Using Hidden Markov Models to Deal with Availability Bias on Line Transect Surveys

D. L. Borchers,[1],* W. Zucchini,[2] M. P. Heide-Jørgensen,[3] A. Cañadas,[4] R. Langrock[1]

[1]Centre for Research into Ecological and Environmental Modelling, The Observatory, Buchanan Gardens,
University of St Andrews, Fife KY16 9LZ, Scotland
[2]Department of Economic Sciences, Georg-August-University Göttingen, Platz der Göttinger Sieben 3, 37073
Göttingen, Germany
[3]Greenland Institute of Natural Resources, Box 570, DK-3900 Nuuk, Greenland
[4]Alnilam Research and Conservation, Candamo 116, 28240 Hoyo de Manzanares, Madrid
*email: dlb@st-andrews.ac.uk

SUMMARY. We develop estimators for line transect surveys of animals that are stochastically unavailable for detection while within detection range. The detection process is formulated as a hidden Markov model with a binary state-dependent observation model that depends on both perpendicular and forward distances. This provides a parametric method of dealing with availability bias when estimates of availability process parameters are available even if series of availability events themselves are not. We apply the estimators to an aerial and a shipboard survey of whales, and investigate their properties by simulation. They are shown to be more general and more flexible than existing estimators based on parametric models of the availability process. We also find that methods using availability correction factors can be very biased when surveys are not close to being instantaneous, as can estimators that assume temporal independence in availability when there is temporal dependence.

KEY WORDS: Availability bias; Detection hazard; Hidden Markov model; Line transect; Wildlife survey.

1. Introduction

Density estimates from line transect (LT) surveys of animals that are not always available for detection while within detectable range are subject to negative bias if animal availability is not accounted for. This is often referred to as "availability bias" (McLaren, 1961; Marsh and Sinclair, 1989). While it may be possible to deal with availability bias using mark-recapture LT (MRLT) methods (see Laake and Borchers, 2004), this does not specifically account for animal availability and is not effective for animals that are unavailable for long periods relative to their time in view. MRLT survey methods are also more difficult to implement than single-observer methods.

A common way of dealing with availability bias on single-observer surveys is to "correct" conventional line transect (CLT) estimators for animal availability. CLT estimators assume that detection probability is 1 for any animal at perpendicular distance zero (i.e., $p(0) = 1$, where $p(x)$ is the probability of detecting an animal that is at perpendicular distance x)—see Buckland et al. (2001). When this assumption holds, CLT estimators provide unbiased estimates of the detection function $p(x)$ and of animal density, but when $p(0) < 1$, they provide unbiased estimates $\hat{g}(x)$ of the shape of $p(x)$ but not of its height, that is, $E(\hat{g}(x)) = p(x)/a$ for some unknown a. CLT estimates of $p(x)$ are "corrected" for availability bias by multiplying them by an estimate of a obtained from independent observations of the availability process (see below).

Availability correction factors are used for estimation from both aerial surveys (e.g., Innes et al., 2002; Pollock et al., 2006) and shipboard surveys (e.g., Forcada et al., 2004; Gomez de Segura et al., 2006). The simplest correction factor $\hat{a}_1$ is the estimated proportion of time an animal is available for detection, which is an estimator of the probability that an animal is available at any randomly chosen instant and is therefore an appropriate correction factor when the survey is instantaneous. The longer animals are within detectable range, the less appropriate this correction factor will be.

Our motivating surveys are an aerial survey of bowhead whales (*Balaena mysticetus*) with perpendicular distance data only and a shipboard survey of beaked whales (*Ziphius cavirostris*) with both forward and perpendicular distance data. The aerial survey was conducted in West Greenland in March and April 2006, flying at approximately 46.3 m/s. In addition to the survey data, time series of animal depths were available from 8 independently tagged animals, and these were used to model the availability process (see below). The shipboard survey was conducted in the Alboran Sea in 2008 and 2009, sailing at approximately 3.5 m/s. In this case we do not have a full time series of depths from which to estimate availability. Instead we have estimates of $E(t_a)$ and $E(t_u)$ (where t_a and t_u are times available and unavailable within a single dive cycle, respectively), together with estimates of their standard errors from focal follows of beaked whales in the region.

After reviewing existing methods (Section 2), we formulate the survey as a hidden Markov model that integrates availability and detection processes (Section 3) and we develop maximum likelihood estimators (Section 4). These are used to estimate detection probability on the aerial and shipboard surveys in Section 5, after which we investigate estimator properties by simulation in Section 6. We extend the method to include hidden semi-Markov models for availability in Section 7 and finish with a summary and discussion in Section 8.

2. Review of Existing Methods

2.1. *Instantaneous Survey Correction Factor*

The simplest correction factor $\hat{a}_1$ mentioned above is an appropriate correction factor for the case in which only animals detected at a single pre-specified forward distance y_0 are included in the sample (e.g., when animals come abeam, at $y_0 = 0$). In this case, the probability that an animal at perpendicular distance x is detected is $a_1 g(x)$, where a_1 is the probability it is available at the selected y_0 and $g(x)$ the probability that an available animal at forward distance y_0 and perpendicular distance x is detected. Hence if $\hat{a}_1$ is an unbiased estimator of a_1 obtained independently of the survey and $\hat{g}(x)$ an unbiased estimator of $g(x)$, then $\hat{p}_1(x) = \hat{a}_1 \hat{g}(x)$ is an unbiased estimator of detection probability $p(x)$.

2.2. *Non-instantaneous Survey Correction*

Now consider the more usual case in which animals are detectable at a range of forward distances, and in which an animal at x is in detectable range for a time $T(x)$. In this case $\hat{a}_1$ is clearly not an appropriate correction factor, since the probability of an animal being available in a time interval $T(x)$ is greater than the probability it is available at an instant. Hence $\hat{p}_1(x)$ will be a negatively biased estimator of $p(x)$. This led McLaren (1961) and Laake et al. (1997) to develop estimators of $p(x)$ that involve multiplying an estimator of detection probability given availability by an estimator of the probability that an animal is available *at least once* in $T(x)$. McLaren's correction factor is based on animals deterministically alternating between being available for an estimated time $\hat{t}_a$ and unavailable for an estimated time $\hat{t}_u$, and is $\hat{a}_2(T(x)) = \{\hat{t}_a + T(x)\}/(\hat{t}_a + \hat{t}_u)$. The correction factor of Laake et al. (1997) is $\hat{a}_3(T(x)) = \left[\hat{E}(t_a) + \hat{E}(t_u)\left\{1 - e^{-T(x)/\hat{E}(t_u)}\right\}\right]/\{\hat{E}(t_a) + \hat{E}(t_u)\}$, where $\hat{E}(t_a)$ and $\hat{E}(t_u)$ are estimates of the average times that animals spend available and unavailable in a single available-unavailable cycle. This estimator is based on the assumption that the availability process is a two-state Markov process in which the expected duration of animal availability periods is $E(t_a)$ and the expected duration of animal unavailability periods is $E(t_u)$. The terms $\hat{E}(t_a)$ and $\hat{E}(t_u)$ are estimators of these expectations, obtained independently of the LT survey.

As Laake et al. (1997) note, $\hat{a}_2(T(x))$ is obtained in the limit from $\hat{a}_3(T(x))$ as $T(x)$ approaches zero, but with $\hat{t}_a$ and $\hat{t}_u$ replaced by $\hat{E}(t_a)$ and $\hat{E}(t_u)$. So for small $T(x)$ the two estimators are virtually indistinguishable but for other $T(x)$, $\hat{a}_3(T(x))$ is clearly better since when $T(x) > \hat{t}_u$, $\hat{a}_2(T(x))$ is greater than 1, and so $\hat{a}_3(T(x))$ should always be used in preference to $\hat{a}_2(T(x))$.

The estimator of $p(x)$ proposed by Laake et al. (1997) is $\hat{p}_3(T(x)) = \hat{a}_3(T(0))\hat{h}(x)$, where $\hat{h}(x)$ is the estimated proba-

bility of detecting an animal that is available within detectable range. Although Laake et al. (1997) did not use it in this way, it has been used by Forcada et al. (2004), Gomez de Segura et al. (2006) and others as a correction factor for CLT estimators, as follows: $\hat{p}_3(x) = \hat{g}(x)\hat{a}_3(T(0))$, where $\hat{g}(x)$ is the estimated detection function from a CLT analysis with $\hat{g}(0) = 1$. We show below that even if $T(0)$ were known (which it is usually not), $\hat{p}_3(x)$ is not an unbiased estimator of $p(x)$ in general.

2.3. *Poisson Availability Estimators*

Skaug and Schweder (1999), Okamura (2003), and Okamura et al. (2003) developed methods for dealing with availability bias that do not use correction factors and which model availability either as a Poisson process or as a deterministic process. The deterministic process is obviously not realistic and we note for reference below that the defining feature of the Poisson process is that it has a constant event hazard.

These models use detection hazard functions that depend on both perpendicular and forward distances, but they do not use forward distance data in estimation. The models we develop below are similar in spirit, but they accommodate availability processes with non-constant event hazards and do use forward detection distances in estimation.

2.4. *Non-Poisson Availability Estimators*

Okamura et al. (2012) developed one such model. It involves two states that correspond to surfacing and diving, respectively, with cues (e.g., blows) assumed to be generated according to a Poisson process when whales are in the surfacing phase. The Poisson rate parameter was estimated from visual observation, and integration over the state process was performed by substituting empirical dive data (i.e., the state process is characterized nonparametrically). While the method is formulated in continuous time, in their application Okamura et al. (2012) use discrete-time depth data from tagged whales with depth records every 60 seconds (after thinning 1-second data to save computer memory). The method uses a result of Skaug and Schweder (1999) to integrate the hazard in the forward detection, conditional on availability, and obtains the expectation over availability by averaging over all possible starting positions of the observed availability sequences. Okamura et al. (2012) say "Although we dealt with the surfacing/diving pattern in a nonparametric way, it would be interesting to develop more parametric methods for the treatment of surfacing/diving patterns". We develop one such method here.

3. A Hidden Markov Model for the Detection Process

We define a Cartesian coordinate system that moves with the observer along a transect line, and has the observer at $(x = 0, y = 0)$ (see Figure 1). Forward distance is denoted y and perpendicular distance from the line, x. At time t animal i is located at (x_i, y_{it}). In common with almost all LT models we assume that animals do not move while within detectable range as the observer moves along the transect line. The data are truncated for estimation at some maximum perpendicular distance W, and at perpendicular distance x the observer searches out to some maximum forward distance $y_{\max}(x)$ (see Figure 1).
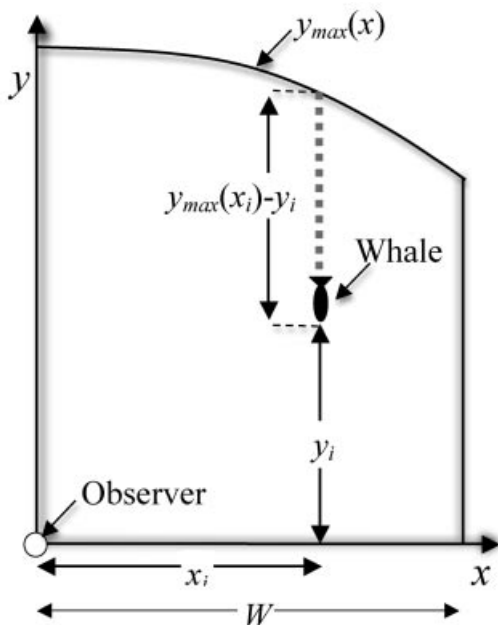
Figure 1. Notation for field of view and location. The vertical arrow is the transect line, with the arrow indicating direction of the observer's movement. The black whale shape shows the location of whale i when it is detected (at forward distance y_i and perpendicular distance x_i). The maximum forward distance that an animal at perpendicular distance x can be detected denoted $y_{\max}(x)$. The whale in the figure "survived" detection for a distance $y_{\max}(x_i) - y_i$. W is the maximum perpendicular distance considered for estimation.

3.1. *Availability Process Models*

Bowhead and beaked whales have availability processes that are more clustered than Poisson. To accommodate this, we model availability using hidden Markov models (HMMs), and later hidden semi-Markov models (HSMMs). Although the availability process is in reality a continuous-time process, here we consider time in discrete units, with equal time between units. We discretize for two reasons, the first being that it is often the case that tags attached to animals to record their depth, and hence availability, gather data at fixed time intervals (one second in the data we consider below), and it is from these data that the availability process model is constructed. Secondly, modelling the availability process as a discrete-time process allows us to take advantage of the substantial body of HMM (and HSMM) literature and software to model the process. And if availability data do not come pre-discretized, a HMM can get arbitrarily close to a continuous-time model simply by shortening the HMM time unit.

We also note that both errors in distance estimation (which are typical of line transect surveys) and animal movement translate into errors in estimating how far in time detections are ahead of observers. On shipboard surveys this time error will very often be much greater than one second (which corresponds to a distance of about 5 m when moving at 10 knots), so that any benefit of modelling the availability process at higher resolution than a second will be swamped by the much greater errors in estimating distance. This is less the case on aerial surveys, where, depending on survey protocol, the time it takes a detected animal to come abeam after detection may be recorded to within a second or less.

For an animal at any x, the process is modelled as a discrete time series with availability $A_t = 1$ if an animal is available at time t and $A_t = 0$ otherwise ($t = 1, \ldots, T(x)$), where $t = 1$ is the first time unit after the animal enters the observer's field of view, and $T(x)$ the last before it leaves.

We allow animals to be in one of m notional hidden states; the state at time t is denoted S_t. Given this state, the probability mass function (pmf) of A_t is $f_{A|S}(A_t|S_t) = \lambda_{S_t}^{A_t}[1 - \lambda_{S_t}]^{1-A_t}$, where λ_{S_t} is the probability of the animal being available, given its hidden state S_t. S_t takes on values $1, \ldots, m$ according to a homogeneous and irreducible Markov chain with $m \times m$ transition probability matrix $\boldsymbol{\Gamma}$, with stationary distribution $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_m)$. $\boldsymbol{\Gamma}$ has elements $\gamma_{j,k}$, which are the transition probabilities from states j to states k ($j, k \in \{1, \ldots, m\}$).

HMMs do not have constant event hazards, but they include a constant-hazard process as a degenerate case: a
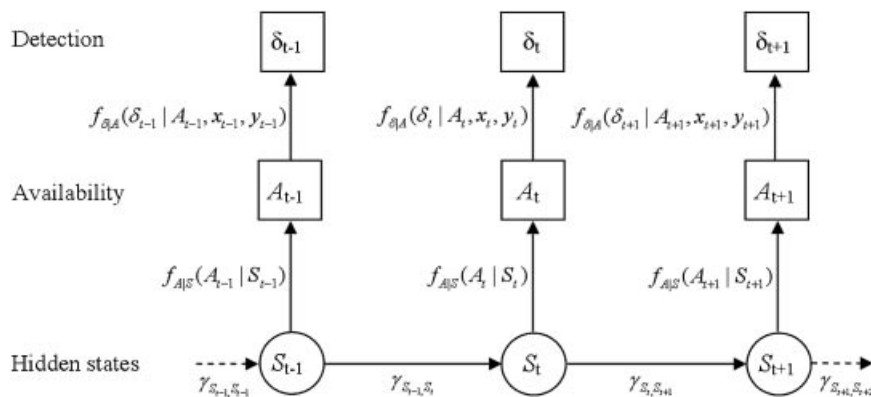
Figure 2. Schematic representation of the hidden Markov model (HMM) for a single animal. For times $t^* = (t-1)$, t, $(t+1)$: S_{t^*} are the hidden states, A_{t^*} are the availability indicators, δ_{t^*} are the detection indicators, S_{t^*}, S_{t^*+1} are the transition probabilities, and x_{t^*} and y_{t^*} are the Cartesian coordinates of the animal relative to the observer. $f_{A|S}(\cdot)$ is the pmf of the availability indicator given state, and $f_{\delta|A}(\cdot)$ is the pmf of the detection indicator given availability and position.

single-state HMM with a binary state-dependent process with parameter λ^* is a constant-hazard process with event rate λ^*. By setting $\lambda^* = \sum_{k=1}^{m} \tau_k \lambda_k$ we create a constant-hazard process with the same mean event rate as an m-state HMM with stationary distribution $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_m)$ and state-dependent event rates $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$. We use this process in our simulation study below.

3.2. *Detection Process Models*

We assume that when $A_t = 0$ animals are not detected, when $A_t = 1$ animals may be detected or missed and the farther an available animal is from the observer, the less likely it is to be detected. To model this, we introduce a random variable δ_t, such that $\delta_t = 1$ if the animal is detected at time t, $\delta_t = 0$ otherwise and we define the pmf of δ_t to be $f_{\delta|A}(\delta_t|A_t, x_t, y_t) = h(x_t, y_t)^{\delta_t}[1 - h(x_t, y_t)]^{1-\delta_t}$ if $A_t = 1$, and 0 if $A_t = 0$. Following Skaug and Schweder (1999), we refer to $h(x, y)$ as the discrete detection hazard; it has a parameter vector $\boldsymbol{\beta}$, which is to be estimated. Note that $h(x, y)$ is the conditional probability of detecting an available animal and is therefore similar to $h(x)$ in the formulation of Laake et al. (1997), but with dependence on forward distance.

4. Model Fitting and Checking

We condition on estimates of $\boldsymbol{\Gamma}$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_m)$ for estimation and include uncertainty in estimating them via a bootstrap procedure (below).

4.1. *Detection Probability*

Given only the state S_t and position (x, y_t) of an animal at time t, the pmf for δ_t is $f_{\delta|S}(\delta_t|S_t, x, y_t) = [\lambda_{S_t} h(x, y_t)]^{\delta_t}[1 - \lambda_{S_t} h(x, y_t)]^{1-\delta_t}$. We can therefore model the time series of observed δs as a hidden Markov model (HMM) with states as above, but with δ_t considered as arising directly from S_t, according to $f_{\delta|S}$.

The positions (x, y_t) of an animal that is within detectable range at time occasions $t = 1, \ldots, T(x)$ are easily calculated using the speed of the vessel. Because detection probability is believed to change after first detection and subsequent detections are typically not recorded, we use data only up to the time of first detection, as is usual on line transect surveys. As a result, the δs for an animal first detected at time t is a series of $(t - 1)$ zeros, followed by a single 1. Given the series of states of the animal, $\boldsymbol{S}_{T(x)} = (S_1, \ldots, S_{T(x)})$, and recalling that t is 1 when the animal enters the detectable range, we can write the probability of first observing the animal at time t as $\Pr(\delta_1, \ldots, \delta_t \mid \boldsymbol{S}_{T(x)}, x) = \prod_{u=1}^{t} f_{\delta|S}(\delta_u|S_u, x, y_u)$, where $\delta_u = 0$ for $u < t$ and $\delta_t = 1$. Adapting a standard HMM formulation (see, e.g., Zucchini and MacDonald, 2009) for our purposes, we can write the unconditional probability that an animal at x is first detected at time t as follows:

$$p(x, y_t) = \boldsymbol{\tau} \left(\prod_{u=1}^{t-1} \boldsymbol{B}(\delta_u = 0, x, y_u) \right) \boldsymbol{B}(\delta_t = 1, x, y_t) \, \boldsymbol{1}', \quad (1)$$

where $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_m)$ is the stationary distribution of the Markov chain, $\boldsymbol{B}(\delta_u, x, y_u) = \boldsymbol{\Gamma} \boldsymbol{f}(\delta_u \mid x, y_u)$ and $\boldsymbol{f}(\delta_u \mid x, y_u) = \mathrm{diag}(f_{\delta|S}(\delta_u|S_u = 1, x, y_u), \ldots, f_{\delta|S}(\delta_u|S_u = m, x, y_u))$, and $\boldsymbol{1}'$ is a column vector of m 1s. The probability of detecting an

animal that is at perpendicular distance x is the complement of the probability of non-detection, namely $p(x) = 1 - \boldsymbol{\tau} \left(\prod_{u=1}^{T(x)} \boldsymbol{B}(\delta_u = 0, x, y_u) \right) \boldsymbol{1}'$.

4.2. *Only Perpendicular Distance Observed*

The detection function $p(x)$ is a function of the availability process parameters $\boldsymbol{\Gamma}$ and $\boldsymbol{\lambda}$ (which are taken as known in the likelihood) and the detection hazard parameters $\boldsymbol{\beta}$. Writing the detection function as $p(x; \boldsymbol{\beta})$, the likelihood for $\boldsymbol{\beta}$, given that n animals were detected at distances $\boldsymbol{x} = (x_1, \ldots, x_n)$ is

$$L_x(\boldsymbol{\beta}|\boldsymbol{x}) = \prod_{i=1}^{n} \frac{p(x_i; \boldsymbol{\beta})\pi(x_i)}{\int_0^W p(x; \boldsymbol{\beta})\pi(x) \, \mathrm{d}x} = \prod_{i=1}^{n} f_x(x; \boldsymbol{\beta}), \quad (2)$$

where $\pi(x) = 1/W$, as is usual with LT surveys.

4.3. *Both Perpendicular and Forward Distances Observed*

Let y_i be the forward distance at which animal i is first detected $(i = 1, \ldots, n)$. For an animal at x, the pdf of y is $f_{y|x}(y|x; \boldsymbol{\beta}) = p(x, y; \boldsymbol{\beta})/p(x; \boldsymbol{\beta})$ and we can write the likelihood for $\boldsymbol{\beta}$ as $L_{xy}(\boldsymbol{\beta} \mid \boldsymbol{x}, \boldsymbol{y}) = L_x(\boldsymbol{\beta}|\boldsymbol{x})L_{y|x}(\boldsymbol{\beta}|\boldsymbol{x}, \boldsymbol{y})$, where $L_{y|x}(\boldsymbol{\beta}|\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} f_{y|x}(x_i, y_i; \boldsymbol{\beta})$. Hence

$$L_{xy}(\boldsymbol{\beta}|\boldsymbol{x}, \boldsymbol{y}) = \prod_{i=1}^{n} \frac{p(x_i, y_i; \boldsymbol{\beta})\pi(x_i)}{\int_0^W p(x; \boldsymbol{\beta})\pi(x) \, \mathrm{d}x}. \quad (3)$$

4.4. *Accommodating Individual Variation in Availability Processes*

It is frequently the case that there is individual heterogeneity in the surfacing process, and there may be temporal heterogeneity and heterogeneity due to environmental factors. From the surveyor's perspective, individual heterogeneity is essentially random, as is heterogeneity due to time or environment if the effects of these covariates on availability are not modelled.

We use the sample of parameters obtained by fitting HMMs to individual whales as an empirical estimator of the joint distribution of individual random effects and draw inferences by marginalizing over these. When separate HMMs have been fitted to availability data from J animals, the marginal likelihoods when only perpendicular distances are recorded, and when perpendicular and forward distances are recorded, are

$$L_x(\boldsymbol{\beta}|\boldsymbol{x}, \{\boldsymbol{\Gamma}, \boldsymbol{\lambda}\}) = \frac{1}{J} \sum_{j=1}^{J} L_x(\boldsymbol{\beta}|\boldsymbol{x}, \boldsymbol{\Gamma}_j, \boldsymbol{\lambda}_j) \quad \text{and} \quad (4)$$

$$L_{xy}(\boldsymbol{\beta}|\boldsymbol{x}, \boldsymbol{y}, \{\boldsymbol{\Gamma}, \boldsymbol{\lambda}\}) = \frac{1}{J} \sum_{j=1}^{J} L_{xy}(\boldsymbol{\beta}|\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\Gamma}_j, \boldsymbol{\lambda}_j), \quad (5)$$

where $\{\boldsymbol{\Gamma}, \boldsymbol{\lambda}\} = (\boldsymbol{\Gamma}_1, \boldsymbol{\lambda}_1, \ldots, \boldsymbol{\Gamma}_J, \boldsymbol{\lambda}_J)$ and $\boldsymbol{\Gamma}_j$ and $\boldsymbol{\lambda}_j$ $(j = 1 \ldots, J)$ are the estimated HMM parameters for animals 1 to J. (The terms inside the summations on the RHS of Equations (4) and (5) are just Equations (2) and (3) with dependence on the HMM parameters made explicit.)

We estimate detection probability as $\hat{p} = \frac{1}{J} \sum_{j=1}^{J} \int_0^W p(x; \hat{\boldsymbol{\beta}} \mid \boldsymbol{\Gamma}_j, \boldsymbol{\lambda}_j) \pi(x) \, dx$, where $\hat{\boldsymbol{\beta}}$ is the value of $\boldsymbol{\beta}$ that maximizes Equation (4) (when only perpendicular distances are available) or (5) (when perpendicular and forward distances are available).

4.5. *Density Estimation*

It is ultimately animal density rather than detection probability that is the object of inference and LT density estimators have the form $\hat{D} = \frac{n}{2WL} \frac{1}{\hat{p}}$, where n is sample size, W is the searched strip half-width, L is line length and $\hat{p}$ is estimated detection probability. Here $\frac{n}{2WL}$ is the observed density of animals and $1/\hat{p}$ can be viewed as a correction factor for animals that are missed. In evaluating estimator bias, we therefore focus on estimators of $1/p$.

If we let $g(x) = p(x)/p(0)$, we can write $1/p$ as $1/\{p(0)g\}$, where $g = \int_0^W g(x) \frac{1}{W} dx$. CDS estimators assume that $p(0) = 1$ so that the bias-corrected CDS estimators of $1/p$ using $\hat{a}_1$ and $\hat{a}_3(T(0))$ are $1/(\hat{a}_1 \hat{g})$ and $1/(\hat{a}_3(T(0)) \hat{g})$. We therefore investigate bias in bias-corrected CDS estimators by considering the bias in $1/\hat{a}_1$ and $1/\hat{a}_3(T(0))$ as estimators of $1/p(0)$.

4.6. *Interval and Variance Estimation*

When time series of availability data are available (as in the aerial survey), variance and confidence interval estimates are obtained by bootstrapping as follows: (1) Resample individuals with replacement; (2) Simulate availability data for each resampled individual's fitted availability HMM, (3) re-estimate HMM availability model parameters from these availability data; (4) resample detection locations with replacement; (5) refit detection hazard model given the refitted availability model.

When we have only $\hat{E}(t_a)$ and $\hat{E}(t_u)$ and an estimate of their variance-covariance matrix, we bootstrap as follows: (1) parametrically resample $\hat{E}(t_a)$ and $\hat{E}(t_u)$ assuming bivariate normality; (2) resample detection locations with replacement; (3) re-estimate detection probability given the resampled $\hat{E}(t_a)$ and $\hat{E}(t_u)$ and resampled detection locations. In the case of the shipboard survey here, we do not have estimates of the covariance of $\hat{E}(t_a)$ and $\hat{E}(t_u)$ and therefore assume them to be independent. In both cases interval estimates are obtained using the percentile method.

4.7. *Goodness of Fit*

Goodness-of-fit was evaluated using Q–Q plots and Kolmogarov–Smirnov (KS) tests separately in the x and y dimensions. In the x-dimension, the ith CDF value used in the KS tests and Q–Q plots is evaluated as $F_{(i)} = F_x(x_{(i)}|\hat{\boldsymbol{\beta}}) = \int_0^{x_{(i)}} f_x(x; \hat{\boldsymbol{\beta}}) dx$, where $x_{(i)}$ is the ith smallest observed x.

To apply a KS test and do Q–Q plots in the y-dimension the CDF values are evaluated as $F_i = F_{y|x}(y_i|x_i; \hat{\boldsymbol{\beta}}) = \int_0^{y_i} f_{y|x}(y|x; \hat{\boldsymbol{\beta}}) dy$, then these and the corresponding empirical CDF values were sorted in order of increasing F_i.

4.8. *Discrete Detection Hazard Models*

Implementation requires functional forms for $h(x, y)$. We consider variants of two forms in the LT literature and we assume that an animal that is available at radial distance zero will be seen (i.e., $h(0, 0) = 1$). This is a much weaker assumption than is made by CLT estimators (which assume

that any animal at perpendicular distance zero will be seen). Our first form is the exponential power model, variants of which were used by Skaug and Schweder (1999), Okamura (2003), and Okamura et al. (2003). The most general form we consider is $h(x, y) = \exp\left\{-\left(\frac{x}{\sigma_x}\right)^{\gamma_x} - \left(\frac{y}{\sigma}\right)^{\gamma}\right\}$. We refer to this as the EP2x model, and consider a number of special cases of it. When the x and y scale parameters are constrained to be equal (i.e., $\sigma_x \equiv \sigma$) we call it the EP2 model. When the two scale parameters are allowed to differ but the shape parameters are equal (i.e., $\gamma_x \equiv \gamma$) we call it the EP1x model, and when the scale parameters and the shape parameters are constrained to be equal (i.e., $\sigma_x \equiv \sigma$ and $\gamma_x \equiv \gamma$) we call it the EP1 model. Skaug and Schweder (1999) used EP1, Okamura (2003), and Okamura et al. (2003) used EP2 and Okamura, Minamakawa, and Kitakado (2006) used EP2x. The final form we consider is the inverse power model (IP model), used by Skaug and Schweder (1999) which is defined as $h(x, y) = \sigma^{\gamma} / (\sigma^2 + x^2 + y^2)^{\gamma/2}$.

We allow the detection hazard functions to depend on a covariate vector through the scale parameter(s), using a log link function. That is, we have $\sigma(z; \boldsymbol{\beta}_z) = \exp(\boldsymbol{\beta}_z^T z)$ and $\sigma_x(z_x; \boldsymbol{\beta}_{x,z}) = \exp(\boldsymbol{\beta}_{x,z}^T z_x))$, where z and z_x are covariate vectors, the latter being specific to σ_x, and $\boldsymbol{\beta}_z$ and $\boldsymbol{\beta}_{x,z}$ are associated parameter vectors, the latter being specific to σ_x.

5. Applications

5.1. *Aerial Bowhead Survey*

5.1.1. The data. No forward distances were recorded on this survey, and only $n = 33$ detections were made. Perpendicular distances and school sizes for all detected groups were recorded. All but three detections were of single animals; the rest were schools of size two. Independently of the survey, electronic depth-recording tags were attached to eight bowhead whales from the population surveyed, generating time series of durations between 2.6 and 53 hours, with depths recorded every second—see Laidre et al. (2007). Following previous practice (Heide-Jørgensen et al., 2007), animals were considered to be available for detection only when within 2 m of the surface. The time series were accordingly converted into binary availability time series. The challenge facing analysts after the survey was how best to estimate detection probability and whale density when there may be substantial availability bias and only perpendicular distance data were recorded on the survey.

To investigate this, we start by fitting HMMs to the tag data and maximizing Equation (4) conditional on the estimated HMM parameters. HMMs were fitted separately to each time series of availability data by numerically maximizing the (standard) HMM likelihood. We considered models with a variety of values for m and found models with $m = 2$ states to be best on the basis of AIC values. Details of the fitted models are given in the Web Appendix.

5.1.2. Detection function estimation. Models with and without school size were fitted and those without school size were preferred on the basis of AIC. AICs for models with school size were larger by about 1 than models without school size. Figure 3 shows the estimated pdfs of perpendicular distance obtained by maximizing Equation (4) with respect to

detection function parameters using models EP1x, EP2, EP1 and IP without school size. The fit of the most flexible model (EP2x) failed to converge. All models fit the perpendicular distance data well, with KS test p-values ranging from 0.969 to 0.997. Model EP1 is preferred on the basis of AIC (Δ AIC = 1.67 for EP1x, 1.98 for EP2 and 0.08 for IP). The estimate of $1/p$ from EP1 is 4.62, with an estimated CV and 95% confidence interval from 999 bootstraps of 20.9% and (3.43, 7.29) respectively. When the estimated availability HMM is treated as known the CV drops to 19.4%.

The models predict very different forward detection distance distributions (Figure 3b), and estimates of $p(0)$ and mean detection probability p that differ by up to 22%. Estimated $p(0)$s are 0.33, 0.33, 0.31 and 0.27 for models EP1, IP, EP2 and EP1x respectively, with estimated CVs of 0.03, 0.09, 0.71, 0.04. Estimated ps are 0.21, 0.22 0.20 and 0.18 for models EP1, IP, EP2 and EP1x respectively, with estimated CVs of 0.03, 0.09, 0.71, 0.04.

When either the scale or shape parameter of the hazard function is allowed to vary separately in x- and y-dimensions, these parameters are estimated with very poor precision in the y-dimension (for model EP1x, $\hat{CV}(\gamma_y) = 2,373\%$ and for model EP2, $\hat{CV}(\sigma_y) = 212\%$) and when both scale and shape are allowed to vary separately in the two dimensions (model EP2x), we are unable to estimate either parameter. The very high CV of $\hat{p}(0)$ in the case of model EP2 arises because $p(0)$ is sensitive to how far ahead animals are detected and of the fitted models, EP2 is the only one that allows a separate scale parameter in the x and y dimensions. The derivative of $\hat{p}(0)$ with respect to the hazard function parameter associated only with x in models EP2 and EP1x is zero, that is, for model EP2, $\partial\hat{p}(0)/\partial\gamma_x = 0$ and for model EP1x, $\partial\hat{p}(0)/\partial\sigma_x = 0$. This confirms what one would expect *a priori*: that when the model parameters relating to the perpendicular distance direction are not tied to those relating to the forward direction, the perpendicular distance data contain no information about $p(0)$. When there are no forward distance data, $\hat{p}(0)$, $\hat{p}$ and $\hat{D}$ are model-dependent: they are determined by the form of the assumed hazard function for a given availability model.

There is large inter-whale variation in estimated detection probabilities. Separate fits of model EP1 using each one of the eight sets of availability data separately produce estimates of $p(0)$ ranging from 0.16 to 0.61, from which it is clear that individual heterogeneity in availability pattern has a very large effect on estimated detection probability and that use of availability data from a single whale could result in very biased estimates.

In addition, estimates are sensitive to the availability model used. When we fit EP1 with availability HMMs that have constant hazards (with the same mean availability rates as the 8 original HMMs) the fit is excellent (KS p-value = 0.99) but $p(0)$ and p are estimated to be 0.97 and 0.66 respectively—three and two times the size of the estimates when using the original HMMs.

5.1.3. Correction factor bias. Figure 4a–c show the biases of $1/\hat{a}_1$ and $1/\hat{a}_3(T(0))$ as estimators of $1/p(0)$. The bias of $1/\hat{a}_3(T(0))$ is shown as a function of the maximum time in view, $T(0)$. In the absence of compelling evidence in favor of any one of the four models fitted to the aerial survey data, we consider this bias assuming each of the fitted models IP, EP1x and EP2 to be true. (We do not consider EP1 as it produced a very similar fit to IP.) We plot %bias as $100(\hat{a}_1^{-1}/p(0)^{-1} - 1)$ and $100(\hat{a}_3(T(0))^{-1}/p(0)^{-1} - 1)$, where $p(0)$ is the intercept of the assumed model. The values of $\hat{a}_1^{-1}$ and $\hat{a}_3(T(x))^{-1}$ used in the figure are the averages of $\hat{a}_1^{-1}$ and $\hat{a}_3(T(x))^{-1}$ over the eight tagged whales.

The bias of $1/\hat{a}_1$ and $1/\hat{a}_3(T(0))$ depends strongly on the true model, and more specifically the true distribution of forward detection distances. The bias using $1/\hat{a}_1$ is virtually zero for model EP1x because in this case the survey is virtually instantaneous. The bias when using $1/\hat{a}_3(T(0))$ can be substantial and either positive or negative, depending on the value assumed for $T(0)$, and in the limit, as true $T(0) \to 0$, both it and $1/\hat{a}_1$ are unbiased. Conversely the longer the true $T(0)$ is, the more arbitrary $1/\hat{a}_3(T(0))$ is as a correction factor. This makes $1/\hat{a}_3(T(0))$ potentially much more dangerous as a correction factor for shipboard surveys, where $T(0)$ is typically much longer than on aerial surveys. We therefore now turn our attention to the shipboard survey data.

5.2. Shipboard Beaked Whale Survey

5.2.1. The data. The shipboard survey has a larger sample size ($n = 73$) and both perpendicular and forward distances to all detections. In addition the height of the observer (ht), the school size (ss) was recorded for every detection.

Data on availability was less comprehensive than was the case with the bowhead survey. Rather than time series of available/unavailable status for a number of whales, we have only the mean time animals are available and unavailable ($\hat{E}(t_a)$ and $\hat{E}(t_u)$) and their standard errors, obtained by focal follows of $J = 57$ animals. We are therefore not able to estimate HMM parameters for the availability process and instead assume that the availability process is a two-state Markov process: λ_1 and λ_2 are set to 0 and 1, respectively, and the state transition probabilities set to $\gamma_{12} = 1/\hat{E}(t_u)$ and $\gamma_{21} = 1/\hat{E}(t_a)$ (see Zucchini and MacDonald, 2009, p133 for justification of this). $\hat{E}(t_a)$ and $\hat{E}(t_u)$ were estimated to be 2 minutes and 26 minutes, with inter-animal coefficients of variation (CVs) 7.9% and 8.5%, respectively.

Focal follows involve visually tracking animals for as long as possible after first detection and recording when they are available and unavailable for detection. This method of obtaining availability data is somewhat less reliable than obtaining it from recorders attached to animals, because it is possible that some availability events may be missed, but it is very much easier and cheaper than attaching tags and sometimes the only feasible option. The approach of Conn et al. (2012) may be useful for modelling availability when availability events are missed.

5.2.2. Detection function estimation. All five models were fitted to the survey data after truncating at a perpendicular distance of 2,000 m, without covariates in the detection function. The best model, on the basis of AIC, was found to be model IP and this was then fitted with σ depending on ht, on ss, on ht and ss and on ht, ss and a $ht \times ss$ interaction. The best model (by AIC) was that with ht and ss but no interaction. Each of the other detection function model forms were then fitted with σ (and in the case of model
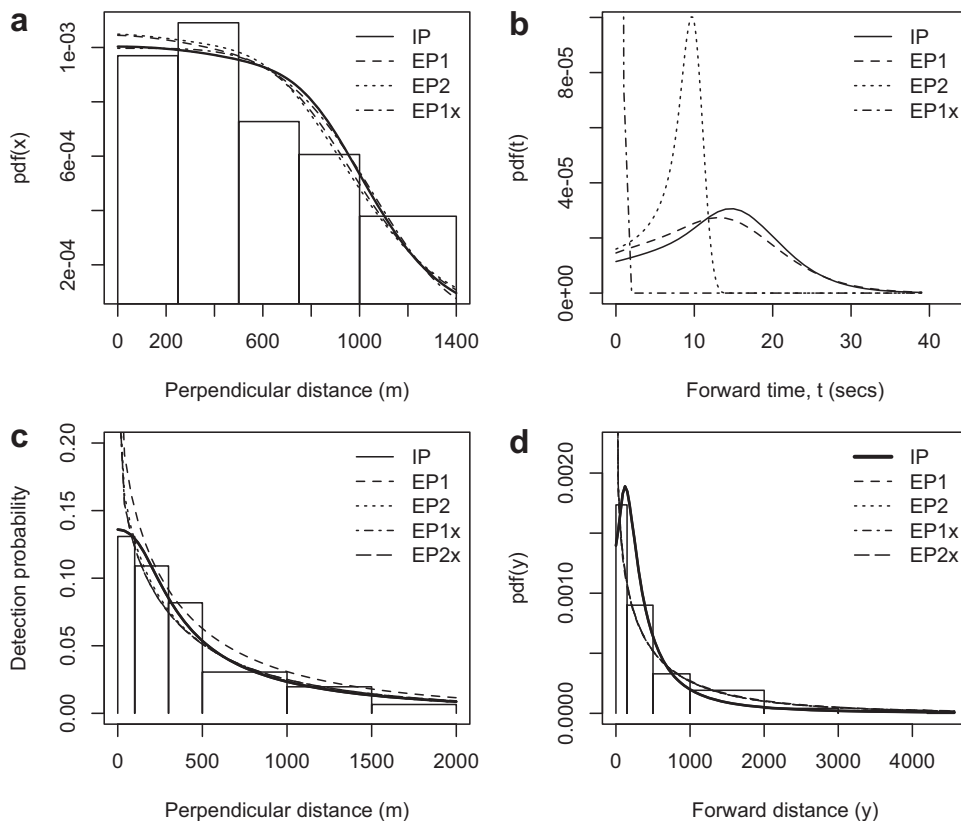
Figure 3. Fitted models to aerial (top) and shipboard (bottom) in the perpendicular distance and forward dimensions. (a) Aerial survey fitted pdfs of perpendicular distance overlaid on the histogram of perpendicular distances of the 33 detected animals. (b) Aerial survey fitted pdfs of forward detection times. (c) Ship survey fitted detection functions overlaid on the histogram of perpendicular distances of the 73 detected animals. (d) Ship survey estimated pdfs of forward detection distance overlaid on the histogram of forward detection distances. (The pdfs for models other than IP are virtually coincident.)

EP2 σ_x) depending on *ht* and *ss* and no interaction. Model IP was found to be best. The differences in AIC statistics and the Kolmogariov-Smirnov goodness of fit p-values in the *x* and *y* dimensions for each combination of covariates when using model IP are given in Web Appendix A. The fit in the perpendicular and forward distance directions is shown in Figure 3c and d. The KS goodness of fit tests for the perpendicular and forward distance fits for the selected model have p-values of 0.94 and 0.83, respectively.

With 999 bootstrap replicates, the CV of $1/\hat{p}$ is estimated to be 31.8% (the point estimate is 25.4) with a 95% confidence interval of (14.7, 45.0). When uncertainty in HMM estimation is neglected (by using the original HMM in all bootstrap resamples), the CV of $1/\hat{p}$ reduces to 30.6%

When model IP is fitted using a constant-hazard HMM with the same mean availability rate as the original availability HMM, the fit is good but $p(0)$ and p are both estimated to be roughly 7 times larger than when using the original HMM.

5.2.3. Correction factor bias. Figure 4d shows the bias of $1/\hat{a}_1$ and $1/\hat{a}_3(T(0))$ as estimators of $1/p(0)$, where truth is taken to be the fitted IP model. It is clear from the figure that the bias correction methods do not in general give unbiased estimates of density or detection probability, and when

animals are within detectable range for a long time relative to the mean time that they are available for detection, as in this shipboard survey, their bias can be very large (between -77% and $+89\%$ here). With an estimated $p(0)$ of 0.136 from the HMM line transect (HMMLT) model, it is also clear that estimating density using a CLT estimator with no availability correction would result in large negative bias (-86.5% bias if the fitted IP model were truth).

6. Simulation Study

We conducted a simulation study to investigate the HMMLT model performance for the aerial and shipboard survey scenarios with a range of sample sizes including those observed ($E[n] = 33$, 73, 200, and 1000). We investigate the performance of the HMMLT estimator with forward distances using the fitted model EP1 to generate perpendicular and forward distance observations for the aerial survey scenarios and model IP for the shipboard scenarios. In both cases we simulate and fit without covariates in the detection hazard functions.

In the simulations $N = E[n]/\hat{p}$ animals were distributed uniformly in a strip of half-width W, where $\hat{p}$ is the estimated detection probability from the model in question. In the
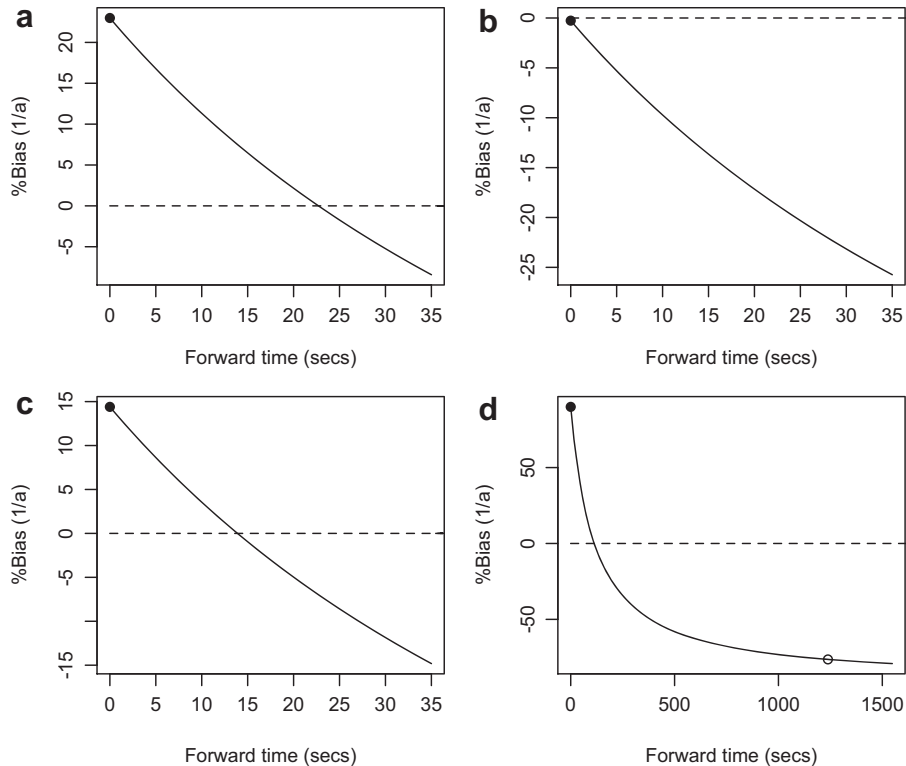
Figure 4. Estimated bias of the availability correction factor $1/\hat{a}_3(T(0))$ (solid line) for (a)–(c) the aerial and (d) the shipboard survey. Plots (a)–(c) are assuming models IP, EP1x and EP2 respectively. Plot (d) assumes model IP for the shipboard survey. The open dot in (d) corresponds to the time in view of the shipboard sighting that was made at greatest forward distance; using this $T(0)$ results in a negative bias of 77%. The solid dots are the bias of the correction factor $1/\hat{a}_1$, which results in biases of 25%, −0.5%, 14%, and 89% in scenarios (a)–(d) respectively.

case of the aerial survey, each animal was assigned an availability process by simulating 8 sets of availability data from the 8 HMMs, refitting HMMs to each simulated dataset and randomly assigning one of these to each whale. In the case of the shipboard survey, each animal was assigned an $E[t_a]$ and $E[t_u]$ by drawing from a normal distribution with parameters corresponding to the original estimated $E[t_a]$ and $E[t_u]$ and their variances. Detections were simulated according to the HMMLT models described above and forward and perpendicular distances of all detections were used to estimate model parameters. This process was repeated 500 times for each scenario.

To investigate the effect of neglecting the uncertainty in the HMM for availability, we also conducted simulations in which the original HMMs for availability and original $E[t_a]$ and $E[t_u]$ were used in estimation ("Fixed" in Table 1). And to investigate the performance of estimators that assume a constant hazard availability process, we also estimated assuming HMMs with parameters set to generate constant-hazard availability, as described above ("ConstHaz" in Table 1). Finally, in the case of the aerial survey where we have availability time series data, we investigate the adequacy of the availability HMM by repeating these simulation scenarios but simulating availability by randomly sampling segments of length greater than

maximum time in view from the 8 time series. Estimated biases, standard errors and root mean squared errors for each scenario are shown in Table 1.

The HMMLT estimators are almost unbiased for all scenarios with availability generated from fitted HMMs, even for small sample sizes. While the HMM availability model for the aerial survey (at least) is not a perfect model—as evidenced by the fact that when availability is generated from the raw availability data the HMM-based estimator is more biased by a couple of percent—it nevertheless performs well, with low bias in all scenarios. The fact that the hazard models used for generation and estimation are identical means that the estimator performance in simulations will be somewhat better than can be expected in practice and we expect that a sample size of 33 is too small for reliable inference. A rule-of-thumb sample size for CLT surveys is around 60 and we expect that similar or larger sample sizes will be required for HMMLT estimators in practice.

Assuming a constant-hazard availability process results in large negative bias in all scenarios. A more clustered availability process (as with the estimated HMMs) allows more animals to have few or no availability events while in view and hence overall detection probability to be lower than would be the case with a constant-hazard availability process.

Table 1

Estimated percentage bias, standard error (se) and root mean squared error (RMSE) of $1/\hat{p}$ for aerial survey scenarios (with Bowhead whale availability model) and shipboard survey scenarios (with Beaked whale availability model), for four expected sample sizes ($E[n]$), using the estimated hazard models EP1 from the aerial survey and IP from the ship survey, respectively

Scenario	$E[n]$	HMM	HMM Avail. Generation			Tag Avail. Generation		
			%Bias	se	RMSE	%Bias	se	RMSE
Aerial	33	Rand	1.05	0.290	0.294	−1.15	0.353	0.357
EP1	73		0.76	0.197	0.200	−1.67	0.225	0.239
	200		0.32	0.114	0.115	−1.93	0.128	0.158
	1000		0.38	0.049	0.052	−1.96	0.058	0.111
	33	Fixed	−0.28	0.299	0.299	−0.91	0.347	0.350
	73		−0.72	0.179	0.182	−1.72	0.225	0.240
	200		−1.03	0.113	0.122	−2.08	0.125	0.160
	1000		−1.17	0.048	0.072	−2.08	0.057	0.120
	33	ConstHaz	−65.7	0.100	3.01	−65.5	0.139	3.17
	73		−66.0	0.068	3.02	−66.0	0.079	3.19
	200		−66.0	0.040	3.02	−64.1	0.045	3.19
	1000		−66.1	0.018	3.02	−66.1	0.021	3.19
Ship	33	Rand	2.95	8.00	8.02			
IP	73		1.81	5.52	5.54			
	200		−0.87	3.01	3.02			
	1000		−0.46	1.32	1.32			
	33	Fixed	3.04	8.70	8.72			
	73		1.12	4.98	4.98			
	200		−0.07	2.89	2.89			
	1000		−0.61	1.32	1.33			
	33	ConstHaz	−85.2	0.89	21.2			
	73		−85.6	0.52	21.3			
	200		−85.8	0.30	21.4			
	1000		−85.8	0.13	21.3			

Results are shown for scenarios in which the availability HMM parameters are re-estimated (Random), treated as known (Fixed) and for the case in which the availability is assumed to be constant-hazard and known (ConstHaz). 'HMM Avail. Generation' means animal availability was generated using the fitted HMM for the scenario; 'Tag Avail. Generation' means animal availability was generated by randomly sampling segments of the original tag data.

Assuming constant-hazard availability runs the risk of severe estimator bias in general.

Failing to incorporate the uncertainty associated with estimating the availability HMMs ("Fixed" scenarios) has very little effect.

7. Hidden Semi-Markov Availability Models

An implicit assumption of HMMs is that the state dwell-times (i.e., the durations the process spends in each state) are geometrically distributed. This is not always biologically reasonable for marine mammals. For example, a state that is associated with a deep dive may have a much lower probability of short duration and higher probability of long duration than can be modelled by a geometric distribution, resulting in a poor fit. In such cases the use of HMMs may lead to biased density estimates. So-called hidden *semi*-Markov models (HSMMs) allow flexibility in the dwell-time distributions by modelling them explicitly using some distribution on the positive integers. Langrock and Zucchini (2011) showed that by expanding each semi-Markovian state into a sufficiently large set of Markovian states sharing the same state-dependent distribution, and structuring the transition probabilities between those states appropriately, one can approximate the desired dwell-time distributions arbitrarily accurately and hence ap-

proximate any HSMM with arbitrary accuracy using a HMM. Using this approach, all methods described in Section 4 above are straightforwardly applicable. In particular, the formulae given for the likelihood of associated detection models remain valid. So when combined with the method of Langrock and Zucchini (2011), the methods of this paper provide a means of accommodating HSMM availability processes in line transect surveys. The added flexibility that HSMMs provide does come with a computational cost, as the HMM representation of a HSMM typically requires many more states than are in the HSMM, making the corresponding HMM transition probability matrix large, which can slow down likelihood computation substantially.

To investigate the utility of HSMMs relative to HMMs, we fitted two-state HSMMs with negative binomial state dwell-time distributions to the bowhead availability series. (We do not have availability time series in the case of the beaked whales.) This leads to much improved AIC values (see Web Appendix A). However, the stationary state probabilities of the HSMMs are very similar to those obtained for the HMMs (cf. Web Appendix A), and as a consequence of this and of the short time the whales are in view on the aerial survey, the resulting estimates of $1/p$ are close to the HMM-based estimates. They differ by only 7%: 4.62 when HMMs are used to

model availability, compared to 4.94 when HSMMs are used. Given the small difference (in the context of a CV of some 32%) and the substantial increase in computation time associated with fitting the HSMM, we have not calculated bootstrap confidence intervals for $1/p$ using the HSMM.

Although we were not able to fit a HSMM in the case of the bowhead shipboard data, we anticipate that the use of HSMMs, as opposed to HMMs, can make a substantial difference to estimates of $1/p$ from shipboard surveys, where animals spend more time within detectable range. In this case it is likely more important to obtain the best possible model of the availability process. We therefore recommend checking the adequacy of the Markovian assumption by fitting HSMMs to availability time series before deciding which availability model to use for estimation of $1/p$.

8. Discussion

Unless animals are within detectable distance for little more than an instant, existing correction factor methods for dealing with availability bias result in biased estimation of detection probability and density. This is because they do not account properly for the facts that the survey is not instantaneous and that detection probability changes while animals are within detectable range. Bias is lower for aerial surveys than shipboard surveys because animals are within viewing range for shorter periods, but bias can still be substantial.

Our estimators are able to estimate density with little or no bias in the presence of stochastic animal availability by using a HMM or HSMM to model availability. Unlike the methods of Skaug and Schweder (1999), Okamura (2003), Okamura et al. (2003), our method uses both forward and perpendicular distance data. We found estimators without forward distance data to be quite heavily model-dependent and advise against their use when animals are not continuously available. We also found that estimators based on constant-hazard availability can be severely biased when the availability hazard is not constant.

Unlike the methods of Skaug and Schweder (1999), Okamura (2003), and Okamura et al. (2003), our methods are based on discretizing time. Our discrete-time formulation allows us to take advantage of the considerable body of HMM and HSMM theory and software for inference. While the actual survey process operates in continuous time, availability data are usually available in discrete time. In any case, continuous time can be approximated arbitrarily closely by finer temporal discretization. And because there are errors in timing and locating detections, there is likely little gain from having a model with temporal resolution that is very much finer than the size of these errors.

Our model is similar in concept to the model of Okamura et al. (2012) although their model is formulated in continuous time. By using discrete-time models for both processes, we are able to take advantage of HMM and HSMM methods both for flexible modelling of the availability process and for efficient integration of the detection and availability processes. Our model accommodates intermittent availability (in which animals alternate between being available and unavailable continuously for some times), instantaneous availability, and a mixture of the two. The model of Okamura et al. (2012) is

designed for instantaneous availability processes only. Being parametric, our model is also able to deal with situations in which availability time series are unavailable but estimates of key availability model parameters (such as expected durations of available and unavailable periods) are available, whereas the model of Okamura et al. (2012) requires time series of availability. In short, our model provides a flexible, extendable alternative to the model of Okamura et al. (2012) for dealing with availability bias on single-platform LT surveys.

The methods of this paper have modest data demands relative to existing methods. Like all methods for dealing with availability bias on single-platform LT surveys, they require independent data on animal availability. To model the availability process reliably, moderately long series of availability data are needed (at least some tens of records long, preferably more). The only additional demand on the LT surveys themselves is that forward distances (or times) be recorded, and on many surveys these would be recorded anyway.

9. Supplementary Materials

Web Appendices referenced in Sections 5 and 7 are available with this paper at the Biometrics website on Wiley Online Library, together with R functions and example code for fitting models when both forward and perpendicular distance data are available.

References

Buckland, S. T., Anderson, D. R., Burnham, K. P., Laake, J. L., Borchers, D. L., and Thomas, L. J. (2001). Introduction to distance sampling, Oxford: Oxford University Press.

Conn, P. B., Johnson, D. S., London, J. M., and Boveng, P. (2012). Accounting for missing data when assessing availability in animal population surveys. *Methods in Ecology and Evolution* **3**, 1039–1046.

Forcada, J., Gazo, M., Aguilar, A., Gonzalvo, J., and Fernandez-Contreras, M. (2004). Bottlenose dolphin abundance in the NW Mediterranean: addressing heterogeneity in distribution. *Marine Ecology Progress Series* **275**, 275–287.

Gomez de Segura, A., Crespo, E. A., Pedraza, S. N., Hammond, P. S., and Raga, J. A. (2006). Abundance of small cetaceans in waters of the Central Spanish Mediterranean. *Marine Biology* **150**, 149–160.

Heide-Jørgensen, M. P., Laidre, K., Borchers, D. L., Samarrra, F., and Stern, H. (2007). Increasing abundance of bowhead whales in West Greenland. *Biology Letters* **3**, 577–580.

Innes, S., Heide-Jørgensen, M. P., Laake, J. L., Laidre, K. L., Cleator, H. J., Richard, P., and Stewart, R. E. A. (2002).

Surveys of belugas and narwhals in the Canadian High Arctic in 1996. *NAMMCO Scientific Publications* **4**, 169–190.

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

Laake, J. L., Calambokidis, J. C., Osmek, S. D., and Rugh, D. J. (1997). Probability of detecting harbor porpoise from aerial surveys: estimating g(0). Journal of Wildlife Management **61**, 63–75.

Laidre, K., Heide-Jørgensen, M. P., and Nielsen, T. (2007). Role of bowhead whale as a predator in West Greenland. *Marine Ecology Progress Series* **346**, 285–297.

Langrock, R. and Zucchini, W. (2011). Hidden markov models with arbitrary state dwell-time distributions. *Computational Statistics and Data Analysis* **55**, 715–724.

Marsh, H. and Sinclair, D. F. (1989). Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *Journal of Wildlife Management* **53**, 1017–1024.

McLaren, I. A. (1961). Methods of determining the numbers and availability of ring seals in the Eastern Canadian Arctic. *Arctic* **14**, 162–175.

Okamura, H. (2003). A line transect method to estimate abundance of long-diving animals. *Fisheries Science* **69**, 1176–1181.

Okamura, H., Kitakado, T., Hiramatsu, K., and Mori, M. (2003). Abundance estimation of diving animals by the double-platform line transect method. *Biometrics* **59**, 512–520.

Okamura, H., Minamakawa, S., and Kitakado, T. (2006). Effect of surfacing patterns on abundance estimates of long-diving animals. *Fisheries Science* **72**, 631–638.

Okamura, H., Minamikawa, S., Skaug, H. J., and Kishiro, T. (2012). Abundance estimation of long-diving animals using line transect methods. *Biometrics* **68**, 504–513.

Pollock, K. H., Marsh, H. D., Lawler, I. R., and Alldredge, M. W. (2006). Estimating animal abundance in heterogeneous environments: An application to aerial surveys for dugongs. *Journal of Wildlife Management* **70**, 255–262.

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

Zucchini, W. and MacDonald, I. (2009). Hidden Markov Models for Time Series: An Introduction Using R. Chapman & Hall/CRC.