3-2010

Double-Observer Line Transect Methods: Levels of Independence

Stephen T. Buckland
Centre for Research into Ecological and Environmental Modelling, University of St Andrews, The Observatory, Buchanan Gardens, St Andrews, Fife KY16 9LZ, Scotland

Jeffrey L. Laake
National Marine Mammal Laboratory, Alaska Fisheries Science Center

David L. Borchers
Centre for Research into Ecological and Environmental Modelling, University of St Andrews, The Observatory, Buchanan Gardens, St Andrews, Fife KY16 9LZ, Scotland

Double-Observer Line Transect Methods: Levels of Independence

Stephen T. Buckland,[1,*] **Jeffrey L. Laake,**[2] **and David L. Borchers**[1]

[1]Centre for Research into Ecological and Environmental Modelling, University of St Andrews, The Observatory,
Buchanan Gardens, St Andrews, Fife KY16 9LZ, Scotland
[2]National Marine Mammal Laboratory, Alaska Fisheries Science Center, NMFS, Seattle, Washington 98115, U.S.A.
*email: steve@mcs.st-and.ac.uk

SUMMARY. Double-observer line transect methods are becoming increasingly widespread, especially for the estimation of marine mammal abundance from aerial and shipboard surveys when detection of animals on the line is uncertain. The resulting data supplement conventional distance sampling data with two-sample mark–recapture data. Like conventional mark–recapture data, these have inherent problems for estimating abundance in the presence of heterogeneity. Unlike conventional mark–recapture methods, line transect methods use knowledge of the distribution of a covariate, which affects detection probability (namely, distance from the transect line) in inference. This knowledge can be used to diagnose unmodeled heterogeneity in the mark–recapture component of the data. By modeling the covariance in detection probabilities with distance, we show how the estimation problem can be formulated in terms of different levels of independence. At one extreme, full independence is assumed, as in the Petersen estimator (which does not use distance data); at the other extreme, independence only occurs in the limit as detection probability tends to one. Between the two extremes, there is a range of models, including those currently in common use, which have intermediate levels of independence. We show how this framework can be used to provide more reliable analysis of double-observer line transect data. We test the methods by simulation, and by analysis of a dataset for which true abundance is known. We illustrate the approach through analysis of minke whale sightings data from the North Sea and adjacent waters.

KEY WORDS: Distance sampling; Double-observer methods; Full independence; Limiting independence; Line transect sampling; Point independence.

1. Introduction

Distance sampling (Buckland et al., 2001) is widely used for estimating animal abundance. In line transect sampling, an observer travels along each of a number of lines, laid out according to some randomized (usually systematic random) scheme, and records each detected animal, together with its perpendicular distance from the line. One of the key assumptions of the method is that animals on the line are certain to be detected.

A number of authors have considered so-called double-observer or double-platform methods to extend line transect sampling to the case that not all animals on the line are detected (e.g., Buckland and Turnock, 1992; Palka, 1995; Alpizar-Jara and Pollock, 1996; Manly, McDonald, and Garner, 1996; Quang and Becker, 1997; Chen, 2000; Innes et al., 2002). The double-observer data can be regarded as two-sample mark–recapture. However, heterogeneity in detection probabilities generates bias in abundance estimates, just as heterogeneity in capture probabilities generates bias in mark–recapture estimates of abundance. Authors have attempted to minimize this bias, for example, by modeling the effects of covariates (Borchers, Zucchini, and Fewster, 1998; Borchers et al., 1998, 2006; Borchers, 1999; Schweder et al., 1999; Laake and Borchers, 2004), or by assuming independence in the detections of instantaneous cues (such as whale blows) rather than of animals (Schweder et al., 1999; Skaug and Schweder, 1999).

In the absence of any heterogeneity in detection probabilities, we might assume that observer j detects any given animal in the surveyed strip with probability p_j, $j = 1, 2$, and that the probability that both observers detect a given animal is $p_{12} = p_1 p_2$. This is the "full independence" assumption. However, in line transect sampling, we allow detection probability to fall off with distance y from the line so that $p_j = p_j(y)$. Thus it is natural to apply the full independence assumption at each distance from the line, so that for an animal at y, we assume $p_{12}(y) = p_1(y)p_2(y)$.

Laake (1999) introduced the concept of "point independence" to reduce the impact of unmodeled heterogeneity in detection probabilities. Knowledge of the distribution of distances allows the full independence assumption to be weakened, as outlined below. (For the moment, we ignore variables other than distance for simplicity.)

A double-observer line transect survey generates both conventional distance sampling data and mark–recapture data. Under the assumption of uniform animal distribution perpendicular to the transect line (achievable by random line placement or systematic placement with a random start), the shape of the probability density function of observed distances is the same as that of the detection function (Buckland et al., 2001, p. 52–53). The mark–recapture data provide additional information on the shape of the detection function based on an assumption of independence of detection probabilities without any assumption about the distribution of perpendicular

distances of animals. If we retain the assumption of uniform perpendicular distance distribution, discrepancies between the shapes can be interpreted as failure of the assumption of independence between detection probabilities.

We diagnose dependence by (a) modeling the shape of observer j's detection function ($p_j(y)$, $j = 1$, 2) under the uniform perpendicular distance assumption, (b) modeling the conditional probability $p_{j|j'}(y)$ that observer j detects an animal at y, given that observer j' detected it ($j = 1$, 2, $j' = 3 - j$), and (c) modeling the covariance in the observers' detection probabilities as a function of y using a function $\delta(y)$ defined below.

For real data, typically $p_{j|j'}(y)$ does not decline as steeply as $p_j(y)$. Hence the full independence assumption ($p_{j|j'}(y) = p_j(y)$) cannot be made at each distance. The reason for this is that at greater distances, only the most detectable animals tend to be recorded, and those that are detected by one observer are therefore more likely to be detected by the other observer. Laake (1999) argued that heterogeneity is less of a problem on the line, where probability of detection is relatively high, than away from the line, so that assuming independence only on the line should yield less biased estimates of abundance. The idea was further developed by Laake and Borchers (2004) and Borchers et al. (2006).

Although we can anticipate less dependence between detections on the line than at greater distances, unless detection on the line is certain, it seems possible that some dependence remains. In this article, we consider levels of independence, and show that the independence assumption can be weakened further by assuming that, as detection probability tends to unity, dependence tends to zero (i.e., independence). We term this "limiting independence."

We illustrate the methods through analyses of data from a shipboard survey of minke whales in the North Sea and adjacent waters.

2. Methods

Suppose detected animals within a strip extending a distance W either side of the line are recorded. We assume that two observers search independently from the same platform, or from two platforms following the same route at almost the same time. We also assume that duplicate detections can be correctly classified, based on time and location of animals or animal cues, for example.

2.1 *Independence Assumptions*

At the simplest level, we might assume that observer 1 detects animals in this covered strip with probability p_1, while observer 2 independently detects animals with probability p_2. In this "full independence" case, an animal is detected by at least one observer with probability $p_\bullet = p_1 + p_2 - p_1p_2$. A Horvitz–Thompson estimator of N_c, the number of animals in the strip, is thus $\hat{N}_c = \sum \frac{1}{p_\bullet} = \frac{n}{p_\bullet}$ where n is the number of animals detected by at least one observer. Note that $n = n_1 + n_2 - n_{12}$, where n_j is the number of animals detected by observer j, $j = 1$, 2, and n_{12} is the number of animals detected by both observers. If we estimate p_j by $\hat{p}_j = n_{12}/n_{j'}$ for $j = 1$, 2, $j' = 3 - j$, and substitute in, we find that $\hat{N}_c = \frac{n_1 n_2}{n_{12}}$, which is the familiar Petersen estimator. This is the full maximum likelihood estimator of N_c (Borchers, Buckland, and Zucchini, 2002, p. 111), or within a single animal of the

maximum likelihood estimator, if we allow for the fact that N_c is integer.

Now suppose that probability of detection is a function of distance y from the line. There may also be dependence on additional covariates $\underline{z}$, although we omit this dependence below, for clarity. Full independence applied at each y gives $p_\bullet(y) = p_1(y) + p_2(y) - p_1(y)p_2(y)$, so that a model is now needed for $p_j(y)$, $j = 1$, 2. We can then proceed to fit the model, and hence to estimate abundance in the covered strip (below).

We would like to relax the full independence assumption. Allowing some degree of dependence ($\delta(y)$), the independence assumption can be expressed more generally such that $p_{12}(y) = \delta(y)p_1(y)p_2(y)$, $p_\bullet(y) = p_1(y) + p_2(y) - \delta(y)p_1(y)p_2(y)$, and $p_{j|j'}(y) = \delta(y)p_j(y)$, $j = 1$, 2, $j' = 3 - j$. The function $\delta(y)$ is related to the covariance $\sigma_{12}(y)$ between detection probabilities $p_1(y)$ and $p_2(y)$ as follows: $\sigma_{12}(y) = [\delta(y) - 1]p_1(y)p_2(y)$. Various alternative expressions can be derived for $\delta(y)$ including

$$\begin{aligned} \delta(y) &= p_{12}(y)/\{p_1(y)p_2(y)\} \\ &= \{p_{1|2}(y) + p_{2|1}(y) - p_{1|2}(y)p_{2|1}(y)\}/p_\bullet(y) \\ &= p_{j|j'}(y)/p_j(y) \end{aligned}$$

for $j = 1$, 2. The latter expressions demonstrate that $\delta(y)$ measures the discrepancy between the conditional detection functions $p_{j|j'}(y)$ derived from the mark–recapture data and the unconditional detection functions $p_j(y)$, which are derived from distance sampling data with the requirement that $p_j(y^*)$ is known for some y^*. For distance sampling with a single observer, the standard assumption is $p_j(0) = 1$. With double observers, this often untenable assumption can be replaced with the assumption of full independence, $\delta(y) = 1$ for all y, or point independence, $\delta(y^*) = 1$ at a specified y^*, usually $y^* = 0$ (Laake and Borchers, 2004). Fitting full independence models to data requires a functional form for $p_j(y)$ and point independence requires the same and a model for $p_{j|j'}(y)$. Neither require a model for $\delta(y)$.

We now relax the assumption that $\delta(y^*) = 1$ at a specified y^*. Instead we assume that we achieve independence in the limit as detection probability tends to one. This requires a model for $\delta(y)$ with the following properties to ensure valid probabilities:

(1) $\delta(y) \leq U(y)$,   where $U(y) = \min\{1/p_1(y), 1/p_2(y)\}$, which ensures that $p_{j|j'}(y) \leq 1$.

(2) $\delta(y) \geq L(y)$, where $L(y) = \max\{0, \frac{p_1(y)+p_2(y)-1}{p_1(y)p_2(y)}\}$, which ensures that $p_\bullet(y) \leq 1$.

If we define $\delta_0(y) = \{\delta(y) - L(y)\}/\{U(y) - L(y)\}$, it is restricted to the unit interval and can be represented by an appropriate functional form such as a logistic. Note also that as $p_j(y) \to 1$, $j = 1$, 2 then $\delta(y) \to 1$.

Using a logistic formulation for $\delta_0(y)$, we can write $\log_e\{\frac{\delta_0(y)}{1-\delta_0(y)}\}$ as some linear function of y. Full and point independence can be derived as special cases of the limiting independence model if we include the following offset $\log_e\{\frac{1-L(y)}{U(y)-1}\}$ in $\delta_0(y)$ which fixes $\delta(y) = 1$. If we consider the following logistic model for limiting independence:

$$\log_e\left\{\frac{\delta_0(y)}{1 - \delta_0(y)}\right\} = \alpha + \beta y + \log_e\left\{\frac{1 - L(y)}{U(y) - 1}\right\}, \quad (1)$$

then $\alpha = 0$ specifies point independence at $y^* = 0$, and $\alpha = \beta = 0$ specifies a full independence model. If $\beta = 0$ and $\alpha \neq 0$, a model with constant dependence for all y can be specified. Models restricted to independence or positive dependence can be achieved by restricting $\alpha \geq 0$, $\beta \geq 0$. Hence this general formulation provides a model selection framework for a range of models with varying degrees of independence.

2.2 *Likelihood*

The full likelihood for double-platform data may be expressed as $\mathcal{L} = \mathcal{L}_n \mathcal{L}_z \mathcal{L}_{y|z} \mathcal{L}_\omega$ where $\mathcal{L}_n$ is the component accounting for variation in total number of animals n detected by at least one observer, $\mathcal{L}_z$ corresponds to any observation-specific covariates $\underline{z}$, $\mathcal{L}_{y|z}$ corresponds to the conditional distribution of distances y, given covariates $\underline{z}$, and $\mathcal{L}_\omega$ corresponds to the mark–recapture data (Laake and Borchers, 2004). $\mathcal{L}_{y|z}$ incorporates the assumption of uniform distribution of animals perpendicular to transect lines. We use just two components of the full likelihood: $\mathcal{L}_{y|z}$ and $\mathcal{L}_\omega$. By doing this, we can avoid making distributional assumptions about n and $\underline{z}$, as estimation is not robust to failure of such assumptions. Instead, we draw inference conditional on n and $\underline{z}$, and use a design-based approach to allow for variation in n. If there are no covariates $\underline{z}$, the full likelihood is $\mathcal{L} = \mathcal{L}_n \mathcal{L}_y \mathcal{L}_\omega$, and we use the second and third components only (in this case, $\mathcal{L}_y$ incorporates the assumption of uniform distribution of animals perpendicular to transect lines). Again for simplicity we consider this latter case; the extensions to include covariates $\underline{z}$ are straightforward.

We have

$$\mathcal{L}_y = \prod_{i=1}^n f_\bullet(y_i) = \prod_{i=1}^n \frac{p_\bullet(y_i)\pi(y_i)}{E(p_\bullet)},$$

where $f_\bullet(y_i)$ is the probability density function (pdf) of detection distances y of animals detected by at least one observer, evaluated at y_i, $p_\bullet(y_i) = p_1(y_i) + p_2(y_i) - \delta(y_i)p_1(y_i)p_2(y_i)$ is the probability that an animal at distance y_i from the line is detected by at least one observer, $\pi(y_i)$ is the unconditional pdf of distances y in the population (whether detected or not), evaluated at y_i, and $E(p_\bullet) = \int_0^w p_\bullet(y)\pi(y)dy$ (Laake and Borchers, 2004, p. 114). Random positioning of the lines (or of a systematic grid of lines) ensures that $\pi(y) = 1/W$.

We also need

$$\mathcal{L}_\omega = \prod_{i=1}^n \frac{\Pr(\underline{\omega}_i \mid y_i)}{p_\bullet(y_i)},$$

where

$$\Pr\{\underline{\omega}_i = (1,0) \mid y_i\} = p_1(y_i)\{1 - p_2(y_i)\delta(y_i)\},$$
$$\Pr\{\underline{\omega}_i = (0,1) \mid y_i\} = p_2(y_i)\{1 - p_1(y_i)\delta(y_i)\},$$
$$\Pr\{\underline{\omega}_i = (1,1) \mid y_i\} = p_1(y_i)p_2(y_i)\delta(y_i).$$

The likelihoods for full, point, and limiting independence only differ in the definition of $\delta(y_i)$. However, if the full independence assumption holds then it is only necessary to use $\mathcal{L}_\omega$ (Borchers et al., 1998) and with the point independence assumption, $\mathcal{L}_\omega$ and $\mathcal{L}_y$ can be maximized independently using models for $p_{j|j'}(y)$ and $p_j(y)$, which separate

into the two respective likelihood components (Borchers et al., 2006). When the likelihood is specified in terms of models for $p_j(y)$ and $\delta(y)$, both components of the likelihood must be maximized jointly.

We assume logistic forms for the detection functions:

$$p_j(y) = \frac{\exp(\lambda_{0j} + \lambda_{1j}y)}{1 + \exp(\lambda_{0j} + \lambda_{1j}y)} \quad \text{for } j = 1 \text{ or } 2. \qquad (2)$$

2.3 *Diagnostic for Reliable Estimation under Limiting Independence*

When fitting limiting independence models, the Hessian matrix is sometimes nearly singular, due to high correlation between the estimates of $p_j(y)$ and $\delta(y)$ at $y = 0$. In these cases, the models are unstable, typically yielding very large abundance estimates and associated variances. We can still usefully calculate Akaike's information criterion (AIC), but if AIC indicates that a limiting independence model is required, then reliable estimation is not possible. To identify such cases, the following diagnostic check was found useful. If the magnitude of the estimated correlation between $\hat{\alpha}$ of equation (1) and $\hat{\lambda}_{0j}$ of equation (2) is found to be large, then estimated abundance should be considered unreliable. The test can be conducted for each of $j = 1$ and $j = 2$, or by arbitrarily choosing one of the two; the two correlations tend to be similar when they are close to ± 1. We defined "large" to be greater than 0.99 in Section 3 and 0.9 in Section 4; choices in the range of 0.9 to 0.99 were found to be effective. Lowering the correlation criterion provides a more conservative approach to avoid overestimation with the only cost being potential underestimation due to the unmodeled dependence.

2.4 *Estimating Abundance*

Given models for $p_1(y)$, $p_2(y)$, and $\delta(y)$, the likelihood conditional on n, $\mathcal{L}_y \mathcal{L}_\omega$, can be maximized, which allows us to estimate $E(p_\bullet)$. Estimated abundance in the covered area is then

$$\hat{N}_c = \sum_{i=1}^n \frac{1}{\hat{E}(p_\bullet)} = \frac{n}{\hat{E}(p_\bullet)}. \qquad (3)$$

This is a Horvitz–Thompson estimator in which the inclusion probabilities have been estimated (Laake and Borchers, 2004, p. 116). When covariates $\underline{z}$ are present, the simplification represented by the second equality does not hold. If the covered area is of size a, and the entire survey region of size A, then estimated abundance in the survey region is

$$\hat{N} = \frac{A}{a}\hat{N}_c = \frac{A}{a}\frac{n}{\hat{E}(p_\bullet)}, \qquad (4)$$

where $a = 2wL$ and L is the total length of transect line.

For our limiting independence model, we cannot use $\text{vâr}(\hat{N}_c)$ as defined in Borchers et al. (2006), because the conditional and unconditional detection functions share parameters under the above formulation. Adapting their result, we have

$$\text{vâr}(\hat{N}_c) = S^2(\hat{\underline{\theta}}) + \hat{\underline{d}}^T \hat{\mathbf{I}}^{-1} \hat{\underline{d}},$$

where $S^2(\hat{\underline{\theta}}) = \sum_{i=1}^n \frac{1-\hat{E}(p_\bullet)}{\{\hat{E}(p_\bullet)\}^2} = \frac{n\{1-\hat{E}(p_\bullet)\}}{\{\hat{E}(p_\bullet)\}^2}$, $\hat{\underline{d}} = \frac{d\hat{N}_c}{d\theta}|_{\hat{\theta}}$, and $-\hat{\mathbf{I}}$ is the matrix of second derivatives of $\ln(\mathcal{L}_y) + \ln(\mathcal{L}_\omega)$, evaluated at $\hat{\underline{\theta}}$, the vector of parameter estimates.

Adapting equation (11) of Marques and Buckland (2003),

$$\mathrm{v\hat{a}r}(\hat{N}) = \left(\frac{A}{a}\right)^2 \left\{ L \sum_{k=1}^K \frac{l_k(\hat{N}_{ck}/l_k - \hat{N}_c/L)^2}{K-1} + \hat{\underline{d}}^T \hat{\mathbf{I}}^{-1} \hat{\underline{d}} \right\}, \quad (5)$$

where $\hat{N}_{ck} = \sum_{i=1}^{n_k} \frac{1}{\hat{E}(p_\bullet)} = \frac{n_k}{\hat{E}(p_\bullet)}$ is estimated abundance for strip k, which has half-width w and length l_k, where $\sum_{k=1}^K l_k = L$.

An alternative to the above is to use the bootstrap, in which bootstrap resamples are generated by sampling the lines with replacement.

If animals occur in clusters, with s_i animals in the ith detected cluster, then the above formula gives estimated cluster abundance, and estimated animal abundance is given by

$$\hat{N} = \frac{A}{a} \frac{\sum_{i=1}^n s_i}{\hat{E}(p_\bullet)}.$$

Variance can be estimated as before, except that now, $S^2(\hat{\underline{\theta}}) = \frac{\{1-\hat{E}(p_\bullet)\}}{\{\hat{E}(p_\bullet)\}^2} \sum_{i=1}^n s_i^2$, $\hat{\underline{d}} = \frac{d\hat{N}_c}{d\underline{\theta}}|_{\hat{\theta}}$ is evaluated using $\hat{N}_c = \frac{\sum_{i=1}^n s_i}{\hat{E}(p_\bullet)}$, and in the formula for variance of $\hat{N}$, $\hat{N}_c = \frac{\sum_{i=1}^n s_i}{\hat{E}(p_\bullet)}$ and $\hat{N}_{ck} = \frac{\sum_{i=1}^{n_k} s_i}{\hat{E}(p_\bullet)}$.

3. Simulation Study

Simulations were conducted to evaluate the performance of the limiting independence model. We simulated a population of $N = 1000$ animals that were uniformly distributed in a strip of width two ($w = 1$) and undefined length. For each of 100 simulation replicates, we generated capture histories for two observers with identical detection probability functions $p_1(y) = p_2(y)$. We used four different logistic models for $p_j(y)$ and two different logistic models for $\delta_0(y)$ to create eight scenarios. For models with a covariate z, the covariate value was generated from a uniform $(0, 1)$ distribution. We fitted the simulated observed data (10, 01, 11 capture histories) with the model that generated the data, and with the equivalent models under the point independence and full independence restrictions. We computed the AIC for each of the fitted models. For model fits where the magnitude of the correlation between $\hat{\alpha}$ and $\hat{\lambda}_{0j}$ exceeded 0.99, results are not reported.

The eight scenarios were as follows. The offset $\log_e\{\frac{1-L(y)}{U(y)-1}\}$ was used in each dependence model to simulate and fit the data. The dependence model $\delta_0(y) = (1 + e^{-1-y})^{-1}$ was used in scenarios 1, 3, 5, and 7, whereas $\delta_0(y) = (1 + e^{-2-2y})^{-1}$ (representing stronger dependence) was used in scenarios 2, 4, 6, and 8. The detection probability model $p_j(y) = (1 + e^{-1.1+3y})^{-1}$, $j = 1$, 2, was used in scenarios 1 and 2, $p_j(y) = (1 + e^{3y})^{-1}$ in scenarios 3 and 4, $p_j(y, \underline{z}) = (1 + e^{0.8417+3y-0.8417z})^{-1}$ in scenarios 5 and 6, and $p_j(y, \underline{z}) = (1 + e^{3y-5z})^{-1}$ in scenarios 7 and 8.

Simulation results appear in Table 1. Full independence and point independence models had substantial negative bias

Table 1

*Mean (standard error in parentheses) of 100 abundance estimates under full independence (FI), point independence (PI), and limiting independence (LI) models for the eight simulation scenarios. The expected capture history frequencies are shown for each scenario. Also shown is p_{AIC}, the proportion of times each model was selected by AIC, and model-averaged (MA) estimates, obtained by taking a weighted average of estimates from the above three models, using AIC weights. Where an LI model was deemed to be parameter redundant ($|\mathrm{correl}(\hat{\alpha}, \hat{\lambda}_{01})| > 0.99$), the model was not considered even if it had the best AIC value, and the weighted average was over the FI and PI models only. The mean and standard error for LI models is across only those runs for which the model was not deemed to be parameter-redundant. The number of runs (Nr) out of 100 contributing to the LI results under each scenario is shown. True abundance is 1000. * Bias significant at 5% level.*

Scenario	Exp. capture history freqs. 10	01	11	FI Mean	p_{AIC}	PI Mean	p_{AIC}	LI Mean	p_{AIC}	Nr	MA Mean
1	101	101	315	607* (3)	0.00	819* (6)	0.23	1074* (14)	0.77	95	1020 (16)
2	33	33	383	450* (1)	0.00	692* (4)	0.04	1037* (18)	0.96	97	1020 (18)
3	89	89	126	499* (4)	0.00	616* (6)	0.93	971 (21)	0.07	60	685* (7)
4	34	34	180	290* (3)	0.00	454* (4)	0.54	963 (21)	0.46	63	664* (21)
5	87	87	135	512* (4)	0.00	650* (6)	0.65	1437* (124)	0.35	76	1021 (96)
6	34	34	189	296* (3)	0.00	465* (5)	0.23	1149* (42)	0.77	86	962 (41)
7	77	77	579	770* (2)	0.00	915* (3)	0.55	1064* (15)	0.45	82	1001 (14)
8	29	29	629	689* (2)	0.00	850* (4)	0.12	1057* (11)	0.88	94	1032* (12)
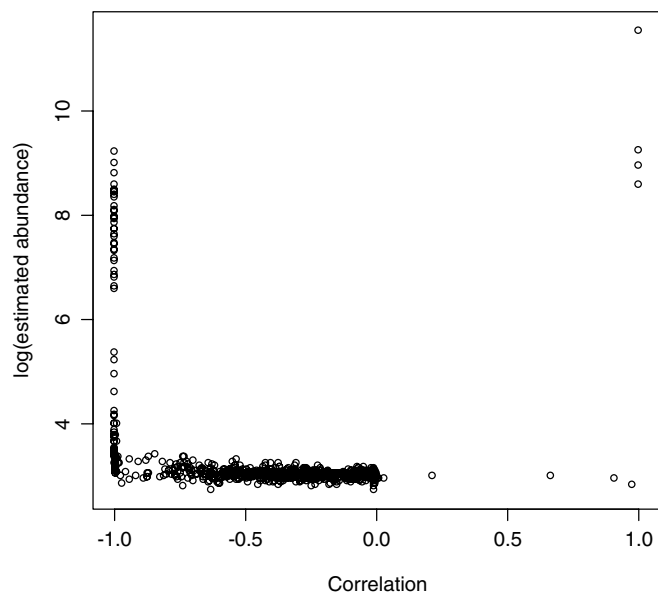
Figure 1. Plot of $\log(\hat{N})$ against $\mathrm{correl}(\hat{\alpha}, \hat{\lambda}_{01})$, illustrating that the very high estimates of abundance from the simulations all arose when $\mathrm{correl}(\hat{\alpha}, \hat{\lambda}_{01})$ was very close to ± 1.

in all scenarios, with full independence models consistently more biased than point independence models. Within a scenario, the bias was remarkably consistent, reflected in the very small standard errors of Table 1, but the bias varied substantially between scenarios. Although the data were simulated from limiting independence models, significant upward bias was found in six of the eight scenarios when the data were analyzed using the true model. However, the size of the bias in most cases was substantially smaller than for point independence models. Model-averaged estimates had low bias, except for scenarios 3 and 4, for which around 40% of analyses under the limiting independence model were rejected due to high correlation between $\hat{\alpha}$ and $\hat{\lambda}_{0j}$. Figure 1 shows the need to reject such cases; all of the very high abundance estimates obtained correspond to a correlation between $\hat{\alpha}$ and $\hat{\lambda}_{0j}$ very close to ± 1. In all cases use of AIC correctly diagnosed the presence of unmodeled heterogeneity, although in some cases it did not differentiate well between point independence and limiting independence scenarios.

4. Stake Data

Laake (1999) used data on a population of wooden stakes of known size to illustrate independence issues in double-observer surveys. We use the same dataset here. The surveys were conducted in 1977 and 1978 (Laake, 1978); as in Laake (1999), we consider only the 1977 data. Multiple observers traversed a 1 km line marked with poles at 100-m intervals and searched a strip of sagebrush-grassland habitat 20 m on either side of the line for 150 wooden stakes that protruded 30 cm above ground. The stakes had a random uniform distribution throughout the 1000 m × 40 m strip. Eight observers separately surveyed the stakes, remaining on the line. Distances from the line were measured accurately by an assistant.

For each pair of observers, we show estimates of abundance in Table 2. Models were fitted corresponding to full independence ($\alpha = \beta = 0$), point independence ($\alpha = 0$, β unconstrained), and limiting independence with $\alpha \geq 0$, $\beta \geq 0$. In each case, three models were fitted: the first with observer as a factor and distance as a covariate, the second with the addition of an interaction term between the two, and the third with the squared distance as an additional covariate, together with interaction terms between observer and the two continuous covariates.

It is clear from Table 2 that models with all three forms of independence are useful for the analysis of these data. Overall performance is remarkably good, with the average of the best estimates (as judged by AIC) coming out close to the true abundance of $N = 150$, as does the average of the model-averaged estimates, using AIC weights (Buckland, Burnham and Augustin 1997).

5. Shipboard Survey of Minke Whales

The second Small Cetacean Abundance in the North Sea and adjacent waters (SCANS II) survey was a multinational survey conducted in 2005 by ship and aircraft to estimate cetacean abundance in the North Sea, Kattegat, Skagerrak, western Baltic, English Channel, and the Celtic Sea. Double-observer line transect survey methods were used because for many species detection of animals on the trackline was expected to be less than unity. Details of the survey and further information can be found at `http://biology.st-andrews.ac.uk/scans2/`. Here we analyze only shipboard survey data on minke whales.

The methods used in the SCANS surveys were designed to break up the dependence between the two observers, by ensuring that they are not simultaneously searching the same patch of sea. A "tracker" scans with high-powered binoculars well ahead of the ship, and tracks detected animals in, to check whether the primary platform, searching with hand-held binoculars and naked eye, detects them (so-called duplicate detections). Previously, we have had no means of testing whether the method is successful in breaking up the dependence between observers.

Using a truncation distance of 700 m, the tracker detected 54 minke groups totaling 62 animals, while the primary platform detected 57 groups totaling 59 animals; 17 groups (19 animals) were detected by both tracker and primary platform.

The full model of Table 3 is defined by

$$p_j(y, z) = \frac{\exp(\lambda_{0j} + \lambda_{1j}y + \lambda_{2j}z)}{1 + \exp(\lambda_{0j} + \lambda_{1j}y + \lambda_{2j}z)}, \qquad (6)$$

$$\delta(y, z) = L(y, z) + \delta_0(y, z)\{U(y, z) - L(y, z)\}, \qquad (7)$$

where

$$L(y, z) = \max\left\{0, \frac{p_1(y, z) + p_2(y, z) - 1}{p_1(y, z)p_2(y, z)}\right\},$$

$$U(y, z) = \min\{1/p_1(y, z), 1/p_2(y, z)\},$$

and

$$\delta_0(y, z) = \frac{\{1 - L(y, z)\}\exp(\alpha + \beta y)}{\{U(y, z) - 1\} + \{1 - L(y, z)\}\exp(\alpha + \beta y)}.$$

Covariate z is sea state (Beaufort).

Table 2
Estimates of abundance for the stake data for all combinations of the eight observers under full independence, FI ($\alpha = \beta = 0$), point independence, PI ($\alpha = 0$, β unconstrained) and limiting independence, LI with $\alpha \geq 0$, $\beta \geq 0$. Covariate model 1 has covariate structure observer + distance, *model 2 has structure* observer * distance, *and model 3 has structure* observer * (distance + distance^2). "*Best*" *corresponds to the model with the smallest AIC (indicated by "#," with standard error in parentheses), and "MA" is the model-averaged estimate, using AIC weights. True abundance is* 150. *LI fits with* $|\mathrm{correl}(\hat{\alpha}, \hat{\lambda}_{01})| > 0.9$ *were not used for the best or model-averaged estimates, and are indicated by "!".*

					Cov model 1			Cov model 2			Cov model 3					
Obs1	Obs2	n_1	n_2	n_{12}	FI	PI	LI	FI	PI	LI	FI	PI	LI	Best	(se)	MA
1	2	81	48	38	136	139	156	132#	134	136	133	138	137	132	(11)	134
1	3	81	68	60	125	146	160	125	154	195#	120	145	193	195	(23)	185
1	4	81	51	42	127	158	215#	126	149	176	123	146	196	215	(23)	172
1	5	81	49	43	127	141#	154	126	136	141!	128	140	142	141	(16)	136
1	6	81	72	58	126#	134	159	126	134	147	125	135	148	126	(10)	133
1	7	81	54	47	120	135	204	118	124	138	122#	125	127	122	(10)	126
1	8	81	84	63	136#	142	156	136	142	155	136	141	157!	136	(10)	140
2	3	48	68	34	147	154	254	141#	150	166	142	154	181	141	(14)	150
2	4	48	51	35	86	131#	140	86	130	139	84	131	139	131	(18)	133
2	5	48	49	38	78	117	124	77	132#	135	75	143	143	132	(32)	135
2	6	48	72	41	105	149	239	105	142#	144	105	153	161	142	(18)	162
2	7	48	54	36	92	130#	145	92	129	137	88	127	138	130	(17)	131
2	8	48	84	46	107	144	197	105	129	141	105	126#	131	126	(11)	132
3	4	68	51	40	129	157	276	124	152	175#	124	148	164	175	(22)	163
3	5	68	49	38	143	164	246	141	156	203#	140	155	220	203	(46)	181
3	6	68	72	55	122	139	177	122	144	166#	122	150	164	166	(17)	155
3	7	68	54	42	130	166	294	127	157	213#	124	143	185	213	(27)	193
3	8	68	84	58	136	147	151!	137	153#	172	135	153	171	153	(16)	153
4	5	51	49	33	105	156	167	105	157#	168	101	156	185	157	(17)	160
4	6	51	72	46	91	140	351!	93	130	178	94	137#	138	137	(15)	144
4	7	51	54	36	99	137#	148	99	137	160	97	133	152	137	(18)	140
4	8	51	84	48	107	141	222	106	124	145#	109	124	130	145	(29)	148
5	6	49	72	41	115	167#	170	115	176	179	114	177	181	167	(27)	171
5	7	49	54	39	91	142#	147	91	143	150	87	143	146	142	(23)	144
5	8	49	84	46	117	148#	163	117	142	156	115	141	158	148	(18)	151
6	7	72	54	48	96	132	214#	97	126	130	99	120	132	214	(39)	164
6	8	72	84	66	106	134	159	106	132#	134	106	133	139	132	(12)	137
7	8	54	84	53	86	152	253#	97	119	96	101	108	108	253	(129)	212
Average					114	144	194	113	140	160	113	140	156	158		153
Std dev.					19	12	55	18	13	23	18	14	25	34		21

The benefits of field methods to break up heterogeneity are immediately apparent from Table 3. AIC favors models with full independence ($\alpha = \beta = 0$), and selects the model with identical detection functions for the two observers, and sea state as a covariate (model 12, Figure 2). Estimation is largely unaffected by whether we assume full independence or point independence. If we also relax the assumption of point independence, AIC values are larger, but estimation is not greatly affected, with the exception of model 2.

It is surprising that AIC favors models which assume the same detection function for the two observers, given that the tracker is searching much further ahead of the ship than the primary platform. However, estimation is barely affected by whether we make this assumption or not. The distribution of distances from the line of detections from the two platforms is clearly very similar out to the truncation distance of 700 m (Figure 2), although beyond this distance, the tracker detects more animals than the primary platform.

6. Discussion

Our methods allow assessment of whether the full independence or point independence assumptions are reasonable. The methods also provide a means of analyzing double-observer surveys without having to assume independence between the observers' detection probabilities, even at distance zero. However, strong dependence between the observers' detection probabilities can lead to unreliable estimation. If possible, field methods should be developed to ensure that there are not animals in the population that are very unlikely to be detected, even if they are on the line. However, this strategy can create problems for identifying duplicate detections, so that in some circumstances, it may be preferable to estimate the proportion of animals that are essentially undetectable. For example in aerial surveys of marine mammals, the observers might record only those animals that are at the surface as they pass abeam, and a separate study might be used to estimate the proportion of animals at the surface at any time.

Table 3

Models fitted to the minke whale survey data. The full model, denoted here by LI (limiting independence), platform ∗ (Dist+Beau), is defined by equations (6) and (7). The estimates $\hat{N}$ and $se(\hat{N})$ were obtained by appropriate extensions of equations (3) and (4) for a stratified design. Correl for LI models with a platform effect corresponds to whichever of $|\mathrm{correl}(\hat{\alpha}, \hat{\lambda}_{01})|$ and $|\mathrm{correl}(\hat{\alpha}, \hat{\lambda}_{02})|$ is closest to 1. PI indicates point independence ($\alpha = 0$) and FI denotes Full Independence ($\alpha = \beta = 0$).

Nos.	Dependence	Covariate model	No. par	$\hat{N}$	$se(\hat{N})$	ΔAIC	Correl
1	LI	Platform ∗ (Dist+Beau)	8	16,912	6663	7.4	0.393
2	LI ($\beta = 0$)	Platform ∗ (Dist+Beau)	7	26,453	18,422	7.0	-0.805
3	PI	Platform ∗ (Dist+Beau)	7	17,156	5359	5.4	–
4	FI	Platform ∗ (Dist+Beau)	6	17,935	5436	3.6	–
5	LI	Platform ∗ (Dist)	6	∗	∗	13.6	-1.000
6	LI ($\beta = 0$)	Platform ∗ (Dist)	5	∗	∗	11.7	-1.000
7	PI	Platform ∗ (Dist)	5	16,192	4446	12.3	–
8	FI	Platform ∗ (Dist)	4	15,195	3311	10.4	–
9	LI	Dist+Beau	5	16,837	7130	3.9	0.073
10	LI ($\beta = 0$)	Dist+Beau	4	16,091	4365	1.9	0.124
11	PI	Dist+Beau	4	17,337	5647	1.9	–
12	FI	Dist+Beau	3	18,173	5534	0.0	–
13	LI	Dist	4	∗	∗	9.8	1.000
14	LI ($\beta = 0$)	Dist	3	∗	∗	7.9	-1.000
15	PI	Dist	3	16,209	4447	8.5	–
16	FI	Dist	2	15,202	3313	6.6	–

∗$|\mathrm{correl}(\hat{\alpha}, \hat{\lambda}_{0j})| > 0.99$ for $j = 1$ and 2 (models 1–8) or $|\mathrm{correl}(\hat{\alpha}, \hat{\lambda}_{0})| > 0.99$ (models 9–16).

Extension of the methods to point transect sampling is straightforward. We now have $\pi(y) = 2y/w^2$ if points are positioned randomly. In equation (4), the covered area a is now $a = K\pi w^2$, where K is the number of points. For vâr($\hat{N}$), we obtain a similar result to equation (5) by adapting equation (3.48) from Marques and Buckland (2004).

It is knowledge of $\pi(y)$ that allows us to weaken the full independence assumption. In principle, the same approach could be applied to conventional mark–recapture models, if $\pi(y)$ were known for some explanatory variable y, although this seems unlikely in most applications.

If there is responsive movement prior to detection so that the distances y available for detection differ in an unknown way from that prior to movement (i.e., from $\pi(y)$), then (a) $\delta(y)$ cannot be interpreted as above and (b) under the assumption of full independence, $\delta(y)$ can be interpreted as a measure of deviation from $\pi(y)$ due to responsive movement. This can be seen from the following. The pdf of observed distance for observer j is $f_j(y) = p_{j|j'}(y)\delta(y)^{-1}\pi(y)/\int p_{j|j'}(y)\delta(y)^{-1}\pi(y)dy$. If we assume full independence, then $p_{j|j'}(y)$ is equal to the unconditional detection function for observer j, and hence $\pi(y)\delta(y)^{-1}$ is proportional to the pdf of y after movement. Note that while the interpretation of $\delta(y)$ is different in this case from the case with no responsive movement, the abundance estimator is still valid.

It is worth noting that although $\delta(y)$ is superficially similar to $\alpha(\underline{z})$ of Chen (1999) and α of Chen and Lloyd (2000), it is in fact quite different. To see this, consider a situation in which distance y and other variables $\underline{u}$ affect detection probability but only y is recorded (in this case $\alpha(\underline{z})$ is a constant). Whereas $\alpha(\underline{z})$ and α quantify the heterogeneity due to y, $\delta(y)$ quantifies the heterogeneity at y due to the unrecorded variables $\underline{u}$. This is an important difference because the formulations of Chen (1999) and Chen and Lloyd (2000) do not accommodate heterogeneity due to the unrecorded variables $\underline{u}$, and it is precisely this heterogeneity that is at issue here. Chen (1999) and Chen and Lloyd (2000) assume that $p_{\bullet}(y) = p_1(y) + p_2(y) - p_1(y)p_2(y)$, whereas we assume that $p_{\bullet}(y) = p_1(y) + p_2(y) - \delta(y)p_1(y)p_2(y)$. It is the discrepancy between the shapes of $f_j(y)$ and $p_{j|j'}(y)$ ($j = 1,2$) that provides the basis for modeling heterogeneity due to the unrecorded variables (together with knowledge of $\pi(y)$); the formulations of Chen (1999) and Chen and Lloyd (2000) model $f_j(y)$ but do not include $p_{j|j'}(y)$ and are therefore unable to exploit the information in the discrepancy between the two. This applies equally to the case in which additional variables $\underline{z}$ are recorded (but $\underline{u}$ remains unrecorded).

We have used AIC to select between models. We have estimated detection functions by maximum likelihood, but abundance is estimated using a Horvitz–Thompson estimator in which the inclusion probabilities have been estimated (by maximum likelihood). As the components of the abundance estimators that are not estimated by maximum likelihood (corresponding to sample size, and to extrapolation from the covered area to the entire survey area) are common across models, it seems not unreasonable to use AIC to select between abundance estimators. However, when some inclusion probabilities are very small, modest error in estimating them can generate large positive bias in the abundance estimate, which would be undetectable by AIC.
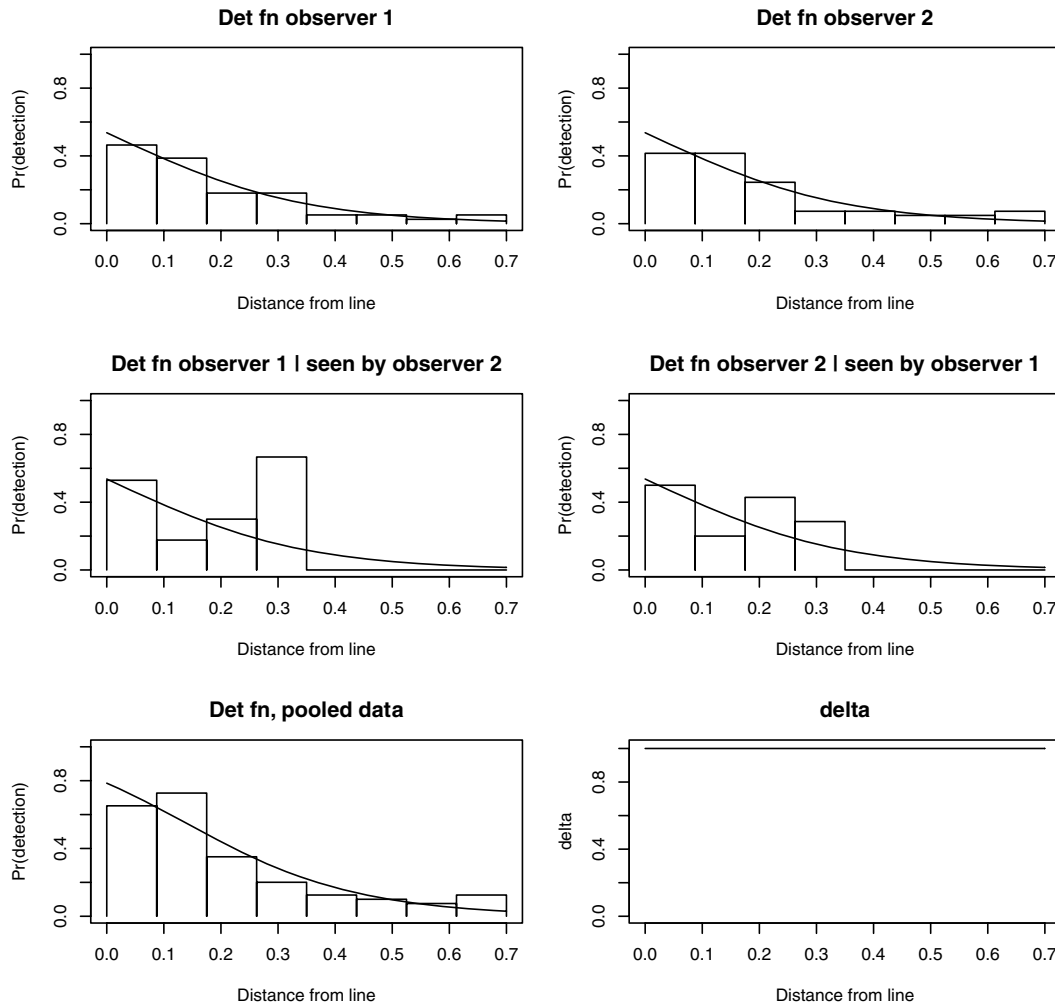
Figure 2. Estimated detection functions for minke whales, model 12 (see Table 3). The top left plot is the estimated unconditional detection function for observer 1, and top right is the estimated unconditional detection function for observer 2. The corresponding conditional detection functions are shown in the center. Under this model, all four of these detection functions are identical, but the data in each plot differ. The estimated detection function for the two observers combined is shown at the bottom left. The bars are: relative frequencies of detections made by observer 1 (top left), relative frequencies of detections made by observer 2 (top right), proportion of observer 2 detections made by observer 1 (middle left), proportion of observer 1 detections made by observer 2 (middle right), and relative frequencies of detections made by at least one observer.

REFERENCES

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

Borchers, D. L. (1999). Composite mark-recapture line transect surveys. In *Marine Mammal Survey and Assessment Methods*, G. W. Garner, S. C. Amstrup, J. L. Laake, B. F. J. Manly, L. L. McDonald, and D. G. Robertson (eds), pp. 115–126. Rotterdam, Netherlands: Balkema.

Borchers, D. L., Buckland, S. T., Goedhart, P. W., Clarke, E. D., and Hedley, S. L. (1998). Horvitz-Thompson estimators for double-platform line transect surveys. *Biometrics* **54,** 1221–1237.

Borchers, D. L., Buckland, S. T., and Zucchini, W. (2002). *Estimating Animal Abundance: Closed Populations*. London: Springer-Verlag.

Borchers, D. L., Laake, J. L., Southwell, C., and Paxton, C. G. M. (2006). Accommodating unmodeled heterogeneity in double-observer distance sampling surveys. *Biometrics* **62,** 372–378.

Borchers, D. L., Zucchini, W., and Fewster, R. M. (1998). Mark-recapture models for line transect surveys. *Biometrics* **54,** 1207–1220.

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. (2001). *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*. Oxford: Oxford University Press.

Buckland, S. T., Burnham, K. P., and Augustin, N. H. (1997). Model selection: An integral part of inference. *Biometrics* **53,** 603–618.

Buckland, S. T. and Turnock, B. J. (1992). A robust line transect method. *Biometrics* **48,** 901–909.

Chen, S. X. (1999). Estimation in independent observer line transect surveys for cluster populations. *Biometrics* **55,** 754–759.

Chen, S. X. (2000). Animal abundance estimation in independent observer line transect surveys. *Environmental and Ecological Statistics* **7,** 285–299.

Chen, S. X. and Lloyd, C. J. (2000). A nonparametric approach to the analysis of two-stage mark-recapture experiments. *Biometrika* **87,** 633–649.

Innes, S., Heide-Jørgensen, M. P., Laake, J. L., Laidre, K. L., Cleator, H., Richard, P., and Stewart, R. E. A. (2002). Surveys of belugas and narwhals in the Canadian high Arctic in 1996. *Scientific Publications of the North Atlantic Marine Mammal Commission* **4,** 169–190.

Laake, J. L. (1978). Line transect estimators robust to animal movement. M.S. Thesis, Utah State University, Logan, Utah.

Laake, J. L. (1999). Distance sampling with independent observers: Reducing bias from heterogeneity by weakening the conditional independence assumption. In *Marine Mammal Survey and Assessment Methods*, G. W. Garner, S. C. Amstrup, J. L. Laake, B. F. J. Manly, L. L. McDonald, and D. G. Robertson (eds), pp. 137–148. Rotterdam, Netherlands: Balkema.

Laake, J. L. and Borchers, D. L. (2004). Methods for incomplete detection at distance zero. In *Advanced Distance Sampling*, S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (eds), pp. 108–189. Oxford: Oxford University Press.

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.

Marques, F. F. C. and Buckland, S. T. (2003). Incorporating covariates into standard line transect analyses. *Biometrics* **59,** 924–935.

Marques, F. F. C. and Buckland, S. T. (2004). Covariate models for the detection function. In *Advanced Distance Sampling*, S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (eds), pp. 31–47. Oxford: Oxford University Press.

Palka, D. (1995). Abundance estimate of the Gulf of Maine harbor porpoise. In *Biology of the Phocoenids*, A. Bjørge and G.P. Donovan (eds), 27–50. Cambridge, U.K.: International Whaling Commission.

Quang, P. X. and Becker, E. F. (1997). Combining line transect and double count sampling techniques for aerial surveys. *Journal of Agricultural, Biological, and Environmental Statistics* **1,** 170–189.

Schweder, T., Skaug, H. J., Langaas, M., and Dimakos, X. K. (1999). Simulated likelihood methods for complex double-platform line transect surveys. *Biometrics* **55,** 678–687.

Skaug, H. J. and Schweder, T. (1999). Hazard models for line transect surveys with independent observers. *Biometrics* **55,** 29–36.