October 2021

Statistical Improvements for Ecological Learning about Spatial Processes

Gaetan L. Dupont
University of Massachusetts Amherst

STATISTICAL IMPROVEMENTS FOR ECOLOGICAL LEARNING ABOUT SPATIAL PROCESSES

A Thesis Presented

by

GAETAN L. B. DUPONT

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE

September 2021

Organismic & Evolutionary Biology

STATISTICAL IMPROVEMENTS FOR ECOLOGICAL LEARNING ABOUT SPATIAL PROCESSES

A Thesis Presented

by

GAETAN L. B. DUPONT

Approved as to style and content by:

Chris Sutherland, Chair

Joseph Elkinton, Member

Toni Lyn Morelli, Member

Daniel Sheldon, Member

R. Craig Albertson, Director Interdisciplinary Graduate Programs Organismic & Evolutionary Biology

"If you have built castles in the air,

your work need not be lost;

that is where they should be.

Now put the foundations under them."

–Henry David Thoreau

Dedication

To my parents: two ice hockey journalists
who eagerly supported my early interest in *bruin tracks*
the same as if it was in *Bruins scores*.

ACKNOWLEDGMENTS

I would like to thank the following people.

Chris Sutherland: for being the chair and my advisor; for seeing and setting these perfect first stepping stones of my career path, and for bringing sophisticated approaches within my reach by teaching me to discern key moves from distracting details.

Joe Elkinton, Toni Lyn Morelli, and Dan Sheldon: for serving on my committee, providing me with space to explore while reminding me to not lose sight of the big picture.

Andy Royle, Ali Nawaz, and Dan Linden: for their support as co-authors, including longer-than-normal coffee breaks in noisy cafes, and later, impromptu Zoom meetings to dig into our newest ideas.

Members of the Sutherland Lab: for all forms of support – sometimes even on the volleyball court! – including entertaining my nagging questions!

The Sabin Snow Leopard Grant Program and Panthera, the Pakistan Snow Leopard and Ecosystem Protection Program, and the Snow Leopard Foundation: for financial support of this research that provided me with a front-row-seat into the world of snow leopard conservation.

Finally, I would like to express my gratitude to my family and friends, especially those who opened my eyes to ecology, ornithology, and statistics, and who encouraged me to pursue my dreams.

ABSTRACT

STATISTICAL IMPROVEMENTS FOR ECOLOGICAL LEARNING ABOUT SPATIAL PROCESSES

SEPTEMBER 2021

GAETAN L. B. DUPONT, B.Sc., CORNELL UNIVERSITY

M.Sc., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Chris Sutherland

Ecological inquiry is rooted fundamentally in understanding population abundance, both to develop theory and improve conservation outcomes. Despite this importance, estimating abundance is difficult due to the imperfect detection of individuals in a sample population. Further, accounting for space can provide more biologically realistic inference, shifting the focus from abundance to density and encouraging the exploration of spatial processes. To address these challenges, Spatial Capture-Recapture ("SCR") has emerged as the most prominent method for estimating density reliably. The SCR model is conceptually straightforward: it combines a spatial model of detection with a point process model of the spatial distribution of individuals, using data collected on individuals within a spatially referenced sampling design. These data are often coarse in spatial and temporal resolution, though, motivating research into improving the quality of the data available for

analysis. Here I explore two related approaches to improve inference from SCR: sampling design and data integration. Chapter 1 describes the context of this thesis in more detail. Chapter 2 presents a framework to improve sampling design for SCR through the development of an algorithmic optimization approach. Compared to pre-existing recommendations, these optimized designs perform just as well but with far more flexibility to account for available resources and challenging sampling scenarios. Chapter 3 presents one of the first methods of integrating an explicit movement model into the SCR model using telemetry data, which provides information at a much finer spatial scale. The integrated model shows significant improvements over the standard model to achieve a specific inferential objective, in this case: the estimation of landscape connectivity. In Chapter 4, I close by providing two broader conclusions about developing statistical methods for ecological inference. First, simulation-based evaluation is integral to this process, but the circularity of its use can, unfortunately, be understated. Second, and often underappreciated: statistical solutions should be as intuitive as possible to facilitate their adoption by a diverse pool of potential users. These novel approaches to sampling design and data integration represent essential steps in advancing SCR and offer intuitive opportunities to advance ecological learning about spatial processes.

Keywords: Spatial capture-recapture; population density; spatial sampling; optimal design; data integration; animal movement; movement modeling

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

1.1 Spatial processes in ecology

Fundamental to ecological inquiry is an understanding of population abundance through time and in space. This has long been recognized and motivates the two main factions of modern ecological research: theory and application. As we witness the anthropogenically initiated decline of countless species globally (coined "the biodiversity crisis"; Pimm et al. 1995; Dirzo and Raven 2003), the need for the most basic information on population size is apparent and essential (Williams et al., 2002). Even a static measurement of total population size is seriously lacking for many wide-ranging and elusive species of conservation concern – like the snow leopard (Jackson et al., 2006) – with multi-national initiatives focused solely on this issue. From a theoretical perspective, many foundational questions relate to changes in population size due to changing demographic rates such as survival and fecundity, stemming from any of a multitude of intrinsic and extrinsic ecological factors (Lack, 1954; MacArthur and Connell, 1967). A focus on population density, then, makes the matter explicit by considering populations in well-defined physical space, bringing much-needed biological realism to any ecological question or objective. Even the ongoing COVID-19 pandemic is a timely example of this: groups

of unmasked and unvaccinated individuals have decidedly higher infection rates, as density directly relates to disease spread (Chu et al., 2020). Such dynamics can be more complicated still: those density-dependent relationships can themselves be dynamic in time and space, as is the case with the house finch conjunctivitis system in which host congregations are actually less likely to facilitate disease spread due to genetically acquired immunity (Hochachka et al., 2021). To more efficiently understand the inner workings of these processes, a bit of math often comes in handy.

One of the classic works on understanding populations in space is the Levins metapopulation model. This mathematical model elegantly posits that patches of habitat can be occupied or not and are subject to colonization or extinction through time (Levins, 1969, 1970). Though the model is simple, its implications are profound: in further extensions, the importance of spatial processes and spatial heterogeneity is apparent as they influence the portion of patches occupied (equilibrium occupancy) and the entire spatial distribution of a species (Hanski, 1999). Further, when expanding to multi-species systems, species coexistence is impossible without spatial heterogeneity (Levins and Culver, 1971; Gravel and Massol, 2020). Even within such a simplified conceptualization of reality, the role of space quickly takes center stage. And real ecosystems are far more complex: how animals interact with each other and their surroundings results from processes of movement, home range selection, local dispersal, spatial stratification, for example, all of which can occur within a single patch (Morales et al., 2010). Encoded in these processes are responses to the environment, whether adapting to thermal tolerances, seeking alternative resources, or otherwise operating in the broader food web and ecosystem at large (Spiegel et al., 2017). Excitingly, we can investigate the majority of these processes using the vast array of statistical methods now available.

1.2 Spatial capture-recapture

To estimate population density, Spatial Capture-Recapture ("SCR") is by far the most prominent and advanced method (Efford, 2004; Borchers and Efford, 2008). In its most basic form, the SCR model combines a spatial model of detection (which is related to space use within a home range or territory), along with a point process model that describes the distribution of individuals in space (Efford, 2004; Borchers and Efford, 2008). Data are collected in the form of observations by an array of detectors with known configuration and relative coordinates (camera traps, hair snares, etc.) that allow for the identification of individuals within some explicitly defined study area (Royle et al., 2018). Despite the rapid uptake in ecology, the model remains largely in its infancy (Royle et al., 2018). To advance SCR and further promote its widespread use, I explore in this thesis two especially fruitful (and highly related) research avenues: sampling design and data integration.

The advantages of SCR are derived from the bottom up, starting with sampling design. To obtain precise estimates of density, one should first look to improve data quality, which requires improved sampling. Even before arriving at a field site, we can stack the deck in our favor. In more detail, the problem of sampling design for SCR can be simply stated: we have some number of detectors available (let's say, camera traps), but where should we place them? Previous research has explored this issue to some success: we want to capture a lot of individuals to understand the extent of variation in a population, but we also want to capture individuals at a lot of traps for a more precise understanding of how individuals use space. To achieve these competing objectives simultaneously, we can simplify trap configuration in SCR to a focus on ensuring trap spacing is not too far and not too close; the Goldilocks paradigm of design. Trap spacing too close minimizes

the probability of encountering individuals, while trap spacing too far minimizes the chance of individuals being captured at multiple traps. From there, previous research has utilized ad hoc approaches. These rely on simple comparisons of the distance between traps to the size of an individual's territory so that traps are just close enough to have 3-to-4 of them within an individual's territory (Sollmann et al., 2012; Efford and Fewster, 2013; Efford and Boulanger, 2019). Though good in theory, the recommendations lack applicability. Regular spacing between traps implies a regular grid of traps, and a regular grid makes for a rectangular study area. But landscapes have far more structural character. For example, rugged terrain carves out inaccessible areas, and governments liberally outline large portions of land for military bases or other forms of restricted access. Unless we intend to sample on a chessboard, there's an obvious need for more flexible approaches to optimize sampling designs and get the most out of the SCR model.

Much like sampling design, data integration also stems from a desire to improve the quality of the data available to the SCR model. Robust inference on density alone is indeed achievable through the use of an expertly designed survey. Still, any further inference on finer-scale processes will be hindered by the otherwise coarse spatial resolution delivered by typical ecological monitoring initiatives. In contrast, telemetry data can provide data at a much finer spatial scale. Collected by deploying a tag on an individual that records the location in space, telemetry data opens up a rich variety of possible movement models and many potential questions (Cooke et al., 2004). What the telemetry data lacks, though, is information on the location of all individuals in the population, which is one of the strengths of SCR, making for a natural integration. Recent advances to combine these two data types into a unified statistical framework are rudimentary. They rely on sampling only a fraction of the telemetry data at some regular time

interval (i.e., "thinning" the tracks) to satisfy SCR's requirement of temporal independence between observations (Royle et al., 2013c; Sollmann et al., 2016; Linden et al., 2018). Unfortunately, this degrades important information encoded in the autocorrelative structure of movement paths. SCR and animal movement modeling continue to grow in popularity (Nathan et al., 2008), motivating the need to integrate the approaches. SCR's population model can serve as the "statistical glue" for including more sophisticated movement processes to answer more detailed ecological questions (Kery and Royle, 2020).

1.3 A thesis roadmap

For the following chapters, I present my work on these two related concepts. Chapter 2 focuses on optimal sampling design for SCR, which brought me to the Sutherland lab group. This work was the product of a collaboration with Andy Royle at the Eastern Ecological Science Center, widely-known for helping to bring hierarchical modeling methods to the modern field of ecology; Ali Nawaz, one of the world's leading experts on snow leopards, who is associated with several of the largest organizations focused on the species; and of course, my advisor, Chris Sutherland, now at the Centre for Research into Ecological and Environmental Modelling at the University of St Andrews, who excels at seamlessly integrating those two perspectives of statistics and conservation. In our study, published in Ecology (Dupont et al. 2021c, Chapter 2), we start by acknowledging that sampling design for SCR can be framed as a straightforward optimization problem that can be solved using existing algorithms. We developed such an algorithm that selects trap locations by maximizing criteria directly related to the SCR model, and, using simulation, we confirmed the robust performance of the approach. Now packaged

into a user-friendly R function, researchers are already using this new framework to design monitoring protocols worldwide (at the time of writing: snow leopards in Pakistan and Mongolia and common leopards in South Africa, among others). Chapter 3 showcases the statistical skillset I have continued to develop during my degree and focuses on data integration to characterize non-Euclidean distance in the SCR model (Sutherland et al., 2015). This project was another exciting collaboration with Chris, and we were accompanied by Dan Linden at the National Oceanic and Atmospheric Administration. Dan's keen focus on the details, paired with his statistical prowess, thoroughly complemented Chris' astute ability to tease out the most important, big-picture takeaways that lend so much toward crafting the story of a paper. I owe much of my developed skillset to many hours spent on Zoom with Chris and Dan, screen-sharing my latest figure as my eyes flickered between their windows. This study, too, has been accepted in Ecology (Dupont et al. 2021a, In press, Chapter 3) and details one of the first attempts at integrating an explicit movement model into the SCR model, allowing all of the available data to be utilized rather than just a fraction of it. Finally, in Chapter 4, I provide a future outlook for research on design and data integration for SCR, pointing the reader to other recent developments, as well as providing a field guide of sorts of things to look for in the coming years that will be most meaningful to these endeavors.

CHAPTER 2

OPTIMAL SAMPLING DESIGN FOR SPATIAL CAPTURE–RECAPTURE

Gates Dupont, J. Andrew Royle, Muhammad Ali Nawaz, Chris Sutherland

2.1 Abstract

Spatial capture-recapture (SCR) has emerged as the industry standard for estimating population density by leveraging information from spatial locations of repeat encounters of individuals. The precision of density estimates depends fundamentally on the number and spatial configuration of traps. Despite this knowledge, existing sampling design recommendations are heuristic and their performance remains untested for most practical applications. To address this issue, we propose a genetic algorithm that minimizes any sensible, criteria-based objective function to produce near-optimal sampling designs. To motivate the idea of optimality, we compare the performance of designs optimized using three model-based criteria related to the probability of capture. We use simulation to show that these designs out-perform those based on existing recommendations in terms of bias, precision, and accuracy in the estimation of population size. Our approach, available as a

function in the R package `oSCR`, allows conservation practitioners and researchers to generate customized and improved sampling designs for wildlife monitoring.

Keywords: SCR, spatial capture-recapture, spatially-explicit capture-recapture, camera traps, density, optimal design, sampling design, spatial sampling, trap spacing, genetic algorithm

2.2 Introduction

The need for conservation managers and practitioners to obtain reliable estimates of population size (Williams et al., 2002) has driven the rapid development of data collection and estimation methods. Capture-recapture (CR), and more recently, spatial capture-recapture (SCR; Efford, 2004; Borchers and Efford, 2008) methods were developed specifically for this purpose and are now routinely applied in ecological research. Concurrently, SCR methods estimate detection, space use, and density by analyzing individual encounter histories while explicitly incorporating auxiliary information from the spatial organization of encounters (Efford, 2004; Royle et al., 2013b). Despite widespread adoption and rapid method development, recommendations about spatial sampling design have received relatively little attention and are arguably heuristic.

The effects of sampling design have been investigated for both CR (Dillon and Kelly 2007; Bondrup-Nielsen 1983) and SCR methods (discussed below). While CR methods aim to balance the number of captures and the number of recaptures, SCR requires a third consideration: the spatial pattern of recaptures at multiple traps. The ability to reliably estimate density is directly related to these considerations: the number of captured individuals n is the sample size; the number of recaptures is

directly related to the baseline detection probability, g_0; and the number and spatial distribution of recaptures are directly related to the spatial scale parameter, σ, as well as the spatial distribution of activity centers. Therefore, improving sampling design has great potential to increase the quality of the data and the precision of parameter estimates.

Several simulation studies evaluating SCR designs have shown that inference is robust to the spatial configuration of traps, as long as some minimum requirements are met: the trap spacing must not be too large relative to individual space use in order to reliably estimate σ, but the array must not be too small such that too few individuals are exposed to capture (Sollmann et al., 2012; Sun et al., 2014; Wilton et al., 2014; Efford and Boulanger, 2019; Tobler and Powell, 2013). Repeated illustrations of this trade-off have lead to recommendations that trap spacing should be approximately two times σ, which maximizes accuracy and minimizes bias of abundance estimates (Sollmann et al., 2012; Efford and Fewster, 2013; Efford and Boulanger, 2019). While most of this research has focused on uniform grids, simulation has also shown that clustered designs can outperform uniform designs (Efford and Fewster, 2013; Sun et al., 2014), particularly for heterogeneously distributed populations (Efford and Fewster, 2013; Wilton et al., 2014). In summary, formalization of the factors that contribute to optimal sampling design for SCR is in it's infancy, described only in generalities. In particular, it is unclear whether existing design heuristics generally hold for spatially-varying density patterns, or in highly-structured landscapes where recommended regular trapping arrays cannot be accommodated, and guidance of generating clustered designs is lacking.

Generally speaking, sampling design for SCR can be conceived as a problem of selecting a subset of all possible trap locations that maximizes some SCR-relevant objective function. Here we develop an analytical framework that directly addresses

this challenge. Our approach generates a near-optimal sampling design with respect to some appropriately defined objective function and information about available resources (traps), a set of all possible trap locations, and information about SCR model parameters. To motivate the idea of optimality, we use simulation to compare the performance of existing recommendation to designs optimized using three model-based criteria related to current thinking about the relationship between data quality and estimator bias and precision. We explore design performances for scenarios where we vary the spatial coverage of traps, the landscape geometry, and deviations from uniform spatial distribution of individuals. Finally, our approach is available as a function in the R package `oSCR`.

2.3 Methods

2.3.1 The standard SCR model

Typically, SCR models have two model components: a spatial model of abundance describing the distribution of individuals characterized by the center of their home range (hereby referred to as an activity center), and a spatial model of detection that relates encounter rates to the distance between the activity center and a trap (e.g., a camera trap). The most basic form assumes a uniform prior for the distribution of activity centers, s_i:

$$s_i \sim \text{ Uniform}(\mathcal{S}),$$

where $\mathcal{S}$, referred to as the state-space, describes all possible locations of activity centers. To facilitate analysis, $\mathcal{S}$ is represented as a uniform grid of points representing the centroids of equal-sized pixels. All individuals within the region,

N, are exposed to capture resulting in the observation of n individuals and hence $n_0 = N - n$ unobserved individuals.

While several formulations of the encounter model exist, we use, without loss of generality, a half-normal encounter model that describes encounter probability as a decreasing function of distance from an individual's activity center s_i:

$$p_{ijk} = g_0 \times \exp(-d(s_i, x_j)^2/(2\sigma^2)), \tag{2.1}$$

where p_{ijk} is the probability of detection of individual i with activity center s_i at trap j during sampling occasion k; $d(s_i, x_j)$ is the distance between the activity center s_i and the trap x_j, and g_0 and σ are the baseline encounter probability and spatial scale parameters, respectively.

2.3.2 Model-based objective functions

From Equation 1, we can use values of g_0 and σ (e.g., from the literature or estimates from a pilot study), to compute the probability that an individual with an activity center s_i is detected in *any* trap in an array $\mathcal{X}$, which we denote as $\bar{p}$:

$$\bar{p}(s_i, \mathcal{X}) = 1 - \prod_{j=1}^{J} \{1 - p(s_i, x_j)\}.$$

The corresponding marginal probability of not being encountered is thus: $\bar{p}_0(s_i, \mathcal{X}) = 1 - \bar{p}(s_i, \mathcal{X})$. Taking the average over all G activity center locations in the landscape $\mathcal{S}$, we can compute the marginal probability of encounter:

$$\bar{p}(\mathcal{X}) = \frac{1}{G} \sum_{s} \bar{p}(s_i, \mathcal{X}).$$

We can also compute the probability of being captured in exactly one trap:

$$\bar{p}_1(s_i, \mathcal{X}) = \bar{p}_0(s_i, \mathcal{X}) \sum_{j=1}^{J} \frac{p(s_i, x_j)}{1 - p(s_i, x_j)}.$$

Finally, the marginal probability of being encountered at more than one trap, i.e., of a spatial recapture is:

$$\bar{p}_m(\mathcal{X}) = \frac{1}{G}\sum_s \{1 - \bar{p}_0(s_i, \mathcal{X}) - \bar{p}_1(s_i, \mathcal{X})\}.$$

Given that the precision of SCR density estimates depends on the total number of individuals captured, n, and the number of spatial recaptures, m (Efford and Boulanger, 2019; Royle et al., 2013b) – $Q_{\bar{p}}$ and $Q_{\bar{p}_m}$ represent logical criteria for optimizing SCR designs (Royle et al. 2013b, Chapter 10). Herein lies one of our novel contributions: we suggest three design criteria: $Q_{\bar{p}} = -\bar{p}(\mathcal{X})$, $Q_{\bar{p}_m} = -\bar{p}_m(\mathcal{X})$, and $Q_{\bar{p}_b} = Q_{\bar{p}} + Q_{\bar{p}_m}$. Importantly, if approximate values of the SCR parameters, g_0 and σ, are available, these objective functions can be evaluated analytically for any number and configuration of traps, providing a metric for efficient identification of optimal SCR designs.

2.3.3 Optimization method

We applied a genetic algorithm (GA) to the task of finding a design that minimizes any criterion, noting that optimality here is with respect to the defined criteria, and in the context of the GA is 'near-optimal' (see Appendix A & Goldberg, 1989). The GA is a random search algorithm which produces multiple generations of solutions, where subsequent generations retain characteristics of top performing solutions from the previous generation. Generations are produced until converging on a near-optimal solution. Wolters (2015) adapted the algorithm to solve a *k-of-n* problem which describes concisely the challenge of the SCR sampling design: the selection of some number of traps, k, in a landscape of n possible locations according to some objective function. We provide a detailed description of the general GA, the *k-of-n* adaptation, and our implementation in the `R` package `oSCR`

in Appendix A and Appendix D.

Conceptually, minimizing the space-filling objective function $Q_{\bar{p}}$ maximizes the expected sample size n. In contrast, minimizing $Q_{\bar{p}_m}$ prioritizes the exposure of individuals to more than one trap and should maximize the number of spatial recaptures m. The third criteria, $Q_{\bar{p}_b}$, attempts to balance $Q_{\bar{p}}$ and $Q_{\bar{p}_m}$.

2.3.4 Design constraints

We were primarily interested in evaluating the performance of SCR designs produced by our framework under a range of biologically-realistic scenarios in an attempt to develop a more general understanding of how performance varies as a function of the following design constraints: *geometry*, defined as the shape of the study area and ease at which a regular square trapping grid can be deployed; *density pattern*, defined as the nature of departure from uniform distribution of individuals; and *effort*, defined as the number of traps available for the design.

Geometry – As has been typical in studies investigating SCR sampling designs, we begin using a square study area with complete accessibility and which lends itself to uniform trapping grids (the *regular area*, Figure 1). To replicate the design challenges posed when generating real-world designs, we also consider an *irregular area* (Figure 1). For this, we use one of the study areas that motivated this work: a large area in Northern Pakistan (3865 km^2) that is the focus of a snow leopard *(Panthera uncia)* camera trapping study, but that has several logistical challenges that determine accessibility (i.e., remoteness, private property, altitude, and slope). To define the complete region of the state-space, we used a 3σ buffer around the trapping extent. The regular area is represented by 24 x 24 landscape with a resolution of 0.5 units, the irregular study area is represented by 89.85 x 133.04 landscape with a resolution of 1.73 units, for a total of 2304 cells in each of the

geometries (Figure 1). While these two state-spaces differ in absolute terms, we insured comparability in relative terms by the definition of area-specific sigma (see below).

Density pattern – Existing investigations of SCR sampling designs typically assume a homogeneous distribution of individuals (but see Efford and Fewster, 2013). Here we formally test the adequacy of designs under specific violations of this assumption. We consider three spatial density patterns: a uniform and two spatially-varying. To generate non-uniform density patterns, we simulated landscapes defined by a parametric Gaussian random field that allows for specification of the degree and range of spatial autocorrelation. Gaussian random fields were generated using the R package, `NLMR` (Sciaini et al., 2018). The values of the simulated landscape were scaled from 0 to 1 and individual activity centers distributed according to the following cell probabilities:

$$\pi_i = \frac{e^{\beta_1 * X_i}}{\sum e^{\beta_1 * X_i}}, \tag{2.2}$$

where X_i is the scaled landscape value at pixel i and β_1 is defined as 1.2 to represent a weak but apparent density pattern. The two inhomogeneous density patterns differ in the scale of spatial autocorrelation. For consistency, we defined this distance in relative terms to the length of the longest side of the state-space: 6% for a *weak* density pattern that produces a patchy landscape, and 100% for a *strong* density pattern produces a landscape with a more contiguous gradient (see Figure 1 for a single realization of the density patterns). Using these three density patterns allows us to evaluate designs through a full range of biological realism, with uniform and strong density patterns representing the polar ends of reality, and the patchy landscape representing the most realistic sampling scenario.

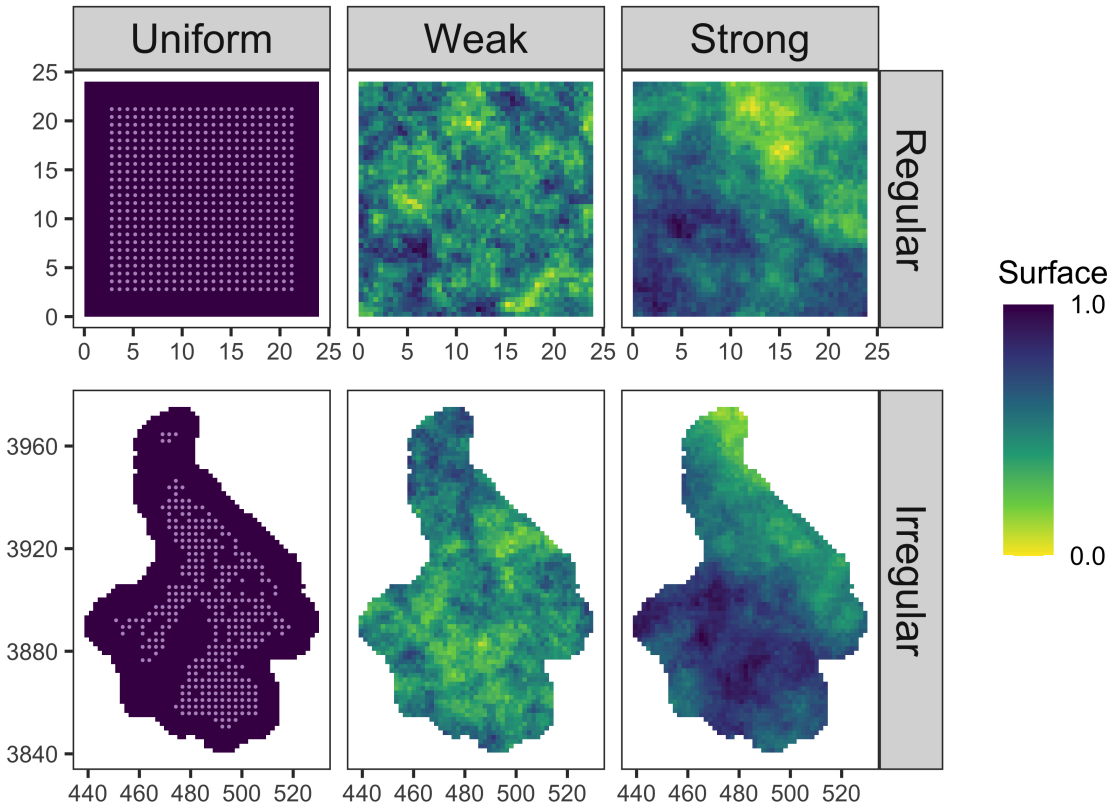

Figure 1. Simulation surfaces for design evaluation
Here we show all possible trap locations overlaid on the uniform landscape for the regular (top) and irregular (bottom) study area geometries alongside a single realization of two (weak: middle, strong: right) of the three (uniform not shown) landscape covariates. For the regular geometry, we tested 12 designs each. For the irregular geometry, we tested 9 designs each. This makes for a total of 63 scenarios.

2.3.5 Design generation

Designs were generated using fixed values of g_0 and σ (see below), a set of potential trap locations, and the number of traps that are available to deploy. It is assumed that the user has knowledge or access to data on information approximate values of SCR parameters, would be able to produce a set of all potential sampling points, and would have some idea of resources (traps) available. For the regular area, we generated $Q_{\bar{p}}$, $Q_{\bar{p}_m}$, and $Q_{\bar{p}_b}$ designs for each of the three levels of effort where there was no restriction on where traps could be placed. In addition, we

generated a regular 2σ design for comparison. For the irregular area in the mountains of Pakistan, we generated only criteria-based designs at each of the three levels of effort (Figure 2). In this case, areas known to be too remote, too high altitude, or too steep to be accessed were removed from the set of potential trap locations. Mirroring real design challenges faced by managers, it was not practical to generate a 2σ grid for the irregular area, and therefore it is not included. This full scenario analysis resulted in a total of 21 designs; 12 designs for the regular area (the three optimized and the 2σ design), and 9 designs for the irregular area (optimized designs only).

2.3.6 Evaluation by simulation

We exposed a population of $N = 300$ individuals to sampling via each of the 21 designs described above. We simulated encounter histories assuming proximity detectors and under the binomial encounter model encounter (Eq.1) with $g_0 = 0.2$, $k = 5$. The two geometries differ in terms of their spatial units so area-specific σ values were chosen such that the number of home ranges required to fill the areas and achieve an equal density was equivalent: $\sigma_{reg} = 0.80$ and $\sigma_{irreg} = 2.59$. We simulated individuals according to the three density patterns described above (Eq.2), resulting in a total of 63 scenarios of interest (three density patterns for each of the 21 design; Appendix C).

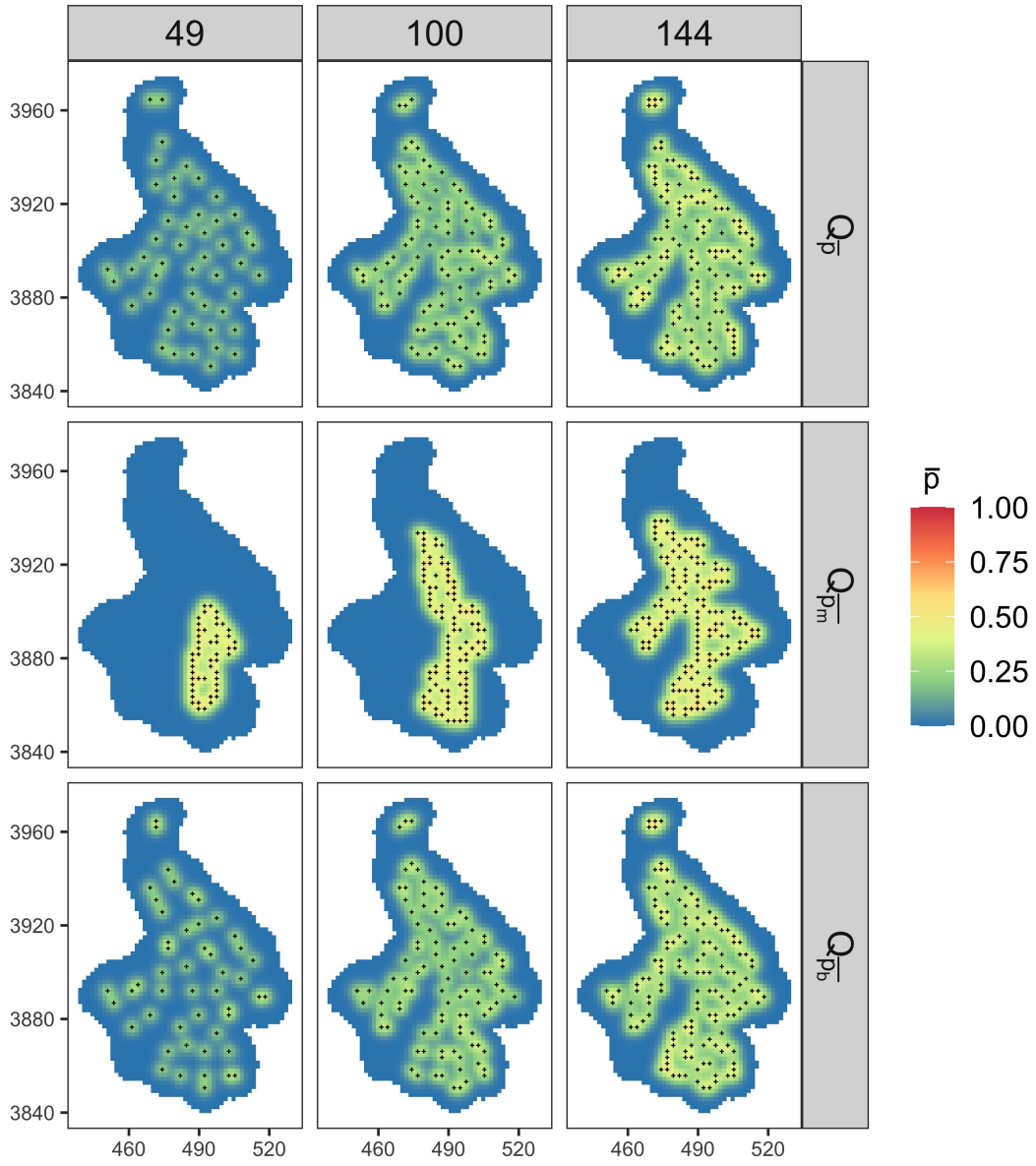

Figure 2. Designs and detection surfaces in the irregular study area
Irregular study area with designs generated using our new framework with three SCR-intuitive, model-based criteria ($Q_{\bar{p}}$, $Q_{\bar{p}_m}$, and $Q_{\bar{p}_b}$), under three levels of effort. 144 traps represents the same number of traps as used to generate a full 2σ grid in a regular study area of the same area. 100 traps is nearly two-thirds as many traps, and 49 is nearly one-third as many traps. Each pixel of the state-space is colored according to the probability of capture, p, for an individual with an activity center at the centroid of the pixel.

For each scenario, we simulated 300 realizations of activity centers. Covariate surfaces were generated randomly using the same seed, again resulting in variation among simulations but consistency across scenarios. In some cases, the realization of activity centers did not provide at least one spatial recapture; we recorded the number of these *failure s* and generated a new realization of activity centers until a single spatial recapture was obtained in order to proceed with model fitting. This only occurred for $Q_{\bar{p}}$ designs with minimum effort, and for less than 5% of the simulations.

We analyzed the resulting encounter history data using a null SCR model ($d.$) and, for spatially structured density scenarios, a density-varying model (d_s). This allowed us to test if accounting for the landscape would improve bias and precision in parameter estimates. For each simulation, and each model, we retained estimates of g_0, σ, and total abundance ($\hat{N}$).

We compared estimates of model parameters to the data-generating values in terms of bias (percent realtive bias, %RB), precision (coefficient of variation, CV), and accuracy (scaled root mean square error, SRMSE). All simulations were conducted in `R`, SCR models were fit using the package `oSCR` (Sutherland et al., 2019), and designs were generated using the `scrdesignGA()` function also in `oSCR` (detailed workflow provided in Appendix D, repository referenced in Appendix I). Design generation and simulations were performed in R version 3.6.1 (R Core Team, 2019).

2.4 Results

We first focus on relative bias. Encouragingly, under the regular-area, homogeneous-density scenario, designs generated using the genetic algorithm perform as well

as existing 2σ recommendations, producing unbiased estimates of abundance for nearly all combinations of design and effort (Figure 3, Table 1). In the case of the irregular geometry with uniform density, $Q_{\bar{p}_m}$ designs perform well for all levels of effort, but performance of $Q_{\bar{p}}$ and $Q_{\bar{p}_b}$ designs declines as the number of traps is reduced, a consequence of widely-spaced traps that result in very few spatial recaptures (Figure 3, Table 1).

Table 1. Percent relative bias of point estimates from tested designs

Percent relative bias of baseline detection (g_0), space use (σ) and total abundance (EN) for each simulation scenario, varying: design criteria *(Design)*, landscape shape *(Geometry, Regular or Irregular)*, the number of traps *(Effort)*, and density patterns *(Density)*. We present results from null ($d.$) and varying density (d_s) models.

			Regular						Irregular					
			g_0		σ		EN		g_0		σ		EN	
Effort	Density	Design	$d.$	d_s	$d.$	d_s	$d.$	d_s	$d.$	d_s	$d.$	d_s	$d.$	d_s
49	uniform	2σ	2.52	–	-0.38	–	0.78	–	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.82	–	-1.00	–	7.27	–	2.27	–	-1.84	–	8.34	–
		$Q_{\bar{p}_m}$	1.33	–	-0.19	–	1.76	–	1.78	–	-0.15	–	0.62	–
		$Q_{\bar{p}_b}$	-0.61	–	-2.06	–	13.32	–	2.53	–	-4.11	–	17.90	–
	weak	2σ	3.16	3.16	-0.62	-0.61	-0.26	-0.05	–	–	–	–	–	–
		$Q_{\bar{p}}$	-0.58	-0.58	0.20	0.25	5.70	5.75	-1.51	-1.51	-1.11	-1.07	9.93	9.89
		$Q_{\bar{p}_m}$	0.08	0.08	0.06	0.11	0.99	1.99	1.15	1.15	-0.27	-0.22	0.07	2.74
		$Q_{\bar{p}_b}$	-2.73	-2.73	-2.36	-2.16	16.12	14.83	0.19	0.19	-1.48	-1.46	13.68	14.09
	strong	2σ	2.26	2.26	-0.47	-0.48	1.82	3.48	–	–	–	–	–	–
		$Q_{\bar{p}}$	1.84	1.84	-0.75	-0.78	6.43	6.55	1.18	1.18	-0.27	-0.32	5.80	6.17
		$Q_{\bar{p}_m}$	2.09	2.09	-0.47	-0.48	1.20	6.82	2.29	2.29	-1.03	-1.01	2.40	9.02
		$Q_{\bar{p}_b}$	0.99	0.99	-3.47	-3.41	14.54	14.10	2.75	2.75	-3.32	-3.26	15.13	15.14
100	uniform	2σ	2.04	–	-0.69	–	0.58	–	–	–	–	–	–	–
		$Q_{\bar{p}}$	2.42	–	-0.61	–	0.90	–	1.42	–	-0.77	–	2.11	–
		$Q_{\bar{p}_m}$	-0.97	–	0.20	–	1.07	–	0.74	–	-0.18	–	0.83	–
		$Q_{\bar{p}_b}$	0.07	–	0.05	–	1.12	–	-0.15	–	-0.51	–	2.55	–
	weak	2σ	-0.13	-0.13	0.15	0.14	-0.34	-0.19	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.61	0.61	-0.27	-0.29	0.95	0.98	0.97	0.97	-0.48	-0.49	1.82	1.89
		$Q_{\bar{p}_m}$	1.68	1.68	-0.77	-0.78	-0.24	0.34	-0.09	-0.09	0.09	0.08	0.34	1.04
		$Q_{\bar{p}_b}$	1.07	1.07	-0.16	-0.18	0.01	0.03	1.23	1.23	-0.30	-0.27	1.06	1.11
	strong	2σ	0.35	0.35	-0.30	-0.30	1.42	1.72	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.18	0.18	-0.93	-0.95	2.89	3.12	1.07	1.07	-0.46	-0.49	0.93	1.40
		$Q_{\bar{p}_m}$	0.64	0.64	-0.04	-0.05	0.90	1.47	1.97	1.97	-0.56	-0.59	-0.44	1.34
		$Q_{\bar{p}_b}$	0.60	0.60	-0.43	-0.43	1.36	1.44	0.21	0.21	-0.05	-0.06	0.40	0.80
144	uniform	2σ	1.32	–	-0.25	–	0.27	–	–	–	–	–	–	–
		$Q_{\bar{p}}$	-1.06	–	0.28	–	1.53	–	0.72	–	0.08	–	-0.27	–
		$Q_{\bar{p}_m}$	0.93	–	-0.28	–	0.88	–	0.53	–	0.00	–	0.75	–
		$Q_{\bar{p}_b}$	0.35	–	-0.07	–	0.90	–	2.12	–	-0.77	–	0.72	–
	weak	2σ	0.49	0.49	-0.33	-0.33	0.41	0.50	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.64	0.64	-0.24	-0.25	0.44	0.47	0.61	0.61	-0.20	-0.20	0.50	0.51
		$Q_{\bar{p}_m}$	1.31	1.31	-0.47	-0.48	-0.39	-0.21	0.03	0.03	0.05	0.04	0.07	0.43
		$Q_{\bar{p}_b}$	-0.02	-0.02	-0.32	-0.33	1.00	0.98	0.77	0.77	-0.25	-0.26	0.93	0.92
	strong	2σ	0.70	0.70	-0.25	-0.25	0.80	1.01	–	–	–	–	–	–
		$Q_{\bar{p}}$	1.35	1.35	-0.31	-0.32	0.32	0.47	-0.13	-0.13	0.21	0.19	0.33	0.66
		$Q_{\bar{p}_m}$	0.14	0.14	0.15	0.14	0.32	0.58	1.74	1.74	-0.55	-0.57	-0.22	0.69
		$Q_{\bar{p}_b}$	1.18	1.18	-0.19	-0.20	-0.03	0.14	-0.59	-0.59	0.12	0.09	0.20	0.62

For scenarios from the regular study area with inhomogeneous density, all designs produced unbiased estimates of abundance, generally. There is a slight bias ($\pm$ 5%) introduced as the number of traps declines, even for the 2σ designs. However, this phenomenon is less apparent in $Q_{\bar{p}_m}$ designs, suggesting improved performance. In the irregular study area, design performance is more dependent on the spatial structure of density. Once again, $Q_{\bar{p}_m}$ designs produced unbiased estimates, and $Q_{\bar{p}}$ and $Q_{\bar{p}_b}$ designs performed poorly with fewer traps (Figure 3, Table 1).

Interestingly, explicitly including the landscape covariate governing spatial variation in density (i.e., d_s rather than $d.$) does not improve performance metrics for any of the designs in any scenario (Figure 3, Table 1), reinforcing the general opinion that SCR models are robust to misspecification of the density model. In fact, fitting the data-generating model for the inhomogeneous cases actually performs worse in low effort scenarios. This suggests that the low numbers of traps do not adequately represent the variation in the landscape, and therefore, the model is unable to reliably estimate the underlying landscape effect (Figure 3, Table 1).

Estimator precision and accuracy generally follow the same patterns as for the bias (Appendix E, Appendix F, Appendix G). Design performance declines as effort decreases for all designs across every scenario. In the regular study area with uniform density, the 2σ and $Q_{\bar{p}_m}$ designs share similar levels of precision, while the $Q_{\bar{p}}$ and $Q_{\bar{p}_b}$ designs with minimal effort are less precise in comparison, with this pattern being magnified in the irregular area. Generally, there is a slight loss of precision in estimates across all designs, but this effect is less apparent for $Q_{\bar{p}_m}$ designs, which maintain their relative equivalency to the standard recommendation, including for the lowest level of effort (when considering comparison across geometries). In scenarios with inhomogenous density, both $Q_{\bar{p}}$ and $Q_{\bar{p}_b}$ designs with minimum effort show precision that is obviously reduced using the null model.

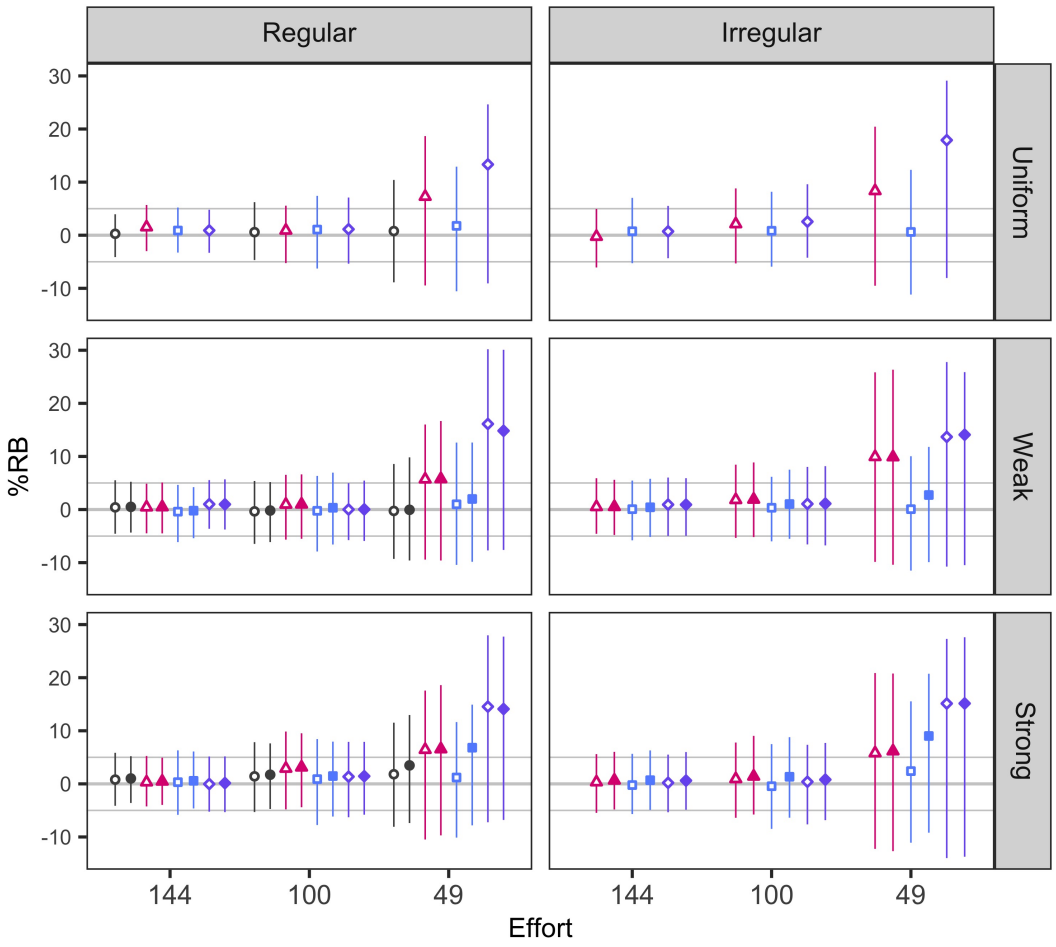

Figure 3. Results from simulation-based evaluations of designs
Percent relative bias (%RB) of estimates of total abundance from the three tested sampling designs under three levels of effort on three density surfaces within two geometries, where estimates are the result of one of two SCR models: density invariant ($d.$, open shapes) or density-varying (d_s, closed shapes). The four designs – 2σ, $Q_{\bar{p}}$, $Q_{\bar{p}_m}$, $Q_{\bar{p}_b}$ – are represented by the four shapes: circles, triangles, squares, and diamonds, respectively. To illustrate estimator precision, vertical lines are 50% confidence intervals, noting that the 50% intervals are proportional to 95% intervals but offer a visual balance of bias and associated variance. The thick horizontal line represents no bias in estimates, with the thin horizontal lines representing an allowable amount of bias ($\pm$ 5%).

However, the density-varying model once again shows no noticeable improvement, and causes a decrease in precision for $Q_{\bar{p}_m}$ designs with the fewest traps.

Overall, designs generated using our proposed framework showed comparable performance to standard recommendations, and critically, these designs are robust

to a variety of constraints that include effort, density signal, and geometry.

2.5 Discussion

In this study, we develop a conceptual and analytical framework for generating near-optimal designs for SCR studies. We suggested three intuitive and statistically-grounded design criteria that can be optimized to produce candidate designs. We demonstrate that designs generated using our framework can perform at least as well as those based on existing heuristics, and further, that the generality and flexibility of our approach means it can be applied to any species or landscape according to logistics and available resources.

It is worth noting that the designs produced using this framework can be considered approximate in terms of specific location, and that the actual, finer-scale site-selection for traps can be informed by knowledge of the species' biology and behavior (e.g., Fabiano et al., 2020). Further, while we develop this framework with camera traps in mind, this method can easily be applied to determine the general location of other non-invasive surveys, wherein the selection of a sampling location instead activates some other form of sampling effort (see Fuller et al. 2016; Sutherland et al. 2018). Importantly, the degree of sampling effort must be maintained among all selected sampling locations.

The designs we created using model-based criteria exhibit their own unique behaviors (Figure 2, Appendix B). The $Q_{\bar{p}}$ criteria generates space-filling designs to maximize the area covered and thereby the expected sample size of unique individuals. As more traps are added, the inner area becomes fully-saturated (such that it is insured that every possible home range will contain at least one trap), and the criteria instead focuses on selecting external traps that patrol the edge of

the trapping extent in order to increase the probability of capture for individuals outside of that area. However, despite the benefit of increasing the sample size (n captured individuals), traps placed too distant from each other fail to generate important spatial recaptures. This is precisely the issue that propagated failures for both $Q_{\bar{p}}$ and $Q_{\bar{p}_b}$ designs with minimum effort (Appendix H).

In contrast, $Q_{\bar{p}_m}$ designs are space-restricting as a result of an inherent trade-off between increasing the number of individuals exposed to capture and having traps close together to insure captures at more than one trap. With fewer traps, however, the effective sampling area is markedly decreased (Figure 2), thereby reducing the sample size. This observation further motivated our evaluations of the designs for inhomogeneous density, which along with the reduced spatial coverage and hence non-representative sampling, is likely responsible for the bias observed in those scenarios, as well as the lower precision.

The $Q_{\bar{p}_b}$ designs can best be described as "clustered space-filling" (Figure 2, Appendix B), as this criteria aims to balance the objectives of $Q_{\bar{p}}$ and $Q_{\bar{p}_m}$, which it can do effectively when provided with a sufficient number of traps. However, as seen with $Q_{\bar{p}}$ designs, the $Q_{\bar{p}_b}$ design performance suffers when too few traps are employed due to even larger distances between traps as a result of clustering, greatly reducing performance even beyond that of $Q_{\bar{p}}$.

More generally, these designs support previous recommendations while also providing new insights. When full effort is possible in the regular area geometry, the $Q_{\bar{p}}$ design fully saturates the trapping extent with some traps to spare in order to meet its objective, while $Q_{\bar{p}_m}$ does not quite fill the trapping area (Figure 2, Appendix B). Interestingly, the 2σ design falls somewhere between these two extents, likely striking an effective balance between the number of captures (as in $Q_{\bar{p}}$) against the number of spatial recaptures (as in $Q_{\bar{p}_m}$), which we also see with $Q_{\bar{p}_b}$

and similar to the effect described by Efford and Boulanger (2019). Despite these differences in spatial configuration, differences in design performance are mostly negligible (Figure 3, Table 1, Appendix E, Appendix F, Appendix G).

As shown by Sun et al. (2014), incorporating trap clustering into sampling designs can be advantageous, as doing so allows for increased likelihood of spatial recaptures to facilitate estimation of the spatial scale parameter, σ. However, the clustered designs proposed by Sun et al. (2014) follow a regular pattern such that there are a limited number of levels of trap spacing, whereas the designs we generated result in a wider distribution of distances between traps. This shifts the importance away from a regular spatial structure of trap configuration to one that is decidedly irregular in order to gain better resolution of movement distances for estimating σ. This is especially useful knowledge and central to generating designs for irregular study areas. Interestingly, this results in designs with smaller effective sampling areas, suggesting that it might be better to reduce the total area covered by the design rather than focus on completely covering the area (within reason). A major insight here is that hierarchical clustering (the selection of approximately 2σ-spaced clusters of traps with further reduced within-cluster spacing) emerges naturally from the $Q_{\bar{p}_m}$ criterion, effectively formalizing the clustering heuristic proposed by Sun et al. (2014).

Our proposed criteria produced designs which perform well, yet there is scope for refinement. With a decrease in effective sampling area, the introduction of bias and imprecision in parameter estimates could be complicated further when the population being sampled has a stronger degree of spatial structuring than we tested here. Designs sampling only areas where individuals are concentrated will result in overestimates of population size and density relative to the whole study area, while those sampling away from concentrated areas will do just the opposite.

This effect is particularly noticeable from the density-varying model (d_s), which generally has relatively lower performance over the fully invariant model as it is including information from nearby traps sampling a landscape that is intrinsically spatially auto-correlated. Advancing this framework to generate designs that explicitly account for the spatial patterns in density as a function of a given landscape is clearly an area for further development, especially if the inferential objective is to estimate density-landscape relationships rather than density or total abundance.

Recently SCR sampling design for multi-species sampling has been considered, with some discussion on how the distribution of trap spacing can allow for better estimates for species with a variety of home range sizes (Rich et al., 2019). However, the design proposed for this purpose lacks a reproducible framework that can be generalized to any biological community. Alternatively, employing our framework for multi-species sampling could be a straightforward approach to this problem, with important implications for the use of SCR to be more easily applied for the study of ecological communities. Again, a highly appealing feature of our $Q_{\bar{p}_m}$ approach is the emergence of designs with much better distribution of trap spacing than under regular designs such as 2σ grids, ideal for sampling groups of species with varying spatial movement ecology.

We considered three criteria that are intuitive in the context of the performance trade off of sample size (n) and spatial recaptures (m). While intuitive, alternative criteria surely exist. For example, Efford and Boulanger (2019) propose an approximation of the variance of density which is related to n and m, and therefore can easily be formulated as an objective function to be optimized in the same way as $Q_{\bar{p}}$ and $Q_{\bar{p}_m}$. Indeed, the function `scrdesignGA()` is designed such that any user-defined objective functions can be used (e.g., Durbach *et al.* 2020 *In press*). We hope that this ability to simultaneously (and efficiently) generate and evaluate

designs based on a variety of design criteria will motivate further research on SCR study design.

Our results show that designs obtained under our proposed criteria perform well relative to design heuristics and can be obtained efficiently as solutions to an optimization problem for arbitrary configurations of possible trapping locations and landscapes, unlike standard recommendations based on 2σ and cluster designs. Both CR and SCR studies are extremely expensive and require substantial effort to conduct, making it imperative that managers are provided with a method to select detector placement before deployment, such as the approach we have presented here. As a result, designs will produce a greater amount of expected information and will lead to more accurate estimates of parameters that describe biological populations of interest, which is critical to global conservation efforts, especially for low density and declining species that are of conservation concern but challenging to monitor.

CHAPTER 3

IMPROVED INFERENCES ABOUT LANDSCAPE CONNECTIVITY FROM SPATIAL CAPTURE-RECAPTURE BY INTEGRATION OF A MOVEMENT MODEL

Gates Dupont, Daniel W. Linden, Chris Sutherland

3.1 Abstract

Understanding how broad-scale patterns in animal populations emerge from individual-level processes is an enduring challenge in ecology that requires investigation at multiple scales and perspectives. Complementary to this need for diverse approaches is the recent focus on integrated modeling in statistical ecology. Population-level processes represent the core of spatial capture-recapture (SCR), with many methodological extensions that have been motivated by standing ecological theory and data integration opportunities. The extent to which these recent advances offer inferential improvements can be limited by the data requirements for quantifying individual-level processes. This is especially true for SCR models

that use non-Euclidean distance to relax the restrictive assumption that individual space use is stationary and symmetrical in order to make inferences about landscape connectivity. To meet the challenges of scale and data quality, we propose integrating an explicit movement model with non-Euclidean SCR for joint estimation of a shared cost parameter between individual and population processes. Here, we define a movement kernel for step selection that uses "ecological distance" instead of Euclidean distance to quantify availability for each movement step in terms of landscape cost. We compare performance of our integrated model to that of existing SCR models using realistic animal movement simulations and data collected on black bears. We demonstrate that an integrated approach offers improvements both in terms of bias and precision in estimating the shared cost parameter over models fit to spatial encounters alone. Simulations suggest these gains were only realized when step lengths were small relative to home range size, and estimates of density were insensitive to whether or not an integrated approach was used. By combining the fine spatiotemporal scale of individual movement processes with the estimation of population density in SCR, integrated approaches such as the one we develop here have the potential to unify the fields of movement, population, and landscape ecology and improve our understanding of landscape connectivity.

Keywords: Animal movement, connectivity, cost function, data-integration, spatial capture-recapture

3.2 Introduction

Current ecological theory frames population processes as emerging from individual behaviors and interactions with local environments that scale up to landscape-

level patterns (Levin, 1992; Morales et al., 2010; Spiegel et al., 2017). Complementary to this perspective are integrative statistical frameworks for the treatment of data from multiple sources, now a prominent research focus in statistical ecology (Plard et al., 2019). Integrated modeling offers an analytical approach that, by combining data for inference across levels of biological organization, facilitates the advance of ecological theory (McClintock et al., 2021). Spatial capture-recapture (SCR; Efford 2004; Royle and Young 2008) presents an unparalleled opportunity for exploration at the crossroads of these endeavors, built upon an implied model of individual space use related to a model of how individuals are distributed in space. However, the model is typically informed by data on spatially-explicit encounter histories that are relatively sparse if not insufficient for characterizing more detailed biological processes as opposed to population-level inferences (e.g., spatial variation in density). Data integration can directly address this challenge, providing the ability to improve inferences about mechanism and creating an explicit link between individual- and population-level processes and landscape-level spatial patterns.

Many methodological extensions of SCR have been motivated by standing ecological theory (see review by Royle et al. 2018), although the extent to which these advances offer inferential improvements is somewhat limited by data requirements. Fortunately, the integration of spatially-rich telemetry data into SCR can enable more complexity in the encounter model and improved inferences about density (Royle et al., 2013c; Sollmann et al., 2016; Linden et al., 2018), as well as extending the types of spatial encounter data that can be used (Sollmann et al., 2013; Tenan et al., 2017; Murphy et al., 2019). However, much of the focus has been on processing telemetry data to match standard SCR assumptions (e.g., temporal independence of detection locations), and, despite the obvious conceptual linkages, the formal integration of an explicit movement model within SCR is surprisingly rare.

Indeed, recognizing and implementing joint movement and SCR models represents the next frontier in spatial population ecology (Royle et al., 2013b; McClintock et al., 2021).

Non-Euclidean SCR models relax a restrictive, though statistically-convenient, assumption that animal home ranges are stationary and symmetrical (Royle et al. 2013a, Sutherland et al. 2015). The implications of this formulation have an even broader impact: the model offers an explicit statistical framework for understanding how landscape structure determines spatial population structure, which is fundamental to a wide range of ecological processes (Tischendorf and Fahrig, 2000; Shaw, 2020). This is accomplished through the direct estimation of landscape resistance, the mathematical inverse of which represents landscape connectivity, a central feature of all spatially-structured populations (Zeller et al., 2012).

Beyond the pervasiveness of structured landscapes, understanding the interplay of landscape connectivity and the emergent spatial population structure at broad spatial scales is especially critical as habitat conversion and fragmentation are increasing globally, with critical implications for the management and conservation of landscapes (Gupta et al., 2019) and for the viability of the majority of extant megafauna (Haddad et al., 2015). However, despite the central role of connectivity in modern conservation science (McRae et al., 2008; Rayfield et al., 2011), there remain few methods for empirically estimating its components, let alone simultaneously describing spatial population structure (Tischendorf et al., 2005; Zeller et al., 2012; Hanks and Hooten, 2013; Howell et al., 2018). By jointly estimating landscape resistance and population density, the non-Euclidean SCR model directly addresses these challenges using standard ecological encounter history data (Morin et al., 2017; Gupta et al., 2019).

It is perhaps optimistic to expect reliable estimation of cost function parameters

using only sparse encounter history data that arise from capture-recapture studies – particularly for elusive, low-density, and wide-ranging species for which estimates of connectivity hold most value (Creel et al., 2020; Crooks et al., 2011, 2017). Small sample sizes of coarse location data limit the space-use complexity of any SCR model, particularly those using non-Euclidean distance, where least-cost paths will have greater uncertainty as the distances in time and space increase between observations. This problem could be mitigated through the integration of fine-scale movement data (e.g., telemetry) and, to accommodate the inevitable temporal autocorrelation, an explicit movement model (Hooten et al., 2017). The integration of explicit movement in the non-Euclidean SCR model should help inform the direction and strength of cost parameters that are challenging to estimate using SCR data alone (Sutherland et al., 2015).

Here we develop a telemetry-integrated, non-Euclidean SCR model that incorporates an explicit movement model, in which the spatial encounter and telemetry likelihoods share parameters of a cost function. We use a straightforward implementation of a weighted distribution for movement (Johnson et al., 2008) where non-Euclidean distance defines availability. We evaluate our model using a novel simulation approach that derives spatial encounter and telemetry data from individual-based movement simulations, and analyze the resulting encounter histories using our integrated SCR-movement cost models, with standard SCR models used for benchmarking. We evaluate the performance of the model under a range of conditions designed to test sensitivity to data quality and the number of telemetered individuals. As a case study, we apply the model to data on an American black bear *(Ursus americanus)* population (Sun, 2014) and discuss these results in the context of the simulations.

3.3 Methods

3.3.1 Spatial capture-recapture with non-Euclidean distance

We begin with a brief description of a standard, constant-density SCR model with a binomial encounter model. Specifically, we have spatially-indexed encounters y_{ij} representing the number of detections of individual i at detector j across K surveys. With no survey-specific variation, we model the encounters as Binomial random variables having an encounter probability p_{ij} conditional on the latent activity center $\mathbf{s}_i$:

$$y_{ij}|\mathbf{s}_i \sim \text{Binomial}(K, p_{ij}).$$

We assume a homogeneous point process for $\mathbf{s}_i$ across the support of the state space $\mathcal{M}$ which will require integration during likelihood estimation. Encounter probability is a decreasing function of the Euclidean distance $\text{d}_{Euc}(\mathbf{x}_j, \mathbf{s}_i)$ between the activity center $\mathbf{s}_i$ and detector $\mathbf{x}_j$:

$$p_{ij} = p_0 \times \exp\left(-\frac{\text{d}_{Euc}(\mathbf{x}_j, \mathbf{s}_i)^2}{2\sigma^2_{\text{det}}} \right),$$

where p_0 is the baseline encounter probability and σ_{det} is the scale parameter of a Gaussian kernel (i.e., the half-normal detection function), both of which are parameters to be estimated. We refer to this model as M_{Euc}.

The encounter function of M_{Euc} is assumed to be proportional to the individual's home range, which implies that home ranges are stationary and symmetric. While statistically convenient, this core assumption is unlikely to hold in reality. Royle et al. (2013a) demonstrated how the Euclidean distance assumption can be relaxed with direct estimation of the parameter of a cost function by modeling encounter probabilities using estimated distances of least-cost paths, $\text{d}_{lcp}(\mathbf{x}_j, \mathbf{s}_i)$, rather than simple measures in Euclidean space. This distance substitution gives rise to a

model, which we refer to as M_{lcp}, where:

$$p_{ij} = p_0 \times \exp\left(- \frac{\mathrm{d}_{lcp}(\mathbf{x}_j, \mathbf{s}_i)^2}{2\sigma^2_{\mathrm{det}}} \right).$$

The cost distance is calculated as follows. First consider a discrete landscape $\mathcal{V}$ of some predefined resolution, where each pixel ν has associated coordinates $\boldsymbol{x}$ and a covariate value, $z(\nu)$. Any path P between two locations (e.g., $\boldsymbol{x}$ and $\boldsymbol{x}'$) consists of M transitions between pairwise adjacent pixels $\{\nu_0, \nu_1, \ldots, \nu_M\}$ and we define a cost function for moving from pixel to pixel as:

$$\mathrm{cost}(\nu_m, \nu_{m+1}) = \frac{\exp(\alpha_1 z(\nu_m)) + \exp(\alpha_1 z(\nu_{m+1}))}{2}$$

Here, α_1 is a cost parameter to be estimated. For a given value of α_1 and all L reasonable paths, the least cost path d_{lcp}, calculated using Dijkstra's Algorithm (Dijkstra et al., 1959), is:

$$\mathrm{d}_{lcp}(\boldsymbol{x}, \boldsymbol{x}') = \min_{P_1,\ldots,P_L} \sum_{m=0}^{M_L} \mathrm{cost}(\nu_m, \nu_{m+1}) \times \mathrm{d}_{Euc}(\nu_m, \nu_{m+1}).$$

Conceptually, the cost surface is a functional transformation of a spatially varying covariate surface using α_1. When $\alpha_1 = 0$, $\mathrm{d}_{lcp}(\mathbf{x}_j, \mathbf{s}_i) = \mathrm{d}_{euc}(\mathbf{x}_j, \mathbf{s}_i)$, and as α_1 increases, cost distance diverges from Euclidean distance. When informed by spatially structured individual encounter data, this divergence of distance metrics provides a mechanism for estimating α_1, i.e., finding the value of the cost parameter that best describes the spatial pattern of the encounter history data.

Regardless of the encounter function used, the latent activity centers must be integrated out of the joint likelihood (over the support of $\mathcal{M}$) to form the marginal likelihood of the encounter data, $[\boldsymbol{y}|\alpha_1, p_0, \sigma_{det}]$. Following Borchers and Efford (2008) we assume population size is a Poisson random variable such that the joint likelihood can enable direct estimation of density as the mean intensity, λ, of a

Poisson point process:

$$\mathcal{L}_{SCR}(\alpha_1, p_0, \sigma_{det}, \lambda|\boldsymbol{y}) = \left\{ \prod_{i=1}^{n} [\boldsymbol{y}_i|\alpha_1, p_0, \sigma_{det}] \right\} \Lambda^n \exp\left(-\Lambda(1-\pi_0)\right)$$

where $\Lambda = \lambda||\mathcal{M}|| = \mathrm{E}(N)$ and π_0 is the probability of an all-zero capture history.

3.3.2 Animal movement with non-Euclidean distance

Here we describe a simple approximation of a spatiotemporal point process model for movement (Hooten et al., 2017) where the probability of an observed location is determined by a redistribution kernel (Moorcroft and Barnett, 2008) conditional on the previous location (i.e., a first-order Markov process) and an attraction to the home range center. This type of weighted distribution approach allows for estimation of resource selection from telemetry (Johnson et al., 2008; Christ et al., 2008) though our model has reduced complexity and focuses on the availability in terms of cost distance. Critically, it is the cost parameter α_1 that is to be estimated from a collection of telemetry locations.

Consider an observed movement trajectory for individual i over T fixed-interval time steps, $\boldsymbol{\mu}_{it} = \boldsymbol{\mu}_{i1}, \boldsymbol{\mu}_{i2}, ..., \boldsymbol{\mu}_{iT}$, where we assume there is no observation error in the observed location (i.e., $u \equiv \mu$). To model movement between steps we define an availability function or redistribution kernel, $k(\boldsymbol{\mu}_{it}|\boldsymbol{\mu}_{i,t-1}, \theta)$, where the relative probability of a location $\boldsymbol{\mu}_{it}$ is a decreasing function of both the cost distance from the previous location $\boldsymbol{\mu}_{i,t-1}$ and the cost distance from the average of all locations $\bar{\boldsymbol{\mu}}_i$ for individual i:

$$k(\boldsymbol{\mu}_{it}|\boldsymbol{\mu}_{i,t-1}, \theta) = \exp\left(-\frac{d_{lcp}(\boldsymbol{\mu}_{it}, \boldsymbol{\mu}_{i,t-1})^2}{2\sigma^2_{step}} - \frac{d_{lcp}(\boldsymbol{\mu}_{it}, \bar{\boldsymbol{\mu}}_i)^2}{2\sigma^2_{home}} \right).$$

The location $\bar{\boldsymbol{\mu}}_i$ is treated as known and is used to capture the effect of central tendency in movement (in this case, within the home range). Spatial scale parameters

represent the variance in movement between observed locations in the movement trajectory (σ_{step}) and across the home range (σ_{home}). We note also that σ_{home} differs from σ_{det} from the SCR model in that there is not an additional detection process occurring for telemetry data as there is for SCR data. Based on this redistribution kernel, the normalized conditional probability of the observed locations becomes:

$$[\boldsymbol{\mu}_{it}|\boldsymbol{\mu}_{i,t-1},\theta] = \frac{k(\boldsymbol{\mu}_{it}|\boldsymbol{\mu}_{i,t-1},\theta)}{\int k(\boldsymbol{\mu}|\boldsymbol{\mu}_{i,t-1},\theta)d\boldsymbol{\mu}}$$

For our discrete landscape from earlier, the locations $\boldsymbol{\mu}$ are associated with pixels $\boldsymbol{\nu}$ and the integral in the denominator above becomes a summation across all $\boldsymbol{\nu} \in \mathcal{V}$:

$$[\nu_{it}|\nu_{i,t-1},\theta] = \frac{k(\nu_{it}|\nu_{i,t-1},\theta)}{\sum_{g=1}^{\mathcal{V}} k(\nu_g|\nu_{i,t-1},\theta)}$$

To accommodate an excess number of steps where pixel location does not change, potentially representing a behavioral state resulting in no movement, we model the probability $\Pr(\nu_{it} \neq \nu_{i,t-1}) = \psi$ that an individual moves from the last known location/pixel. The likelihood for the movement model therefore has the following form:

$$\mathcal{L}_{move}(\alpha_1,\sigma_{step},\sigma_{home},\psi;\boldsymbol{\nu}) = \prod_{i=1}^{N_{tel}}\prod_{t=2}^{T}\psi\left(\frac{k(\nu_{it}|\nu_{i,t-1},\theta)}{\sum_{g=1}^{\mathcal{V}^*} k(\nu_g|\nu_{i,t-1},\theta)}\right)^{I(\nu_{it}\neq\nu_{i,t-1})} \times (1-\psi)^{I(\nu_{it}=\nu_{i,t-1})}$$

Here, the denominator calculates availability for pixels $\nu_g \in \mathcal{V}^*$, representing all pixels in the discrete landscape not including $\nu_{i,t-1}$. Note, we do not model the first location ($t = 1$) under the assumption that this loss of information is small with a long time series of telemetry locations (Johnson et al., 2008), although the first location could be modeled as: $\mu_{i1} \sim Normal(\bar{\mu}_i, \sigma_{home})$, in discrete space.

3.3.3 Integrating the movement model into SCR

The joint SCR-movement likelihood combines the cost-based SCR likelihood and cost-based movement likelihood through shared estimation of the cost param-

eter α_1. We refer to this model as iM$_{lcp}$, where "i" stands for *integrated*:

$$\mathcal{L}_{iM_{lcp}}(\alpha_1, p_0, \sigma_{det}, \lambda, \sigma_{step}, \sigma_{home}, \psi; \boldsymbol{y}, \boldsymbol{\nu}) = \mathcal{L}_{SCR} \times \mathcal{L}_{move}.$$

Note that the joint estimation of cost assumes that the effect of landscape resistance is equivalent across scales of movement, whether steps between telemetry locations or the long-term utilization of a home range resulting in trap encounters. An assumption of consistency should be reasonable since cost paths are scaled by distance and calculated as accumulations of pixels with the same resolution.

The joint likelihood treats data from spatial encounter histories and telemetry as independent, i.e., there is no overlap in identities of passively-detected and telemetered individuals. Though this is not always the case in practice, it does create flexibility in the model to allow for sampling from different time periods, and importantly, should have little consequence on the primary biological inference objectives, namely: cost and density (Royle et al., 2013c).

The the joint likelihood function is provided in Appendix L. Model parameters are estimated using maximum likelihood methods (Borchers and Efford, 2008) in `R` (R Core Team, 2019). Least cost paths are computed using Dijkstra's Algorithm (Dijkstra et al., 1959) implemented in the `R` package `gdistance` (Etten, 2017).

3.3.4 Evaluation by simulation

Our primary evaluations of the model were conducted through simulation, testing sensitivity to data quality and comparing performance against the standard Euclidean and non-Euclidean SCR-only models.

We begin by defining the SCR state space ($\mathcal{M}$), which is the area within which all individual activity centers are simulated and the area for which density is estimated. We defined a discrete set of 25 x 25 points representing the centers of 1 x

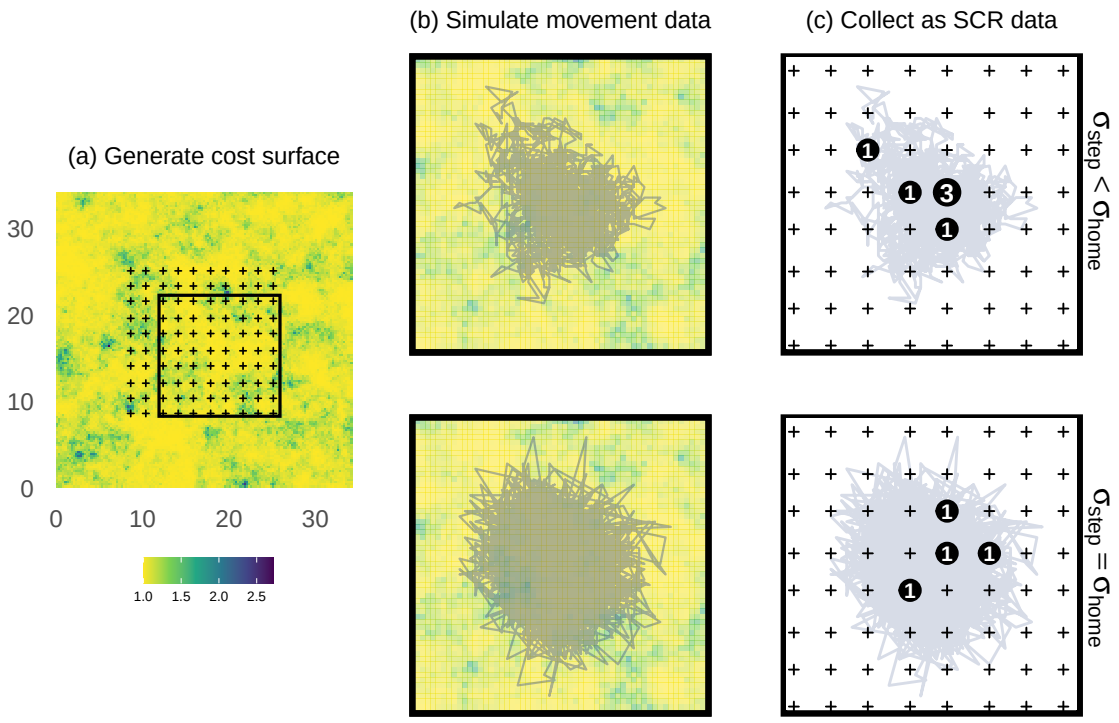

Figure 4. Outline of movement simulations

General outline of realistic movement simulations, with subfigures a-c representing consecutive steps in the workflow. (a) First, using a Gaussian field model we generate a cost surface, where yellow is low cost and dark purple is high cost. Proximity detectors (traps) are black crosses, and the black bounding box indicates the inset region shown in figures b and c. (b) Activity centers are generated randomly (not shown) and are used to simulate movement tracks, with one show in dark blue. These tracks differ in the value of σ_{step}, either small (top) or large (bottom), and relative to σ_{home} as shown on the right side. (c) Movement tracks are thinned and recorded as trap captures, collected as SCR data, as shown by the black circles with white text for the number of captures at that trap.

1 unit pixels. Activity centers can be located at any pixel within $\mathcal{M}$, but to allow unconstrained movement trajectories, especially for those activity centers located close to the edges, we define a second landscape, $\mathcal{V}$, comprising all theoretically possible movement trajectories given our simulation settings, such that $||\mathcal{V}|| > ||\mathcal{M}||$. We define $\mathcal{V}$ as a grid of 137 x 137 points representing the centers of each 0.25 unit x 0.25 unit pixels in the movement landscape, noting that the resolution of $\mathcal{V}$ is finer than $\mathcal{M}$ to allow finer scale movement trajectories that better approximate

continuous space. We generated a cost covariate landscape by assigning pixel values in $\mathcal{V}$ generated under a Gaussian random field model with weak autocorrelation (maximum autocorrelation range = 6 pixels), which was then scaled from 0 to 1 (Figure 4a). Gaussian random fields were generated in `R` using the package `NLMR` (Sciaini et al., 2018).

We simulated movement trajectories and subsequently derived encounter data as emergent properties of the movement data based on proximity to an array of traps. To do so, we randomly generated $N = 100$ activity centers, with replacement, from $\mathcal{M}$. Then, using the model described above and the availability landscape $\mathcal{V}$, we simulated movement tracks for every individual using the activity center as the initial location. Movement paths were generated for $T = 2160$ times steps (or "fixes") approximating hourly fixes over a hypothetical three-month period (Figure 4b). Parameters of the data-generating movement model were chosen to reflect biologically-realistic parameter combinations and are provided in Table 2.

Spatial encounter histories were then derived from the simulated movement paths based on a 10 x 10 trapping grid located in the center of the study area (Figure 4c). If a movement step contained a trap, it was 'detected' with probability 0.1, where the binomial thinning ensures that the distribution of recaptures and spatial recaptures were realistic when compared to empirical SCR data (see Appendix J: Figure 9 for scenario-specific summaries of the spatial encounter histories). Per standard design recommendations, detectors spaced 1.75 units apart (approx. $2\sigma_{home}$) leaving a 3 unit ($3\sigma_{home}$) buffer to the edge of $\mathcal{M}$.

This approach of deriving spatial encounter histories from individual movement data differs from the traditional approach of simulating data directly from the SCR model. We believe our approach provides a valuable simulation framework for exploring SCR model performance and sensitivity under a range of potential

Table 2. Model parameters and simulation values

Parameters of the integrated model along with pre-defined values used for the simulation study, where parameter name, symbol, and biological interpretation are shown in each row. Numbers in bold highlight the main difference between the two simulation scenarios, small-σ_{step} and large-σ_{step}.

Parameter	Symbol	Small-σ_{step}	Large-σ_{step}	Biology
Cost	α_1	1	1	Resistance of the landscape to movement
Baseline detection	p_0	–	–	Detection probability at the activity center in the home range
Detection spatial scale	σ_{det}	2.5	2.5	Scaling factor determining shape of the detection kernel
Home range spatial scale	σ_{home}	2.5	2.5	Scaling factor determining size of the home range
Density	λ	0.16	0.16	Number of individuals in the state-space
Step distance spatial scale	σ_{step}	**0.625**	**2.5**	Scaling factor determining the size of the movement kernel
Probability moved	ψ	0.9	0.9	The probability that an individual moves, i.e., that it leaves the pixel

assumption violations and, as demonstrated here, movement characteristics.

We were interested in exploring model performance as a function of two features of telemetry data: the number of telemetered individuals used in the analysis, and ratio of the spatial variance parameters. First, to evaluate whether the performance of the integrated model was sensitive to the number of telemetered individuals included, we considered the inclusion of 1, 3, and 5 individuals. Second, we evaluated model performance with respect to the variance ratios, which we motivate conceptually as representing potential differences in either species propensity for movement or the telemetry fix rate. Specifically, we tested two ratios: the spatial scale of the steps σ_{step} being one quarter that of the home range space use parameter σ_{home}

(the "small-σ_{step} scenario", $\sigma_{step} = 0.25\sigma_{home}$), and σ_{step} and a "large-σ_{step} scenario" where the ratio was 1 ($\sigma_{step} = \sigma_{home}$).

We simulated 100 data sets under each σ_{step}-ratio scenario. For each, the activity centers, corresponding movement paths (i.e., telemetry data), and the cost surface were generated randomly. Individuals were subject to detection via the trapping array resulting in n individuals with non-zero encounter histories (i.e., SCR data). We analyzed the SCR and telemetry data using the integrated non-Euclidean model where the 1, 3, or 5 movement tracks were randomly selected from all 100 individuals. As a comparison, we also analyzed each data set without telemetry data, i.e., an SCR-only data set, using the non-Euclidean SCR model (M_{lcp}), which is a special case with 0 telemetered individuals. In summary, we fit two SCR variants – not integrated and integrated – to each data set, and for the integrated model consider 1, 3, and 5 telemetered individuals, i.e., 4 comparisons for each σ_{step}-ratio scenario. We note that, rather than model the detections as a function of the spatiotemporal proximity to the movement path as per the data generating procedure, SCR detections are modelled as function of distance to the latent activity center. The former requires an additional conditional structure and individuals that appear in both the telemetry and SCR data. While dependency is straightforward (e.g., Linden et al., 2018) a latent movement model for the encounter data would be a computationally intensive extension for our model, which we do not consider.

For the 8 simulation sets, we evaluated model performance by computing percent relative bias for each parameter. For model M_{lcp} in the large–σ_{step} scenario, we removed 4 simulation results due to poor model fit. All simulations and analyses were conducted in `R` version 4.0.3 (R Core Team, 2019), and are available in our Open Science Framework repository (Dupont et al., 2021b).

3.3.5 Case study: New York black bears

As a demonstration of the model in practice and its relative performance, we analyze data from a study of black bears in western New York (Sun, 2014). The data consists of 103 hair snares from a non-invasive genetic mark-recapture study, which resulted in spatial encounter histories for 33 individuals. In addition to the trapping, three bears were fitted with GPS collars in the same study area during the same time period (Figure 6a). These data are described in detail in Sun (2014).

We used percent forest ("% forest") as the cost covariate, but note that our model can easily accommodate the simultaneous estimation of multiple cost parameters associated with multiple landscapes. We derived % forest from the National Land Cover Database (NLCD) (Yang et al., 2018) by aggregating the 30-m raster to a 500 m resolution and calculating the proportion of forest pixels therein (Figure 6a). NLCD data were accessed using the `R` package FedData (Bocinsky, 2016). We defined the state-space as a 4968 km^2 rectangular area with a $3\sigma_{det}$ buffer around the trapping array, based on the σ_{det} estimated in previous analyses of these data (Royle et al., 2013c; Morin et al., 2017), and a resolution approximating σ_{step}, calculated from the raw fix locations using Euclidean distance.

We analyzed the bear data using the four models described above (see also: Appendix J: Table 9). Our primary comparison was between the existing model M_{lcp} and our integrated model, iM_{lcp}. For more thorough reference and comparison, we included all applicable null models. To explore the effect of individual variation, we also fit the integrated model including only a single individual at a time. These analyses were conducted in `R` and all code is provided in our Open Science Framework repository (Dupont et al., 2021b).

After obtaining model results, we followed the procedure of Morin et al. (2017)

to produce surfaces of realized density (i.e., the estimated density of activity centers conditional on the observed data) and density-weighted connectivity (i.e., expected pixel use based on estimated cost and weighted by realized density).

3.4 Results

3.4.1 Evaluation by simulation

We first report results from the small-σ_{step} scenario. The standard non-Euclidean SCR model (M_{lcp}) produced a biased estimate of cost (%RB =–17) with relatively low precision (Figure 5, Appendix J: Table 10). In contrast, and consistent with the remarkable robustness of SCR, estimates of density were unbiased (Figure 5). In terms of estimating the cost parameter, integrating telemetry data improved performance considerably: data from a single individual alone reduced the bias to just –3%) and markedly improved the precision (Figure 5, Appendix J: Table 10). Telemetry data from more individuals resulted in further improvements (Figure 5, Appendix J: Table 10).

Thus, the integrated model consistently produced unbiased and precise estimates of both the data-generating movement parameters and density (Figure 5 and Appendix J: Table 10). Moreover, the performance (both bias and precision) improved with the addition of more telemetered individuals (Figure 5 and Appendix J: Table 10).

In the large-σ_{step} scenario, and consistent with the small-σ_{step} scenario, estimates of density were unbiased irrespective of whether integration was used, although precision was marginally poorer (Figure 5). What was strikingly different in the large-σ_{step} scenario, where distances between locations were larger, and hence num-

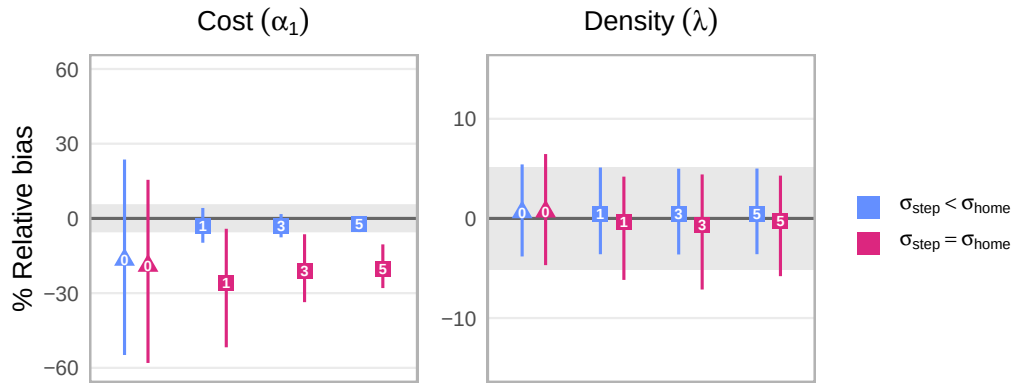

Figure 5. Results from simulation-based evaluations

Percent relative bias from simulation results. To illustrate estimator precision, vertical lines are 50% confidence intervals. 50% intervals are proportional to 95% intervals but offer a visual balance of bias and associated variance. Blue and fuschia shapes represent the small- and large-σ_{step} scenarios, respectively. Triangles represent model M_{lcp} and squares represent model iM_{lcp}, with an increasing number of telemetered individuals (inset number). Gray shaded area shows an acceptable amount of bias (±5%). Full results in Appendix J: Table 10.

ber of possible least cost paths was greater, was the inability to reliably estimate the cost parameter. In all cases, regardless of number of telemetered individuals included, the integrated model produced biased estimates of α_1, ranging from –26% to –20% (Figure 5, Appendix J: Table 10).

Table 3. Case study results

Results from American black bears analysis. From left to right: fitted model (see Methods: Case study for reference), all estimated parameters showing point estimates with standard errors directly below in parentheses, and the negative likelihood value.

model	α_1	log(λ)	σ_{det}	p_0	ψ	σ_{step}	σ_{home}	–loglik	AIC
M_{Euc}	-	–3.843	1.238	–2.768	-	-	-	124.32	254.64
	-	(0.254)	(0.139)	(0.361)	-	-	-		
M_{lcp}	0.226	–3.845	1.284	–2.736	-	-	-	124.27	256.54
	(0.753)	(0.254)	(0.209)	(0.377)	-	-	-		
iM_{Euc}	-	–3.843	1.238	–2.768	–0.666	–0.195	1.914	3874.45	7762.90
	-	(0.254)	(0.139)	(0.361)	(0.048)	(0.020)	(0.164)		
iM_{lcp}	–1.172	–3.841	0.812	–2.882	–0.666	–0.624	1.439	3835.42	7684.84
	(0.111)	(0.255)	(0.138)	(0.368)	(0.048)	(0.044)	(0.165)		

3.4.2 Case study: New York black bears

The results from the black bears case study were consistent with the simulation results suggesting no appreciable influence of cost or telemetry integration on density estimates (Table 3). Most importantly in the context of this study, the integrated model (iM_{lcp}) captured an effect of cost that was both stronger and more precise (estimate: –1.17 [SE: 0.11] vs 0.23 [SE: 0.75]; Table 3) than the SCR-only model (M_{lcp}). In fact, the standard model suggests no compelling evidence of a resistance effect due to forest whereas the integrated model reveals a strong and highly plausible negative effect of decreased resistance as forest cover increases.

Finally, fitting the integrated model to each of the 3 individuals separately indicated some heterogeneity in the estimated cost parameter but consistency in the direction and relative magnitude of the effects (Appendix J: Figure 10).

3.5 Discussion

Integrated data modeling represents a promising statistical framework for testing and advancing theory about ecological processes across multiple scales. Here, we illustrated an integration of trap encounters and telemetry locations that embedded an explicit movement model within spatial capture-recapture, leveraging their conceptual linkages. Together, movement data provide high-resolution information on step selection, and encounter history data allow for landscape-scale inference about spatial patterns of density, facilitating an improved understanding of how landscape-scale patterns of density and connectivity emerge from individual-level processes. Specifically, we provide an intuitive and flexible statistical framework for integrating telemetry and encounter history data; we highlight the value of an integrated modelling approach for improved inferences about the processes of density, space use, and landscape connectivity simultaneously.

Our method for integrating fine-scale telemetry data offers key performance improvements over the original non-Euclidean SCR model, and suggests that these improvements are related to the quantity and quality of the telemetry data. Here we considered including data on only a few individuals (1, 3, 5), and observed precision gains for the cost parameter that increased with the number of telemetered individuals added. In contrast, estimates of density did not benefit from the integration, likely due to the lack of spatial variation in density (e.g., Linden et al., 2018). Notably, performance improvements were only realized in scenarios where the spatial scale of consecutive fixes is much less than the spatial scale of individual space use; here, the integrated model is both unbiased and more precise than the non-integrated version. When the distance between fixes is relatively large, there is a loss of precision in both density and cost and estimates of cost are negatively

biased. The latter is likely due to uncertainty about the exact path taken between two locations at larger distances, which degrades the effect size of any cost relationships. In practice, the movement integrated model is recommended when telemetry data are available, motivated by both the conceptual and empirical advantages presented here, but we note that efforts should be made to select fix rates that result in short step lengths relative to the size of the home range (i.e., $\sigma_{step} < \sigma_{home}$). In general, and when individual heterogeneity in movement is low, fix rate appears to be more important than the number of collared individuals in terms of estimator improvements.

It is perhaps optimistic to expect accurate inference about fine-scale movement processes from coarse-scale SCR data (Sutherland et al., 2015). Indeed, this was evident from both the simulations and the bear case study, where cost parameters are not well estimated without the integration of high-resolution telemetry data via an explicit movement model. The addition of even a single telemetered individual resulted in the recovery of the cost parameters with accuracy and precision, a pattern noted in the context of SCR models that integrate resource selection functions (Royle et al., 2013c). Though it is not advisable to rely on a single animal for inferences due to individual heterogeneity in movement (Revilla and Wiegand, 2008), the bear case study did indicate consistency across individuals in the direction and magnitude of the cost estimate (Appendix J: Figure 10). Either way, the benefits of the non-Euclidean model may only be realized with large amounts of spatial recaptures (which is unlikely in practice), or through integrated approaches such as we develop here. Estimation of landscape connectivity is an important endeavour in wildlife conservation, yet is not well developed, especially using empirical observations (Zeller et al., 2012). Our integrated model provides an important formal statistical link between individual data and estimates of landscape connectivity

(Figure 6).

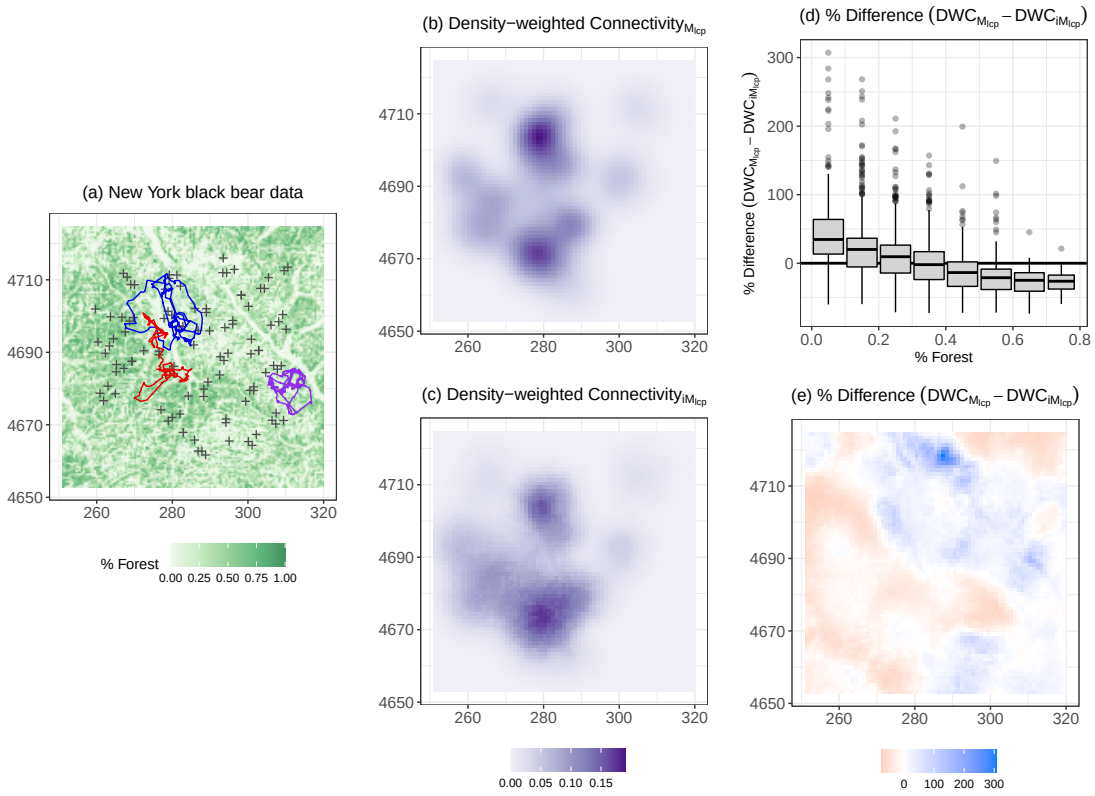

Figure 6. Case study data and results

Black bear data and results. (a) The % forest covariate surface, black crosses are hair snares and colored paths are movement tracks for three individuals. Density-weighted connectivity (DWC) from model M_{lcp} (c), and from model iM_{lcp} (d). Boxplot representation of the percent difference between DWC from model M_{lcp} and model iM_{lcp}. Finally, (e) is a spatial representation of the percent DWC differences (i.e., c minus d).

We demonstrated clear improvements in SCR model performance with telemetry integration, as has been shown previously (Royle et al., 2013c; Sollmann et al., 2013; Linden et al., 2018). These other approaches formalized telemetry integration by joint estimation of shared parameters, particularly the detection scale parameter (see our exploration on sharing the space use parameter within this model presented in Appendix K). Royle et al. (2013c) additionally relied on diluting or thinning their high-resolution telemetry data to match the coarser scale of spatial

encounter histories rather than integrating an explicit movement model. Though thinning telemetry data removes autocorrelation to satisfy independence assumptions, it removes by definition the conditional dependence in the observed locations that directly reflects fine-scale decisions about movement. Retaining the sequential structure of movement paths by explicitly modeling those decision processes yields improved insights about animal movement (McClintock et al., 2021). By also leveraging the less obvious connection between SCR and movement models, namely the existence of a shared cost function, we allowed those benefits to be realized at the population and landscape level.

In contrast to conventional SCR simulation approaches, where encounter history data are simulated from the expected encounter model, we allowed encounter data to emerge as a property of simulated animal movement paths. We did not fit this data-generating model (due to the additional computational complexity), but used it for simulation as it offers some specific benefits. Movement paths are allowed to evolve from a wide range of well-established and flexible models that can be parameterized according to known species characteristics (Hooten et al., 2017). As a result, the emergent encounter data arise as a property of the movement model rather than the specified SCR model, arguably a more preferable challenge for model evaluation with greater ecological realism. This provides a flexible simulation structure for better understanding how model performance varies under a wider range of scenarios that includes considerations specific to the system, the species, or sampling. For example, we investigated model performance for two scenarios related to the ratio of the variance parameters σ_{step} and σ_{home}, which we motivate as being representative of potential differences in either species propensity for movement or the telemetry fix rate, but ultimately that give rise to different data structures (Figure 4b). In terms of density estimation, these differences had

little effect, apart from a marginal increase in precision. However, the models were sensitive to the characteristics of the data when estimating parameters related to movement. In the large ratio scenario, estimates of the cost parameter (α_1, Figure 5) and the movement scale parameter (σ_{step}, Appendix J: Table 10) were systematically biased low whereas the estimated SCR spatial scale parameter (σ_{det}) was always higher (Appendix J: Table 10).

Our simulation structure also allowed us to consider the consequences of incorrect parameter sharing, motivated by the high prevalence in the SCR literature of interpreting σ_{det} as an estimate of space use, rather than explicitly defining it as a detection range which is *proportional* to space use. Under this interpretation, there is likely to be a temptation to assume some equivalency (i.e, $\sigma_{det} = \sigma_{home}$, which we show produces negatively biased estimates of cost and density (Appendix K). Depending on the assumed model of individual movement, the variance of a bivariate normal kernel estimated from the location data may not be equivalent to that as estimated by the encounter histories using SCR. Ultimately, our simulations highlight that strict biological interpretations of movement parameters from standard SCR models should generally be avoided.

Our movement model addresses a fraction of what is possible among the diverse and complex animal movement processes for which models have been developed (Hooten et al., 2017). Our implementation of the weighted distribution approach makes a restrictive assumption that telemetry fixes have regular time intervals, and while it did not affect simulations it did likely affect the black bear case study. We expect that the consequence was similar to that for spatial encounter data alone, namely that landscape cost was slightly underestimated when hourly steps were diluted by steps with longer time periods between. This problem could be explored through simulation or addressed more comprehensively by the movement model

likelihood (Johnson et al., 2008). Relatedly, the first order Markov relationship between steps ignores the history of movements prior to the step in time $t-1$, and this covariance across time could be estimated (Johnson et al., 2008). Finally, we did not model the latent movement of individuals from the encounter data (i.e., the data-generating model) which could inform far more spatial and temporal complexity in the encounter model at the expense of computational burdens. Our approach was intended to illustrate the clear benefits of including a straightforward movement model when integrating telemetry into SCR using a likelihood approach that can be easily implemented by extending existing software (e.g., Sutherland et al., 2019). Current applications of integrated models are simply the tip of the iceberg and will expand as technology and algorithm development continue (McClintock et al., 2021).

CHAPTER 4

DISCUSSION

4.1 Summary and synthesis

Spatial capture-recapture is a promising approach to study spatial and temporal dynamics in biological populations, creating a powerful opportunity to leverage the mechanistic model for more detailed ecological learning about spatial processes, as demonstrated in this thesis. These advanced methods allow us to formalize hypotheses about landscape connectivity (Sutherland et al., 2015), spatial structure (Royle et al., 2015), resource selection (Royle et al., 2013c), and demographic rates (Gardner et al., 2010). The challenge of utilizing these advanced methods relies on improving the quality of the data available, which is readily achievable through a focus on sampling design and data integration.

In Chapter 2, I illustrated an algorithmic approach to optimize sampling design using criteria directly related to the inner workings of the SCR model (Dupont et al., 2021c). Compared to the pre-existing recommendations that focus on grid-based designs, these optimal designs performed just as well but with far more flexibility to account for available resources and challenging sampling scenarios. This approach represents a fundamental shift in the design conceptualization process for SCR by formalizing it within an optimization procedure applicable to any species in any

study area. Relieving the onus of trap placement selection from practitioners and researchers with this solution will result in better designs with expected increases in data quality, especially critical for species of conservation concern on a global scale.

In Chapter 3, I demonstrated the integration of an explicit movement model into the non-Euclidean SCR model to achieve a specific inferential objective, the estimation of landscape connectivity (Dupont et al., 2021a). Data integration shares a similar motivation to sampling design but requires more nuanced approaches for modeling. For example, previous demonstrations use only a fraction of the available data to avoid addressing assumptions of independence. In contrast, the approach I presented is one of the first descriptions of a method for integrating an explicit model that allows for the comprehensive inclusion of auxiliary information into the SCR model. The integrated model showed significant improvement over the standard, non-integrated SCR model in terms of bias and precision for estimating landscape resistance and connectivity. In doing so, the modeling framework provides a more insightful approach grounded in the ecological process of animal movement.

4.2 Future directions

These novel approaches to sampling design and data integration represent essential steps in advancing SCR while creating opportunities for more sophisticated methods. Some extensions of this work have already commenced. For example, Durbach et al. (2021) formulated a design criterion based on a metric proposed by Efford and Boulanger (2019) and adopted the algorithmic approach from Chapter 2 to optimize trap locations accordingly. This new criterion could generate designs

that accounted for expected spatial variation in density, representing a substantial advancement for SCR and highlights the flexibility of the framework in Chapter 2. Other adaptations involving additional criteria should be pursued, such as for multiple species sampling. I have begun preliminary work with colleagues on this: the approach is made straightforward by the algorithm's flexibility, but it requires further evaluation before broad-scale field implementation. Improving sampling design in any of these directions, whether single- or multi-species, delivers more data to the SCR model and facilitates our ability to answer more sophisticated ecological questions.

Data integration offers opportunities for improvement that are equally exciting. First and foremost, more complicated movement models exist and could be employed, especially those designed to address more directly the typical nuances of telemetry data (such as irregular fixes or location error) or more complex processes of animal movement. One obvious opportunity would be a more direct treatment of the scales of selection (short-term steps and long-term home range). Although the integrated model in Chapter 3 relies on the assumption of equivalence between these two scales of selection with reasonable justification, there could be scenarios that violate this assumption. One method for addressing this would be the Markov chain Monte Carlo ("MCMC") step selection models developed by Michelot et al. (2019) (see also: Michelot et al., 2020). These models describe multi-scale selection resulting from the same underlying habitat selection process by leveraging the conceptual similarity between animal movement and MCMC algorithms. Regarding ecological advancements, a population model with an explicit movement model could present one approach for direct inference on differences in selection on territory versus during dispersal and relative effects of landscape characteristics during those distinct ecological processes (Cushman et al., 2013). Previous attempts have

been limited or have used ad hoc approaches rather than formal estimation procedures (Zeller et al., 2012).

Additional research opportunities exist at the crossroads of these two endeavors. Movement corridor design would be a natural fit, leveraging an extension of the spatial selection algorithm from Chapter 2 with the model for landscape resistance from Chapter 3. For example, previous research has focused on corridor design (Peterman, 2018) and reserve design (Gupta et al., 2019), but not through an integrated population-movement model. Similarly, species reintroduction programs focus on the ideas of optimal placement and connectivity (Rout et al., 2007), the decision-making process for which could benefit from the development of a formal optimization approach. Interestingly, such an approach could serve at the scale of a population or a metapopulation, utilizing a connectivity surface of either density-weighted connectivity (Morin et al., 2017) or occupancy-weighted connectivity (Meyer et al., 2020). Finally, pairing multi-species sampling methods with species-specific models to estimate connectivity could illuminate new insights into how communities share corridors for movement and whether there is convergence on specific habitat types and structures across species. This could lead to advances in both community and landscape ecology, especially to understand better the utility of umbrella species in conservation planning (Simberloff, 1998).

4.3 On developing methods for ecology

I close by providing two broader conclusions about developing statistical methods for ecology. First, simulation-based evaluation is integral to this process, but the circularity of its use can, unfortunately, be understated. In Chapter 2, I adopted the typical approach where the data-generating model matched the model

for estimation. In contrast, my evaluations in Chapter 3 employed a unique data-generating model, considerably more sophisticated than the estimation model. The estimation model served as an approximation for the data-generating model, preferable for its relative simplicity and efficiency. This approach to model evaluation highlighted critical nuances of the model, which could be investigated more directly given the quality of the simulated data. For example, highlighting the nuances of direct interpretation of SCR model parameters would not have been possible under the typical approach. Though in scenarios using simpler models, as in Chapter 2, the circularity between simulation and estimation should be less problematic.

A counterargument related to that circularity leads me to my second conclusion. In Chapter 3, the existence of the more sophisticated model for data generation means that it could be adapted for estimation, likely improving inference. This assertation is undoubtedly true. Readily overlooked, though, is that statistical solutions should be intuitive at some level to facilitate their widespread use. The SCR model is a perfect example: it is a conceptually simple hierarchical model, reasonably straightforward to specify in software like JAGS, making it easily understood by a diverse pool of potential users. Importantly, such simplicity should also be matched in SCR model extensions when possible. Chapter 2 leveraged the conceptual clarity of the SCR model, reflected by the model-based criteria, rather than utilizing an alternative approach that could be less familiar to users. Chapter 3 used a relatively simple movement model but leveraged that model's similarity to the components of the SCR model. Indeed, the capabilities of the most advanced models are impressive – and in many ways, the work presented in Chapters 2 and 3 would be considered advanced by some audiences. Intuitive approaches, though, are critical in encouraging the widespread use of new methods, promoting the advancement of ecological inference.

APPENDIX A

GENETIC ALGORITHM DETAILS

The design-generation function discussed here, *scrdesignGA()*, serves as a wrapper around the *k-of-n* genetic algorithm implemented by the R function, *kofnGA()* (Wolters, 2015). The genetic algorithm that drives the design-generation process conducts a random search for solutions until it converges on a near-optimal solution. This algorithm is in the broader class of evolutionary algorithms, and following this, its components are named in similar terms.

The algorithm starts by generating a random set of possible solutions, which is the initial population of 'offspring' (in this context, designs). The size of the population is constant throughout the process, and is determined by the user via the *popsize* parameter. The offspring are then evaluated according to an objective function, resulting in a 'fitness' value for each offspring. Some proportion of the offspring are then allowed to 'breed' – determined via the *keepbest* parameter – which in effect is a mechanism that shares the 'genetic' material of the most fit offspring (i.e., the beneficial components) in order to make the next 'generation' of offspring. This process repeats for some number of generations, predefined by the *ngen* parameter. The *k-of-n* component of the specific genetic algorithm that we employ adds location-switching functionality that allows flexibility such that some number of traps, k, can be selected from n possible trap locations, fitting neatly into the SCR design process.

While the three parameters mentioned above are important, we have found that only *ngen* is critical for parameter tuning of the algorithm. The algorithm should be parameterized to allow for a sufficient number of generations to reach (near-)convergence, which can be found via visual inspection by plotting the output object as illustrated in Appendix D. Beyond that, and assuming the algorithm has converged on a (near-)optimal design, we have found that the genetic algorithm is not particularly sensitive to sensible values for the other two tuning parameters, *popsize* and *keepbest*, with the biggest differences relating to efficiency, which is case-specific and not critical overall.

See the main text of this manuscript for details regarding the required SCR components for design-generation, and for more details on the genetic algorithm we direct the reader to Wolters (2015), as cited in the main text.

APPENDIX B

REGULAR GEOMETRY DESIGNS

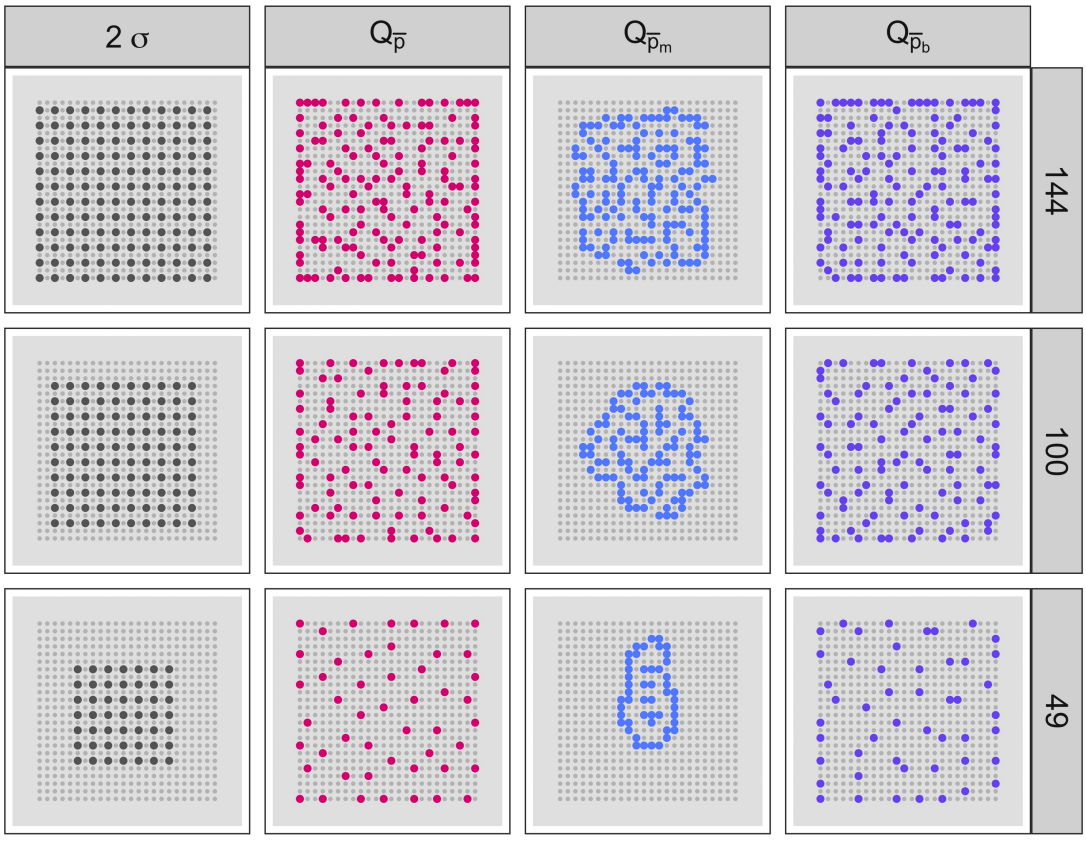

Figure 7. Regular geometry designs
Full set of designs in the regular study area, including the 2σ designs as well as the criteria-based designs. From top to bottom, rows represent 144, 100, and 49 traps included in the design, respectively.

A P P E N D I X C

VIGNETTE OF CHAPTER 2 SIMULATION DESIGN

Vignette of simulation structure. Each row details a single scenario, for which we generated 300 realizations of activity centers, generated detection histories for the specified sampling design, and estimated SCR parameters using two models.

Table 4. Vignette of simulation structure in the regular study area

Regular study area

Scenario	Geometry	Effort	Design	Density	Model
1	regular	144	2σ	uniform	$d.$
2				weak	$d.; d_s$
3				strong	$d.; d_s$
4		100		uniform	$d.$
5				weak	$d.; d_s$
6				strong	$d.; d_s$
7		49		uniform	$d.$
8				weak	$d.; d_s$
9				strong	$d.; d_s$
10		144	$Q_{\bar{p}}$	uniform	$d.$
11				weak	$d.; d_s$
12				strong	$d.; d_s$
13		100		uniform	$d.$
14				weak	$d.; d_s$
15				strong	$d.; d_s$
16		49		uniform	$d.$
17				weak	$d.; d_s$
18				strong	$d.; d_s$
19		144	$Q_{\bar{p}_m}$	uniform	$d.$
20				weak	$d.; d_s$
21				strong	$d.; d_s$
22		100		uniform	$d.$
23				weak	$d.; d_s$
24				strong	$d.; d_s$
25		49		uniform	$d.$
26				weak	$d.; d_s$
27				strong	$d.; d_s$
28		144	$Q_{\bar{p}_b}$	uniform	$d.$
29				weak	$d.; d_s$
30				strong	$d.; d_s$
31		100		uniform	$d.$
32				weak	$d.; d_s$
33				strong	$d.; d_s$
34		49		uniform	$d.$
35				weak	$d.; d_s$
36				strong	$d.; d_s$

Table 5. Vignette of simulation structure in the irregular study area

Irregular study area					
Scenario	Geometry	Effort	Design	Density	Model
37	irregular	144	$Q_{\bar{p}}$	uniform	$d.$
38				weak	$d.; d_s$
39				strong	$d.; d_s$
40		100		uniform	$d.$
41				weak	$d.; d_s$
42				strong	$d.; d_s$
43		49		uniform	$d.$
44				weak	$d.; d_s$
45				strong	$d.; d_s$
46		144	$Q_{\bar{p}_m}$	uniform	$d.$
47				weak	$d.; d_s$
48				strong	$d.; d_s$
49		100		uniform	$d.$
50				weak	$d.; d_s$
51				strong	$d.; d_s$
52		49		uniform	$d.$
53				weak	$d.; d_s$
54				strong	$d.; d_s$
55		144	$Q_{\bar{p}_b}$	uniform	$d.$
56				weak	$d.; d_s$
57				strong	$d.; d_s$
58		100		uniform	$d.$
59				weak	$d.; d_s$
60				strong	$d.; d_s$
61		49		uniform	$d.$
62				weak	$d.; d_s$
63				strong	$d.; d_s$

APPENDIX D

EXAMPLE CODE FOR GENERATING DESIGNS

Example code implemented in R to generate and plot designs using the optimal sampling design algorithm `scrdesignGA()` for the area of interest in Pakistan. For more information on spatial data preparation and design generation, please see:

`https://bookdown.org/chrissuthy/SCR-design-book/`

```
library(oSCR)
library(kofnGA)
#----Load data from Pakistan----
data(pakistan)
# This loads the statespace:  pakSS
# and the possible trap locations:  pakTT
#----Plot this data----
plot(pakSS, asp=1, pch=16, col="grey")
points(pakTT, pch=20, cex = 0.5)
#----Run the design-finding algo----
testdesign <- scrdesignGA(
statespace = pakSS, alltraps = pakTT, # Study area components
ntraps = 25, # Number of available traps
beta0 = 0.2*5, # Expected data (where beta0 = log(g0 * k))
sigma = 3, crit = 1, # Expected data
popsize = 50, keepbest = 5, ngen = 50) # GA settings
 # FOR BEST DESIGNS: ngen should be around 1500+
 # See Appendix A for a brief discussion of the tuning parameters
#----Plot the results----
par(mfrow=c(1,3))  # Setup plotting area
plot(testdesign, which=4)  # plots all 3 diagnostic plots
par(mfrow=c(1,1))  # Reset plotting area
```

APPENDIX E

COEFFICIENT OF VARIATION SIMULATION RESULTS

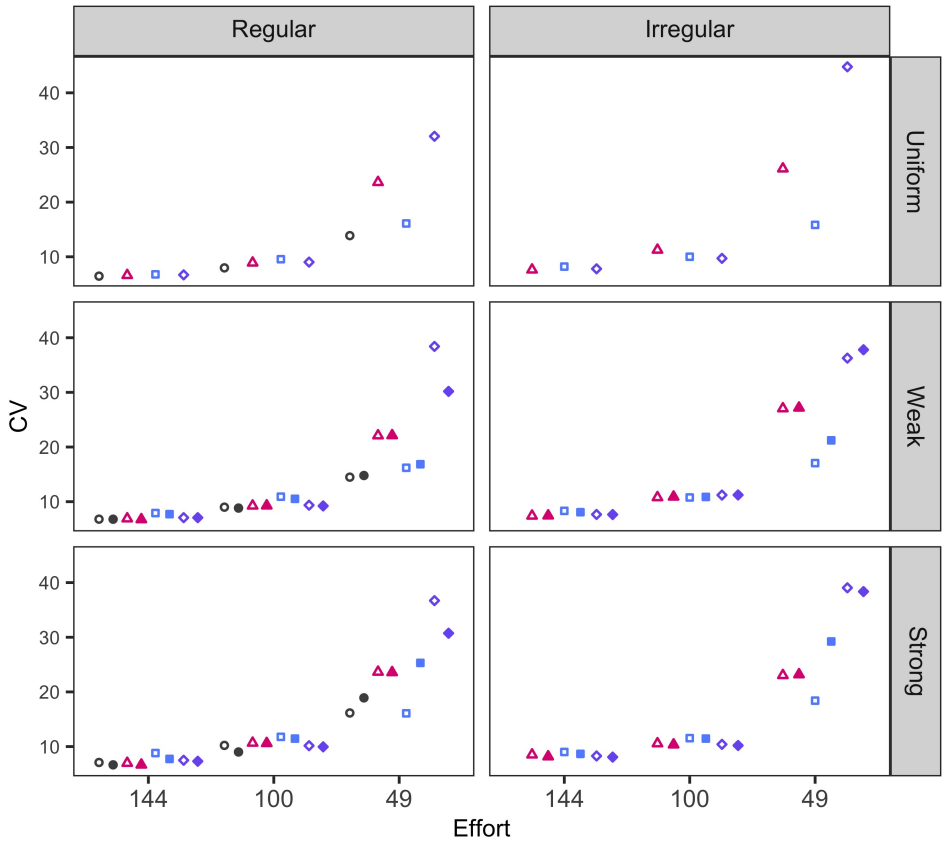

Figure 8. Coefficient of variation from simulation evaluations
Coefficient of variation (precision; CV) of estimates of total abundance from the four tested sampling designs under three levels of effort on three density surfaces within two geometries, where estimates are the result of one of two SCR models: density invariant ($d_.$, open shapes) or density-varying (d_s, closed shapes). The four designs – 2σ, $Q_{\bar{p}}$, $Q_{\bar{p}_m}$, $Q_{\bar{p}_b}$ – are represented by the four shapes: circles, triangles, squares, and diamonds respectively.

A P P E N D I X F

COEFFICIENT OF VARIATION TABLE

Coefficient of variation (precision; CV) of baseline detection (g_0), space use (σ) and total abundance (EN) for each of the 63 simulation scenarios, in which we varied the design criteria *(Design)*, the shape and accessibility of the landscape *(Geometry, Regular or Irregular)*, the number of traps *(Effort)*, and the underlying density patterns *(Density)*. Results from the null model ($d.$) are reported for all scenarios, and models from the data generating density model (d_s) are reported for scenarios with spatially varying density.

Table 6. Coefficient of variation from simulation evaluations

			Regular						Irregular					
			g_0		σ		EN		g_0		σ		EN	
Effort	Density	Design	*d.*	d_s	*d.*	d_s	*d.*	d_s	*d.*	d_s	*d.*	d_s	*d.*	d_s
49	uniform	2σ	19.68	–	7.28	–	13.87	–	–	–	–	–	–	–
		$Q_{\bar{p}}$	22.82	–	9.56	–	23.64	–	25.34	–	11.10	–	26.12	–
		$Q_{\bar{p}_m}$	16.19	–	7.61	–	16.10	–	18.31	–	8.14	–	15.84	–
		$Q_{\bar{p}_b}$	22.40	–	12.97	–	32.05	–	22.01	–	14.40	–	44.76	–
	weak	2σ	18.29	18.29	7.33	7.35	14.49	14.79	–	–	–	–	–	–
		$Q_{\bar{p}}$	23.05	23.05	9.86	9.87	22.09	22.12	21.94	21.94	11.52	11.45	27.04	27.17
		$Q_{\bar{p}_m}$	14.91	14.91	6.48	6.49	16.20	16.85	18.80	18.80	8.05	8.04	17.06	21.22
		$Q_{\bar{p}_b}$	21.20	21.20	12.59	12.20	38.41	30.18	25.28	25.28	13.85	14.01	36.26	37.80
	strong	2σ	19.24	19.24	7.62	7.61	16.15	18.91	–	–	–	–	–	–
		$Q_{\bar{p}}$	22.45	22.45	10.27	10.24	23.65	23.57	24.71	24.71	10.58	10.59	23.03	23.18
		$Q_{\bar{p}_m}$	17.43	17.43	7.59	7.63	16.08	25.31	19.90	19.90	8.07	8.06	18.40	29.23
		$Q_{\bar{p}_b}$	21.21	21.21	12.88	12.78	36.70	30.74	24.21	24.21	15.01	14.90	39.05	38.36
100	uniform	2σ	12.77	–	5.02	–	7.97	–	–	–	–	–	–	–
		$Q_{\bar{p}}$	13.02	–	6.02	–	8.90	–	14.50	–	6.07	–	11.26	–
		$Q_{\bar{p}_m}$	11.02	–	4.54	–	9.56	–	13.00	–	5.49	–	10.01	–
		$Q_{\bar{p}_b}$	13.74	–	5.51	–	9.03	–	13.74	–	5.87	–	9.71	–
	weak	2σ	12.72	12.72	5.16	5.16	9.00	8.83	–	–	–	–	–	–
		$Q_{\bar{p}}$	13.55	13.55	5.75	5.71	9.25	9.27	14.15	14.15	6.30	6.29	10.78	10.88
		$Q_{\bar{p}_m}$	12.42	12.42	5.02	5.03	10.92	10.52	13.99	13.99	5.84	5.84	10.77	10.87
		$Q_{\bar{p}_b}$	14.82	14.82	5.79	5.76	9.36	9.22	14.99	14.99	6.33	6.31	11.21	11.23
	strong	2σ	13.20	13.20	5.28	5.28	10.22	9.01	–	–	–	–	–	–
		$Q_{\bar{p}}$	14.79	14.79	6.05	6.07	10.66	10.60	14.77	14.77	6.27	6.25	10.56	10.35
		$Q_{\bar{p}_m}$	10.74	10.74	4.69	4.69	11.78	11.46	12.42	12.42	5.33	5.33	11.54	11.46
		$Q_{\bar{p}_b}$	14.74	14.74	5.58	5.57	10.16	9.95	15.18	15.18	6.80	6.79	10.41	10.19
144	uniform	2σ	10.86	–	4.24	–	6.46	–	–	–	–	–	–	–
		$Q_{\bar{p}}$	10.50	–	4.14	–	6.62	–	10.78	–	4.70	–	7.63	–
		$Q_{\bar{p}_m}$	9.68	–	3.95	–	6.78	–	10.69	–	4.20	–	8.21	–
		$Q_{\bar{p}_b}$	11.29	–	4.43	–	6.69	–	11.66	–	4.74	–	7.80	–
	weak	2σ	10.42	10.42	3.94	3.95	6.80	6.79	–	–	–	–	–	–
		$Q_{\bar{p}}$	11.00	11.00	4.42	4.43	6.90	6.75	11.58	11.58	4.81	4.83	7.41	7.44
		$Q_{\bar{p}_m}$	8.97	8.97	3.88	3.88	7.92	7.72	10.78	10.78	4.15	4.14	8.32	8.08
		$Q_{\bar{p}_b}$	10.20	10.20	4.17	4.15	7.08	7.08	11.91	11.91	4.94	4.93	7.66	7.65
	strong	2σ	10.32	10.32	4.20	4.21	7.09	6.65	–	–	–	–	–	–
		$Q_{\bar{p}}$	9.83	9.83	4.46	4.45	6.98	6.66	11.23	11.23	4.58	4.56	8.51	8.17
		$Q_{\bar{p}_m}$	9.47	9.47	4.04	4.03	8.81	7.74	10.65	10.65	4.54	4.54	9.01	8.66
		$Q_{\bar{p}_b}$	11.56	11.56	4.34	4.32	7.50	7.31	11.97	11.97	4.48	4.49	8.28	8.07

A P P E N D I X G

SCALED ROOT MEAN SQUARE ERROR TABLE

Scaled root mean square error (accuracy; SRMSE) of baseline detection (g_0), space use (σ) and total abundance (EN) for each of the 63 simulation scenarios, in which we varied the design criteria *(Design)*, the shape and accessibility of the landscape *(Geometry, Regular or Irregular)*, the number of traps *(Effort)*, and the underlying density patterns *(Density)*. Results from the null model ($d.$) are reported for all scenarios, and models from the data generating density model (d_s) are reported for scenarios with spatially varying density.

Table 7. Scaled root mean square error from simulation evaluations

			Regular						Irregular					
			g_0		σ		EN		g_0		σ		EN	
Effort	Density	Design	$d.$	d_s	$d.$	d_s	$d.$	d_s	$d.$	d_s	$d.$	d_s	$d.$	d_s
49	uniform	2σ	0.20	–	0.07	–	0.14	–	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.23	–	0.10	–	0.26	–	0.26	–	0.11	–	0.29	–
		$Q_{\bar{p}_m}$	0.16	–	0.08	–	0.16	–	0.19	–	0.08	–	0.16	–
		$Q_{\bar{p}_b}$	0.22	–	0.13	–	0.39	–	0.23	–	0.14	–	0.56	–
	weak	2σ	0.19	0.19	0.07	0.07	0.14	0.15	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.23	0.23	0.10	0.10	0.24	0.24	0.22	0.22	0.11	0.11	0.31	0.31
		$Q_{\bar{p}_m}$	0.15	0.15	0.06	0.06	0.16	0.17	0.19	0.19	0.08	0.08	0.17	0.22
		$Q_{\bar{p}_b}$	0.21	0.21	0.12	0.12	0.47	0.38	0.25	0.25	0.14	0.14	0.43	0.45
	strong	2σ	0.20	0.20	0.08	0.08	0.17	0.20	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.23	0.23	0.10	0.10	0.26	0.26	0.25	0.25	0.11	0.11	0.25	0.25
		$Q_{\bar{p}_m}$	0.18	0.18	0.08	0.08	0.16	0.28	0.20	0.20	0.08	0.08	0.19	0.33
		$Q_{\bar{p}_b}$	0.21	0.21	0.13	0.13	0.44	0.38	0.25	0.25	0.15	0.15	0.47	0.47
100	uniform	2σ	0.13	–	0.05	–	0.08	–	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.14	–	0.06	–	0.09	–	0.15	–	0.06	–	0.12	–
		$Q_{\bar{p}_m}$	0.11	–	0.05	–	0.10	–	0.13	–	0.05	–	0.10	–
		$Q_{\bar{p}_b}$	0.14	–	0.06	–	0.09	–	0.14	–	0.06	–	0.10	–
	weak	2σ	0.13	0.13	0.05	0.05	0.09	0.09	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.14	0.14	0.06	0.06	0.09	0.09	0.14	0.14	0.06	0.06	0.11	0.11
		$Q_{\bar{p}_m}$	0.13	0.13	0.05	0.05	0.11	0.11	0.14	0.14	0.06	0.06	0.11	0.11
		$Q_{\bar{p}_b}$	0.15	0.15	0.06	0.06	0.09	0.09	0.15	0.15	0.06	0.06	0.11	0.11
	strong	2σ	0.13	0.13	0.05	0.05	0.10	0.09	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.15	0.15	0.06	0.06	0.11	0.11	0.15	0.15	0.06	0.06	0.11	0.11
		$Q_{\bar{p}_m}$	0.11	0.11	0.05	0.05	0.12	0.12	0.13	0.13	0.05	0.05	0.11	0.12
		$Q_{\bar{p}_b}$	0.15	0.15	0.06	0.06	0.10	0.10	0.15	0.15	0.07	0.07	0.10	0.10
144	uniform	2σ	0.11	–	0.04	–	0.06	–	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.10	–	0.04	–	0.07	–	0.11	–	0.05	–	0.08	–
		$Q_{\bar{p}_m}$	0.10	–	0.04	–	0.07	–	0.11	–	0.04	–	0.08	–
		$Q_{\bar{p}_b}$	0.11	–	0.04	–	0.07	–	0.12	–	0.05	–	0.08	–
	weak	2σ	0.10	0.10	0.04	0.04	0.07	0.07	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.11	0.11	0.04	0.04	0.07	0.07	0.12	0.12	0.05	0.05	0.07	0.07
		$Q_{\bar{p}_m}$	0.09	0.09	0.04	0.04	0.08	0.08	0.11	0.11	0.04	0.04	0.08	0.08
		$Q_{\bar{p}_b}$	0.10	0.10	0.04	0.04	0.07	0.07	0.12	0.12	0.05	0.05	0.08	0.08
	strong	2σ	0.10	0.10	0.04	0.04	0.07	0.07	–	–	–	–	–	–
		$Q_{\bar{p}}$	0.10	0.10	0.04	0.04	0.07	0.07	0.11	0.11	0.05	0.05	0.09	0.08
		$Q_{\bar{p}_m}$	0.09	0.09	0.04	0.04	0.09	0.08	0.11	0.11	0.05	0.05	0.09	0.09
		$Q_{\bar{p}_b}$	0.12	0.12	0.04	0.04	0.07	0.07	0.12	0.12	0.04	0.04	0.08	0.08

A P P E N D I X H

TABLE OF PROBLEMATIC SIMULATIONS

Table 8. Problematic simulations reporting

Percent of simulations failed for each scenario due to a lack of spatial recaptures. For each scenario, we simulated encounter histories until 300 of those were acceptable (i.e., included at least one individual captured in more than one trap) and recorded the number of times that threshold was not met, which we recorded as a 'failure', expressed here as a percentage failed across all simulations within each scenario. Scenarios without failures are not included.

Scenario	Geometry	Design	Density	Effort	% Failed
28	regular	$Q_{\bar{p}}$	uniform	49	3.23
29	regular	$Q_{\bar{p}}$	weak	49	3.85
30	regular	$Q_{\bar{p}}$	strong	49	4.46
34	regular	$Q_{\bar{p}_b}$	uniform	49	0.33
35	regular	$Q_{\bar{p}_b}$	weak	49	0.33
55	irregular	$Q_{\bar{p}}$	uniform	49	1.96
56	irregular	$Q_{\bar{p}}$	weak	49	1.32
57	irregular	$Q_{\bar{p}}$	strong	49	1.64

A P P E N D I X I

SIMULATION CODE FOR CHAPTER 2

Code and data used for the analysis are publicly available:

`https://github.com/GatesDupont/scr_design_sims`.

The R code, found in the R subdirectory, operates within the file structure of this repository (outlined below). Simulations are performed in the file `6_sims.r` , which pulls other files from the directory that are conveniently compiled into the R data file: `workspace/sims_ws.RData`. The simulations file is currently parameterized to run all simulations simultaneously, and for sake of efficiency, the simulations are distributed across several cores.

File	Description
R	Sub-directory containing code for design generations and simulations.
R/0_functions.R	R code containing functions for simulations, including the simulator() function, which contains the data-generating model.
R/1_SS_regular.R	R code to generate the regular, square statespace.
R/2_SS_irregular.R	R code to generate the irregular statespace.
R/3_designs_regular.R	R code to generate the designs for the regular area.
R/4_designs_irregular.R	R code to generate the designs for the irregular area.
R/5_sims_gather_ws.R	R code for gathering data into a single workspace for the simulations.
R/6_sims.R	R code to run the simulations. Outputs csv file containing rows for each simulation run.
R/7_evaluate	R code to analyze the results of the simulation.
statespaces	Sub-directory containing csv files of the statespaces generated in R scripts 1 and 2.
traps	Sub-directory containing csv files of the possible trapping locations for the regular and irregular statespaces.
designs	Sub-directory containing csv files of all of the designs evaluated in the simulations.
workspace	Sub-directory containing an RData file of the workspace compiled in R script 5.
it_out	Sub-directory that is used to print out each iteration of the simulations.
output	Sub-directory that is used to print out plots of each simulation as well as the final results file.

A P P E N D I X J

ADDITIONAL MATERIALS FROM CHAPTER 3

Table 9. Model components for all models from the evaluations

Model	Data	Distance	Evaluations
M_{Euc}	SCR	Euclidean	Case study
M_{lcp}	SCR	Least-cost path	Case study + Simulations
iM_{Euc}	SCR+Telem	Euclidean	Case study
iM_{lcp}.	SCR+Telem	Least-cost path	Case study + Simulations

Table 10. Full simulation results of relevant parameters

Full simulation results of relevant parameters from both σ_{step} scenarios, using percent relative bias (%RB). Relative bias for σ_{det} and σ_{home} are both relative to the data-generating value for σ_{home}.

			Small-σ_{step}	Large-σ_{step}
Parameter	Model	n_{tel}	%RB	%RB
α_1	M	0	-17	-19
α_1	iM	1	-3	-26
α_1	iM	3	-3	-21
α_1	iM	5	-2	-20
λ	M	0	1	1
λ	iM	1	0	0
λ	iM	3	0	-1
λ	iM	5	0	0
σ_{det}	M	0	-31	-22
σ_{det}	iM	1	-29	-23
σ_{det}	iM	3	-29	-22
σ_{det}	iM	5	-28	-22
σ_{home}	iM	1	3	-1
σ_{home}	iM	3	3	-1
σ_{home}	iM	5	3	0
σ_{step}	iM	1	-1	-8
σ_{step}	iM	3	-1	-7
σ_{step}	iM	5	-1	-7
ψ	iM	1	0	0
ψ	iM	3	0	0
ψ	iM	5	0	0

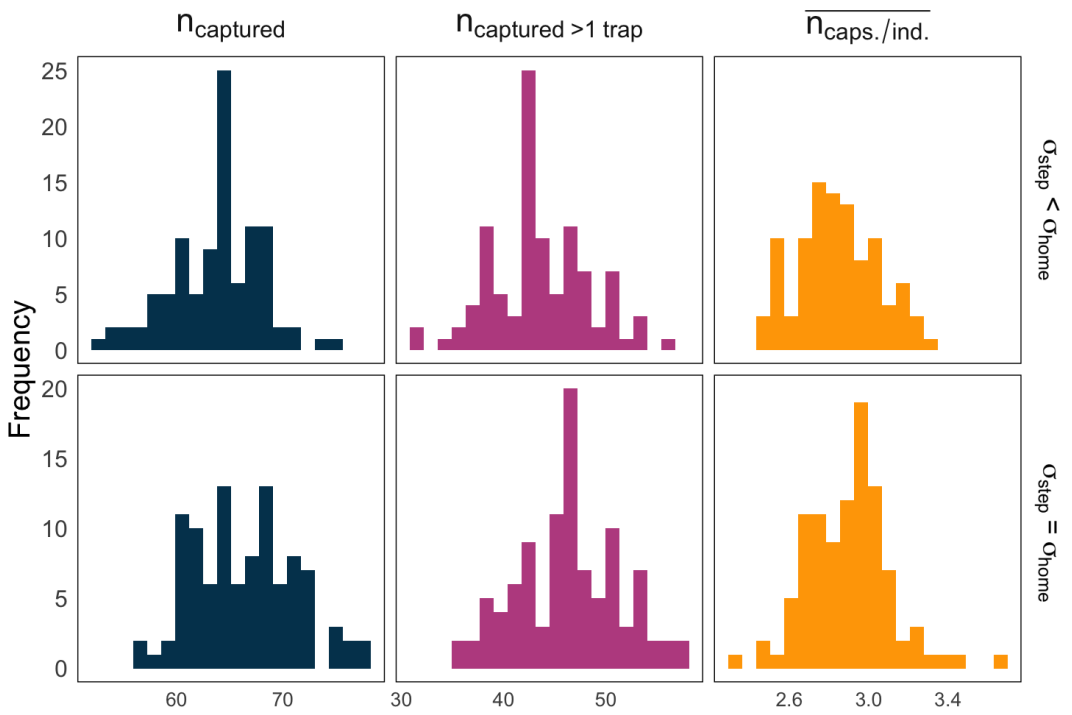

Figure 9. Summary of encounter data recorded from simulated tracks Summary of spatial encounter histories from 100 simulations. From left to right, horizontal facets show: number of individuals captured, number of individuals captured on more than one trap, and mean number of captures per individual. From top to bottom, vertical facets show the two scenarios: small-σ_{step} and large-σ_{step}, respectively.

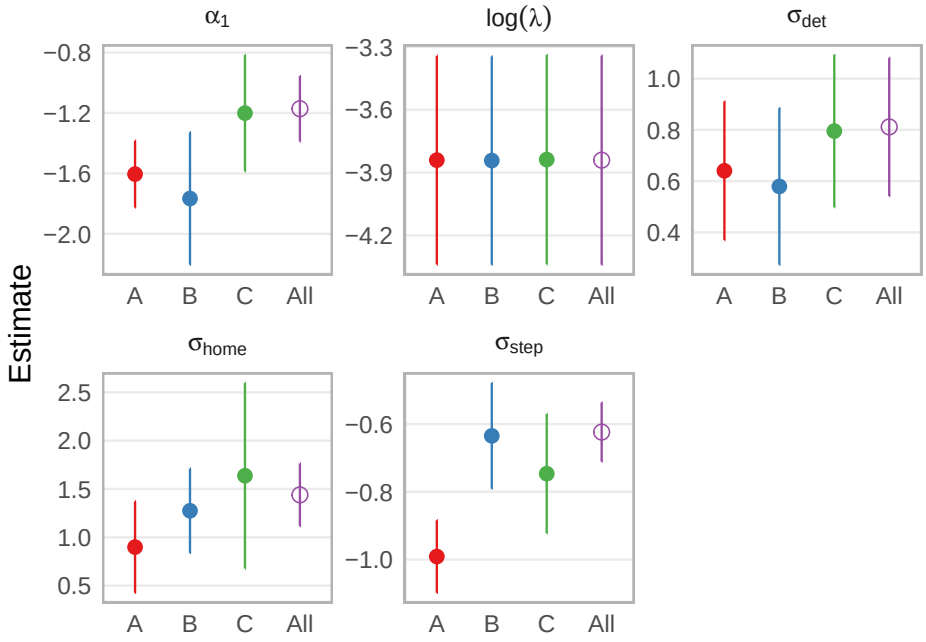

Figure 10. Parameters estimates from case study

Results from exploring the effect of individual variation on parameter estimates in the integrated model. Each closed circles represents parameter estimates from a single model fit with one of the individuals (here: A, B, or C), and the open circles represent parameter estimates from a model fit pooling all three individuals. Error bars represent 95% confidence intervals. Each plot title denotes the model parameter.

A P P E N D I X K

PARAMETER CONSISTENCY

Parameter sharing across data types and associated sub-models is widely promoted as offering improved inferences about complex processes, but less attention has been paid to understanding the consequences of misspecifying the sharing structure (Tenan et al., 2017; Murphy et al., 2019). The results reported in the main text arise from an integrated model with a single shared parameter, the cost parameter α, and three independent scale or variance parameters, σ_{step} and σ_{home} in the movement model, and σ_{det} in the SCR model).

We note that σ_{det} is the spatial scale of *detection* in the SCR model and is *proportional to* space use, whereas σ_{home} in the movement model is a direct estimate of space use. However, given the frequent interpretation of σ_{det} as an estimate of space use that is subsequently converted to an estimate of home range size, some may argue there is a conceptual equivalence between the two parameters under this interpretation. Therefore, in addition to the model evaluations reported in the main text, we also investigated the consequences of making such an assumption by sharing a single 'space use' parameter across the SCR and movement models, and hence forcing a $\sigma_{det} = \sigma_{home}$ constraint.

Here we present the results of a simulation that includes evaluating a formulation of the integrated model with a second shared parameter, name the scale parameter: $\sigma_{det} = \sigma_{home}$. Following the model naming convention in the main text, we refer to the integrated model with independent scale parameters as iM''_{lcp} and the model with a shared scale parameter as iM'_{lcp} (Appendix K: Table 11).

Model performance: simulations

Assuming that the SCR detection range and the spatial scale parameter in the movement model were equivalent resulted in substantial bias in estimates of density, which increased with the addition of more telemetered individuals (Appendix K: Table 12, Appendix K:Figure 11). In the small-σ_{step} scenario, density was underestimated by between –7% ($n_{tel} = 1$) and –9% ($n_{tel} = 5$) on average when the scale parameters were shared, but was unbiased when scale parameters were treated as independent (0% for $n_{tel} = 1$; Appendix K: Table 12, Appendix

K:Figure 11). This bias arises because the shared parameter is an average of the emergent SCR spatial scale parameter and the true movement parameter. For example, for $n_{tel} = 1$, the average percent relative bias of σ_{det} and σ_{home} (relative to σ_{home}) for the independent model was –29% and –3%, respectively, compared to –9% when the parameter is shared (Appendix K: Table 12, Appendix K:Figure 11). Furthermore, the estimate for the SCR parameter was always lower than the equivalent movement model parameter (σ_{det} vs. σ_{home}; Appendix K: Table 12, Appendix K:Figure 11). As more individuals are added, estimates are more influenced by the data-rich movement model and the bias increases. Thus, while iM'_{lcp} model improves estimation of cost compared to model M_{lcp}, the misspecification of the encounter model renders density estimation significantly less useful. The results from the large-σ_{step} scenario were similar in that assuming parameter consistency again resulted in bias that ranged from –7% ($n_{tel} = 1$) to –9% ($n_{tel} = 5$; Appendix K: Table 12, Appendix K:Figure 11). In all cases in this scenario, regardless of number of telemetered individuals or whether or not parameter consistency is assumed, the integrated model produced biased estimates of the cost parameter α_1, ranging from –26% to –20% for iM''_{lcp} and –24% to –20% for iM'_{lcp} model; Appendix K: Table 12, Appendix K:Figure 11).

Model performance: empirical example

In our empirical example, and further supporting our findings, model iM''_{lcp} (the independent parameterization of σ_{det}) always outperformed the shared parameterization of model iM'_{lcp}, again identified by the large AIC differences (Appendix K: Table 13). The shared model, iM'_{lcp}, also appears to underestimate density (–3.956 vs –3.841; Appendix K: Table 13), though to a lesser degree than seen in the simulation study (Appendix K: Table 12, Appendix K:Figure 11).

Conclusions

Sharing the spatial scale parameter σ_{det} inflated its estimate within the SCR sub-model and introduced bias into the density estimator. The technical solution was straightforward: including an independent scale parameter for the telemetry data (σ_{home}) completely remedied the bias, producing an accurate estimate of density simultaneously with cost. Thus, while the integration of an explicit movement model is a logical extension of the SCR model, our results raise three important issues. First, caution should be exercises when making assumptions about parameter sharing across data types or sub models. Second, independent verification of consistency should be conducted where possible (see also Tenan et al. 2017). Third, the interpretation of the spatial scale parameter should be considered carefully, for example when converting σ_{det} to an approximation of home range size.

Table 11. Model components for all models from the evaluations

Model	Data	Distance	σ_{det}	Evaluations
M_{Euc}	SCR	Euclidean	Standard	Case study
M_{lcp}	SCR	Least-cost path	Standard	Case study + Simulations
iM'_{Euc}	SCR+Telem	Euclidean	Shared	Case study
iM''_{Euc}	SCR+Telem	Euclidean	Independent	Case study
iM'_{lcp}	SCR+Telem	Least-cost path	Shared	Case study + Simulations
iM''_{lcp}	SCR+Telem	Least-cost path	Independent	Case study + Simulations

Table 12. Full results of relative bias

Full simulation results of relevant parameters from both σ_{step} scenarios, showing percent relative bias (%RB). Lower and upper values are from 95% quantiles. The column σ_{det} notes the model sharing structure. Relative bias for σ_{det} and σ_{home} are both relative to the data-generating value for σ_{home}.

			Small-σ_{step}	Large-σ_{step}
Parameter	σ_{det}	n_{tel}	%RB	%RB
α_1	standard	0	-17	-19
α_1	shared	1	-2	-24
α_1	independent	1	-3	-26
α_1	shared	3	-3	-21
α_1	independent	3	-3	-21
α_1	shared	5	-2	-20
α_1	independent	5	-2	-20
λ	standard	0	1	1
λ	shared	1	-7	-7
λ	independent	1	0	0
λ	shared	3	-9	-7
λ	independent	3	0	-1
λ	shared	5	-9	-7
λ	independent	5	0	0
σ_{det}	standard	0	-31	-22
σ_{det}	shared	1	-9	-4
σ_{det}	independent	1	-29	-23
σ_{det}	shared	3	-3	-1
σ_{det}	independent	3	-29	-22
σ_{det}	shared	5	-1	0
σ_{det}	independent	5	-28	-22
σ_{home}	independent	1	3	-1
σ_{home}	independent	3	3	-1
σ_{home}	independent	5	3	0
σ_{step}	shared	1	-1	-7
σ_{step}	independent	1	-1	-8
σ_{step}	shared	3	-1	-7
σ_{step}	independent	3	-1	-7
σ_{step}	shared	5	-1	-7
σ_{step}	independent	5	-1	-7
ψ	shared	1	0	0
ψ	independent	1	0	0
ψ	shared	3	0	0
ψ	independent	3	0	0
ψ	shared	5	0	0
ψ	independent	5	0	0

Table 13. Case study results

Results from American black bears analysis. From left to right: fitted model, all seven estimated parameters showing point estimates with standard errors directly below in parentheses, and the negative likelihood value.

model	α_1	$\log(\lambda)$	σ_{det}	p_0	ψ	σ_{step}	σ_{home}	–loglik
M_{Euc}	–	–3.8428	1.2377	–2.7677	–	–	–	124.3217
	–	(0.2545)	(0.1386)	(0.3608)	–	–	–	
M_{lcp}	0.2261	–3.8453	1.2845	–2.7356	–	–	–	124.2723
	(0.7533)	(0.2544)	(0.2086)	(0.3768)	–	–	–	
iM'_{Euc}	–	–3.9861	1.6937	–3.5626	–0.6663	–0.1955	–	3880.029
	–	(0.2475)	(0.0857)	(0.3071)	(0.0479)	(0.0197)	–	
iM''_{Euc}	–	–3.8428	1.2377	-2.7677	–0.6663	–0.1949	1.9137	3874.447
	–	(0.2545)	(0.1386)	(0.3608)	(0.0479)	(0.0197)	(0.1642)	
iM'_{lcp}	–1.2414	–3.9560	1.2049	–3.6727	–0.6663	–0.6492	–	3840.379
	(0.1223)	(0.2527)	(0.0980)	(0.3145)	(0.0479)	(0.0485)	–	
iM''_{lcp}	–1.172	–3.8407	0.8116	–2.8823	–0.6663	–0.6237	1.4391	3835.423
	(0.1104)	(0.2547)	(0.1376)	(0.3676)	(0.0479)	(0.0445)	(0.1654)	

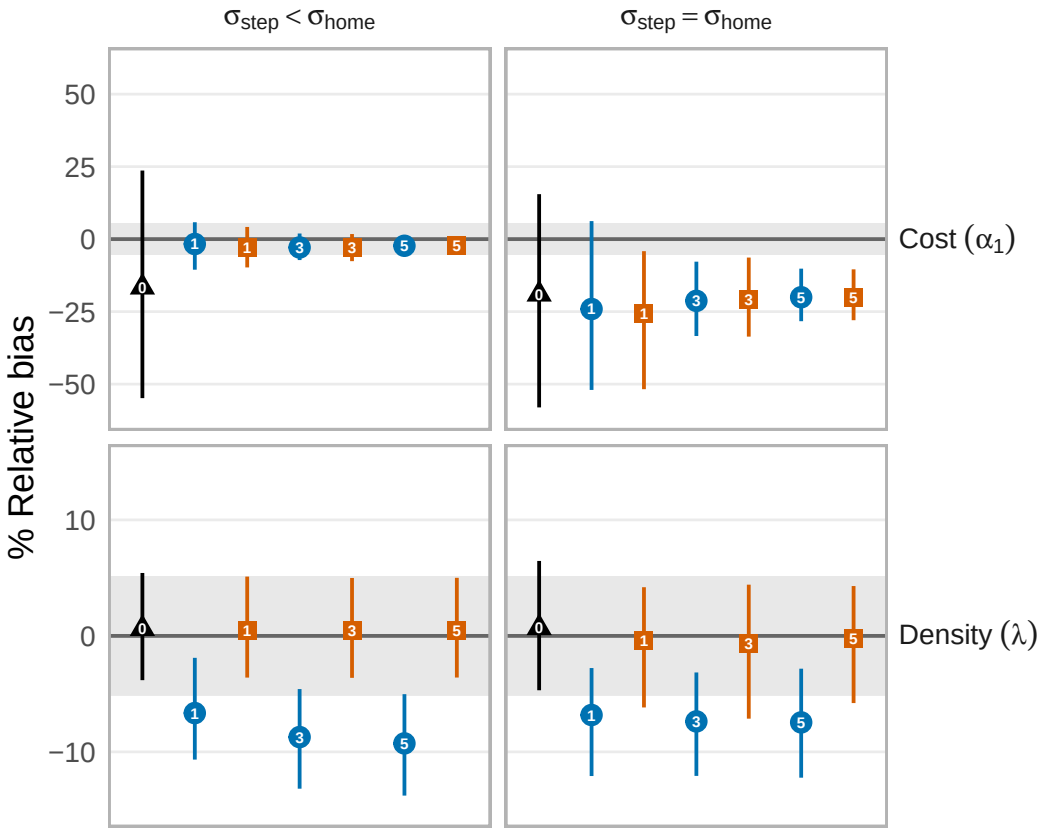

Figure 11. Percent relative bias from simulation results

To illustrate estimator precision, vertical lines are 50% confidence intervals, noting that the 50% intervals are proportional to 95% intervals but offer a visual balance of bias and associated variance. Black triangles represent model M_{lcp}. Blue circles represent model iM'_{lcp}, and orange squares represent model iM''_{lcp}, with an increasing number of telemetered individuals included (inset number). Gray shaded area shows an acceptable amount of bias (±5%). Columns left-to-right represent the small- and large-σ_{step} scenarios, respectively, and rows top-to-bottom represent cost and density estimates, respectively.

A P P E N D I X L

LIKELIHOOD FUNCTION FOR THE INTEGRATED MODEL

R code for the likelihood function. Maximum likelihood optimization was carried out using `nlm` and minimizing the negative log-likelihood.

Authors
Gates Dupont
University of Massachusetts–Amherst
gdupont@umass.edu

Daniel W. Linden
NOAA National Marine Fisheries Service,
Greater Atlantic Regional Fisheries Office
danlinden@gmail.com

Chris Sutherland
Centre for Research into Ecological and Environmental Modelling,
University of St Andrews
css6@st-andrews.ac.uk

```
iM_lcp <- function(
    param, share_sigma = FALSE,
    dist=c("circ","lcp")[3], mod=c("exp","gauss")[2],
    teldata, spatdata = NULL, landscape = NULL, use.sbar=FALSE,
    scr_y = NULL, K = NULL, trap_locs = NULL, scr_ss = NULL,
    prj = NULL){

    #alpha1: cost parameter
    #sigma_step: spatial scale (steps)
    #psi: pr(move from current cell)
    #sigma_det: spatial scale (detection; shared OR SCR only)
    #sigma_home: spatial scalec(range)
    #p0: baseline encounter probability
    #log_lambda: population density

    # Basics
    nguys <- length(teldata)
    ll <- rep(NULL,nguys)
    np <- ncol(spatdata[[1]])-2

    # Some starting parameters
    alpha1 <- matrix(param[1], nrow = nguys, ncol = np, byrow = TRUE)
    sigma_step <- rep(exp(param[2]), nguys)
    psi <- rep(plogis(param[3]), nguys)
    sigma_det <- exp(param[4])

    #----S C R----

    # SCR starting parameters
    a1_scr <- as.numeric(unique(alpha1))
    p0 <- plogis(param[5]) # Probability, on link scale
    log_lambda <- exp(param[6])
    # Shouldn't be negative, so on link scale

    # Expanding K for later use
    if(length(K)==1) K<- rep(K,nrow(trap_locs))
    # Generalized for irregular sampling periods

    G <- coordinates(scr_ss) # Pixels
    nG <- nrow(G) # Number of pixels

    # Cost distance pieces
    cost <- exp(a1_scr*landscape) # Cost surface w/ proposed parameter
    tr1 <- transition(cost,transitionFunction=function(x) 1/mean(x),
                      directions=16)
    tr1CorrC <- geoCorrection(tr1,type="c",multpl=FALSE,scl=FALSE)
    D <- costDistance(tr1CorrC,trap_locs,G) # Cost distance

    # Half-normal encounter model
    probcap <- p0 * exp(-D^2/(2*sigma_det^2))
    # rows = traps, col = pixels

    # Detection matrix, rows = traps, cols = pixels
    Pm <- matrix(NA,nrow=nrow(probcap),ncol=ncol(probcap))

    # Encounter historiess, augemented with 0 row
    # to estimate probability of an uncaptured individual
    # being in one of those pixels
    if(!is.na(dim(scr_y)[3])){
      scr_y <- apply(scr_y, 1:2, sum)
    }

    ymat <- rbind(scr_y,rep(0,ncol(scr_y)))

    # Loop through encounter histories (inidividuals + extra)
```

```
# Calculate likelihood of that individual's activity center
# being at each pixel
lik.marg <- rep(NA,nrow(ymat))
for(i in 1:nrow(ymat)){
  Pm[1:length(Pm)] <-  (dbinom(
      rep(ymat[i,],nG),
      rep(K,nG),
      probcap[1:length(Pm)],
      log=TRUE))
   lik.cond <-  exp(colSums(Pm))
  # Likelihood of number of captures
  # conditional on s_i at pixel G_i
   lik.marg[i] <-  sum( lik.cond*(1/nG) )
  # Likelihood averaged across all pixels
}

# Remaining likelihood pieces
nv <- c(rep(1, length(lik.marg) - 1), 1)
atheta <- 1 - lik.marg[nrow(ymat)]
nind <- nrow(ymat) - 1
part1 <- nind*log(sum(log_lambda*nG))-sum(log_lambda*nG)*atheta
part2 <- sum(nv[1:nind] * log(lik.marg[1:nind]))
scr_out <- part1 + part2

#---- Movement model -----

sigma_home <- ifelse(share_sigma==TRUE,sigma_det,exp(param[7]))

#loop over individuals
ll <- numeric(nguys)
for(ind in 1:nguys){

  ss <- spatdata[[ind]][,1:2]
  # SPATDATA AS COORDINATES OF RASTER (still the bigger box)

  tel.locs <- as.matrix(teldata[[ind]])
  pixels <- raster::extract(rasterFromXYZ(cbind(ss, z=1)),
                            tel.locs, cellnumbers=T)[,1]
  moved <- 1 - as.numeric(diff(pixels) == 0)

  if(dist=="circ"){
    D <- as.matrix(gdistance::commuteDistance(
                          tr1CorrC,
                          ss[1:(nrow(ss)-1),]))
    D <- D/ncell(cost)
    D <- D[pixels,]
  }
  if(dist=="lcp"){
    D_ss <- gdistance::costDistance(tr1CorrC,  ss, ss)
    D <- D_ss[pixels,]
  }

  if(use.sbar){

    dsbar <- D_ss[which(spatdata[[ind]][,3] == 1),]

    dsbar <- matrix(dsbar, nrow=nrow(D),
                    ncol=length(dsbar), byrow=TRUE)

    if(mod=="gauss"){
      kern<- {exp(-D*D/(2*sigma_step[ind]*sigma_step[ind]) -
      dsbar*dsbar/(2*sigma_home*sigma_home))}
    }
    if(mod=="exp"){
      kern<- {exp(-D/(2*sigma_step[ind]*sigma_step[ind]) -
      dsbar*dsbar/(2*sigma_home*sigma_home))}
```

```
      }
    }else{
      if(mod=="gauss"){
        kern <- exp(-D*D/(2*sigma_step[ind]*sigma_step[ind]))
      }
      if(mod=="exp"){
        kern <- exp(-D/(2*sigma_step[ind]*sigma_step[ind]))
      }
    }
    kern[cbind(1:nrow(kern), pixels )] <- 0
    # *cannot move to same pixel bc CONDITIONAL on moved
    kern <- kern/rowSums(kern)
    probs <- kern[cbind(1:(nrow(kern)-1), pixels[-1] )]
    # remove terminus because it has no movement outcome
    part1 <- moved*log(psi[ind]) + (1-moved)*log(1-psi[ind])
    part2 <- log(probs[moved==1])
    ll[ind]<- sum(part1) + sum(part2)
  }
  nll <- -1*(sum(ll)+scr_out)
  nll
}
```

BIBLIOGRAPHY

Bocinsky, R. K. (2016). *FedData: Functions to Automate Downloading Geospatial Data Available from Several Federated Data Sources*. R package version 2.0.4.

Bondrup-Nielsen, S. (1983). Density estimation as a function of live-trapping grid and home range size. *Canadian Journal of Zoology*, 61(10):2361–2365.

Borchers, D. L. and Efford, M. G. (2008). Spatially Explicit Maximum Likelihood Methods for Capture-Recapture Studies. *Biometrics*, 64(2):377–385.

Christ, A., Ver Hoef, J., and Zimmerman, D. L. (2008). An animal movement model incorporating home range and habitat selection. *Environmental and Ecological Statistics*, 15(1):27–38.

Chu, D. K., Akl, E. A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H. J., El-harakeh, A., Bognanni, A., Lotfi, T., Loeb, M., et al. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of sars-cov-2 and covid-19: a systematic review and meta-analysis. *The Lancet*, 395(10242):1973–1987.

Creel, S., Merkle, J., Mweetwa, T., Becker, M. S., Mwape, H., Simpamba, T., and Simukonda, C. (2020). Hidden markov models reveal a clear human footprint on the movements of highly mobile african wild dogs. *Scientific reports*, 10(1):17908–17908.

Crooks, K. R., Burdett, C. L., Theobald, D. M., King, S. R., Di Marco, M., Rondinini, C., and Boitani, L. (2017). Quantification of habitat fragmentation reveals extinction risk in terrestrial mammals. *Proceedings of the National Academy of Sciences*, 114(29):7635–7640.

Crooks, K. R., Burdett, C. L., Theobald, D. M., Rondinini, C., and Boitani, L. (2011). Global patterns of fragmentation and connectivity of mammalian carnivore habitat. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1578):2642–2651.

Cushman, S. A., McRae, B., Adriaensen, F., Beier, P., Shirley, M., and Zeller, K. (2013). Biological corridors and connectivity [chapter 21]. *In: Macdonald, DW; Willis, KJ, eds. Key Topics in Conservation Biology 2. Hoboken, NJ: Wiley-Blackwell. p. 384-404.*, pages 384–404.

Dijkstra, E. W. et al. (1959). A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271.

Dillon, A. and Kelly, M. J. (2007). Ocelot Leopardus pardalis in Belize: the impact of trap spacing and distance moved on density estimates. *Oryx*, 41(4):469–477.

Dirzo, R. and Raven, P. H. (2003). Global state of biodiversity and loss. *Annual review of Environment and Resources*, 28(1):137–167.

Dupont, G., Linden, D. W., and Sutherland, C. (2021a). Improved inferences about landscape connectivity from spatial capture-recapture by integration of a movement model. *Ecology*, 00(0):000. *In Press*.

Dupont, G., Linden, D. W., and Sutherland, C. (2021b). Supplement: Improved inferences about landscape connectivity from spatial capture-recapture by integration of a movement model. Open Science Framework. https://osf.io/684p9/.

Dupont, G., Royle, J. A., Nawaz, M. A., and Sutherland, C. (2021c). Optimal sampling design for spatial capture–recapture. *Ecology*, 102(3):e03262.

Durbach, I., Borchers, D., Sutherland, C., and Sharma, K. (2021). Fast, flexible alternatives to regular grid designs for spatial capture–recapture. *Methods in Ecology and Evolution*, 12(2):298–310.

Efford, M. (2004). Density estimation in live-trapping studies. *Oikos*, 106(3):598–610.

Efford, M. G. and Boulanger, J. (2019). Fast evaluation of study designs for spatially explicit capture–recapture. *Methods in Ecology and Evolution*, 10(9):1529–1535.

Efford, M. G. and Fewster, R. M. (2013). Estimating population size by spatially explicit capture-recapture. *Oikos*, 122(6):918–928.

Etten, J. v. (2017). R package gdistance: distances and routes on geographical grids. *Journal of Statistical Software*.

Fabiano, E. C., Sutherland, C., Fuller, A. K., Nghikembua, M., Eizirik, E., and Marker, L. (2020). Trends in cheetah (Acinonyx jubatus) density in north-central Namibia. *Population Ecology*, pages 1438–390X.12045.

Fuller, A. K., Sutherland, C. S., Royle, J. A., and Hare, M. P. (2016). Estimating population density and connectivity of American mink using spatial capture-recapture. *Ecological Applications*, 26(4):1125–1135.

Gardner, B., Royle, J. A., Wegan, M. T., Rainbolt, R. E., and Curtis, P. D. (2010). Estimating Black Bear Density Using DNA Data From Hair Snares. *Journal of Wildlife Management*, 74(2):318–325.

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley Professional.

Gravel, D. and Massol, F. (2020). Toward a general theory of metacommunity ecology. In *Theoretical Ecology*, pages 195–220. Oxford University Press.

Gupta, A., Dilkina, B., Morin, D. J., Fuller, A. K., Royle, J. A., Sutherland, C., and Gomes, C. P. (2019). Reserve design to optimize functional connectivity and animal density. *Conservation Biology*, 33(5):1023–1034.

Haddad, N. M., Brudvig, L. A., Clobert, J., Davies, K. F., Gonzalez, A., Holt, R. D., Lovejoy, T. E., Sexton, J. O., Austin, M. P., Collins, C. D., et al. (2015). Habitat fragmentation and its lasting impact on earth's ecosystems. *Science advances*, 1(2):e1500052.

Hanks, E. M. and Hooten, M. B. (2013). Circuit theory and model-based inference for landscape connectivity. *Journal of the American Statistical Association*, 108(501):22–33.

Hanski, I. (1999). *Metapopulation ecology*. Oxford University Press.

Hochachka, W. M., Dobson, A. P., Hawley, D. M., and Dhondt, A. A. (2021). Host population dynamics in the face of an evolving pathogen. *Journal of Animal Ecology*.

Hooten, M. B., Johnson, D. S., McClintock, B. T., and Morales, J. M. (2017). *Animal movement: statistical models for telemetry data*. CRC press.

Howell, P. E., Muths, E., Hossack, B. R., Sigafus, B. H., and Chandler, R. B. (2018). Increasing connectivity between metapopulation ecology and landscape ecology. *Ecology*, 99(5):1119–1128.

Jackson, R. M., Roe, J. D., Wangchuk, R., and Hunter, D. O. (2006). Estimating snow leopard population abundance using photography and capture-recapture techniques. *Wildlife Society Bulletin*, 34(3):772–781.

Johnson, D. S., Thomas, D. L., Ver Hoef, J. M., and Christ, A. (2008). A general framework for the analysis of animal resource selection from telemetry data. *Biometrics*, 64(3):968–976.

Kery, M. and Royle, J. A. (2020). *Applied Hierarchical Modeling in Ecology: Analysis of Distribution, Abundance and Species Richness in R and BUGS: Volume 2: Dynamic and Advanced Models*. Academic Press.

Lack, D. (1954). *The natural regulation of animal numbers*. Oxford Clarendon Press.

Levin, S. A. (1992). The problem of pattern and scale in ecology: the Robert H. MacArthur award lecture. *Ecology*, 73(6):1943–1967.

Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *American Entomologist*, 15(3):237–240.

Levins, R. (1970). Extinction. In Gerstenhaber, M., editor, *Some Mathematical Questions in Biology. Lectures on Mathematics on the Life Sciences. Volume 2.*, pages 77–107. American Mathematical Society.

Levins, R. and Culver, D. (1971). Regional coexistence of species and competition between rare species. *Proceedings of the National Academy of Sciences*, 68(6):1246–1248.

Linden, D. W., Sirén, A. P., and Pekins, P. J. (2018). Integrating telemetry data into spatial capture–recapture modifies inferences on multi-scale resource selection. *Ecosphere*, 9(4):e02203.

MacArthur, R. and Connell, J. (1967). *The biology of populations*. Wiley, New York.

McClintock, B. T., Abrahms, B., Chandler, R., Conn, P. B., Converse, S. J., Emmet, R., Gardner, B., Hostetter, N. J., and Johnson, D. S. (2021). An integrated path for spatial capture-recapture and animal movement modeling. *Ecology*, 00(0):0000.

McRae, B. H., Dickson, B. G., Keitt, T. H., and Shah, V. B. (2008). Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, 89(10):2712–2724.

Meyer, N. F., Moreno, R., Sutherland, C., de la Torre, J. A., Esser, H. J., Jordan, C. A., Olmos, M., Ortega, J., Reyna-Hurtado, R., Valdes, S., et al. (2020). Effectiveness of panama as an intercontinental land bridge for large mammals. *Conservation Biology*, 34(1):207–219.

Michelot, T., Blackwell, P. G., Chamaillé-Jammes, S., and Matthiopoulos, J. (2020). Inference in mcmc step selection models. *Biometrics*, 76(2):438–447.

Michelot, T., Blackwell, P. G., and Matthiopoulos, J. (2019). Linking resource selection and step selection models for habitat preferences in animals. *Ecology*, 100(1).

Moorcroft, P. R. and Barnett, A. (2008). Mechanistic home range models and resource selection analysis: a reconciliation and unification. *Ecology*, 89(4):1112–1119.

Morales, J. M., Moorcroft, P. R., Matthiopoulos, J., Frair, J. L., Kie, J. G., Powell, R. A., Merrill, E. H., and Haydon, D. T. (2010). Building the bridge between animal movement and population dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1550):2289–2301.

Morin, D. J., Fuller, A. K., Royle, J. A., and Sutherland, C. (2017). Model-based estimators of density and connectivity to inform conservation of spatially structured populations. *Ecosphere*, 8(1):e01623.

Murphy, S. M., Wilckens, D. T., Augustine, B. C., Peyton, M. A., and Harper, G. C. (2019). Improving estimation of puma *(Puma concolor)* population density: clustered camera-trapping, telemetry data, and generalized spatial mark-resight models. *Scientific reports*, 9(1):1–13.

Nathan, R., Getz, W. M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D., and Smouse, P. E. (2008). A movement ecology paradigm for unifying organismal movement research. *Proceedings of the National Academy of Sciences*, 105(49):19052–19059.

Peterman, W. E. (2018). Resistancega: An r package for the optimization of resistance surfaces using genetic algorithms. *Methods in Ecology and Evolution*, 9(6):1638–1647.

Pimm, S. L., Russell, G. J., Gittleman, J. L., and Brooks, T. M. (1995). The future of biodiversity. *Science*, 269(5222):347–350.

Plard, F., Fay, R., Kéry, M., Cohas, A., and Schaub, M. (2019). Integrated population models: powerful methods to embed individual processes in population dynamics models. *Ecology*, 100(6):e02715.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rayfield, B., Fortin, M.-J., and Fall, A. (2011). Connectivity for conservation: a framework to classify network measures. *Ecology*, 92(4):847–858.

Revilla, E. and Wiegand, T. (2008). Individual movement behavior, matrix heterogeneity, and the dynamics of spatially structured populations. *Proceedings of the National Academy of Sciences*, 105(49):19120–19125.

Rich, L. N., Miller, D. A., Muñoz, D. J., Robinson, H. S., McNutt, J. W., and Kelly, M. J. (2019). Sampling design and analytical advances allow for simultaneous density estimation of seven sympatric carnivore species from camera trap data. *Biological Conservation*, 233:12–20.

Rout, T. M., Hauser, C. E., and Possingham, H. P. (2007). Minimise long-term loss or maximise short-term gain?: Optimal translocation strategies for threatened species. *Ecological Modelling*, 201(1):67–74.

Royle, J. A., Chandler, R. B., Gazenski, K. D., and Graves, T. A. (2013a). Spatial capture–recapture models for jointly estimating population density and landscape connectivity. *Ecology*, 94(2):287–294.

Royle, J. A., Chandler, R. B., Sollmann, R., and Gardner, B. (2013b). *Spatial capture-recapture*. Academic Press.

Royle, J. A., Chandler, R. B., Sun, C. C., and Fuller, A. K. (2013c). Integrating resource selection information with spatial capture–recapture. *Methods in Ecology and Evolution*, 4(6):520–530.

Royle, J. A., Fuller, A. K., and Sutherland, C. (2018). Unifying population and landscape ecology with spatial capture–recapture. *Ecography*, 41(3):444–456.

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(2):1–11.

Royle, J. A. and Young, K. V. (2008). A hierarchical model for spatial capture-recapture data. *Ecology*, 89(8):2281–2289.

Sciaini, M., Fritsch, M., Scherer, C., and Simpkins, C. E. (2018). Nlmr and landscapetools: An integrated environment for simulating and modifying neutral landscape models in r. *Methods in Ecology and Evolution*, 9(11):2240–2248.

Shaw, A. K. (2020). Causes and consequences of individual variation in animal movement. *Movement ecology*, 8(1):1–12.

Simberloff, D. (1998). Flagships, umbrellas, and keystones: is single-species management passé in the landscape era? *Biological conservation*, 83(3):247–257.

Sollmann, R., Gardner, B., and Belant, J. L. (2012). How Does Spatial Study Design Influence Density Estimates from Spatial Capture-Recapture Models? *PLoS ONE*, 7(4):e34575.

Sollmann, R., Gardner, B., Belant, J. L., Wilton, C. M., and Beringer, J. (2016). Habitat associations in a recolonizing, low-density black bear population. *Ecosphere*, 7(8):e01406.

Sollmann, R., Gardner, B., Parsons, A. W., Stocking, J. J., McClintock, B. T., Simons, T. R., Pollock, K. H., and O'Connell, A. F. (2013). A spatial mark–resight model augmented with telemetry data. *Ecology*, 94(3):553–559.

Spiegel, O., Leu, S. T., Bull, C. M., and Sih, A. (2017). What's your move? movement as a link between personality and spatial dynamics in animal populations. *Ecology letters*, 20(1):3–18.

Sun, C. (2014). Estimating black bear population density in the southern black bear range of new york with a non-invasive, genetic, spatial capture-recapture study. *Thesis, Cornell University.*

Sun, C. C., Fuller, A. K., and Royle, J. A. (2014). Trap Configuration and Spacing Influences Parameter Estimates in Spatial Capture-Recapture Models. *PLoS ONE*, 9(2):e88025.

Sutherland, C., Fuller, A. K., and Royle, J. A. (2015). Modelling non-euclidean movement and landscape connectivity in highly structured ecological networks. *Methods in Ecology and Evolution*, 6(2):169–177.

Sutherland, C., Fuller, A. K., Royle, J. A., Hare, M. P., and Madden, S. (2018). Large-scale variation in density of an aquatic ecosystem indicator species. *Scientific Reports*, 8(1):8958.

Sutherland, C., Royle, J. A., and Linden, D. W. (2019). oSCR: A spatial capture–recapture R package for inference about spatial ecological processes. *Ecography*, 42(9):1459–1469.

Tenan, S., Pedrini, P., Bragalanti, N., Groff, C., and Sutherland, C. (2017). Data integration for inference about spatial processes: A model-based approach to test and account for data inconsistency. *PloS one*, 12(10):e0185588.

Tischendorf, L. and Fahrig, L. (2000). How should we measure landscape connectivity? *Landscape ecology*, 15(7):633–641.

Tischendorf, L., Grez, A., Zaviezo, T., and Fahrig, L. (2005). Mechanisms affecting population density in fragmented habitat. *Ecology and Society*, 10(1).

Tobler, M. W. and Powell, G. V. (2013). Estimating jaguar densities with camera traps: Problems with current designs and recommendations for future studies. *Biological Conservation*, 159:109–118.

Williams, B. K., Nichols, J. D., and Conroy, M. J. (2002). *Analysis and Management of Animal Populations : Modeling, Estimation, and Decision Making*. Academic Press, San Diego, CA, first edition.

Wilton, C. M., Puckett, E. E., Beringer, J., Gardner, B., and Eggert, L. S. (2014). Trap Array Configuration Influences Estimates and Precision of Black Bear Density and Abundance. *PLoS ONE*, 9(10):111257.

Wolters, M. A. (2015). A Genetic Algorithm for Selection of Fixed-Size Subsets with Application to Design Problems. *Journal of Statistical Software*, 68.

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S. M., Case, A., Costello, C., Dewitz, J., Fry, J., et al. (2018). A new generation of the united

states national land cover database: Requirements, research priorities, design, and implementation strategies. *ISPRS Journal of Photogrammetry and Remote Sensing*, 146:108–123.

Zeller, K. A., McGarigal, K., and Whiteley, A. R. (2012). Estimating landscape resistance to movement: a review. *Landscape ecology*, 27(6):777–797.