A model for digital aerial surveys of marine mammals without individual identification

D. L. Borchers[1,*], P. Nightingale[1], B.C. Stevenson[2], and R.M. Fewster[2]

[1]University of St Andrews, Centre for Research into Ecological and Environmental Modelling,

St Andrews, Fife, UK

[2]Department of Statistics, University of Auckland, Private Bag 92019,

Auckland, New Zealand

*email: dlb@st-andrews.ac.uk

SUMMARY: Yadda, yadda, yadda, ...

KEY WORDS: Keywords: Double-observer, Mark-recapture, line transect, availability bias, movement model, Poisson process

This paper has been submitted for consideration for publication in *Biometrics*

1. Introduction

Aerial surveys of marine mammals are generally much cheaper than shipboard surveys. We anticipate that aerial surveys with human observers will increasingly be replaced by unmanned aerial vehicle (UAV) surveys with digital video or still cameras as observers. Such surveys present some new statistical challenges. In this paper we address these challenges and develop a method of estimating animal density from a UAVs carrying cameras.

Animals are within detectable range of an observer on an aircraft for only a short time, because aircraft move fast. Some animals are missed from an aircraft because they are underwater and unobservable while within detection range. This is known as the "$g(0) < 1$" problem in the line transect literature **add citation** and the bias that results from ignoring this problem is often called "availability bias" **add citation**. UAVs with digital cameras will tend to miss more animals than for aircraft with humans because the cameras typically have much smaller fields of view: perhaps 100m either side of the survey transect line, compared with 1000m or more for human observers.

Mark-recapture distance sampling (MRDS) methods **add citation** are commonly used to deal with the "$g(0) < 1$" problem. These methods are a variety of capture-recapture method, in which two independently searching observers constitute two "capture occasions" and animals detected by both constitute "recaptures".

MRDS methods on aerial surveys suffer from the fact that two observers on the same aircraft search the same patch of sea at the same time so that animals tend to be available or unavailable to both observers at the same time. This induces positive correlation in the detection events between the observers, and negative bias in density estimates **add citation**. (For those familiar with capture-recapture models, you can think of the availability state of an animal as an individual-level latent variable: when there is unmodelled individual heterogeneity on capture-recapture surveys, density estimates are biased.)

There are two ways in which this problem has been dealt with. The first is to model the latent process, i.e., the availability process. For example, **(citations)** modelled latent availability using a hidden Markov model, while **(citations)** modelled it using a Markov modulated Poisson processes. The second way is to separate the times that the two observers search the same region, either by having one observer search far ahead of the aircraft and the other close to it, or by using two aircraft in tandem, or having a single aircraft circle back over its path a second time (see Hiby and Lovell, 1998, (need to add others too?)).

However, separating the observers in this way has a cost: it introduces uncertainty into capture histories.

MRDS methods require that every detected individual be identified as having been detected by observer 1 only (capture history (1,0)), observer 2 only (capture history (0,1)), or both observers (capture history (1,1)). Because marine mammals are generally not individually identifiable from a fast-moving aerial survey platform, it is often not possible to obtain these capture histories without error.

The greater the time separation between observers, the more severe is the problem of errors in capture histories, but the less severe the problem of dependence in detections.

Separating UAV "observers" (cameras) by using UAVs in tandem or circling back on themselves is much less feasible than doing the same with human observers on aircraft. This is because field conditions are typically such that it is not possible for the aircraft doing the second pass to follow exactly the path of the aircraft doing the first pass. When observer field of view is around 2,000m wide (as is not unusual with human observers), having a second pass that deviates by 200m from the path of the first pass still leaves considerable overlap between the regions surveyed by each observer. It would leave no overlap at all for UAVs with a field of view 200m wide.

We seek a method that allows a single UAV to carry two cameras. This limits the extent to

which the two observers can be separated in time and requires that we deal with dependence between detections resulting from cameras passing in quick succession. We develop a method that does not require capture histories to be assigned to any individuals and deals with dependence by modelling the availability process.

Our method is similar to that of Stevenson et al. (2018) in that it has the same data requirements and does not require capture histories. It differs in that we do inference by maximum likelihood rather than using an approximation to the Palm likelihood. Unlike the method of Stevenson et al. (2018), our method becomes computationally infeasible when there are very many plausible pairings of detections by different observers. It is also similar to that of Hiby and Lovell (1998), which suffers from the same computational restriction. We investigate its utility for aerial UAV surveys and compare it to the method of Stevenson et al. (2018).

2. The model

2.1 *The movement model*

Two observers move along a transect line, one behind the other moving at v m/s, separated by a distance vl. Animals may move between the time that the first and second observers pass them.

We model animal movement in time t, as a bivariate normal distribution with mean $(0,0)$ and variance $\Sigma(t) = \sigma^2 t \boldsymbol{I}$, where $\boldsymbol{I}$ is a 2×2 identity matrix. This model is consistent with movement that is Brownian motion. It is straightforward to introduce directionality by replacing the zero off-diagonal elements of $\boldsymbol{I}$ with a correlation parameter ρ, although we do not do that here.

2.1.1 *Forward movement.* In Appendix 8 we show that when the two observers are moving at a speed v separated by time lag l, the pdf of the time between the first and second observers

encounter an animal (T), is

$$f_T(t) \;=\; \frac{vl \exp\left\{-\frac{v^2(l-t)^2}{2\sigma^2 t}\right\}}{\sqrt{2\pi\sigma^2 t^3}}. \tag{1}$$

The animal will have moved a forward distance $v(t-l)$ in this time, with negative distance being movement in the opposite direction to that of the observers, and positive distance being in the same direction.

2.2 *Availability models*

There are two reasons that animals that are within the survey strip at some point between the passing of the first and second observers may not be available for detection. The first is that they are invisible to the observers because they are diving. The second is that they are invisible because they are not within the strip at the time they are passed. We construct models for each of these processes below.

2.2.1 *The up/down availability model.*   We assume animals unavailability due to diving is governed by a two-state continuous time Markov process with transition intensity matrix parameters such that the time in state 1 (the near-surface state) is an exponential random variable with expected value κ, and the time in state 2 (the diving state) is an exponential random variable with expected value $\tau - \kappa$, where τ is the expected dive cycle duration:

The Markov transition rate matrix $\boldsymbol{Q}$ is

$$\boldsymbol{Q} \;=\; \begin{pmatrix} -\frac{1}{\kappa} & \frac{1}{\kappa} \\[2mm] \frac{1}{\tau-\kappa} & -\frac{1}{\tau-\kappa} \end{pmatrix} \tag{2}$$

from which it follows that the stationary distribution of the Markov chain is

$$\underline{\pi} \;=\; \left(\frac{\kappa}{\tau}, \frac{\tau-\kappa}{\tau}\right). \tag{3}$$

The transition probability matrix for transitions between states at time separation t is $\boldsymbol{U}(t) = \exp(\boldsymbol{Q}t)$.

2.2.2 *The in/out availability model.* Suppose both observers search a distance w either side of the transect line. We consider a strip of half-with $b > w$, such that there is zero, or neglibly small probability that an animal outside this strip when the first observer passes, moves to within w of the transect in the time between the two observers coming abeam of the animal. For each animal that is within b of the transect when observer 1 passes it, we define the binary random variable $Z(t)$ to be 1 if the animal is within w of the transect at time t after the first observer passes it, and is zero otherwise. For notational convenience, we define $\bar{Z}_j = 1 - Z_j$. Under the above movement model for $X(t)$, the probability $\mathbb{P}(Z(t) = 0|Z(0) = 1)$, that an animal moves from inside the searched strip to outside it in an interval t, and the probability $\mathbb{P}(Z(t) = 1|Z(0) = 0)$, that it moves from outside the searched strip to inside it, are

$$p_{IO}(t) = \mathbb{P}(Z(t) = 0|Z(0) = 1) \quad = \quad \frac{1}{w} \int_0^w \Phi(x - w; \sigma^2 t) + \Phi(-x - w; \sigma^2 t) dx \qquad (4)$$

$$p_{OI}(t) = \mathbb{P}(Z(t) = 1|Z(0) = 0) \quad = \quad \frac{1}{b - w} \int_w^b \Phi(x - w; \sigma^2 t) - \Phi(-x - w; \sigma^2 t) dx \quad (5)$$

where $\Phi(\cdot; \sigma^2 t)$ is the cumulative distribution function of a normal random variable with mean zero and variance $\sigma^2 t$.

We model the movement in and out of the survey strip as a two-state Markov process with transition probabilities at time t after being passed by the first observer that are given by Eqns (4) and (5):

$$\boldsymbol{M}(t) \quad = \quad \begin{pmatrix} 1 - p_{IO}(t) & p_{IO}(t) \\ p_{OI}(t) & 1 - p_{OI}(t) \end{pmatrix} \qquad (6)$$

2.2.3 *The combined availability model.* The possibility of being in the survey strip or outside it and on the surface or below it, gives rise to four states that animals can occupy: (up and in), (up and out), (down and in) and (down and out), which we now number 1 to 4

in that order. Assuming that being up or down is independent of being in or out, transitions between these states at a time separation t is governed by the transition probability matrix $\boldsymbol{\Gamma}(t) = \boldsymbol{U}(t) \otimes \boldsymbol{M}(t)$.

Assuming uniform animal distribution with respect to the transect line at the time that the first observer passes animals, the probability that an animal that is within b of the line, is in the searched strip is w/b. Combined with the stationary distribution π, this gives the following initial state distribution:

$$\boldsymbol{\delta} = \left(\frac{\kappa}{\tau} \frac{w}{b}, \; \frac{\kappa}{\tau} \left[1 - \frac{w}{b} \right], \; \left[\frac{\tau - \kappa}{\tau} \right] \frac{w}{b}, \; \left[\frac{\tau - \kappa}{\tau} \right] \left[1 - \frac{w}{b} \right] \right). \tag{7}$$

3. A Markov modulated Bernoulli process model

We assume that the probability that an animal is in each state at the time the first observer passes over it is given by the stationary distribution $\boldsymbol{\delta}$, and hence that its state distribution after a waiting time t, when the second observer passes it, is $\boldsymbol{\delta}\boldsymbol{\Gamma}(t)$.

Each observer records binary variable $X_{ij}(t)$, which is 1 when the animal i is detected by observer j at time t and is zero otherwise. We model $X_{ij}(t)$ as a state-dependent Bernoulli random variable with parameter $p_j(c) = \Pr(X_{ij}(t) = 1 | C_i(t) = c)$ where $C_i(t)$ is the state of animal i at time t and $c \in (1, 2, 3, 4)$.

It follows that $X_{ij}(t_{ij})$ $(j = 1, 2)$ are observations from a Markov modulated Bernoulli process (MMBP) at times t_{i1} and t_{i2}. It is convenient to define the time at which the first observer passes animal i (t_{i1}) to be zero, so that the time at which the second observer passes is t_{i2} (t in the development of the previous section).

It is convenient for the hidden Markov model formulation to arrange detection probabilities

in a detection probability matrix. For observer j, this matrix is

$$
\boldsymbol{P}\{x_{ij}(t)\} \;=\; \begin{array}{cccc} \text{up,in} & \text{up,out} & \text{down,in} & \text{down,out} \end{array}
$$

$$
\boldsymbol{P}\{x_{ij}(t)\} \;=\; \begin{pmatrix} \text{Bern}(x_{ij}(t); p_j(1)) & 0 & 0 & 0 \\ 0 & 1 - x_{ij}(t) & 0 & 0 \\ 0 & 0 & \text{Bern}(x_{ij}(t); p_j(3)) & 0 \\ 0 & 0 & 0 & 1 - x_{ij}(t) \end{pmatrix} \tag{8}
$$

where $\text{Bern}(x_{ij}(t); p_j(c) \equiv p_j(c)^{x_{ij}(t)}[1 - p_j(c)]^{1-x_{ij}(t)})$. In the above matrix we allow a more general case than we deal with below, in which it is possible to detect an animal that is within the survey strip and not in the near-surface state. In what follows we assume that only animals in the near-surface state can be detected so that $p_j(3) = 0$.

Conditional on t_{i2}, and remembering that t_{i1} is defined to be 0, the probability of observing $\{x_{i1}(0), x_{i2}(t_{i2})\}$, is

$$
p\{x_{i1}(0), x_{i2}(t_{i2}) | t_{i2}\} \;=\; \boldsymbol{\delta P}\{x_{i1}(0)\}\boldsymbol{\Gamma}(t_{i2})\boldsymbol{P}\{x_{i2}(t_{i2})\}\boldsymbol{1} \tag{9}
$$

where $\boldsymbol{1}$ is a column vector of ones. The joint probability of the second observer passing animal i at t_{i2} and observing $\{x_{i1}(0), x_{i2}(t_{i2})\}$ is $p\{x_{i1}(0), x_{i2}(t_{i2}|t_{i2})\}f_T(t_{i2})$.

For brevity, we now write $p\{x_{i1}(0) = 0, x_{i2}(t_{i2}) = 1|t_{i2}\}$ as $p_{01}(t_{i2})$, we write $p\{x_{i1}(0) = 1, x_{i2}(t_{i2}) = 0|t_{i2}\}$ as $p_{10}(t_{i2})$, and we write $p\{x_{i1}(0) = 1, x_{i2}(t_{i2}) = 1\}$ as $p_{11}(t_{i2})$.

Writing the three observable capture histories as $\omega_1 = (01), \omega_2 = (10), \omega_3 = (11)\}$, we can write the probability of observing each of these as

$$
\tilde{p}_{\omega_k} = E_t[p_{\omega_k}(t)] \;=\; \int p_{\omega_k}(t)f_T(t)dt. \tag{10}
$$

for $k = 1, 2, 3$.

4. The survey model

We assume that the number and locations of animals along the transect at the time of the first or second observer passing are governed by a nonhomogeneous Poisson process (NHPP) with rate paramter $D(s)$ at location s. We define the observed location of detected animal i to be its location at the time the first observer passes it, or if it is not detected by the first observer, then at the time that the second observer passes it.

Let $\boldsymbol{s}_\omega = (\boldsymbol{s}_{\omega_1}, \boldsymbol{s}_{\omega_2}, \boldsymbol{s}_{\omega_3})$ be the observed locations along the transect of detections with capture history $\omega_1 = (01), \omega_2 = (10), \omega_3 = (11)$. These arise from thinned NHPPs with thinning probabilities $\tilde{p}_{\omega_k}$ $(k = 1, 2, 3)$, so that for a survey along transects of total length L, the pdf of locations $\boldsymbol{s}_\omega$ is

$$f(\boldsymbol{s}_\omega) = \left[\prod_{i=1}^{n_\omega} \tilde{p}_\omega D(s_{\omega i}) \right] e^{-\int_0^L \tilde{p}_\omega D(s) ds}. \tag{11}$$

where n_ω is the number of individuals with capture history ω.

In the case of $\omega = (11)$, we also observe the time $t_{(11)i}$ between the two observers passing the ith animal with capture history (11), for $i = 1, \ldots, n_{(11)}$. (We do not observer these for capture histories (01) and (10).) The pdf for the ith of these is the conditional pdf of time between passing, given that the capture history is (11). This is

$$f_{T|11}(t_{(11)i}, \omega_i = 11) = \frac{f_T(t_{(11)i}) p_{11}(t_{(11)i})}{\tilde{p}_{11}}. \tag{12}$$

We assume that the $t_{(11)i}$s are independent $(i = 1, \ldots, n_{(11)})$.

Given the capture histories $\boldsymbol{\omega}$, the likelihood is therefore as follows (with model parameters suppressed for simplicity of presentation):

$$\mathcal{L}(\boldsymbol{s}|\boldsymbol{\omega}) = \left\{ \prod_\omega \left[\prod_{i=1}^{n_\omega} \tilde{p}_\omega D(s_{\omega i}) \right] e^{-\int_0^L \tilde{p}_\omega D(s) ds} \right\} \prod_{i=1}^{n_{(11)}} f_{T|11}(t_{(11)i}|\omega_i = 11)$$

$$= e^{-\int_0^L \tilde{p}_\cdot D(s) ds} \left\{ \prod_\omega \tilde{p}_\omega^{n_\omega} \prod_{i=1}^{n_\omega} D(s_{\omega i}) \right\} \prod_{i=1}^{n_{(11)}} f_{T|11}(t_{(11)i}|\omega_i = 11) \tag{13}$$

where $\boldsymbol{\omega}$ is the set of capture histories of each detection, $\prod_\omega$ means the product over $\omega \in \{(10), (01), (11)\}$, $\tilde{p}_\cdot = \sum_\omega \tilde{p}_\omega$, and $\boldsymbol{s} = (\boldsymbol{s}_{(10)}, \boldsymbol{s}_{(01)}, \boldsymbol{s}_{(11)})$.

4.1 Homogeneous Poisson case

In the homogeneous Poisson case, when $D(s) = D$

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{s}|\boldsymbol{\omega}) &= e^{-\tilde{p}.DL}D^n \left(\prod_{\omega \in \{10,01,11\}} \tilde{p}_\omega^{n_\omega} \right) \prod_{i=1}^{n_{(11)}} f_{T|11}(t_{(11)i}|\omega_i = 11) \\
&= e^{-\tilde{p}.DL}D^n \left(\prod_{\omega \in \{10,01\}} \tilde{p}_\omega^{n_\omega} \right) \prod_{i=1}^{n_{(11)}} f_T(t_{(11)i})p_{11}(t_{(11)i}) \qquad (14)
\end{aligned}
$$

where $n = \sum_\omega n_\omega$.

4.2 Model parameters

The model has four kinds of parameters:

Density parameters: In the case of the homogenous Poisson process there is one parameter, θ such that $D = e^\theta$. When density varies with some covariates, θ is replaced by a linear predictor involving a parameter vector.

Dive cycle parameters: The two-state dive cycle model described above involves the mean time in the near-surface state, μ_1, and the mean dive cycle length, τ, which are linked to parameters γ_1 and γ_τ via a log link: $\mu_1 = e^{\gamma_1}$ and $\tau = e^{\gamma_\tau}$.

Movement parameters: The movement model has one parameter, σ, which we model using a log link: $\sigma = e^\phi$.

Detection parameters: Assuming that animals are only detectable when in state $c = 1$ (up,in), we have two Bernoulli parameters to model: $p_1(1)$ and $p_2(1)$. These can be modelled using logit link functions and if the observers are identical digital detectors it may be reasonable to assume these two probabilities are identical, i.e. $p_1(1) = p_2(1) = p = e^\beta/(1+e^\beta)$.

As is the case with density parameters, the other three kinds of parameters can be modelled as functions of suitable covariates by replacing the relevant scalar parameter on the link scale with a suitable linear predictor involving the covariates.

We focus on the constant density model in the rest of this paper. With no covariates, it has five parameters, which we write as $\boldsymbol{\theta}^* = (\theta, \gamma_1, \gamma_\tau, \phi, \beta)$. Stevenson et al. (2018) show

that not all of these parameters are identifiable. They assume that $p = e^\beta/(1 + e^\beta) = 1$, and we do the same. This is reasonable for digital aerial surveys conducted in low sea states if we define the near-surface state to be "breaking the surface" (a state that is easily observed). The field of view of digital cameara is such that objects towards the periphery of the image are as easily detected as objects in the centre of the image so that a detection function of the sort used on line transect surveys, i.e., one that drops off with distance from the line, is not needed.

Stevenson et al. (2018) show that even in this case, only two of $(\theta, \gamma_1, \gamma_\tau)$ are identifiable. Like Stevenson et al. (2018) and Hiby and Lovell (1998), we specify the mean dive cycle duration, $\tau = e^{\gamma_\tau}$ on the basis of estimates external to the survey. In what follows, we assume certain detection of animals that are on the surface when in view, use external estimates to set τ, and estimate the density, D, the mean time in surface state, μ_1, and the Brownian movement parameter, σ, That is, we estimate $\boldsymbol{\theta} = (\theta, \gamma_1, \phi))$. We do this using the locations along the transect line of detections by the first observer and the locations along the transect line of detections by the second observer (and known observe speed) without assigning capture histories to any individual.

5. Unknown capture histories

If we knew which animals had capture histories (10), (01) and (11) (i.e., we were given $\boldsymbol{\omega}$), we could estimate the model parameters by maximising the likelihood Equation (14) with respect to $\boldsymbol{\theta}$.

However, we don't know which animals have which capture histories. Following Hiby and Lovell (1998) we marginalise over the (unknown) capture histories by summing the likelihoods for all possible capture history combinations. We refer to each combination as a "pairing" since once the pairs of detections with capture history (11) have been determined, this

determines the capture histories of all detections – because we know for each of the other detections which one of the two observers made the detection.

If we call the mth set of pairings $\boldsymbol{\omega}^{(m)}$, we can consider $\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{s}|\boldsymbol{\omega}^{(m)})$ to be the conditional likelihood, given $\boldsymbol{\omega}^{(m)}$, and we obtain the unconditional likelihood for the parameters $\boldsymbol{\theta}$ as

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{s}) \;\; = \;\; \sum_{m=1}^{M} \mathcal{L}(\boldsymbol{\theta}; \boldsymbol{s}|\boldsymbol{\omega}^{(m)}) \tag{15}$$

where M is the number of possible pairings.

While this likelihood is easy to write down, it is challenging to evaluate because we need to enumerate all possible $\boldsymbol{\omega}^{(m)}$s, and for any but very small sample sizes, the number M of possible pairings is very large ($M >> 10^9$ for even moderate sample sizes).

We use constraint programming methods (see Section 5.2) to efficiently find all possible pairings, and subdivision of $\boldsymbol{s}$ into subsets between which pairings of detections by different observers is impossible, to reduce the number of possible pairings.

5.1 *Subdivision of $\boldsymbol{s}$*

We subdivide $\boldsymbol{s}$ by "cutting" the transect line immediately after detections by observer j for which the distance to the next detection by observer $(3 - j)$ is greater than a maximum possible distance that an animal could have moved between the two observers passing over it (d_{max}). This distance d_{max} must be decided by the user in the light of knowledge of the target species movement speed and behaviour. Users can investigate what d_{max} is reasonable by doing inference at a range of plausible values to find where estimates become insensitive to d_{max}. The cost of setting d_{max} too large is in computational speed; the cost of setting d_{max} too small is positive bias in estimation of D since too small a d_{max} will result in some animals with true capture history (11) being assigned capture history (01) or (10).

Having divided the transect line into C segments, we enumerate the possible parings $\boldsymbol{\omega}^{(m_c)}$ for segments $c = 1, \ldots, C$ (with M_c possible pairings in segment c) and calculate

the likelihood as

$$\mathcal{L}(\boldsymbol{\theta}; \boldsymbol{s}) \;\; = \;\; \prod_{c=1}^{C} \sum_{m_c=1}^{M_c} \mathcal{L}(\boldsymbol{\theta}; \boldsymbol{s}|\boldsymbol{\omega}^{(m_c)}) \tag{16}$$

When d_{max} is substantially smaller than most of the distances between detections by different observers, segmentation can lead to a massive reduction in computation time, making an otherwise computationally infeasible likelihood evaluation quite feasible. If d_{max} is not substantially smaller than a large fraction of the distances between detections by different observers, and sample sizes are not very small, computation of the likelihood may be so time consuming as to render the method practically useless.

5.2 *Enumerating all $\boldsymbol{\omega}^{(m)}s$*

To enumerate the possible pairings within one segment efficiently we define a simple *constraint satisfaction problem* (CSP) (Russell and Norvig, 2014, Chapter 6). A CSP is a triple $\mathcal{X} = \langle \mathcal{V}, \mathcal{D}, \mathcal{C} \rangle$. The CSP $\mathcal{X}$ has a set of decision variables $\mathcal{V}$, each of which has a set of possible values that it may take, called its *domain*, where $\mathcal{D}(x)$ is the domain of $x \in \mathcal{V}$. In addition a CSP has a set of constraints $\mathcal{C}$ that restrict the combinations of values that may be taken by the variables. A constraint $c \in \mathcal{C}$ is a relation defined on a set of variables scope$(c) \subseteq \mathcal{V}$. A *solution* is an assignment of values to variables such that each variable is assigned a value from its domain, and all constraints are satisfied.

We define a CSP for a segment as follows. Two detections by different observers may be paired iff the distance between them is less than or equal to d_{max}. For each set $\{i, j\}$ of two observations that may be paired we define one decision variable $b_{i,j}$ with domain $\{0, 1\}$. Variable $b_{i,j}$ is equal to 1 in a solution iff the two observations are paired.

Suppose we have two sets, $s_1 = \{i, j\}$ and $s_2 = \{k, l\}$, where $s_1 \cap s_2 \neq \emptyset$, i may be paired with j, and k may be paired with l. The two sets cannot both be paired simultaneously because they share an observation. In all such cases we add the constraint $b_{i,j} = 0 \lor b_{k,l} = 0$ to prevent both sets being paired simultaneously.

We use a backtracking search procedure with forward checking (Russell and Norvig, 2014, Chapter 6) to enumerate all solutions to the CSP. The set of solutions to the CSP correspond one-to-one to the set of valid pairings within the segment. When a solution is found, the part of the likelihood pertaining to that pairing is calculated, avoiding the need to store the set of pairings and allowing efficient calculation of $\sum_{m_c=1}^{M_c} \mathcal{L}(\boldsymbol{\theta}; \boldsymbol{s}|\boldsymbol{\omega}^{(m_c)})$.

5.3 *Interval estimation*

We estimated the variances of parameters using the inverse of the Hessian obtained in the fitting process. Confidence intervals for the parameters D, σ and μ_1 were obtained using the inverse log transformation of confidence intervals for θ, ϕ and γ_1, assuming normality of the maximum likelihood estimators of these parameters.

6. Application

We developed our estimation method in anticipation of digital aerial survey data becoming widely used, but do not currently have any digital aerial survey data of marine mammals. We therefore estimate density from the same data as were used by Stevenson et al. (2018) – data from an aerial survey with human observers, from the periods when the aircraft circled back over its transect (in which the second "observer" is the second pass) after a lag of $k = 248$ seconds.

Like Stevenson et al. (2018), we use a searched strip half-width of $W = 0.125$ km and a "buffer" of $b = 2$ km beyond this, beyond which we assume no animal could enter the searched strip between the fist and second observer passing it. They obtained the following estimates (95% confidence intervals in brackets): $\hat{D} = 1.05$ (0.84, 1.60) pods per km2, $\hat{\sigma}_{palm} = 0.15$ km, and the expected proportion of time in the surface state $\hat{\alpha} = 0.86$ (0.56, 1.00)). The corresponding estimates using our method are $\hat{D} = 1.24$ (0.97,1.6) pods per km2, $\hat{\sigma}_{248} =$

$\hat{\sigma}\sqrt{248} = 0.18$ (0.15,0.22) km, and $\hat{\alpha} = \hat{\mu}_1/\tau = 0.73$ (0.57,0.94). The estimate $\hat{\sigma}_{248} = 0.18$ corresponds to a mean rate of movement of 0.82 m/s.

Stevenson et al. (2018) estimated the coefficients of variation (CVs) of $\hat{D}$, $\hat{\sigma}_{palm}$ and $\hat{\alpha}$ to be 19%, 16% and 13%, respectively. The corresponding estimated CVs from our method (involving $\hat{\sigma}_{248}$ instead of $\hat{\sigma}_{palm}$) are 13%, 10% and 13%, respectively.

The estimates from the two methods are broadly consistent, but as we can't evaluate their relative merits on the basis of a single survey with unknown density, we investigated the performance of our density estimator by simulation.

7. Simulation study

With lag $k = 0$, we can only estimate the proportion of animals that are missed if we know, or can separately estimate, the proportion of time that animals are available. If the lag is long enough that there is no correlation between the availability of an animal to each observer, then the probability of recapture is the product of the probability of being available and the probability of detection, given availability (see Hayes and Buckland, 1983), and we do not need an availability model to estimate animal density. If we also assume certain detection given availability ($p = 1$), then we can estimate the proportion of time that animals are available ($\alpha = \mu_1/\tau$), without an availability model. Note that α does not determine $\mathbf{\Gamma}(k)$ – there is an infinite number of $\mathbf{\Gamma}(k)$s that generate the same α. (For example animals could switch between being available and unavailable very quickly and be available a proportion α of the time on average, or they could switch very slowly and be available the same proportion of the time on average.)

We would like to survey with two video cameras on a single aircraft (manned or not) and in this case it may not be possible to separate the cameras' fields of view in the forward direction by a large enough distance to ensure independence between detections by each camera. We

are therefore interested in estimating density when the lag l is so short that detections are not independent between the two observers. If we neglect the correlation between detections in this case, estimates of density will be negatively biased. But our Markov process model for availability implicitly models the corelation, allowing the possibility of unbiased estimation of density with short lags.

Here we investigate the performance of our density estimator for short lags. We start by looking at the correlation between detections as a function of lag, and the extent to which short lags inflate the detection pobability of the second observer given detection/non-detection by the first observer. In all simulation scenarios we set $W = 0.125$ km, $b = 5 \times \sigma \sqrt{k}$ km, and total transect length the same as that in the application above, namely $L = 1,100$ km.

Our data consist of observations of Bernoulli random variables X_1 and X_2, each of which is 1 when animals are available and zero when they are not. The correlation between these variables when availability is governed by a two-state hidden Markov model with transition probability matrix $\mathbf{\Gamma}(k) = \exp(\boldsymbol{Q}k)$ with off-diagonal elements $\gamma_{12}(k)$ and $\gamma_{21}(k)$, is $\mathrm{corr}(X_1 X_2) = (1 - \gamma_{12}(1) - \gamma_{21}(1))^k$, while $\mathbb{P}(X_2 = 1 | X_1 = 1, k) = \gamma_{11}(k)$ and $\mathbb{P}(X_2 = 1 | X_1 = 0) = \gamma_{21}(k)$.

Figure **??** shows the correlation between the observers' detections, as a function of the lag (l) when $\alpha = 0.73$, $\tau = 100$ and $\sigma = 0.18/\sqrt{248}$. This is a scenario consistent with the estimates we obtained in the previous section, although for convenience we have standardised dive cycle length to be 100.. Figures **??** and **??** show the percentages by which $\mathbb{P}(X_2 = 1 | X_1 = 1, k)$ and $\mathbb{P}(X_2 = 1 | X_1 = 0)$ are greater than α for $\alpha = (20\%, 30\%, 40\%, 50\%, 60\%, 70\%, 80\%)$.

[Figure 1 about here.]

[Figure 2 about here.]

[Figure 3 about here.]

On the bases of these figures, we decided to simulate with lag l equal 10%, 20%, 50% and 80% of dive cycle length and for α equal to 10%, 20%, 50% and 80%. We do this for three values of σ, coresponding to mean animal speed of 1.5 m/s (the value estimated by Hiby and Lovell, 1998), 0.95 m/s (the value estimated by Westgate et al., 1995), and 0.65 m/s (the value estimated by Stevenson et al., 2018, ,which being lower than that which we estimated above, broadens the range of our simulations beyond that were we to use our estimate of σ). In all cases we simulate with true density $D = 1.24$, as estimated in the previous section using our method.

Simulation results are shown in Table 1.

[Table 1 about here.]

Figure 4 shows the relative bias and CV of the maximum likelihood estimator (MLE) of this paper, for each of the simulation scenarios, as a function of propotion of time avaialable (α) and lag (l). Figue 5 compares the relative biases and CVs of the MLE and Palm estimators for each of the simulation scenarios.

[Figure 4 about here.]

[Figure 5 about here.]

Tentative conclusions from this (need larger number simulations to comment further):

- The MLE is relatively unbiased, except when α is small, and in this case bias gets worse as lag l gets larger. (This seems weird to me - I'd have thought that bias is lowest when l is large! What is going on here?)
- Similarly, percentage CV is worst when α is small and lag l is big. (Again weird!)
- Both MLE and Palm estimators probably almost unbiased.

- The MLE seems a little more precise than the Palm estimator. Interpretation: Likelihood contains some information that is lost in the Palm likelihood approximation.

- (Want to include SCR estimator when do fuller set of simulations - to investigate information loss due to lack of recapture identities.)

8. Discussion

- Could estimate (D, p, σ) instead of (D, μ_1, σ), if unsure that see all on surface, but then need to assume know (μ_1, τ).

- Varying lag makes one more parameter identifiable (in principle at least). Would be good to expand on this some.

- Can put covariates into any of the parameters. *Does this make parameters identifiable? Varying lag is effectively this.*

- With known capture histories this is a kind of SCR transect search with a continuous time movement model and an availability model. Mention difference betwee this and SCR M_b model (and refer to Ben's paper)?

- If know capture histories, can estimate movement and dive cycle parameters.

- Slower moving observers: need to take account of first passage time (which ccr method does not).

- We provide a theoretical structure on which one can build more general modesl for digital survey data, without requiring recapture identification. We anticipate that this could result in very substantial estimation efficiency by allowing the estimation process to be automated. To automate the process we need only an adequate automatic identifier of the species in question in the video stream from each "observer" separately. If false negatives affect only the availability process (e.g. remove animals underwater and partially visible), the only cost of using a strict identification criterion in order to avoid false positives, is reduced sample size. If, however, false negatives affect the detection process (e.g. remove

some individuals because although they were as available as possible, a wave broke over them as the observer passed) then the assumption of $p = 1$ may be violated and bias may ensue. Surveying at different lags will in principle allow estimation of p (in additon to D, σ and μ_1) so that it will likely be possible to automate inference from digital surveys by using automated object identification criteria that are sufficiently strict so as to reduce the probability of false positives to virtually zero, providing that the survey involves effort and detections at more than one lag.

Received ?? 2018. Revised 20??

Appendix Derivation of $f_T(t)$

Note to self: need to simplify this - see Rachel's suggestions.

One can view the poblem of finding the pdf of the time between the first and second observers passing an animal as a the problem of finding the first passage time to a point at distance vl from the origin, of a particle following Brownian motion starting at the origin, with constant drift velocity v and diffusion parameter σ. We can write the process as $Y_t = \sigma B_t + vt$, where B is a Brownian motion and Y_t is the animal's displacement at time t.

We use the result (**citation**) that for $X_t = B_t + ct$, where B is a Brownian motion, a "boundary" at $a > 0$, and a constant drift rate $c \in \mathbb{R}$, the pdf of $T = \inf\{t : X_t = a\}$ is

$$f_T(t) = \frac{a \exp\left\{\frac{-(a-ct)^2}{2t}\right\}}{\sqrt{2\pi t^3}}. \tag{A.1}$$

In our case, we have the "boundary" (the second observer) at a distance vl from the starting position of the animal, approaching it at speed v, and the animal moving according to brownian motion with parameter σ, so that $Y_t = \sigma B_t + vt$. If we rescale distance to be in units of σ, i.e. $X_t = Y_t/\sigma$, we have $X_t = B_t + \frac{v}{\sigma}t$ and $a = vl/\sigma$. Hence the pdf of time to the second observer passing, given that the second observer passes a fixed point at a time l

later than the first, is

$$f_T(t) \quad = \quad \frac{\frac{vl}{\sigma}\exp\left\{\frac{-(vl/\sigma - vt/\sigma)^2}{2t}\right\}}{\sqrt{2\pi t^3}} \quad = \quad \frac{vl\exp\left\{\frac{-v^2(l-t)^2}{2\sigma^2 t}\right\}}{\sqrt{2\pi\sigma^2 t^3}}. \tag{A.2}$$

Appendix Relationship between σ and mean animal speed

We're considering 1-dimensional movement here. Let $X(t)$ be the location of the animal relative to its starting position after time t. Then, assuming it starts at $X(0) = 0$, $X(t) \sim N(0, \sigma\sqrt{t})$ and $Z(t) = X(t)/[\sigma\sqrt{t}] \sim N(0,1)$. Let $U(t) = \sqrt{Z(t)^2}$. Then[1] $U(t) \sim \chi(1)$, with expected value $E[U(t)] = \sqrt{2}\Gamma(1)/\Gamma(1/2) = 0.7978846$. Hence an animal moving according to Brownian motion, with parameter σ, is expected to move a distance 0.7978846σ per unit time, i.e. have an average speed of 0.7978846σ. Conversely, to model Brownian animal movement in which animals have an average speed of v, we need $\sigma = v/0.7978846$. For example, to model animal movement with average speed of 0.95m/s, as in Stevenson et al. (2018), we require $\sigma = 0.95/0.7978846$m/s, or $0.95/797.8846 \approx 0.00119$km/s.

References

Hayes, R. J. and Buckland, S. T. (1983). Radial-distance models for the line-transect method. *Biometrics* **39,** 29–42.

Hiby, L. and Lovell, P. (1998). Using aircraft in tandem formation to estimate abundance of harbour porpoise. *Biometrics* **54,** 1280–1289.

Russell, S. and Norvig, P. (2014). *Artificial Intelligence A Modern Approach, Third Edition.* Pearson.

Stevenson, B. C., Borchers, D. L., and Fewster, R. M. (2018). Cluster capture-recapture to account for identification uncertainty on aerial suveys of animal populations. *Biometrics* **??,** ???–???

---

[1]See https://math.stackexchange.com/questions/1059938/whats-the-expectation-of-square-root-of-chi-square-variable

Westgate, A. J., Read, A., Beggren, P., Koopman, H. N., and Gaskin, D. E. (1995). Diving behaviour of harbour porpoises, *Phocoena phocoena. Canadian Journal of Fisheries and Aquatic Sciences* **52,** 1064–1073.
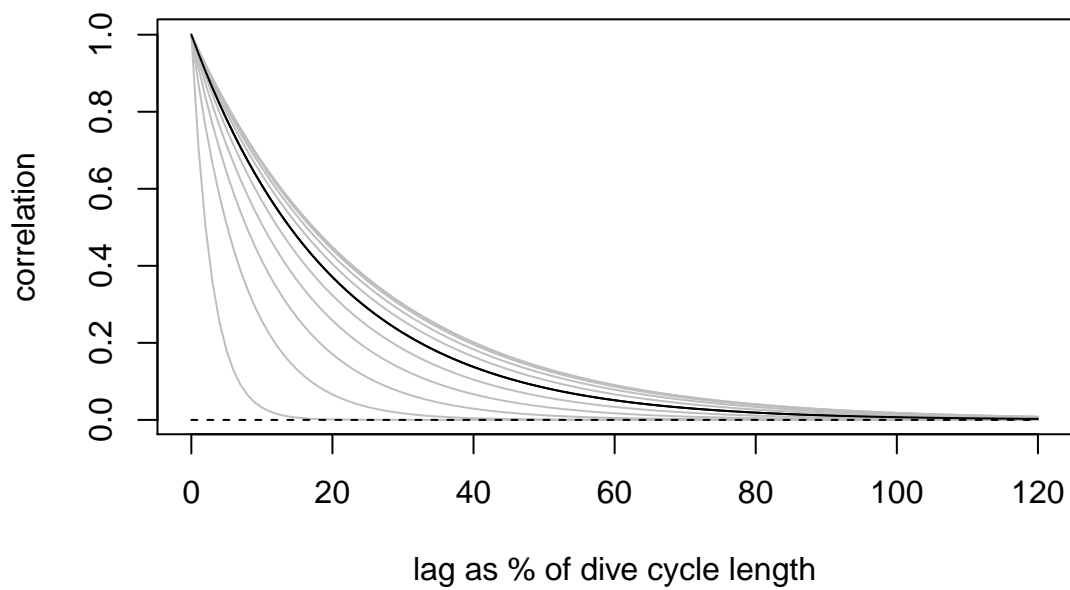
Figure 1. Correlation between detections by the two observes as a function of lag, for mean proportion of time available (α) and mean dive cycle duration of 100 seconds (dark line). Gray lines show the correlation for mean proportions of time available of $47\%, 52\%, 57\%, 62\%, 67\%, 77\%, 82\%, 87\%, 92\%$ and, 97% seconds. The black line is for $\alpha = 72\%$
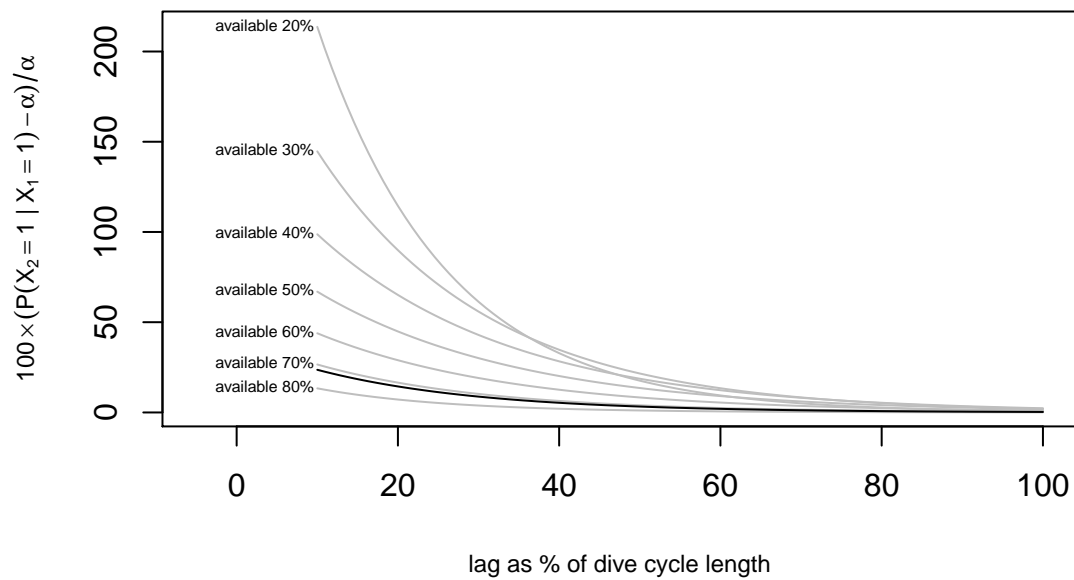
Figure 2. The percentage by which the probability of observer 2 detecting an animal given that observer 1 detected it, is greater than the proportion of time animals are available (α). This is shown for lags from 10 to 100 percent of the dive cycle length. The dark line is for $\alpha = 0.72$.

Figure 3. The percentage by which the probability of observer 2 detecting an animal given that observer 1 did not detect it, is greater than the proportion of time animals are available (α). This is shown for lags from 10 to 100 percent of the dive cycle length. The dark line is for $\alpha = 0.72$.
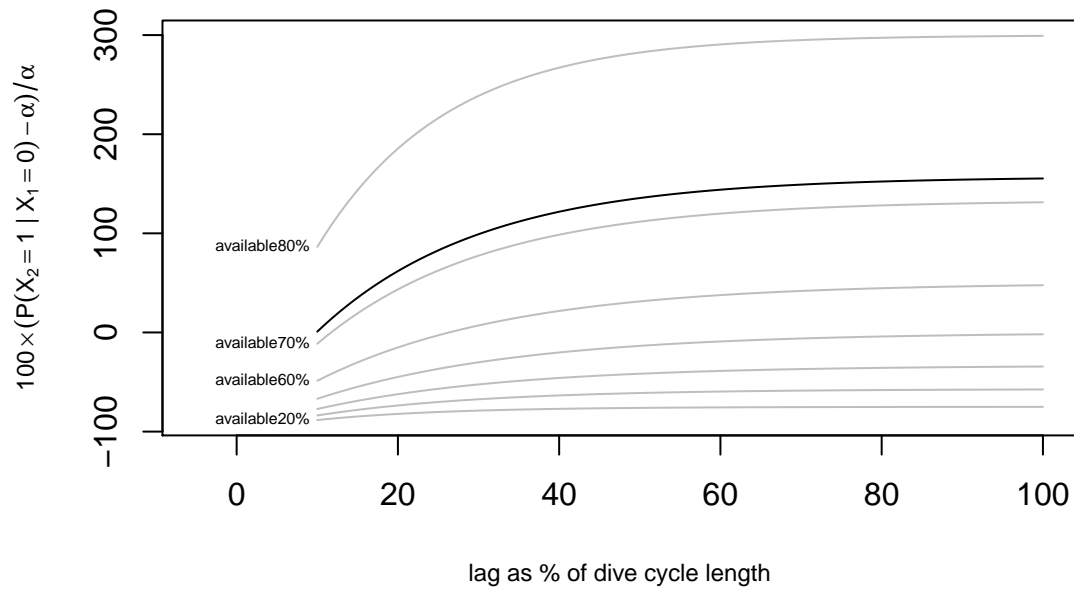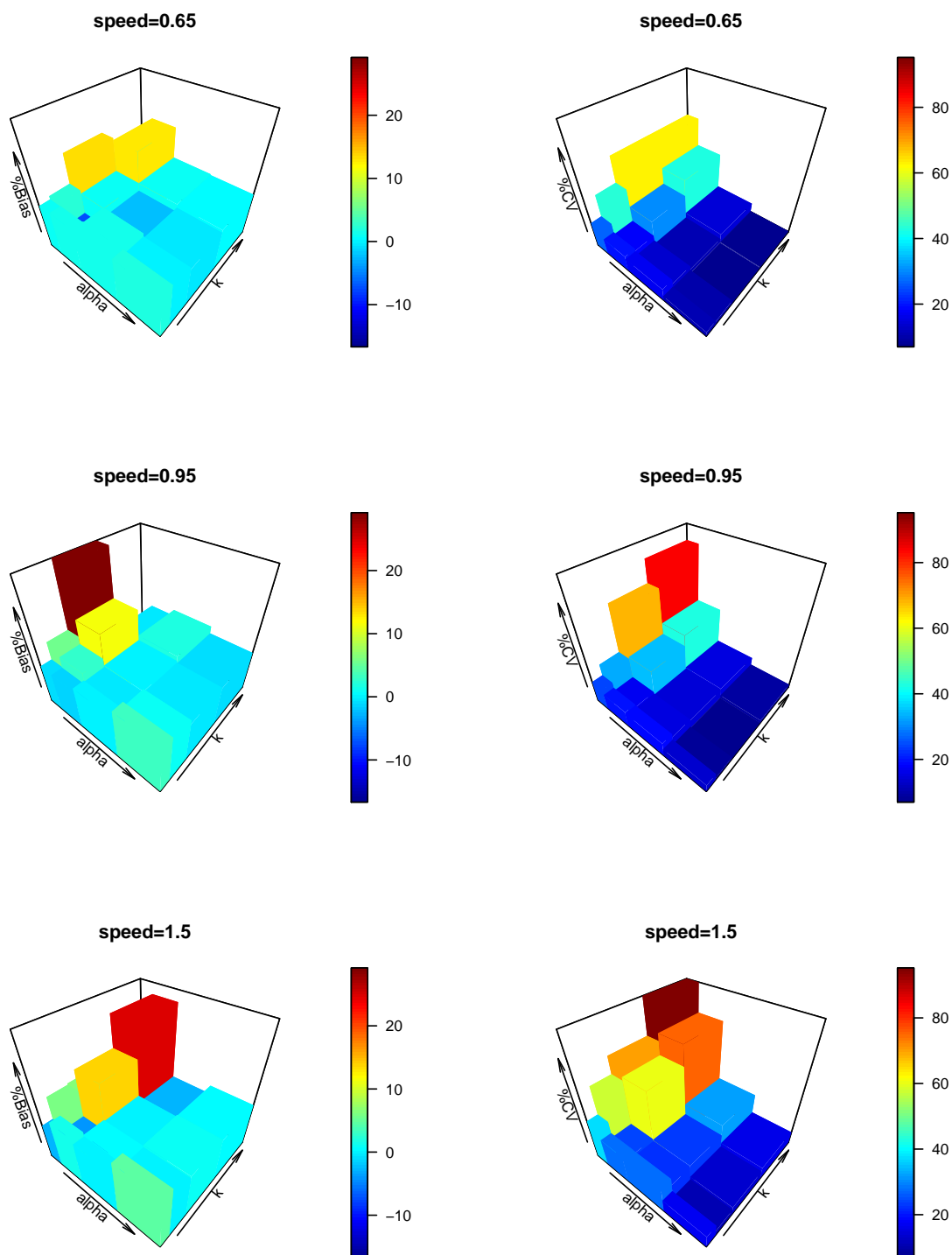
Figure 4. The percentage bias (first column) and percentage CV (second column) of $\hat{D}$ as a function of proportion of time available (α) and lag (l), for each of the three animal speeds considered.

Figure 5. The percentage bias (left) and percentage CV (right) of the MLE and Palm estimators of $\hat{D}$ over all simulation scenarios. Each point corresponds to one simulation scenario; scenaios with $k = 80$ are marked with diamonds, those with $\alpha = 0.1$ are marked with crosses. The bias and CV are the same for the two methods along the lines.

Table 1
Preliminay sims with Nsim=300.

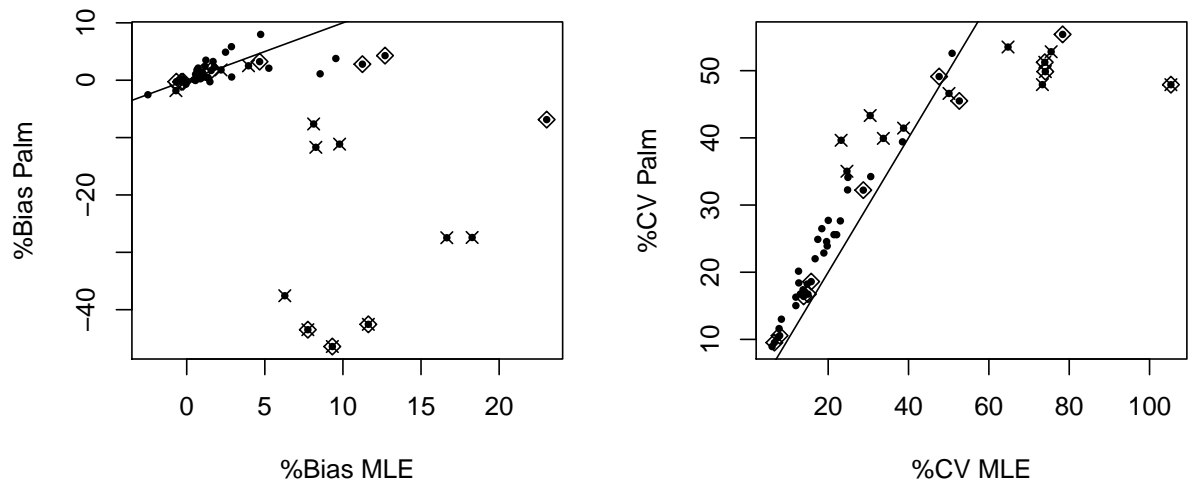|	alpha	l	speed	%Bias	%CV	%Cover
1	0.10	10.00	0.65	2.20	23.22	0.96
2	0.10	10.00	0.95	3.96	24.67	0.94
3	0.10	10.00	1.50	-0.69	30.45	0.95
4	0.10	20.00	0.65	8.13	38.77	0.95
5	0.10	20.00	0.95	8.27	33.73	0.97
6	0.10	20.00	1.50	9.79	50.08	0.96
7	0.10	50.00	0.65	16.65	75.51	0.83
8	0.10	50.00	0.95	18.27	64.76	0.83
9	0.10	50.00	1.50	6.28	73.32	0.76
10	0.10	80.00	0.65	7.76	74.05	0.77
11	0.10	80.00	0.95	11.63	73.92	0.75
12	0.10	80.00	1.50	9.32	105.32	0.66
13	0.20	10.00	0.65	0.73	18.42	0.97
14	0.20	10.00	0.95	1.48	20.02	0.96
15	0.20	10.00	1.50	-2.49	24.91	0.97
16	0.20	20.00	0.65	1.27	18.92	0.96
17	0.20	20.00	0.95	2.88	19.75	0.93
18	0.20	20.00	1.50	0.94	23.03	0.96
19	0.20	50.00	0.65	5.26	30.59	0.96
20	0.20	50.00	0.95	9.55	38.51	0.93
21	0.20	50.00	1.50	8.54	50.85	0.97
22	0.20	80.00	0.65	11.25	47.62	0.91
23	0.20	80.00	0.95	12.70	52.64	0.94
24	0.20	80.00	1.50	23.04	78.36	0.88
25	0.50	10.00	0.65	0.66	16.75	0.97
26	0.50	10.00	0.95	1.18	19.61	0.93
27	0.50	10.00	1.50	-0.27	24.83	0.92
28	0.50	20.00	0.65	1.38	13.75	0.96
29	0.50	20.00	0.95	1.16	14.80	0.94
30	0.50	20.00	1.50	1.23	21.44	0.91
31	0.50	50.00	0.65	0.56	11.96	0.96
32	0.50	50.00	0.95	-0.27	14.18	0.94
33	0.50	50.00	1.50	0.87	22.10	0.92
34	0.50	80.00	0.65	-0.66	13.88	0.95
35	0.50	80.00	0.95	1.57	15.77	0.95
36	0.50	80.00	1.50	4.67	28.74	0.94
37	0.80	10.00	0.65	1.69	11.95	0.92
38	0.80	10.00	0.95	2.86	12.62	0.94
39	0.80	10.00	1.50	4.73	17.42	0.88
40	0.80	20.00	0.65	0.57	7.78	0.95
41	0.80	20.00	0.95	0.85	8.35	0.96
42	0.80	20.00	1.50	2.49	12.68	0.92
43	0.80	50.00	0.65	0.47	6.04	0.98
44	0.80	50.00	0.95	-0.05	7.47	0.97
45	0.80	50.00	1.50	1.76	13.07	0.95
46	0.80	80.00	0.65	-0.35	6.61	0.99
47	0.80	80.00	0.95	-0.24	7.92	0.96
48	0.80	80.00	1.50	1.02	14.96	0.89