Received: 20 November 2018 | Accepted: 28 August 2019

DOI: 10.1111/biom.13150

Biometrics A JOURNAL OF THE INTERNATIONAL BIOMETRIC SOCIETY WILEY

BIOMETRIC METHODOLOGY

A spatial open-population capture-recapture model

Murray G. Efford | **Matthew R. Schofield**

Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand

Correspondence
Murray G. Efford, Department of Mathematics and Statistics, University of Otago, P.O. Box 56, Dunedin 9054, New Zealand.
Email: murray.efford@otago.ac.nz

Abstract

A spatial open-population capture-recapture model is described that extends both the non-spatial open-population model of Schwarz and Arnason and the spatially explicit closed-population model of Borchers and Efford. The super-population of animals available for detection at some time during a study is conceived as a two-dimensional Poisson point process. Individual probabilities of birth and death follow the conventional open-population model. Movement between sampling times may be modeled with a dispersal kernel using a recursive Markovian algorithm. Observations arise from distance-dependent sampling at an array of detectors. As in the closed-population spatial model, the observed data likelihood relies on integration over the unknown animal locations; maximization of this likelihood yields estimates of the birth, death, movement, and detection parameters. The models were fitted to data from a live-trapping study of brushtail possums (*Trichosurus vulpecula*) in New Zealand. Simulations confirmed that spatial modeling can greatly reduce the bias of capture-recapture survival estimates and that there is a degree of robustness to misspecification of the dispersal kernel. An R package is available that includes various extensions.

KEYWORDS

dispersal, population growth rate, Pradel-Link-Barker models, recruitment, spatially explicit capture-recapture, survival

1 | INTRODUCTION

Open-population capture-recapture models are used widely in population ecology and wildlife management to estimate animal abundance and related parameters (Seber, 1982; Williams *et al*., 2002). Schwarz and Arnason (1996) consolidated work by Jolly (1965), Seber (1965), and Crosbie and Manly (1985) to allow estimation of time-specific survival, recruitment, population growth rate, and population size by maximum likelihood. The Jolly-Seber-Schwarz-Arnason (JSSA) model is the foundation for later work on the joint estimation of these parameters (eg, Link and Barker, 2005; Pledger *et al*., 2010; Schofield and Barker, 2016). The JSSA model is non-spatial in that it is parameterized in terms of population size N, does not address the spatial distribution of animals, and does not model capture as a spatial process. Spatially explicit capture-recapture (SECR) methods have developed separately; these were designed to estimate the density of a spatially distributed but demographically closed population, sampled with an array of passive detectors such as traps (Efford, 2004; Borchers and Efford, 2008; Royle *et al*., 2014; Borchers and Fewster, 2016).

The potential benefits of generalizing SECR to demographically open populations are widely recognized. These include relaxing the assumption of closure when sampling is protracted, allowing for temporal variation in the spatial extent of sampling, distinguishing mortality from emigration, and reducing bias due to spatially

induced individual heterogeneity in capture probability. Several studies have addressed aspects of the problem (Gardner *et al*., 2010; Ergon and Gardner, 2014; Schaub and Royle, 2014; Raabe *et al*., 2014; Chandler and Clark, 2014; Whittington and Sawaya, 2015; Gardner *et al*., 2018). Most have relied on a Bayesian formulation and model fitting by MCMC with either conditioning on the first detection, as in the Cormack-Jolly-Seber model (Lebreton *et al*., 1992), or data augmentation. Glennie *et al*. (2019) recently provided a hidden-Markov formulation that could be fit by maximizing the likelihood.

This paper describes a spatial open-population model extending both the non-spatial open-population model of Schwarz and Arnason (1996) and the spatially explicit closed-population model of Borchers and Efford (2008). The model accommodates, but does not require, the "robust" sampling design of Pollock (1982) in which multiple secondary samples are nested within each primary sample. Section 2 introduces the data structure and notation and separately describes the spatial and non-spatial precursor models. Section 3 describes the open-population spatial model, initially for static home-range centers and then with between-session movement according to a dispersal kernel. Section 4 deals with model fitting and Section 5 presents an example. Simulations are used in Section 6 to evaluate spatial and non-spatial models with respect to bias in survival estimates and to assess the effect of misspecifying the movement kernel in spatial models. Section 7 has some concluding remarks.

TABLE 1 Major notation

General	
n	Number of individuals observed in the course of study
J	Number of primary sampling sessions
S	Number of secondary sampling sessions
Spatial	
K	Number of detectors
$\boldsymbol{\omega}$	Spatial detection histories of all observed animals
$\boldsymbol{\omega}_i$	Spatial detection history of observed animal i
$\boldsymbol{\theta}$	Vector of parameters for the spatial detection process
$\boldsymbol{\nu}$	Vector of parameters for spatial density process
$\mathbf{x}$	Location in-plane (vector of x-y coordinates)
$D(\mathbf{x})$	Density at $\mathbf{x}$ of animals available for detection in at least one session
$p_{sk}(\mathbf{x})$	Probability animal centered at $\mathbf{x}$ is detected at detector k in secondary session s
$p_{\cdot}(\mathbf{x})$	Probability animal centered at $\mathbf{x}$ is detected at at least once during study
Non-spatial	
$\boldsymbol{\omega}'$	Collapsed (non-spatial) detection histories of observed animals
$\boldsymbol{\omega}_i'$	Non-spatial detection history of observed animal i
$\boldsymbol{\theta}'$	Vector of parameters for non-spatial detection process
N	Number of animals available for detection in at least one session
Population turnover	
ϕ_j	Probability animal available for detection at time j is also available at time $j+1$
β_j	Probability member of the superpopulation first available for detection at time $j+1$
b, d	Indices of times that an animal is first and last available for detection

2 | DATA AND EXISTING MODELS

The major notation is shown in Table 1 and explained in the following sections.

2.1 | Data

We consider data sets resulting from spatial sampling of an animal population with passive detectors (typically traps, DNA hair snags or cameras) placed at multiple known locations across a study area. Detectors are operated for J primary sessions between which there may be partial turnover of the population. Within each primary session, there may be one or more secondary sessions, sampling intervals closely spaced in time between which there is no turnover. In each secondary session, newly detected animals are marked individually (if not distinguishable by natural marks) and released; redetections are recorded. The spacing of detectors is not critical, but it should allow some animals to be redetected at more than one detector; placement of detectors may vary over time.

This sampling design results in a set of "detection histories" $\boldsymbol{\omega}$ for the animals detected at least once. The history $\boldsymbol{\omega}_i$ of a detected individual i comprises a matrix of observations on particular secondary sessions (rows s) at particular detectors (columns k). The observation ω_{isk} in each cell of the matrix $\boldsymbol{\omega}_i$ may be binary (detected or not) or a count of the number of detections. A trap is a particular type of detector which allows an animal to be detected no more than once per secondary session. The term "capture" is natural when traps are used for sampling, but "detection" is more general.

2.2 | Closed-population SECR model

Data as described may be analyzed in the closed-population SECR paradigm of Borchers and Efford (2008) and Efford *et al*. (2009a) by maximizing an observed data likelihood formed as the product of components, one for each primary

session j. In this section, we focus on a single session j with S_j secondary sessions and n_j non-null histories $\boldsymbol{\omega}_{\cdot j}$. Each session-specific component has two parts, one for the probability of observing n_j individuals in that session, and the other for the probability conditional on n_j of observing the particular set of detection histories, that is, $L(\theta, \nu \mid \boldsymbol{\omega}_{\cdot j}) = \Pr(n_j \mid \theta, \nu)\ \Pr(\boldsymbol{\omega}_{\cdot j} \mid n_j, \theta)$, where θ is a vector of detection parameters and the vector ν parameterizes the intensity of an inhomogeneous Poisson process D. The probability of observing exactly n_j animals may be treated as either binomial or Poisson (Borchers and Efford, 2008).

Individuals are assumed to be detected independently of each other, so the probability of $\boldsymbol{\omega}_{\cdot j}$ may be modeled as

$$\Pr(\boldsymbol{\omega}_{\cdot j} \mid n_j, \theta) \propto \prod_{\omega_{ij} \in\, \boldsymbol{\omega}_{\cdot j}} \int \Pr(\omega_{ij} \mid \omega_{ij} > 0, \mathbf{x})\, f(\mathbf{x} \mid \omega_{ij} > 0) d\mathbf{x}, \quad (1)$$

using $\omega_{ij} > 0$ to indicate a non-null detection history and f for the spatial distribution of individual home-range centers $\mathbf{x}$ (we omit a multinomial constant that does not depend on the parameters). The integration marginalizes over the unknown locations of individuals. In this paper, integration is over the real plane $\mathbb{R}^2$, unless otherwise indicated. The appropriate form for $\Pr(\omega_{ij} \mid \omega_{ij} > 0, \mathbf{x})$ depends on the properties of the detectors (Borchers and Efford, 2008; Efford *et al*., 2009b). For concreteness, we focus on binary proximity detectors, for which

$$\Pr(\omega_{ij} \mid \omega_{ij} > 0, \mathbf{x}) = p_{\cdot}(\mathbf{x})^{-1} \prod_{s=1}^{S_j} \prod_{k=1}^{K} p_{sk}(\mathbf{x})^{\omega_{isk}} \{1 - p_{sk}(\mathbf{x})\}^{1-\omega_{isk}}, \quad (2)$$

with $p_{\cdot}(\mathbf{x}) = 1 - \prod_{s=1}^{S_j} \prod_{k=1}^{K} \{1 - p_{sk}(\mathbf{x})\}$. The dependence of detection on the distance of $\mathbf{x}$ from each detector k is captured by the function p_{sk} that may take various forms. For example, the cumulative hazard of detection may be modeled as a half-normal function of distance $h(d) = h_0 \exp\{-d^2/(2\sigma^2)\}$; then $p_{sk}(\mathbf{x}) = 1 - \exp\{-h(\|\mathbf{x} - \mathbf{k}\|)\}$ where $\mathbf{k}$ is the location of detector k. The distribution of detected animals is given by $f(\mathbf{x} \mid \omega_{ij} > 0) = D(\mathbf{x}) p_{\cdot}(\mathbf{x}) / \int D(\mathbf{x}) p_{\cdot}(\mathbf{x}) d\mathbf{x}$.

A multi-session closed-population approach discards information from the recaptures of individuals between primary sessions. When some animals persist between primary sessions, the assumption of independence is violated, potentially leading to underestimation of sampling variance. This is a further reason for using an open population model, even when population turnover (loss and recruitment between primary sessions) is not of interest in itself.

2.3 | Open-population non-spatial model (JSSA)

The same data may be analyzed with the JSSA model if we discard location information and reduce each detection history to a binary vector for each individual. The frequencies of the reduced capture histories $\boldsymbol{\omega}'$ are modeled as arising from a multinomial distribution with one cell for each secondary session. Cell probabilities combine a demographic process (determining the primary sessions in which each animal is alive and available for capture) and a detection process (probability of capture given alive).

The JSSA observed data likelihood has broadly the same structure as the SECR likelihood: $L(\theta, N \mid \boldsymbol{\omega}') = \Pr(n \mid \theta, N) \Pr(\boldsymbol{\omega}' \mid n, \theta)$, where N is the superpopulation size, defined as the number of animals available for capture in at least one primary session. Breaking this down, $\Pr(n \mid \theta, N) = \binom{N}{n} p^n (1-p)^{N-n}$ for binomial n, where p is the probability a member of the superpopulation is detected at least once, and $\Pr(\boldsymbol{\omega}' \mid n, \theta) \propto \prod_{\omega_i' \in \boldsymbol{\omega}'} \Pr(\boldsymbol{\omega}_i' \mid \boldsymbol{\omega}_i' > 0)$.

An individual-based formulation of the JSSA likelihood (Schofield and Barker, 2008; Pledger *et al*., 2010) uses a summation over the birth and death times consistent with each detection history. The probability that an animal is first available for detection at b and last available for detection at d is $\Pr(b, d) = \beta_{b-1}\left(\prod_{j=b}^{d-1} \phi_j\right)(1 - \phi_d)$, where β_j is the probability that a member of the notional superpopulation is first available for detection at session $j+1$, and ϕ_j is the conditional probability that it is last available at session j. For individual i first observed in session f_i and last observed in session l_i,

$$\begin{aligned} \Pr(\boldsymbol{\omega}_i' \mid \boldsymbol{\omega}_i' > 0) &= \sum_{b=1}^{f_i} \sum_{d=l_i}^{J} \left\{ \Pr(b, d) \Pr(\boldsymbol{\omega}_i' \mid b, d, \boldsymbol{\omega}_i' > 0) \right\} \\ &= p_{\cdot}^{-1} \sum_{b=1}^{f_i} \sum_{d=l_i}^{J} \left\{ \Pr(b, d) \prod_{j=b}^{d} p_j^{\omega_{ij}'} (1 - p_j)^{1-\omega_{ij}'} \right\}, \end{aligned} \quad (3)$$

where p_j is the probability of detection if available for detection in session j and $p_{\cdot} = 1 - \sum_{b=1}^{J} \sum_{d=b}^{J} \{\Pr(b, d) \prod_{j=b}^{d} (1 - p_j)\}$. For clarity, we follow Pledger *et al*. (2010) in presenting the model for a single sample in each primary session; the extension to multiple secondary sessions is straightforward.

3 | SPATIOTEMPORAL MODEL

Two steps are required to adapt the JSSA model for spatial data $\boldsymbol{\omega}$. First, we conceptualize the superpopulation as a

spatial point process, in which animals are distributed independently in the plane with density $D(\mathbf{x})$. Second, we substitute a spatial detection model, such as (2) for the non-spatial detection model (3). We start here by specifying the model for stationary home-range centers, and then adapt it for between-session movement governed by a probabilistic kernel.

The spatiotemporal likelihood has the same general form as the SECR model, $L(\theta, \nu|\boldsymbol{\omega}) = \Pr(n|\theta, \nu)\Pr(\boldsymbol{\omega}|n, \theta)$, but the vector θ includes parameters for turnover (survival and recruitment), and $D(\mathbf{x})$ refers to the superpopulation rather than a time-specific density. The conditional likelihood

$$L_\omega(\boldsymbol{\omega}|n) \propto \prod\nolimits_{\omega_i \in \omega} \sum_{b=1}^{f_i} \sum_{d=l_i}^{J} \Pr(b, d)\Pr(\boldsymbol{\omega}_i|b, d, \omega_i > 0)$$

has the same form as in the non-spatial JSSA model.

3.1 | Stationary home ranges

We assume initially that each animal is fixed at its unknown initial location $\mathbf{x}$. We marginalize over $\mathbf{x}$ using

$$\Pr(\boldsymbol{\omega}_i|b, d, \omega_i > 0) = \int \Pr(\boldsymbol{\omega}_i|b, d, \omega_i > 0, \mathbf{x})\, f(\mathbf{x}|b, d, \omega_i > 0)d\mathbf{x}.$$

The probability that an animal centered at $\mathbf{x}$ and available from b to d is detected at least once is given by $p_\cdot(\mathbf{x}|b, d) = 1 - \prod_{s=S_f(b)}^{S_l(d)} \prod_{k=1}^{K} \{1 - p_{sk}(\mathbf{x})\}$ where s indexes the secondary sessions that an animal was available for detection, and $S_f(j)$ and $S_l(j)$ refer, respectively, to the first and last secondary sessions in primary session j. The spatial distribution of detected animals is given by $f(\mathbf{x}|b, d, \omega_i > 0) = p_\cdot(\mathbf{x}|b, d)D(\mathbf{x}) / \int p_\cdot(\mathbf{x}|b, d)D(\mathbf{x})d\mathbf{x}$, and hence

$$\Pr(\boldsymbol{\omega}_i|b, d, \omega_i > 0) = \frac{\int \Pr(\boldsymbol{\omega}_i|b, d, \mathbf{x})D(\mathbf{x})d\mathbf{x}}{\int p_\cdot(\mathbf{x}|b, d)D(\mathbf{x})d\mathbf{x}}. \quad (4)$$

A Poisson model for the number of detected animals n has the rate parameter

$$\lambda(\theta, \nu) = \sum_{b=1}^{J} \sum_{d=b}^{J} \Pr(b, d) \int p_\cdot(\mathbf{x}|b, d)D(\mathbf{x})d\mathbf{x}. \quad (5)$$

3.2 | Movement between primary sessions

Dispersal, defined as a shift in the home-range center between primary sessions, is a potentially important demographic process alongside in situ recruitment and mortality. Dispersal is conveniently modeled with a two-dimensional kernel centered on the origin. If the kernel is radially symmetrical, then the probability density of a movement from $\mathbf{x}_{j-1}$ to $\mathbf{x}_j$ is a function of distance r. It is tempting to rely on the Gaussian (bivariate normal) kernel $k(r) = (1/2\pi\alpha_g^2)\exp(-r^2/2\alpha_g^2)$ with scale parameter α_g, but dispersal kernels in biology are commonly fat-tailed (Nathan *et al*., 2012). A simple alternative with fatter tails is the Laplace (negative exponential) kernel $k(r) = (1/2\pi\alpha_l^2)\exp(-r/\alpha_l)$ with the scale parameter α_l. The distribution of location in session $j-1$ is projected forwards to session j by convolving the initial distribution with the kernel k

$$\begin{aligned} f(\mathbf{x}_j|\mathbf{x}_{j-1}) &= f(\mathbf{x}_{j-1})^* k \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} k(u, v) f(x_{j-1} - u, y_{j-1} - v)du\, dv. \end{aligned}$$

In a model with movement, the numerator of (4) entails multiple integrations over the unknown location of animal i at each sampling occasion. Thus

$$\begin{aligned} &\int \Pr(\boldsymbol{\omega}_i|b, d, \mathbf{x})D(\mathbf{x})d\mathbf{x} \\ &= \int \cdots \int \Pr(\boldsymbol{\omega}_{i,b}|\mathbf{x}_b)D(\mathbf{x}_b) \\ &\quad \times \prod_{j=b+1}^{d} \Pr(\boldsymbol{\omega}_{i,j}|\mathbf{x}_j)\, f(\mathbf{x}_j|\mathbf{x}_{j-1})d\mathbf{x}_b \ldots \mathbf{x}_d. \end{aligned} \quad (6)$$

For the movement model, the denominator of (4) is the complement of (6) evaluated for a null history.

4 | IMPLEMENTATION

4.1 | Discretization

In the development, we considered animal locations across the entire plane ($\mathbb{R}^2$), but in applications the integration is replaced by summation over discrete pixels in a finite area A. This is computationally convenient and allows the inclusion of biological detail by, for example, dropping nonhabitat pixels from the summation. In a continuous habitat, the extent of A is arbitrary, although using a region that is too small relative to the spatial scale of detection results in a positive bias in estimated density. Models with dispersal between primary sessions admit the possible immigration of more distant animals and require that a larger region is considered if edge effects are to be avoided.

A further benefit of discretization is that an approximation to the multiple integral (6) may be

evaluated by the forward algorithm commonly used for hidden Markov models (Zucchini and MacDonald, 2009)

$$\begin{aligned}\int \Pr(\boldsymbol{\omega}_i|b,d,\mathbf{x})\,D(\mathbf{x})\,d\mathbf{x} \approx{}& \sum_{\mathbf{x}_d} f(\mathbf{x}_d|\mathbf{x}_{d-1})\Pr(\boldsymbol{\omega}_{i,d}|\mathbf{x}_d)\\ &\times \sum_{\mathbf{x}_{d-1}} f(\mathbf{x}_{d-1}|\mathbf{x}_{d-2})\Pr(\boldsymbol{\omega}_{i,d-1}|\mathbf{x}_{d-1})\ldots\\ &\times \sum_{\mathbf{x}_{b+1}} f(\mathbf{x}_{b+1}|\mathbf{x}_b)\Pr(\boldsymbol{\omega}_{i,b+1}|\mathbf{x}_{b+1})\\ &\times \sum_{\mathbf{x}_b} D(\mathbf{x}_b)\Pr(\boldsymbol{\omega}_{i,b}|\mathbf{x}_b),\end{aligned}$$

evaluated from right to left.

To limit computation, we used a discrete movement kernel with the same pixel size as the habitat and truncated the kernel at an arbitrary radius. Cell probabilities were calculated from the continuous model (eg, bivariate normal) evaluated at the centroid of each pixel and normalized across the kernel.

4.2 | Parameterization of recruitment

Recruitment in the multinomial JSSA model is naturally parameterized in terms of the entry probabilities β_j (Schwarz and Arnason, 1996). However, the β_j lack a direct biological interpretation, as they depend on the superpopulation size N or density $D(\mathbf{x})$, which increase with study duration. More useful alternatives are per capita recruitment f_j (Pradel, 1996; Link and Barker, 2005) and the finite rate of population increase $\lambda_j = f_j + \phi_j$ (Pradel, 1996; Schwarz and Arnason, 1996; Link and Barker, 2005; Pledger *et al*., 2010), and others are possible, including time-specific density D_j (Table 2). The parameter f_j may be computed recursively from the β_j and ϕ_j using the intermediate quantity $d_{j+1} = d_j\phi_j + \beta_j$ where $d_1 \equiv \beta_0$ and $f_j = \beta_j/d_j$ (Link and Barker, 2005). Conversely, the model may be expressed in terms of the alternative recruitment parameter, with the β_j computed by the inverse recursion for use in $\Pr(b,d)$.

Models that relate f_j or λ_j to covariates or apply a temporal constraint (eg, constancy over time) have a direct biological interpretation, unlike simple models for β_j. Parameterizations of recruitment for the non-spatial model transfer directly to the spatial model except that the state variable density replaces population size (Table 2). We encountered fewer numerical problems when using the f_j or β_j parameterizations than with λ_j or D_j.

TABLE 2 Parameterizations of recruitment in JSSA models, assuming no losses on capture

β_j	Entry probability
$d_{j+1} = d_j\phi_j + \beta_j$; $d_1 \equiv \beta_0$	Used in following recursions
$f_j = \beta_j/d_j$	Per capita recruitment
$\lambda_j = \phi_j + f_j$	Finite population growth rate
$\gamma_j = \phi_j/\lambda_j$	Seniority (Pradel, 1996)
κ_j	Link and Barker (2005, p. 48)
$D_j = d_j D$	Time-specific density
$B_j = D_{j+1} - D_j\phi_j$	Density of births

Note: ϕ_j is the *per capita* survival from time j to time $j+1$. D_j and B_j are specific to homogeneous spatial models (superpopulation density D).

4.3 | Variable sampling interval

The interval between primary sessions may vary in duration. It then aids interpretation and modeling to express demographic rates in relation to a constant unit of time, or as instantaneous rates (Ergon *et al*., 2018). The realized session-specific rate is usually a nonlinear function of the underlying standardized rate. For survival, there is a straightforward power relationship: $\phi'_j = \phi_j^{\Delta t_j}$ where Δt_j is the duration from the start of session j to the start of session $j+1$. The likelihood is evaluated using ϕ'_j in place of ϕ_j. The corresponding adjustment for recruitment under the various parameterizations in Table 2 is clear only in the case of a finite population growth rate ($\lambda'_j = \lambda_j^{\Delta t_j} = (\phi_j + f_j)^{\Delta t_j}$). For the per capita growth rate parameterization we suggest $f'_j = \lambda'_j - \phi'_j$.

Adjustment of movement parameters for a varying interval is unproblematic in the case of a continuous Gaussian kernel because the kernel remains Gaussian under convolution, with variance proportional to elapsed time $\left(\alpha'_{g,j} = \sqrt{\alpha^2_{g,j}\Delta t_j}\right)$. However, this is only an approximation when the kernel is discretized (eg, Slone, 2011; Chipperfield *et al*., 2013) and may not apply to other shapes of kernel.

4.4 | Conditioning on n

The detection and turnover parameters of non-spatial JSSA models may be estimated by maximizing the likelihood conditional on n (Schwarz and Arnason, 1996), and the parameterizations of Pradel (1996) and Link and Barker (2005) use only the conditional likelihood. Schofield and Barker (2016) showed that the conditional estimates of survival and recruitment are identical to the full estimates when n is Poisson because n is then ancillary. There is no name in general use for the survival-and-recruitment model conditional on n. We suggest Pradel-Link-Barker or "PLB" for the non-spatial model, recognizing the contributions of Pradel (1996) and

Link and Barker (2005), and "spatial PLB" for its spatial analog.

Many properties of the spatial analogs of the PLB and JSSA models derive directly from their non-spatial equivalents (eg, alternative parameterizations of recruitment). The spatial PLB formulation may be sufficient if the research question concerns only demographic rates. In our experience, spatial PLB models are easier to fit than spatial JSSA models, and estimates from a spatial PLB fit are good starting values for maximizing the full likelihood. Density may be estimated from a spatial PLB fit as a derived parameter using the Horvitz-Thompson estimator (Borchers and Efford, 2008).

4.5 | Software

The R (R Core Team, 2018) package openCR (Efford, 2019) implements PLB and JSSA models in various parameterizations with the spatial extensions described here.

5 | EXAMPLE

Brushtail possums *Trichosurus vulpecula* are 2–4 kg arboreal marsupials introduced to New Zealand, where their consumption of native biota is a significant conservation problem. A long-term study of possum dynamics was undertaken in native forest of the Orongorongo Valley near Wellington. Females bear a single young about May and carry it in their pouch until late in the calendar year. Traps were set across about 14 ha of forest at 30-m spacing for five consecutive nights three times each year at significant points in the annual cycle (February, when young of the previous year are first caught independently of their mothers, June, after the birth pulse, and September, before any young leave the pouch). Independent animals were ear marked and released.

We analyzed a 2-year subset of these data from 1996–1997. Details of the data and analysis are in Web Appendix A. The spatial model fitted better when allowance was made for home-range shifts between primary sessions; the reduction in deviance was greater for the Laplace kernel (Δdeviance = 229) than the Gaussian kernel (Δdeviance = 117). Models that allowed for between-session movement gave higher estimates of the survival rate than the static model, although the effect was strong only in early sessions (Table 3). The estimated scale of between-session movement was small in this example (bivariate normal $\widehat{\alpha}_g \leq 20$ m) and allowing for such movement had only a small effect on the estimated scale of detection, reducing $\widehat{\sigma}$ from 31–40 m to 27–38 m. Despite the strong difference in fit between the Gaussian and Laplace movement kernels, estimates of shared parameters (ϕ, D, λ_0, and σ) were almost identical.

6 | SIMULATIONS

6.1 | Assessment of bias in survival estimates from spatial sampling

Spatial sampling with an array of fixed detectors in a more extensive population inevitably exposes some individuals to greater risk of detection than others. This source of individual heterogeneity causes bias in $\widehat{\phi}$ from non-spatial models; we expect the effect to be absent in spatial models. Simulations were conducted to compare survival estimates from constant-parameter spatial and non-spatial PLB models.

The simulation scenario comprised a square grid of 100 traps at spacing σ; 200 activity centers were placed at random in a square region extending distance w beyond the traps in each cardinal direction, where w was varied from 0 to 5σ in increments of 0.5σ. True $\phi = 0.8$ and $\lambda = 1.0$; centers were stationary. Trapping was simulated for five primary sessions, each comprising five secondary sessions; 100 replicate simulations were performed for each level of w. Detection on each secondary occasion was governed by a half-normal hazard function $h(r) = 0.5 \exp\{-r^2/(2\sigma^2)\}$.

The non-spatial PLB estimator of ϕ was strongly biased for realistic levels of heterogeneity that arise when the target population extends well beyond the detectors ($w > 2\sigma$) (Figure 1). Spatial PLB estimates of ϕ were largely immune to this source of bias.

6.2 | Misspecification of the movement model

Simulations were also performed to evaluate the robustness of estimates to misspecification of the movement kernel. We generated data using a continuous Laplace kernel with scale $\alpha_l = \sigma$, and fitted static, Gaussian and Laplace movement models. Fitted kernels were discretized and truncated at $3\alpha_l$ or $6\alpha_l$ (see Web Appendix D for details). Results are summarized in Figure 2. Fitting a static model led to strong bias in all directly estimated parameters ($\phi, f, \sigma, \lambda_0$), echoing a result of Glennie *et al.* (2019). However, estimates of the derived parameter λ appeared essentially unbiased (negative bias in $\widehat{\phi}$ was balanced by positive bias in $\widehat{f}$).

Modeling movement between primary sessions almost eliminated biases in $\widehat{\phi}$ and $\widehat{f}$, even when the movement model was misspecified or used a small, truncated kernel. Misspecification of the movement kernel (Gaussian vs Laplace) had little effect on estimates of the detection parameters σ and λ_0, but estimates of these parameters were biased when the kernel was heavily truncated (radius $3\alpha_l$). Estimates of the movement parameter from the generating model were nearly unbiased when the

TABLE 3 Non-spatial and spatial Jolly-Seber-Schwarz-Arnason analyses of robust-design brushtail possum data set

		j	Non-spatial	Static	Normal	Exponential
(a) Apparent survival $\widehat{\phi}$						
1996	Feb	1	0.61 (0.50, 0.71)	0.62 (0.50, 0.72)	0.68 (0.58, 0.77)	0.67 (0.57, 0.76)
	Jun	2	0.31 (0.22, 0.42)	0.32 (0.23, 0.43)	0.37 (0.27, 0.48)	0.37 (0.27, 0.48)
	Sep	3	0.73 (0.62, 0.83)	0.77 (0.65, 0.86)	0.79 (0.68, 0.87)	0.79 (0.68, 0.87)
1997	Feb	4	0.65 (0.52, 0.77)	0.82 (0.67, 0.91)	0.82 (0.69, 0.91)	0.82 (0.69, 0.91)
	Jun	5	0.46 (0.32, 0.60)	0.65 (0.46, 0.80)	0.65 (0.46, 0.80)	0.65 (0.46, 0.81)
	Sep	6	...	...	...	...
(b) Density $\widehat{D}$ (ha^{-1})						
1996	Feb	1	...	13.4 (11.7, 15.2)	12.1 (10.6, 13.7)	12.2 (10.7, 13.9)
	Jun	2	...	12.2 (10.8, 13.9)	11.4 (10.1, 13.0)	11.5 (10.2, 13.0)
	Sep	3	...	9.4 (8.2, 10.8)	9.1 (7.9, 10.4)	9.2 (8.0, 10.5)
1997	Feb	4	...	9.1 (7.9, 10.5)	8.9 (7.8, 10.2)	9.0 (7.9, 10.3)
	Jun	5	...	8.7 (7.5, 10.0)	8.5 (7.4, 9.7)	8.6 (7.4, 9.8)
	Sep	6	...	7.9 (6.7, 9.3)	7.7 (6.6, 9.0)	7.8 (6.7, 9.1)
(c) Baseline detection hazard $\widehat{\lambda_0}$						
1996	Feb	1	...	0.07 (0.06, 0.08)	0.10 (0.09, 0.12)	0.10 (0.09, 0.12)
	Jun	2	...	0.10 (0.08, 0.11)	0.11 (0.09, 0.12)	0.11 (0.10, 0.13)
	Sep	3	...	0.06 (0.05, 0.07)	0.06 (0.05, 0.07)	0.07 (0.06, 0.09)
1997	Feb	4	...	0.07 (0.06, 0.09)	0.08 (0.07, 0.10)	0.09 (0.07, 0.10)
	Jun	5	...	0.08 (0.07, 0.09)	0.09 (0.08, 0.11)	0.10 (0.09, 0.12)
	Sep	6	...	0.15 (0.13, 0.18)	0.19 (0.16, 0.23)	0.20 (0.17, 0.24)
(d) Detection scale $\widehat{\sigma}$ (m)						
1996	Feb	1	...	34.5 (32.5, 36.7)	27.9 (26.1, 29.9)	27.4 (25.6, 29.3)
	Jun	2	...	32.9 (31.1, 34.8)	31.0 (29.2, 32.8)	30.0 (28.3, 31.9)
	Sep	3	...	39.4 (36.9, 42.0)	38.1 (35.5, 40.8)	35.1 (32.5, 37.9)
1997	Feb	4	...	40.2 (37.9, 42.7)	37.5 (35.1, 40.1)	36.7 (34.4, 39.2)
	Jun	5	...	39.2 (36.9, 41.7)	35.6 (33.1, 38.2)	34.3 (32.1, 36.8)
	Sep	6	...	31.1 (29.3, 33.1)	27.1 (25.4, 29.0)	26.5 (24.9, 28.2)
(e) Movement kernel scale (m) $\widehat{\alpha}_g$, $\widehat{\alpha}_l$						
1996	Feb	1	...	...	19.7 (16.4, 23.7)	10.3 (8.4, 12.8)
	Jun	2	...	...	4.2 (1.7, 10.0)	8.2 (6.1, 10.8)
	Sep	3	...	...	19.0 (15.3, 23.6)	10.3 (8.1, 13.0)
1997	Feb	4	...	...	10.4 (4.7, 22.8)	8.2 (6.1, 11.0)
	Jun	5	...	...	15.9 (12.4, 20.5)	7.8 (6.2, 9.9)
	Sep	6	...	...	...	...

Note: Annual rates of apparent survival ϕ, density D, baseline detection hazard λ_0, and two spatial scale parameters σ and α_g (95% confidence interval in parentheses). In the three spatial models, possums either retain the same activity center between sessions ("static") or shift randomly according to a Gaussian ("normal") or Laplace ("exponential") kernel.

fitted kernel was large enough and of the correct shape (discretized Laplace kernel radius $6\alpha_l$).

7 | DISCUSSION

The spatial open-population model relaxes some assumptions of non-spatial open-population and spatial closed-population models (Table 4). This enables parameters of interest (population growth rate λ, survival rate ϕ and density D) to be estimated with less bias and better confidence interval coverage (Table 4). Web Appendix C illustrates the benefits from open-population spatial modeling when there is population turnover (nonclosure) or variation in the area sampled. Varying proximity to detectors is a major reason for the violation of the fourth assumption in Table 4 (homogeneity of detection probability) and this component of heterogeneity is included in spatial models, greatly reducing bias in PLB survival estimates (Figure 1). Against these benefits must be weighed the extra complexity of spatial open population models, which can lead to numerical difficulties, and some limitations in the present models that are discussed below (see also Gardner *et al.*, 2018).

The present model is in most respects the same as that developed independently by Glennie *et al.* (2019). This fact is masked by their use of a hidden Markov formulation, whereas we emphasize continuity with

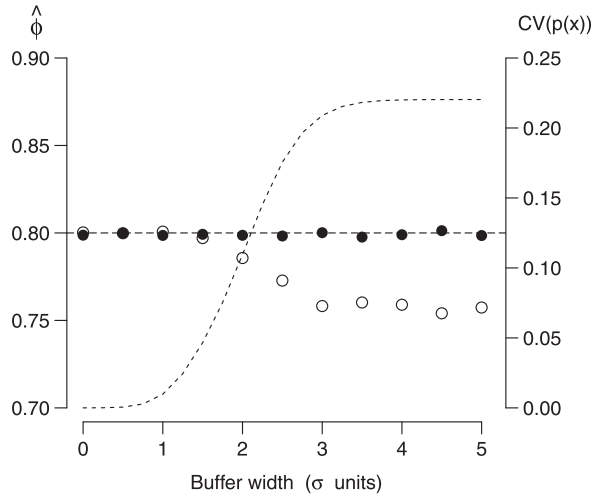

FIGURE 1 Survival rate ϕ estimated by fitting non-spatial (○) and spatial (●) PLB models to simulated data from populations confined to varying buffered areas about a trap array and therefore differing in heterogeneity of detection probability. Average of 100 replicates. Horizontal line indicates true ϕ. Dashed curve shows increase in heterogeneity of detection probability per primary session ($\mathrm{CV}(p(\mathbf{x}))$) with increasing buffer width (weighted by probability of detection; see Web Appendix B for details).
PLB, Pradel-Link-Barker

previous closed-spatial and open-non-spatial models. If activity centers are assumed stationary the models yield the same estimates. The only substantive differences are in the implementation of the movement submodel; that of Glennie *et al.* (2019) relies on a Gaussian movement kernel, whereas our model allows kernels of arbitrary shape up to the radius of truncation. Their Gaussian implementation is elegant and efficient, but other kernel shapes may fit better, as in our possum example, and may be required for biological accuracy.

Many extensions to the present model are possible. Individual heterogeneity may be accommodated with finite mixtures (Pledger *et al.*, 2010). If temporary unavailability is suspected, then the model might be extended to estimate within-primary-session transitions between the available and unavailable states (Kendall and Bjorkland, 2001). We consider inhomogeneous recruitment below. We emphasize two insights from the literature on non-spatial open-population models that apply also to spatial models. First, various parameterizations of recruitment yield the same likelihood, the choice being a matter of modeling convenience. Second, survival, recruitment, and movement may be estimated from a model that

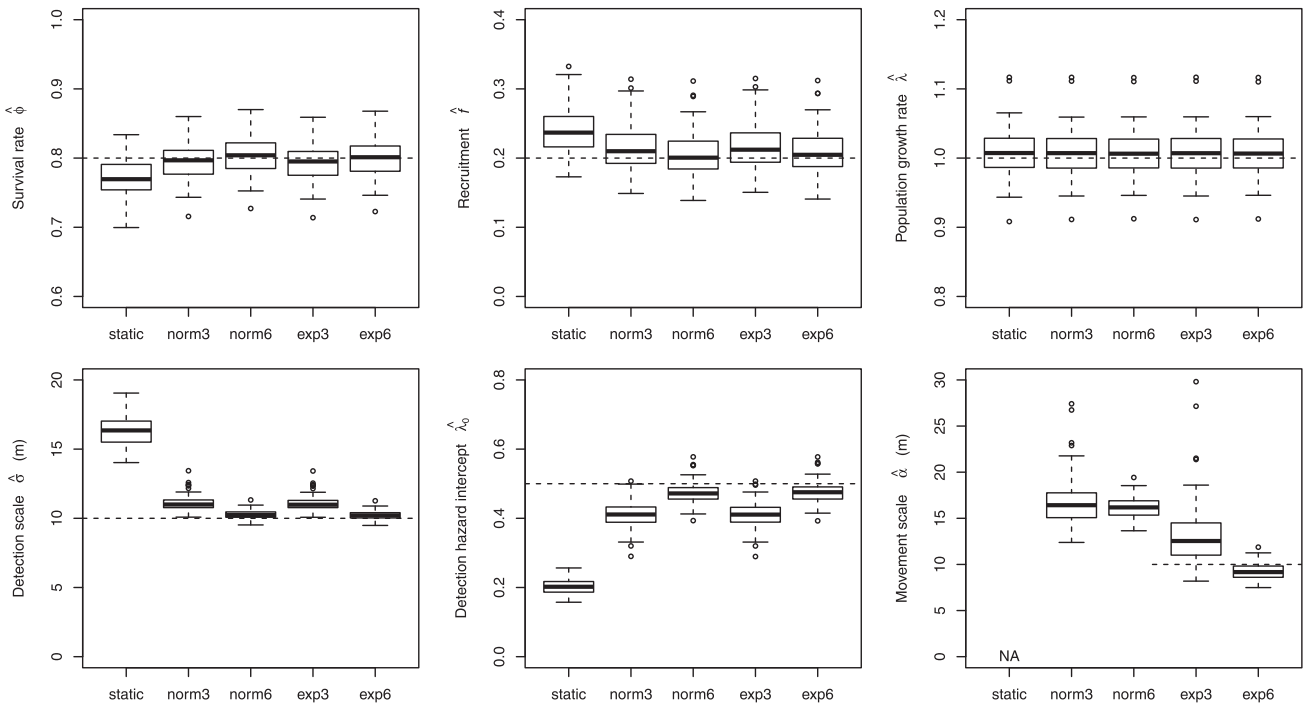

FIGURE 2 Distribution of estimates from data simulated with a continuous Laplace (negative exponential) movement kernel. Each panel presents the estimates for one parameter from a static model and four-movement models using discrete kernels with radius $3\alpha_l$ or $6\alpha_l$ and either Gaussian ("norm3," "norm6") or Laplace ("exp3," "exp6") shape. Dashed lines indicate the true value of each parameter. The static and Gaussian models are misspecified. Parameter α refers to α_g or α_l as appropriate to model (100 replicates). NA, not estimated

TABLE 4 Consequences of assumption failure when compromise models are applied to spatial data from open populations.

Model class	Assumption	Consequence if violated
Spatial, closed	Demographic closure Independence of sessions	Positive bias in $\widehat{D}$ Poor CI coverage $\widehat{\lambda}$
Non-spatial, open	Constant study area Homogeneous detection probability	Spurious trend $\widehat{N}$, bias $\widehat{\lambda}$ Negative bias in $\widehat{\phi}$ and $\widehat{N}$

conditions on the number of individuals observed (the PLB formulation) without the need to estimate the density of the superpopulation.

7.1 | Separating mortality and emigration

A further motivation for spatial capture-recapture models has been to distinguish between in situ mortality and emigration from the study area. A parametric model of movement between primary sessions, fitted using recaptures within the study area, potentially predicts the rate at which marked animals leave the study area, leading to an unbiased estimate of in situ survival (Ergon and Gardner, 2014; Schaub and Royle, 2014). Conversely, failure to allow for between-session movement in spatial capture-recapture models can result in biased estimation of all parameters (Glennie *et al*., 2019).

Our model, like that of Glennie *et al*. (2019), provides the means to separate movement from other demographic processes in open population capture-recapture studies. The operational robustness of the models has yet to be fully tested. On the positive side, our estimates of brushtail possum survival and population density were robust to the choice of dispersal kernel, and simulations with misspecified kernels (Figure 2, Web Appendix D) supported this result, even when the scale of movement was poorly estimated.

However, there are reasons for caution. The results are potentially sensitive to the dispersal kernel, particularly the weight in the tail (Fujiwara *et al*., 2006; Schaub and Royle, 2014), and dispersal distances may greatly exceed the spatial scale of sampling, leaving the tail undocumented. Capture-recapture on local detector arrays is generally not well-suited to quantifying dispersal movements (for counter examples see Zeng and Brown, 1987; Gilroy *et al*., 2012; Taylor *et al*., 2015). We may expect the scale of movement to be estimated less well than other demographic parameters (survival and recruitment), but it appears that estimates of survival and recruitment benefit from the inclusion of movement scale as a nuisance parameter, even when it is poorly estimated (Figure 2).

7.2 | Spatial CJS model

The widely used survival-only Cormack-Jolly-Seber (CJS) model results from conditioning on times of first detection and not modeling first detection (eg, Link and Barker, 2005). A direct spatial analog of the CJS model is challenging to define; a model for the latent initial locations of detected animals is needed to model detection and nondetection on subsequent occasions, yet the CJS model discards initial detections. If the distribution of detected animals is naively assumed to be uniform, then estimates of ϕ will be biased, as demonstrated in Web Appendix E. The problem does not arise when initial locations are observed (Schaub and Royle, 2014). The alternative sometimes adopted in spatial models nominally derived from CJS is to model initial detections, but that is outside the CJS formulation (Web Appendix E). The spatial PLB model conditions on n rather than times of the first detection, and is both widely applicable and more informative than spatial CJS because it also estimates recruitment.

7.3 | Density surfaces and non-Poisson recruitment

We have presented a spatial model in which turnover (survival, recruitment, and movement) is independent of location. The extensions needed to allow dependence on location is simple in principle. However, it may be difficult to articulate these with a spatial density model. Spatially inhomogeneous birth and death processes can be expected to result in a spatially inhomogeneous equilibrium distribution of animals; density-independent dispersal tends to flatten the distribution. A plausible inhomogeneous model could project forward an initial density surface under spatially inhomogeneous birth, death, and dispersal processes. The spatial distribution of the superpopulation under a particular sampling scheme would then bear only an indirect relation to the time-specific or equilibrium distributions, which would be obtained as derived attributes of the fitted model. In the context of open-population modeling, it may be desirable, both mathematically and

biologically, to see the parameters of demographic turnover as primary, and the pattern of density as secondary and derived.

7.4 | Parameter identifiability, model selection, and overdispersion

Parameter redundancy and non-identifiability is a significant headache for the application of classical open population models (eg, Catchpole and Morgan, 1997). This applies also with spatial open population models, with the added caveat that the scale of spatial sampling should be appropriate (great enough to span some individual activity distributions, fine enough to allow recaptures of an animal at multiple points). In the non-spatial JSSA model, there is no opportunity for recaptures within a secondary session, and hence some session-specific detection probabilities are unidentifiable or confounded with survival. This is true also of spatial models with one secondary session per primary session when the detectors are traps. However, session-specific detection parameters may be estimated, when detectors allow multiple detections per secondary session, as with binary and count proximity detectors (Efford *et al*., 2009b).

The likelihood framework facilitates information-theoretic model selection by criteria such as AIC. Overdispersion is likely to be an issue when animals live in groups, suggesting a role for quasi-likelihood adjustment using an empirical estimate of overdispersion.

ACKNOWLEDGMENTS

We thank Richard Glennie for providing additional detail on Glennie *et al*. (2019). We thank Torbjørn Ergon, two anonymous reviewers, and the associate editor for comments that substantially improved the paper.

ORCID

Murray G. Efford http://orcid.org/0000-0001-5231-5184
Matthew R. Schofield http://orcid.org/0000-0003-1481-2766

REFERENCES

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

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

Catchpole, E.A. and Morgan, B.J.T. (1997) Detecting parameter redundancy. *Biometrika*, 84, 187–196.

Chandler, R.B. and Clark, J.D. (2014) Spatially explicit integrated population models. *Methods in Ecology and Evolution*, 5, 1351–1360.

Chipperfield, J.D., Holland, E.P., Dytham, C., Thomas, C.D. and Hovestadt, T. (2013) On the approximation of continuous dispersal kernels in discrete-space models. *Methods in Ecology and Evolution*, 2, 668–681.

Crosbie, S.F. and Manly, B.F.J. (1985) Parsimonious modelling of capture-mark-recapture studies. *Biometrics*, 41, 385–398.

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

Efford, M.G. (2019) openCR: Open population capture-recapture models. R package version 1.4.1. Available at: https://CRAN.R-project.org/package=openCR

Efford, M.G., Borchers, D.L. and Byrom, A.E. (2009a) Density estimation by spatially explicit capture-recapture: likelihood-based methods. In: Thomson, D.L., Cooch, E.G. and Conroy, M.J. (Eds.) *Modeling Demographic Processes in Marked Populations*. New York: Springer, pp. 255–269.

Efford, M.G., Dawson, D.K. and Borchers, D.L. (2009b) Population density estimated from locations of animals on a passive detector array. *Ecology*, 90, 2676–2682.

Ergon, T., Borgon, Ø., Nater, C.R. and Vindenes, Y. (2018) The utility of mortality hazard rates in population analyses. *Methods in Ecology and Evolution*, 9, 2046–2056.

Ergon, T. and Gardner, B. (2014) Separating mortality and emigration: modelling space use, dispersal and survival with robust-design spatial capture-recapture data. *Methods in Ecology and Evolution*, 5, 1327–1336.

Fujiwara, M., Anderson, K.E., Neubert, M.G. and Caswell, H. (2006) On the estimation of dispersal kernels from individual mark-recapture data. *Environmental and Ecological Statistics*, 13, 183–197.

Gardner, B.J., Reppucci, 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.

Gardner, B., Sollmann, R., Kumar, N.S., Jathanna, D. and Karanth, K.U. (2018) State space and movement specification in open population spatial capture-recapture models. *Ecology and Evolution*, 8, 10336–10344.

Gilroy, J.J., Virzi, T., Boulton, R.L. and Lockwood, J.L. (2012) A new approach to the "apparent survival" problem: estimating true survival rates from mark-recapture studies. *Ecology*, 93, 1509–1516.

Glennie, R., Borchers, D.L., Murchie, M., Harmsen, B.J. and Foster, R.J. (2019) Open population maximum likelihood spatial capture-recapture. *Biometrics*. https://doi.org/10.1111/biom.13078

Hines, J.E. and Nichols, J.D. (2002) Investigations of potential bias in the estimation of λ using Pradel's (1996) model for capture-recapture data. *Journal of Applied Statistics*, 29, 573–587.

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

Kendall, W.L. and Bjorkland, R. (2001) Using open robust design models to estimate temporary emigration from capture-recapture data. *Biometrics*, 57, 1113–1122.

Lebreton, J.D., Burnham, K.P., Clobert, J. and Anderson, D.R. (1992) Modeling survival and testing biological hypotheses using marked animals—a unified approach with case-studies. *Ecological Monographs*, 62, 67–118.

Link, W.A. and Barker, R.J. (2005) Modeling association among demographic parameters in analysis of open-population capture-recapture data. *Biometrics*, 61, 46–54.

Nathan, R., Klein, E., Robledo-Arnuncio, J.J. and Revilla, E. (2012) Dispersal kernels: review. In: Clobert, J. (Ed.) *Dispersal Ecology and Evolution*. Oxford: Oxford University Press, pp. 187–210.

Pledger, S., Pollock, K.H. and Norris, J.L. (2010) Open capture-recapture models withheterogeneity: II. Jolly-Seber model. *Biometrics*, 66, 883–890.

Pollock, K.H. (1982) A capture-recapture design robust to unequal probability of capture. *Journal of Wildlife Management*, 46, 752–757.

Pradel, R. (1996) Utilization of capture-mark-recapture for the study of recruitment and population growthrate. *Biometrics*, 52, 703–709.

R Core Team. (2018) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raabe, J.K., Gardner, B. and Hightower, J.E. (2014) A spatial capture-recapture model to estimate fish survival and location from linear continuous monitoring arrays. *Canadian Journal of Fisheries and Aquatic Sciences*, 71, 120–130.

Royle, J.A., Chandler, R.B., Sollmann, R. and Gardner, B. (2014) *Spatial capture-recapture*. Waltham, MA: Academic Press.

Schaub, M. and Royle, J.A. (2014) Estimating true instead of apparent survival using spatial Cormack-Jolly-Seber models. *Methods in Ecology and Evolution*, 5, 1316–1326.

Schofield, M.R. and Barker, R.J. (2008) A unified capture-recapture framework. *Journal of Agricultural, Biological, and Environmental Statistics*, 13, 458–477.

Schofield, M. and Barker, R. (2016) 50-year-old curiosities: ancillarity and inference in capture-recapture models. *Statistical Science*, 31, 161–174.

Schwarz, C.J. and Arnason, A.N. (1996) A general methodology for the analysis of capture-recapture experiments in open populations. *Biometrics*, 52, 860–873.

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

Seber, G.A.F. (1982) *The estimation of animal abundance and related parameters*, 2nd edition. London: Griffin.

Slone, D.H. (2011) Increasing accuracy of dispersal kernels in grid-based population models. *Ecological Modelling*, 222, 573–579.

Taylor, C.M., Lank, D.B. and Sandercock, B.K. (2015) Using local dispersal data to reduce bias in annual apparent survival and mate fidelity. *Condor*, 117, 598–608.

Whittington, J. and Sawaya, M.A. (2015) A comparison of grizzly bear demographic parameters estimated from non-spatial and spatial open population capture-recapture models. *PLOS ONE*, 10, e0134446.

Williams, B.K., Nichols, J.D. and Conroy, M.J. (2002) *Analysis and management of animal populations*. San Diego, CA: Academic Press.

Zeng, Z. and Brown, J.H. (1987) A method for distinguishing dispersal from death in mark-recapture studies. *Journal of Mammalogy*, 68, 656–665.

Zucchini, W. and MacDonald, I.L. (2009) Hidden Markov models for time series: an introduction using R, *Monographs in Applied Statistics and Probability*. Boca Raton, FL: CRC Press, p. 110.

SUPPORTING INFORMATION

Web Appendices referenced in Sections 5 to 7 are available with this paper at the Biometrics website on Wiley Online Library. The models described here may be fitted with the R package openCR available at https://CRAN.R-project.org/package=openCR.

How to cite this article: Efford MG, Schofield MR. A spatial open-population capture-recapture model. *Biometrics*. 2020;76:392–402. https://doi.org/10.1111/biom.13150