BIOMETRICS 00, 1–??
?? 2018
DOI: 10.1111/j.1541-0420.2005.00454.x

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.

Bit of ambiguity here b/c ~~MRDS~~ $g(0)<1$ embraces both availability & perception bias, and MRDS is usually used to address just perception bias.

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?)). McKenzie & Clements.

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

(*) Not just along-transect, as the model given is 2-dim'l?

We model the distance that an animal moves in the along-transect (or "forward") direction and perpendicular direction, 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 parameer ρ, although we do not do that here.

Ben uses log l for what's called k here.

2.1.1 *Forward movement.* In Appendix 8 A we show that when ~~the~~ 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 of an animal (T), is

Could say T = time at which the forward coordinate of ~~plane~~ observer 2 equals ~~matches~~ that of whale: there might not be an "encounter".

Remove "|k" in both places

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

The animal will have moved a forward distance $v(t-k)$ 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.

Need to establish bufferzone before this step

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

I get $p_{IO}(t) = \frac{1}{2W}\int_{-W}^{W} \mathbb{P}(\text{out}\,|\,\text{start @ }x)\,dx = \frac{1}{W}\int_0^W \mathbb{P}(\text{out}\,|\,x)\,dx = \frac{1}{W}\int_0^W \{\Phi(x-W)+\Phi(-(x+W))\}\,dx$

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

$-W-x$?

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

Hmm... Strictly speaking $p_{OI}(t)=0$ unless we restrict the x's from which the whale is allowed to start.

Lifted from old notes without checking – needs checking.

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

Might need to use a bufferzone argument like Ben did.

2.2 *Availability models*

a more complicated expression of the sample space... suspect Ben's bufferzone is easier.

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 constuct 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, $\tau = \mu_1 + \mu_2$.

Ben uses K for μ_1 and $\tau - K$ for μ_2.

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

$$\boldsymbol{Q} = \begin{pmatrix} q_{11} & -q_{11} \\ -q_{22} & q_{22} \end{pmatrix} \tag{4}$$

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

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

The transition probability matrix for transitions between states at time separation t is $\boldsymbol{U}(t) = \exp(\boldsymbol{Q}t)$. → does this equate to $\mathbb{P}(U_2|U_1) = \frac{K}{\tau} + \frac{\tau-K}{\tau} \exp\{-(\frac{1}{K} + \frac{1}{\tau-K})t\}$ as in Ben's paper?

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):

Could save space by moving the 2.1.2 material so all the in-out treatment is together here.

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

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

To obtain the initial state distribution, we consider a "buffer" of width $b\sigma T$ either side

Ben's b includes W : this b is added on separately.

of the searched strip (of width W). We choose b such that there is zero, or close 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.

Using buffer bσT means it depends on the parameter σ. Could cause problems b/c large σ will suddenly make buffer larger?

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:

spaces

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

τ not μ ← (*)

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~~ observer/camera passes over it is given by the stationary distribtion $\boldsymbol{\delta}$, 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} (t in the development of the previous section).

→ Indicate that this is a convenience for the HMM formulation (the matrix form)

We and define detection probability matrix for observer j to be

ie prob = 0 if $X_{ij} = 1$, prob = 1 if $X_{ij} = 0$

Use lower-case x instead of $X_{ij}(t)$

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

Why is this a matrix?

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 $p_j(2) = 0$ (P_j(3) = 0 ?) and element $(3,3)$? is $1 - X_{ij}(t)$. Why aren't all elts 0 except P_{11} ?

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

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

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 $p_{11}(t_{i2}|k)$.

Remove "$|k$" from everything.

shift the brief notation earlier.

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 (for the three ωs) of

shift this to end of previous section.

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

so that for a survey along transects of total length L, the pdf of locations $\boldsymbol{s}_\omega$ is

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

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

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

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):

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

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

4.1 *Homogeneous Poisson case*

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

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

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 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 $f_{T|k}(t|k)$

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 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.

We use the result **(citation)** that for $X_t = B_t + ct$, where B is a Brownian motion, a
location (boundary might confuse b/c it isn't really a boundary?)
"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 vk from the starting position of the animal, approaching it at speed v, and the animal moving according
already said, so unclear if this is derivation.
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

Nice, but took me a bit of working out – not very obvious. Could say:
(Plane 2's position @ time t) $= -vk + vt$; (Whale's posn @ time t) $= \sigma W_t$ Brownian motion
For sighting, require plane 2's posn = whale's posn
$\Rightarrow -vk + vt = \sigma W_t$
$\Rightarrow \frac{vk}{\sigma} = \frac{vt}{\sigma} - W_t = \frac{vt}{\sigma} + B_t$ where $B_t = -W_t$ also B.M.
then call on A.1, no need for rescaling etc.

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 $f_{T|k}(t|k)$

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 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.

We use the result **(citation)** that for $X_t = B_t + ct$, where B is a Brownian motion, a
location (boundary might confuse b/c it isn't really a boundary?)
"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 vk from the starting position of the animal, approaching it at speed v, and the animal moving according
already said, so unclear if this is derivation.
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

Nice, but took me a bit of working out — not very obvious. Could say:
(Plane 2's position @ time t) $= -vk + vt$; (Whale's posn @ time t) $= \sigma W_t$ → Brownian motion
For sighting, require plane 2's posn = whale's posn
$\Rightarrow -vk + vt = \sigma W_t$
$\Rightarrow \frac{vk}{\sigma} = \frac{vt}{\sigma} - W_t = \frac{vt}{\sigma} + B_t$ where $B_t = -W_t$ also B.M.
then call on A.1, no need for rescaling etc.