[-\lambda, \lambda]$, and the estimator for λ is

$$\hat{\lambda} = \sqrt{3(\hat{m}_2 - \sigma_z^2)}.$$

The proposed estimators for $f_x(0)$ and D that correct for measurement errors are

$$\hat{f}_{xm_1}(0) = \frac{1}{2\hat{\lambda}\Gamma(1 + 1/\hat{p})} \quad \text{and} \quad \hat{D}_{m_1} = \frac{n\hat{f}_{xm_1}(0)}{L}. \tag{14}$$

Assume that $n/N \to P_0$ in probability when $L \to \infty$, where P_0 is the probability of sighting an animal from the transect. By repeatedly using the linearization method (Serfling, 1980, p. 118), it may be shown that $\hat{D}_{m_1}$ is asymptotically $\mathrm{N}(D, v_{D_1}/L)$ distribued, where v_{D_1} is a function of $\lambda, p, \sigma_z^2, m_2, m_4$, and m_{z4}.

An explicit formula for v_{D_1} is not given here because the formula is quite complicated as a result of applying the linearization method many times. Also, an estimate of the variance may be achieved by using the bootstrap. Therefore, the asymptotic normality only serves for theoretical purposes. For example, confidence intervals for D may be established by looking at the normal tables but the variance is obtained via the bootstrap.

It may be shown by the delta method that both $\hat{f}_{xm_1}(0)$ and $\hat{D}_{m1}$ are asymptotically unbiased as the survey effort L increases, that is,

$$\mathrm{E}\{\hat{f}_{xm_1}(0)\} = f_x(0) + o(1) \quad \text{and} \quad \mathrm{E}\{\hat{D}_{m_1}\} = D + o(1). \tag{15}$$

3.2 *p is Known*

The value of the shape parameter p may be known in some applications, for example, $p = 1$ or 2, which corresponds to the exponential or the normal detection functions, respectively. In this case, the method of moments requires knowledge of σ_z^2 only, and it uses the estimating equation imposed on the second moment, that is,

$$\hat{m}_2 = \lambda^2\Gamma(3/p)/\Gamma(1/p) + \sigma_z^2.$$

The estimator for λ is

$$\hat{\lambda} = \sqrt{(\hat{m}_2 - \sigma_z^2)\Gamma(1/p)/\Gamma(3/p)},$$

and corresponding estimators for $f_x(0)$ and D are

$$\hat{f}_{xm_2}(0) = \frac{1}{2\hat{\lambda}\Gamma(1+1/p)} \quad \text{and} \quad \hat{D}_{m_2} = \frac{n\hat{f}_{xm_2}(0)}{L}. \tag{16}$$

Using the linearization method again, we can show that $\hat{D}_{m_2}$ is asymptotically normally distributed with mean D and variance v_{D_2}/L, where v_{D_2} is a function of λ, p, m_2, and σ_z^2. Again we recommend using the bootstrap to estimate v_{D_2}. The asymptotic unbiasedness of $\hat{f}_{xm_2}(0)$ and $\hat{D}_{m_2}$ holds as well.

From the results presented in Sections 3.1 and 3.2, we see that the method of moments estimators for both $f_x(0)$ and D take account of the errors in measurements, as indicated by the results that they are asymptotically unbiased and normally distributed.

4. A Numerical Example

In this section, we demonstrate how to correct for measurement errors for the tuna data set used in Chen (1996). The data set contains 162 perpendicular sighting distances of detected tuna schools with a survey effort of 10,361 nautical miles.

Measurement calibration was not carried out for the aerial survey to gain information on measurement errors. This lack of information on measurement errors is common for line transect surveys. There are few data sets with information on the error distribution. This is perhaps because there has been little research on the problem and because the effect of measurement errors was largely unknown and methods for correcting the errors were not available.

To demonstrate the effect of measurement errors, three levels of σ_z ($\sigma_z = 0.5, 1.0$, and 1.5, which should be well within the real range of the parameter) were assumed. We also assumed normality for the error distribution, even though knowledge of σ_z^2 and m_{z4} was enough for the proposed estimators. The second and the fourth sample moments of the tuna data were $\hat{m}_2 = 22.37$ and $\hat{m}_4 = 1774.39$, respectively. These, together with the assumed σ_z values and the fact that $m_{z4} = 3\sigma_z^4$, provided values for $Q(\hat{m}_2, \hat{m}_4)$, which in turn gave estimates for p. The estimates for λ and D follow immediately and are displayed in Table 1. The uncorrected estimates using the method of moments were also calculated for comparison. The standard errors for abundance estimates were obtained by the bootstrap based on 200 resamples.

Table 1 shows that the error-corrected estimates for D are higher than the uncorrected estimates due to negative bias associated with the latter. The table reveals that the smaller uncorrected estimates for D are due to the larger estimates for p, as measurement errors flatten the probability density f_x. When $\sigma = 0.5$, there is little difference between the error-corrected and uncorrected estimates as the estimated C value is very close to 1. This implies that the effect of the measurement error is small at this level of measurement error. The relative difference between the uncorrected and corrected estimates is actually only 1%. When $\sigma_z = 1.0$ and 1.5, the differences between the corrected and uncorrected estimates are quite large, as the estimated C value is getting smaller.

5. Simulation results

In this section, we present some simulation results designed to evaluate the ability of the proposed method of moments estimators to eliminate the effect of measurement errors. The results are all based on 2000 simulations. In each simulation, N points were generated uniformly within a rectangular area with length L and width $2w$ to simulate the spatial distribution of an animal population. We fixed $D = 0.15$ and $w = 10$ and chose $N = 400$ and 800, respectively. The transect

Table 1

Error-corrected and uncorrected estimates for tuna data

	σ_z	$\hat{p}$	$\hat{\lambda}$	$\hat{C}$	$\hat{D}$	SE($\hat{D}$)
Uncorrected	0.0	1.605	5.817	1.00	0.001499	0.000177
Corrected	0.5	1.599	5.766	0.983	0.001512	0.000166
	1.0	1.578	5.611	0.946	0.001552	0.000203
	1.5	1.541	5.341	0.892	0.001627	0.000212

length L was determined according to $L = N/(2wD)$. The expected number of observed points is then $\mathrm{E}(n) = N/\{2wf_x(0)\}$, and n is approximately Poisson distributed. The double exponential power series detection function

$$g(x) = \exp\left\{-\left(\frac{|x|}{\lambda}\right)^p\right\}$$

was used to detect the N simulated points. Thus, f_x was within the double exponential power series family and $f_x(0) = \{2\lambda\Gamma(1 + 1/p)\}^{-1}$. The scale parameter λ was fixed at 2.0, whereas the shape parameter p had values 1.0, 2.0, and 3.0, respectively. For each successful detection, an $\mathrm{N}(0, \sigma_z^2)$ measurement error was added to the perpendicular sighting distance, with $\sigma_z^2 = 0.5$ and 1.0, respectively.

The proposed estimators for $f_x(0)$ and D were evaluated under two circumstances. One assumed that the shape parameter p was unknown. The other assumed that p was known. Estimates that ignore measurement errors, which were labelled as uncorrected estimates, were also provided for comparison. When p was known, uncorrected estimates using the method of moments were obtained by assigning a zero value to both σ_z and m_{z4} in the estimating equations (10) and (11). The maximum likelihood estimates were not provided in this case, as the calculation involved with the digamma and trigamma functions and the relevant software was not available. When p was known, both the method of moments and the maximum likelihood estimates for λ were obtained for the uncorrected estimates for $f_x(0)$ and D, respectively. The maximum likelihood estimate for λ is $\{p\,\Sigma\,Y_i^p/n\}^{1/p}$, given by Pollock (1978).

Table 2 contains both corrected and uncorrected point estimates together with their standard errors for $f_x(0)$ and D when p was unkown. Table 3 contains estimates obtained assuming p known. The values of the reliability ratio C are also given. The simulation results show that the proposed method of moments estimators for $f_x(0)$ and D performed quite well in correcting for measurement errors except when $p = 1$. The bias of the uncorrected estimates increases as the variance of the measurement error σ_z^2 increases. Reductions in the bias of the proposed error-corrected estimates were achieved by increasing the population size N; in contrast, similar reductions were not necessarily observed for the uncorrected estimates. In fact, there were increases

Table 2

Point estimates for D with unknown shape parameter p together with their standard errors and relative bias. The estimates with a subscript c correct for measurement errors, whereas those with subscripts u are ordinary method of moments estimates, which ignore measurement errors. The values of the reliability ratio C are also given. The true density D = 0.15.

	$p = 1.0$		$p = 2.0$		$p = 3.0$	
	$N = 400$	$N = 800$	$N = 400$	$N = 800$	$N = 400$	$N = 800$
			$\sigma_z = \mathbf{0.5}$			
$\hat{D}_c$	0.119	0.121	0.149	0.149	0.15	0.15
Standard error	0.03	0.02	0.03	0.02	0.03	0.02
Relative bias	−20%	−19%	−0.7%	−0.7%	0%	0%
$\hat{D}_u$	0.117	0.119	0.141	0.141	0.142	0.141
Standard error	0.03	0.02	0.03	0.02	0.03	0.02
Relative bias	−22%	−20%	−6%	−6%	−5.3%	−6%
C	0.828		0.943		0.977	
			$\sigma_z = \mathbf{1.0}$			
$\hat{D}_c$	0.119	0.121	0.152	0.150	0.159	0.155
Standard error	0.03	0.02	0.06	0.04	0.06	0.04
Relative bias	−20%	−19%	1.3%	0%	6%	3.3%
$\hat{D}_u$	0.112	0.115	0.134	0.126	0.139	0.129
Standard error	0.03	0.02	0.06	0.03	0.05	0.03
Relative bias	−25.3%	−23.3%	−10.7%	−16%	−7.3%	−14%
C	0.699		0.816		0.867	

Table 3

Point estimates for D with known shape parameter p, together with their standard errors and relative bias. The estimates with a subscript c correct for the measurement error, whereas those with subscripts uml and umm are ordinary maximum likelihood and method of moments estimates, respectively, which ignore measurement errors. The values of the reliability ratio C are also given. The true density D = 0.15.

	$p = 1.0$		$p = 2.0$		$p = 3.0$	
	$N = 400$	$N = 800$	$N = 400$	$N = 800$	$N = 400$	$N = 800$
			$\sigma_z = 0.5$			
$\hat{D}_c$	0.162	0.160	0.153	0.151	0.152	0.151
Standard error	0.020	0.02	0.02	0.02	0.02	0.02
Relative bias	8%	6.67%	2%	0.67%	1.3%	0.67%
$\hat{D}_{uml}$	0.152	0.151	0.143	0.142	0.139	0.138
Standard error	0.02	0.01	0.02	0.01	0.02	0.01
Relative bias	1.33%	0.67%	−4.7%	−5.3%	−7.3%	−8%
$\hat{D}_{umm}$	0.159	0.158	0.143	0.142	0.140	0.139
Standard error	0.02	0.02	0.02	0.01	0.02	0.01
Relative bias	6%	5.33%	−4.7%	−5.3%	−6.7%	−7.3%
C	0.828		0.943		0.977	
			$\sigma_z = 1.0$			
$\hat{D}_c$	0.162	0.161	0.154	0.152	0.155	0.152
Standard error	0.03	0.02	0.03	0.02	0.03	0.02
Relative bias	8%	7.3%	2.7%	1.33%	3.3%	1.3%
$\hat{D}_{uml}$	0.142	0.141	0.124	0.123	0.116	0.115
Standard error	0.02	0.01	0.02	0.01	0.02	0.01
Relative bias	−5.3%	−6%	−17.3%	−18%	−22.7%	−23.3%
$\hat{D}_{umm}$	0.151	0.150	0.124	0.123	0.117	0.117
Standard error	0.02	0.02	0.02	0.01	0.02	0.01
Relative bias	0.67%	0%	−17.3%	−18%	−22%	−22%
C	0.699		0.816		0.867	

in the bias for the uncorrected estimates, which happened when the simulated bias agreed with the true bias. The bias of the uncorrected estimates was large when the reliability ratio C was small and was small when C was close to 1. The knowledge of the shape parameter improved the performance of both corrected and uncorrected estimators. However, the bias associated with the uncorrected estimates was still very obvious. The proposed estimators did not work well when $p = 1$ for an exponential detection function as, in this case, the second and the fourth sample moments $\hat{m}_2$ and $\hat{m}_4$ were subject to large variations.

6. Discussion

We have discussed the effect of measurement errors in the estimation of animal abundance in line transect surveys by showing that there is a systematic bias that cannot be reduced by increasing the survey effort L. The systematic bias was demonstrated by the theoretical analysis, the numerical example, and the simulation results. Estimators that correct for the bias have been proposed based on the method of moments. The simulation results show that the proposed estimators helped to remove bias, at least in cases where the detection function had a shoulder.

The effect of measurement errors depends on the amount of measurement errors involved. When the amount is relatively large, estimators that correct for the measurement errors should be used. The proposed bias-corrected estimators require knowledge of the second and fourth moments of the error distribution, which to a large extent must be obtained by measurement calibration. Almost all of the existing line transect data sets lack information on measurement errors. This is because the effects of measurement errors were unknown and methods for correcting them were not available. We hope that this paper will lead to the collection of error information in future line transect surveys when measurement errors are substantial.

ACKNOWLEDGEMENTS

The author thanks two referees and the associate editor for helpful comments and suggestions which have improved the presentation of the paper.

RÉSUMÉ

Un préquis important de l'analyse des données en transect pour l'estimation de l'abondance animale est que la distance des animaux repères mesurée perpendiculairement au transect, soit connue avec exactitude. Cependant dans beaucoup d'études de ce type l'estimation de cette distance est entachée d'erreur. Dans ce papier nous mesurons l'effet de cette erreur sur l'estimation de l'abondance animale et démontrons l'existence d'un biais systématique qui ne peut être compense simplement par l'augmentation de l'effort d'observation. Des estimateurs corrigés sont proposés par la méthode des moments en assumant une fonction de détection en séries exponentielles et la connaissance de seulement quelques moments de la distribution de l'erreur.

REFERENCES

Buckland, S. T., Anderson, D. R., Burnham, K. P., and Laake, J. L. (1993). *Distance Sampling.* London: Chapman and Hall.

Carroll, R. J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *Journal of the American Statistical Association* **83,** 1184–1186.

Carroll, R. J., Ruppert, D., and Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models.* London: Chapman and Hall.

Chen, S. X. (1996). Kernel estimates for density of a biological population using line transect sampling. *Journal of the Royal Statistical Society, Series C* **45,** 135–150.

Fuller, W. A. (1987). *Measurement Error Models.* New York: Wiley.

Gates, C. E., Evans, W., Gober, D. R., Guthery, F. S., and Grant, W. E. (1985). Line transect estimation of animal densities from large data sets. In *Game Harvest Management,* S. L. Beasom and S. F. Roberson (eds), 37–50. Kingsville: Texas A&I University.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions,* Volume 1, 2nd edition. New York: Wiley.

Pollock, K. H. (1978). A family of density estimators for line-transect sampling. *Biometrics* **34,** 475–478.

Schweder, T. (1996). A note on a buoy-sighting experiment in the North Sea in 1990. *Report of the International Whaling Commission* **46,** 383–385.

Schweder, T., Hagen, G., Helgeland, J., and Koppervik, I. (1996). Abundance estimation of northeastern Atlantic minke whales. *Report of the International Whaling Commission* **46,** 391–405.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics.* New York: Wiley.

Stefanski, L. A. and Carroll, R. J. (1990). Deconvoluting kernel density estimators. *Statistics* **21,** 169–184.

Received September 1996; *revised April and August* 1997; *accepted October* 1997.

APPENDIX

Proof of the Lemma

Define $S(q) = \log T(1/q) = \log \Gamma(5q) + \log \Gamma(q) - 2 \log \Gamma(3q)$. Let $\psi(q)$ be the derivative of $\log \Gamma(q)$. A useful expansion for $\psi(q)$ is

$$\psi(q) = -C_{\mathrm{E}} - q^{-1} + q \sum_{k=1}^{\infty} \frac{1}{k(q+k)}, \tag{A.1}$$

where C_{E} is Euler's constant. The derivative of $S(q)$ is

$$S^{'}(q) = 5\psi(5q) + \psi(q) - 6\psi(3q).$$

From (A.1),

$$S^{'}(q) = 25q \sum_{k=1}^{\infty} \frac{1}{k(5q+k)} + q \sum_{k=1}^{\infty} \frac{1}{k(q+k)} - 18q \sum_{k=1}^{\infty} \frac{1}{k(3q+k)}$$

$$= 8q \sum_{k=1}^{\infty} \frac{k}{(5q+k)(q+k)(3q+k)} > 0,$$

which implies that $S(q)$ is monotonically increasing. So $S(1/p) = \log T(p)$ is monotonically decreasing. Therefore, $T(p)$ is monotonically decreasing, too.

As $\lim_{p\to\infty} \Gamma(1/p)/\Gamma(3/p) = 3$ and $\lim_{p\to\infty} \Gamma(5/p)/\Gamma(3/p) = 3/5$, we have

$$\lim_{p\to\infty} T(p) = 9/5 = 1.8.$$