BIOMETRICS 57, 732–742
September 2001

Measurement Errors in Line Transect Surveys Where Detectability Varies with Distance and Size

Song Xi Chen

Department of Statistics and Applied Probability,
National University of Singapore, Singapore 117543
*email: stacsx@nus.edu.sg

and

Ann Cowling

CSIRO Marine Research, G.P.O. Box 1538, Hobart, Tasmania 7001, Australia
**email: ann.cowling@marine.csiro.au

SUMMARY. When using bivariate line transect methods to estimate the biomass density of a tightly clustered biological population, it is generally assumed that both the perpendicular distance from the trackline to the cluster and the cluster size, or biomass, are measured without error. This is unlikely to be the case in practice. In this article, assuming additive mean zero errors in distance and multiplicative errors in size, we develop an estimator of density that corrects for these errors. We use the method of moments for the case of gamma cluster size, randomly placed transect lines, and the generalized exponential detection function. We derive results that show that it may not be necessary to correct for errors in distance or size when the distance and size estimates are not biased. When the size estimates are biased, the biomass density estimate has approximately the same bias as the size estimates. The work is illustrated in the context of annual aerial surveys for juvenile southern bluefin tuna in the Great Australian Bight.

KEY WORDS: Detectability; Line transect surveys; Measurement error; Method of moments; Parametric models; Size bias.

1. Introduction

Overfishing has led to a serious decline in the population of the commercially valuable southern bluefin tuna (SBT) over the last 30 years. The current parental biomass is estimated to be only 5–8% of the 1960 level (Anonymous, 1996). It is possible that the stock will not recover, even with current catch limits. Management measures include monitoring the juvenile stock using line transect surveys, which have been used for many years to estimate the abundance of animal populations (Burnham, Anderson, and Laake, 1980; Buckland et al., 1993).

During the southern summer, juvenile SBT migrate to the warm waters of the Great Australian Bight. On warm and calm days, they form schools at the surface that can be detected from the air from up to 20 nautical miles away. Commercial aerial spotting planes have operated over the Bight since the 1960s, directing vessels to the most profitable schools. Pilots and spotters have considerable skill in identifying fish species and assessing both the size of SBT schools and the size of fish in the schools.

Since 1991, Australia and Japan (through CSIRO and the National Research Institute of Far Seas Fisheries) have collaboratively conducted annual aerial surveys over the summer range of juvenile SBT. The main indices of abundance derived from the line transect surveys are the number of SBT sightings per 1000 square nautical miles (school density) and the surface biomass of SBT per 1000 square nautical miles (biomass density).

Many animal populations aggregate naturally into clusters—flocks, herds, schools. A cluster is a relatively tight gathering of animals compared with the looser association observed in a localized region of high density. Under the assumption that the probability of detecting a cluster on the transect line is one, the only covariates affecting the densities of clusters and biomass are the perpendicular distance from the transect line to the cluster and the cluster size. The probability of detecting a cluster decreases with distance from the transect line, and at any distance, the probability of detecting a cluster increases with cluster size. The detection function $g(x, s)$ gives the probability of detection at any perpendicular distance x and cluster size s. It is generally assumed that perpendicular distances and cluster sizes are measured exactly, but this may not be possible in all line transect surveys.

In the SBT surveys, the cluster, known as a sighting, is either a single school or a close grouping of schools. A plane is flown along a transect line and two experienced spotters

in the plane search for surface schools of SBT in the 180° segment to the front of the plane. The perpendicular distance from the sighting to the transect line, x, is estimated from three waypoints recorded by a Global Positioning System. The plane leaves the transect line to determine the species of fish in each school and estimate sighting size. The sighting size, s, is estimated by the average of the two spotters' estimates of the sum of the sizes of each school in the sighting.

Two planes are used during the surveys. There are differences in school-size estimates between spotters and between planes. Calibration studies conducted over three seasons using both planes have found that the estimates from one plane are systematically higher.

There are measurement errors in both x and s as a result of measurement errors in the three waypoints, rounding errors in the spotters' size estimates, which are usually given to the nearest 5 tons, and natural limitations in the spotters' ability to estimate school size precisely. It would be extremely costly to collect sufficient ground-truthing data to accurately estimate the errors in the spotters' estimates. The alternative is to evaluate the extent to which errors in x and s may affect the surface abundance indices. If they have a strong effect, it may be necessary to ground-truth the size estimates.

In this article, we assume that the survey is conducted along the line $x = 0$, so that $-\infty < x < \infty$. Clearly, $s > 0$. The usual line transect estimator of mean number of animals per unit area is

$$\hat{D}_0 = L^{-1} n \hat{f}(0),$$

where L is the total length of transect line searched, n is the total number of animals detected, and $\hat{f}(x)$ is the estimated density of the sighting distances x (cf., Burnham et al., 1980). An estimator of biomass density in clustered populations is

$$\hat{D}_1 = L^{-1} n \hat{\beta}(0), \tag{1.1}$$

where $\hat{\beta}(0) = \int s\hat{f}(0,s)ds$ and $\hat{f}(x,s)$ is the estimated density of clusters at (x,s) (Quang, 1991).

The bias in the density estimates due to measurement errors is quite distinct from the biases due to assuming that the probability of detection on the trackline is one, which was first described in Alpizar-Jara and Pollock (1996), Borchers (1996), and Manly, McDonald, and Garner (1996). This is fairly intuitive if we write the biomass density estimate as $D_1 = D_0 \mathrm{E}(S \mid x = 0)$. Errors in the assumption about the probability of detection on the trackline affect mainly D_0, errors in x affect both terms, and measurement errors in size affect only $\mathrm{E}(S \mid x = 0)$.

Although nonparametric estimators for $\beta(0)$ have been developed by Quang (1991), Chen (1996), and Mack and Quang (1998), it is difficult to correct measurement errors using a nonparametric approach (see the discussion in Chen (1998) based on the work of Carroll and Hall (1988), Stefanski and Carroll (1990), and others). The work to date on correcting measurement errors in line transect surveys has therefore used a parametric approach.

Chen (1998) and Alpizar-Jara et al. (1998) describe methods of estimating D_0 when there are additive measurement errors in x and parametric detection functions. Assuming that the errors are symmetric about zero with known second and fourth moments and that the detection function comes from the double exponential power series family, Chen develops an estimator of D_0 using the method of moments. Assuming that the errors have a nonzero mean possibly depending on x with known variance and that the detection function has a known parametric form, Alpizar-Jara et al. use calibration to correct the systematic measurement error effects and use Cook and Stefanski's (1994) SIMEX algorithm to correct the random component.

In this article, we develop an estimator of D_1 taking a parametric approach. The method of moments is less parametric than maximum likelihood since it does not require assumptions about the distribution of the errors in x and s, although it does require knowledge of the first few moments of the distributions. The method of moment estimators are also much more mathematically tractable in this problem.

In Section 2, we develop estimators that correct for measurement errors based on the method of moments. We use the parametric framework for size-biased line transect surveys developed by Drummer and McDonald (1987). In Section 3, we apply these methods to data collected in the 1993–1998 SBT surveys and compare the corrected and uncorrected biomass density estimates. In Section 4, we conduct a simulation study to assess the accuracy of the estimates. In Section 5, we derive theoretical results that explain the numerical results, and in Section 6, we discuss all the results and their implications for line transect surveys.

2. Correcting for Measurement Errors

Suppose a line transect survey of a clustered population is conducted along the line $x = 0$ and observed values of perpendicular distance and school size $\{(Y_i, T_i)\}_{i=1}^n$ are recorded, corresponding to true distances and school sizes $\{(X_i, S_i)\}_{i=1}^n$. We assume an additive model for errors in X_i and a multiplicative model for errors in S_i, i.e., that

$$Y_i = X_i + \epsilon_{xi} \quad \text{and} \quad T_i = S_i \epsilon_{si}, \tag{2.1}$$

where ϵ_{xi} and ϵ_{si} are measurement errors in X_i and S_i, respectively. We shall also assume that the errors in x have a symmetric distribution. It is reasonable to assume independence between (X_i, S_i) and ϵ_{xi}, between (X_i, S_i) and ϵ_{si}, and between ϵ_{xi} and ϵ_{si}.

Let f^* be the probability density function of (x,s) before considering detection or nondetection. As the transect lines are allocated randomly in the survey region, we put

$$f^*(x) = (2w)^{-1} I(-w \le x \le w)$$

and let $w \to \infty$. We assume that $f^*(s)$ is gamma(ν_1, ν_2). Further assuming that X and S are independent, we have

$$f^*(x,s) = (2w)^{-1} f^*(s) I(-w \le x \le w).$$

We assume the bivariate detection function of Drummer and McDonald (1987), which generalizes the univariate exponential power series detection function of Pollock (1978),

$$g(x,s) = \exp\left(-\left|\frac{x}{\lambda s^\alpha}\right|^p\right), \qquad -\infty < x < \infty, \tag{2.2}$$

where λ is a scale parameter, p is a shape parameter, and $\alpha > 0$ describes the relationship between size and detectability. Figure 1 shows plots of the detection function for various values of α, p, and λ.

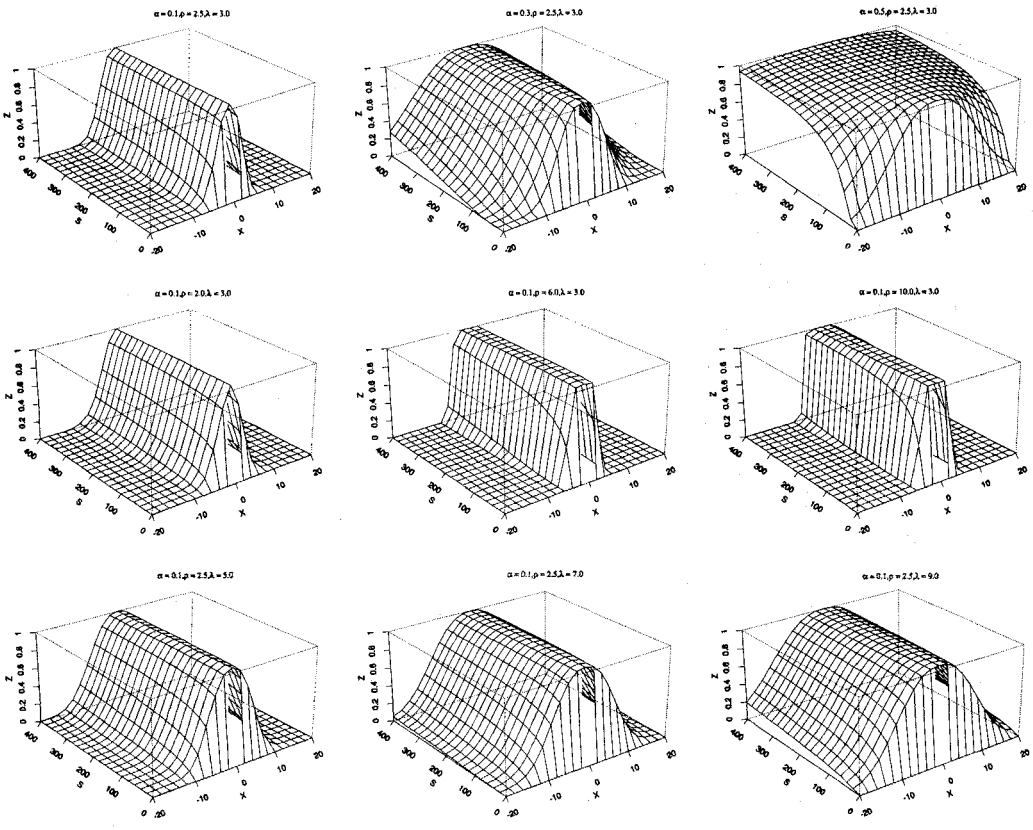

Figure 1. Effect of α, p, and λ on the shape of the detection function.

Then

$$f(x,s) = \frac{g(x,s)f^*(x,s)}{\int\int g(x,s)f^*(x,s)dxds} = \frac{g(x,s)f^*(s)}{\int\int g(x,s)f^*(s)dxds}$$

and $\int_{-\infty}^{\infty} g(x,s)dx = 2\int_0^{\infty} \exp\{-(x/\lambda s^{\alpha})^p\}dx = 2\lambda s^{\alpha}\Gamma(1+1/p)$, so

$$\begin{aligned} f(x,s) &= \frac{e^{-|x/\lambda s^{\alpha}|^p} s^{\nu_1-1} e^{-s/\nu_2}}{2\lambda\Gamma(1+1/p)\int_0^{\infty} s^{\nu_1+\alpha-1}e^{-s/\nu_2}ds} \\ &= \frac{e^{-|x/\lambda s^{\alpha}|^p} s^{\nu_1-1} e^{-s/\nu_2}}{2\lambda\Gamma(1+1/p)\Gamma(\nu_1+\alpha)\nu_2^{\nu_1+\alpha}}. \end{aligned} \tag{2.3}$$

Hence,

$$f(s) = \frac{s^{\nu_1+\alpha-1}e^{-s/\nu_2}}{\Gamma(\nu_1+\alpha)\nu_2^{\nu_1+\alpha}},$$

and so the observed school size

$$S \sim \text{gamma}(\nu_1+\alpha, \nu_2). \tag{2.4}$$

There is an increase of α in the shape parameter because larger clusters are more likely to be detected.

Under the above parametric model,

$$\beta(0) = \int_0^{\infty} sf(0,s)ds = \frac{\Gamma(\nu_1+1)}{2\lambda\Gamma(1+1/p)\Gamma(\nu_1+\alpha)\nu_2^{\alpha-1}}.$$

We now derive moment estimators of the unknown parameters $(\nu_1, \nu_2, \lambda, p, \alpha)$ in $\beta(0)$, assuming that ϵ_x is symmetric and the first two nonzero moments of ϵ_x and ϵ_s are known. Recall that, because of measurement errors, we observe Y and T instead of X and S. Define for integers j and k

$$\bar{T}_j = n^{-1}\sum_{i=1}^{n} T_i^j,$$

$$\bar{Y}_j = n^{-1}\sum_{i=1}^{n} Y_i^j,$$

and

$$\overline{YT}_{jk} = n^{-1}\sum_{i=1}^{n} Y_i^j T_i^k.$$

In Appendix 1, we show that

$$\mathrm{E}\left(X^j S^k\right) = \frac{\lambda^j\Gamma\{(j+1)/p)\}\Gamma(\nu_1+\alpha+j\alpha+k)\nu_2^{j\alpha+k}}{\Gamma(1/p)\Gamma(\nu_1+\alpha)}, \tag{2.5}$$

where j and k are positive integers and j is even.

From (2.1) and writing $\mu_{s1} = \mathrm{E}(\epsilon_s)$, $\mu_{s2} = \mathrm{E}(\epsilon_s^2)$, $\mu_{x4} = \mathrm{E}(\epsilon_x^4)$, and $\sigma_x^2 = \mathrm{var}(\epsilon_x)$, we have

$$\begin{aligned} \mathrm{E}(T) &= \mathrm{E}(S)\mu_{s1} \\ \mathrm{E}\left(T^2\right) &= \mathrm{E}\left(S^2\right)\mu_{s2} \\ \mathrm{E}\left(Y^2\right) &= \mathrm{E}\left(X^2\right) + \sigma_x^2 \\ \mathrm{E}\left(Y^2T\right) &= \left\{\mathrm{E}\left(X^2S\right) + \mathrm{E}(S)\sigma_x^2\right\}\mu_{s1} \end{aligned}$$

$$\mathrm{E}\left(Y^4\right) = \mathrm{E}\left(X^4\right) + 6\mathrm{E}\left(X^2\right)\sigma_x^2 + \mu_{x4}.$$

Replacing $\mathrm{E}(T)$, $\mathrm{E}(T^2)$, $\mathrm{E}(Y^2)$, $\mathrm{E}(Y^2T)$, and $\mathrm{E}(Y^4)$ by their sample counterparts and substituting the results given in (2.5), we have the following estimating functions for $(\nu_1, \nu_2, \lambda, p, \alpha)$:

$$\bar{T} = (\nu_1 + \alpha)\nu_2\mu_{s1}, \tag{2.6}$$

$$\bar{T}_2 = (\nu_1 + \alpha)(\nu_1 + \alpha + 1)\nu_2^2\mu_{s2}, \tag{2.7}$$

$$\bar{Y}_2 = \frac{\lambda^2\Gamma(3/p)\Gamma(\nu_1 + 3\alpha)\nu_2^{2\alpha}}{\Gamma(1/p)\Gamma(\nu_1 + \alpha)} + \sigma_x^2, \tag{2.8}$$

$$\overline{YT}_{21} = \left\{\frac{\lambda^2\Gamma(3/p)\Gamma(\nu_1 + 3\alpha + 1)\nu_2^{2\alpha+1}}{\Gamma(1/p)\Gamma(\nu_1 + \alpha)} + \sigma_x^2(\nu_1 + \alpha)\nu_2\right\}\mu_{s1}, \tag{2.9}$$

$$\bar{Y}_4 = \frac{\lambda^4\Gamma(5/p)\Gamma(\nu_1 + 5\alpha)\nu_2^{4\alpha}}{\Gamma(1/p)\Gamma(\nu_1 + \alpha)} + 6\sigma_x^2(\bar{Y}_2 - \sigma_x^2) + \mu_{x4}. \tag{2.10}$$

From (2.6) and (2.7), we have

$$\hat{\nu}_2 = \left(\mu_{s1}^2\bar{T}_2 - \mu_{s2}\bar{T}^2\right)/\mu_{s1}\mu_{s2}\bar{T} \tag{2.11}$$

and

$$\hat{\nu}_1 + \hat{\alpha} = \mu_{s2}\bar{T}^2/\left(\mu_{s1}^2\bar{T}_2 - \mu_{s2}\bar{T}^2\right). \tag{2.12}$$

Estimators for α and ν_1 are obtained from (2.8), (2.9), and (2.12) as

$$\hat{\alpha} = \frac{\mu_{s2}\bar{T}(\overline{YT}_{21} - \bar{Y}_2\bar{T})}{2\left(\bar{Y}_2 - \sigma_x^2\right)\left(\mu_{s1}^2\bar{T}_2 - \mu_{s2}\bar{T}^2\right)}, \tag{2.13}$$

$$\hat{\nu}_1 = \frac{\mu_{s2}\bar{T}\left(3\bar{Y}_2\bar{T} - 2\bar{T}\sigma_x^2 - \overline{YT}_{21}\right)}{2\left(\bar{Y}_2 - \sigma_x^2\right)\left(\mu_{s1}^2\bar{T}_2 - \mu_{s2}\bar{T}^2\right)}. \tag{2.14}$$

Manipulating (2.8) and (2.10), we obtain as estimating function for p

$$T(p) = A, \tag{2.15}$$

where

$$A = \frac{\left\{\bar{Y}_4 - \mu_{x4} - 6\sigma_x^2\left(\bar{Y}_2 - \sigma_x^2\right)\right\}\Gamma^2(\hat{\nu}_1 + 3\hat{\alpha})}{\left(\bar{Y}_2 - \sigma_x^2\right)^2\Gamma(\hat{\nu}_1 + \hat{\alpha})\Gamma(\hat{\nu}_1 + 5\hat{\alpha})} \tag{2.16}$$

and

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

As shown in Chen (1998), $T(p)$ is a monotonic decreasing function and $\lim_{p\to\infty} = 9/5$, so there is a unique solution for p if $A > 9/5$. When $A \le 9/5$, $\hat{p}$ is assigned a large value, implying a constant detection function. Occasional low values of A occur in finite samples when the data show a flat shoulder near $x = 0$.

Substituting the estimates into (2.8), an estimator for λ is

$$\hat{\lambda} = \sqrt{\frac{\left(\bar{Y}_2 - \sigma_x^2\right)\Gamma(1/\hat{p})\Gamma(\hat{\nu}_1 + \hat{\alpha})}{\Gamma(3/\hat{p})\Gamma(\hat{\nu}_1 + 3\hat{\alpha})\hat{\nu}_2^{2\hat{\alpha}}}}. \tag{2.17}$$

When the shape parameter p is known, the estimating function (2.10) is not required. The estimators for ν_1, ν_2, α, and λ given in (2.14), (2.11), (2.13), and (2.17) remain unchanged. The only minor change is that the known value p is used instead of $\hat{p}$ in (2.17).

As all the parameter estimators given above are smooth functions of independent and identically distributed means, they are asymptotically normally distributed, and it can be shown that the estimators are asymptotically unbiased for the parameters (Serfling, 1980). Also,

$$\hat{\beta}(0) = \frac{\Gamma(\hat{\nu}_1 + 1)}{2\hat{\lambda}\Gamma(1 + 1/\hat{p})\Gamma(\hat{\nu}_1 + \hat{\alpha})\hat{\nu}_2^{\hat{\alpha}-1}}$$

is a smooth function of the estimators, so it is asymptotically normally distributed and has $\beta(0)$ as its asymptotic mean. This ensures that

$$\hat{D}_1 = L^{-1}n\hat{\beta}(0)$$

is an asymptotically unbiased estimator for D_1 and has an asymptotic normal distribution. The latter can be used to construct confidence intervals for D_1. As the variance is quite complicated to derive theoretically because of the smooth functions involved, the bootstrap should be used to evaluate the variance.

3. SBT Surveys

In this section, we apply the estimation method of Section 2 to data collected in the 1994–1998 SBT line transect surveys in order to assess the effect of measurement errors on the density estimates from the surveys. All distances are in nautical miles and school sizes are in tons.

In the 1998 survey season, experiments were conducted in which two planes independently estimated the location of detected schools and the size of each school. From these data, estimates of various moments of the measurement errors in x and s were obtained as $\hat{\sigma}_x = 0.71$, $\hat{\mu}_{s1} = 1.59$, $\hat{\mu}_{s2} = 3.09$. In the analyses below, we assumed that errors in distance were normally distributed so that $\mu_{x4} = 3(\sigma_x^2)^2$. This assumption was borne out by the data.

In the analysis of the SBT survey data in this article, we have assumed that all data were collected by the same plane for which $\mu_{s1} \neq 1$, and so the observed school size T is biased for the true school size S. We define the mean squared deviation (MSD) of the error in size ϵ_s as

$$\mathrm{E}(\epsilon_s - 1)^2 = \mathrm{var}(\epsilon_s) + \{\mathrm{E}(\epsilon_s) - 1\}^2.$$

We compare the effects of errors with the same MSD but with different values of $\mathrm{E}(\epsilon_s) - 1$.

The effects of correcting various measurement errors on the parameter and density estimates are shown in Table 1. We show uncorrected estimates, estimates corrected for measurement errors in distance but not size, estimates corrected for measurement error in distance and a hypothetical measurement error in size with the same MSD as the observed error in the tuna surveys but with $\mu_{s1} = 1$, and estimates corrected for measurement error in distance and the measurement error in size observed in the tuna surveys.

The uncorrected estimates were obtained using $\sigma_x = 0.00$, $\mu_{x4} = 0.00$, $\mu_{s1} = 1.00$, and $\mu_{s2} = 1.00$ in equations (2.14), (2.11), (2.13), (2.15), and (2.17), solving equation (2.15) by

Table 1

Parameter estimates from the SBT surveys 1994–1998 *with corrections for measurement errors. This series of estimates is one of several series of estimates derived from this data and should not be interpreted in isolation.*

Error[a]	Year	n	10^{-4} L	$10\hat{\nu}_1$	$10^{-2}\hat{\nu}_2$	$\hat{p}$	$\hat{\lambda}$	$10\hat{\alpha}$	$10^2\hat{D}_0$	$10\hat{D}_1$	$10^{-1}\hat{\mathrm{E}}(S)$
U	1994	289	2.02	4.04	1.93	1.58	4.00	0.91	1.52	1.18	7.82
	1995	294	2.08	2.25	3.21	1.98	3.38	1.21	1.84	1.33	7.21
	1996	185	1.82	3.29	2.96	2.51	4.03	1.33	0.94	0.92	9.71
	1997	189	1.28	4.18	2.22	2.43	5.45	0.42	1.34	1.24	9.28
	1998	146	1.19	1.68	4.07	2.10	4.61	0.29	1.49	1.02	6.83
C_1	1994	289	2.02	4.02	1.93	1.56	3.88	0.93	1.55	1.20	7.77
	1995	294	2.08	2.20	3.21	1.96	3.25	1.26	1.91	1.35	7.07
	1996	185	1.82	3.25	2.96	2.53	3.92	1.36	0.96	0.92	9.61
	1997	189	1.28	4.16	2.22	2.45	5.37	0.43	1.36	1.25	9.25
	1998	146	1.19	1.66	4.07	2.10	4.49	0.30	1.53	1.04	6.77
C_2	1994	289	2.02	13.97	0.56	1.81	1.43	3.24	1.49	1.16	7.77
	1995	294	2.08	6.14	1.15	2.65	1.16	3.51	1.85	1.30	7.07
	1996	185	1.82	10.71	0.90	4.71	0.86	4.49	0.93	0.89	9.61
	1997	189	1.28	13.67	0.68	2.59	3.40	1.43	1.34	1.24	9.25
	1998	146	1.19	3.87	1.75	2.14	3.73	0.70	1.52	1.03	6.77
C_3	1994	289	2.02	5.52	0.89	1.59	3.53	1.28	1.54	0.75	4.89
	1995	294	2.08	2.92	1.52	2.05	2.90	1.67	1.90	0.84	4.44
	1996	185	1.82	4.43	1.36	2.70	3.36	1.86	0.95	0.58	6.04
	1997	189	1.28	5.67	1.03	2.47	5.12	0.59	1.35	0.79	5.82
	1998	146	1.19	2.12	2.00	2.10	4.39	0.39	1.53	0.65	4.26

[a] U, uncorrected estimates; C_1, estimates corrected for observed errors in distance but uncorrected for errors in size; C_2, estimates corrected for observed errors in distance and hypothetical errors in size; C_3, estimates corrected for observed errors in distance and observed errors in size.

Brent's method. Estimates corrected for measurement error in distance but not size were obtained using $\sigma_x = 0.71$, $\mu_{x4} = 3(\sigma_x^2)^2$, $\mu_{s1} = 1.00$, and $\mu_{s2} = 1.00$ in the same equations. Estimates corrected for error in distance and the hypothetical error in size were obtained using $\sigma_x = 0.71$, $\mu_{x4} = 3(\sigma_x^2)^2$, $\mu_{s1} = 1.00$, and $\mu_{s2} = 1.91$. Estimates corrected for the errors observed in the survey were obtained using $\sigma_x = 0.71$, $\mu_{x4} = 3(\sigma_x^2)^2$, $\mu_{s1} = 1.59$, and $\mu_{s2} = 3.09$.

Table 1 shows that correction for only additive mean zero measurement errors in distance has little effect on either the parameter estimates or the density estimates. Correction for multiplicative errors in size with $\mu_{s1} = 1$ has a large effect on the estimates of $(\nu_1, \nu_2, p, \lambda, \alpha)$ but little effect on the density estimates. When $\mu_{s1} \neq 1$, correction for errors in size has a great effect on both the estimates of $(\nu_1, \nu_2, p, \lambda, \alpha)$ and the biomass density D_1.

4. Simulation Study

A simulation study was conducted to compare the relative root mean squared error, relative bias, and relative standard error of the uncorrected estimates and of estimates with corrections for the case when $\mu_{s1} = 1$ and for the case when $\mu_{s1} \neq 1$. In each case, the parameters ν_1, ν_2, p, λ, α and the biomass density D_1 were estimated in 500 simulations using various corrections for measurement errors; we compare uncorrected estimates, estimates corrected for measurement errors in distance but not size, estimates corrected for measurement error in distance and an incorrect measurement error in size, and estimates corrected for measurement error in distance and the correct measurement error in size. The simulation results are given in Tables 2 and 3.

The parameter values used in the study were chosen to be similar to the parameter estimates obtained from the SBT data. They were

$$(\nu_1, \nu_2, p, \lambda, \alpha, D_0) = (0.4, 200, 2.5, 3.0, 0.1, 0.015).$$

The values of the measurement errors used in the simulation study were based on those observed in the tuna surveys. We used $\sigma_x = 0.71$ and $\mu_{x4} = 3(\sigma_x^2)^2$ in both cases. In case 1, the errors in size were generated and estimated using $(\mu_{s1}, \mu_{s2}) = (1.00, 2.00)$. In case 2, the errors in size were generated and estimated using $(\mu_{s1}, \mu_{s2}) = (1.58, 3.09)$, as observed in the tuna surveys.

In each simulation, a realization of a homogeneous Poisson process with intensity 0.015 was generated on a 40 × 2000 rectangle to represent the locations of the sightings, although in this study, only the perpendicular distances of the sightings from the transect line were used. A gamma(ν_1, ν_2) random variable was generated to represent the size of each sighting. The detection function in (2.2) was used to determine which of the sightings were detected from a transect line of length 2000 placed along the center of the rectangle.

Measured perpendicular distances were obtained by adding $\mathrm{N}(0, \sigma_1^2)$ random variables to the generated perpendicular distances. Measured sighting sizes were obtained by multiplying the generated sighting sizes by gamma$(\mu_{s1}^2/(\mu_{s2} - \mu_{s1}^2), (\mu_{s2} - \mu_{s1}^2)/\mu_{s1})$ random variables.

Table 2

Simulation study: case 1, $E(\epsilon_s = 1)$. *Percentage relative root mean squared error (RMSE), percentage relative bias (RB), and percentage relative standard deviation (RSD) of parameter estimates from* 500 *simulations, with corrections for measurement errors.*[a]

		p Estimated				p Known			
		U	C_1	C_2	C_3	U	C_1	C_2	C_3
$\hat{\nu}_1$	RMSE	53.9	54.5	35.3	62.6	54.6	55.3	36.1	60.8
	RB	−51.4	−52.1	−17.4	29.8	−52.1	−52.8	−18.8	27.3
	RSD	16.1	16.1	30.7	55.0	16.3	16.4	30.8	54.4
$\hat{\nu}_2$	RMSE	177.9	177.9	83.8	52.9	181.7	181.7	86.5	54.7
	RB	140.0	140.0	43.3	−5.0	141.9	141.9	44.5	−4.2
	RSD	109.8	109.8	71.7	52.7	113.5	113.5	74.2	54.5
$\hat{p}$	RMSE	25.5	32.7	42.3	98.0	—	—	—	—
	RB	3.6	7.6	13.6	33.7	—	—	—	—
	RSD	25.3	31.8	40.0	92.0	—	—	—	—
$\hat{\lambda}$	RMSE	37.8	34.0	25.3	29.7	39.4	34.8	27.5	32.7
	RB	35.0	30.5	13.9	−2.9	37.0	31.8	14.4	−2.8
	RSD	14.3	15.0	21.2	29.6	13.7	14.2	23.4	32.6
$\hat{\alpha}$	RMSE	66.8	65.4	61.9	92.0	66.2	64.8	60.7	88.6
	RB	−59.6	−57.0	−26.1	15.9	−59.0	−56.4	−25.4	16.0
	RSD	30.2	32.1	56.1	90.7	30.1	32.0	55.1	87.2
$\hat{D}_1$	RMSE	22.3	23.8	23.8	23.8	19.5	20.3	20.1	20.0
	RB	1.2	3.3	3.0	2.6	−1.1	1.4	2.1	3.2
	RSD	22.2	23.5	23.6	23.6	19.5	20.2	20.0	19.7

[a] U, uncorrected estimates; C_1, estimates corrected for observed errors in distance but uncorrected for errors in size; C_2, estimates corrected for observed errors in distance and hypothetical errors in size; C_3, estimates corrected for observed errors in distance and observed errors in size.

There were occasional difficulties in estimating p in the simulations. As stated in Section 2, equation (2.15) cannot be solved for p when $A \leq 9/5$. If this occurred during the simulations, the program terminated. For this reason, in Tables 2 and 3, we also show the results of simulations in which p was known to be 2.5.

The results in Table 2 show that, when $\mu_{s1} = 1$, there is very little difference between the corrected and uncorrected estimates of biomass density D_1, although correction has a strong effect on each of the parameters ν_1, ν_2, p, λ, and α.

The results in Table 3 show that, when $\mu_{s1} \neq 1$, there is a need to correctly estimate and adjust for μ_{s1} if the biomass density D_1 is to be well estimated.

The results in both tables show that, when p is known, the estimates of D_1 have a lower MSE due to a reduction in variance. There is little change in the bias of D_1, which is very low in both cases. Similarly, there is little effect on the estimates of ν_1, ν_2, p, λ, and α.

5. Interpretation

At first glance, recalling that $D_1 = D_0 E(S \mid x = 0)$, it seems obvious that, if there is a multiplicative bias in the size estimates, the density estimate will be biased by approximately the same amount. However, the effect of measurement errors is not as simple as this, and unbiased measurement errors in covariates can cause considerable bias in an overall estimate (Carroll, Ruppert, and Stefanski, 1995). Chen (1998) found that unbiased measurement errors in distance had a noticeable effect on the estimates of D_0 in univariate distance-only line transect surveys. Moreover, the effect of multivariate measurement errors is not as clear cut as the effect of univariate errors. Carroll et al. (1995) also found this for measurement error in multiple linear regression.

When $\mu_{s1} = 1$, little difference was observed between the corrected and uncorrected estimates of D_1 in both the SBT surveys example and the simulation study. In Appendix 2, we present an analysis that explains the lack of difference.

The most important result in that appendix is that, for values of the parameters similar to those observed in the tuna survey, equation (A.2) in Appendix 2 reduces to

$$\hat{\beta}_u(0) \simeq \mu_{s1}\, \beta(0),$$

where $\hat{\beta}_u(0)$ is the uncorrected version of $\hat{\beta}_u 0$.

Thus, the uncorrected density estimates are biased only when $\mu_{s1} \neq 1$. For parameter values very different from those in the tuna survey, μ_{s1} will remain an important factor, but it is possible that the other parameters might make a contribution to the uncorrected density estimate.

It should be noted that this result is a property of the functional form of the bivariate detection function used in this article rather than an inherent property of the method of moments estimator derived here.

6. Discussion

As far as we know, this is the first work to examine the problem of correcting for measurement errors in size in line transect surveys. One reason for this may be the difficulty of estimating possible errors in size.

Table 3

Simulation study: case 2, $E(\epsilon_s \neq 1)$. Percentage relative root mean squared error (RMSE), percentage relative bias (RB), and percentage relative standard deviation (RSD) of parameter estimates from 500 simulations, with corrections for measurement errors.

		p Estimated				p Known			
		U	C_1	C_2	C_3	U	C_1	C_2	C_3
$\hat{\nu}_1$	RMSE	28.7	29.6	63.4	34.0	26.9	27.7	63.4	32.7
	RB	−17.8	−19.0	43.5	9.4	−16.7	−17.9	45.5	10.9
	RSD	22.5	22.7	46.1	32.7	21.1	21.2	44.1	30.9
$\hat{\nu}_2$	RMSE	125.8	125.8	32.8	31.6	122.3	122.3	31.4	29.6
	RB	109.2	109.2	−13.5	−1.4	107.4	107.4	−14.4	−2.4
	RSD	62.4	62.4	29.9	31.5	58.6	58.6	28.0	29.5
$\hat{p}$	RMSE	28.4	40.0	81.3	51.0	—	—	—	—
	RB	7.9	13.3	31.6	19.4	—	—	—	—
	RSD	27.3	37.7	74.9	47.1	—	—	—	—
$\hat{\lambda}$	RMSE	21.9	19.8	27.7	20.3	23.5	20.9	29.4	21.9
	RB	13.4	8.4	−13.8	−0.2	14.7	9.2	−13.6	0.2
	RSD	17.3	17.9	24.0	20.3	18.4	18.8	26.1	21.9
$\hat{\alpha}$	RMSE	47.9	47.7	84.8	57.9	47.0	46.4	79.8	54.2
	RB	−26.1	−21.4	38.1	5.7	−28.4	−23.8	34.4	2.7
	RSD	40.1	42.6	75.8	57.6	37.4	39.8	72.0	54.2
$\hat{D}_1$	RMSE	63.6	67.0	25.8	19.5	61.5	65.5	23.6	14.9
	RB	56.5	59.4	13.2	0.0	57.1	61.1	16.7	1.9
	RSD	29.2	31.0	22.2	19.5	22.9	23.7	16.6	14.8

[a] U, uncorrected estimates; C_1, estimates corrected for observed errors in distance but uncorrected for errors in size; C_2, estimates corrected for observed errors in distance and hypothetical errors in size; C_3, estimates corrected for observed errors in distance and observed errors in size.

In this article, we used an additive model for errors in distance and a multiplicative model for errors in size. The method of moments can be used in a similar way if an additive model is used for errors in size. The estimators based on additive size errors have not been included in this article since the multiplicative model is generally thought more appropriate for errors in size.

The models and distributions in this analysis are mathematically tractable, allowing the joint distribution in (2.3) and its moments in (2.5) to be derived quite simply. Moreover, the bivariate detection function used in this article is very general, allowing a wide range of shapes of detection function to be fitted. The gamma distribution for the observed sizes in (2.4) is flexible and gives a good fit to the observed sizes in this example. It would be difficult to use maximum likelihood in this analysis because it is very hard to derive the distribution of (Y, T) from (2.1) and (2.3) even if distributions for the errors are assumed. Thus, in this analysis, the method of moments has a major advantage.

The parametric approach used in this article might be inappropriate when the distributional assumptions do not hold. However, nonparametric approaches to adjusting for measurement error are complex and hindered by the slow rate of convergence of the estimators. Nevertheless, further work in this area might prove fruitful.

The analytical results in Appendix 2 explain why the data analysis and simulations indicated that biomass density estimates need not be adjusted for additive mean zero measurement errors in distance or measurement errors in size unless $\mu_{s1} \neq 1$ for parameter configurations close to those observed in the SBT surveys. We do not know how far these results can be applied to other surveys with different parameter configurations.

The bivariate detection function used in this article has a certain inherent resistance to measurement errors in size when $\mu_{s1} = 1$. This resistance may not be shared by other bivariate detection functions since it is a property of the functional form rather than the method of moments used for estimation in this article. Chen (1998) found that the univariate detection function from which this bivariate detection function is generalized was not very resistant to measurement errors.

These are important results for line transect surveys. In surveys where detectability varies with distance and size, there may be significant bias in the biomass density estimate if the size estimates are biased. It is therefore important to assess whether there is measurement error in size.

In line transect surveys with multiple observers, the size estimates of the different observers should be ground-truthed. If, e.g., all observers underestimate size, the biomass density may be underestimated. However, if the mean size of the pooled observers is unbiased, there may be no need to correct for errors in size.

In the SBT survey, it is highly probable that the observers will underestimate cluster size. First, echo sounders and sonars on fishing vessels show that schools of SBT may extend up to 40–60 m below the surface, well below the visibility limits of the observers. Second, commercial spotters tend to underestimate school size: fishing boats are not pleased

when spotters direct them to a school of fish that is smaller than the spotters' estimates but are happy when the school is larger than the estimate. When commercial spotters discuss the feedback from the fishing vessels, they report that their estimates are very often too small and very seldom too large.

In the SBT survey, the gamma distribution was used for cluster size because the size of a sighting of SBT (in tons) is continuous and positive. In many line transect surveys, clusters are groups of animals, and so discrete distributions such as the Poisson distribution truncated above zero would be more appropriate. It would be interesting to know to what extent these results generalize to the discrete case.

Acknowledgements

This research is part of a collaborative research project to monitor the abundance of southern bluefin tuna that is being conducted between Australia (through the CSIRO) and Japan (through the National Research Institute of Far Seas Fisheries of the Japan Fisheries Agency). The research was funded by the Australian government and tuna industry through AFMA and FRDC and the Japanese government and tuna industry through JAMARC.

Résumé

Lorsque l'on utilise les méthodes de la ligne de transect bivariée pour estimer la densité de biomasse d'une population biologique fortement agrégée, il est généralement supposé que la distance entre la ligne de transect et l'agrégat d'une part et la taille ou la biomasse de l'agrégat d'autre part sont mesurées sans erreur. Ce n'est que très rarement le cas en pratique. Dans cet article, nous supposons que les erreurs sont additives et de moyenne nulle pour les distances et multiplicatives pour la taille et nous développons un estimateur de la densité qui corrige ces erreurs. Nous utilisons la méthode des moments dans le cas de taille d'agrégats suivant une loi Gamma, de lignes de transect placées aléatoirement et d'une fonction de détection de type exponentielle. Nous dérivons des résultats qui montrent qu'il n'est pas nécessaire de corriger pour les erreurs sur la distance ou la taille lorsque les estimations de la distance ou de la taille ne sont pas biaisées. Lorsque les estimations de la taille sont biaisées, l'estimation de la densité de biomasse a approximativement le même biais que les estimations de la taille. Le travail est illustré dans le contexte de recensements annuels aériens de juvéniles de thon rouge dans la Grand Golfe australien (Great Australian Bight).

References

Alpizar-Jara, R. and Pollock, K. P. (1996). A combination line transect and capture–recapture sampling model for multiple observers in aerial surveys. *Journal of Environmental and Ecological Statistics* **3,** 311–327.

Alpizar-Jara, R., Stefanski, L. A., Pollock, K. H., and Laake, J. L. (1998). *Assessing the effects of measurement errors in line transect sampling.* Technical Report, Institute of Statistics, Mimeograph Series 2508, North Carolina State University, Raleigh.

Anonymous. (1996). Report of the second meeting of the scientific committee of the Commission for the Conservation of Southern Bluefin Tuna. Technical Report, August 26–September 5, Commission for the Conservation of Southern Bluefin Tuna, Hobart, Australia.

Borchers, D. L. (1996). Line transect abundance estimation with uncertain detection on the trackline. Ph.D. thesis, University of Cape Town, Cape Town, South Africa.

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

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

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). Studying school size effects in line transect sampling using the kernel method. *Biometrics* **52,** 1283–1294.

Chen, S. X. (1998). Measurement errors in line transect surveys. *Biometrics* **54,** 899–908.

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

Drummer, T. D. and McDonald, L. L. (1987). Size bias in line transect sampling. *Biometrics* **43,** 13–21.

Mack, Y. P. and Quang, P. X. (1998). Kernel methods in line and point transect sampling. *Biometrics* **54,** 606–619.

Manly, B. F. J., McDonald, L. L., and Garner, G. W. (1996). Maximum likelihood estimation for the double-count method with independent observers. *Journal of Agricultural, Biological, and Environmental Statistics* **1,** 170–189.

Pollock, K. H. (1978). A family of density estimators for line-transect sampling. *Biometrics* **47,** 475–478.

Quang, P. X. (1991). A nonparametric approach to size-biased line transect sampling. *Biometrics* **47,** 269–279.

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 April 2000. *Revised December* 2000.
Accepted January 2001.

Appendix 1

We need to show that

$$\mathrm{E}\left(X^j S^k\right) = \frac{\lambda^j \Gamma\{(j+1)/p\}\Gamma(\nu_1+\alpha+j\alpha+k)\nu_2^{j\alpha+k}}{\Gamma(1/p)\Gamma(\nu_1+\alpha)},$$

where j and k are positive integers and j is even.

From (2.3),

$$\mathrm{E}\left(X^j S^k\right) = \int_{-\infty}^{\infty}\int_{0}^{\infty} x^j s^k f(x,s)dsdx$$

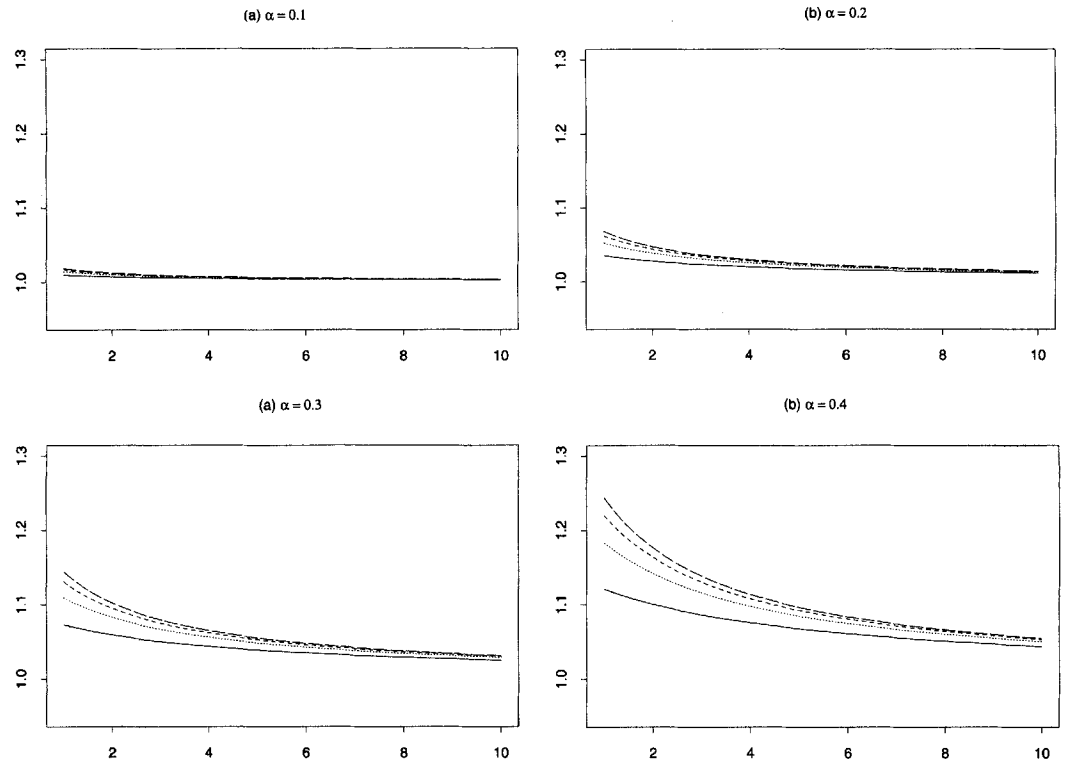

Figure 2. Effect of α, $\nu_1 + \alpha$, and η on r_1 assuming $\sigma_x^2 = 0$. Horizontal axis is $\nu_1 + \alpha$, vertical axis is r_1. Solid line is $\eta = 1.25$, dotted line is $\eta = 1.5$, short dashed line is $\eta = 1.5$, and long dashed line is $\eta = 2.5$.

$$
\begin{aligned}
&= \frac{1}{\lambda\Gamma(1+1/p)\Gamma(\nu_1+\alpha)\nu_2^{\nu_1+\alpha}} \\
&\quad \times \int_0^\infty \left\{ \int_0^\infty x^j e^{-(x/\lambda s^\alpha)^p} dx \right\} \\
&\qquad \times s^{\nu_1+k-1} e^{-s/\nu_2} ds \\
&= \frac{1}{\lambda\Gamma(1+1/p)\Gamma(\nu_1+\alpha)\nu_2^{\nu_1+\alpha}} \\
&\quad \times \int_0^\infty \left\{ \int_0^\infty \frac{(\lambda s^\alpha)^{j+1}}{p} t^{(j+1)/p-1} e^{-t} dt \right\} \\
&\qquad \times s^{\nu_1+k-1} e^{-s/\nu_2} ds \\
&= \frac{\lambda^{j+1}\Gamma\{(j+1)/p\}}{\lambda p\Gamma(1+1/p)\Gamma(\nu_1+\alpha)\nu_2^{\nu_1+\alpha}} \\
&\quad \times \int_0^\infty s^{\alpha(j+1)+\nu_1+k-1} e^{-s/\nu_2} ds \\
&= \frac{\lambda^j\Gamma\{(j+1)/p\}\Gamma(\nu_1+\alpha+j\alpha+k)\nu_2^{j\alpha+k}}{\Gamma(1/p)\Gamma(\nu_1+\alpha)}.
\end{aligned}
$$

Appendix 2

In this appendix, we study the impact of ignoring measurement error on the estimate of $\beta(0)$ and thus on D_1. An analysis of the impact on $f(0)$ and D_0 can be formulated similarly. Throughout this appendix, we use $\hat{\theta}_u$ to denote the estimator of a parameter θ that ignores both ϵ_x and ϵ_s. Similarly, we use θ_u to denote a constant to which $\hat{\theta}_u$ converges in probability.

All uncorrected estimates are obtained by putting $\mu_{s1} = \mu_{s2} = 1$ and $\sigma_x^2 = 0$ in the relevant equations from (2.11) to (2.14).

First, consider the uncorrected estimate of ν_2, $\hat{\nu}_{2u} = (\bar{T}_2 - \bar{T}^2)/\bar{T}$. Applying the strong law of large numbers to the sample moments, it can be shown that

$$
\begin{aligned}
\hat{\nu}_{2u} &\to \left\{ \mathrm{E}\left(T^2\right) - \mathrm{E}^2(T) \right\} / \mathrm{E}(T) \\
&= \nu_2\{\mu_{s2}/\mu_{s1} + (\nu_1+\alpha)(\mu_{s2}/\mu_{s1} - \mu_{s1})\},
\end{aligned}
$$

where, throughout this section, $\to$ means convergence almost surely.

Writing $\eta = \mu_{s2}/\mu_{s1}^2$, $\gamma = \eta + (\nu_1+\alpha)(\eta-1)$, and $\delta = \sigma_x^2/\mathrm{E}(X^2)$, we have $\hat{\nu}_{2u} \to \nu_2\gamma\mu_{s1}$. Similarly, $\hat{\nu}_{1u} \to \nu_1/\gamma + \alpha\delta/\{\gamma(1+\delta)\}$ and $\hat{\alpha}_u \to \alpha/\{\gamma(1+\delta)\}$.

Note that $\eta > 1$ and $\gamma > 1$ provided ϵ_s is not degenerate, and so $\hat{\nu}_{2u}$ overestimates ν_2 and $\hat{\nu}_{1u}$ and $\hat{\alpha}_u$ underestimate ν_1 and α, respectively. This is observed in Tables 1, 2 and 3.

Putting $\mu_{s1} = \mu_{s2} = 1$ and $\sigma_x^2 = 0$ in (2.16), we have

$$
A_u = \frac{\bar{Y}_4\Gamma^2(\hat{\nu}_{1u} + 3\hat{\alpha}_u)}{(\bar{Y}_2)^2\Gamma(\hat{\nu}_{1u}+\hat{\alpha}_u)\Gamma(\hat{\nu}_{1u}+5\hat{\alpha}_u)},
$$

and $\hat{p}_u$ is obtained by solving $T(p_u) = A_u$.

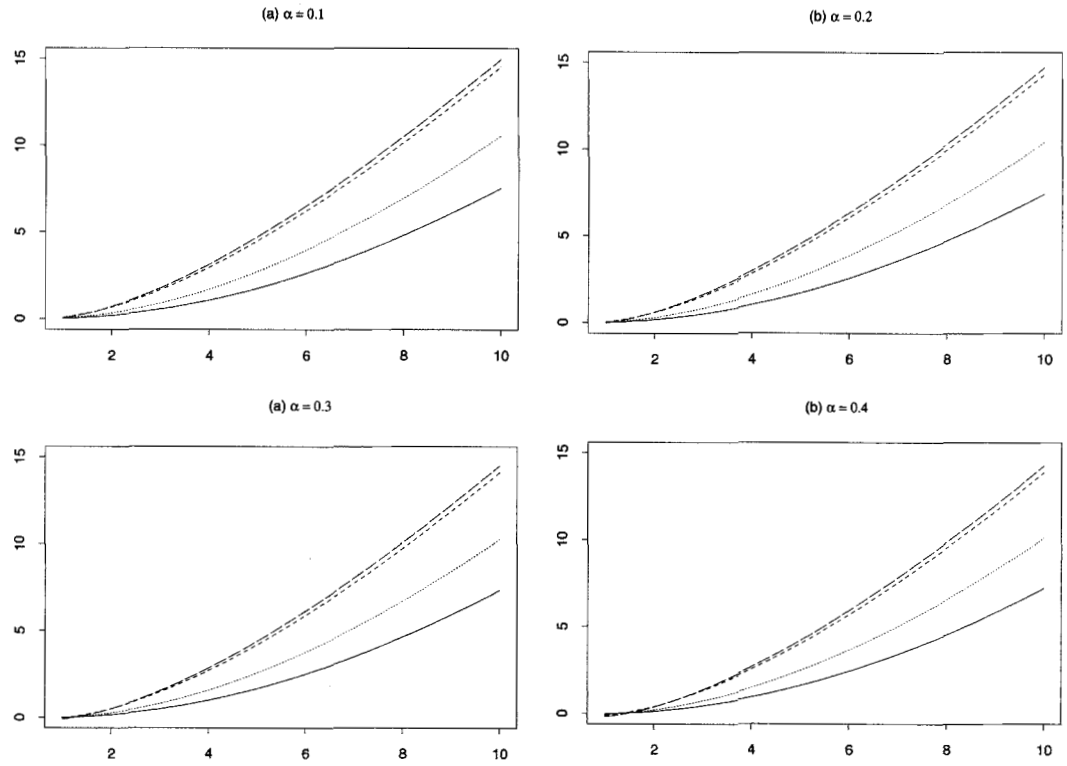

Figure 3. Effect of α, $\nu_1 + \alpha$, and η on $\log(r_3)$ assuming $\sigma_x^2 = 0$. Horizontal axis is $\nu_1 + \alpha$, vertical axis is r_1. Solid line is $\eta = 1.05$, dotted line is $\eta = 1.1$, short dashed line is $\eta = 1.5$, and long dashed line is $\eta = 2.5$.

If ϵ_x is normally distributed, as assumed in the data analysis and the simulations,

$$A_u \to B\frac{\Gamma^2(\nu_{1u} + 3\alpha_u)}{\Gamma(v_{1u} + \alpha_u)\Gamma(v_{1u} + 5\alpha_u)},$$

where

$$B = \frac{T(p)\Gamma(\nu_1 + 5\alpha)\Gamma(\nu_1 + \alpha)\Gamma^{-2}(\nu_1 + 3\alpha) + 6\delta + 3\delta^2}{(1+\delta)^2}. \tag{A.1}$$

If ϵ_x is not normally distributed, there will be additional terms in $\mu_{x4}/\mathrm{E}(X^2)$.

It is likely that δ will be small in well-implemented surveys because the designers would adopt procedures to measure distances accurately. This is the case for the tuna aerial survey and for the configuration used in the simulation. When δ is small relative to the other part of the numerator in (A.1),

$$B \approx T(p)\Gamma(\nu_1 + 5\alpha)\Gamma(\nu_1 + \alpha)\Gamma^{-2}(\nu_1 + 3\alpha).$$

Thus, A_u converges to a quantity close to $T(p)r_1(\nu_1, \alpha, \nu_{1u}, \alpha_u)$, where

$$r_1(\nu_1, \alpha, \nu_{1u}, \alpha_u) = \frac{\Gamma(v_1 + 5\alpha)\Gamma^2(v_{1u} + 3\alpha_u)\Gamma(v_1 + \alpha)}{\Gamma(v_{1u} + 5\alpha_u)\Gamma^2(v_1 + 3\alpha)\Gamma(v_{1u} + \alpha_u)}.$$

If the configuration of $(\nu_1 + \alpha, \alpha, \eta)$ is such that $r_1(\nu_1 + \alpha, \alpha, \eta) \approx 1$, then $\hat{p}_u \approx \hat{p}$. Figure 2 shows r_1 for $\eta = 1.25, 1.5, 1.75$, and 2.0, $\alpha = 0.1, 0.2, 0.3$, and 0.4, and $\nu_1 + \alpha \in (0.25, 10)$. The figure shows that r_1 is quite flat and close to one for small α values. When α is large (0.3 and 0.4), it differs from one by a big margin for small $\nu_1 + \alpha$ but converges to one as $\nu_1 + \alpha$ increases. So, if $\nu_1 + \alpha$ is fairly large or α is small, then $r_1 \approx 1$. This means that failing to correct measurement errors has little effect on $\hat{p}$. Because r_1 is always larger than one, the uncorrected estimate is slightly lower than the corrected estimate. This can be observed in Tables 1 and 2.

To see the effect on λ, note from (2.17) that

$$\hat{\lambda}_u^2 \to (1+\delta)\lambda^2 \frac{\Gamma(3/p)\Gamma(1/p_u)\Gamma(v_{1u} + \alpha_u)\Gamma(v_1 + 3\alpha)v_2^{2\alpha}}{\Gamma(3/p_u)\Gamma(1/p)\Gamma(v_1 + \alpha)\Gamma(v_{1u} + 3\alpha_u)v_{2u}^{2\alpha_u}}.$$

Most importantly, the final effect on $\beta(0)$ is

$$\hat{\beta}_u(0) \to \mu_{s1}\,\beta(0)\,r_2(p, p_u)r_3(\nu_1, \alpha, \nu_{1u}, \alpha_u)/\sqrt{1+\delta}, \tag{A.2}$$

where

$$r_2(p, p_u) = \frac{p_u}{p}\sqrt{\frac{\Gamma(3/p_u)\Gamma^3(1/p)}{\Gamma(3/p)\Gamma^3(1/p_u)}}$$

and

$$r_3(\nu_1, \alpha, \nu_{1u}, \alpha_u, \gamma) = \gamma\sqrt{\frac{\Gamma^3(v_1 + \alpha)\Gamma(v_{1u} + 3\alpha_u)}{\Gamma^3(v_{1u} + \alpha_u)\Gamma(v_1 + 3\alpha)}}.$$

Clearly, if $p_u \approx p$, then $r_2 \approx 1$. Figure 3 shows $\log(r_3)$ for $\eta = 1.25, 1.5, 1.75$, and 2.0, $\alpha = 0.1, 0.2, 0.3$, and 0.4, and $\nu_1 + \alpha \in (0.25, 10)$. It appears that values of α in this range have

little effect on $\log(r_3)$. If $\nu_1 + \alpha < 2$ and δ is small, then $r_3 \approx 1$ as well. In this case, μ_{s1} is the most influential parameter in (A.2). When $\mu_{s1} \neq 1$, the effect of measurement error becomes significant, as shown in the simulations and data analysis. The parameters ν_1, α, and μ_{s2} have only a secondary influence on $\hat{\beta}_u(0)$ through the ratios r_1, r_2, and r_3.