Estimating occupancy and abundance using aerial images with imperfect detection

Perry J. Williams[1,2], Mevin B. Hooten[3,2], Jamie N. Womble[4,5] and Michael R. Bower[4]

[1]Colorado Cooperative Fish and Wildlife Research Unit, Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO

[2]Department of Statistics, Colorado State University, Fort Collins, CO

[3]U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Research Unit, Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO

[4]National Park Service, Southeast Alaska Inventory and Monitoring Network, Juneau, AK

[5]National Park Service, Glacier Bay Field Station, Juneau, AK

May 5, 2017

10.1111/2041-210X.12815

Summary

1. Species distribution and abundance are critical population characteristics for efficient management, conservation, and ecological insight. Point process models are a powerful tool for modeling distribution and abundance, and can incorporate many data types, including count data, presence-absence data, and presence-only data. Aerial photographic images are a natural tool for collecting data to fit point process models, but aerial images do not always capture all animals that are present at a site. Methods for estimating detection probability for aerial surveys usually include collecting auxiliary data to estimate the proportion of time animals are available to be detected.

2. We developed an approach for fitting point process models using an N-mixture model framework to estimate detection probability for aerial occupancy and abundance surveys. Our method uses multiple aerial images taken of animals at the same spatial location to provide temporal replication of sample sites. The intersection of the images provide multiple counts of individuals at different times. We examined this approach using both simulated and real data of sea otters (*Enhydra lutris kenyoni*) in Glacier Bay National Park, southeastern Alaska.

3. Using our proposed methods, we estimated detection probability of sea otters to be 0.76, the same as visual aerial surveys that have been used in the past. Further, simulations demonstrated that our approach is a promising tool for estimating occupancy, abundance, and detection probability from aerial photographic surveys.

4. Our methods can be readily extended to data collected using unmanned aerial vehicles, as technology and regulations permit. The generality of our methods for other aerial surveys depends on how well surveys can be designed to meet the assumptions of N-mixture

models.

Key words abundance estimation, aerial photographic surveys, availability bias, detection bias, detection probability, occupancy estimation, perception bias, point process models, sea otters

Introduction

Aerial surveys are an important tool for estimating abundance and distribution of vertebrate populations. Methods for design and data analysis of aerial surveys have been developed to accommodate visual observations where observers count animals from aircraft (Siniff & Skoog 1964, Goddard 1967; 1969, Jolly 1969a;b, Watson *et al.* 1969, Caughley & Goddard 1972, Pennycuick & Western 1972, Caughley 1974, Certain & Bretagnolle 2008), and for photographic survey methods (Leedy 1948, Leonard & Fish 1974, Boyd 2000, Bechet *et al.* 2004, Buckland *et al.* 2012, Conn *et al.* 2015, Ver Hoef & Jansen 2014). Undercounting animals from aircraft presents a major estimation problem with both visual and photographic aerial surveys (Graham & Bell 1969, Caughley 1974). Animals are undercounted because they are not available to be counted (e.g., underwater as in Lukacs *et al.* 2010, termed *availability bias*), or observers miss animals that are available to be counted (termed *perception bias*; Marsh & Sinclair 1989). Aerial images improve perception bias, but not necessarily availability bias (Leonard & Fish 1974, Gibbs *et al.* 1988, Bayliss & Yeomans 1990, Frederick *et al.* 2003). For example, many seabirds and marine mammals are virtually certain to be detected in images if they are at the surface of the water, but animals may be diving beneath the surface of the water and unavailable to be photographed (Buckland *et al.* 2012, Conn *et al.* 2014). Aerial images alone typically do not provide sufficient information for estimating availability, and auxiliary information is usually required to estimate absolute abundance. For example, activity budgets or time spent diving underwater can be estimated from telemetry devices including VHF transmitters, satellite-linked transmitters, or time-depth recorders (Bechet *et al.* 2004, Heide-Jørgensen *et al.* 2007, Conn *et al.*

2014). Often, aerial image data are easily obtainable, but auxiliary data may be more challenging to acquire due to financial, logistical, or regulatory constraints, precluding estimation of availability. Even when auxiliary information can be collected, there is often a disparate scale of inference between auxiliary data and aerial image data, potentially introducing a variety of statistical challenges for modeling and inference (Gotway & Young 2002). Further, it is difficult to determine whether correction factors based on behavioral data (e.g., dive times) are appropriate because corrections may not apply to animals engaged in different activities such as feeding or resting; activities that might be difficult to characterize from an aircraft (Hiby & Lovell 1998).

Another method for estimating availability is to use multiple aircraft in tandem, where observers in each aircraft count animals independently (Hiby & Lovell 1998). Using aircraft in tandem is twice as expensive as using one aircraft, and transects of each aircraft might not overlap due to error in GPS locations and misaligned flight paths. Further, methods for using aircraft in tandem have relied on identifying individual animals (i.e., duplicates seen by each aircraft) which is often problematic (Hiby & Lovell 1998).

In light of these constraints, we describe a point process model that leverages an N-mixture framework for simultaneously estimating detection probability, occupancy, and abundance from aerial images (Royle 2004). The N-mixture model fits naturally within the framework of a point process model. The N-mixture model is advantageous in that it does not rely on data auxiliary to aerial images. Data required to fit N-mixture models can be collected from a single aircraft, and individual animals do not need to be uniquely identified. In addition, aerial photographic images provide a permanent record that is available for independent verification, may be used for automated detection, and allows for quantification of habitat covariates (Martin *et al.* 2012, McNabb *et al.* 2016). Photographic sampling methods can also be extended to unmanned aerial vehicles, which are relatively new low-cost platforms that can be used to quantify wildlife and their habitats (Hodgson *et al.* 2013, Sweeney *et al.* 2015).

An ecological application motivating the methods we present involves the use of aerial photographic survey methods for estimating the distribution and abundance of sea otters (*Enhydra lutris kenyoni*) in Glacier Bay National Park (GBNP), southeastern Alaska. Sea otters were recently identified as a vital sign for long-term monitoring in GBNP because of their role as a keystone species and their influence in structuring nearshore marine communities (Estes & Palmisano 1974). Data on sea otter abundance in GBNP were formerly collected using design-based, visual aerial surveys, where observers counted sea otters along randomly selected transects (Bodkin & Udevitz 1999, Williams *et al.* 2017). Detection probability for the design-based survey was estimated by conducting intensive searches at 469 randomly-selected locations from the design-based survey that contained sea otters. At these random sites, observers first conducted the design-based survey along the transect and counted individuals along the strip. The plane then deviated off the transect to conduct intensive searches which entailed circling a group of otters five times at a given speed and altitude (Bodkin & Udevitz 1999). The number of circles was based on the estimated dive duration or aerobic dive limit of sea otters. The additional survey effort allowed observers to obtain more precise counts of sea otters, including individuals that were underwater and not available for detection during the original design-based phase. The National Park Service is developing an aerial survey method that extends previous survey methods with the specific objectives of 1) improving safety and reducing risk associated with aerial surveys, 2) decreasing cost and optimizing efficiency, 3) increasing the number of pilots capable of conducting the surveys, 4) creating a permanent record that can be independently verified, 5) reducing observer bias, 6) quantifying associated habitat covariates from imagery, 7) developing a platform and survey design that is capable of being extended to unmanned aerial vehicles, and 8) improving precision of occupancy and abundance estimates. The use of aerial images improves objectives 1–7, relative to the original surveys. However, objective 8 relies on accurate and precise estimation of detection probability. Thus, we developed field-based methods and an associated statistical framework for simultaneously estimating occupancy, abundance, and detection probability of animals using only aerial images, where a subset of images overlap in space. We demonstrate our approach by first applying our framework to simulated data of sea otters in GBNP. We conducted a pilot study in which we assessed our ability to collect temporally replicated images of spatially referenced sites containing sea otters. Finally, we estimated abundance and detection probability of our sampled sites.

Methods

Individuals in a population exist as points in space and time and therefore can be modeled as a spatio-temporal point process (Fig. 1; Hefley & Hooten 2016). A point process is a stochastic process that governs the location of a set of points $\{\mathbf{s}_i\}$ in some set $D \subset \mathbb{R}^d$ (Moller & Waagepetersen 2003, Diggle 2013, Cressie & Wikle 2011). We consider the two-dimensional space $D \subset \mathbb{R}^2$ that describes a study area of interest during time interval $[0, T]$. We let $\mathbf{s} = (\text{latitude, longitude})'$ represent any 2×1 vector of coordinates in D, and the set of coordinates $\{\mathbf{s}_i \ \forall i\}$ represents the locations of animals in D. A point process $N(\cdot)$ is characterized by counting the number of points (e.g., animals) that belong to various measurable subsets $A \subset D \subset \mathbb{R}^2$ (e.g., A represents the subset of area in D captured by aerial images). Let $N(A, t)$ represent a count of the true number of animals in A during time t. A fundamental quantity of interest of a point process is the expected abundance in A. The expected abundance can be calculated in terms of a locally integrable intensity function $\lambda(\mathbf{s}, t)$. The intensity function describes the expected abundance of an infinitely small area $d\mathbf{s}$ centered at point $\mathbf{s}$. The expected abundance of the bounded subset A from time t to $t + \Delta t$ can be obtained by integrating the intensity function over A from time t to $t + \Delta t$,

$$\lambda(A, t) = E(N(A, t)) = \int_A \int_t^{t+\Delta t} \lambda(\mathbf{s}, t) dt d\mathbf{s} < \infty.$$

If $\lambda(\mathbf{s}, t)$ varies in space and/or time, and is deterministic, and we assume $N(A_1, t)$ and $N(A_2, t)$ are independent whenever $A_1 \cap A_2 = \emptyset$, where $\cap$ represents the intersection, then $N(A, t)$ is an inhomogeneous Poisson point process and

$$N(A, t) \sim \text{Poisson}(\lambda(A, t)) \tag{1}$$

(Moller & Waagepetersen 2003, Cressie & Wikle 2011). Similarly, if $\lambda(\mathbf{s}, t)$ is a stochastic process, then $N(A, t)$ is known as a Cox process (Cox 1955). We consider inhomogeneous Poisson processes for the remainder of the paper; a number of authors provide details on more general point processes that could be used in our framework (e.g., Moller & Waagepetersen 2003, Illian *et al.* 2008, Cressie & Wikle 2011, Banerjee *et al.* 2014, Baddeley *et al.* 2015). An important derived quantity of Eq. 1 is the probability that $N(A, t) > 0$ (i.e., the occupancy probability, $\phi(A, t)$). Useful distribution models predict both occupancy and abundance (Oppel *et al.* 2012). The spatio-temporal occupancy probability is

$$\phi(A, t) = P(N(A, t) > 0) = 1 - e^{-\lambda(A,t)} \tag{2}$$

(see Hefley & Hooten 2016, Williams *et al.* 2017).

A set of n aerial images taken at locations $\{\mathbf{c}_i\}_{i=1}^{n}$ capture information on bounded subregions A_i, and therefore can be used to characterize the point process (Cressie & Wikle 2011). We denote counts of animals on an image of site A_i taken during time t as $y(A_i, t)$, where the area captured by the aerial image is denoted as $|A_i|$, and in practice $\sum_{i=1}^{n} |A_i| << |D|$. Counts of individuals are usually obtained by examining images post-flight, and summing the number of individuals within the image (see *Field and lab methods*, below). Counts often contain false negatives; some proportion of animals within A_i will not appear on images or are missed when investigators count the animals. Therefore, we adopt the most commonly used model for false negatives,

$$y(A_i,t) \sim \text{Binomial}(N(A_i,t), p(A_i,t)), \tag{3}$$

where $p(A_i,t)$ is the detection probability, potentially varying in space and time. If occupancy is a state variable of interest, Eq. 3 reduces to $I_{\{y(A_i,t)>0\}} \sim \text{Bernoulli}(\phi(A_i,t))$, where $I_{\{y(A_i,t)>0\}}$ is an indicator function that equals one when $y(A_i,t) > 0$, and zero otherwise (Hefley & Hooten 2016). Note that $p(A_i,t)$ is a composite parameter of both the probability of an observer counting an individual on the image, conditional on it being available to be counted ($p_1(A_i,t)$), and the probability an individual is available to be counted ($p_2(A_i,t)$). That is, $p(A_i,t) = p_1(A_i,t)p_2(A_i,t)$. If aerial images have sufficient resolution such that the observer detection probability $p_1(A_i,t) = 1$, then $p(A_i,t) = p_2(A_i,t)$. For our application, we assume $p(A_i,t) = p_2(A_i,t)$. When this assumption is not valid, other techniques could be used to estimate $p_1(A_i,t)$ when individuals are counted on images (e.g., double observer methods as used in Buckland *et al.* 2012). Following the terminology of Berliner (1996), Eq. 3 is a data model and Eq. 1 is a process model (e.g., an inhomogeneous Poisson process model), and the hierarchical formulation of the model is

$$\begin{aligned} y(A_i,t) &\sim \text{Binomial}(N(A_i,t), p(A_i,t)), \\ N(A_i,t) &\sim \text{Poisson}(\lambda(A_i,t)), \\ \lambda(A_i,t) &= \int_{A_i}\int_t^{t+\Delta t} \lambda(\mathbf{c},t)dtd\mathbf{c} \\ \log(\lambda(\mathbf{c},t)) &= \mathbf{x}(\mathbf{c},t)'\boldsymbol{\beta}, \end{aligned} \tag{4}$$

where $\mathbf{x}(\mathbf{c},t)$ is a vector of covariates for locations $\mathbf{c}$, time t, and $\boldsymbol{\beta}$ is a vector of parameters to be estimated.

To estimate detection probability, p, we assume a subset of A_i and $A_k (k \in \mathscr{I}; i \neq k)$

intersect (Fig. 1). That is, an aerial image of subregion A_i taken at time $j = j_1$ overlaps an image of subregion A_k taken at time $j = j_1 + \Delta j$, where Δj is sufficiently small so it can be assumed that the point pattern realization is static, and in practice, $\Delta j << \Delta t$ (c.f. *primary* and *secondary* sampling periods *sensu*, Kendall & Nichols 1995, for t and j, respectively). We can view the union of A_i and A_k as 3 distinct bounded sites; $A_i \cap A_k^c$, $A_i^c \cap A_k$, and $A_i \cap A_k$ (Fig. 1), where "c" represents the complement. Counts of individuals (or alternatively, occupancy status) in the subregion $A_i \cap A_k$ can be obtained from each image (Fig. 1). Counting individuals from each image in the intersection $A_i \cap A_k$ provides temporal replication, $y(A_i \cap A_k, j, t)$, $j = 1, ..., J$, and can be used to estimate p in an N-mixture model framework (Royle 2004). That is, assuming Δj is sufficiently small to ensure the population being sampled is closed with respect to movement, mortality, and recruitment, and conditional on $N(A_i \cap A_k, t)$, $y(A_i \cap A_k, j, t)$ may be viewed as independent and identically distributed binomial random variables

$$y(A_i \cap A_k, j, t) \sim \text{Binomial}(N(A_i \cap A_k, t), p(A_i \cap A_k, t))$$

(Royle 2004). Thus, the Poisson process assumes that when $A_1 \cap A_2 = \emptyset$, $N(A_1)$, and $N(A_2)$ are independent and Poisson, conditional on $\lambda(A_1)$ and $\lambda(A_2)$, respectively. The N-mixture model assumes that when $A_1 \cap A_2 \neq \emptyset$, counts of individuals in the intersection are independent and binomial, conditional on $N(A_1 \cap A_2)$ and $p(A_1 \cap A_2)$.

It is not necessary to collect intersecting images at all spatial locations, but collecting intersecting images in a variety of environmental conditions provides sufficient information to estimate how detection probability may vary in response to spatial or temporal covariates. For example, heterogeneity in detection probability could be modeled as

$$\text{logit}(p(A_i \cap A_k, t)) = \mathbf{w}(A_i \cap A_k, t)'\boldsymbol{\alpha}, \tag{5}$$

where $\mathbf{w}(A_i \cap A_k, t)$ are covariates associated with detection probability collected at $A_i \cap A_k$ at time t, and $\boldsymbol{\alpha}$ are parameters to be estimated. To simplify notation in what follows, we assume that all sites (i.e., A_i, A_k, $A_i \cap A_k$, $A_i^c \cap A_k$, $A_i \cap A_k^c$) are represented with i.

The parameters of our model (p, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$) can be estimated using either Bayesian methods (Royle & Dorazio 2008), or maximum likelihood methods (Royle 2004). Assuming a Bayesian hierarchical specification of the model, the full Bayesian posterior distribution of Eq. 4 and 5 is

$$[\mathbf{N}, \boldsymbol{\alpha}, \boldsymbol{\beta} | \mathbf{Y}] \propto \prod_{i=1}^{n} \prod_{t=1}^{T} \left\{ \prod_{j=1}^{J_i} \left\{ [y(A_i, j, t) | N(A_i, t), \boldsymbol{\alpha}] \right\} [N(A_i, t) | \boldsymbol{\beta}] \right\} [\boldsymbol{\alpha}, \boldsymbol{\beta}], \tag{6}$$

where we use the square-bracket notation $[a|b]$ to represent the probability density or mass function of variable a given variable b (Gelfand & Smith 1990).

Application: sea otters in Glacier Bay National Park

We conducted a simulation study to evaluate our model (Appendix S1). After our simulation study, we developed and implemented field methods to assess the logistics of collecting the necessary data required to fit our model (Appendix S2). Although Eqs. 4 and 6 are described in sufficient generality to incorporate relevant spatio-temporal processes for time $t = 1, \ldots, T$, to simplify demonstration, our example assumes $t = 1$.

Simulated data

We simulated sea otter abundance data across GBNP (Appendix S1) using the model described in Eq. (6), and included an intercept and four covariates to associate abundance to local conditions in GBNP. The covariates were ocean depth, distance to shore, slope of the ocean floor, and shoreline complexity. We based the relationship (positive or negative) between abundance intensity and parameter values on Williams *et al.* (2017). We also allowed detection probability to vary in space. We simulated values of $w(\mathbf{s}_i)$ using a Bernoulli distribution with success probability equal to 0.5. Specifically,

$$\log(\lambda(\mathbf{s}_i)) = -2.4 + 0.75\text{depth}(\mathbf{s}_i) - 0.75\text{distance}(\mathbf{s}_i) + 0.4\text{slope}(\mathbf{s}_i) - 0.4\text{complexity}(\mathbf{s}_i),$$

$$\text{logit}(p(\mathbf{s}_i)) = 1 + 0.5w(\mathbf{s}_i).$$

We simulated 8,895 images from 50 transects placed randomly across GBNP, and selected 100 random locations containing sea otters where one additional replicate image was taken (Appendix S1). We then fit a Bayesian hierarchical N-mixture model to the simulated data. We assumed vague prior distributions for all parameters. After fitting the model, we compared the estimated posterior distributions to the parameter values that were used to simulate the data. We also plotted the true expected abundance ($\lambda(\mathbf{s})$) and occupancy ($\phi(\mathbf{s})$) and the estimated expected abundance and occupancy (Fig. 3). All posterior distributions had good coverage of true parameter values (Fig. 2), and the estimated expected abundance and occupancy probability represented the truth well (Fig. 3).

Field and lab methods

We developed a pilot study to assess the ability of obtaining intersecting aerial images of groups of sea otters in GBNP. We obtained aerial digital imagery to estimate abundance and detection of sea otters. Three separate aerial photographic surveys were conducted during July and September 2016 in Glacier Bay from a de Havilland Canada DHC-2 Beaver single-engine high-winged aircraft (Ward Air Inc., Juneau, Alaska). The aircraft was flown at approx. 213–250 m at 157–166 km/hr. Overlapping digital photographic images of sea otter groups were taken directly under the plane using a vertically-aimed digital camera (Nikon D810, 36.3 megapixel) with an 85 mm focal length lens (Zeiss F/1.4 ZF.2). The camera was attached to a tripod head and mounted to a plywood platform that was secured in the belly porthole of the aircraft. The camera captured an image every second, using a digital timer (Nikon MC36) that was attached to the camera and operated by the primary observer. A second observer monitored movement of individual sea otters to determine if any sea otters were dispersing or moving in or out of the photograph footprint. After one transect was complete (i.e., one survey occasion), the pilot attempted to fly the same transect (using a combination of a GPS and visual cues) to obtain replicate images. At one site, otters dispersed in varying directions after the initial photograph, and therefore, we removed it from analysis because it would potentially violate the closure assumption. The speed–altitude combination did not appear to have an observable influence on sea otter behavior.

An onboard global positioning system (GPS; Garmin 76 CSX), with an external antenna, was used to record the track line and position of the plane (latitude, longitude, altitude) at 1-sec intervals. Each digital image (7,360×4,912 pixel JPG) covered approximately 90 m $\times$ 60 m at the surface of the water with a 1.64 cm pixel resolution. The rate of image capture and dimensions of the photographed area provided a linear mosaic of overlapping images with approximately 30 m of overlap between adjacent images.

The latitude, longitude, and altitude from the track line were downloaded and written to the EXIF headers of each digital image to permanently embed the location data into each image using RoboGeo v6.3. All images were reviewed using ACDSee Pro 9 (ACD Systems International, Incorporated). From each survey occasion at each site, one image was selected based on its clarity and location of sea otters with respect to the boundary of the photographed area. Although images within one survey occasion created a linear mosaic of intersecting images and could be used as temporal replicates, we selected one image from each survey occasion to use in our analysis. Selecting one image from each survey occasion increased Δj between temporal replicates, increasing the probability that individual sea otters would make state transitions from available to unavailable (or *vice versa*; i.e., so subsequent images were more likely to be independent). Images where a raft of sea otters was in the center of the image were preferred to help ensure closure assumptions were reasonably met. The best image for each group of sea otters and sampling occasion was selected and imported into Count Clusters (Dynamic Ventures, Incorporated, Cupertino, California), a custom software program that can be used for counting objects in digital images. An experienced observer marked each sea otter in the image. The total number of individuals per image as well as attribute data from each digital image including date, time, latitude, longitude, and altitude were exported to a file for analysis. All analyses were conducted in R version 3.2.3 (R Core Team 2013).

Our field-based methods were associated with a pilot study with limited spatial coverage, thus, we did not attempt to estimate abundance for all of GBNP. We collected 60 images from 20 locations containing sea otters (Table 1). We focused our estimates of detection probability and abundance in the sites we surveyed. We visited each site multiple times. However, visiting each site multiple times is not required, and monitoring designs can be made more efficient by coupling the information from sites with multiple visits and sites that are visited once, provided the same standards that are used to collect and analyze overlapping photographs are used in all photographs.

The estimated posterior distributions for detection probability and abundance for these data are provided in Fig. 4. The mean of the posterior distribution, optimal for squared-error loss (e.g., Williams & Hooten 2016), for detection probability from the aerial image data from 20 sites equaled 0.76, the same as the mean of the posterior distribution from the original design-based survey of sea otters estimated from 469 intensively searched sites (Williams *et al.* 2017). We assessed model fit using Bayesian p-values (Hobbs & Hooten 2015). We used the χ^2 goodness-of-fit discrepancy function for calculating Bayesian p-values (Gelman *et al.* 2014). The Bayesian p-value was 0.52, suggesting no lack of model fit.

Discussion

We presented a spatio-temporal point process model, in combination with a novel application of N-mixture models fitted to digital aerial survey data when detection probability is < 1 due to animals being unavailable for detection. We used this framework to simultaneously estimate detection probability, occupancy probability, and total abundance based on an intensity surface that is a realization of a continuous spatio-temporal inhomogeneous Poisson process. Applying our model to both simulated data and real data on sea otters collected during a pilot study demonstrated that this framework is a promising tool for estimating occupancy, abundance, and detection probability from aerial image surveys. Additionally, the spatio-temporal point process model is sufficiently flexible to accommodate count data, presence-absence data, presence only data (Dorazio 2014, Fithian *et al.* 2015, Hefley & Hooten 2016), and spatio-temporal dependence in ecological processes (Cressie & Wikle 2011).

Model estimates based on simulated data recovered true parameters well. This was not surprising, as previous simulation studies have shown that N-mixture models usually perform well for estimating abundance and detection probability for a variety of conditions with varying level of detection probability, few replicate temporal counts, and few sites (see: Kéry *et al.* 2005, Hunt *et al.* 2012, Couturier *et al.* 2013, Yamaura 2013, Dennis *et al.* 2015, McCaffery *et al.* 2016). The novel application in our simulation was the use of the intersection of two overlapping spatial sites as temporal replication. Although the formulation of the model has been used for other purposes, our application extends it to new situations involving survey design. For example, if known-radius point counts (e.g., Henry *et al.* 2015) are conducted to collect data to estimate abundance, sites could be chosen such that neighboring sites intersect to make temporal replication more efficient than replicating visits to sites over multiple days.

The general applicability of these methods for aerial images depends on whether sample sites are closed with respect to movement, mortality, and recruitment. Mortality and recruitment are unlikely during the course of an aerial survey for most populations. However, the movement assumption may not be valid for highly mobile animals. In temporally replicated counts, we assume that the set of animals that occupy a site (but not necessarily observed in the site), is unchanged. Thus, if animals move in or out of the area in the time difference Δj between images, the closure assumption will be violated and estimates of detection probability could be biased depending on how animals disperse in and out of sites. If dispersal is random, bias will likely be small. If animals systematically disperse away from a site after an initial survey, perhaps in response to the aircraft, then bias may be large, unless this dispersal can be modeled. For example, if it is possible to uniquely identify some individuals in repeated images, movement models could be used to explicitly account for animal movement among photographs (e.g., Royle & Young 2008, Hooten *et al.* 2017).

Another assumption of the N-mixture model is that organisms are detected independently of each other (Royle & Dorazio 2008, Martin *et al.* 2011, Dorazio *et al.* 2013). This assumption may be violated if behavior among organisms is correlated, and affects their probability of being detected (e.g., manatees surfacing for air in groups; Martin *et al.* 2011). Martin *et al.* (2011) developed an extension to the N-mixture model for accounting for correlated behavior and non-independent detection of individuals. Our goodness-of-fit evaluation for the sea otter data suggested no lack of model fit, and therefore, there was no evidence that this assumption was violated. However, alternative models (e.g., negative binomial Ver Hoef & Boveng 2007) can be implemented when violations, or lack of model fit occur.

There are at least two design considerations that can help prevent violations in the closure assumption. The first consideration applies mainly to populations that congregate in groups (e.g., rafts of sea otters, flocks of birds, pods of whales, rafts of pinipeds), or are relatively immobile among replicate surveys, and is to use a camera, lens, and altitude combination that produces images that have a larger footprint (c.f., plot size; Efford & Dawson 2012). A larger footprint may provide a buffer around a group of animals, requiring more time for individuals to move out of the footprint (or move from outside the footprint into the footprint). In our application, sea otter movement between subsequent photographs was small, relative to the footprint of the photographs we used. Further, we selectively chose photographs with groups of sea otters located in the center of photographs, reducing the opportunity sea otters had to leave the area captured by photographs. Thus, although it is possible that there was some movement into or out of areas captured by replicate photographs, our survey design minimized this possibility, (which was corroborated by observations from the secondary observer), and any violation of this assumption was small (the secondary observer never witnessed it), and likely negligible for estimating abundance and detection probability of sea otters in GBNP. The second consideration is to reduce time between survey occasions (i.e., decrease Δj) limiting the time animals have to move out of the footprint. During our pilot survey, it required approximately 2–3 minutes to fly over a group of sea otters twice. However, because we acquired several images each time we flew over sea otters, and neighboring images contained overlap, we could have used these intersecting regions to provide temporal replication, minimizing the probability that any otter moved out of the area captured by the intersecting images. However, using multiple images within a transect may result in neighboring images that are not independent, and therefore we selected one image from each transect. Additionally, multiple cameras can be mounted on an aircraft such that one faces forward and one faces backward, programmed on a timer such that they capture an image of the same area at different times in the same flight pass. Similarly, cameras could be placed side-to-side to increase horizontal size of footprints (see Conn *et al.* 2016, Fig. 2, for a picture of this setup). Extensions that use this framework for video surveys are also possible.

Another design consideration for the application of N-mixture models to aerial survey data is the time it requires for animals to switch between the states of unavailable to available. That is, whether Δj is sufficiently large so that intersecting photographs are independent, conditional on $N(A_i)$. Sea otters are good candidates for these models because they are relatively shallow divers with short dive durations (mean dive duration was 85 seconds; Bodkin *et al.* 2004). Further, the calculated aerobic dive limit (cADL) for sea otters is < 5 mins (cADL for juveniles = 3.62 mins; cADL for adults = 4.82 mins; Thometz *et al.* 2015). Gibbs *et al.* (1988) used aerial images to estimate the number of great blue heron (*Ardea herodias*) nests. Many nests were unavailable for detection due to vegetation obstructing nests from the line-of-sight of the camera. Because it is unlikely that vegetation would change within the course of a survey, allowing additional nests to be identified in subsequent images, the estimate of availability for these data might be biased high, and therefore, nest counts biased low.

Aerial images are often taken along irregular flight paths that do not necessarily form a rectangular grid that partitions a domain of interest, making traditional design-based estimates of abundance difficult (Fig. 1; Ver Hoef & Jansen 2014). Our proposed methods used a model-based approach for estimating abundance based on a spatial point process that can be integrated, resulting in a Poisson regression model that matches the scale of the data, and subsequently, could be incorporated in an N-mixture model. In our application, the abundance intensity was determined by the parameters ($\boldsymbol{\beta}$) using generalized linear regression. The model-based approach provides additional flexibility, compared to design-based estimates, that allows incorporation of spatial, temporal, or spatio-temporal autocorrelation (Cressie 1993, Cressie & Wikle 2011, Diggle 2013), and computationally efficient methods for fitting them (Hooten *et al.* 2013, Ver Hoef & Jansen 2014, Hefley *et al.* 2017).

In our study, we achieved temporal replication by identifying a group of sea otters at a location, and then conducting multiple flights over the group and taking images. We used this method to reduce the possibility of spatial displacement by sea otters. Any flight plan could be selected to obtain temporal replication of sites, provided it reasonably meets the assumptions of the model. Additionally, the precision and robustness of parameter estimates, with respect to the number of replicate sites conducted, and the amount of overlap obtained in photographs, can be evaluated using a simulation that is specific to individual study systems.

Finally, if detection probability is likely to change between survey periods (e.g., each year), temporal replication of images can be incorporated into each survey using randomization, or model-based optimization (e.g., Wikle & Royle 1999; 2005, Hooten *et al.* 2009). Alternatively, if detection probability is not likely to change through time, a pilot study could be conducted to examine availability bias, and then used as an informative prior distribution for future aerial surveys, precluding the necessity to conduct replicate surveys during each sampling period.

Acknowledgments

We thank Andy Royle and three anonymous reviewers for valuable insight about this work. Funding was provided by the National Park Service Inventory & Monitoring Program and Glacier Bay National Park & Preserve Marine Management Fund. Heather Coletti, Dan Esler, George Esslinger, and Dan Monson provided technical and logistical support. We are grateful to Louise Taylor-Thomas for processing imagery and Dennis Lozier for piloting the aircraft. Yitzchak Ehrlich provided expertise and assistance with image software. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Data accessibility

All data used in the manuscript are reported in Table 1.

Author contributions

P.J.W., M.B.H., J.N.W., and M.R.B. designed the research. J.N.W. and M.R.B. organized field studies. J.N.W. and P.J.W. collected data. P.J.W. and M.B.H. contributed new analytic tools. P.J.W. developed simulations and analyzed data. P.J.W., M.B.H., J.N.W., and M.R.B. wrote the paper.

References

Baddeley, A., Rubak, E. & Turner, R. (2015) *Spatial point patterns: methodology and applications with R*. CRC Press, Boca Raton FL, USA.

Banerjee, S., Carlin, B.P. & Gelfand, A.E. (2014) *Hierarchical modeling and analysis for spatial data*. CRC Press, Boca Raton FL, USA.

Bayliss, P. & Yeomans, K. (1990) Use of low-level aerial photography to correct bias in aerial survey estimates of magpie goose and whistling duck density in the northern territory. *Australian Wildlife Research* **17**, 1–10.

Bechet, A., Reed, A., Plante, N., Giroux, J.F. & Gauthier, G. (2004) Estimating the size of the greater snow goose population. *Journal of Wildlife Management* **68**, 639–649.

Berliner, L.M. (1996) Hierarchical Bayesian time series models. *Maximum Entropy and Bayesian Methods* (eds. K.M. Hanson & R.N. Silver), pp. 15–22, Kluwer Academic Publishers, Dordrecht, NL.

Bodkin, J.L., Esslinger, G.G. & Monson, D.H. (2004) Foraging depths of sea otters and implications to coastal marine communities. *Marine Mammal Science* **20**, 305–321.

Bodkin, J.L. & Udevitz, M.S. (1999) An aerial survey method to estimate sea otter abundance. *Marine Mammal Survey and Assessment Methods* (eds. G.W. Garner, S.C. Amstrup, J.L. Laake, B.F.J. Manly, L.L. McDonald & D.G. Robertson), pp. 13–26, A. A. Balkema, Leiden, NL.

Boyd, W.S. (2000) A comparison of photo counts versus visual estimates for determining the size of snow goose flocks. *Journal of Field Ornithology* **71**, 686–690.

Buckland, S.T., Burt, M.L., Rexstad, E.A., Mellor, M., Williams, A.E. & Woodward, R. (2012) Aerial surveys of seabirds: the advent of digital methods. *Journal of Applied Ecology* **49**, 960–967.

Caughley, G. (1974) Bias in aerial survey. *The Journal of Wildlife Management* **38**, 921–933.

Caughley, G. & Goddard, J. (1972) Improving the estimates from inaccurate censuses. *The Journal of Wildlife Management* **36**, 135–140.

Certain, G. & Bretagnolle, V. (2008) Monitoring seabirds population in marine ecosystem: the use of strip-transect aerial surveys. *Remote Sensing of Environment* **112**, 3314–3322.

Conn, P.B., Johnson, D.S., Ver Hoef, J.M., Hooten, M.B., London, J.M. & Boveng, P.L. (2015) Using spatiotemporal statistical models to estimate animal abundance and infer ecological dynamics from survey counts. *Ecological Monographs* **85**, 235–252.

Conn, P.B., Moreland, E.E., Regehr, E.V., Richmond, E.L., Cameron, M.F. & Boveng, P.L. (2016) Using simulation to evaluate wildlife survey designs: polar bears and seals in the chukchi sea. *Royal Society Open Science* **3**, 150561.

Conn, P.B., Ver Hoef, J.M., McClintock, B.T., Moreland, E.E., London, J.M., Cameron, M.F., Dahle, S.P. & Boveng, P.L. (2014) Estimating multispecies abundance using automated detection systems: ice-associated seals in the bering sea. *Methods in Ecology and Evolution* **5**, 1280–1293.

Couturier, T., Cheylan, M., Bertolero, A., Astruc, G. & Besnard, A. (2013) Estimating abundance and population trends when detection is low and highly variable: a comparison of three methods for the hermann's tortoise. *The Journal of Wildlife Management* **77**, 454–462.

Cox, D.R. (1955) Some statistical methods connected with series of events. *Journal of the Royal Statistical Society. Series B (Methodological)* **2**, 129–164.

Cressie, N. (1993) *Statistics for spatial data*. Wiley-Interscience, New York, USA.

Cressie, N. & Wikle, C.K. (2011) *Statistics for spatio-temporal data*. John Wiley & Sons, Hoboken, New Jersey, USA.

Dennis, E.B., Morgan, B.J. & Ridout, M.S. (2015) Computational aspects of n-mixture models. *Biometrics* **71**, 237–246.

Diggle, P.J. (2013) *Statistical analysis of spatial and spatio-temporal point patterns*. CRC Press, Boca Raton FL, USA.

Dorazio, R.M. (2014) Accounting for imperfect detection and survey bias in statistical analysis of presence-only data. *Global Ecology and Biogeography* **23**, 1472–1484.

Dorazio, R.M., Martin, J. & Edwards, H.H. (2013) Estimating abundance while accounting for rarity, correlated behavior, and other sources of variation in counts. *Ecology* **94**, 1472–1478.

Efford, M.G. & Dawson, D.K. (2012) Occupancy in continuous habitat. *Ecosphere* **3**, 1–15.

Estes, J.A. & Palmisano, J.F. (1974) Sea otters: their role in structuring nearshore communities. *Science* **185**, 1058–1060.

Fithian, W., Elith, J., Hastie, T. & Keith, D.A. (2015) Bias correction in species distribution models: pooling survey and collection data for multiple species. *Methods in Ecology and Evolution* **6**, 424–438.

Frederick, P.C., Hylton, B., Heath, J.A. & Ruane, M. (2003) Accuracy and variation in estimates of large numbers of birds by individual observers using an aerial survey simulator. *Journal of Field Ornithology* **74**, 281–287.

Gelfand, A.E. & Smith, A.F. (1990) Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398–409.

Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. (2014) *Bayesian data analysis*, vol. 2. Chapman & Hall/CRC Boca Raton, FL, USA.

Gibbs, J.P., Woodward, S., Hunter, M.L. & Hutchinson, A.E. (1988) Comparison of techniques for censusing great blue heron nests. *Journal of Field Ornithology* **59**, 130–134.

Goddard, J. (1967) The validity of censusing black rhinoceros populations from the air. *African Journal of Ecology* **5**, 18–23.

Goddard, J. (1969) Aerial census of black rhinoceros using stratified random sampling. *African Journal of Ecology* **7**, 105–114.

Gotway, C.A. & Young, L.J. (2002) Combining incompatible spatial data. *Journal of the American Statistical Association* **97**, 632–648.

Graham, A. & Bell, R. (1969) Factors influencing the countability of animals. *East African Agricultural and Forestry Journal* **34**, 38–43.

Hefley, T.J., Broms, K.M., Brost, B.M., Buderman, F.E., Kay, S.L., Scharf, H.R., Tipton, J.R., Williams, P.J. & Hooten, M.B. (2017) The basis function approach for modeling autocorrelation in ecological data. *Ecology* **98**, 632–646.

Hefley, T.J. & Hooten, M.B. (2016) Hierarchical species distribution models. *Current Landscape Ecology Reports* **1**, 87–97.

Heide-Jørgensen, M.P., Laidre, K., Borchers, D., Samarra, F. & Stern, H. (2007) Increasing abundance of bowhead whales in west greenland. *Biology Letters* **3**, 577–580.

Henry, E.H., Haddad, N.M., Wilson, J., Hughes, P. & Gardner, B. (2015) Point-count methods to monitor butterfly populations when traditional methods fail: a case study with miami blue butterfly. *Journal of Insect Conservation* **19**, 519–529.

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

Hobbs, N.T. & Hooten, M.B. (2015) *Bayesian models: a statistical primer for ecologists*. Princeton University Press, Princeton, New Jersey, USA.

Hodgson, A., Kelly, N. & Peel, D. (2013) Unmanned aerial vehicles (uavs) for surveying marine fauna: a dugong case study. *Plos One* **8**, e79556.

Hooten, M.B., Garlick, M.J. & Powell, J.A. (2013) Computationally efficient statistical differential equation modeling using homogenization. *Journal of Agricultural, Biological and Environmental Statistics* **18**, 405–428.

Hooten, M.B., Johnson, D.S., McClintock, B.T. & Morales, J.M. (2017) *Animal movement: statistical models for telemetry data*. CRC Press, Boca Raton FL, USA.

Hooten, M.B., Wikle, C.K., Sheriff, S.L. & Rushin, J.W. (2009) Optimal spatio-temporal hybrid sampling designs for ecological monitoring. *Journal of Vegetation Science* **20**, 639–649.

Hunt, J.W., Weckerly, F.W. & Ott, J.R. (2012) Reliability of occupancy and binomial mixture models for estimating abundance of golden-cheeked warblers (*Setophaga chrysoparia*). *The Auk* **129**, 105–114.

Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. (2008) *Statistical analysis and modelling of spatial point patterns*, vol. 70. John Wiley & Sons, Hoboken, New Jersey, USA.

Jolly, G. (1969a) Sampling methods for aerial censuses of wildlife populations. *East African Agricultural and Forestry Journal* **34**, 46–49.

Jolly, G. (1969b) The treatment of errors in aerial counts of wildlife populations. *East African Agricultural and Forestry Journal* **34**, 50–55.

Kendall, W.L. & Nichols, J.D. (1995) On the use of secondary capture-recapture samples to estimate temporary emigration and breeding proportions. *Journal of Applied Statistics* **22**, 751–762.

Kéry, M., Royle, J.A. & Schmid, H. (2005) Modeling avian abundance from replicated counts using binomial mixture models. *Ecological Applications* **15**, 1450–1461.

Leedy, D.L. (1948) Aerial photographs, their interpretation and suggested uses in wildlife management. *The Journal of Wildlife Management* **12**, 191–210.

Leonard, R.M. & Fish, E.B. (1974) An aerial photographic technique for censusing lesser sandhill cranes. *Wildlife Society Bulletin* **2**, 191–195.

Lukacs, P.M., Kissling, M.L., Reid, M., Gende, S.M. & Lewis, S.B. (2010) Testing assumptions of distance sampling on a pelagic seabird. *The Condor* **112**, 455–459.

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

Martin, J., Edwards, H.H., Burgess, M.A., Percival, H.F., Fagan, D.E., Gardner, B.E., Ortega-Ortiz, J.G., Ifju, P.G., Evers, B.S. & Rambo, T.J. (2012) Estimating distribution of hidden objects with drones: From tennis balls to manatees. *PLoS One* **7**, e38882.

Martin, J., Royle, J.A., Mackenzie, D.I., Edwards, H.H., Kery, M. & Gardner, B. (2011) Accounting for non-independent detection when estimating abundance of organisms with a Bayesian approach. *Methods in Ecology and Evolution* **2**, 595–601.

McCaffery, R., Nowak, J.J. & Lukacs, P.M. (2016) Improved analysis of lek count data using n-mixture models. *The Journal of Wildlife Management* **80**, 1011–1021.

McNabb, R.W., Womble, J.N., Prakash, A., Gens, R. & Haselwimmer, C.E. (2016) Quantification and analysis of icebergs in a tidewater glacier fjord using an object-based approach. *Plos One* **11**, e0164444.

Moller, J. & Waagepetersen, R.P. (2003) *Statistical inference and simulation for spatial point processes*. CRC Press, Boca Raton FL, USA.

Oppel, S., Meirinho, A., Ramírez, I., Gardner, B., OConnell, A.F., Miller, P.I. & Louzao, M. (2012) Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. *Biological Conservation* **156**, 94–104.

Pennycuick, C. & Western, D. (1972) An investigation of some sources of bias in aerial transect sampling of large mammal populations. *African Journal of Ecology* **10**, 175–191.

R Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Royle, J.A. (2004) N-mixture models for estimating population size from spatially replicated counts. *Biometrics* **60**, 108–115.

Royle, J.A. & Dorazio, R.M. (2008) *Hierarchical Modeling and Inference in Ecology: The Analysis of Data From Populations, Metapopulations, and Communities*. Academic Press, Cambridge, Massachusetts, USA.

Royle, J.A. & Young, K.V. (2008) A hierarchical model for spatial capture-recapture data. *Ecology* **89**, 2281–2289.

Siniff, D.B. & Skoog, R.O. (1964) Aerial censusing of caribou using stratified random sampling. *The Journal of Wildlife Management* **28**, 391–401.

Sweeney, K.L., Helker, V.T., Perryman, W.L., LeRoi, D.J., Fritz, L.W., Gelatt, T.S. & Angliss, R.P. (2015) Flying beneath the clouds at the edge of the world: using a hexacopter to supplement abundance surveys of steller sea lions (*Eumetopias jubatus*) in alaska. *Journal of Unmanned Vehicle Systems* **4**, 70–81.

Thometz, N.M., Murray, M.J. & Williams, T.M. (2015) Ontogeny of oxygen storage capacity and diving ability in the southern sea otter (*Enhydra lutris nereis*): costs and benefits of large lungs. *Physiological and Biochemical Zoology* **88**, 311–327.

Ver Hoef, J.M. & Boveng, P.L. (2007) Quasi-Poisson vs. negative binomial regression: how should we model overdispersed count data? *Ecology* **88**, 2766–2772.

Ver Hoef, J.M. & Jansen, J.K. (2014) Estimating abundance from counts in large data sets of irregularly spaced plots using spatial basis functions. *Journal of Agricultural, Biological and Environmental Statistics* **20**, 1–27.

Watson, R., Parker, I. & Allan, T. (1969) A census of elephant and other large mammals in the mkomazi region of northern Tanzania and southern Kenya. *African Journal of Ecology* **7**, 11–26.

Wikle, C.K. & Royle, J.A. (1999) Space: time dynamic design of environmental monitoring networks. *Journal of Agricultural, Biological and Environmental Statistics* **4**, 489–507.

Wikle, C.K. & Royle, J.A. (2005) Dynamic design of ecological monitoring networks for non-gaussian spatio-temporal data. *Environmetrics* **16**, 507–522.

Williams, P.J. & Hooten, M.B. (2016) Combining statistical inference and decisions in ecology. *Ecological Applications* **26**, 1930–1942.

Williams, P.J., Hooten, M.B., Womble, J.N., Esslinger, G.G., Bower, M. & Hefley, T.J. (2017) An integrated data model to estimate spatio-temporal occupancy, abundance, and colonization dynamics. *Ecology* **98**, 328–336.

Yamaura, Y. (2013) Confronting imperfect detection: behavior of binomial mixture models under varying circumstances of visits, sampling sites, detectability, and abundance, in small-sample situations. *Ornithological Science* **12**, 73–88.

Supporting Information

Appendix S1

Script to simulate sea otter abundance, simulate sampling data, fit N-mixture model to sampling data, and examine output.

Appendix S2

Script to fit N-mixture model to sea otter data, and examine output.

Table 1: Counts of sea otters from aerial images taken at 20 sites in Glacier Bay National Park, Alaska. Sampling occasion refers to the number of times a site was flown over in an aircraft and a picture was taken of the same group of sea otters.

	Sampling occasion				
	1	2	3	4	5
Site 1	20	17	15	15	
Site 2	60	62	58	55	
Site 3	15	16	15	15	
Site 4	8	12			
Site 5	9	10			
Site 6	19	20	19	19	
Site 7	17	17	17	13	
Site 8	52	53			
Site 9	162	171			
Site 10	37	40			
Site 11	144	138			
Site 12	21	25	17		
Site 13	20	19	18	18	
Site 14	86	83	87	91	
Site 15	47	46			
Site 16	21	20			
Site 17	19	12			
Site 18	2	1	1		
Site 19	83	85	83	85	83
Site 20	55	48	52		

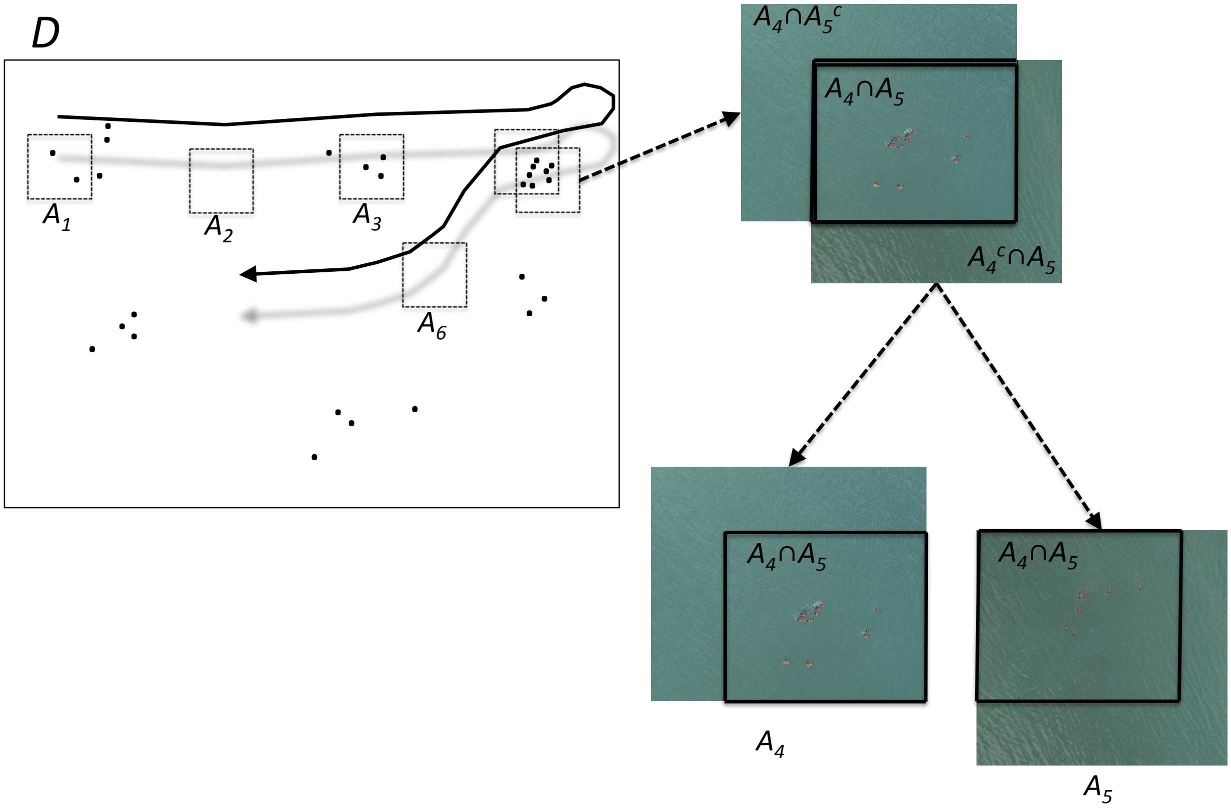

Figure 1: Depiction of aerial survey of domain D, with image locations $A_1, \ldots, A_n$. Overlapping images provide temporal replication at sites where the images intersect.

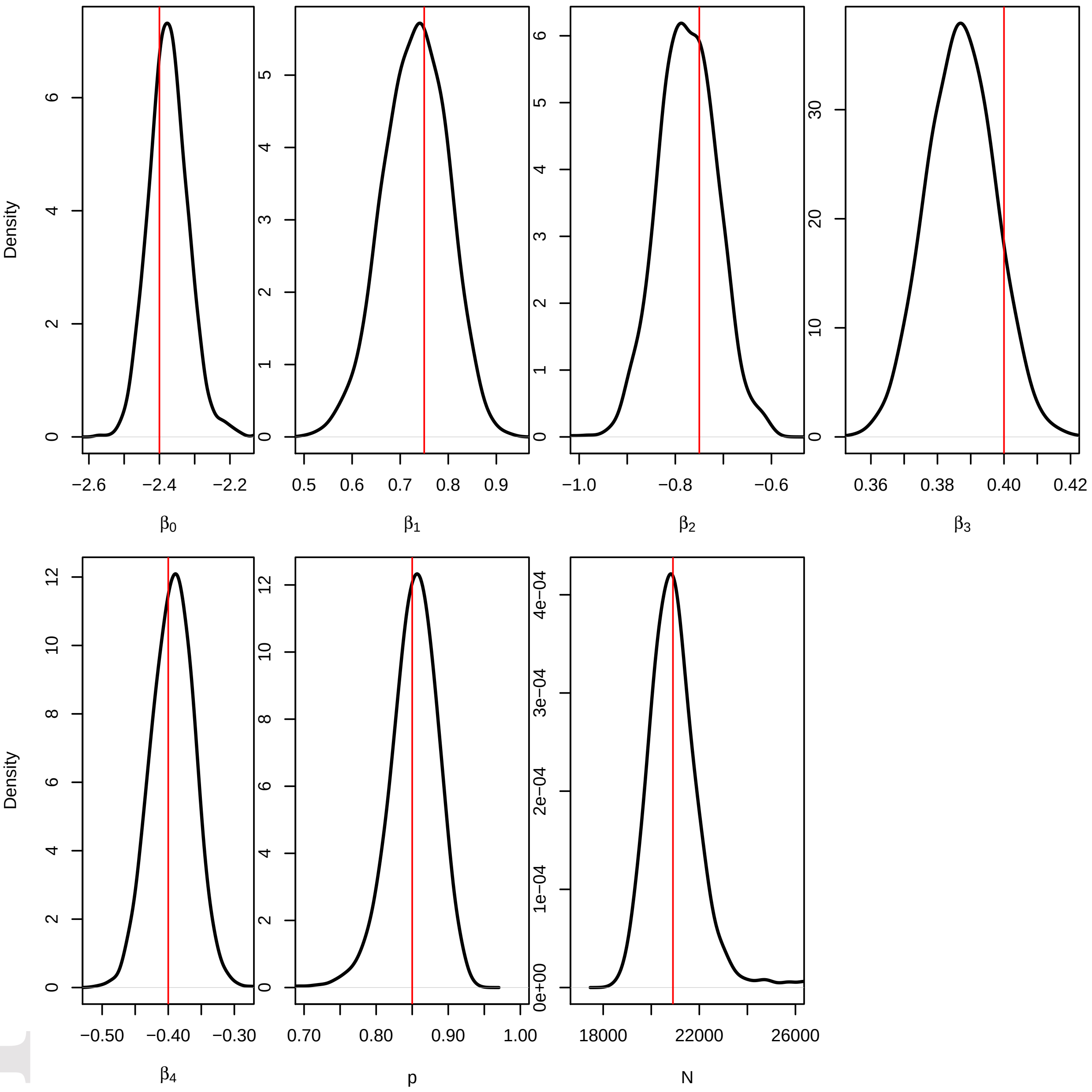

Figure 2: Marginal estimated posterior distributions (black lines) of parameters in aerial sea otter abundance model fit to simulated data. Red lines indicate values used to simulate data (see Appendix S1).

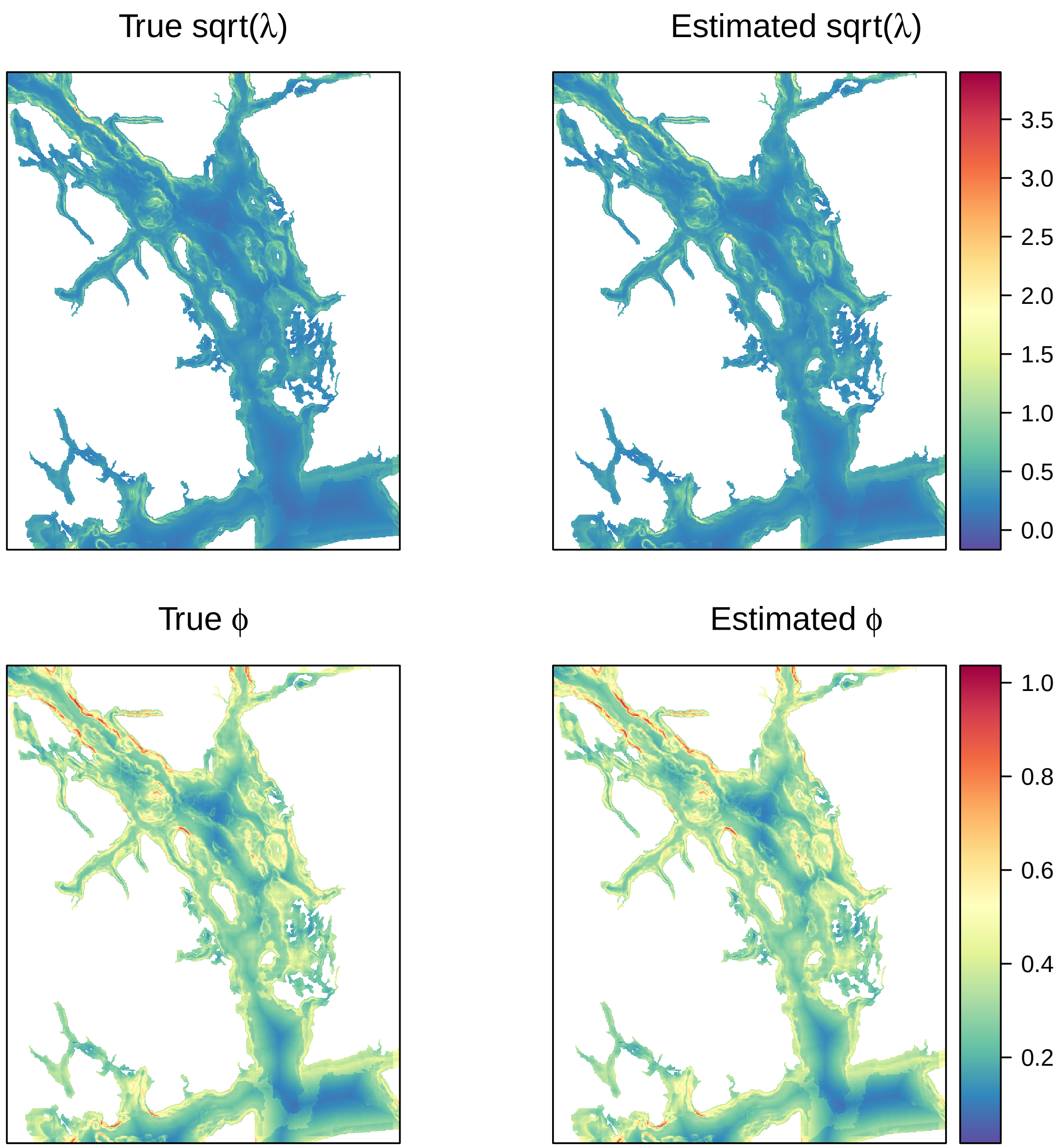

Figure 3: Left: True expected abundance ($\lambda(\mathbf{s})$; top) and occupancy probability ($\phi(\mathbf{s})$; bottom) used to simulate abundance data for sea otters in Glacier Bay National Park, southeastern Alaska, USA. Right: The estimated expected abundance and occupancy probability using simulated aerial photographs. The square-root of $\lambda(\mathbf{s})$ was used to highlight spatial variation.

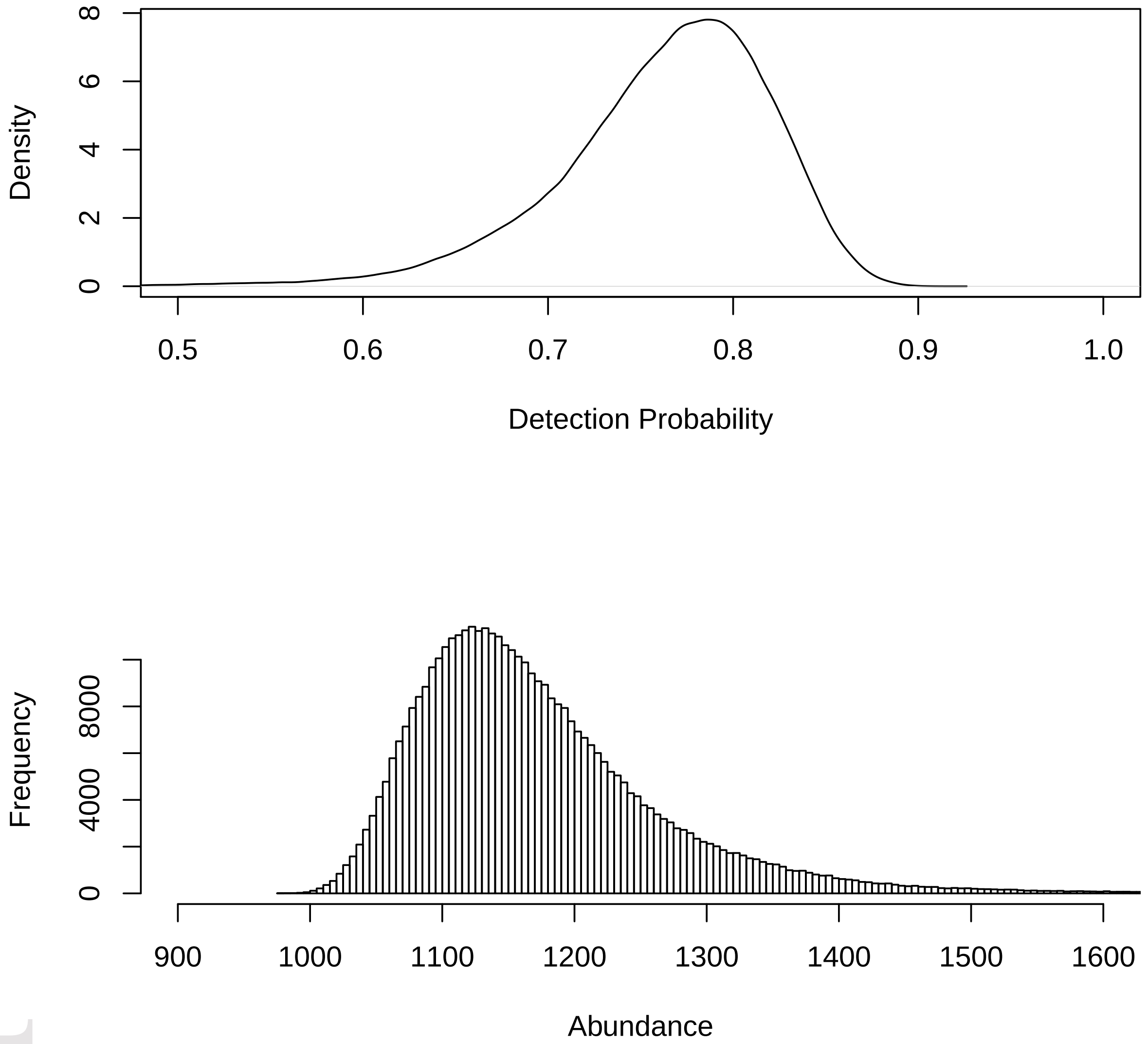

Figure 4: Marginal estimated posterior distributions of detection probability and abundance from aerial image data collected at 20 sites containing sea otters in Glacier Bay National Park, Alaska (see Appendix S2).