Statistical Science
2016, Vol. 31, No. 2, 219–232
DOI: 10.1214/16-STS557

Spatial Capture–Recapture Models

David Borchers and Rachel Fewster

Abstract. There has been a rapid growth in spatial capture–recapture (SCR) methods in the last decade. This paper provides an overview of existing SCR models and suggestions on how they might develop in future. The core of the paper is a likelihood framework that synthesises existing SCR models. This is used to illustrate similarities and differences between models.

The key difference between conventional capture–recapture models and SCR models is that the latter include a spatial point process model for individuals' locations and allow capture probability to depend on location. This extends the kinds of inferences that can be drawn from capture–recapture surveys, allowing them to address questions of a fundamentally spatial nature, relating to animal distribution, habitat preference, movement patterns, spatial connectivity of habitats and dependence of demographic parameters on spatial variables.

Key words and phrases: Capture–recapture, competing risks, detection hazard, Poisson process, spatial modelling.

1. INTRODUCTION

This review draws together the various spatial capture–recapture models in the literature into a common analytic framework. [We will call them "Spatial capture–recapture" (SCR) rather than "Spatially explicit capture–recapture" (SECR) methods, for brevity.] Design issues are not considered, nor are diagnostic or model selection issues. The focus is on the models, and to a lesser extent methods of inference. As the majority of existing SCR models deal with closed population estimation, open population SCR Jolly–Seber and Cormack–Jolly–Seber methods are reviewed only briefly in the last section of the paper. We assume that the reader is familiar with non-spatial CR models.

SCR models are distinguished from non-spatial CR models by the fact that they include individuals' locations as latent or hidden variables. This makes them hierarchical. They can be considered to be state space models of a sort, insofar as the locations of individuals in the population characterise the population state at any time, and the observation process depends on this state.

The paper is structured as follows: Section 2 introduces spatial point process models, Section 3 deals with spatial observation models, Section 4 develops likelihood functions and deals briefly with inference methods. Various extensions to the basic SCR models are covered in Sections 5 through 9, followed by a discussion and summary in Section 10.

2. SPATIAL POINT PROCESS MODELS

Consider a population of individuals occupying some finite two-dimensional region. The column vector $\mathbf{s}_i$ denotes the Cartesian coordinates of the location of individual i in the population. We treat individuals' locations as realisations of some spatial point process and arrange these vectors in an $N \times 2$ matrix $\mathbf{S}_N = (\mathbf{s}_1, \ldots, \mathbf{s}_N)'$ in which the number of rows (N) is random and the ith row is $\mathbf{s}_i'$. For a given value of N, we denote this matrix $\mathbf{S}_N|N$ to make conditioning on N explicit. The first n rows are the locations of the n individuals that are detected on a survey, and we denote these $\mathbf{S}_n$.

The notation "[ ]" is used to denote a probability distribution, so that $[\mathbf{S}_N, N] = [\mathbf{S}_N|N][N]$, for example, where $[N]$ is the probability mass function of N

D. Borchers is Professor of Statistics, Centre for Research into Ecological and Environmental Modelling, The Observatory, University of St Andrews, St Andrews, KY16 9LZ, Scotland (e-mail: dlb@st-andrews.ac.uk). R. Fewster is Associate Professor, Department of Statistics, University of Auckland, Private Bag 92019, Auckland, New Zealand (e-mail: r.fewster@auckland.ac.nz).

and $[\mathbf{S}_N|N]$ is the probability density function of individuals' locations, given that there are N individuals. To avoid more complicated notation than is essential, we do not distinguish in our notation between random variables and realisations of random variables. The distinction is usually apparent from the context.

We can think of $\mathbf{S}_N$ as arising from a spatial point process with two sub-levels:

$[N]$: The lowest level is a stochastic model for individual abundance, N.

$[\mathbf{S}_N|N]$: Given N, this is a stochastic model for the locations of the N individuals.

Most SCR models used to date assume a homogeneous, uncorrelated distribution of individuals in space. This is a computationally convenient but a biologically unrealistic model for most populations. Following Borchers and Efford (2008), we deal with the case in which individuals occur in the survey region according to a nonhomogeneous Poisson process (NHPP). This covers almost all SCR spatial models currently in the literature (Reich and Gardner, 2014, is an exception; see Section 2.1) and allows variation in density in space, but not dependence between the locations of individuals. We assume that individuals occur in the plane according to an NHPP with intensity $\lambda(\mathbf{s})$ at location $\mathbf{s}$. As $\lambda(\mathbf{s})$ is the population density at $\mathbf{s}$, which must be non-negative, dependence on $\mathbf{s}$ is naturally modelled using a log link function, that is,

$$\lambda(\mathbf{s}) = e^{\mathbf{x}'\boldsymbol{\beta}}, \tag{1}$$

where $\mathbf{x}$ is a column vector of explanatory variables. This would include $\mathbf{s}$ or spatially referenced variables evaluated at $\mathbf{s}$. To accommodate non-monotonic dependence on explanatory variables, Borchers and Kidney (2014) proposed use of regression splines in the linear predictor $\mathbf{x}'\boldsymbol{\beta}$ and this method is implemented in the maximum likelihood package "`secr`" (Efford, 2013).

With this model, N has a Poisson distribution

$$[N] = \frac{\Lambda^N e^{-\Lambda}}{N!}, \tag{2}$$

where $\Lambda = \int \lambda(\mathbf{s})\, d\mathbf{s}$, where integration is over the subset of $\mathbb{R}^2$ comprising the survey region. Conditional on N, and noting that the order of the rows in $\mathbf{S}_N$ is arbitrary so that there are $N!$ ways to get the same set of $\mathbf{s}_i$s,

$$[\mathbf{S}_N|N] = N! \prod_{i=1}^{N} \frac{\lambda(\mathbf{s}_i)}{\Lambda}. \tag{3}$$

The joint distribution of $\mathbf{S}_N$ and N constitutes an NHPP:

$$[\mathbf{S}_N, N] = e^{-\Lambda} \prod_{i=1}^{N} \lambda(\mathbf{s}_i). \tag{4}$$

2.1 More Complex Spatial Point Process Models

A somewhat limiting feature of an NHPP is that while it accommodates spatial variation in density, it does not accommodate spatial correlation. Using a Neyman–Scott process (see Illian et al., 2009, pp. 374–379) is one way to model this; another is to use a log-Gaussian Cox process (LGCP) (see Illian et al., 2009, pp. 381–382). This is obtained by modelling the intensity of the NHPP, $\lambda(\mathbf{s})$, as a log-linear function of a Gaussian random field $\{Z(\mathbf{s})\}$ (which has spatial correlation structure), for example, $\lambda(\mathbf{s}_i) = e^{Z(\mathbf{s}_i)}$.

One can also usefully extend an NHPP by adding "marks", that is, features of points, to give a marked NHPP. This allows individual-level characteristics that may vary in space to be modelled. One such feature might, for example, be group size, where group locations are governed by an NHPP or LGCP.

While LGCPs can deal with spatial clustering, they do not readily accommodate spatial avoidance, as would be typical of territorial species for example. Candidate spatial models for this situation include a Strauss process and other "pair potential" spatial models (see Illian et al., 2009, pp. 141–142). Reich and Gardner (2014) developed an SCR model using a Strauss process with constant density in the survey region (unlike the NHPP, which allows varying density).

3. SPATIAL DETECTION MODELS

The observation process in SCR studies depends on the realisation of the spatial point process: which individuals you detect where, depends on where the individuals are located. It also obviously depends on the detector locations. We denote the detector locations $\mathbf{L} = (\mathbf{l}_1, \ldots, \mathbf{l}_K)'$, where $\mathbf{l}_k$ $(k = 1, \ldots, K)$ is a column vector of Cartesian coordinates of the kth detector's location. Individuals are more likely to be detected by a detector that was close (by some measure) to its location ($\mathbf{s}_i$ for individual i) than a detector that was far away, so the observation process is a function of the distance measure d_{ik} of the individual's location, $\mathbf{s}_i$, from the detector's location, $\mathbf{l}_k$. For brevity, we do not include $\mathbf{L}$ explicitly in expressions for the observation process.

It is useful to formulate the observation process as a hierarchy too, conditional on the realisation $\mathbf{S}_N$ of the

spatial point process. To do this, we define for individual i ($i = 1, \ldots, N$) the following hierarchy of random variables:

- δ_i is 1 if the individual was detected at all in the course of the survey, and is zero otherwise,
- δ_{ij} is 1 if the individual was detected on occasion j, and is zero otherwise,
- δ_{ijk} is 1 if the individual was detected on occasion j by detector k, and is zero otherwise,
- n_{ijk} is the number of times that the individual was detected by detector k on occasion j, and
- t_{ijkm} is the time of the mth detection of the individual at detector k on occasion j.

Notice that the minimal data required for capture–recapture of any sort is δ_{ij}, while the minimal data required for spatial capture–recapture is δ_{ijk} (together with the locations of the detectors). The counts n_{ijk} provide additional information over and above that contained in δ_{ijk}, and detection times t_{ijkm} provide additional information over and above that. So any capture–recapture study that recorded *which* detectors made detections (and the locations of all detectors) is amenable to analysis using SCR methods. This means that any survey that recorded these data, even if it occurred prior to the advent of SCR methods, can be analysed using SCR methods.

We assume that individuals are detected independently. While this may not always be the case (see Section 3.2.1 below), inference becomes more difficult when it is not. The probability of observing an individual at any detector on a survey, given its location $\mathbf{s}$, is $p(\mathbf{s})$, and with independent detections the probability distribution of $\Delta = (\delta_1, \ldots, \delta_N)$ is

$$[\Delta|\mathbf{S}_N] = \prod_{i=1}^{N} p(\mathbf{s}_i)^{\delta_i}\{1 - p(\mathbf{s}_i)\}^{1-\delta_i}. \tag{5}$$

Before considering capture history models, we take a diversion to look at ways of parameterising the inclusion probability function $p(\mathbf{s})$.

3.1 Detection Hazard and Encounter Rate Models

For brevity, $p(\mathbf{s}_i)$ is written as p_i and it is the probability of individual i being detected on some occasion by some detector, while p_{ij} is the probability that it is detected on occasion j by some detector, and p_{ijk} is the probability that it is detected on occasion j by detector k.

It is useful to formulate the detection process for each individual at each detector in terms of a rate per unit time at which detections are expected to occur. We denote this rate for individual i on occasion j at detector k at time t by $h_{ijk}(t)$. When more than one detection is possible within an occasion (with proximity detectors—see below), the expected number of encounters of individual i on occasion j by detector k in the time period 0 to T_j, is the cumulative hazard over a time period, which we denote $H_{ijk} = \int_0^{T_j} h_{ijk}(u)\,du$. And whether or not more than one detection is possible per occasion, the probability of individual i "surviving" detection (i.e., failing to be detected) by detector k over a time period of length T_j on occasion j is $S_{ijk} = e^{-H_{ijk}}$. (Bold $\mathbf{S}$ indicates a matrix of Cartesian coordinates of individuals, while non-bold S indicates a detection survival function.)

Detection probability is the complement of the survival function S_{ijk}:

$$p_{ijk} = 1 - S_{ijk} = 1 - e^{-H_{ijk}}. \tag{6}$$

With independent detections over time, detection times $[\mathbf{t}_{ijk} = (t_{ijk1}, \ldots, t_{ijkm_{ijk}})$ for the m_{ijk} detections of individual i on occasion j by detector $k]$ are observations from an NHPP in time. As with the spatial NHPP intensity function, $\lambda()$, a log link function can be used to make the detection rate function $h()$ depend on covariates:

$$h_{jk}(\mathbf{s}_i, t) = e^{\mathbf{x}'_{ijk}\boldsymbol{\theta}}. \tag{7}$$

Here, $\mathbf{x}_{ijk}$ is a column vector of explanatory variables for individual i that includes $\mathbf{s}_i$, and if encounter rate varies with occasion or time, j or t. For example, with $\mathbf{x}'_{ijk} = (1, \|\mathbf{s}_i - \mathbf{l}_k\|)$ and $\boldsymbol{\theta} = (\theta_0, \theta_d)'$ the detection rate is a function of the distance $d_{ik} = \|\mathbf{s}_i - \mathbf{l}_k\|$ of the point $\mathbf{s}_i$ from the detector location $\mathbf{l}_k$, which for now we assume is the same for all j.

All SCR models are of type M_h in the classification scheme of Otis et al. (1978), that is, they have heterogeneity in detection probability at the level of the individual. This is because with SCR models, individuals' detection probabilities depend on their locations. There is scope for variety within each of the model types of Otis et al. (1978). For example, one could make the intercept (θ_0 above) and/or the range parameter (θ_d above) depend on occasion and/or behavioural response and/or additional individual covariates other than distance d_{ik}. In short, SCR models greatly extend the variety of possible model types, so that while the classification of Otis et al. (1978) is useful at a gross level, it is inadequate to describe the rich variety of possible SCR models. The ability to model detection

probability in more complex ways with SCR models does not of course mean that data will be adequate to support more complex models.

Not all SCR detection probability models in the literature use the hazard rate formulation, and most ignore T_j. Notwithstanding this, we can rewrite any detection probability model in terms of a detection hazard function for a given T. For example, the half-normal detection function $p_{ijk} = g_0 \exp\{-d_{ik}^2/\sigma^2\}$ $(0 \le g_0 \le 1; \sigma > 0)$ which is probably the most widely-used form, can be rewritten in a hazard rate form (with hazard h_{ijk} independent of time t) by letting $h_{ijk} = -\log(1 - g_0 \exp\{-d_{ik}^2/\sigma^2\})/T$. An advantage of formulating it in this manner is that it provides a way of modelling the dependence of capture probability on the length of time that individuals are at risk of detection (see Efford, Borchers and Mowat, 2013). The analyst has the option of modelling at the detection function level (e.g., $p_{ijk} = g_0 \exp\{-d_{ik}^2/\sigma^2\}$) or at the hazard level (e.g., $h_{ijk} = h_0 \exp\{-d_{ik}^2/\sigma_h^2\}$) and it is straightforward to switch between the two.

If individuals are detected independently, then the probability of detecting individual i on at least one detector on occasion j, and of detecting the individual at all are, respectively,

$$p_{ij} = 1 - e^{-H_{ij}} \quad \text{and} \quad p_i = 1 - e^{-H_i}, \tag{8}$$

where $H_i = \sum_j H_{ij}$ is the total detection hazard of the K detectors over the study for individual i and $H_{ij} = \sum_k H_{ijk}$ is the total detection hazard on occasion j for this individual. For notational brevity, $1 - p_{ij}$ and $1 - p_i$ are usually written as S_{ij} and S_i below.

3.2 Models for Capture Histories and Associated Responses

We write the capture history of individual i at detector k over the J occasions of the survey as $\boldsymbol{\omega}_{ik} = (\delta_{i1k}, \ldots, \delta_{iJk})$, and its capture history over all detectors is $\boldsymbol{\omega}_i = (\boldsymbol{\omega}_{i1}, \ldots, \boldsymbol{\omega}_{iK})$. We collect together all individuals' capture histories in $\Omega = (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_N)$. The form that its probability distribution takes depends on the kind of detectors used. Following Efford, Borchers and Byrom (2009), it is useful to distinguish between three kinds of detector:

Single-catch traps: these detain individuals when they detect them and are taken out of action by the detection process. A cage trap is an example of a single-catch trap.
Multi-catch traps: these detain individuals when they detect them but do not "fill up". Mist nests are examples of multi-catch traps.
Proximity detectors: these do not detain individuals and do not fill up. A camera trap is an example of a proximity detector.

It is also useful to distinguish between three kinds of response (i.e., kinds of detection):

Binary: only records of whether or not individuals were detected at each detector on each occasion are available.
Counts: records of how many times each individual was detected at each detector on each occasion are available.
Times: records of the times of each detection of each individual at each detector on each occasion are available.

Capture–recapture data are traditionally of the binary kind, but some detectors can generate counts or times of detection. Different kinds of detectors can generate different kinds of response, and impose different independence constraints on detections. These are summarised in Table 1.

All detector types involve independence between captures of the same individual on different occasions (given $\mathbf{s}_i$ and the covariates $\mathbf{x}_{ijk}$: $j = 1, \ldots, J; k =$

TABLE 1

Types of independence within occasions, and types of responses possible for different detector types. "(Yes)" indicates that this kind of response is possible in principle, but whether or not it is available depends on the particular detector used

Detector type	Independence		Possible response type		
	Between detectors within individual?	Between individuals at each detector?	Binary?	Count?	Time?
Single-catch	No	No	Yes	No	(Yes)
Multi-catch	No	Yes	Yes	No	(Yes)
Proximity	Yes	Yes	Yes	(Yes)	(Yes)

$1, \ldots, K$). We define $\delta_{ij} = 1$ if an individual was captured on occasion j and $\delta_{ij} = 0$ otherwise, and define the vector $\boldsymbol{\delta}_i = (\delta_{i1}, \ldots, \delta_{iJ})$. It follows that

$$[\boldsymbol{\delta}_i | \delta_i, \mathbf{s}_i] = \left\{ \frac{1}{p_i} \prod_j p_{ij}^{\delta_{ij}} S_{ij}^{1-\delta_{ij}} \right\}^{\delta_i}. \tag{9}$$

3.2.1 *Binary data*. With single-catch and multi-catch traps, more than one capture occasion is required to generate capture–recapture data and on occasion j individual i generates a binary response at detector k: δ_{ijk} which is 1 if the individual was captured on the occasion by the detector, and zero otherwise.

Detections between individuals within occasions are not independent in the case of single-catch traps (since catching one individual will in general reduce the capture probabilities for others, by taking one trap out of action). A general expression for the probability distribution of capture histories remains to be developed for this case.

With **multi-catch traps**, detections between individuals are independent, but detections between traps are not: once an individual is caught by one trap it cannot be caught by any others. In this sense, traps "compete" to catch individuals, and a competing risks survival model formulation is appropriate. So for a survey of J occasions, the conditional probability of observing capture history $\boldsymbol{\omega}_i$, given that individual i was detected on at least one occasion, is a product of multinomials with index 1:

$$[\boldsymbol{\omega}_i | \boldsymbol{\delta}_i, \mathbf{s}_i]_{\text{multi}} = \prod_{j=1}^{J} \prod_{k=1}^{K} \left(\frac{H_{ijk}}{H_{ij}} \right)^{\delta_{ijk}}. \tag{10}$$

With **proximity detectors**, in which detections of the same individual at different detectors are independent events, the capture history distribution, conditional on $\boldsymbol{\delta}_i$, is a product of Bernoulli distributions

$$[\boldsymbol{\omega}_i | \boldsymbol{\delta}_i, \mathbf{s}_i]_{\text{prox}} = \prod_{j=1}^{J} \left(\frac{1}{p_{ij}} \prod_{k=1}^{K} p_{ijk}^{\delta_{ijk}} S_{ijk}^{1-\delta_{ijk}} \right)^{\delta_{ij}}. \tag{11}$$

3.2.2 *Count data*. Some proximity detectors, such as camera traps, record how many times individuals were detected. If independence of detections can be assumed, then given $\boldsymbol{\omega}_{ik}$ and $\delta_i = 1$, we can write the probability of obtaining the count history $\mathbf{n}_i = (\mathbf{n}_{i1}, \ldots, \mathbf{n}_{iK})$, where $\mathbf{n}_{ik} = (n_{i1k}, \ldots, n_{iJk})$ for detector k and n_{ijk} is the number of times it was detected on occasion j by detector k, is the product of zero-truncated Poisson distributions, as follows (remembering that $p_{ijk} = 1 - e^{-H_{ijk}}$):

$$[\mathbf{n}_i | \boldsymbol{\omega}_i, \mathbf{s}_i]_{\text{prox}} = \prod_{j=1}^{J} \prod_{k=1}^{K} \left(\frac{1}{p_{ijk}} \frac{H_{ijk}^{n_{ijk}} e^{-H_{ijk}}}{n_{ijk}!} \right)^{\delta_{ijk}}. \tag{12}$$

3.2.3 *Event time data*. Some proximity detectors, like camera traps, record the exact times of detection or capture. In this case, given $\mathbf{n}_i$, and assuming independent detection between detectors, the p.d.f. of detection times $\mathbf{t}_i = (\mathbf{t}_{i11}, \ldots, \mathbf{t}_{i1K}, \ldots, \mathbf{t}_{iJ1}, \ldots, \mathbf{t}_{iJK})$, where $\mathbf{t}_{ijk} = (t_{ijk1}, \ldots, t_{ijkn_{ijk}})$, is

$$\begin{aligned} [\mathbf{t}_i | \mathbf{n}_i, \mathbf{s}_i]_{\text{prox}} &= \prod_{j=1}^{J} \prod_{k=1}^{K} [\mathbf{t}_{ijk} | n_{ijk}, \mathbf{s}_i] \\ &= \prod_{j=1}^{J} \prod_{k=1}^{K} n_{ijk}! \prod_{r=1}^{n_{ijk}} \frac{h_{ijk}(t_{ijkr})}{H_{ijk}} \end{aligned} \tag{13}$$

(where we define the product over r to be 1 if $n_{ijk} = 0$).

3.3 Spatial Capture History Models

The appropriate form for the capture history model is readily constructed by using the relevant conditional distributions above as building blocks. For example, with multi-catch traps and binary data (omitting $\boldsymbol{\theta}$ for brevity),

$$\begin{aligned} &[\Omega, \Delta | \mathbf{S}_N] \\ &\quad = [\Delta | \mathbf{S}_N] \prod_{i=1}^{N} [\boldsymbol{\delta}_i | \delta_i, \mathbf{s}_i] [\boldsymbol{\omega}_i | \boldsymbol{\delta}_i, \mathbf{s}_i]_{\text{multi}}. \end{aligned} \tag{14}$$

For proximity detectors with count data,

$$\begin{aligned} &[\mathbf{n}, \Delta | \mathbf{S}_N] \\ &\quad = [\Delta | \mathbf{S}_N] \prod_{i=1}^{N} [\boldsymbol{\delta}_i | \delta_i] [\boldsymbol{\omega}_i | \boldsymbol{\delta}_i, \mathbf{s}_i]_{\text{prox}} [\mathbf{n}_i | \boldsymbol{\omega}_i, \mathbf{s}_i]_{\text{prox}}, \end{aligned} \tag{15}$$

where $\mathbf{n} = (\mathbf{n}_1, \ldots, \mathbf{n}_N)$. And for proximity detectors with detection times

$$\begin{aligned} [\mathbf{t}, \Delta | \mathbf{S}_N] &= [\Delta | \mathbf{S}_N] \prod_{i=1}^{N} [\boldsymbol{\delta}_i | \delta_i] [\boldsymbol{\omega}_i | \boldsymbol{\delta}_i, \mathbf{s}_i]_{\text{prox}} \\ &\quad \cdot [\mathbf{n}_i | \boldsymbol{\omega}_i, \mathbf{s}_i]_{\text{prox}} [\mathbf{t}_i | \mathbf{n}_i, \mathbf{s}_i]_{\text{prox}}, \end{aligned} \tag{16}$$

where $\mathbf{t} = (\mathbf{t}_1, \ldots, \mathbf{t}_N)$.

For use below, note that the products in equations (14)–(16) can be factorised into a component for the n detected individuals (with $\delta_i = 1$) and the

$N-n$ undetected individuals. For example, the components for the detected individuals in the cases of equations (14) and (15) are

$$[\Omega_n|\mathbf{S}_n] = \prod_{i=1}^{n} [\boldsymbol{\delta}_i|\delta_i, \mathbf{s}_i][\boldsymbol{\omega}_i|\boldsymbol{\delta}_i, \mathbf{s}_i]_{\text{multi}} \tag{17}$$

and

$$\begin{aligned}&\big[(\mathbf{n}_1, \ldots, \mathbf{n}_n)|\mathbf{S}_n\big] \\ &= \prod_{i=1}^{n} [\boldsymbol{\delta}_i|\delta_i][\boldsymbol{\omega}_i|\boldsymbol{\delta}_i, \mathbf{s}_i]_{\text{prox}}[\mathbf{n}_i|\boldsymbol{\omega}_i, \mathbf{s}_i]_{\text{prox}},\end{aligned} \tag{18}$$

where $\Omega_n = (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_n)$ is the capture histories of detected individuals, and $\mathbf{S}_n = (\mathbf{s}_1, \ldots, \mathbf{s}_n)$ is their locations.

3.4 Data and Detection Model Hierarchy

SCR extends non-spatial CR by adding a spatially-referenced capture indicator, δ_{ijk}, for each detector within occasions. Depending on the detector type, we may also have capture frequencies, n_{ijk}, and perhaps a set of capture times, t_{ijkr} as well. There may be additional data layers on top of this also; see Section 7.1.

Our aim in constructing detection models in the hierarchical way that we have done above, is pedagogical rather than practical. We have tried to show how the various SCR detection models that appear in the literature are related to one another, how count data models are binary data models supplemented with counts, for example, and how time-to-detection models are count models with temporal data. While this formulation does let you construct any SCR detection model by using the appropriate building blocks, $[\boldsymbol{\delta}_i|\delta_i]$, $[\boldsymbol{\omega}_i|\boldsymbol{\delta}_i, \mathbf{s}_i]_{\text{prox}}$, etc., as in equations (14) to (16), this is probably not the most straightforward way of doing so if your aim is to get an expression for the detection model rather than to show how models are related. For example, equation (18) reduces to

$$[\Omega_n|\mathbf{S}_n] = \prod_{i=1}^{n}\prod_{j=1}^{J}\prod_{k=1}^{K} \frac{H_{ijk}^{n_{ijk}} e^{-H_{ijk}}}{n_{ijk}!}. \tag{19}$$

This is just a Poisson regression model. Neither $\boldsymbol{\delta}_i$ nor $\boldsymbol{\omega}_i$ appear in it, and you do not need $[\boldsymbol{\delta}_i|\delta_i]$, $[\boldsymbol{\omega}_i|\boldsymbol{\delta}_i, \mathbf{s}_i]_{\text{prox}}$ to obtain it.

4. INFERENCE METHODS

Three approaches to SCR inference have been developed. The first, inverse prediction, was developed by Efford (2004) and although Bayesian and maximum likelihood methods have to a large extent replaced it, it remains the only method that can explicitly deal with single-catch traps.

We do not cover single-catch trap likelihoods. Distiller and Borchers (2015) developed an approximate SCR MLE for the case in which the capture times are known, although Efford, Borchers and Byrom (2009) showed that use of a multi-catch SCR MLE for single-catch trap data works remarkably well for estimating overall density in many situations. This was confirmed by Distiller and Borchers (2015) although they found that in some such situations inferences about spatial distribution (rather than overall density) are biased.

4.1 Complete-Data and Semi-Complete-Data Likelihoods

Because Bayesian inference tends to proceed from what is sometimes called a "complete data likelihood" (see Schofield and Barker, 2016, for example) in this volume, which here means the likelihood as if we observed $\mathbf{S}_N$, we start with this. Two cases have been considered, one in which the distribution of $\mathbf{S}_N$ is modelled using equation (4), and one in which the conditional distribution of $\mathbf{S}_N|N$ given N, is modelled using equation (3). We refer to the former as the "Poisson model" and the latter as the "Binomial model".

King et al. (2016) developed what they call a "semi-complete-data likelihood", which here means the likelihood as if we observed the locations $\mathbf{S}_n$ of the n detected individuals, but not the locations of the undetected individuals. They developed this only for the Binomial model, but we consider it below for the Poisson model as well. Bayesian inference with a semi-complete-data likelihood has some advantages over Bayesian inference with a complete-data likelihood, as outlined below. Maximum likelihood estimation proceeds from the "observed data likelihood", and involves marginalising a semi-complete-data likelihood over the (unobserved) locations of detected individuals.

We illustrate complete-data and semi-complete-data likelihoods below using the capture histories Ω as the response. The cases in which one has count data or detection time data are obtained similarly, but with $\mathbf{n}$ or $\mathbf{t}$ in place of Ω.

4.1.1 *Complete data likelihoods*. Recall that the spatial point process model has a parameter vector $\boldsymbol{\beta}$ and the observation model has parameter vector $\boldsymbol{\theta}$. If we observed $\mathbf{S}_N$, then the likelihood for the parameters $(\boldsymbol{\beta}, \boldsymbol{\theta})$, given Ω, $\mathbf{S}_N$ and $\mathbf{X}$ can be written as follows

(note that once we know Ω, we know Δ, but for model description it is useful to condition on Δ in the likelihood):

$$\mathcal{L}_\lambda = [\mathbf{S}_N|\boldsymbol{\beta}] \times [\Delta|\mathbf{S}_N;\boldsymbol{\theta}][\Omega|\Delta,\mathbf{S}_N;\boldsymbol{\theta}]. \tag{20}$$

The spatial point process model is in the first term and the observation model is in the second and third.

For the Binomial model, in which we condition on N, the complete-data likelihood is as follows:

$$\mathcal{L}_N = [\mathbf{S}_N|N;\boldsymbol{\beta}] \times [\Delta|\mathbf{S}_N;\boldsymbol{\theta}][\Omega|\Delta,\mathbf{S}_N;\boldsymbol{\theta}]. \tag{21}$$

4.1.2 *Semi-complete-data likelihoods*. Noting that detecting individuals from an NHPP with probability $p(\mathbf{s})$ at $\mathbf{s}$, amounts to thinning the NHPP, the likelihood component for $\mathbf{S}_n$ is a thinned NHPP:

$$[\mathbf{S}_n;\boldsymbol{\beta},\boldsymbol{\theta}] = e^{-\tilde{\Lambda}} \prod_{i=1}^{n} \lambda_i p_i, \tag{22}$$

where $\lambda_i = \lambda(\mathbf{s}_i)$ and $\tilde{\Lambda} = \int \lambda(\mathbf{s}) p(\mathbf{s})\, d\mathbf{s}$. As there is no uncertainty about δ for detected individuals, the only other randomness in what is observed is in the capture histories of detected animals, $\Omega_n = (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_n)$, and so the Poisson model semi-complete-data likelihood involves the product of just two components:

$$\mathcal{L}_{\tilde{\lambda}} = [\mathbf{S}_n|\boldsymbol{\beta},\boldsymbol{\theta}][\Omega_n|\mathbf{S}_n;\boldsymbol{\theta}]. \tag{23}$$

To get the Binomial model semi-complete-data likelihood, we write the product $[\mathbf{S}_N|N;\boldsymbol{\beta}][\Delta|\mathbf{S}_N;\boldsymbol{\theta}]$ as

$$\begin{aligned}&[\mathbf{S}_N|N;\boldsymbol{\beta}][\Delta|\mathbf{S}_N;\boldsymbol{\theta}]\\ &\quad = N!\left\{\prod_{i=1}^{n} f_i p_i\right\}\left\{\prod_{i=n+1}^{N} f_i(1-p_i)\right\},\end{aligned} \tag{24}$$

where $f_i = f(\mathbf{s}_i) = \frac{\lambda(\mathbf{s}_i)}{\Lambda}$. Noting that $\mathbf{s}_{n+1}, \ldots, \mathbf{s}_N$ are i.i.d. so that integrals can be taken inside the second product above, we integrate the locations of the unobserved individuals out of this equation. Dividing by $(N-n)!$ to deal with the fact that without their locations in the likelihood, the $N-n$ unobserved individuals are indistinguishable from one another, we get the Binomial model semi-complete-data likelihood:

$$\begin{aligned}\mathcal{L}_{\tilde{N}} \propto \frac{N!}{(N-n)!}&\left\{\prod_{i=1}^{n} f_i p_i\right\}\\ &\cdot \{1-E(p)\}^{N-n}[\Omega_n|\mathbf{S}_n;\boldsymbol{\theta}],\end{aligned} \tag{25}$$

where $E(p) = \int f(\mathbf{s}) p(\mathbf{s})\, d\mathbf{s}$.

4.2 Bayesian Inference

Starting with Royle and Young (2008), Bayesian SCR methods have focussed on the Binomial model complete-data likelihood. With this approach $\mathbf{S}_N$ is treated as a parameter vector, and since N is unknown, the parameter space has unknown dimension. While reversible jump MCMC is one way of dealing with this, it has not been used to date. Instead, Bayesian approaches for SCR use data augmentation with regular MCMC, adding a vector of i.i.d. latent binary variables $z_1, \ldots, z_M$, with $M > N$ and $z_i = 1$ indicates that individual i is part of the population exposed to capture. It is assumed that $[z_i] = \phi^{z_i}(1-\phi)^{1-z_i}$, independently for all individuals, with ϕ, rather than N, being the focus of inference. There is an element of arbitrariness about the magnitude of M, but Royle et al. (2014) (p. 92) say that "results will be insensitive to choice of M, unless the data set is sufficiently small that parameters are weakly defined".

Another way of dealing with the fact that the complete-data likelihood has unknown parameter dimension, is to use one of the semi-complete-data likelihoods instead. These do not involve the locations of the undetected individuals and are therefore of fixed and known parameter dimension. This approach was developed for the Binomial model by King et al. (2016), who found it to be much more efficient than using the complete-data likelihood and data augmentation. No Bayesian inference method has been developed to date for the Poisson model semi-complete-data likelihood equation (23), but there is no obstacle to doing so.

The semi-complete-data likelihood approach exploits the fact that the capture histories for all undetected individuals are the same, reducing the dimension of the parameter space by $(N-n)$ (which can be a large number) and, crucially, converting the problem from one with unknown parameter dimension to one with known dimension, at the cost of having to evaluate a single two-dimensional numerical integral. This integral is $\int f(\mathbf{s}) p(\mathbf{s})\, d\mathbf{s}$ in the case of the Binomial model; see equations (24) and (25). In the case of the Poisson model, it is the integral $\int \lambda(\mathbf{s}) p(\mathbf{s})\, d\mathbf{s}$.

Although the parameter space for individual i's location, $\mathbf{s}_i$, is continuous, comprising a subset of $\mathbb{R}^2$ that includes all individual locations from which detection by at least one of the K detectors is possible, all Bayesian and frequentist inference approaches that have been implemented to date discretise this space into an $m_x \times m_y$ grid of points, which we will denote $\mathbb{S}$, and they allow only $\mathbf{s}_i \in \mathbb{S}$. (The Bayesian SCR

literature tends to call $\mathbb{S}$ a "state space", while the frequentist literature tends to call it a "mask".) The cost of too coarse a discretisation (m_x and/or m_y too small) is a poor approximation to the continuous space in which individuals are really located, and possibly biased inference. The advantage of coarser discretisation is computational speed. To implement both Bayesian and frequentist methods, one also needs to decide on the extent of $\mathbb{S}$, that is, a maximum distance of points from any detector that is to be included in $\mathbb{S}$. To avoid bias, this distance must not be shorter than the maximum distance from which an individual could be detected by a detector. As this is not known, or at best only roughly known, some experimenting with different sizes of $\mathbb{S}$ is usually done prior to doing inference, to find something like a minimum size and resolution of $\mathbb{S}$ beyond which estimates are insensitive to change.

4.3 Maximum Likelihood Inference

Maximum likelihood inference in the presence of latent variables typically proceeds by marginalising over the latent variables, and in the case of SCR this involves integrating $\mathbf{s}_1, \ldots, \mathbf{s}_n$ out of the Poisson model or Binomial model semi-complete-data likelihood functions (23) or (25). This is the approach that was developed for SCR maximum likelihood inference by Borchers and Efford (2008).

Integration is typically done numerically by defining a discrete location space $\mathbb{S}$ as above, and using something as simple as a trapezoidal rule, although more sophisticated numerical integration methods might also be used.

Maximum likelihood estimates (MLEs) of $\boldsymbol{\theta}$ and $\boldsymbol{\beta}$ (and with Binomial models, N) are obtained by numerically maximising this marginal likelihood with respect to these parameters. Interval estimates can usually be obtained using the inverse of the Hessian matrix obtained from numerical maximisation, by profile likelihood, or by bootstrap (see Section 7.2 below for a situation in which bootstrap is necessary). A number of authors, including Borchers and Efford (2008), Efford, Dawson and Borchers (2009) and Efford (2011) have found that using the inverse of the Hessian to estimate variance gives coverage probabilities close to the nominal 95% in the studies they considered.

5. ADDITIONAL INDIVIDUAL RANDOM EFFECTS

Borchers and Efford (2008) developed MLE methods with an additional latent variable by adapting the finite mixture model developed by Pledger (2000) in the context of non-spatial CR surveys, for SCR surveys. Royle et al. (2014) propose a model with a continuous, normally distributed random effect. The motivation for such models is as a mechanism for dealing with unobservable (and usually unknown) sources of heterogeneity in detection probability: the latent variable is a mechanism for modelling heterogeneity in capture probabilities that cannot be explained by observed covariates. We will denote the latent variable u, or $\mathbf{u}$ if it is a vector.

5.1 Partially Observed Latent Variables

Sometimes a latent variable is observed when individuals are detected but not otherwise. Sex is an example. Royle et al. (2015) developed a Binomial model and MLE for this case, while Efford (2013) implemented a Poisson model and MLE in the `R` package `secr`. These models are constructed as described for the latent variable above, but marginalisation is over only the latent variables associated with undetected individuals, or if the variable is unobserved for some detected individuals, for these and the undetected individuals.

5.2 The Conditional Likelihood Approach

An alternative to extending the hierarchical SCR model to incorporate a new latent variable layer, is to perform inference in two steps: (1) estimate detection probabilities p_{ijk}, and hence p_i for each detected individual from a conditional likelihood (conditional on individuals having been detected). (2) Use a Horvitz–Thompson-like estimator to estimate N (and/or density) conditional on the n estimates of $p_1, \ldots, p_n$. See Borchers and Efford (2008) for details.

The advantage of this method is that it does not require specification of $[\mathbf{u}]$. When u is univariate and binary, like sex, specifying a suitable form for $[\mathbf{u}]$ is straightforward. But when u is a vector, it may be difficult to know what a suitable form for the multivariate distribution $[\mathbf{u}]$ is, and in this case a conditional likelihood approach may be appealing.

6. MODELLING DETECTION BY MOVEMENT

Because individuals are detected by virtue of their movement with detectors like camera traps, these detectors allow inferences to be drawn about individuals' activity patterns, movements and habitat use.

6.1 Integrating Telemetry Data

6.1.1 *Telemetry data for detection probability.* A relatively simple way in which telemetry data can be

used with SCR survey data is to use an SCR detection function model to parameterise the telemetry data. One might, for example, assume that animals distribute themselves about their home range centre according to a bivariate normal distribution (as did Sollmann et al., 2013a, for example), and use this same bivariate normal model as an SCR detection function model.

6.1.2 *Telemetry data and resource selection*. Royle et al. (2013) proposed a model that integrates telemetry data and camera trap data for the case in which individuals are identifiable from photographs, and used it to draw inferences about habitat use and connectivity. This model was criticised by Efford (2014) because although the resource selection model includes a term accounting for the quality of the resources available to the individual in the vicinity of its activity centre, the detection hazard does not. This detection hazard implies that two individuals equidistant from a given detector, one with very good resources in its immediate vicinity, the other with very poor resources in its immediate vicinity are equally likely to travel to the detector—and this contradicts the resource selection model.

6.2 Non-Euclidian Distance

The SCR detection models described above use Euclidian distance d_{ik} to measure the distance of individuals' activity centres from detectors, but when detection is by virtue of movement, as is the case with traps, this may be too simple a model. Sutherland, Fuller and Royle (2015) developed an SCR model that uses a least-cost distance metric to reflect habitat preference in the way animals move, and that moving the same distance through different environments has different costs. A parameter determining the extent to which movement is affected by habitat covariates is estimated with other SCR parameters. In addition to accommodating habitat preference in the SCR detection function, the model allows inferences to be drawn about space usage patterns and habitat connectivity.

6.3 Alternative Detection Function Parameterisations

Efford and Mowat (2014) argue on biological grounds that the range and intercept of the detection function p_{ijk} will tend to be negatively correlated. (Individuals that travel further on average will tend to spend less time at any point and, therefore, have lower detection probability at any point, including at their activity centres, than individuals that travel less.) They therefore propose a reparameterisation of the detection hazard, in which one parameter involves the product of the intercept parameter and the range parameter of the hazard function [i.e., $(\sigma^2 h_0)$, where $h_{ijk} = h_0 \exp\{-d_{ik}^2/2\sigma^2\}$] and the other is the range parameter (σ^2). The negative correlation between h_0 and σ^2 results in $(\sigma^2 h_0)$ being much more stable than either of its component parameters, and not highly correlated with σ^2. In a broadly similar vein, Efford et al. (2015) note that there is evidence that home range size, encapsulated in the half-normal detection function range parameter σ, is inversely proportional to root density $\sqrt{D}$ across a range of densities and propose using a parameter that is the product of these two.

7. ACOUSTIC SURVEYS AND SUPPLEMENTARY LOCATION DATA

Whereas most SCR surveys involve detection and identification of individuals, some involve detection of what we will call "cues", that is, evidence of the presence of individuals by virtue of something they produce. Dung surveys are an example: because genetic identification of individuals from dung is possible, SCR methods can be used to estimate animal density without ever detecting any animals, but detecting their dung. This is useful because for some species dung can be much easier to detect and to survey than the animals producing it.

Similarly, SCR methods can be used to estimate density by detecting vocalisations without ever detecting the vocalising individuals. This allows SCR methods to be used with vocalising species that are difficult to see but easy to hear. Examples include various species of songbird (see Efford, Dawson and Borchers, 2009, Efford and Dawson, 2009, e.g.), frogs (see Stevenson et al., 2015, e.g.) vocalising primates such as gibbons (see Borchers et al., 2014, e.g.) and some cetacean species (see Marques et al., 2010, e.g.).

7.1 Supplementary Location Data and Acoustic Detection Functions

A somewhat unique feature of acoustic detections is that they often come with supplementary information about the location of the vocalising individuals, in the form of precise times of arrival of sound at detectors (when detectors are synchronised microphones), received signal strength (again with microphones), estimated angles to sound source (with human detectors and some types of microphone) and/or estimated distances to sound sources (in the case of human detectors). These data can be used by incorporating probability models that are conditional on sound source location, $\mathbf{s}_i$, for each kind of supplementary location data,

into likelihoods. Dawson and Efford (2009), Borchers et al. (2015) and Stevenson et al. (2015) developed and used additional probability models for received signal strength, Borchers et al. (2015) and Stevenson et al. (2015) for time-of-arrival data, and Borchers et al. (2015) for estimated angles and estimated distances. Incorporating such data can improve estimator precision substantially and in some cases reduce bias (see Borchers et al., 2015, Stevenson et al., 2015, for example).

Efford, Dawson and Borchers (2009) and Stevenson et al. (2015) developed detection functions specifically for acoustic data, which are formulated in terms of received signal strength exceeding some threshold value. One can envisage that the detection models with non-Euclidian distance developed by Sutherland, Fuller and Royle (2015) will be very useful for acoustic SCR surveys in which acoustic propagation is habitat-dependent: in the same way that animals will move preferentially through certain habitats, sound propagates better through some habitats than others.

7.2 Acoustic Identification and Independence

There are problems that arise on acoustic surveys that do not arise when individuals themselves are detected. Primary among these is the difficulty of individual identification from acoustic signals. While it may be relatively easy to identify "recaptures" of the same vocalisation on different detectors, it may be very difficult to identify individuals themselves from their calls. In this case, one can estimate call density from the SCR survey and separately estimate call rate in order to estimate individual density (see Stevenson et al., 2015).

Another problem is lack of independence of call source locations. Because the locations of different calls from the same individual are not independent if the individual's location is not independent for each call (as is virtually always the case), care needs to be taken to incorporate this lack of independence when obtaining interval estimates when we do not know which calls are from which individuals. Stevenson et al. (2015) describes a bootstrap method for obtaining valid interval estimates in this case.

8. INCOMPLETE CAPTURE HISTORIES

We consider the case in which a known number, n, of individuals (regarded as a random sample from the population) carry individual marks and the rest of the population cannot be individually identified. A survey with these characteristics is often called a "mark-resight" survey.

While SCR models and inference with a marked sample of the population are possible with single-catch, multi-catch and proximity detectors generating binary, count or detection time data, only SCR with proximity detectors and count data has been considered in the literature. This is no doubt because (a) with single- and multi-catch traps you get the opportunity to mark all captured individuals so "mark-resight" methods are redundant, (b) the kinds of proximity detectors that generate only binary data (hair snares and dung surveys, e.g.) generally allow identification of all captured individuals, and (c) models with detection times have seldom been used to date. Here, we review only Binomial models for proximity detectors with count data.

For this scenario, the part of the likelihood associated with the known number n of individuals with known identities is

$$\mathcal{L}_n \propto \prod_{i=1}^{n} f_i \prod_{j=1}^{J} \prod_{k=1}^{K} \frac{H_{ijk}^{n_{ijk}} e^{-H_{ijk}}}{n_{ijk}!}. \tag{26}$$

In addition to the recapture data, total counts u_{jk} $(j = 1, \ldots, J; k = 1, \ldots, K)$ of unmarked/unidentifiable individuals are observed at each detector on each occasion (without knowing how many individuals the counts correspond to). The joint distribution of $\mathbf{u} = (u_{11}, \ldots, u_{JK})$ and $\mathbf{S}_u = (\mathbf{s}_{n+1}, \ldots, \mathbf{s}_N)$ is easy to write down but computationally intractable for all but the smallest and simplest datasets. It is

$$[\mathbf{u}, \mathbf{S}_u] = \sum_{\mathbf{n} \in \mathcal{N}} \prod_{i=n+1}^{N} f_i \prod_{j=1}^{J} \prod_{k=1}^{K} \frac{H_{ijk}^{n_{ijk}} e^{-H_{ijk}}}{n_{ijk}!}, \tag{27}$$

where $\mathcal{N}$ is the set of all realisations of $\mathbf{n} = (n_{(n+1)11}, \ldots, n_{NJK})$ such that for all j, k, $\sum_i n_{ijk} = u_{jk}$. The complete-data Binomial model likelihood is the product of equations (26) and (27).

Chandler and Royle (2013) developed a Bayesian data augmentation method for drawing inferences about N and other parameters from equation (27) alone (without any capture history data), assuming uniform f_i. This method relies entirely on spatial clustering of detections informing inferences about abundance and is therefore likely to be sensitive to assumptions about both f_i and about (lack of) spatial variation in detection functions, although this has not been investigated. They applied the method to an acoustic survey of birds using 105 detectors, spaced such that individuals are virtually certain to be detected at more than one detector. While they were able to obtain posterior estimates of N, the credible intervals ranged from the

minimum possible number (i.e., 1) to roughly six times the posterior median, and posterior estimates were sensitive to assumed priors (results are summarised in Table 18.2 of Royle et al., 2014). This method has not been shown to produce practically useful estimates of abundance.

However, when a sample of the population is marked [and the product of equations (26) and (27) is used as the likelihood], estimator performance can improve markedly, so that this method is potentially very useful for a wide variety of species—for any population that is not naturally marked, but in which a sufficient number of randomly sampled individuals can be marked. Table 2 of Chandler and Royle (2013) summarises the improvement in estimator performance in a simulated population of 75 individuals as between about 7% and 50% of the population is marked, with RMSE decreasing by about 60%.

Building on work of Chandler and Royle (2013), Sollmann et al. (2013b, 2013a) integrated telemetry data and associated likelihood (some marked individuals also had radio collars; see also Section 6.1 above) with the above likelihood for mark-resight data. Credible interval half-widths were estimated to be about 8% of the posterior median in the case of Sollmann et al. (2013b) and about 50% in the case of Sollmann et al. (2013a), which involved only three radio-collared individuals.

9. OPEN POPULATION MODELS

9.1 Modelling with a Series of Unlinked SCR Estimates

With a series of SCR surveys, each providing an estimate of density or abundance, the simplest way to model change in population size across the duration of the surveys is to fit an empirical smooth of density or abundance over time. This approach is characterised by the fact that recapture data do not persist between the series of SCR surveys (or if they do, are not used). SCR data are used to obtain a series of abundance estimates, without linking detections across the series. Borchers and Efford (2008) did this assuming a log link between density and time, which restricts change to being monotonic (always increasing or always decreasing) but one can quite easily allow for non-monotonic change too. Following a method proposed by Borchers and Kidney (2014) for spatial smoothing, Efford (2013) implemented temporal smoothing using regression splines.

9.2 More Integrated Open Population SCR Modelling

A more elegant and more powerful way to model population dynamics with SCR data is to incorporate recapture data across the whole period of interest. Relatively little work has been done on open population SCR models of this sort. Royle et al. (2014) and Gardner et al. (2010) consider Jolly–Seber (Jolly, 1965, Seber, 1965) SCR (JSSCR) models, while Ergon and Gardner (2013) and Royle et al. (2014) developed Cormack–Jolly–Seber (Cormack, 1964, Jolly, 1965, Seber, 1965) SCR (CJSSCR) models. Two key differences between non-spatial CJS and CJSSCR models are (a) that movement, and hence emigration, is modelled explicitly and separately from survival (see Ergon and Gardner, 2013), which requires a spatio-temporal model for how locations change between surveys, and (b) that survival can be made dependent on location. This adds complexity, but it also adds versatility in the ability to explicitly model survival as a function of location.

Complete-data JSSCR likelihood functions, which model recruitment as well as survival, can be developed in a similar way, and although they did not write the likelihood down, Gardner et al. (2010) developed an open population model that accommodates birth and death (albeit without spatial dependence), as did Royle et al. (2015). Both used Bayesian data augmentation with MCMC methods for inference. Bayesian MCMC methods are well suited to open population models because they allow sampling from the posterior without having to marginalise over latent variables. With closed population SCR methods this is not really an advantage because the latent variable structure is usually low dimensional and simple enough that marginalisation is quick and easy. But because of the time series nature of open population models, marginalising is more difficult, and Bayesian MCMC methods provide a powerful general tool for inference in such cases.

9.3 Movement and Transience

Inference with a realistic movement model can be challenging. Gardner et al. (2010) avoid this problem by assuming no movement across surveys while Royle et al. (2015) assume independence in locations of individuals across surveys. These are the two relatively easy cases: constant locations and temporally independent locations.

For many studies, a more realistic model is one in which individuals' locations vary in time but there

is temporal dependence in any individual's location across time. Royle et al. (2015) note the similarity between such a model and multi-state CR models, while also noting the fact that SCR involves continuous space (albeit approximated by a large number of discrete points), whereas multi-state CR models involve very few discrete states, and this makes inference for SCR open population models with temporally correlated locations considerably more challenging. Gardner et al. (2010) and Royle et al. (2015) proposed a diffusion process model for movement, Ergon and Gardner (2013) proposed what they describe as a more general dispersal model for their CJSSCR model, which they fit using MCMC methods, and Royle, Fuller and Sutherland (2016) developed a Bayesian SCR estimator of abundance and movement, modelling movement as Markovian transience or dispersion. Royle, Fuller and Sutherland (2016) found their estimator to perform well, and interestingly also found that a maximum likelihood estimator that assumes no movement in locations estimated abundance with little bias, although it gave biased estimates of detection function range parameters. They therefore caution against drawing inferences about movement from detection function range parameters from this model when there may be movement of individuals' activity centres. A more realistic model for many species might be an Ornstein–Uhnlenbeck process (Ornstein and Uhnlenbeck, 1930), which unlike a diffusion process, has a centre of attraction, but this remains to be investigated with SCR models.

10. DISCUSSION AND SUMMARY

To the extent that all CR surveys involve captures at given spatial locations, all CR surveys are SCR surveys. SCR methods can do everything that non-spatial CR methods can do, and more besides (estimating density, modelling spatial distribution and dealing with unmodelled heterogeneity due to heterogeneity in individual locations, for example). Non-spatial CR methods are a special case of SCR methods, the special case being that spatial information has been discarded or has not been recorded.

One important practical aspect of SCR surveys that we have omitted is the spatial extent of individual detectors. We have treated the detectors as points above, but SCR methods for detectors that are lines or areas of polygons have also been developed. These involve somewhat different detection function models; see Royle and Young (2008) and Efford (2011) for details and examples.

Because Bayesian methods deal so easily with hierarchical models, they seem naturally suited to SCR inference. But for most SCR models, marginalising over the individual random effects (and individual locations, in particular) is numerically straightforward and fast, and frequentist inference methods have proved to be faster than Bayesian methods. Bayesian methods come into their own in more complex scenarios, like those of Sections 8 and 9.2, where the difficulty of marginalising presents very substantial challenges for frequentist methods.

Existing methods have only begun to explore the potential for drawing inferences about the spatio-temporal processes of populations from SCR data. Likely future developments include the following:

Integrated movement modelling: There is a rapidly growing body of research developing and fitting realistic models for modelling animal movement trajectories from individual radio tag and GPS tag data. There is opportunity here to integrate individual movement models and continuous-time SCR models to improve inference at both population and individual levels.

Recapture uncertainty: While there has been a lot of recent research activity dealing with this for non-spatial CR, (see Link, 2003, Wright et al., 2009, Link et al., 2010, McClintock et al., 2014a, 2014b, Fewster, Stevenson and Borchers, 2016, in this volume, for example) no methods have as yet been developed for dealing with uncertainty about recapture identity on SCR surveys. The increasing use of automated methods of species and individual identification from visual or acoustic recordings is likely to require new methods of dealing with this uncertainty.

Better spatial and spatio-temporal modelling: SCR methods have only just begun to realise their potential to inform realistically complex spatial and spatio-temporal models, and we can expect substantial developments in this area in the near future.

In summary, SCR methods have added a new dimension to CR, both literally and figuratively. A simple illustration of the sea change that this induces is the fact that reliable SCR inference is possible using data from a single occasion with proximity detectors (in which case capture histories are purely spatial), whereas a survey with a single occasion is an anathema to non-spatial CR.

In the same vein, a fundamental way in which SCR changes CR is that by explicitly including space, it provides a new tool for addressing questions of a fundamentally spatial nature such as species distribution, habitat preference, movement patterns, spatial connectivity and spatial aspects of population dynamics.

REFERENCES

BORCHERS, D. L. and EFFORD, M. G. (2008). Spatially explicit maximum likelihood methods for capture–recapture studies. *Biometrics* **64** 377–385. MR2432407

BORCHERS, D. L. and KIDNEY, D. (2014). Flexible density surface estimation for spatially explicit capture–recapture surveys. Technical report, Univ. St Andrews.

BORCHERS, D. L., DISTILLER, G., FOSTER, R., HARMSEN, B. and MILAZZO, L. (2014). Continuous-time spatially explicit capture–recapture, with an application to a jaguar camera-trap survey. *Methods in Ecology & Evolution* **5** 565–665.

BORCHERS, D. L., STEVENSON, B. C., KIDNEY, D., THOMAS, L. and MARQUES, T. A. (2015). A unifying model for capture–recapture and distance sampling surveys of wildlife populations. *J. Amer. Statist. Assoc.* **110** 195–204. MR3338496

CHANDLER, R. B. and ROYLE, J. A. (2013). Spatially explicit models for inference about density in unmarked or partially marked populations. *Ann. Appl. Stat.* **7** 936–954. MR3113496

CORMACK, R. (1964). Estimates of survival from the sighting of marked animals. *Biometrika* **51** 429–438.

DAWSON, D. K. and EFFORD, M. G. (2009). Bird population density estimated from acoustic signals. *Ecology* **46** 1201–1209.

DISTILLER, G. and BORCHERS, D. L. (2015). A spatially explicit capture–recapture estimator for single-catch traps. *Ecol. Evol.* **5** 5075–5087.

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

EFFORD, M. G. (2011). Estimation of population density by spatially explicit capture–recapture analysis of data from area searches. *Ecology* **92** 2202–2207.

EFFORD, M. G. (2013). secr: Spatially explicit capture–recapture. Version 2.7.0. Univ. Otago, Available at http://www.otago.ac.nz/density.

EFFORD, M. G. (2014). Bias from heterogeneous usage of space in spatially explicit capture–recapture analyses. *Methods in Ecology & Evolution* **5** 599–602.

EFFORD, M. G., BORCHERS, D. L. and BYROM, A. E. (2009). Density estimation by spatially explicit capture–recapture: Likelihood-based methods. In *Modeling Demographic Processes in Marked Populations* (D. L. Thompson, E. G. Cooch and M. J. Conroy, eds.) 255–269. Springer, New York.

EFFORD, M. G., BORCHERS, D. L. and MOWAT, G. (2013). Varying effort in capture–recapture studies. *Methods in Ecology & Evolution* **4** 629–636.

EFFORD, M. G. and DAWSON, D. K. (2009). Bird population density estimated from acoustic signals. *J. Appl. Ecol.* **46** 1201–1209.

EFFORD, M. G., DAWSON, D. K. and BORCHERS, D. L. (2009). Population density estimated from locations of individuals on a passive detector array. *Ecology* **90** 2676–2682.

EFFORD, M. G. and MOWAT, G. (2014). Compensatory heterogeneity in spatially explicit capture–recapture data. *Ecology* **95** 1341–1348.

EFFORD, M. G., DAWSON, D. K., JHALA, Y. V. and QURESHI, Q. (2015). Density-dependent home-range size revealed by spatially explicit capture–recapture. *Ecography* **38** 1–13.

ERGON, T. and GARDNER (2013). Separating mortality and emigration: Modelling space use, dispersal and survival with robust-design spatial capture–recapture data. *Methods in Ecology & Evolution* **5** 1327–1336.

FEWSTER, R. M., STEVENSON, B. C. and BORCHERS, D. L. (2016). Trace-contrast models for capture–recapture without capture histories. *Statist. Sci.* **31** 245–258.

GARDNER, B., REPUCCI, J., LUCHERINI, M. and ROYLE, J. A. (2010). Spatially explicit inference for open populations: Estimating demographic parameters from camera-trap studies. *Ecology* **91** 3376–3383.

ILLIAN, J., PENTTINEN, A., STOYAN, H. and STOYAN, D. (2009). *Statistical Analysis and Modelling of Spatial Point Patterns*. Wiley, New York.

JOLLY, G. M. (1965). Explicit estimates from capture–recapture data with both death and immigration-stochastic model. *Biometrika* **52** 225–247. MR0210227

KING, R., MCCLINTOCK, B. T., KIDNEY, D. and BORCHERS, D. L. (2016). Capture–recapture abundance estimation using a semi-complete data likelihood approach. *Ann. Appl. Stat.* **10** 264–285. MR3480496

LINK, W. A. (2003). Nonidentifiability of population size from capture–recapture data with heterogeneous detection probabilities. *Biometrics* **59** 1123–1130. MR2019822

LINK, W. A., YOSHIZAKI, J., BAILEY, L. L. and POLLOCK, K. H. (2010). Uncovering a latent multinomial: Analysis of mark-recapture data with misidentification. *Biometrics* **66** 178–185. MR2756704

MARQUES, T. A., THOMAS, L., MARTIN, S. W., MELLINGER, D. K., JARVIS, S., MORRISSEY, R. P., CIMINELLO, C. and DIMARZIO, N. (2010). Spatially explicit capture recapture methods to estimate minke whale abundance from data collected at bottom mounted hydrophones. *Journal of Ornithology* **152** S445–S455.

MCCLINTOCK, B. T., BAILEY, L. L., DREHER, B. P. and LINK, W. A. (2014a). Probit models for capture–recapture data subject to imperfect detection, individual heterogeneity and misidentification. *Ann. Appl. Stat.* **8** 2461–2484. MR3292505

MCCLINTOCK, B. T., HILL, J. M., FRITZ, L., CHUMBLEY, K., LUXA, K. and DIEFENBACH, D. R. (2014b). Mark-resight abundance estimation under incomplete identification of marked individuals. *Methods in Ecology & Evolution* **5** 1295–1304.

ORNSTEIN, L. S. and UHNLENBECK, G. E. (1930). On the theory of Brownian motion. *Phys. Rev.* **36** 823–841.

OTIS, D. L., BURNHAM, K. P., WHITE, G. C. and ANDERSON, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs* **62** 1–135.

PLEDGER, S. (2000). Unified maximum likelihood estimates for closed capture–recapture models using mixtures. *Biometrics* **56** 434–442.

REICH, B. J. and GARDNER, B. (2014). A spatial capture–recapture model for territorial species. *Environmetrics* **25** 630–637. MR3295555

ROYLE, A. J., FULLER, A. K. and SUTHERLAND, C. (2016). Spatial capture–recapture models allowing Markovian transience of dispersal. *Population Ecology* **58** 53–62.

ROYLE, J. A. and YOUNG, K. V. (2008). A hierarchical model for spatial capture–recapture data. *Ecology* **89** 2281–2289.

ROYLE, A. J., CHANDLER, R. B., SUN, C. C. and FULLER, A. K. (2013). Integrating resource selection information with spatial capture–recapture. *Methods in Ecology and Evolution* **4** 520–530.

ROYLE, J. A., CHANDLER, R. B., SOLLMANN, R. and GARDNER, B. (2014). *Spatial Capture–Recapture*. Academic Press, San Diego, CA.

ROYLE, J. A., SUTHERLAND, C., FULLER, A. K. and SUN, C. C. (2015). Likelihood analysis of spatial capture–recapture models for stratified or class structured populations. *Ecosphere* **6** 1–11.

SCHOFIELD, M. R. and BARKER, R. J. (2016). 50-year-old curiosities: Ancillarity and inference in capture–recapture models. *Statist. Sci.* **31** 161–174.

SEBER, G. A. F. (1965). A note on the multiple-recapture census. *Biometrika* **52** 249–259. MR0210228

SOLLMANN, R., GARDNER, B., CHANDLER, R. B., SHINDLE, D. B., ONARATO, D. P., ROYLE, J. A. and O'CONNELL, A. F. (2013a). Using multiple data sources provides density estimates for endangered Florida panther. *J. Appl. Ecol.* **50** 961–968.

SOLLMANN, R., GARDNER, B., PARSONS, A. W., STOCKING, J. J., MCCLINTOCK, B. T., SIMONS, T. R., POLLOCK, K. J. and O'CONNELL, A. F. (2013b). A spatial mark-resight model augmented with telemetry data. *Ecology* **94** 553–559.

STEVENSON, B. C., BORCHERS, D. L., ALTWEGG, R., SWIFT, R. J., GILLESPIE, D. M. and MEASEY, G. J. (2015). A general framework for animal density estimation from acoustic detections across a fixed microphone array. *Methods in Ecology & Evolution* **6** 38–48.

SUTHERLAND, C., FULLER, A. K. and ROYLE, J. A. (2015). Modelling non-Euclidean movement and landscapte connectivity in highly structured ecological networks. *Methods in Ecology & Evolution* **6** 169–177.

WRIGHT, J. A., BARKER, R. J., SCHOFIELD, M. R., FRANTZ, A. C., BYROM, A. E. and GLEESON, D. M. (2009). Incorporating genotype uncertainty into mark-recapture-type models for estimating abundance using DNA samples. *Biometrics* **65** 833–840. MR2649856