

A Markov modulated Bernoulli process model for digital aerial surveys of marine mammals without individual identification

D. L. Borchers^{1,*}, P. Nightingale¹, B.C. Stevenson², and R.M. Fewster²

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

St Andrews, Fife, UK

²Department 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, unknown recaptures, line transect, availability bias, movement model, Poisson process

1. Introduction

Aerial surveys of marine mammals are generally much cheaper than shipboard surveys. However, animals are within detectable range of an observer on an aircraft for much less time than they are to an observer on a ship, because aircraft move much faster than ships. More animals tend to be missed from an aircraft because more are underwater and unobservable while within detection range. This is known as the “ $g(0) < 1$ ” problem in the line transect literature **add citation**. The bias that results from ignoring this problem is often called “availability bias” **add citation**.

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”.

The utility of MRDS methods on aerial surveys is hampered by the fact that both observers 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 correlation in the resulting density estimate **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.

One of the ways of dealing with this problem is to model the latent variable - as is done with spatial capture-recapture methods **add citation**, where the locations of animals’ activity centres are the latent variables. Here the availability process is latent and we avoid bias by modelling it.

Another problem with MRDS methods on aerial surveys is that they require capture

histories, i.e. they require that every detected individual be recognised as having been detected by observer 1 only, observer 2 only, or both observers. 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. A central feature of our method that does not require capture histories.

Our method is developed in anticipation of digital detectors (high-definition video or stills cameras) being used increasingly in place human observers on aerial surveys. It has very models data requierments. It uses , and develop a method that is designed to estimate density from these data without the need to assign capture histories to individuals. Our method is similar to that of **CCR citation** in that it has the same data requirements and does not need any capture histories. It differs in that we do inference by maximum likelihood rather than using an approximation to the likelihood. Unlike the method of **CCR citation**, our method becomes computationally infeasible when there are very many plausible pairings of detections by different observers. We investigate the utility of our method and compare it to that of **CCR citation** when this is not the case. It is similar to that of Hiby and Lovell (1998), but is more general and extendable.

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 vk . Animals may move between the time that the first and second observers pass them.

We model the distance that an animal moves in the along-transect (or “forward”) direction ($B(t)$) and perpendicular direction ($X(t)$), in time t , as a bivariate normal distribution with mean $(0, 0)$ and variance $\Sigma(t) = \sigma^2 t \mathbf{I}$, where \mathbf{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 \mathbf{I} with a correlation parameter ρ , although we do not do that here.

2.1.1 Forward movement. In Appendix 7 we show that when the two observers are moving at a speed v separated by time k , the pdf of the time between the first and second observers encounter an animal (T), is

$$f_{T|k}(t|k) = \frac{vk \exp \left\{ -\frac{v^2(k-t)^2}{2\sigma^2 t} \right\}}{\sqrt{2\pi\sigma^2 t^3}}. \quad (1)$$

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

2.1.2 Movement in and out of the survey strip. Suppose both observers search a distance W either side of the transect line. Under the above movement model for $X(t)$, the probabilities that an animal moves from inside the strip to outside the strip in an interval t ($p_{IO}(t)$), and from outside the strip to inside the strip $p_{OI}(t)$, are

$$p_{IO}(t) = 2 \int_0^W \Phi(-(x - W); \sigma^2 t) + \Phi(-(x + W); \sigma^2 t) dx \quad (2)$$

$$p_{OI}(t) = 2 \int_W^\infty \Phi(W - x; \sigma^2 t) - \Phi(3W - x; \sigma^2 t) dx \quad (3)$$

This needs checking.

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

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 q_{11} and q_{22} such that the time in state 1 (the near-surface state) is an exponential random variable with expected value $\mu_1 = -\frac{1}{q_{11}}$, and the time in state 2 (the diving state) is an exponential random variable with expected value $\mu_2 = -\frac{1}{q_{22}}$. It is convenient to parameterise this Markov process in terms of the expected time in the near-surface state μ_1 and the expected dive-cycle length, $\tau = \mu_1 + \mu_2$.

The Markov transition rate matrix \mathbf{Q} is

$$\mathbf{Q} = \begin{pmatrix} q_{11} & -q_{11} \\ -q_{22} & q_{22} \end{pmatrix} \quad (4)$$

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

$$\pi = \left(\frac{q_{22}}{q_{22} + q_{11}}, \frac{q_{11}}{q_{11} + q_{22}} \right) = \left(\frac{\mu_1}{\tau}, \frac{\mu_2}{\tau} \right). \quad (5)$$

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

2.2.2 The in/out availability model. 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 (2) and (3):

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

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 $\mathbf{\Gamma}(t) = \mathbf{U}(t) \otimes \mathbf{M}(t)$.

To obtain the initial state distribution, we consider a “buffer” of width $b\sigma T$ either side of the searched strip (of width W). We choose b such that there is zero, or nearly to zero, probability of an animal beyond this buffer moving into the survey strip in a time T , and we choose T such that there is zero, or nearly zero, probability of the time between observers passing an animal being greater than 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 $b\sigma T$ of the line, is in the survey strip is $\phi = W/(W + b\sigma T)$. Combined with the stationary distribution π , this gives the following initial state distribution:

$$\boldsymbol{\delta} = \frac{1}{\mu} (\mu_1\phi, \mu_1(1 - \phi), \mu_2\phi, \mu_2(1 - \phi)). \quad (7)$$

3. A Markov modulated Bernoulli process model

We assume that the probability that an animal is in each state at the time the first aircraft 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 $\boldsymbol{\delta}\mathbf{\Gamma}(t)$.

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} is the time difference between the first and second observer passing (t in the development of the previous section).

We and define detection probability matrix for observer j to be

$$\mathbf{P}\{X_{ij}(t)\} = \begin{pmatrix} \text{up,in} & \text{up,out} & \text{down,in} & \text{down,out} \\ \text{Bern}(X_{ij}(t); p_j(1)) & 0 & 0 & 0 \\ 0 & 1 - X_{ij}(t) & 0 & 0 \\ 0 & 0 & \text{Bern}(X_{ij}(t); p_j(3)) & 0 \\ 0 & 0 & 0 & 1 - X_{ij}(t) \end{pmatrix} \quad (8)$$

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 $p_j(2) = 0$ and element (3, 3) is $1 - X_{ij}(t)$.

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

$$p\{X_{i1}(0), X_{i2}(t_{i2}) | t_{i2}\} = \boldsymbol{\delta} \mathbf{P}\{X_{i1}(0)\} \boldsymbol{\Gamma}(t_{i2}) \mathbf{P}\{X_{i2}(t_{i2})\} \mathbf{1} \quad (9)$$

where $\mathbf{1}$ 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 $p\{X_{i1}(0), X_{i2}(t_{i2}) | t_{i2}\} f_{T|k}(t_{i2} | k)$. For brevity, we write $p\{X_{i1}(0) = 0, X_{i2}(t_{i2}) = 1 | t_{i2}\} f_{T|k}(t_{i2} | k)$ as $p_{01}(t_{i2} | k)$, $p\{X_{i1}(0) = 1, X_{i2}(t_{i2}) =$

$0|t_{i2}\}f_{T|k}(t_{i2}|k)$ is written as $p_{10}(t_{i2}|k)$, and $p\{X_{i1}(0) = 1, X_{i2}(t_{i2}) = 1|t_{i2}\}f_{T|k}(t_{i2}|k)$ as $p_{11}(t_{i2}|k)$.

4. The survey probability 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 $\mathbf{s}_\omega = (\mathbf{s}_{\omega_1}, \mathbf{s}_{\omega_2}, \mathbf{s}_{\omega_3})$ be the observed locations along the transect of detections with capture history $\omega_1 = (01), \omega_2 = (10), \omega_3 = (11)\}$. These arise from a thinned NHPP with a thinning probability (for each ω) of

$$\tilde{p}_\omega(k) = E_t[p_\omega(t|k)] = \int p_\omega(t|k)f_{T|k}(t|k)dt, \quad (10)$$

so that for a survey along transects of total length L , the pdf of locations \mathbf{s}_ω is

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

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

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

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

Given the capture histories $\boldsymbol{\omega}$, the likelihood is therefore as follows (with model parameters

suppressed for simplicity of presentation):

$$\begin{aligned}\mathcal{L}(\mathbf{s}|\boldsymbol{\omega}) &= \left\{ \prod_{\omega} \left[\prod_{i=1}^{n_{\omega}} \tilde{p}_{\omega}(k) D(s_{\omega i}) \right] e^{-\int_0^L \tilde{p}_{\omega}(k) D(s) ds} \right\} \prod_{i=1}^{n_{(11)}} f_{T|k,11}(t_{(11)i}|k, \omega_i = 11) \\ &= e^{-\int_0^L \tilde{p} \cdot (k) D(s) ds} \left\{ \prod_{\omega} \tilde{p}_{\omega}(k)^{n_{\omega}} \prod_{i=1}^{n_{\omega}} D(s_{\omega i}) \right\} \prod_{i=1}^{n_{(11)}} f_{T|k,11}(t_{(11)i}|k, \omega_i = 11)\end{aligned}\quad (13)$$

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

4.1 Homogeneous Poisson case

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

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

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 $\theta^* = (\theta, \gamma_1, \gamma_\tau, \phi, \beta)$. **CCR citation** 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 camera 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.

CCR citation show that even in this case, only two of $(\theta, \gamma_1, \gamma_\tau)$ are identifiable. Like **cite CCR** 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 $\theta = (\theta, \gamma_1, \phi)$. 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 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 ω), we could estimate the model parameters by maximising the likelihood Equation Equation (14) with respect to θ .

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

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

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 $\omega^{(m)}$ s, and for any but very small sample sizes, the number M of possible pairings is very large ($M \gg 10^9$ for even moderate sample sizes).

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

5.1 Subdivision of \mathbf{s}

We subdivide \mathbf{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 pairings $\omega^{(m_c)}$ for segments $c = 1, \dots, C$ (with M_c possible pairings in segment c) and calculate the likelihood as

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

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 $\omega^{(m)}_s$

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 $\text{scope}(c) \subseteq \mathcal{V}$. 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 $b_{i,j} = 0 \vee b_{k,l} = 0$ 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 $\sum_{m_c=1}^{M_c} \mathcal{L}(\theta; \mathbf{s} | \omega^{(m_c)})$.

6. Applications

6.1 Lag and correlation

These are just preliminary thoughts, written here so I don't forget them.

The optimal lag depends on the correlation between detections by the two observers (*in some way I have not yet figured out*): With lag $k = 0$, there is 100% correlation there is no information in the data on the availability process. If the lag is such that there is no correlation between the availability of an animal to each observer, then the availability process is irrelevant since the probability of a recapture is a combination of the probability of being available and the probability of detection, given availability (see Hayes and Buckland (1983)).

Figure 1 shows the correlation as a function of k when $\mu_c = 138$ and $\delta_1 = 28/138$ (our test scenario). It is calculated using Equation (2.10) for covariance of a hidden Markov model, obtained from Zucchini and MacDonald (2009):

$$\text{cov}(X_t X_{t+k}) = \sum_{i,j=1}^2 E(X_i | S = i) E(X_j | S = i) \delta_i \gamma_{ij}(k). \quad (17)$$

For $u = j, k$, X_u is 1 if observer u detects an animal, and is zero otherwise, and $X_u | S$

is a Bernoulli random variable with parameter equal to the probability of observer u detecting an animal in state S . The term $\gamma_{ij}(k)$ is the top left element of $\mathbf{\Gamma}^k$. In our case Equation (17) reduces to $\text{cov}(X_t X_{t+k}) = p^2 \delta_1 \gamma_{11}(k)$ and the correlation is $\text{cov}(X_t X_{t+k}) = p^2 \delta_1 \gamma_{11}(k) / p^2 \delta_1 \gamma_{11}(0) = \gamma_{11}(k)$.

[Figure 1 about here.]

7. Discussion

- With known capture histories this is a kind of SCR transect search with a continuous time movement model and an availability model. Would improve it if extend theory to include 2D observed locations.
- If know capture histories, can estimate movement and dive cycle parameters.
- Add covariates to any of the models
- Slower moving observers: need to take account of first passage time (which ccr method does not).
- Add more dive states, e.g. split “near-surface” state into “breaking surface” and “near-surface”. Probably infeasible as this adds another 4 parameters. Probably only feasible with known capture histories.

Received ?? 2018. Revised 20??

Appendix Derivation of $f_{T|k}(t|k)$

One can view the problem 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 vk 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 .

Currently this is based on these two pages:

[https://math.stackexchange.com/questions/1053294/density-](https://math.stackexchange.com/questions/1053294/density-of-first-hitting-time-of-brownian-motion-with-drift)

[of-first-hitting-time-of-brownian-motion-with-drift](https://math.stackexchange.com/questions/1053294/density-of-first-hitting-time-of-brownian-motion-with-drift) and

[https://math.stackexchange.com/questions/179210/derivation-of-wiener-process-](https://math.stackexchange.com/questions/179210/derivation-of-wiener-process-first-passage-times-using-probability-generating-fu)
[first-passage-times-using-probability-generating-fu](https://math.stackexchange.com/questions/179210/derivation-of-wiener-process-first-passage-times-using-probability-generating-fu)

From the first of the above sites:

Suppose that $X_t = B_t + ct$, where B is a Brownian motion, c is a constant. Set $H_a = \inf\{t : X_t = a\}$ for $a > 0$. It follows that for $c \in \mathbb{R}$, the density of H_a is

$$f_{H_a}(t) = \frac{a \exp\left\{\frac{-(a-ct)^2}{2t}\right\}}{\sqrt{2\pi t^3}}. \quad (\text{A.1})$$

In our case, we have an animal approaching the “boundary” (second observer) that is a distance vk from the starting position of the animal, 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 $a = vk/\sigma$. Hence the pdf of time to the second observer passing, given that the second observer passes a fixed point at a time k later than the first, is

$$f_{T|k}(t|k) = \frac{\frac{vk}{\sigma} \exp\left\{\frac{-(vk/\sigma - vt/\sigma)^2}{2t}\right\}}{\sqrt{2\pi t^3}} = \frac{vk \exp\left\{\frac{-v^2(k-t)^2}{2\sigma^2 t}\right\}}{\sqrt{2\pi \sigma^2 t^3}}. \quad (\text{A.2})$$

Appendix Summary of inputs and outputs

Evaluation of the above likelihood requires the following inputs:

θ_1 : the density parameter,

θ_2 : a parameter governing transition from state 1 (surface) to state 2 (underwater),

μ_c : the expected dive cycle duration,

$p(1)$: the probability that observer 1 detects an animal that is on the surface,

$p(2)$: the probability that observer 2 detects an animal that is on the surface,

θ_3 : the log of animal movement speed,

k : the time separation between observers,

L : the total distance flown by the observers,

s_{11}, \dots, s_{1n_1} : the distances along the flight path at which the n_1 animals detected by observer 1 were detected,

s_{21}, \dots, s_{2n_2} : the distances along the flight path at which the n_2 animals detected by observer 2 were detected,

d_{max} : the distance beyond which detections by different observers are certainly detections of different animals,

δ_{si} : the distance (in observer-seconds) that the i th animal moves between the first and second observer passing it,

$t_i = k + \delta_{si}$: the time between first and second observer passing animal i .

from which it should return the value of the likelihood function $\mathcal{L}(k)$.

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

Online supplementary material

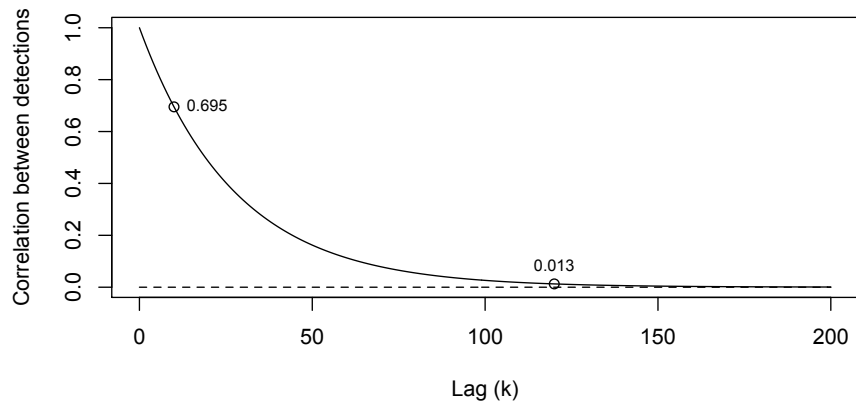


Figure 1. Correlation of detections at various lags (k) for an availability process with mean availability cycle length of 138 and mean proportion of time available of $28/138$. Correlations at $k = 10$ and $k = 120$ are shown explicitly.