A hierarchical point process model for spatial capture-recapture data

Wei Zhang\(^1,\ast\), Joseph D. Chipperfield\(^2,3,\ast\), Janine B. Illian\(^4\), Pierre Dupont\(^2\), Cyril Milleret\(^2\), Perry de Valpine\(^1\), Richard Bischof\(^2\)

\(^1\)Department of Environmental Science, Policy and Management, University of California Berkeley, Berkeley, USA

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

\(^3\)Norwegian Institute for Nature Research, Hyteknologisenteret, Thormøhlensgate 55, Bergen, Norway

\(^4\)School 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 10,000km$^2$ 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
& Young (2008) proposed a SCR model to investigate temporary emigration of individuals. In their case, data are collected by an exhaustive area-search for individuals within a delineated sample unit, thus individuals can be detected anywhere in the unit. Similarly, Royle et al. (2011a) developed a model for ‘search-encounter’ SCR data that are collected by surveyors searching along transects.

Despite their variety (Royle et al., 2014; Efford, 2020), existing detector-based or search-encounter models do not adequately cover all common SCR sampling processes. Non-invasive genetic sampling (NGS) is now commonly used for collecting SCR data (Burgar et al., 2018; Ferreira et al., 2018; López-Bao et al., 2018). NGS in SCR studies involves the collection of DNA sources such as shed hair, faeces, saliva, and urine at fixed detector or area searches, followed by individual identification (Lamb et al., 2019). When NGS is implemented by searching a given area or along transects, paths taken by searchers can be recorded and used as a direct measure of effort in space and time. In praxis, this is not always possible, practical, or reliable. The model described in this manuscript is motivated by our experiences with the Scandinavian large carnivore monitoring program (Bischof et al., 2019), one of the world’s most extensive NGS projects (as of December 2019: over 35,000 samples from 6,000 individual bears (*Ursus arctos*), wolves (*Canis lupus*), and wolverines). Although field protocols prescribe the recording of GPS tracks during the performance of NGS by Norwegian and Swedish government employees, due to technical limitations and logistic constraints, these cannot always be linked unambiguously with detections. Furthermore, the region monitored is vast (over 500,000 km$^2$) and a substantial portion of genetic samples are obtained via opportunistic detections by members of the public (hunters, hikers, berry pickers, etc.). As a consequence, no direct measure of effort is available and investigators wanting to account for spatial heterogeneity in detection probability are left with proxies for effort such as distance to roads and snow cover (Bischof et al., 2019). In cases such as this, detections, at least in theory, are possible at any location within the general area that humans could visit. An approach that uses 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 $x$ gives the probability that there is one point (one detection or activity center in
our case) falling within an infinitesimal volume centered around $x$ (p.51, Chiu et al., 2013). The
detection intensity for one individual at one point in the study area is a (typically decreasing) func-
tion 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 ap-
proximation 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 unob-
served. 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 \( i = 1, \ldots, N \).
- \( D_{i,o} \): the number of detections of individual \( i \) across the entire region \( o \).
- \( \tilde{o} \): the entire region for the placement of activity centers (habitat selection).
- \( o \): the entire region where individuals can be detected; \( o \) may or may not be the same as \( \tilde{o} \).
• \( \tilde{w}_h \): a sub-window of \( \tilde{o} \); \( h = 1, \ldots, H \).

• \( w_l \): a sub-window of \( o \); \( l = 1, \ldots, L \).

• \( s_i \): location of the activity center of individual \( i \); \( s_i \in \tilde{o}; i = 1, \ldots, N \).

• \( d_{i,j} \): location of the \( j \)-th detection for individual \( i \); \( d_{i,j} \in o; i = 1, \ldots, N, j = 1, \ldots, M_i \).

• \( \beta \): a vector of parameters that determine the intensity function of the point process model for habitat selection.

• \( \theta \): a vector of parameters that determine the baseline detection function.

• \( \sigma \): a vector of parameters that determine the (general) detection kernel function.

• \( \sigma \): a scalar parameter that determines the (specific) isotropic multivariate Gaussian detection kernel function, which is given in Appendix II.

• \( \tilde{\lambda}(s \mid \beta) \): the intensity function of the point process model for habitat selection.

• \( \lambda(d \mid s, \theta, \sigma) \): the intensity function of the point process model for detection.

• \( \tilde{\Lambda} (\tilde{o} \mid \beta) \): the expected number of activity centers within the region \( \tilde{o} \).

• \( \Lambda (o \mid s_i, \theta, \sigma) \): the expected number of detections of individual \( i \) within the region \( o \).

• \( \tilde{\Lambda} (\tilde{w} \mid \beta) \): the expected number of activity centers within sub-window \( \tilde{w} \).

• \( \Lambda (w \mid s_i, \theta, \sigma) \): the expected number of detections of individual \( i \) within sub-window \( w \).

For convenience, we use \( \int_{\Omega} f(x) dx \) to denote the multi-dimensional integral of \( f(x) \) over the region \( \Omega \), which can be \( o, \tilde{o}, w \) or \( \tilde{w} \).
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 $\tilde{\lambda}(s | \beta)$ is a constant across $\tilde{o}$, 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, $s_1, \ldots, s_N$, over the entire region $\tilde{o}$ is

$$
\log P(s_1, \ldots, s_N | \beta) = -\tilde{\Lambda}(\tilde{o} | \beta) + \sum_{i=1}^{N} \log \tilde{\lambda}(s_i | \beta),
$$

(1)

where

$$
\tilde{\Lambda}(\tilde{o} | \beta) = \int_{\tilde{o}} \tilde{\lambda}(s | \beta)ds
$$

(2)

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

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

\[ \lambda(d \mid s, \theta, \sigma) = b(d \mid \theta) \tau(d \mid s, \sigma), \]  

where \( d \) denotes the location of the detection, \( s \) denotes the activity center, \( \tau(d \mid s, \sigma) \) is a kernel describing the distance decay relationship of the detection intensity from the activity center, and \( b(d \mid \theta) \) is a function describing the baseline detection intensity at the detection location. The function \( \tau(d \mid s, \sigma) \) 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 \( b(d \mid \theta) \) 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, \( d_{i,1} \ldots d_{i,M_i} \), of individual \( i \) given
that its activity center is $s_i$ is:

$$\log \mathbb{P}(d_{i,1}, \ldots, d_{i,M_i} \mid s_i, \theta, \sigma) = - \Lambda(o \mid s_i, \theta, \sigma) + \sum_{j=1}^{M_i} \log \lambda(d_{i,j} \mid s_i, \theta, \sigma),$$  

where

$$\Lambda(o \mid s_i, \theta, \sigma) = \int_o \lambda(x \mid s_i, \theta, \sigma) \, dx$$

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:

$$\log \mathbb{P}(d_{1,1}, \ldots, d_{1,M_1}, \ldots, d_{N,1}, \ldots, d_{N,M_N} \mid s_1, \ldots, s_N, \theta, \sigma) = \sum_{i=1}^{N} \log \mathbb{P}(d_{i,1}, \ldots, d_{i,M_i} \mid s_i, \theta, \sigma)$$

$$= - \sum_{i=1}^{N} \Lambda(o \mid s_i, \theta, \sigma) + \sum_{i=1}^{N} \sum_{j=1}^{M_i} \log \lambda(d_{i,j} \mid s_i, \theta, \sigma).$$

**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 $\mathbb{P}(D_{i,o} = 0 \mid s_i, \theta, \sigma)$ for individual $i$ is

$$\mathbb{P}(D_{i,o} = 0 \mid s_i, \theta, \sigma) = \exp \{- \Lambda(o \mid s_i, \theta, \sigma)\}.$$
The void probability depends on the unknown activity center \( s_i \) 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 \textit{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 \( s_i \) is

\[
\mathbb{P}(s_i | \beta) = \frac{\tilde{\lambda}(s_i | \beta)}{\Lambda(\tilde{o} | \beta)}. \tag{8}
\]

Thus, given there are \( N \) activity centers in the region of interest \( \tilde{o} \), the log-probability density for activity centers \( s_1, \ldots, s_N \) is

\[
\log \mathbb{P}(s_1, \ldots, s_N | \beta, N) = -N \log \Lambda(\tilde{o} | \beta) + \sum_{i=1}^{N} \log \tilde{\lambda}(s_i | \beta). \tag{9}
\]

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

\[
p^* = \mathbb{P}(D_{i,o} = 0 | \theta, \sigma, \beta) = \int_{\tilde{o}} \mathbb{P}(s_i | \beta) \mathbb{P}(D_{i,o} = 0 | s_i, \theta, \sigma) ds_i
\]

\[
= \frac{1}{\Lambda(\tilde{o} | \beta)} \int_{\tilde{o}} \tilde{\lambda}(s_i | \beta) \exp \{-\Lambda(\tilde{o} | s_i, \theta, \sigma)\} ds_i. \tag{10}
\]

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 $\tilde{\lambda}(s | \beta)$. Then the population size $N$ follows a Poisson distribution:

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

where the expected population size $\tilde{\Lambda}(\tilde{o} | \beta)$ 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, $s_1, \ldots, s_{N^*}$, follow a binomial point process, whose intensity function is given in (8).

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

$$\log \mathbb{P}(d_{i,1}, \ldots, d_{i,M_i} | s_i, \theta, \sigma) - \log(1 - p^*), \quad (13)$$

where $\log \mathbb{P}(d_{i,1}, \ldots, d_{i,M_i} | s_i, \theta, \sigma)$ 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 $1 - p^*$. 

13
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 $\tilde{o}$ is divided into a set of $H$ non-overlapping habitat sub-windows $\{\tilde{w}_1, \ldots, \tilde{w}_H\}$ such that $\tilde{w}_i \cap \tilde{w}_j = \emptyset$ if $i \neq j$ and $\tilde{w}_1 \cup \cdots \cup \tilde{w}_H = \tilde{o}$. Similarly, for cell-based detection covariates we divide the detection region $o$ into a set of of $L$ non-overlapping detection sub-windows $w_1, \ldots, w_L$ such that $w_i \cap w_j = \emptyset$ if $i \neq j$ and $w_1 \cup \cdots \cup w_L = o$. 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 $\tilde{w}_h, h = 1, \ldots, H$.

It follows that

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

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

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

where $\zeta_{il}$ is the value of the $i$-th detection covariate within sub-window $w_l$.

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 extend; 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 \( \tilde{o} \) to be a 10km\( \times \)10km square, which included a buffer of 0.6km around the detection region \( o \), a 8.8km\( \times \)8.8km square. The region \( \tilde{o} \) was divided into 100 equally-sized sub-windows and the region \( o \) was divided into 25 equally-sized sub-windows. We considered two
Figure 1: Violin plots showing the distribution of relative estimation error (defined as \((\text{estimate} - \text{true value}) / \text{true value}\)) of the population size \(N\) based on 100 replicates for each scenario. The value of \((\beta_1, \theta_1)\) 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.

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 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 \(\beta_0\) and \(\theta_0\) were set to be 1.0 and 2.0 respectively. The values of the slope parameters \(\beta_1\) and \(\theta_1\) 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.
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/orhttp://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 $\mathbf{a}$ was divided into $L = 195$ sub-windows each of size 20km $\times$ 20km and the habitat selection region $\mathbf{\tilde{a}}$ was divided into $H = 40$ sub-windows each of size 60km $\times$ 60km. The region $\mathbf{\tilde{a}}$ covers the entire region $\mathbf{a}$ 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 $\zeta_i, i = 1, \ldots, 5$ were used for modelling baseline detection. Indicator variables $\zeta_1$ and $\zeta_2$ take value 1 if the detection location is in Oppland and Sør-Trøndelag, respectively, and take value 0 otherwise. In addition, covariates $\zeta_3, \zeta_4,$ and $\zeta_5$ 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.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Mean</th>
<th>SD</th>
<th>2.5%</th>
<th>97.5%</th>
<th>R-hat</th>
<th>ESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>137.19</td>
<td>12.25</td>
<td>115</td>
<td>163</td>
<td>&lt; 1.001</td>
<td>5,761</td>
</tr>
<tr>
<td>$\beta_0$</td>
<td>-21.04</td>
<td>0.15</td>
<td>-21.35</td>
<td>-20.76</td>
<td>1.002</td>
<td>2,626</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.64</td>
<td>0.11</td>
<td>0.42</td>
<td>0.86</td>
<td>&lt; 1.001</td>
<td>5,008</td>
</tr>
<tr>
<td>$\theta_0$</td>
<td>-18.38</td>
<td>0.83</td>
<td>-20.05</td>
<td>-16.79</td>
<td>&lt; 1.001</td>
<td>13,771</td>
</tr>
<tr>
<td>$\theta_1$</td>
<td>-18.28</td>
<td>0.21</td>
<td>-18.71</td>
<td>-17.88</td>
<td>&lt; 1.001</td>
<td>12,538</td>
</tr>
<tr>
<td>$\theta_2$</td>
<td>-17.99</td>
<td>0.14</td>
<td>-18.27</td>
<td>-17.71</td>
<td>1.002</td>
<td>2,980</td>
</tr>
<tr>
<td>$\theta_3$</td>
<td>0.25</td>
<td>0.09</td>
<td>0.08</td>
<td>0.43</td>
<td>&lt; 1.001</td>
<td>5,120</td>
</tr>
<tr>
<td>$\theta_4$</td>
<td>-0.13</td>
<td>0.09</td>
<td>-0.31</td>
<td>0.05</td>
<td>&lt; 1.001</td>
<td>11,540</td>
</tr>
<tr>
<td>$\theta_5$</td>
<td>0.14</td>
<td>0.10</td>
<td>-0.06</td>
<td>0.34</td>
<td>&lt; 1.001</td>
<td>12,452</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>5.16</td>
<td>0.21</td>
<td>4.76</td>
<td>5.59</td>
<td>1.001</td>
<td>5,121</td>
</tr>
</tbody>
</table>

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 modeling 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, $\lambda(s | \beta)$ 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 (Cinlar & Agnew, 1968). This means that detections from a group of non-uniquely-identifiable individuals can be modeled 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


Appendix I: cell-based habitat selection

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

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

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 $N_{\tilde{w}}$ within a subregion $\tilde{w} \subset \tilde{o}$ follows a Poisson distribution:

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

where

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

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

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

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

$$\tilde{\Lambda}(\tilde{o} | \beta) = \sum_{h=1}^{H} \tilde{\Lambda}(\tilde{w}_h | \beta) = \sum_{h=1}^{H} \tilde{\lambda}_h V_h. \quad (20)$$
Appendix II: cell-based detection

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

$$ b(d | \theta) = \sum_{l=1}^{L} b_l \mathbb{I}(d \in w_l). \quad (21) $$

A common detection decay kernel is the isotropic multivariate Gaussian kernel

$$ \tau(d | s_i, \sigma) = \exp \left\{ -\frac{1}{2\sigma^2} \sum_{j=1}^{D} (d_j - s_{ij})^2 \right\} $$

$$ = (2\pi)^\frac{D}{2} \prod_{j=1}^{D} \phi \left( \frac{d_j - s_{ij}}{\sigma} \right), \quad (22) $$

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, $\tau(d | s_i, \sigma)$ is proportional to the probability density function of a multivariate Gaussian distribution with mean equal to the activity centre $s_i$ of individual $i$, and with a diagonal variance-covariance matrix whose diagonal elements are all set to $\sigma^2$, representing the isotropic variance of the detection decay kernel. Thus, the parameter $\sigma$ regulates the home range size of the species of interest. In this paper, we do not consider covariates for $\sigma$. However, it is immediate that $\sigma$ 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

\[
\lambda \left( d \mid s_i, \theta, \sigma \right) = (2\pi)^{\frac{D}{2}} \left\{ \sum_{l=1}^{L} b_l \mathcal{I} \left( d \in w_l \right) \right\} \prod_{j=1}^{D} \phi \left( \frac{d_j - s_{ij}}{\sigma} \right).
\]  

(24)

**Resulting marginal void probability**

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

\[
D_{i,w} \sim \text{Poisson} \left\{ \Lambda \left( w \mid s_i, \theta, \sigma \right) \right\},
\]

(25)

where

\[
\Lambda \left( w \mid s_i, \theta, \sigma \right) = \int_w \lambda \left( x \mid s_i, \theta, \sigma \right) dx
\]

(26)

is the expected number of detections within \(w\). Then the expected number of detections \(\Lambda \left( w_l \mid s_i, \theta, \sigma \right)\) of individual \(i\) within sub-window \(w_l\) is

\[
\Lambda \left( w_l \mid s_i, \theta, \sigma \right) = (2\pi)^{\frac{D}{2}} b_l \prod_{j=1}^{D} \int_{w_{lj1}}^{w_{lj2}} \phi \left( \frac{x_j - s_{ij}}{\sigma} \right) dx_j
\]

(27)

\[
= (2\pi)^{\frac{D}{2}} \sigma^{D} b_l \prod_{j=1}^{D} \left[ \Phi \left( \frac{w_{lj2} - s_{ij}}{\sigma} \right) - \Phi \left( \frac{w_{lj1} - s_{ij}}{\sigma} \right) \right],
\]

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

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

(28)
According to Equation (10), the marginal void probability in the case of cell-based habitat selection and detection is

\[
P(D_{i,o} = 0 \mid \beta, \theta, \sigma, i \leq N) = \frac{1}{\Lambda(\tilde{o} \mid \beta)} \int_{\tilde{o}} \tilde{\lambda}(s_i \mid \beta) \exp \{-\Lambda(o \mid s_i, \theta, \sigma)\} \, ds_i
\]

\[
= \frac{1}{\Lambda(\tilde{o} \mid \beta)} \sum_{h=1}^{H} \tilde{\lambda}_h \int_{\tilde{w}_h} \exp \{-\Lambda(o \mid s_i, \theta, \sigma)\} \, ds_i,
\]

where \(\tilde{\Lambda}(\tilde{o} \mid \beta)\) is given in Equation (20), and the definite integral \(\int_{\tilde{w}_h} \exp \{-\Lambda(o \mid s_i, \theta, \sigma)\} \, ds_i\) can be calculated using the grid integration of Equation (28).