A hierarchical point process model for spatial capture-recapture data

Wei Zhang^{1,*}, Joseph D. Chipperfield^{2,3,*}, Janine B. Illian⁴, Pierre Dupont², Cyril Milleret², Perry de Valpine¹, Richard Bischof²

¹Department of Environmental Science, Policy and Management, University of California Berkeley, Berkeley, USA

²Faculty of Life Sciences and Natural Resource Management, Norwegian University of Life Sciences, Høgskoleveien 12, Ås, Norway

³Norwegian Institute for Nature Research, Hyteknologisenteret, Thormøhlensgate 55, Bergen, Norway

⁴School of Mathematics and Statistics, University of Glasgow, Glasgow, UK

Correspondence

Wei Zhang

Department of Environmental Science, Policy and Management

University of California, Berkeley, 130 Mulford Hall, Berkeley, CA, USA

Email: weizhang0721@berkeley.edu

Running headline

A hierarchical point process SCR model

*The authors contributed equally to the manuscript

Abstract

1. Spatial capture-recapture (SCR) is a popular method for estimating the abundance and density of wildlife populations. A standard SCR model consists of two sub-models: one for the activity centers of individuals and the other for the detections of each individual conditional on its activity center. So far, the detection sub-model of most SCR models is designed for sampling situations where fixed trap arrays are used to detect individuals.
2. Non-invasive genetic sampling (NGS) is widely applied in SCR studies. Using NGS methods, one often searches the study area for potential sources of DNA such as hairs and faeces, and records the locations of these samples. To analyse such data with SCR models, investigators usually impose an artificial detector grid and project detections to the nearest detector. However, there is a trade-off between the computational efficiency (fewer detectors) and the spatial accuracy (more detectors) when using this method.
3. Here, we propose a point process model for the detection process of SCR studies using NGS. The model better reflects the spatially continuous detection process and allows all spatial information in the data to be used without approximation error. As in many SCR models, we also use a point process model for the activity centers of individuals. The resulting hierarchical point process model enables estimation of total population size without imputing unobserved individuals via data augmentation, which can be computationally cumbersome. We write custom distributions for those spatial point processes and fit the SCR model in a Bayesian framework using Markov chain Monte Carlo in the R package nimble.
4. Simulations indicate good performance of the proposed model for parameter estimation. We demonstrate the application of the model in a real-life scenario by fitting it to NGS data of female wolverines (*Gulo gulo*) collected in three counties of Norway during the winter of

2018/19. Our model estimates that the density of female wolverines is 9.53 (95% CI: 8–11) per 10,000km² in the study area.

Keywords: Area search, Binomial point process, Continuous sampling, NIMBLE, Non-invasive genetic sampling, Poisson point process, Spatial capture-recapture, Wolverine

Introduction

Capture-recapture has a long history for estimating demographic parameters of wildlife populations such as abundance, survival rate, and density or relative abundance, which play a key role in wildlife conservation and management. Spatial capture-recapture (SCR), an extension to capture-recapture, has enjoyed rapid growth in the past decade. SCR models (see Borchers & Fewster, 2016, for a review) extend traditional capture-recapture models by incorporating the activity centers of individuals into the modelling framework as latent variables. The probability that an individual is detected by a detector (e.g., a trap) depends on the distance between the activity center of the individual and the detector. SCR models can therefore estimate spatially-explicit abundance of a population.

Standard SCR models are hierarchical and comprise two components: one for modelling the activity centers of individuals, and the other for modelling the detections of each individual conditional on its activity center. It is common to model the activity centers by a spatial point process since Efford (2004) used one for the first SCR model; see Chandler & Royle (2013); Dorazio (2013); Borchers *et al.* (2014) and Borchers & Marques (2017) for examples. A spatial point process describes the distribution of points in space, with both the number and locations of points being random. These models have been widely used for analysing spatial data in diverse fields such as histology, epidemiology, and seismology among many others (see Baddeley *et al.*, 2006).

The detection model in SCR depends on the type of detectors used for data collection. Most detection models to date were developed for sampling situations in which the set of possible detection locations is fixed. This is the case when the data are collected using restraining devices such as live traps (see Powell & Proulx, 2003, for examples) or non-invasive techniques such as camera traps (Trolliet *et al.*, 2014; Glover-Kapfer *et al.*, 2019), hair snares (Henry *et al.*, 2011), and acoustic arrays (Thomas & Marques, 2012). Other detection models have also been proposed for cases where detections are not necessarily associated with fixed locations. For example, Royle

72 & Young (2008) proposed a SCR model to investigate temporary emigration of individuals. In
 73 their case, data are collected by an exhaustive area-search for individuals within a delineated
 74 sample unit, thus individuals can be detected anywhere in the unit. Similarly, Royle *et al.* (2011a)
 75 developed a model for ‘search-encounter’ SCR data that are collected by surveyors searching along
 76 transects.

77 Despite their variety (Royle *et al.*, 2014; Efford, 2020), existing detector-based or search-
 78 encounter models do not adequately cover all common SCR sampling processes. Non-invasive
 79 genetic sampling (NGS) is now commonly used for collecting SCR data (Burgar *et al.*, 2018;
 80 Ferreira *et al.*, 2018; López-Bao *et al.*, 2018). NGS in SCR studies involves the collection of DNA
 81 sources such as shed hair, faeces, saliva, and urine at fixed detector or area searches, followed by
 82 individual identification (Lamb *et al.*, 2019). When NGS is implemented by searching a given area
 83 or along transects, paths taken by searchers can be recorded and used as a direct measure of effort in
 84 space and time. In praxis, this is not always possible, practical, or reliable. The model described in
 85 this manuscript is motivated by our experiences with the Scandinavian large carnivore monitoring
 86 program (Bischof *et al.*, 2019), one of the world’s most extensive NGS projects (as of December
 87 2019: over 35,000 samples from 6,000 individual bears (*Ursus arctos*), wolves (*Canis lupus*), and
 88 wolverines). Although field protocols prescribe the recording of GPS tracks during the performance
 89 of NGS by Norwegian and Swedish government employees, due to technical limitations and logistic
 90 constraints, these cannot always be linked unambiguously with detections. Furthermore, the region
 91 monitored is vast (over 500,000 km²) and a substantial portion of genetic samples are obtained
 92 via opportunistic detections by members of the public (hunters, hikers, berry pickers, etc.). As
 93 a consequence, no direct measure of effort is available and investigators wanting to account for
 94 spatial heterogeneity in detection probability are left with proxies for effort such as distance to
 95 roads and snow cover (Bischof *et al.*, 2019). In cases such as this, detections, at least in theory, are
 96 possible at any location within the general area that humans could visit. An approach that uses
 97 actual detection locations would be preferable to the typical approach of projecting detections to

an artificial detection grid (Russell *et al.*, 2012; Lonsinger *et al.*, 2018; López-Bao *et al.*, 2018) as the latter: (a) means an approximation of detection locations, (b) involves aggregation of detection information, and (c) forces investigators to trade off precision for computational efficiency (Milleret *et al.*, 2018).

To deal with SCR data collected through area searches, we propose a hierarchical point process modelling framework. As in most SCR models in the literature, we model the activity centers of individuals by a spatial point process, which in our case is an inhomogenous Poisson point process. We assume these activity centers do not change during the study, although it is possible to extend the framework proposed here to include movement. A second Poisson point process is used to model the detections of an individual conditional on its activity center.

A Poisson point process is determined by an intensity function. Evaluating the intensity function at point $\mathbf{x}$ gives the probability that there is one point (one detection or activity center in our case) falling within an infinitesimal volume centered around $\mathbf{x}$ (p.51, Chiu *et al.*, 2013). The detection intensity for one individual at one point in the study area is a (typically decreasing) function of the distance between the individual’s activity center and that point. The expected number of detections for one individual in the study area or any subregion is obtained by integrating the intensity function over the area or subregion. By eliminating the need for an artificial grid of detectors, the point process model represents more closely the spatially-continuous data collection process. All spatial information in the data can be used without additional post-collection approximation error. We fit the proposed hierarchical model in a Bayesian framework using Markov chain Monte Carlo (MCMC).

Estimating total population size when fitting SCR models using MCMC leads to technical challenges because it involves estimating the number of individuals that were completely unobserved. Most authors do this using data augmentation (Royle *et al.*, 2007; Royle & Dorazio, 2012), in which unobserved individuals, their activity centers, and their state (real or virtual) are imputed as part of the MCMC posterior sampling. This can be computationally costly, especially when

detection rates are low and hence there may be many unobserved individuals. A more efficient solution is to construct a ‘semi-complete data likelihood’, which is the product of a complete data likelihood of observed individuals and a marginal likelihood of unobserved individuals obtained by integrating out their activity centers (King *et al.*, 2016). Then, posterior sampling for model parameters and the activity centers of observed individuals can be achieved using MCMC and the semi-complete data likelihood without data augmentation. This is achieved by deriving the marginal detection probability, which is calculated by numerical integration.

We investigate performance of the proposed model and estimation method for recovering parameters using a simulation study. We also apply the model to analysis of NGS data of female wolverines collected during the winter of 2018/19 in three counties of Norway. Due to the spatial scale of the study area and the nature of the detection sampling, these data are difficult to analyze with other approaches for modelling detection.

Materials and methods

Notation

A set of notation used in following sections is summarized here:

- N : the population size.
- N^* : the number of individuals detected at least once.
- D : the dimension of the spatial domain (study area): $D = 2$.
- M_i : the observed number of detections for individual $i = 1, \dots, N$.
- $D_{i,\mathbf{o}}$: the number of detections of individual i across the entire region $\mathbf{o}$.
- $\tilde{\mathbf{o}}$: the entire region for the placement of activity centers (habitat selection).
- $\mathbf{o}$: the entire region where individuals can be detected; $\mathbf{o}$ may or may not be the same as $\tilde{\mathbf{o}}$.

- 146 • $\tilde{\mathbf{w}}_h$: a sub-window of $\tilde{\mathbf{o}}$; $h = 1, \dots, H$.
- 147 • $\mathbf{w}_l$: a sub-window of $\mathbf{o}$; $l = 1, \dots, L$.
- 148 • $\mathbf{s}_i$: location of the activity center of individual i ; $\mathbf{s}_i \in \tilde{\mathbf{o}}$; $i = 1, \dots, N$.
- 149 • $\mathbf{d}_{i,j}$: location of the j -th detection for individual i ; $\mathbf{d}_{i,j} \in \mathbf{o}$; $i = 1, \dots, N$, $j = 1, \dots, M_i$.
- 150 • $\boldsymbol{\beta}$: a vector of parameters that determine the intensity function of the point process model
- 151 for habitat selection.
- 152 • $\boldsymbol{\theta}$: a vector of parameters that determine the baseline detection function.
- 153 • $\boldsymbol{\sigma}$: a vector of parameters that determine the (general) detection kernel function.
- 154 • σ : a scalar parameter that determines the (specific) isotropic multivariate Gaussian detection
- 155 kernel function, which is given in Appendix II.
- 156 • $\tilde{\lambda}(\mathbf{s} \mid \boldsymbol{\beta})$: the intensity function of the point process model for habitat selection.
- 157 • $\lambda(\mathbf{d} \mid \mathbf{s}, \boldsymbol{\theta}, \boldsymbol{\sigma})$: the intensity function of the point process model for detection.
- 158 • $\tilde{\Lambda}(\tilde{\mathbf{o}} \mid \boldsymbol{\beta})$: the expected number of activity centers within the region $\tilde{\mathbf{o}}$.
- 159 • $\Lambda(\mathbf{o} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma})$: the expected number of detections of individual i within the region $\mathbf{o}$.
- 160 • $\tilde{\Lambda}(\tilde{\mathbf{w}} \mid \boldsymbol{\beta})$: the expected number of activity centers within sub-window $\tilde{\mathbf{w}}$.
- 161 • $\Lambda(\mathbf{w} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma})$: the expected number of detections of individual i within sub-window $\mathbf{w}$.

162 For convenience, we use $\int_{\Omega} f(\mathbf{x}) d\mathbf{x}$ to denote the multi-dimensional integral of $f(\mathbf{x})$ over the region

163 Ω , which can be $\mathbf{o}$, $\tilde{\mathbf{o}}$, $\mathbf{w}$ or $\tilde{\mathbf{w}}$.

Habitat selection

The activity center of an individual represents the center of the individual's home range, around which the individual moves. Spatial point processes have been widely used for modelling the placement of activity centers since Efford (2004); see Illian *et al.* (2007) for an introduction to spatial point processes. In most studies, such as Chandler & Royle (2013), a homogenous Poisson point process is used, whose intensity function $\tilde{\lambda}(\mathbf{s} | \boldsymbol{\beta})$ is a constant across $\tilde{\mathbf{o}}$, the entire region of interest. Although less widely used, inhomogenous Poisson point processes have also gained some applications in modelling habitat selection (e.g., Dorazio, 2013; Borchers *et al.*, 2014; Borchers & Marques, 2017). Using an inhomogeneous Poisson process, it is possible to model a spatially varying intensity surface, which allows for a higher density of activity centers in areas with higher habitat quality, and to study second-order habitat selection (Johnson, 1980).

Following Illian *et al.* (2007, p.121), the log-probability density of there being N activity centers, $\mathbf{s}_1, \dots, \mathbf{s}_N$, over the entire region $\tilde{\mathbf{o}}$ is

$$\log \mathbb{P}(\mathbf{s}_1, \dots, \mathbf{s}_N | \boldsymbol{\beta}) = -\tilde{\Lambda}(\tilde{\mathbf{o}} | \boldsymbol{\beta}) + \sum_{i=1}^N \log \tilde{\lambda}(\mathbf{s}_i | \boldsymbol{\beta}), \quad (1)$$

where

$$\tilde{\Lambda}(\tilde{\mathbf{o}} | \boldsymbol{\beta}) = \int_{\tilde{\mathbf{o}}} \tilde{\lambda}(\mathbf{s} | \boldsymbol{\beta}) d\mathbf{s} \quad (2)$$

is the expected number of activity centers over $\tilde{\mathbf{o}}$.

Detection intensity

Key to formulating a model for the detection process in SCR is to construct a detection intensity function, which may be related to the observed data in different ways. For example, if one individual can be detected by one detector multiple times, then the number of detections for the individual at that detector can be modelled by a Poisson distribution with mean equal to the intensity (see López-Bao *et al.*, 2018, for example). In cases where an individual can only be detected once

by a given detector during one sampling period, the intensity can be defined as the probability of detecting the individual at that detector. The detection outcome for the individual at that detector follows a Bernoulli distribution (see Royle *et al.*, 2011b, for example). In addition, to improve computational efficiency the Bernoulli detection model can be adequately approximated by a binomial model with the detection locations aggregated at coarser spatial scales (Milleret *et al.*, 2018).

For most current SCR models and the model we develop below, the detection intensity is of the following form

$$\lambda(\mathbf{d} \mid \mathbf{s}, \boldsymbol{\theta}, \boldsymbol{\sigma}) = b(\mathbf{d} \mid \boldsymbol{\theta}) \tau(\mathbf{d} \mid \mathbf{s}, \boldsymbol{\sigma}), \quad (3)$$

where $\mathbf{d}$ denotes the location of the detection, $\mathbf{s}$ denotes the activity center, $\tau(\mathbf{d} \mid \mathbf{s}, \boldsymbol{\sigma})$ is a kernel describing the distance decay relationship of the detection intensity from the activity center, and $b(\mathbf{d} \mid \boldsymbol{\theta})$ is a function describing the baseline detection intensity at the detection location. The function $\tau(\mathbf{d} \mid \mathbf{s}, \boldsymbol{\sigma})$ describes the decreasing chance of detecting an individual as the distance from the activity center increases and is therefore a functional description of the home range of the individual. The function $b(\mathbf{d} \mid \boldsymbol{\theta})$ represents the heterogeneity in detection probability that might result from differences in one or more aspects of the following: (1) sampling effort, (2) the effects of the landscape on individual detectability, and (3) trap efficiency such as the type of camera traps (Efford *et al.*, 2013).

Spatial point process model for detection

When SCR data are collected by searching the study area, the location of each detection can be any point in $\mathcal{O}$. Here, we model the detections of each individual using a spatial point process. A class of spatial point process models that serves as an initial framework for this purpose is the inhomogeneous Poisson process.

The log-probability density for the M_i detection locations, $\mathbf{d}_{i,1} \dots \mathbf{d}_{i,M_i}$, of individual i given

that its activity center is $\mathbf{s}_i$ is:

$$\log \mathbb{P}(\mathbf{d}_{i,1}, \dots, \mathbf{d}_{i,M_i} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}) = -\Lambda(\mathbf{o} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}) + \sum_{j=1}^{M_i} \log \lambda(\mathbf{d}_{i,j} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}), \quad (4)$$

where

$$\Lambda(\mathbf{o} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}) = \int_{\mathbf{o}} \lambda(\mathbf{x} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}) d\mathbf{x} \quad (5)$$

is the expected number of detections of individual i over the entire detection region. We assume that detections of each of the N individuals are conditionally independent given their respective activity centers. It follows that the log-probability density for the detections of all individuals in the population is the sum of the individual log-probability densities:

$$\begin{aligned} \log \mathbb{P}(\mathbf{d}_{1,1}, \dots, \mathbf{d}_{1,M_1}, \dots, \mathbf{d}_{N,1}, \dots, \mathbf{d}_{N,M_N} \mid \mathbf{s}_1, \dots, \mathbf{s}_N, \boldsymbol{\theta}, \boldsymbol{\sigma}) \\ = \sum_{i=1}^N \log \mathbb{P}(\mathbf{d}_{i,1}, \dots, \mathbf{d}_{i,M_i} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}) \\ = - \sum_{i=1}^N \Lambda(\mathbf{o} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}) + \sum_{i=1}^N \sum_{j=1}^{M_i} \log \lambda(\mathbf{d}_{i,j} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}). \end{aligned} \quad (6)$$

Marginal void probability

In all SCR studies, there are some individuals in the population that are never detected. The probability of an individual in the population never being detected throughout the study plays a key role in formulating our hierarchical modelling framework below. For spatial point processes, this is related to the so-called ‘void probability’, the probability that no points fall within a given region. From Equation (25), the void probability $\mathbb{P}(D_{i,\mathbf{o}} = 0 \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma})$ for individual i is

$$\mathbb{P}(D_{i,\mathbf{o}} = 0 \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}) = \exp \{-\Lambda(\mathbf{o} \mid \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma})\}. \quad (7)$$

The void probability depends on the unknown activity center $\mathbf{s}_i$ of individual i , so it cannot be used for our purpose directly. We need the marginal void probability, which is derived below.

A known result of the Poisson point process is that conditioning on the total number of points in a given region yields a binomial point process (p.69, Illian *et al.*, 2007), a point process with a fixed number of points. For a binomial point process, the probability density of one point's location is the intensity of the corresponding Poisson process evaluated at that point divided by the integral of the intensity function over the entire region. It follows that given individual i exists in the population, the probability density that it has activity center $\mathbf{s}_i$ is

$$\mathbb{P}(\mathbf{s}_i | \boldsymbol{\beta}) = \frac{\tilde{\lambda}(\mathbf{s}_i | \boldsymbol{\beta})}{\tilde{\Lambda}(\tilde{\mathbf{o}} | \boldsymbol{\beta})}. \quad (8)$$

Thus, given there are N activity centers in the region of interest $\tilde{\mathbf{o}}$, the log-probability density for activity centers $\mathbf{s}_1, \dots, \mathbf{s}_N$ is

$$\log \mathbb{P}(\mathbf{s}_1, \dots, \mathbf{s}_N | \boldsymbol{\beta}, N) = -N \log \tilde{\Lambda}(\tilde{\mathbf{o}} | \boldsymbol{\beta}) + \sum_{i=1}^N \log \tilde{\lambda}(\mathbf{s}_i | \boldsymbol{\beta}). \quad (9)$$

From Equations (7) and (8), it follows that the marginal void probability p^* that individual i is never detected is

$$\begin{aligned} p^* &= \mathbb{P}(D_{i,\mathbf{o}} = 0 | \boldsymbol{\theta}, \boldsymbol{\sigma}, \boldsymbol{\beta}) = \int_{\tilde{\mathbf{o}}} \mathbb{P}(\mathbf{s}_i | \boldsymbol{\beta}) \mathbb{P}(D_{i,\mathbf{o}} = 0 | \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}) d\mathbf{s}_i \\ &= \frac{1}{\tilde{\Lambda}(\tilde{\mathbf{o}} | \boldsymbol{\beta})} \int_{\tilde{\mathbf{o}}} \tilde{\lambda}(\mathbf{s}_i | \boldsymbol{\beta}) \exp\{-\Lambda(\mathbf{o} | \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma})\} d\mathbf{s}_i. \end{aligned} \quad (10)$$

It is immediate that the marginal void probability is the same for all individuals in the population. Calculating p^* seems complicated; however, for many commonly-used habitat selection and detection intensity functions, it is possible to express the formula for calculating p^* either in terms of elementary functions or non-elementary functions with computationally tractable numerical approximations. One example of the Gaussian kernel function is given in the Appendix.

Hierarchical point process model

Combining the models for habitat selection and detection described above results in a hierarchical point process SCR model. As shown in Figure [No.], the full model consists of four parts:

- The activity centers of individuals are modelled by a Poisson point process with intensity $\tilde{\lambda}(\mathbf{s} | \boldsymbol{\beta})$. Then the population size N follows a Poisson distribution:

$$N \sim \text{Poisson} \left\{ \tilde{\Lambda}(\tilde{\boldsymbol{\alpha}} | \boldsymbol{\beta}) \right\}, \quad (11)$$

where the expected population size $\tilde{\Lambda}(\tilde{\boldsymbol{\alpha}} | \boldsymbol{\beta})$ is given in Equation (2).

- Conditional on population size N , the number of individuals detected, N^* , follows a binomial distribution:

$$N^* \sim \text{Binomial}(N, 1 - p^*), \quad (12)$$

where p^* is the marginal void probability given in (10).

- Conditional on the number of detected individuals N^* , the activity centers of the detected individuals, $\mathbf{s}_1, \dots, \mathbf{s}_{N^*}$, follow a binomial point process, whose intensity function is given in (8).

- Conditional on the activity center $\mathbf{s}_i$ of individual $i = 1, \dots, N^*$, the detections $\mathbf{d}_{i,1}, \dots, \mathbf{d}_{i,M_i}$ follow a Poisson point process with intensity $\lambda(\mathbf{d} | \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma})$. The log-probability density for these detections is:

$$\log \mathbb{P}(\mathbf{d}_{i,1}, \dots, \mathbf{d}_{i,M_i} | \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma}) - \log(1 - p^*), \quad (13)$$

where $\log \mathbb{P}(\mathbf{d}_{i,1}, \dots, \mathbf{d}_{i,M_i} | \mathbf{s}_i, \boldsymbol{\theta}, \boldsymbol{\sigma})$ is given in Equation (4). Note that Equation (13) contains the term $-\log(1 - p^*)$ since here we condition on the fact that the individual is detected, which has probability $1 - p^*$.

Cell-based framework for covariates

SCR models often incorporate covariates to improve the accuracy of inferences. For example, available resources are usually considered when evaluating how animals distribute themselves in space (Gaillard *et al.*, 2010). Using our point process model, it is straightforward to use covariates for either the habitat selection process or the detection process. This can be achieved by formulating the intensity of the inhomogeneous Poisson process as a function of covariates.

While in some cases it is possible to derive the environmental covariates as spatial functions, it is more common for environmental data to be retrieved, processed, and distributed as cell-based discrete approximations to the continuous process. To handle such covariates for habitat selection, the entire region of interest $\tilde{\mathbf{o}}$ is divided into a set of H non-overlapping habitat sub-windows $\{\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_H\}$ such that $\tilde{\mathbf{w}}_i \cap \tilde{\mathbf{w}}_j = \emptyset$ if $i \neq j$ and $\tilde{\mathbf{w}}_1 \cup \dots \cup \tilde{\mathbf{w}}_H = \tilde{\mathbf{o}}$. Similarly, for cell-based detection covariates we divide the detection region $\mathbf{o}$ into a set of L non-overlapping detection sub-windows $\mathbf{w}_1, \dots, \mathbf{w}_L$ such that $\mathbf{w}_i \cap \mathbf{w}_j = \emptyset$ if $i \neq j$ and $\mathbf{w}_1 \cup \dots \cup \mathbf{w}_L = \mathbf{o}$. The value of each covariate is constant within each sub-window and is related to the intensity value of the sub-window through some link function. In principal, any non-negative function is valid for this purpose. Here, we consider the log-linear model as an example for both habitat selection and detection. Define $\tilde{\lambda}_h$ to be the habitat selection intensity value of sub-window $\tilde{\mathbf{w}}_h, h = 1, \dots, H$. It follows that

$$\log(\tilde{\lambda}_h) = \beta_0 + \sum_i \beta_i \xi_{ih}, \quad (14)$$

where ξ_{ih} is the value of the i -th habitat selection covariate in sub-window $\tilde{\mathbf{w}}_h$. Similarly, define b_l to be the baseline detection intensity of sub-window $\mathbf{w}_l, l = 1, \dots, L$. Then we have

$$\log(b_l) = \theta_0 + \sum_i \theta_i \zeta_{il}, \quad (15)$$

where ζ_{il} is the value of the i -th detection covariate within sub-window $\mathbf{w}_l$.

In situations where orthogonally neighbouring sub-windows share the same values for the

covariates, it may be possible to merge these sub-windows into one. This minimises the number of sub-windows and thus reduces the amount of computation needed for model fitting. The size of the habitat and detection sub-windows should be chosen according to computational considerations but also at the scale relevant to the ecological and observation processes (Johnson, 1980; Gaillard *et al.*, 2010).

Using the grid cells above, we can rewrite the intensity functions of the Poisson point processes for habitat selection and detection as two piecewise constant functions; see Appendix I and II for details. This reduces the calculation of the definite integrals in (2) and (5) to discrete summations. As a result, calculating the marginal void probability (10) is simplified to some extent; see Appendix III.

Software implementation

We fit the model via Bayesian MCMC in the R package nimble (de Valpine *et al.*, 2017). nimble supports nearly the same modelling language as JAGS and WinBUGS but allows extensions with new functions and distributions. We used these capabilities to write both Poisson and binomial point process distributions and also marginal void probability calculations in nimble, allowing them to be automatically incorporated into MCMC sampling.

Results

Simulation study

We designed a simulation study to investigate the relative bias and frequentist coverage of credible intervals for calibrated Bayes interpretation (Little, 2006) under a range of covariate effects. We set the habitat region $\tilde{\mathbf{o}}$ to be a 10km×10km square, which included a buffer of 0.6km around the detection region $\mathbf{o}$, a 8.8km×8.8km square. The region $\tilde{\mathbf{o}}$ was divided into 100 equally-sized sub-windows and the region $\mathbf{o}$ was divided into 25 equally-sized sub-windows. We considered two

Figure 1: Violin plots showing the distribution of relative estimation error (defined as (estimate - true value) / true value) of the population size N based on 100 replicates for each scenario. The value of (β_1, θ_1) in scenarios 1–9 is $(-1, -1)$, $(0, -1)$, $(1, -1)$, $(-1, 0)$, $(0, 0)$, $(1, 0)$, $(-1, 1)$, $(0, 1)$, and $(1, 1)$ respectively. The value below each violin plot gives the credible interval coverage.

spatial covariates in the simulations, one for modelling the habitat selection intensity and the other for the baseline detection intensity. Values of both covariates were simulated from a uniform distribution: $\text{Uniform}(-1, 1)$. The intercept parameters β_0 and θ_0 were set to be 1.0 and 2.0 respectively. The values of the slope parameters β_1 and θ_1 were chosen from set $\{-1, 0, 1\}$ so that there were nine scenarios defined by different combinations for the two parameters. An isotropic multivariate Gaussian function with $\sigma = 0.2$ was used for the detection kernel.

For each of the nine scenarios, we simulated 100 datasets under the hierarchical point process model. The model was then fit to the data using Bayesian MCMC in nimble. MCMC sampling was run for 110,000 iterations with the first 10,000 discarded as burn-in. Estimation results of

the population size N are shown in Fig. 1. It can be seen that almost unbiased inference with roughly a nominal value for the credible interval coverage was obtained for all the simulations. Similar conclusions can be drawn from the results of other parameters shown in the supplementary material. We therefore conclude that the proposed model works well for parameter estimation.

Wolverine data analysis

To demonstrate the application of the proposed model, we applied it to part of the wolverines data available in the Scandinavian large carnivore database, Rovbase 3.0 (<http://rovbase.se> or <http://rovbase.no/>). This database is jointly used by Norway and Sweden to record NGS data, dead recoveries, GPS search tracks, and observations of wolverines and other large carnivores. Here, we restricted our analysis to the data of female wolverines collected from December 2018 to June 2019 in Norwegian counties of Hedmark, Oppland, and parts of Sør-Trøndelag (Fig. 2; Flagstad et al. 2004, Brøseth et al. 2010, Gervasi et al. 2015). The dataset is composed of 228 scat- and hair-based DNA samples from 72 female wolverines.

As shown in Fig. 2, the detection region $\mathbf{o}$ was divided into $L = 195$ sub-windows each of size $20\text{km} \times 20\text{km}$ and the habitat selection region $\tilde{\mathbf{o}}$ was divided into $H = 40$ sub-windows each of size $60\text{km} \times 60\text{km}$. The region $\tilde{\mathbf{o}}$ covers the entire region $\mathbf{o}$ and a surrounding buffer, allowing the activity centers of individuals to be located outside the searched area. The number of known wolverine dens was used as a covariate for modelling habitat selection. Five covariates $\zeta_i, i = 1, \dots, 5$ were used for modelling baseline detection. Indicator variables ζ_1 and ζ_2 take value 1 if the detection location is in Oppland and Sør-Trøndelag, respectively, and take value 0 otherwise. In addition, covariates ζ_3, ζ_4 , and ζ_5 denote respectively the recorded length of GPS tracks logged by searchers, the average percentage of snow cover (MODIS at 0.1 degrees resolution, www.neo.sci.gsfc.nasa.gov, accessed 2019-10-11), and the average distance to the nearest primary and secondary roads. The isotropic multivariate Gaussian kernel function was used for detection

To fit the data, we ran eight MCMC chains for 20,000 iterations and removed the first 5,000

Figure 2: Distribution of non-invasive DNA samples (red crosses) of female wolverines collected during winter 2018/19 in Norway. The area that was searched for samples is represented by the red polygon on the map in the upper left corner. Detection sub-windows (20km×20km each) are represented by the dark green grid cells. Habitat selection sub-windows (60km×60km each) are represented by the larger light green grid cells.

Table 1: Estimation results for the hierarchical point process model fitted to the female wolverine data collected in eastern Norway during winter 2018/2019. ESS represents the effective sample size.

Parameters	Mean	SD	2.5%	97.5%	R-hat	ESS
N	137.19	12.25	115	163	< 1.001	5,761
β_0	-21.04	0.15	-21.35	-20.76	1.002	2,626
β_1	0.64	0.11	0.42	0.86	< 1.001	5,008
θ_0	-18.38	0.83	-20.05	-16.79	< 1.001	13,771
θ_1	-18.28	0.21	-18.71	-17.88	< 1.001	12,538
θ_2	-17.99	0.14	-18.27	-17.71	1.002	2,980
θ_3	0.25	0.09	0.08	0.43	< 1.001	5,120
θ_4	-0.13	0.09	-0.31	0.05	< 1.001	11,540
θ_5	0.14	0.10	-0.06	0.34	< 1.001	12,452
σ	5.16	0.21	4.76	5.59	1.001	5,121

iterations as burn-in. The model was assumed to have reached convergence when all R-hat values were ≤ 1.1 and visual inspections of trace plots revealed good mixing (Brooks & Gelman, 1998). Mean parameter estimates, standard deviations, 95% credible intervals, R-hat values, and effective sample sizes are shown in Table 1. It is estimated that on average there are 9.53 (95% CI: 8–11) female wolverines per 10,000 km² within the sampled area.

Discussion

We developed a hierarchical point process model for SCR data collected by exhaustive area searches. The development of the model was motivated by the NGS data of large carnivores in Scandinavia, where individuals can, at least in theory, be detected anywhere within the sampled area. Fitting such data via the artificial detector grid approach leads to spatial information loss and computational efficiency issues (Milleret *et al.*, 2018). The search-encounter model of Royle *et al.* (2011a) is not suitable in this context because auxiliary information (i.e. sampling paths) on sampling effort is not completely available.

Our model is a reasonable reflection of the NGS data collection process and all spatial in-

formation in the data can be used without additional approximation error. Spatial covariates can be incorporated to model the intensity functions of the inhomogeneous Poisson point processes for both habitat selection and detection. In addition, we formulated the hierarchical model using the void probability approach, which makes it possible to fit the model via MCMC in a Bayesian paradigm without needing to use data augmentation. This improves the efficiency of model fitting and eliminates the implementation challenges associated with data augmentation. We demonstrated the promising performance of the proposed model for parameter estimation with a set of simulations and the wolverine data analysis.

The approach presented here can be extended to work for several more advanced SCR models. First, we could extend the modelling framework to open populations by adding an extra layer to the current model to account for the changes in population size and structure over time. It follows that the population size N will be conditional on the population model used rather than the intensity function for habitat selection. The distribution of activity centers then becomes a binomial point process instead of Poisson. In this case, $\tilde{\lambda}(\mathbf{s} | \boldsymbol{\beta})$ becomes a function describing the relative density of activity centers rather than the absolute density. Second, SCR data with group identities (e.g., Chandler & Royle, 2013; Augustine *et al.*, 2018) can be handled easily using our point process detection model. Suppose that there are two sets of spatial points, each of which is generated from a Poisson point process. Then points in the union of the two sets follow another Poisson point process, whose intensity is the sum of the intensities of the two separate Poisson processes (Çinlar & Agnew, 1968). This means that detections from a group of non-uniquely-identifiable individuals can be modelled as a single Poisson point process. It is immediate that the framework presented in this manuscript can be applied directly to SCR data with group identities.

Acknowledgements

The study was funded by the Norwegian Environment Agency (Miljødirektoratet), the Swedish Environmental Protection Agency (Naturvårdsverket), the Research Council of Norway (NFR

286886; project WildMap), and the Peder Sather Grant. We thank Henrik Brøseth for comments on the manuscript.

Authors' contributions

JDC, PD, CM, and RB conceived the ideas with input from JBI and PdV. WZ, JDC, and JBI formulated the mathematical framework of the model. WZ, JDC, and PdV implemented the model in NIMBLE. PD, CM, and RB performed the analysis of the wolverines data. WZ and JDC led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data availability

The NGS wolverine data and R code used in this manuscript will be available on GitHub (<https://github.com/wzhang721>) after the manuscript is accepted.

References

- Augustine, B.C., Royle, J.A., Kelly, M.J., Satter, C.B., Alonso, R.S., Boydston, E.E. & Crooks, K.R. (2018) Spatial capture-recapture with partial identity: an application to camera traps. *The Annals of Applied Statistics* **12**, 67–95.
- Baddeley, A., Gregori, P., Mateu, J., Stoica, R. & Stoyan, D. (2006) *Case Studies in Spatial Point Process Modeling*. Springer.
- Bischof, R., Milleret, C., Dupont, P., Chipperfield, J., Brøseth, H. & Kindberg, J. (2019) Rovquant: estimating density, abundance and population dynamics of bears, wolverines and wolves in Scandinavia.
- Borchers, D., Distiller, G., Foster, R., Harmsen, B. & Milazzo, L. (2014) Continuous-time spatially

explicit capture-recapture models, with an application to a jaguar camera-trap survey. *Methods in Ecology and Evolution* **5**, 656–665.

Borchers, D. & Fewster, R. (2016) Spatial capture–recapture models. *Statistical Science* **31**, 219–232.

Borchers, D.L. & Marques, T.A. (2017) From distance sampling to spatial capture–recapture. *AStA Advances in Statistical Analysis* **101**, 475–494.

Brooks, S. & Gelman, A. (1998) General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* **7**, 434–455.

Burgar, J.M., Stewart, F.E., Volpe, J.P., Fisher, J.T. & Burton, A.C. (2018) Estimating density for species conservation: comparing camera trap spatial count models to genetic spatial capture–recapture models. *Global Ecology and Conservation* **15**, e00411.

Chandler, R.B. & Royle, J.A. (2013) Spatially explicit models for inference about density in unmarked or partially marked populations. *The Annals of Applied Statistics* **7**, 936–954.

Chiu, S.N., Stoyan, D., Kendall, W.S. & Mecke, J. (2013) *Stochastic Geometry and its Applications*. John Wiley & Sons, Ltd.

Çinlar, E. & Agnew, R.A. (1968) On the superposition of point processes. *Journal of the Royal Statistical Society, Series B* **30**, 576–581.

de Valpine, P., Turek, D., Paciorek, C.J., Anderson-Bergman, C., Lang, D.T. & Bodik, R. (2017) Programming with models: Writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics* **26**, 403–413.

Dorazio, R.M. (2013) Bayes and empirical bayes estimators of abundance and density from spatial capture-recapture data. *PLoS ONE* **8**, e84017.

416 Efford, M. (2004) Density estimation in live-trapping studies. *Oikos* **106**, 598–610.

417 Efford, M. (2020) *secr: Spatially explicit capture-recapture models*. URL [https://CRAN.](https://CRAN.R-project.org/package=secr)

418 [R-project.org/package=secr](https://CRAN.R-project.org/package=secr), r package version 4.2.2.

419 Efford, M.G., Borchers, D.L. & Mowat, G. (2013) Varying effort in capture-recapture studies.

420 *Methods in Ecology and Evolution* **4**, 629–636.

421 Ferreira, C.M., Sabino-Marques, H., Barbosa, S., Costa, P., Encarnação, C., Alpizar-Jara, R.,

422 Pita, R., Beja, P., Mira, A., Searle, J.B., Paupério, J. & Alves, P.C. (2018) Genetic non-invasive

423 sampling (gNIS) as a cost-effective tool for monitoring elusive small mammals. *European Journal*

424 *of Wildlife Research* **64**, e46.

425 Gaillard, J.M., Hebblewhite, M., Loison, A., Fuller, M., Powell, R., Basille, M. & Van Moorter,

426 B. (2010) Habitat–performance relationships: finding the right metric at a given spatial scale.

427 *Philosophical Transactions of the Royal Society B: Biological Sciences* **365**, 2255–2265.

428 Glover-Kapfer, P., Soto-Navarro, C.A. & Wearn, O.R. (2019) Camera-trapping version 3.0: current

429 constraints and future priorities for development. *Remote Sensing in Ecology and Conservation*

430 **5**, 205–223.

431 Henry, P., Henry, A. & Russello, M.A. (2011) A noninvasive hair sampling technique to obtain

432 high quality DNA from elusive small mammals. *Journal of Visualized Experiments* **49**, e2791.

433 Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. (2007) *Statistical Analysis and Modelling of*

434 *Spatial Point Patterns*. John Wiley & Sons, Ltd.

435 Johnson, D.H. (1980) The comparison of usage and availability measurements for evaluating re-

436 source preference. *Ecology* **61**, 65–71.

437 King, R., McClintock, B.T., Kidney, D. & Borchers, D. (2016) Capture–recapture abundance

estimation using a semi-complete data likelihood approach. *The Annals of Applied Statistics* **10**, 264–285.

Lamb, C.T., Ford, A.T., Proctor, M.F., Royle, J.A., Mowat, G. & Boutin, S. (2019) Genetic tagging in the Anthropocene: scaling ecology from alleles to ecosystems. *Ecological Applications* **29**, e01876.

Little, R.J. (2006) Calibrated Bayes: a Bayes/frequentist roadmap. *The American Statistician* **60**, 213–223.

Lonsinger, R., Lukacs, P., Gese, E., Knight, R. & Waits, L. (2018) Estimating densities for sympatric kit foxes (*vulpes macrotis*) and coyotes (*canis latrans*) using noninvasive genetic sampling. *Canadian Journal of Zoology* **96**, 1080–1089.

López-Bao, J.V., Godinho, R., Pacheco, C., Lema, F.J., García, E., Llana, L., Palacios, V. & Jiménez, J. (2018) Towards reliable population estimates of wolves by combining spatial capture-recapture models and non-invasive DNA monitoring. *Scientific Reports* **8**, e2177.

Milleret, C., Dupont, P., Brøseth, H., Kindberg, J., Royle, J.A. & Bischof, R. (2018) Using partial aggregation in spatial capture recapture. *Methods in Ecology and Evolution* **9**, 1896–1907.

Powell, R.A. & Proulx, G. (2003) Trapping and marking terrestrial mammals for research: Integrating ethics, performance criteria, techniques, and common sense. *ILAR Journal* **44**, 259–276.

Royle, J.A., Chandler, R.B., Sollman, R. & Gardner, B. (2014) *Spatial Capture-Recapture*. Academic Press, Waltham, United States.

Royle, J.A., Kéry, M., Gautier, R. & Schmid, H. (2007) Hierarchical spatial models of abundance and occurrence from imperfect survey data. *Ecological Monographs* **77**, 465–481.

Royle, J.A., Kéry, M. & Guélat, J. (2011a) Spatial capture-recapture models for search-encounter data. *Methods in Ecology and Evolution* **2**, 602–611.

- 461 Royle, J.A., Magoun, A.J., Gardner, B., Valkenburg, P. & Lowell, R.E. (2011b) Density esti-
462 mation in a wolverine population using spatial capture-recapture models. *Journal of Wildlife*
463 *Management* **75**, 604–611.
- 464 Royle, J.A. & Dorazio, R.M. (2012) Parameter-expanded data augmentation for Bayesian analysis
465 of capture–recapture models. *Journal of Ornithology* **152**, 521–537.
- 466 Royle, J.A. & Young, K.V. (2008) A hierarchical model for spatial capture–recapture data. *Ecology*
467 **89**, 2281–2289.
- 468 Russell, R.E., Royle, J.A., Desimone, R., Schwartz, M.K., Edwards, V.L., Pilgrim, K.P. & Mck-
469 elvey, K.S. (2012) Estimating abundance of mountain lions from unstructured spatial sampling.
470 *The Journal of Wildlife Management* **76**, 1551–1561.
- 471 Thomas, L. & Marques, T.A. (2012) Passive acoustic monitoring for estimating animal density.
472 *Acoustics Today* **8**, 35–44.
- 473 Trollet, F., Huynen, M.C., Vermeulen, C. & Hambuckers, A. (2014) Use of camera traps for
474 wildlife studies: a review. *Biotechnology, Agronomy, Society and Environment* **18**, 446–454.

Appendix I: cell-based habitat selection

Note that $\tilde{\lambda}_h$ is the habitat selection intensity value of sub-window $\tilde{\mathbf{w}}_h, h = 1, \dots, H$. It follows that the habitat selection intensity function $\tilde{\lambda}(\mathbf{s} | \boldsymbol{\beta})$ can be written as

$$\tilde{\lambda}(\mathbf{s} | \boldsymbol{\beta}) = \sum_{h=1}^H \tilde{\lambda}_h \mathcal{I}(\mathbf{s} \in \tilde{\mathbf{w}}_h), \quad (16)$$

where $\mathcal{I}(\cdot)$ is the usual indicator function that equals 1 if the condition $\cdot$ is true and 0 otherwise.

Provided that an inhomogenous Poisson point process is used for habitat selection, the number of activity centers $N_{\tilde{\mathbf{w}}}$ within a subregion $\tilde{\mathbf{w}} \subset \tilde{\mathbf{o}}$ follows a Poisson distribution:

$$N_{\tilde{\mathbf{w}}} \sim \text{Poisson} \left\{ \tilde{\Lambda}(\tilde{\mathbf{w}} | \boldsymbol{\beta}) \right\}, \quad (17)$$

where

$$\tilde{\Lambda}(\tilde{\mathbf{w}} | \boldsymbol{\beta}) = \int_{\tilde{\mathbf{w}}} \tilde{\lambda}(\mathbf{s} | \boldsymbol{\beta}) d\mathbf{s} \quad (18)$$

is the expected number of activity centers located within $\tilde{\mathbf{w}}$. According to Equation (18), the expected number of activity centers within sub-window $\tilde{\mathbf{w}}_h$ is

$$\tilde{\Lambda}(\tilde{\mathbf{w}}_h | \boldsymbol{\beta}) = \tilde{\lambda}_h V_h, \quad (19)$$

where V_h is the volume of sub-window $\tilde{\mathbf{w}}_h$. It follows that the expected total number of activity centers across the entire region $\tilde{\mathbf{o}}$ is

$$\tilde{\Lambda}(\tilde{\mathbf{o}} | \boldsymbol{\beta}) = \sum_{h=1}^H \tilde{\Lambda}(\tilde{\mathbf{w}}_h | \boldsymbol{\beta}) = \sum_{h=1}^H \tilde{\lambda}_h V_h. \quad (20)$$

Appendix II: cell-based detection

Note that b_l denotes the baseline detection intensity of sub-window $\mathbf{w}_l, l = 1, \dots, L$. Then, similarly to the habitat selection case above, the baseline detection intensity function can be written as

$$b(\mathbf{d} | \boldsymbol{\theta}) = \sum_{l=1}^L b_l \mathcal{I}(\mathbf{d} \in \mathbf{w}_l). \quad (21)$$

A common detection decay kernel is the isotropic multivariate Gaussian kernel

$$\begin{aligned} \tau(\mathbf{d} | \mathbf{s}_i, \sigma) &= \exp \left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^D (d_j - s_{ij})^2 \right\} \\ &= (2\pi)^{\frac{D}{2}} \prod_{j=1}^D \phi \left(\frac{d_j - s_{ij}}{\sigma} \right), \end{aligned} \quad (22)$$

where $\phi(x)$ is the standard normal probability density function

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{x^2}{2} \right). \quad (23)$$

Here, $\tau(\mathbf{d} | \mathbf{s}_i, \sigma)$ is proportional to the probability density function of a multivariate Gaussian distribution with mean equal to the activity centre $\mathbf{s}_i$ of individual i , and with a diagonal variance-covariance matrix whose diagonal elements are all set to σ^2 , representing the isotropic variance of the detection decay kernel. Thus, the parameter σ regulates the home range size of the species of interest. In this paper, we do not consider covariates for σ . However, it is immediate that σ can be modelled as a regression of relevant environmental and/or individual covariates allowing for different home range sizes in different environments or for species that have different home ranges at different life stages or between different sexes.

Inserting the cell-based baseline detection function (21) and the Gaussian decay kernel (22)

into Equation (3) yields the following composite detection intensity function

$$\lambda(\mathbf{d} \mid \mathbf{s}_i, \boldsymbol{\theta}, \sigma) = (2\pi)^{\frac{D}{2}} \left\{ \sum_{l=1}^L b_l \mathcal{I}(\mathbf{d} \in \mathbf{w}_l) \right\} \prod_{j=1}^D \phi\left(\frac{d_j - s_{ij}}{\sigma}\right). \quad (24)$$

Resulting marginal void probability

Similar to the point process for activity centers, the number of detections $D_{i,\mathbf{w}}$ of individual i within a subregion $\mathbf{w} \subset \mathbf{o}$ follows a Poisson distribution:

$$D_{i,\mathbf{w}} \sim \text{Poisson} \{ \Lambda(\mathbf{w} \mid \mathbf{s}_i, \boldsymbol{\theta}, \sigma) \}, \quad (25)$$

where

$$\Lambda(\mathbf{w} \mid \mathbf{s}_i, \boldsymbol{\theta}, \sigma) = \int_{\mathbf{w}} \lambda(\mathbf{x} \mid \mathbf{s}_i, \boldsymbol{\theta}, \sigma) d\mathbf{x} \quad (26)$$

is the expected number of detections within $\mathbf{w}$. Then the expected number of detections $\Lambda(\mathbf{w}_l \mid \mathbf{s}_i, \boldsymbol{\theta}, \sigma)$ of individual i within sub-window $\mathbf{w}_l$ is

$$\begin{aligned} \Lambda(\mathbf{w}_l \mid \mathbf{s}_i, \boldsymbol{\theta}, \sigma) &= (2\pi)^{\frac{D}{2}} b_l \prod_{j=1}^D \int_{w_{lj1}}^{w_{lj2}} \phi\left(\frac{x_j - s_{ij}}{\sigma}\right) dx_j \\ &= (2\pi)^{\frac{D}{2}} \sigma^D b_l \prod_{j=1}^D \left[\Phi\left(\frac{w_{lj2} - s_{ij}}{\sigma}\right) - \Phi\left(\frac{w_{lj1} - s_{ij}}{\sigma}\right) \right], \end{aligned} \quad (27)$$

where w_{lj1} and w_{lj2} are the lower and upper bounds of the j -th dimension of sub-window $\mathbf{w}_l$, and $\Phi(x)$ is the standard Gaussian cumulative distribution function. Then the expected number of detections of individual i across the entire region $\mathbf{o}$ is

$$\Lambda(\mathbf{o} \mid \mathbf{s}_i, \boldsymbol{\theta}, \sigma) = (2\pi)^{\frac{D}{2}} \sigma^D \sum_{l=1}^L b_l \prod_{j=1}^D \left[\Phi\left(\frac{w_{lj2} - s_{ij}}{\sigma}\right) - \Phi\left(\frac{w_{lj1} - s_{ij}}{\sigma}\right) \right]. \quad (28)$$

510 According to Equation (10), the marginal void probability in the case of cell-based habitat selection
511 and detection is

$$\begin{aligned}\mathbb{P}(D_{i,\mathbf{o}} = 0 \mid \boldsymbol{\beta}, \boldsymbol{\theta}, \sigma, i \leq N) &= \frac{1}{\tilde{\Lambda}(\tilde{\mathbf{o}} \mid \boldsymbol{\beta})} \int_{\tilde{\mathbf{o}}} \tilde{\lambda}(\mathbf{s}_i \mid \boldsymbol{\beta}) \exp\{-\Lambda(\mathbf{o} \mid \mathbf{s}_i, \boldsymbol{\theta}, \sigma)\} d\mathbf{s}_i \\ &= \frac{1}{\tilde{\Lambda}(\tilde{\mathbf{o}} \mid \boldsymbol{\beta})} \sum_{h=1}^H \tilde{\lambda}_h \int_{\tilde{\mathbf{w}}_h} \exp\{-\Lambda(\mathbf{o} \mid \mathbf{s}_i, \boldsymbol{\theta}, \sigma)\} d\mathbf{s}_i,\end{aligned}\tag{29}$$

512 where $\tilde{\Lambda}(\tilde{\mathbf{o}} \mid \boldsymbol{\beta})$ is given in Equation (20), and the definite integral $\int_{\tilde{\mathbf{w}}_h} \exp\{-\Lambda(\mathbf{o} \mid \mathbf{s}_i, \boldsymbol{\theta}, \sigma)\} d\mathbf{s}_i$
513 can be calculated using the grid integration of Equation (28).