Estimating demographic rates to improve monitoring of highly mobile species

Robert Emmet

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Beth Gardner, Chair

Robert Long

Aaron Wirsing

Program Authorized to Offer Degree:

Quantitative Ecology and Resource Management, College of the Environment

University of Washington

Abstract

Estimating demographic rates to improve monitoring of highly mobile species

Robert Emmet

Chair of the Supervisory Committee:
Beth Gardner
School of Environmental and Forest Sciences, College of the Environment

Estimating demographic rates of wildlife species, such as survival and fecundity, is crucial for monitoring wildlife populations and informing management of these species. Monitoring highly mobile species is especially challenging, as their life histories and behaviors (e.g., migration) can affect inference on demographic rates and ultimately render monitoring less effective. Species' movements can expose them to a variety of hazards and opportunities, creating spatial and temporal variation in demographic rates that must be accounted for in models. Furthermore, the movement behaviors of many highly mobile species can violate key assumptions of the standard statistical models used to estimate demographic rates, so that new monitoring frameworks and models need to be designed to minimize violations of model assumptions or relax those assumptions. In this dissertation, I used several case studies to demonstrate how novel models and monitoring frameworks can improve demographic rate estimation, ecological inference, and population monitoring capabilities for highly mobile species.

This dissertation is dedicated to two people: my grandfather Richard S. Emmet II, who inspired me to start this dissertation, and my fiancée Isabel Storey, who encouraged and supported me as I finished it.

Acknowledgements

I could not have written this dissertation without the support and advice of numerous people and institutions. Funding for this research was provided by the Charlotte Martin Foundation, Woodland Park Zoo, Wilburforce Foundation, USGS, and USFWS. Part of this research was funded by a U.S. Geological Survey Northwest Climate Adaptation Science Center award G17AC000218 to Robert Emmet. Of the many people to thank, first of all, I would like to thank the coauthors of these chapters, including but not limited to Pat Devers, Briana Abrahms, Lindsey Rich, and Ben Augustine. These papers were made stronger by your contributions to them and they couldn't have been written without you. I would also like to thank the numerous collaborators who gave feedback on these projects at various stages, including Paula Mackay, Jeff Lewis, Keith Aubry, Cathy Raley, Jocelyn Akins, Steph Williams, Nathan Hostetter, Sarah Converse, Devin Johnson, Paul Conn, Richard Chandler, and Brett McClintock. Your advice and ideas helped push this dissertation in new directions. Thank you for L. Van der Weyde and R.H. Walker for helping to identify individual wild dogs. My colleagues in the Quantitative Ecology and Quantitative Conservation Labs and the QERM program at the University of Washington provided invaluable feedback and support throughout the dissertation writing process. Thanks especially to Olivia Sanderfoot and Sarah Bassing for providing advice, helping me weather the numerous ups and downs of research, and reviewing countless pieces of code and writing over the years. Thanks also to Tim Essington and Erica Owens for helping me figure out the ins and outs of the QERM program. My Carleton friends, mentors, and mentees have been a joy to work with and reconnect with throughout my graduate school career. Thank you again and always to Laura Chihara, Katie St. Clair, and Ian Taylor for your mentorship and career guidance. Thanks to Jessalyn Ayars, Emily Schulenberg, and Aya Klos, for being such excellent Carleton students to work with. I know that your contributions to quantitative biology/ecology will be invaluable. Thanks to Katie Koza, Ben Strauss, and many other lifelong Carleton friends, for your infinite good humor and support. An especially big thank you to my committee, Aaron Wirsing, Andy Royle, and Kristin Laidre, for your invaluable feedback on my dissertation. It has been an honor and a privilege to work with you. Another big thank you goes to my co-advisors, Beth Gardner and Robert Long. Beth and Robert, thank you for always being ready with excellent advice and for challenging me to think outside the box. It would take up too much space to list all of things you have done for me during graduate school, but I am forever grateful for all of your support and feedback over the years. Finally, I could not have finished this dissertation without my parents, my grandparents, my brother, and my fiancée, Isabel Storey. I love you all and I hope you enjoy reading what I've been working on for the past five years.

1 Table of Contents

List of Tables and Figures (excluding appendices)

2 Introduction

Monitoring species to obtain information on abundance, survival, and other demographic rates is a key component of conservation. Monitoring is commonly used to detect changes in status of vulnerable or endangered species (Ellis, Ivan and Schwartz, 2014; Steenweg et al., 2016) or to determine an acceptable level of harvest or indirect mortality for a population or species (Johnson et al., 2015). Monitoring can also be used to improve decision-making via an adaptive management framework (Johnson et al., 2015), and can simultaneously be used to improve ecological inference (Nichols, Kendall and Boomer, 2019). Given the potentially serious consequences of using inaccurate population status projections to inform conservation decisions, it is crucial that monitoring programs be designed to obtain accurate estimates of demographic rates.

However, obtaining accurate estimates of demographic rates from monitoring is difficult for many species, particularly highly mobile species. While there seems to be no universal definition of a highly mobile species, I will use a working definition of highly mobile species as species that must make long-distance movements (i.e., regional to continental scales) to successfully forage or reproduce (cf. Runge et al. 2014). This definition includes species that migrate seasonally, species that disperse at some point in their life cycle, or species that are nomadic, and thus encompasses species from a wide variety of taxa (Robinson et al., 2009; Runge et al., 2014). Regardless of the type of movement that makes a species highly mobile, these species have one thing in common: their long-distance movements can bias estimates of demographic rates obtained from common wildlife sampling approaches.

Species' movements can cause bias in demographic rates at multiple stages of the monitoring process. First, regardless of the biases that could be caused by inadequate monitoring design or

model misspecification, animal movement can lead to enormous spatial and temporal variation in demographic rates that must be correctly accounted for in monitoring design and model specification. Movement is a crucial ecological process for many species, and movement decisions can affect individual fitness and survival, population persistence, and even biodiversity (Morales et al., 2010; Jeltsch et al., 2013). For example, individuals that move to new areas may face different levels of exposure to hazards (e.g., hunting, intraspecific competition), leading to variation in survival (Krementz et al., 1987; Duriez et al., 2009; Sawyer et al., 2019). Similarly, immigration and emigration can cause variation in abundance in a given area over time, even if no births or deaths occur during that time (Dail and Madsen, 2011). Variation in demographic rates caused by movement must be adequately accounted for to avoid bias.

Second, incomplete understanding of species' movements and habitat needs may lead to monitoring designs, and subsequent management decisions, that miss key habitats or stages in the species' life cycle. For instance, failure to model a migratory species' full annual cycle may lead researchers and managers to misidentify the factors contributing to a species' decline (Marra et al., 2015). If habitats such as stopover sites are not known to be important for a species due to inadequate monitoring, these habitats may be improperly managed, resulting in potentially catastrophic consequences for the species (Runge et al., 2014). For many species, monitoring must occur across multiple habitat types and during multiple times of the species' life cycle in order to accurately estimate demographic rates and monitor population status (Marra et al., 2015).

Finally, movement can cause bias in demographic rates if the data collected on highly mobile species do not meet the assumptions of the model used to estimate demographic rates. For example, many hierarchical models of species occurrence, abundance, and survival assume that

animals do not leave or enter the study area (geographic closure). Many highly mobile species traverse ranges much larger than our study areas. Animal movement has been shown to bias estimates of abundance (Dail and Madsen, 2011), occupancy (Rota et al., 2009; Hayes and Monfils, 2015; Berigan et al., 2018; Neilson et al., 2018), detection (Stewart et al., 2018), and survival (Ergon and Gardner, 2014; Schaub and Royle, 2014; Gardner et al., 2018). Often, in models with strict closure assumptions, movement outside of study areas is misidentified as a lack of detection (assuming perfect availability for detection) or mortality, leading to bias in true detection, survival, abundance, or other parameters. More generally, violation of both closure and other assumptions can lead to biases in demographic rates, limiting the effectiveness of monitoring as a tool for conservation.

Because inherent variation caused by movement, poor monitoring design, and violation of model assumptions can lead to biases and negative conservation outcomes, it is crucial that we improve monitoring of highly mobile species. Fortunately, adjustments to both monitoring design and model specification can correct biases caused by animal movement. Use of full annual cycle models can help identify key habitats and times of year for population dynamics of migratory species (Robinson et al., 2016). Models can be specified to account for temporary emigration, thus reducing the strictness of, or even eliminating, geographic closure assumptions (MacKenzie et al., 2003; Schaub and Royle, 2014). However, much work remains to be done.

My dissertation focused on several case studies of developing models to improve monitoring of highly mobile species. In particular, each chapter involved extending currently existing models of demographic rates to further account for movement and the variation and bias in demographic rates caused by movement. Chapters 1 and 2 involved development of a modeling and monitoring framework for wolverines in Washington's Cascades that is robust to wolverines'

low detectability and long-distance movements. In Chapter 1, I developed an occupancy model that accounts for a species' high mobility using multi-scale modeling, and in Chapter 2 I extended the preliminary findings of Chapter 1 to create a survey design for wolverine monitoring in the Cascades. In Chapter 3, I used a two-season survival model of American black ducks that accounts for variation in survival and band recovery probabilities by season and location to investigate the relationship between severe winter weather and survival in American black ducks. Finally, in Chapter 4, I implemented a spatial capture-recapture model for group-living species, allowing researchers to simultaneously monitor multiple metrics of population status in group-living species.

References

Berigan, W.J., G.M. Jones, S.A. Whitmore, R.J. Gutiérrez, and M.Z. Peery. 2018. "Cryptic wide-ranging movements lead to upwardly-biased occupancy in a territorial species." *Journal of Applied Ecology* 56 (2): 470–480.

Dail, D., and L. Madsen. 2011. "Models for estimating abundance from repeated counts of an open metapopulation." *Biometrics* 67 (2): 577–587.

Duriez, O., S.A. Sæther, B.J. Ens, R. Choquet, R. Pradel, R.H. Lambeck, and M. Klaassen. 2009. "Estimating survival and movements using both live and dead recoveries: A case study of oystercatchers confronted with habitat change." *Journal of Applied Ecology* 46 (1): 144–153.

Ellis, M.M., J.S., Ivan, and M.K. Schwartz. 2014. "Spatially explicit power analyses for occupancy-based monitoring of wolverine in the U.S. Rocky Mountains." *Conservation Biology* 28 (1): 52–62.

Ergon, T., and B. Gardner. 2014. “Separating mortality and emigration: modelling space use, dispersal and survival with robust-design spatial capture-recapture data.” *Methods in Ecology and Evolution* 5 (12): 1327–1336.

Gardner, B., R. Sollmann, N.S. Kumar, D. Jathanna, and K.U. Karanth. 2018. “State space and movement specification in open population spatial capture – recapture models.” *Ecology and Evolution* 8 (20): 10336–10344.

Hayes, D.B. and M.J. Monfils. 2015. “Occupancy modeling of bird point counts: Implications of mobile animals.” *Journal of Wildlife Management* 79 (8): 1361–1368.

Jeltsch, F., D. Bonte, G. Pe’er, B. Reineking, P. Leimgruber, N. Balkenhol, B. Schröder, C.M. Buchmann, T. Mueller, N. Blaum, D. Zurell, K. Böhning-Gaese, T. Wiegand, J.A. Eccard, H. Hofer, J. Reeg, U. Eggers, and S. Bauer. 2013. “Integrating movement ecology with biodiversity research - exploring new avenues to address spatiotemporal biodiversity dynamics.” *Movement Ecology* 1 (6): 1–13.

Johnson, F.A., G.S. Boomer, B.K. Williams, J.D. Nichols, and D.J. Case. 2015. “Multilevel learning in the adaptive management of waterfowl harvests: 20 years and counting.” *Wildlife Society Bulletin* 39 (1): 9–19.

Krementz, D.G., M.J. Conroy, J.E. Hines, and H.F. Percival. 1987. “Sources of variation in survival and recovery rates of American black ducks.” *Journal of Wildlife Management* 51 (4): 689–700.

MacKenzie, D.I., J.D. Nichols, J.E. Hines, M.G. Knutson, and A.B. Franklin. 2003. "Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly." *Ecology* 84 (8): 2200–2207.

Marra, P.P., E.B. Cohen, S.R. Loss, J.E. Rutter, and C.M. Tonra. 2015. "A call for full annual cycle research in animal ecology." *Biology Letters* 11 (8): 20150552.

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

Neilson, E.W., T. Avgar, A.C. Burton, K. Broadley, and S. Boutin. 2018. "Animal movement affects interpretation of occupancy models from camera-trap surveys of unmarked animals." *Ecosphere* 9 (1): e02092.

Nichols, J.D., W.L. Kendall, and G.S. Boomer. 2019. "Accumulating evidence in ecology: once is not enough." *Ecology and Evolution* 9 (24): 13991–14004.

Robinson, R.A., H.Q. Crick, J.A. Learmonth, I.M. Maclean, C.D. Thomas, F. Bairlein, M.C. Forchhammer, C.M. Francis, J.A. Gill, B.J. Godley, J. Harwood, G.C. Hays, B. Huntley, A.M. Hutson, G.J. Pierce, M.M. Rehfisch, D.W. Sims, M. Begoña Santos, T.H. Sparks, D.A. Stroud, and M.E. Visser. 2009. "Travelling through a warming world: Climate change and migratory species." *Endangered Species Research* 7 (2): 87–99.

Robinson, O.J., C.P. McGowan, P.K. Devers, R.W. Brook, M. Huang, M. Jones, D.G. Mcauley, and G. Zimmerman. 2016. “A full annual cycle modeling framework for American black ducks.” *Natural Resource Modeling* 29 (1): 159–174.

Rota, C.T., R.J. Fletcher, R.M. Dorazio, and M.G. Betts. 2009. “Occupancy estimation and the closure assumption.” *Journal of Applied Ecology* 46 (6): 1173–1181.

Runge, C.A., T.G. Martin, H.P. Possingham, S.G. Willis, and R.A. Fuller. 2014. “Conserving mobile species.” *Frontiers in Ecology and the Environment* 12 (7): 395–402.

Sawyer, H., C.W. Lebeau, T.L. McDonald, W. Xu, and A.D. Middleton. 2019. “All routes are not created equal: An ungulate’s choice of migration route can influence its survival.” *Journal of Applied Ecology* 56 (8): 1860–1869.

Schaub, M. and J.A. Royle. 2014. “Estimating true instead of apparent survival using spatial Cormack-Jolly-Seber models.” *Methods in Ecology and Evolution* 5 (12): 1316–1326.

Steenweg, R., J. Whittington, M. Hebblewhite, A. Forshner, B. Johnston, D. Petersen, B. Shepherd, and P.M. Lukacs. 2016. “Camera-based occupancy monitoring at large scales: power to detect trends in grizzly bears across the Canadian Rockies.” *Biological Conservation* 201: 192–200.

Stewart, F.E.C., J.T. Fisher, A.C. Burton, and J.P. Volpe. 2018. “Species occurrence data reflect the magnitude of animal movements better than the proximity of animal space use.” *Ecosphere* 9 (2): e02112.

3 Chapter 1: Modeling multi-scale occupancy for monitoring rare and highly mobile species

Publication history: A version of this chapter is published as: Emmet, R.L., R.A. Long, and B. Gardner. "Modeling multi-scale occupancy for monitoring rare and highly mobile species." Ecosphere 12 (7): e03637.

3.1 Introduction

Understanding the distribution and abundance of species, and how these population attributes change over time, is paramount to the study of ecology. Since their inception, occupancy models have been applied to the study of distribution, habitat associations, and relative abundance of species (MacKenzie et al. 2002). Researchers have extended occupancy models to examine dynamic occurrence patterns (including local colonization and extinction) with respect to covariates (MacKenzie et al. 2003), community richness (Dorazio and Royle 2005), and interspecies interactions (Rota et al. 2016). For some species, especially those which maintain territories, occupancy can be directly proportional to the number of individuals in a study area (MacKenzie et al. 2005). Thus, occupancy has been used for monitoring populations (Pavlacky et al. 2012, Steenweg et al. 2016, Linden et al. 2017, Berigan et al. 2018), and in some cases occupancy can be used as an index of abundance (MacKenzie and Nichols 2004), or to infer abundance from detection frequencies (Royle and Nichols 2003).

Despite widespread use and constant development, occupancy models remain difficult to fit for rare and highly mobile species. Occupancy models may greatly overestimate occupancy probabilities for species that occur at low densities and have large home ranges (Neilson et al. 2018, Stewart et al. 2018). This is especially true if a species leaves or enters sampling units during a primary sampling period (e.g., temporary emigration), thus violating the assumption of geographic closure required of standard occupancy models (Rota et al. 2009). When the closure

assumption is violated, occupancy probability is often re-interpreted as the probability of "use" (MacKenzie et al. 2002); however, this simply represents the probability that a species was present at a site at least once during the study period. Moreover, standard occupancy models may underutilize continuously collected data from survey methods such as track surveys, scat surveys, and camera traps. These data are often discretized into arbitrarily defined "encounter histories" under a standard occupancy model (Gompper et al. 2006, Kays and Slauson 2008, Guillera-Arroita et al. 2011, Burton et al. 2015). This discretization can potentially lead to overestimates of occupancy when detection probabilities are low and resulting detections are sparse (Guillera-Arroita et al. 2011), as the resulting loss of data may lead to underestimates of detection probability, thus inflating occupancy probability.

When they occur together, violations of the closure assumption and low detection probability among species with large home ranges and low density can lead to overestimates of occupancy and ultimately a non-linear occupancy-abundance relationship (Neilson et al. 2018). Multi-scale models separately estimate coarse-scale occupancy and fine-scale use in sample units, allowing researchers to estimate presence over multiple temporal or spatial scales (Nichols et al. 2008, Bled, Nichols and Altwegg 2013), which may have advantages over standard single-season occupancy models for monitoring changes in abundance. Occupancy is more linearly related to abundance over shorter time periods than over longer time periods (Ellis et al. 2014, Steenweg et al., 2018). Accounting for site use with multi-scale models (Nichols et al. 2008) may result in a more linear occupancy-abundance relationship, improving the power of occupancy models to monitor rare, highly mobile species. However, a model has not yet been developed to allow for continuous-time detections in a multi-scale occupancy model framework, allowing researchers to

address closure assumption violations and low detection probability for rare, highly mobile species simultaneously.

Our motivation for developing a novel multi-scale occupancy model was to enable monitoring of wolverines (*Gulo gulo*) in the Cascade Mountains of Washington, USA using camera traps. Wolverines are naturally recolonizing portions of their former range in the continental U.S., including the Washington Cascades (Aubry, McKelvey and Copeland 2007), but have recently been considered for listing under the Endangered Species Act due in part to their apparent reliance on persistent spring snow cover (Copeland et al. 2010). Thus, researchers are interested in monitoring changes in wolverine abundance over time. Occupancy models have been suggested as one approach for monitoring wolverine populations; however, Ellis et al. (2014) showed that occupancy models fail to detect trends in abundance for wolverines, unless changes are very large, and only detect trends if large numbers of sites are surveyed. Our objective was to develop a model to improve detection of trends in wolverine populations using data from camera traps, which is the most practical data collection method currently available for this low-density, highly mobile species.

We aimed to address two complications that arise when using occupancy models to monitor highly mobile species: (1) low detection probability of such species and (2) the tendency of these species to leave all but the largest sampling sites during a sampling season, which may violate the geographic closure assumption and makes interpretation of occupancy probabilities difficult. We propose a multi-scale occupancy model to address these issues. We incorporate a continuous-time detection process (Guillera-Arroita et al. 2011), which allows researchers to use all available data and avoid creating arbitrary detection periods, into a two-scale, latent occupancy process allowing changes in species' availability for detection (site "use") at an

occupied site over time (Nichols et al. 2008). We compared our model with a discrete-time multi-scale model in a case study of wolverines and in simulations. We conducted a power analysis to test our model's ability to detect trends in abundance. We compared power to detect trends (1) using a multi-scale metric derived from occupancy and use parameters and (2) using occupancy alone, under multiple scenarios of population change and sampling effort.

3.2 Methods

3.2.1 Continuous-time multi-scale occupancy model

In this hierarchical model, the occupancy process describes species presence at sites (i.e., sampling units) at two temporal scales: among primary sampling occasions (hereafter primary occasions) and within secondary sampling occasions (hereafter secondary occasions). Site occupancy is estimated at the temporal scale of primary occasions, and site use is estimated at the temporal scale of secondary occasions. We define *occupancy* in the asymptotic sense (Rota et al. 2009, Efford and Dawson 2012); that is, a site is considered occupied if any individual of the target species passed through it at any time during a primary occasion. We define shorter-term *use* such that a site is considered used if an individual is sufficiently close to the detector or survey route at a site to have a non-zero probability of detection during a given secondary occasion (Nichols et al. 2008, Bled et al. 2013). What we call use is sometimes called small-scale occupancy and represents the immediate presence, either spatially or temporally depending on the type of multi-scale model, of a species conditional on the species ever having been present (i.e., conditional on asymptotic occupancy; Nichols et al. 2008). When occupancy and use are defined in this way, it is possible for a site to be occupied but never used, e.g., if an individual is present at least once in a sampling unit but never close enough to a detector to have a non-zero probability of being detected (Bled et al. 2013). Moreover, under our definition of use, the same

individual can be detected at multiple sites during a secondary occasion, which is consistent with how use is defined in single-season occupancy models. The occurrence of a single individual at multiple sites would, however, violate the assumption of independent sites in occupancy models, so spacing of sites should be designed to minimize overlap of individuals between sites, if possible.

We assume geographic closure within secondary occasions. That is, if a site is used during a given secondary occasion, we assume that one or more individuals must have a non-zero probability of detection at that site throughout that secondary occasion. However, a site's use status is allowed to change between secondary occasions, which is what distinguishes a temporal multi-scale occupancy model from a single-scale model. While our model is conceptually similar to a single-scale, dynamic occupancy model, in which the modeler defines multiple primary occasions over the course of the study period, these models differ in several ways. First, our model allows for straightforward estimation of asymptotic occupancy, which single-scale dynamic occupancy models do not estimate across primary occasions. Second, single-scale dynamic occupancy models assume that occupancy is constant over the entire primary occasion and do not allow for variation in how much the site is actually used. The temporal multi-scale model allows for changes in presence at the secondary occasion or "use" level, allowing this process of short-term changes in presence to be separated from long-term occurrence (i.e., asymptotic occupancy). Researchers may define the length and number of primary and secondary occasions *a priori*, depending on the species' movement ecology (and corresponding likelihood of closure assumption violations) and the length of the time periods of interest (e.g., denning seasons). However, in order to estimate probability of use, each primary occasion must contain at least two secondary occasions.

A given site *i* has latent occupancy status z_i, which is occupied with probability ψ_i:

$z_i \sim Bernoulli(\psi_i)$,

such that $z_i = 1$ implies that a site is occupied.

Use (u_{ib}) is also latent and conditional on occupancy (Nichols et al. 2008); a site *i* is used in a given secondary occasion *b*, given it is occupied, with probability θ_{ib}:

$$u_{ib} \sim Bernoulli(z_i \theta_{ib}).$$

Note that sites that are not occupied cannot be used.

Finally, when a site is occupied and used within a secondary occasion (i.e., where u_{ib} = 1), the number of detection events in secondary occasion b at site i, y_{ib}, is:

$$y_{ib} \sim Poisson(u_{ib} \lambda_{ib} L_{ib})$$

where L_{ib} is the length of time of deployment of a camera at site i during secondary occasion b. The continuous-time detection process is a Poisson process and can be parametrized either in terms of the number of detections or the time between detections. We use the former parametrization to describe our model, as it is straightforward, requiring only counts of detections and a measure of effort expended at sites (e.g., time that cameras are deployed and functioning). The parameter for this Poisson process is often called a hazard rate (Ergon, Borgan and Vindenes 2018), and is the rate, per unit time or space, at which events (detections) occur. Thus, λ_{ib} is a detection rate rather than the expected number of detections.

Occupancy and use can be parametrized using local colonization and extinction probabilities. If there were more than one primary occasion, occupancy at site i in primary occasion j can be modeled as a function of the previous primary occasion's occupancy probability such that:

$$\psi_{ij} = (1 - \psi_{i(j-1)})\gamma + \psi_{i(j-1)}(1 - \epsilon)$$

where γ is the probability that a site not occupied in primary occasion $j - 1$ is occupied in primary occasion j and ϵ is the probability that a site occupied in primary occasion $j - 1$ is not occupied in primary occasion j. This implementation can also be reduced to a trend model:

$$\psi_{ij} = \rho\psi_{i(j-1)}$$

where ρ is a parameter controlling the degree of autocorrelation between occupancy at site i in primary occasion j and occupancy at site i in primary occasion $j - 1$.

Each of the parameters indexed by i can be modeled to include site-specific covariates and each parameter indexed by b can be modeled with secondary-occasion-specific covariates. For the sake of simplicity, we include no covariates or spatial or temporal processes in our model, but investigate mean occupancy ψ_0, use θ_0, and detection λ_0 in place of ψ_i, θ_{ib}, and λ_{ib}, respectively. However, it is possible to extend this model to relax certain assumptions, such as the assumption of independent detections in Poisson processes (Guillera-Arroita et al. 2011).

We derive a metric, the average occupancy probability ψ_0 multiplied by the average use probability θ_0, which is defined as the probability that a site is both occupied during a given primary occasion and used during a given secondary occasion. This metric, called "instantaneous occupancy" (sensu Efford and Dawson 2012) is essentially equivalent to an occupancy probability if each secondary occasion were treated as a season and a multi-season occupancy

model were fit within a primary sampling occasion. This is very similar to the probability that a site is occupied and that the immediate sampling area is also occupied in multi-method occupancy models (Nichols et al. 2008, Fisher and Bradbury 2014, Whittington et al. 2015), except that it accounts for multiple scales of occupancy and use in time, not in space. As such, this model allows researchers to estimate occurrence at multiple scales without more than a single detector (e.g., a camera) per site, by substituting time for space.

3.2.2 Wolverine case study

Wolverines generally occur at very low densities across their range and have large home ranges (Inman et al. 2012). In the Washington Cascades wolverines are naturally recolonizing parts of their former range (Aubry et al. 2007), and there is, therefore, interest in determining the best methods for monitoring changes in the population's abundance. Recent focus has been on using baited camera traps to estimate occupancy as a surrogate for abundance (MacKenzie and Nichols 2004).

We used data collected in the Cascades Mountains of Washington State to estimate and interpret the parameters in our model. The data were detections of wolverines at 25 baited camera stations in a study area of about 48,000 km^2 operated from December 1st, 2016 to March 31st, 2017. Each station was deployed before December 1st, 2016. Stations accessible by snowmobile during winter months (n = 18) were baited with a deer leg attached to a target tree, and a camera trap was directed at the target tree. These "accessible" stations were revisited monthly—on or around January 1st, February 1st, and March 1st—when they were rebaited, images were retrieved, and batteries replaced. The stations were removed on or around March 31st. Stations in locations that were not accessible (n = 7) during winter were outfitted with a camera trap and automated scent

dispenser that negated the need for rebaiting visits. These sites were deployed for the full 4-month survey duration.

We divided the survey into four 30-day secondary occasions. To decrease temporal dependence in detections, we retained only detections that occurred greater than 60 minutes apart at each survey station. We then counted the number of independent detections in each secondary period at each site, and used these counts to create an encounter history matrix for the continuous-time multi-scale model. Secondary occasions are still discretized in time, but the detections within those occasions are modeled in continuous time. We calculated the survey effort, L, for each survey site as the number of days that the camera was operational. We also fit a discrete-time, temporal multi-scale occupancy model, which is a modified version of the model proposed by Bled et al. (2013), without spatial autocorrelation or explicit colonization and extinction probabilities due to the sparsity of the data and the collection of data over a relatively short time period (one season). The discrete-time model we used estimated average occupancy and use probabilities over all sites, using multiple "repeat surveys" (tertiary sampling occasions) per secondary occasion. The discrete-time model includes an estimate of per-visit detection probability, p_0. This probability can be approximated, in the continuous-time model, using the formula $p_0 = 1 - \exp\frac{-\lambda_0 L}{K}$, where L is the amount of survey effort (e.g., time) that is expended at a site, and K is the number of repeat surveys (Guillera-Arroita et al. 2011). We binned detections into binary detection-non-detection records within each secondary occasion using 10-day periods. An effort covariate could be included to account for variation in sampling; however, we did not do this as only two out of these 300 tertiary occasions contained fewer than 10 days of sampling. While we could construct more, shorter tertiary occasions to avoid throwing out detections, constructing detection histories with more survey occasions can increase computation

time, while ultimately approaching but not outperforming inference under a continuous-time model (Guillera-Arroita et al. 2011).

We fit all models in a Bayesian framework, and used Uniform(0,1) priors for mean occupancy, use, colonization, extinction, and detection parameters (both the discrete-time detection probability and continuous-time detection rate). We fit models using JAGS version 4.2.0 (Plummer 2003) with the R package jagsUI version 1.5.1 (Kellner 2016) in R version 4.0.1 (R Core Team 2019). We ran 3 chains in parallel of length 50,000 for each model, with a burn-in of 20,000 iterations and a thin rate of 10 (due to autocorrelation in chains). We checked each model for convergence using the Gelman-Rubin convergence diagnostic.

3.2.3 Simulation

We ran simulations under a range of occupancy, use, and detection values. To approximate the survey design used to survey wolverines in the Cascades, we simulated occupancy data over one 120-day primary occasion, with four equal-length (30-day) secondary occasions nested within this primary occasion. If a site was used within a secondary occasion, we simulated detections according to a continuous-time Poisson process. For fitting discrete-time models, we sorted detections within a secondary occasion into three repeat surveys (tertiary occasions of 10 days each) and created a detection history for each secondary occasion.

We chose occupancy, use, and per-visit detection probabilities (p_0 above) to represent a range of possible values. For each of our three parameters, we simulated data under probabilities 0.2, 0.5, and 0.8. To simulate continuous-time data, we backtransformed per-visit detection probabilities into detection rates. For each combination of these parameters, we then simulated 200 data sets, each with surveys from 200 sites, and fit continuous-time and discrete-time models.

We fit all models in a Bayesian framework, and used Uniform(0,1) priors for mean occupancy, use, and detection parameters. We fit models using JAGS (Plummer 2003) with R package jagsUI (Kellner 2016, R Core Team 2019). We ran 3 chains in parallel of length 5,000 for each model, with a burn-in of 2,000 iterations and a thin rate of 2. We checked each model for convergence using the Gelman-Rubin convergence diagnostic, and recorded the computation time for each iteration. We calculated root mean square error (RMSE) and mean relative bias in each scenario, using the posterior mean estimates from each simulated data set to calculate RMSE and bias.

3.2.4 Power analysis

To test the ability of the continuous-time multi-scale model to detect changes in abundance over a longer time frame, we conducted a spatially explicit power analysis. We simulated population growth, movement, and detection of a hypothetical population of wolverines in an area roughly equivalent to the northern part of Washington's Cascades (north of I-90) in size (about 36 225-km^2 grid cells) using parameters based on empirically informed wolverine movement rates and space use (Appendix S1). We created 12 combinations of changes in abundance (increase and decrease of 25% and 50% over 10 years) and proportion of study area sampled (0.2, 0.5, 0.8), and simulated 100 data sets under each combination. We delineated four secondary periods, each 30 days in length, and 10 primary periods representing 10 years of survey effort. We constructed detection histories for all sampled sites, and fit continuous-time multi-scale models with both implicit (i.e., estimating occupancy independently in each primary occasion) and explicit (i.e., estimating colonization and extinction) occupancy dynamics between years. We fit linear models to derived posterior mean occupancy and instantaneous occupancy estimates in each simulation

replicate (Appendix S1), and summed the number of replicates in which a significant trend was detected to calculate simulated power in each scenario.

3.3 Results

3.3.1 Case study

There were 58 independent detections (defined using a 60-minute gap) at nine of the 25 sites in Washington state during our designated survey period. Of the nine sites with detections, four sites had only a single detection. Posterior mean estimates of long-term, asymptotic occupancy (ψ_0) were 0.619 (95% Bayesian credible interval (BCI): 0.374-0.987) and 0.717 (95% BCI: 0.367-0.984) for the continuous-time and discrete-time models, respectively, showing very similar results between models, and relatively high wolverine occupancy across the region over the 4-month sampling period (Table 3.1). However, estimates of use (immediate presence of a species given asymptotic occupancy), detection, and instantaneous occupancy (the product of use and asymptotic occupancy) were markedly different between models. Estimates from the continuous-time model were more precise than those from the discrete-time model (Table 3.1). Additionally, the discrete-time model's estimate of the probability that a wolverine is immediately present during a secondary occasion, given that the site is occupied (use), was about 50% higher ($\theta_0 = 0.318$; 95% BCI: 0.11-0.803) than the estimate from the continuous-time model ($\theta_0 = 0.216$; 95% BCI: 0.099-0.392). The estimates of instantaneous occupancy probability displayed similar patterns, because instantaneous occupancy is derived from use (Table 3.1).

Detection probabilities were very different between continuous-time and discrete-time models, with continuous-time detection probability (which is approximated, as it is converted from a detection rate) close to 0.8 and discrete-time detection probability close to 0.2. The continuous-

time detection probability is higher because the continuous-time model includes all detection events as data (n=58 detection events), not just detection-non-detection records (n=13 detections after creating 10-day tertiary occasions), and thus could include large clusters of detections that lead to higher estimates of detection probability (Guillera-Arroita et al. 2011). This difference in detection likely explains the differences in use and instantaneous occupancy estimates and their relative precision (Table 3.1).

3.3.2 Simulation

Model performance was generally good for both the continuous-time and discrete-time models. Both models estimated detection probability well; the continuous-time model provided more precise and less biased estimates of detection probability in general (overall RMSE was 0.028 for continuous-time detection and 0.035 for discrete-time detection; Appendix S2: Figure S1). The bias and precision of occupancy estimates were sensitive to use; while occupancy generally had small relative bias, with the largest bias occurring under low-use scenarios (Figure 3.1; Appendix S2: Table S2). Bias and precision of occupancy, however, were largely not sensitive to changes in detection probability, unless use was low (Figure 3.1; Appendix S2: Table S2). Use and instantaneous occupancy were almost always less biased under the continuous-time model than under the discrete-time model (Figure 3.2; Table 3.2; Appendix S2: Tables S3-S4), and bias in these quantities decreased substantially as use increased. The continuous-time and discrete-time models yield similar results partially because we simulated independent detections. In this case, detections would be distributed uniformly over time within a given secondary occasion, and thus converting to discrete-time models would generally result in less data loss than if highly clustered detections were discretized (as is common with real data, e.g., at baited camera traps).

3.3.3 Power analysis

Power to detect trends in abundance, using either occupancy or instantaneous occupancy as an index, was low in most scenarios. The highest power we documented was 63% to detect a 50% decline in abundance, using occupancy from an explicit dynamic occupancy model, over 10 years when 80% of cells were sampled (equivalent to 28 cameras over 7,125 km^2, or about 0.4 cameras per 100 km^2). Explicit (with local extinction and colonization rates) and implicit dynamic occupancy models differed in power when using instantaneous occupancy. When we used an explicit dynamic occupancy model, instantaneous occupancy provided lower power than occupancy in ninr out of 12 scenarios, only providing higher power when the population was increasing rapidly (50% in 10 years). Conversely, when we used an implicit dynamic occupancy model, instantaneous occupancy provided higher power than occupancy in every scenario, increasing power between 33% and 67% depending on the scenario (Figure 3.3).

3.4 Discussion

We developed a new continuous-time multi-scale occupancy model to estimate occupancy and use parameters while accounting for imperfect detection. We found that this model improved inference relative to a discrete-time multi-scale occupancy model in our wolverine case study, decreasing standard deviation of estimated use probabilities by upwards of 50% (Table 3.1), and also provided more conservative use and instantaneous occupancy estimates. Our model also suggested that asymptotic occupancy of wolverines in the Cascades is high, but short-term use of any given area is low, consistent with estimates of moderate occurrence probability from previous studies conducted in the region (Lukacs et al. 2020). The difference in precision between continuous-time and discrete-time models is due to the need to discard >75% of detection events to create detection histories for the discrete-time model (Guillera-Arroita et al.

2011). Given that discrete-time occupancy models of highly mobile species such as wolverines may bin detections into detection histories as long as 30 days to estimate detection probability (Lukacs et al. 2020), such differences in results may occur often. Inference from discrete-time models could be improved by creating longer detection histories with more tertiary occasions, each of shorter length (e.g., 10 3-day tertiary occasions instead of three 10-day tertiary occasions). However, even when shorter tertiary occasions are used, it is likely that the continuous-time model will slightly outperform a discrete-time model in terms of precision (Guillera-Arroita et al. 2011). Short tertiary occasions can also lead to increased computation time and extremely small detection probability which can be difficult to estimate, and they run a greater risk of violating assumptions of independent detections relative to longer tertiary occasions. While continuous-time models also assume independent detections, they offer computational advantages over discrete-time models, even in the presence of violated independence assumptions. Thus, there may be a tradeoff between increasing precision and violating independence assumptions when considering a continuous-time model or shorter tertiary occasions. This should be investigated by fitting both a continuous-time model and discrete-time models of varying tertiary occasion lengths if possible (Guillera-Arroita et al. 2011).

Based on simulation results, the continuous-time model performed well when detection rate and occupancy probability were moderate to high. The continuous-time model generally provided less biased estimates of use and instantaneous occupancy than the discrete-time model, particularly when detection rate was low. This corroborates the findings of Guillera-Arroita et al. (2011) that parameter estimates can be more accurate and precise for continuous-time models than for discrete-time models when little data is available. Moreover, computation times for

continuous-time models were generally lower than those for discrete-time models (0.187 minutes for continuous-time vs. 0.365 minutes for discrete-time per iteration, on average, on a 64-bit operating system with 8 GB of memory, using 3 of 4 available cores).

In addition to a fast and robust continuous-time detection process, our model estimates occupancy at multiple temporal scales. Using a temporal multi-scale model offers several potential advantages over a spatial multi-scale model, particularly for camera traps. First, using multiple temporal scales to separately estimate occupancy and use more closely matches the idea that occupancy and instantaneous occupancy occur on different temporal rather than spatial scales (Efford and Dawson 2012). Second, as mentioned previously, using different temporal scales to separately estimate occupancy and use allows the user to fit a multi-scale model without deploying multiple detectors per sampling unit (Nichols et al. 2008), saving time and money, and potentially allowing more sampling units to be surveyed. Finally, our model allows users to characterize changing presence at sites over long sampling periods, while still estimating an overall measure of the proportion of sites visited (i.e., asymptotic occupancy). This model formulation may be especially useful for species for which long sampling periods are required to achieve sufficient detections to accurately estimate parameters, as the model only assumes geographic closure over secondary occasions. This decreases the likelihood of violating closure assumptions within secondary occasions, as the model requires individuals to have a non-zero probability of detection over shorter time periods than if closure were assumed for entire primary occasions.

Despite our attempt to address closure issues with this model, the closure assumption is still likely to be violated within secondary occasions as well as within primary occasions for highly mobile species such as wolverines. The seriousness of the effects of these violations on

parameter estimates will depend on both the nature of the violation and the length of the secondary occasions. Species such as wolverines might be detectable at a specific camera station intermittently, but still be present in the survey site as a whole (Fisher and Bradbury 2014, Inman et al. 2012), or use the site only once during dispersal (Moriarty et al. 2009). Our case study results suggest that this may be true for wolverines in the Cascades, as both the continuous-time and discrete-time model results show that wolverines visit 60-70% of the Cascades study area over longer periods, but only consistently use 10-20% of the area at any given time (Table 3.1). Because highly mobile species such as wolverines are likely to temporarily leave all but very large sampling units, the length of secondary occasions should be species- and context-specific to minimize closure violations. Furthermore, we caution that the interpretation of the use parameter, and not the occupancy parameter, will change if the closure assumption is violated. The definitions of occupancy and use we adopt differ from the definitions commonly used to describe standard, single-season occupancy models, in which "use" is the suggested interpretation of occupancy if the closure assumption is violated (i.e., this interpretation is what we term asymptotic occupancy or occupancy).

The length of secondary occasion that minimizes violations of the closure assumption is likely strongly related to site size and the species' home range relative to site size. In particular, a secondary occasion should be no longer than the time it would take any individuals present within a site (sampling unit) to leave that site. This amount of time will be affected by where the individuals' activity centers are within the site (e.g., individuals with activity centers far inside a site will take longer to leave the site than individuals with activity centers at the edge of a site), and by the individuals' speed and movement behavior. Unfortunately, many of these factors are difficult to estimate. Even defining a site is difficult for camera trap studies; not only can camera

traps vary enormously in their effective detection radii (Hofmeester et al. 2019), but also, for baited camera traps like those used in wolverine surveys, sites are of indeterminate size, making interpretation of occupancy estimates difficult (Efford and Dawson 2012). We suggest fitting the model with varying lengths of secondary occasions to assess the model parameters' sensitivity to this length, and recommend considering how the spatial scale of the target species' movements affects the optimal length of secondary occasion. Where available, telemetry data could be used to define appropriate secondary occasion lengths for different site sizes.

Our power analysis revealed another advantage of the multi-scale model: not only does the continuous-time, multi-scale model provide relatively unbiased estimates of instantaneous occupancy, but these instantaneous occupancy estimates can also provide higher power to detect changes in abundance than occupancy alone when an implicit dynamic occupancy model must be used (i.e., a single-season survey design). On the one hand, when it is possible to use an explicit dynamic occupancy model, the ability to estimate colonization and extinction parameters allows researchers to explicitly investigate the processes driving changes in occupancy over time (MacKenzie et al. 2003), in addition to monitoring changes in abundance using occupancy as an index. On the other hand, when not every site can be sampled at regular intervals, or when there are not enough sites to estimate colonization and extinction parameters (McKann, Gray, and Thogmartin 2013), the combination of implicit dynamic occupancy models and instantaneous occupancy as a metric offers a viable alternative for detecting changes in abundance. However, our power analysis results suggest that power to detect changes in abundance of rare and highly mobile species remains low regardless of the choice of model structure and metrics. Power likely remains low because of low sample size (number of sites sampled); the number of sites sampled annually in our power analysis is no greater than 29 sites. We echo the findings of Ellis et al.

(2014) that increasing the number of sites sampled will increase power to detect changes in abundance using occupancy as a metric. Other improvements in metrics such as scaling occupancy by home range size might strengthen estimated occupancy-abundance relationships and improve power to detect changes in abundance as well (Rogan et al. 2019).

In situations when it is too expensive or logistically difficult to collect abundance data, or when occupancy surveys cannot be carried out at regular intervals, the continuous-time multi-scale occupancy model and its resulting instantaneous occupancy parameter can provide an index of abundance with detection non-detection data alone. Our model's ability to estimate probabilities of occurrence at multiple temporal scales with only one detector per site, and its strong performance under low data availability, makes it potentially useful for surveying elusive species in remote areas, such as wolverines (Ellis et al. 2014, Bischof et al. 2020), fishers (Linden et al. 2017), martens (Sheehy et al. 2018), and lynx (Whittington et al. 2015). This model can even be extended to permit the "use" parameter to vary by secondary occasion, such that researchers can investigate differences in short-term use between, for example, denning and non-denning seasons (Nichols et al. 2008). A model with time-varying use parameters would be more conceptually similar to a single-scale dynamic occupancy model, allowing instantaneous occupancy and use to vary over time while keeping asymptotic occupancy constant.

References

Aubry, K.B., K.S. McKelvey, and J.P. Copeland. 2007. "Distribution and broadscale habitat relations of the wolverine in the contiguous United States." *Journal of Wildlife Management* 71 (7): 2147–2158.

Berigan, W.J., G.M. Jones, S.A. Whitmore, R.J. Gutiérrez, and M.Z. Peery. 2018. "Cryptic wide-ranging movements lead to upwardly-biased occupancy in a territorial species." *Journal of Applied Ecology* 56 (2): 470–480.

Bischof, R., C. Milleret, P. Dupont, J. Chipperfield, M. Tourani, A. Ordiz, P. de Valpine, D. Turek, J.A. Royle, O. Gimenez, Ø. Flagstad, M. Åkesson, L. Svensson, H. Brøseth, and J. Kindberg. 2020. "Estimating and forecasting spatial population dynamics of apex predators using transnational genetic monitoring." *Proceedings of the National Academy of Sciences* 117 (48): 30531–30538.

Bled, F., J.D. Nichols, and R. Altwegg. 2013. "Dynamic occupancy models for analyzing species' range dynamics across large geographic scales". *Ecology and Evolution* 3 (15): 4896–4909.

Burton, A.C., E.W. Neilson, D. Moreira, A. Ladle, R. Steenweg, J.T. Fisher, E. Bayne, and S. Boutin. 2015. "Wildlife camera trapping: a review and recommendations for linking surveys to ecological processes." *Journal of Applied Ecology* 52 (3): 675–685.

Copeland, J.P., K.S. McKelvey, K.B. Aubry, A. Landa, J. Persson, R.M. Inman, J. Krebs, E. Lofroth, H. Golden, J.R. Squires, A.J. Magoun, M.K. Schwartz, J. Wilmot, C.L. Copeland, R.E. Yates, I. Kojola, R. May. 2010. "The bioclimatic envelope of the wolverine (Gulo gulo): do climatic constraints limit its geographic distribution?" *Canadian Journal of Zoology* 88 (3): 233–246.

Dorazio, R.M., and J.A. Royle. 2005. "Estimating size and composition of biological communities by modeling the occurrence of species." *Journal of the American Statistical Association* 100 (470): 389–398.

Efford, M.G., and D.K. Dawson. 2012. “Occupancy in continuous habitat.” *Ecosphere* 3 (4): Article 32.

Efford, M.G., D.K. Dawson, Y.V. Jhala, and Q. Qureshi. 2016. “Density-dependent home-range size revealed by spatially explicit capture-recapture.” *Ecography* 39 (7): 676–688.

Ellis, M.M., J.S. Ivan, and M.K. Schwartz. 2014. “Spatially explicit power analyses for occupancy-based monitoring of wolverine in the U.S. Rocky Mountains.” *Conservation Biology* 28 (1): 52–62.

Ergon, T., Ø. Borgan, and Y. Vindenes. 2018. “The utility of mortality hazard rates in population analyses.” *Methods in Ecology and Evolution* 9: 2046–2056.

Fisher, J.T., and S. Bradbury. 2014. “A multi-method hierarchical modeling approach to quantifying bias in occupancy from noninvasive genetic tagging studies”. *Journal of Wildlife Management* 78 (6): 1087–1095.

Gompper, M.E., R.W. Kays, J.C. Ray, S.D. Lapoint, D.A. Bogan, and J.R. Cryan. 2006. “A comparison of noninvasive techniques to survey carnivore communities in northeastern North America.” *Wildlife Society Bulletin* 34 (4): 1142–1151.

Guillera-Arroita, G., B.J. Morgan, M.S. Ridout, and M. Linkie. 2011. “Species occupancy modeling for detection data collected along a transect.” *Journal of Agricultural, Biological, and Environmental Statistics* 16 (50): 301–317.

Hofmeester, T.R., J.P. Cromsigt, J. Odden, H. Andrén, J. Kindberg, and J.D. Linnell. 2019. “Framing pictures: a conceptual framework to identify and correct for biases in detection

probability of camera traps enabling multi-species comparison." *Ecology and Evolution* 9 (4): 2320–2336.

Inman, R.M., M.L. Packila, K.H. Inman, A.J. McCue, G.C. White, J. Persson, B.C. Aber, M.L. Orme, K.L. Alt, S.L. Cain, J.A. Fredrick, B.J. Oakleaf, S.S. Sartorius. 2012. "Spatial ecology of wolverines at the southern periphery of distribution." *Journal of Wildlife Management* 76 (4): 778–792.

Kays, R.W. and Slauson, K. 2008. "Remote cameras." Pages 110-140 *in* R.A. Long, P. MacKay, J.C. Ray, and W.J. Zielinski, editors. Noninvasive survey methods for carnivores. Island Press, Washington, D.C., USA.

Kellner, K. 2016. "jagsUI: a wrapper around 'rjags' to streamline 'JAGS' analyses." Version 1.5.1. R package.

Linden, D.W., A.K. Fuller, J.A. Royle, and M.P. Hare. 2017. "Examining the occupancy-density relationship for a low-density carnivore." *Journal of Applied Ecology* 54 (6): 2043–2052.

Lukacs, P.M., D.E. Mack, R.M. Inman, J.A. Gude, J.S. Ivan, R.P. Lanka, J.C. Lewis, R.A. Long, R. Sallbanks, Z. Walker, S. Courville, S. Jackson, R. Kahn, M.K. Schwartz, S.C. Torbit, J.S. Waller, and K. Carroll. 2020. "Wolverine occupancy, spatial distribution, and monitoring design." *Journal of Wildlife Management* 84 (5): 841-851.

MacKenzie, D.I., and J.D. Nichols. 2004. "Occupancy as a surrogate for abundance estimation." *Animal Biodiversity and Conservation* 27 (1): 461–467.

MacKenzie, D.I., J.D. Nichols, J.E. Hines, M.G. Knutson, and A.B. Franklin. 2003. "Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly." *Ecology* 84 (8): 2200–2207.

MacKenzie, D.I., J.D. Nichols, G.B. Lachman, S. Droege, J.A. Royle, and C.A. Langtimm. 2002. "Estimating site occupancy rates when detection probabilities are less than one." *Ecology* 83 (2): 2248–2255.

MacKenzie, D.I., J.D. Nichols, N. Sutton, K. Kawanishi, and L.L. Bailey. 2005. "Improving inferences in population studies of rare species that are detected imperfectly." *Ecology* 86 (5): 1101–1113.

McKann, P.C., B.R. Gray, and W. Thogmartin. 2013. "Small sample bias in dynamic occupancy models." *Journal of Wildlife Management* 77 (1): 172-180.

Moriarty, K.M., W.J. Zielinski, A.G. Gonzales, T.E. Dawson, K.M. Boatner, C.A. Wilson, F.V. Schlexer, K.L. Pilgrim, J.P. Copeland, M.K. Schwartz. 2009. "Wolverine confirmation in California after nearly a century: native or long-distance immigrant?" *Northwest Science* 83 (2): 154–162.

Neilson, E.W., T. Avgar, A.C. Burton, K. Broadley, and S. Boutin. 2018. "Animal movement affects interpretation of occupancy models from camera-trap surveys of unmarked animals." *Ecosphere* 9 (1): e02092.

Nichols, J.D., L.L. Bailey, A.F. O'Connell Jr., N.W. Talancy, E.H. Campbell Grant, A.T. Gilbert, E.M. Annand, T.P. Husband, and J.E. Hines. 2008. "Multi-scale occupancy estimation

and modelling using multiple detection methods." *Journal of Applied Ecology* 45 (5): 1321–1329.

Pavlacky, D.C., J.A. Blakesley, G.C. White, D.J. Hanni, and P.M. Lukacs. 2012. "Hierarchical multi-scale occupancy estimation for monitoring wildlife populations." *Journal of Wildlife Management* 76 (1): 154–162.

Plummer, M. 2003. "JAGS: a program for analysis of Bayesian graphical models using Gibbs Sampling." Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), Vienna, Austria.

R Core Team. 2019. "R: a language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria.

Rogan, M.S., G.A. Balme, G. Distiller, R.T. Pitman, J. Broadfield, G.K.H. Mann, G.M. Whittington-Jones, L.H. Thomas, and M.J. O'Riain. 2019. "The influence of movement on the occupancy-density relationship at small spatial scales." *Ecosphere* 10 (8): e02807.

Rota, C.T., M.A. Ferreira, R.W. Kays, T.D. Forrester, E.L. Kalies, W.J. McShea, A.W. Parsons, and J.J. Millspaugh. 2016. "A multispecies occupancy model for two or more interacting species." *Methods in Ecology and Evolution* 7 (10): 1164–1173.

Rota, C.T., R.J. Fletcher, R.M. Dorazio, and M.G. Betts. 2009. "Occupancy estimation and the closure assumption." *Journal of Applied Ecology* 46 (6): 1173–1181.

Royle, J.A., and J.D. Nichols. 2003. "Estimating abundance from repeated presence-absence data or point counts." *Ecology* 84 (3): 777–790.

Sheehy, E., C. Sutherland, C. O'Reilly, and X. Lambin. 2018. "The enemy of my enemy is my friend: native pine marten recovery reverses the decline of the red squirrel by suppressing grey squirrel populations." *Proceedings of the Royal Society B: Biological Sciences* 285 (1874): 20172603.

Steenweg, R., M. Hebblewhite, J. Whittington, P.M. Lukacs, and K.S. McKelvey. 2018. "Sampling scales define occupancy and underlying occupancy-abundance relationships in animals." *Ecology* 99 (1): 172–183.

Steenweg, R., J. Whittington, M. Hebblewhite, A. Forshner, B. Johnston, D. Petersen, B. Shepherd, and P.M. Lukacs. 2016. "Camera-based occupancy monitoring at large scales: power to detect trends in grizzly bears across the Canadian Rockies." *Biological Conservation* 201: 192–200.

Stewart, F.E.C., J.T. Fisher, A.C. Burton, and J.P. Volpe. 2018. "Species occurrence data reflect the magnitude of animal movements better than the proximity of animal space use." *Ecosphere* 9 (2): e02112.

Whittington, J., K. Heuer, B. Hunt, M. Hebblewhite, and P.M. Lukacs. 2015. "Estimating occupancy using spatially and temporally replicated snow surveys." *Animal Conservation* 18 (1): 92–101.

Data Availability Statement

The data that support the findings of portions of this study are available from the Washington Department of Fish and Wildlife. Restrictions apply to the availability of these data, which were

used under license for this study. Please contact the department directly (Jeff Lewis at jeffrey.lewis@dfw.wa.gov) with data requests.

3.5 Tables and Figures

Table 3.1. Posterior mean estimates and standard errors (se) for results of continuous-time and discrete-time multi-scale occupancy models of wolverine distribution in the Cascades.

	Continuous-time		Discrete-time	
	mean	se	mean	se
Occupancy	0.619	0.179	0.717	0.174
Use	0.216	0.077	0.318	0.173
Detection probability	0.798	0.043	0.244	0.105
Instantaneous occupancy	0.125	0.037	0.218	0.116

Table 3.2. Mean relative bias of instantaneous occupancy probabilities (percent) from continuous-time and discrete-time multi-scale occupancy models for all occupancy and use scenarios when detection probability is 0.2.

Use	Occupancy	Continuous-time	Discrete-time
0.2	0.2	35.576	43.335
0.5	0.2	8.206	8.909
0.8	0.2	-2.078	-2.719
0.2	0.5	15.340	18.255
0.5	0.5	3.484	6.056
0.8	0.5	-1.201	-1.433
0.2	0.8	6.984	11.823
0.5	0.8	0.886	2.981
0.8	0.8	-2.584	-2.763

Figure 3.1.. Posterior mean estimates of occupancy probability from all simulation scenarios. True occupancy probability increases from the left panel (probability 0.2) to the right panel (probability 0.8) and is also marked by horizontal black lines in each panel. The dark, medium, and light gray boxplots represent scenarios with true use probability of 0.2, 0.5, and 0.8, respectively.

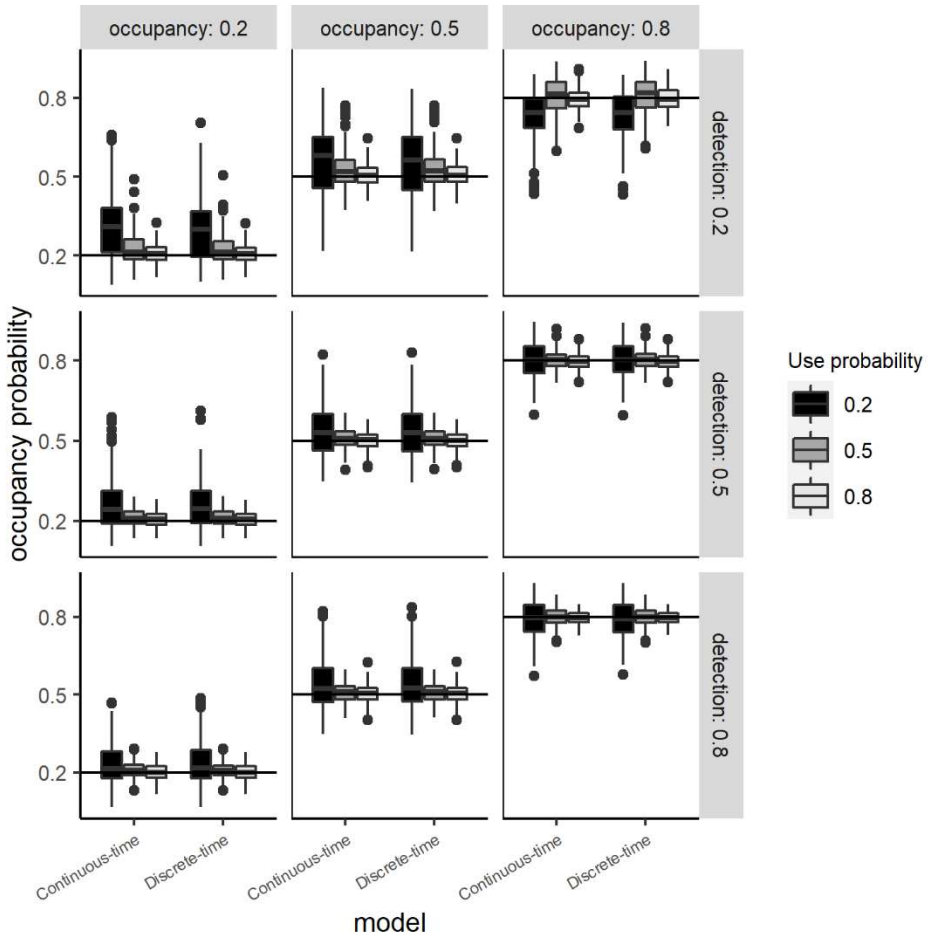


Figure 3.2. Posterior mean estimates of use probability from all simulation scenarios where true per-visit detection probability was 0.2. True occupancy probability increases from the left panel (probability 0.2) to the right panel (probability 0.8). The dark, medium, and light gray boxplots represent scenarios with true use probability of 0.2, 0.5, and 0.8, respectively, and black lines represent the true use values of 0.2, 0.5, and 0.8 as well.

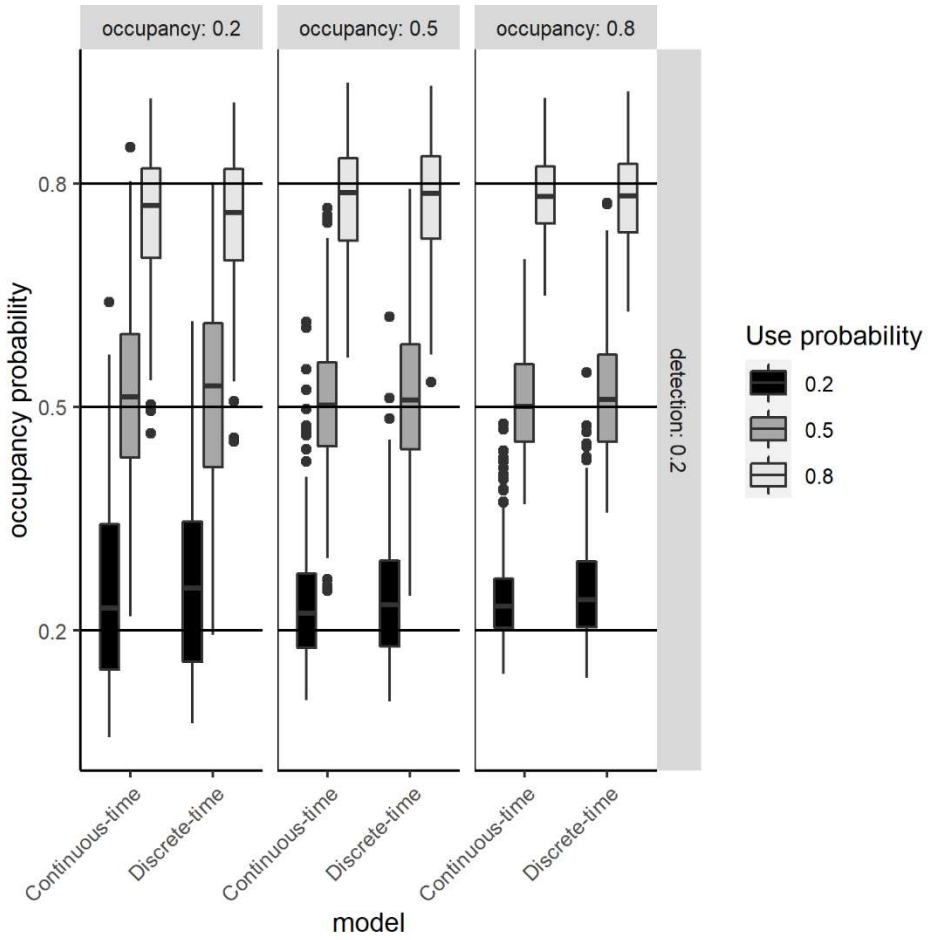
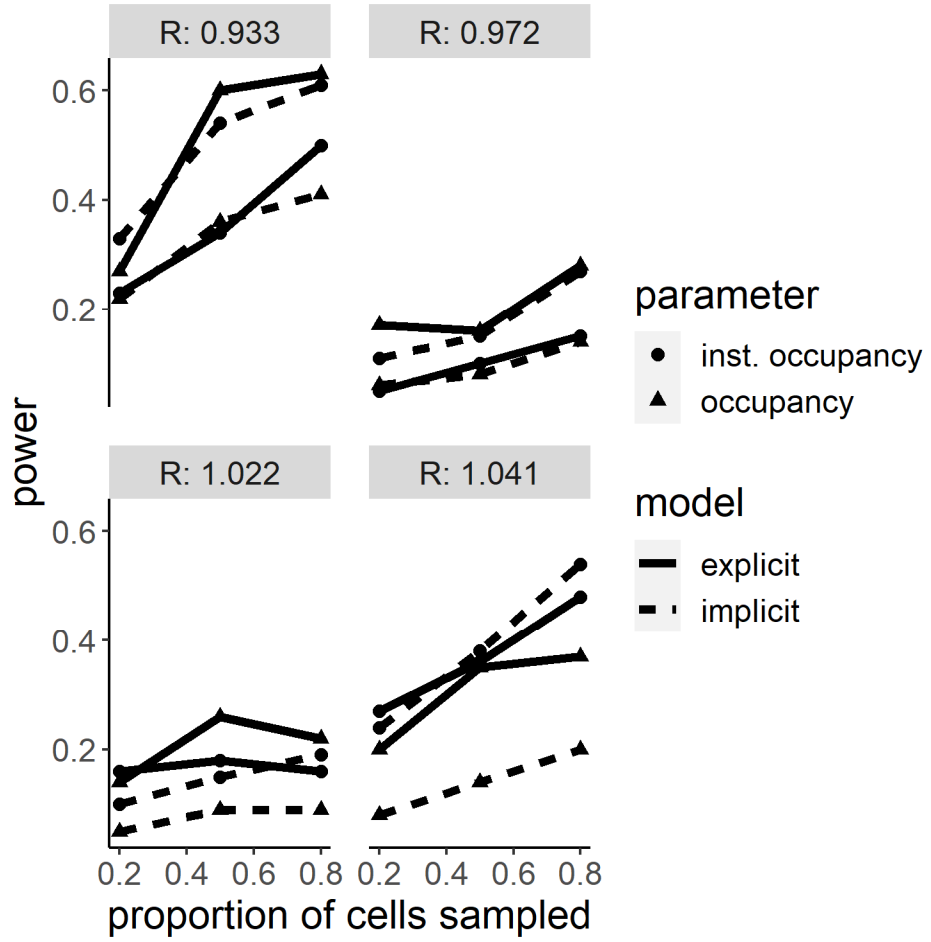


Figure 3.3. Estimates of power from 12 simulated scenarios of different population change and sampling area. The top two panels represent population declines and the bottom two panels represent population growth. The x-axis is the proportion of cells sampled and the y-axis is the power to detect a trend in abundance. The triangles (occupancy) and circles (instantaneous (inst.) occupancy) are shown for each scenario under the explicit dynamic occupancy model (solid lines), and implicit dynamic occupancy model (dashed lines).



4 Chapter 2: Developing an occupancy-based monitoring framework for wolverines in Washington's Cascades

Publication history: This chapter is not currently published or in review at a peer-reviewed journal.

4.1 Introduction

Wolverines (*Gulo gulo*) are naturally recolonizing much of their former range in the continental U.S., including the Cascades in Washington (Aubry, McKelvey, and Copeland 2007). However, they have recently been considered for listing under the Endangered Species Act (USFWS 2010) owing in part to their apparent selection for persistent spring snow cover (Copeland et al. 2010). It has been hypothesized that this habitat component is obligate for wolverines (Copeland et al. 2010), with spring snow cover providing increased reproductive denning success and carrion refrigeration (WDFW 2015, Inman et al. 2012). Because of this demonstrated reliance on cold, snowy environments, wolverines could be an indicator species for climate change.

Spring snow levels have already declined in the North Cascades in recent years (Mote et al. 2005), and spring snow cover may decrease further under predicted climate change scenarios (McKelvey et al. 2011). Such decreases across its southern range are expected to reduce overall wolverine distribution and connectivity in the continental U.S. by increasing isolation of suitable wolverine habitat (McKelvey et al. 2011). Washington's Cascades may be one of only a few "climate refuges" remaining for wolverines in the continental U.S. under some future climate scenarios (McKelvey et al. 2011). It is therefore crucial for natural resource managers in the Cascades to be able to detect long-term changes in the wolverine population (e.g., range contractions or expansions) so that they can identify and manage risks or benefits to the population.

Multiple studies suggest that changes in population and distribution can be detected using presence-absence data, even with unmarked individuals (Whittington et al. 2015; MacKenzie and Nichols 2004). In particular, occupancy modeling has been used to quantify changes in distribution, or, indirectly, abundance (Steenweg et al. 2016). However, Ellis, Ivan, and Schwartz (2014) have recently shown that occupancy-based monitoring may be insufficient to detect population declines of wolverines in the U.S. Rocky Mountains. They used a standard implicit dynamic occupancy model to assess power to detect trends. Their model measures how "asymptotic" occupancy (the probability that one or more individuals of a species is present and available for detection at a site at least once) changes between seasons, while assuming that shorter-term presence remains constant within a season. While Ellis, Ivan, and Schwartz (2014) found that long-term "asymptotic" occupancy may be insufficient to detect trends in abundance, Steenweg et al. (2018) have suggested that short-term presence, rather than long-term occupancy, may better track abundance.

To investigate whether short-term presence provides greater power to detect trends in abundance than does long-term occupancy, we a developed a continuous-time multi-scale occupancy modeling framework (Nichols et al. 2008; Whittington et al. 2015) designed to monitor rare species that are surveyed over long time periods (e.g., months). Our multi-scale occupancy model separately estimates asymptotic occupancy (called "occupancy") and short-term presence conditional on occupancy (called "use), and can include a derived parameter,"instantaneous occupancy", which is the probability that a given location is both occupied in the long term and used during a short-term, pre-defined survey occasion. In a previous study, we found that instantaneous occupancy can provide higher power to detect trends in abundance than asymptotic occupancy when an implicit dynamic occupancy model is used (i.e., occupancy is estimated

independently in each year over several years), while asymptotic occupancy provides higher power when an explicit dynamic occupancy model is used (i.e., when occupancy is estimated using a combination of colonization and extinction probabilities; Emmet et al. in press). However, we also found that power to detect trends in abundance was not particularly high in any of the sampling scenarios we examined. Because occupancy may not have a linear relationship with abundance, while estimated occupancy should have a roughly linear relationship with true occupancy, it may be worth revisiting under what sampling designs changes in distribution can be detected using occupancy models before reconsidering how to track changes in abundance.

Even after choosing an appropriate occupancy model for monitoring, however, researchers must still decide how best to carry out monitoring. In particular, researchers must decide on several aspects of sampling design, such as the number of sites to visit or the length of the sampling period (MacKenzie and Royle 2005). For instance, for rare species such as wolverines, MacKenzie and Royle (2005) suggest that "it is more efficient to survey more sampling units less intensively, while for a common species fewer sampling units should be surveyed more intensively." For multi-species survey efforts, including those involving rare species (Robinson, Cushman, and Lucid 2017), researchers may benefit from sampling more sites (rotating cameras if necessary) rather than leaving cameras stationary for a longer time (Robinson, Cushman, and Lucid 2017; Si, Kays, and Ding 2014). If a species is especially rare or highly clustered in distribution, adaptive cluster sampling (ACS) may be an appropriate sampling design that can be analyzed using occupancy models (Pacifici et al. 2016).

We developed a power analysis to compare different survey designs for wolverines in a large landscape. We simulated a range of distribution change and sampling scenarios for wolverines,

and calculated the ability of each simulated survey design to detect trends in occupancy. By examining trade-offs between cost and power, we developed several recommendations for a wolverine monitoring framework in the Washington Cascades.

4.2 Methods

4.2.1 Occupancy model

We used a previously developed continuous-time, dynamic multi-scale occupancy model (Emmet et al. in press). This model distinguishes between occupancy (whether a wolverine is present at all during a given primary sampling occasion) and use (whether a wolverine is present during a particular, shorter secondary sampling occasion) at separate time scales, where use is nested within occupancy. The detection process is a temporal Poisson point process, such that the number of detections at an occupied site that is used during a given secondary sampling occasion is a Poisson random variable (Guillera-Arroita et al. 2011). This allows researchers to use all available detections, rather than discarding detection data and simply reporting "detection" or "non-detection" during a given sampling occasion. We used an explicit dynamic occupancy model for this analysis, such that during each primary sampling occasion, occupancy was estimated as a function of the occupancy probability in previous primary sampling occasion and local site colonization and extinction probabilities (MacKenzie et al. 2003). We fit this model in a maximum likelihood framework using the optim() function in R.

4.2.2 Occupancy simulation

We simulated changes in occupancy over a study area that consisted of a grid of hexagonal cells consisting of >50% suitable wolverine habitat according to a habitat suitability layer derived from the results of Inman et al. (2012). This area spanned Washington state in the north-south direction and covered much of Washington's Cascade Mountains. In a given simulation scenario,

occupancy growth or decline was simulated over 11 years (11 120-day primary sampling occasions with 4 30-day secondary occasions each) to ensure that, at minimum, we could simulate the implementation of 3 surveys at 5-year intervals (years 1, 6, and 11). The starting occupancy probability, ψ_0, was used to simulate occupancy in cells designated as >50% suitable habitat. We simulated starting occupancy probabilities of 0.4 and 0.6. To investigate the effects of spatial patterns in occupancy changes, we simulated starting occupancy and changes in occupancy in one of two ways. In the "all cells change" scenario, we simulated starting occupancy and changes in occupancy randomly in all cells designated as >50% suitable habitat. In the "fringe cells change" scenario, we divided the study area into two sections. The northern section, corresponding roughly to the area of Washington's Cascade Mountains north of I-90 or a 100-mile radius around Harts Pass, started as containing 90% of the occupied cells in the study area, with the other 10% in the southern section. In this scenario, changes in occupancy were simulated in the southern section first, with changes only occurring in the northern section after the southern section was completely filled with or emptied of occupied cells. We simulated four different occupancy change scenarios: changes of -50%, -25%, 25%, and 50% over the first 10 years. We simulated annual changes in occupancy by calculating the number of occupied cells to be added or removed each year for each 10-year growth rate, and adding or removing that number of cells each year.

4.2.3 Detection simulation

First, we overlayed a grid of hexagonal cells of a fixed area on the landscape. Each of these grid cells' centers could hold a camera, which corresponds to an individual "site" for the purposes of occupancy modeling. We simulated two scenarios for the number of sites sampled in each survey: 40, 700, or 100. We also simulated two different grid cell sizes: 150 or 225 km^2.

Next, we simulated the placement of cameras. We simulated two sampling design scenarios: random sampling and generalized random-tessellation (GRTS) sampling. GRTS sampling can provide a more spatially balanced and efficient sample than random sampling (Stevens and Olsen 2004), and GRTS sampling is the design used by Lukacs et al. (2020) for a multi-state wolverine survey. Thus, we wanted to compare the monitoring capabilities of GRTS versus random sampling designs for long-term wolverine monitoring. In the random sampling design, camera sites were randomly selected at the beginning of the 11-year monitoring period. In the GRTS sampling design, camera sites were initially placed according to a GRTS sample across cells. Our sampling frame consisted of all grid cells with greater than or equal to 50% habitat suitability, similar to that of Lukacs et al. (2020).

We simulated detections under the assumptions of the continuous-time, multi-scale occupancy model, simulating use and detection over four 30-day secondary sampling occasions during each survey year. We simulated two scenarios each of use and detection; specifically, we simulated use probabilities of 0.2 and 0.8, and detection rates that would correspond to 10-day detection probabilities of ~0.2 and 0.8, respectively.

4.2.4 Power analysis

We created scenarios of combinations of population growth rate, proportion of cells sampled, grid cell size (which affects camera spacing and the absolute number of cells), and sampling design. For each scenario, we simulated 100 replicate data sets, fit multi-scale occupancy models to each data set, and recorded whether or not a statistically significant trend at the $\alpha = 0.1$ level was detected (in the same direction as the change in occupancy) in the occupancy parameters when a linear model was fit to them. We also checked for statistically significant trends in instantaneous occupancy parameters. To simulate different survey frequency patterns, we fit

linear models to different subsets of the data, corresponding to scenarios where sampling occurs every year, once every 3 years (years 1, 4, 7, and 10), once every 5 years (years 1, 6, and 11), or "2 years on, 2 years off" (years 1, 2, 5, 6, 9, and 10). Then, for each scenario, we calculated the proportion of replicates in which a trend was detected (power).

4.3 Results

We found that estimation of occupancy probabilities was poor when use was low (0.2); in these scenarios, 71% of the models we fit estimated at least one occupancy probability to be greater than 0.999, suggesting that these models lacked sufficient data to estimate occupancy probability. Thus, we interpreted only the results for scenarios in which use was high (0.8); in these scenarios, only 3% of the models we fit returned occupancy estimates greater than 0.999. Moreover, all of the models converged under high-use scenarios.

The variables considered can be divided into two types: those that researchers can control to some extent (cell size, number of cameras deployed, sampling design, detection rate, and survey frequency, though detection rate in particular can be increased but never fully controlled) and those that researchers cannot control (starting occupancy, spatial patterns of change, and use probability). Beginning with the variables that researchers can control, the most important variable for increasing power was the number of years of sampling (Figure 4.1). For instance, moving from sampling once every 5 years to sampling 2 years on, 2 years off, would increase power from 50% to 80%. Such a move would also provide the highest power short of sampling annually (of the scenarios considered). Most of the analyses of variables within researchers' control were consistent with the results of Ellis, Ivan, and Schwartz (2014). The number of cameras deployed had a substantial effect on power; power increased from 63% with 40 cameras to 73% with 70 cameras, and increased less steeply to 79% with 100 cameras. Cell size had little

impact on power (69% power for 150 km^2 versus 74% power for 225 km^2). This is consistent with previous results from Ellis, Ivan, and Schwartz (2014), which only show substantial declines in power for much larger grid cell sizes (such as 1,000 km^2). Similarly, the survey design had negligible impact on power (72% power for GRTS versus 72% power for random). Finally, increasing detection rate from low (equivalent to a 10-day detection probability of 0.2) to high (equivalent to a 10-day detection probability of 0.8) increased power from 54% to 89%.

Of the variables that researchers cannot control, by far the most influential was use probability. Increasing use from low (0.2) to high (0.8) increased power from 13% to 72%. The magnitude and direction of change in occupancy also mattered; power was higher when detecting increases rather than decreases in occupancy, and higher for larger changes (Figure 4.1). Neither starting occupancy nor spatial patterns of change had large effects on power. Increasing starting occupancy from 0.4 to 0.6 only increased power from 67% to 76%, while moving from an "all cells change" scenario to a "fringe cells change" scenario decreased power from 75% to 68%.

Comparing trend detection using occupancy and instantaneous occupancy, we found that occupancy outperformed instantaneous occupancy, but that the results they yielded were fairly similar when detection was high and surveys were infrequent (e.g., every 3 or 5 years; Figure 4.1). Instantaneous occupancy did sometimes outperform occupancy when use was low, though the high failure rate of models in these scenarios is likely a factor in this result.

4.4 Discussion

The results of the power analysis suggested that increasing the number of years of sampling was one of the most important factors in increasing power to detect trends in abundance. This was a major difference between our results and those of Ellis, Ivan, and Schwartz (2014), who found a greater effect of increasing the number of cells sampled than an effect of changing from

sampling every other year to sampling every year. It was surprising that sampling for 6 years (sampling for 2 years, then skipping 2 years) was almost as effective, in terms of power, as sampling for 10 years (sampling every year). Like Ellis, Ivan, and Schwartz (2014), we also found that increasing the number of cameras deployed and the detection rate of surveys can drastically increase power to detect changes. Broadly speaking, our results confirm those of Ellis, Ivan, and Schwartz (2014).

It is possible that our power analysis results are partially an artefact of the design of our simulations. We used a Type I error threshold of 0.1, which does increase the probability of a false positive (i.e., detecting a change in occupancy when none occurred) even as it increases power. Given the vulnerability of the wolverine population in Washington, we assumed that it would be better to detect a false trend and initiate further monitoring or overcommit management resources than to miss a change in distribution if it had occurred. We simulated detections directly from the multi-scale, continuous-time occupancy model developed in Chapter 1, which does make the power analysis sensitive to certain assumptions. For example, just as in Chapter 1 we found that inference about occupancy suffered when use was low, estimation and trend detection were difficult in this analysis when use was low as well. This would likely be worsened by violations of the closure assumption at the secondary occasion level, which would result in even lower use in practice. Additionally, the use probabilities and detection rates that we simulated are likely affected by a number of factors that we did not account for in our simulations, such as species' mobility (and how that changes with abundance or occupancy). However, we expect that our results are mostly valid given their similarity to the results of Ellis, Ivan, and Schwartz (2014) and Steenweg et al. (2016). Ellis, Ivan, and Schwartz (2014) and Steenweg et al. (2016) simulated the changing population metric and detections differently and

used a different method for detecting trends, yet still highlighted the importance of number of cameras deployed relative to other sampling considerations. Any artefacts of our results likely do not affect the general conclusion that the number of years and sites sampled are important factors in power to detect trends in abundance or occupancy, and that achieving high power to detect trends in abundance or even occupancy using occupancy models may involve prohibitively large sampling effort.

Both our power analyses (this analysis, Emmet et al. in press) and previous occupancy-based power analyses (Ellis, Ivan, and Schwartz 2014; Steenweg et al. 2016) show mixed results in terms of the ability of occupancy models to detect trends in abundance or even occupancy. There may be numerous reasons that changes in abundance and even occupancy can be difficult to detect using occupancy models. Our results and those of Steenweg et al. (2016) suggest that data availability can be a major factor. Our results reflect previous analyses by McKann, Gray, and Thogmartin (2013) and Steenweg et al. (2016) that suggest that it can be difficult to estimate occupancy and changes in occupancy with fewer than 60 cameras per deployment; even 70 cameras per deployment only yielded 60-70% power to detect changes of +/-25% in occupancy, which is close to 80% but still not necessarily sufficient. The use parameter also greatly affects inference and power, as it controls the number of detections available for inference and can result in biased occupancy estimates when use is low (Emmet et al. in press). Unfortunately, this use parameter is roughly a measure of how often a species is immediately present at a site, given that it ever visits the site, and thus is largely out of researchers' control. Our results suggest that surveys should be timed for when use is expected to be high, which might be when individuals shrink their home ranges and are thus visiting each area in their home range more frequently (e.g., during denning for some species).

One factor making occupancy-based monitoring difficult is that occupancy itself may be difficult to interpret under the sampling design assumed by Emmet et al. (in press) and Ellis, Ivan, and Schwartz (2014), namely a design with attractants at sites. Efford and Dawson (2012) note that attractants make it difficult or even impossible to determine the size of a "site" for the purposes of occupancy models. Without knowledge of how attraction to sites varies over space and time, it may be impossible to compare occupancy from different areas or times of year if attractiveness of sites varies. More research is needed on factors affecting effective site size at sites with attractants. Design factors such as the use of an automated scent lure dispenser (R. Long, unpubl. data) would increase the stability of effective site size over time, as regularly dispensed scent lure would not decay or be consumed over time in the way that a deer carcass or other attractants would. In an occupancy modeling framework, factors that affect the detection ability of scent detection dogs, such as wind direction and precipitation, might also be considered as detection covariates for occupancy models where sites have attractants (Reed et al. 2011). Modeling scent plumes (Bouchet and Meeuwig 2015) or using auxiliary data to account for the effects of attractants on animal movement and behavior (Stewart, Volpe, and Fisher 2019) may also help define how effective "sites" vary over space and time at sites with attractants. Occupancy models largely fail to yield consistent definitions of occupancy over time and space, and scaling issues affecting the occupancy-abundance relationship and the use of attractants may make it difficult to use estimates of occupancy that are linearly, or even consistently, related to abundance or true occupancy.

We did not consider any models of unmarked or marked abundance in this power analysis, but such models could be useful for detecting trends in abundance of wolverines if certain limitations are addressed. Some of these models, such as N-mixture models, assume that sites are

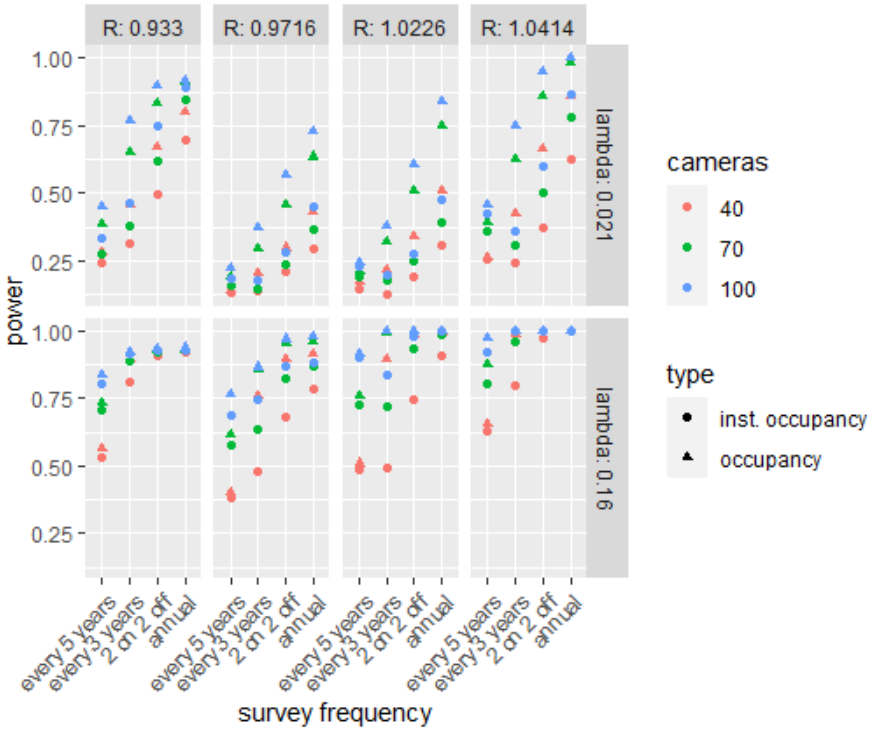
independent of one another. For highly mobile species such as wolverines, even a carefully thought-out sampling design with sites spaced far apart cannot guarantee that sites will be independent. In such cases, spatial models of occupancy or unmarked abundance (Lukacs et al. 2020; Zhao et al. 2016) may be appropriate to account for lack of independence between sites placed closer together. Many models of unmarked abundance, including N-mixture models (Royle 2004), random encounter and staying time (REST) models (Nakashima, Fukasawa, and Samejima 2018), and time- or space-to-event models (Moeller, Lukacs, and Horne 2018), assume that detections within a site are independent of one another. This may not be the case at sites with attractants, as animals could show a varying behavioral response to attractants over time; for instance, they may visit a site less often as an attractant decays, or as they grow accustomed to the attractant's scent and realize it is not an actual resource. The resulting heterogeneity in detection could lead to bias in some estimates of abundance (Barker et al. 2018; Moeller, Lukacs, and Horne 2018). Time-varying models of detection like inhomogeneous Poisson process models could account for variation in detection over time due to behavioral changes or decay of attractants; such an extension for time-varying detection has already been implemented for occupancy models (Guillera-Arroita et al. 2011). However, many models of unmarked abundance do not include such an extension (Moeller, Lukacs, and Horne 2018; Nakashima, Fukasawa, and Samejima 2018). Models of marked abundance such as spatial capture-recapture (SCR) would require a substantially different design than models of unmarked abundance or occupancy models, and feasible survey designs within Washington's Cascades may not yield sufficient spatial recaptures to estimate abundance or detect trends in abundance. Models of unmarked abundance may prove useful for monitoring trends in abundance of rare and

highly mobile species if sampling issues for rare species (such as the use of attractants) can be accounted for in these models.

In light of our results and those of previous power analyses, our recommendations for increasing power to detect trends in abundance are much the same as those of Ellis, Ivan, and Schwartz (2014). Increasing the number of survey years and increasing the number of cameras both increase power, though not necessarily sufficiently to detect small changes in abundance (Ellis, Ivan, and Schwartz 2014) or occupancy. Some tradeoffs in survey frequency and number of cameras may be necessary or even desirable. For instance, deployments of 70 cameras every 3 years can detect moderate (+/- 25%) changes in occupancy almost as well as deployments of 40 cameras on a "2 years on, 2 years off" schedule. Alternatively, other survey design changes that can increase detection rate or probability (e.g., automated scent lure dispensers or improved camera models) may lead to increases in power. Even with these improvements, however, both multi-season and multi-scale occupancy models, as we and Ellis, Ivan, and Schwartz (2014) currently formulate them, may not be sufficient to detect trends in occupancy without substantial sampling effort. Emerging or novel sampling methods (e.g., community science, eDNA) and models (e.g., models of unmarked abundance, multispecies models) may be required to detect trends in occupancy or abundance.

4.5 Tables and Figures

Figure 4.1. Estimates of power for different survey frequencies across scenarios of occupancy change rate (R, corresponding to -50%, -25%, 25%, and 50% change left to right), detection rate (lambda, corresponding to 10-day detection probabilities of 0.2 and 0.8 top and bottom), and cameras, with estimator types (inst. occupancy=instantaneous occupancy) shown as different shapes.



References

Aubry, K.B., K.S. McKelvey, and J.P. Copeland. 2007. "Distribution and broadscale habitat relations of the wolverine in the contiguous United States." *Journal of Wildlife Management* 71 (7): 2147–58.

Barker, R.J., M.R. Schofield, W.A. Link, and J.R. Sauer. 2018. "On the reliability of N-mixture models for count data." *Biometrics* 74 (1): 369–77.

Bouchet, P.J., and J.J. Meeuwig. 2015. "Drifting baited stereo-videography: a novel sampling tool for surveying pelagic wildlife in offshore marine reserves." *Ecosphere* 6 (8): 137.

Copeland, J.P., K.S. McKelvey, K.B. Aubry, A. Landa, J. Persson, R.M. Inman, J. Krebs, E. Lofroth, H. Golden, J.R. Squires, A.J. Magoun, M.K. Schwartz, J. Wilmot, C.L. Copeland, R.E. Yates, I. Kojola, I., and R. May. 2010. "The bioclimatic envelope of the wolverine (Gulo gulo):

do climatic constraints limit its geographic distribution?" *Canadian Journal of Zoology* 88 (3): 233–46.

Efford, M.G., and D.K. Dawson. 2012. "Occupancy in continuous habitat." *Ecosphere* 3 (4): Article 32.

Ellis, M.M., J.S. Ivan, and M.K. Schwartz. 2014. "Spatially explicit power analyses for occupancy-based monitoring of wolverine in the U.S. Rocky Mountains." *Conservation Biology* 28 (1): 52–62.

Emmet, R.L., R.A. Long, and B. Gardner. 2021. "Modeling multi-scale occupancy for monitoring rare and highly mobile species." *Ecosphere* 12 (7): e03637.

Guillera-Arroita, G., B.J.T. Morgan, M.S. Ridout, and M. Linkie. 2011. "Species occupancy modeling for detection data collected along a transect." *Journal of Agricultural, Biological, and Environmental Statistics* 16 (50): 301–17.

Inman, R.M., M.L. Packila, K.H. Inman, A.J. McCue, G.C. White, J. Persson, B.C. Aber, M.L. Orme, K.L. Alt, S.L. Cain, J.A. Fredrick, B.J. Oakleaf, and S. S. Sartorius. 2012. "Spatial ecology of wolverines at the southern periphery of distribution." *Journal of Wildlife Management* 76 (4): 778–92.

Lukacs, P.M., D.E. Mack, R.L. Inman, J.A. Gude, J.S. Ivan, R.P. Lanka, J.C. Lewis, R.A. Long, R. Sallabanks, Z. Walker, S. Courville, S. Jackson, R. Kahn, M.K. Schwartz, S.C. Torbit, J.S. Waller, and K. Carroll. 2020. "Wolverine occupancy, spatial distribution, and monitoring design." *Journal of Wildlife Management* 84 (5): 841–51.

MacKenzie, D.I., and J.D. Nichols. 2004. “Occupancy as a surrogate for abundance estimation.” *Animal Biodiversity and Conservation* 27 (1): 461–67.

MacKenzie, D.I., J.D. Nichols, J.E. Hines, M.G. Knutson, and A.B. Franklin. 2003. “Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly.” *Ecology* 84 (8): 2200–2207.

MacKenzie, D.I., and J.A. Royle. 2005. “Designing occupancy studies: General advice and allocating survey effort.” *Journal of Applied Ecology* 42 (6): 1105–14.

McKann, P.C., B.R. Gray, and W.E. Thogmartin. 2013. “Small sample bias in dynamic occupancy models.” *Journal of Wildlife Management* 77 (1): 172–80.

McKelvey, K.S., J.P. Copeland, M.K. Schwartz, J.S. Littell, K.B. Aubry, J.R. Squires, S.A. Parks, M.M. Elsner, and G.S. Mauger. 2011. “Climate change predicted to shift wolverine distributions, connectivity, and dispersal corridors.” *Ecological Applications* 21 (8): 2882–97.

Moeller, A.K., P.M. Lukacs, and J.S. Horne. 2018. “Three novel methods to estimate abundance of unmarked animals using remote cameras.” *Ecosphere* 9 (8): e02331.

Mote, P.W., A.F. Hamlet, M.P. Clark, and D.P. Lettenmaier. 2005. “Declining mountain snowpack in western North America.” *Bulletin of the American Meteorological Society* 86 (1): 39–50.

Nakashima, Y., K. Fukasawa, and H. Samejima. 2018. “Estimating animal density without individual recognition using information derivable exclusively from camera traps.” *Journal of Applied Ecology* 55: 735–44.

Nichols, J.D., L.L. Bailey, A.F. O'Connell Jr., N.W. Talancy, E.H. Campbell Grant, A.T. Gilbert, E.M. Annand, T.P. Husband, and J.E. Hines. 2008. "Multi-scale occupancy estimation and modelling using multiple detection methods." *Journal of Applied Ecology* 45 (5): 1321–9.

Pacifici, K., B.J. Reich, R.M. Dorazio, and M.J. Conroy. 2016. "Occupancy estimation for rare species using a spatially-adaptive sampling design." *Methods in Ecology and Evolution* 7 (3): 285–93.

Reed, S.E., A.L. Bidlack, A. Hurt, and W.M. Getz. 2011. "Detection distance and environmental factors in conservation detection dog surveys." *Journal of Wildlife Management* 75 (1): 243–51.

Robinson, L., S.A. Cushman, and M.K. Lucid. 2017. "Winter bait stations as a multispecies survey tool." *Ecology and Evolution* 7 (17): 6826–38.

Royle, J.A. 2004. "N-Mixture Models for Estimating Population Size from Spatially Replicated Counts." *Biometrics* 60 (1): 108–15.

Si, X., R.W. Kays, and P. Ding. 2014. "How long is enough to detect terrestrial animals? Estimating the minimum trapping effort on camera traps." *PeerJ* 2: e374.

Steenweg, R., M. Hebblewhite, J. Whittington, P.M. Lukacs, and K.S. McKelvey. 2018. "Sampling scales define occupancy and underlying occupancy-abundance relationships in animals." *Ecology* 99 (1): 172–83.

Steenweg, R., J. Whittington, M. Hebblewhite, A. Forshner, B. Johnston, D. Petersen, B. Shepherd, and P.M. Lukacs. 2016. "Camera-based occupancy monitoring at large scales: power to detect trends in grizzly bears across the Canadian Rockies." *Biological Conservation* 201: 192–200.

Stevens, D.L., and A.R. Olsen. 2004. “Spatially balanced sampling of natural resources.” *Journal of the American Statistical Association* 99 (465): 262-278.

Stewart, F.E.C., J.P. Volpe, and J.T. Fisher. 2019. “The Debate About Bait: A Red Herring in Wildlife Research.” *The Journal of Wildlife Management* 83 (4): 985–92.

Whittington, J., K. Heuer, B. Hunt, M. Hebblewhite, and P.M. Lukacs. 2015. “Estimating occupancy using spatially and temporally replicated snow surveys.” *Animal Conservation* 18 (1): 92–101.

Zhao, Q., E. Silverman, K. Fleming, and G.S. Boomer. 2016. “Forecasting waterfowl population dynamics under climate change- does the spatial variation of density dependence and environmental effects matter?” *Biological Conservation* 194: 80–88.

5 Chapter 3: Modeling effects of severe winter weather on survival and band recoveries of American black ducks

Publication history: This chapter is to be submitted for review at a peer-reviewed journal.

5.1 Introduction

Migratory birds using different migration routes, breeding areas, and non-breeding areas can experience vastly different pressures across their annual cycle (Runge et al. 2014). Variation in migratory movement choices, and resulting differences in exposure to risks and access to resources, can create spatial variation in demographic rates (Krementz et al. 1987; Duriez et al. 2009; Zhao, Boomer, and Kendall 2018) and processes such as density dependence (Zhao et al. 2016; Roy, McIntire, and Cumming 2016). Failing to account for the resulting spatial variation could bias estimates of demographic rates (Pollock and Raveling 1982) or lead to inaccurate understanding of drivers of demographic rates. Moreover, models of migratory species often do not model demographic rates over multiple seasons, and monitoring may not occur throughout the annual cycle (Marra et al. 2015). Seasonal variation in habitat use and risk factors can exert complex, sometimes competing pressures on demographic rates (Rushing, Ryder, and Marra 2016). Carryover effects further complicate population dynamics of migratory species, as conditions during one season can affect fitness and demographic processes during subsequent seasons (Haramis et al. 1986; Sedinger et al. 2011; Rockwell et al. 2017). Thus, models of demographic rates could be missing key spatial variation and seasonal factors driving changes in populations of migratory species (Marra et al. 2015).

Understanding spatial and temporal variation in demographic rates is crucial for the management of many harvested migratory species including American black ducks (*Anas rubripes*, hereafter black ducks), as these species face complex pressures that affect survival and other demographic rates (Diefenbach, Nichols, and Hines 1988; Duriez et al. 2009; Zhao, Boomer, and Kendall

2018) and require different management strategies throughout different portions of the annual cycle and species' range (Robinson et al. 2016). For example, Robinson et al. (2016) found that black duck populations would likely increase more if habitat management increased carrying capacity in the eastern US and western Canada than in eastern Canada or the interior US. Black ducks declined in abundance during the 20th century, and their recovery has been limited (Conroy, Miller, and Hines 2002). Researchers have hypothesized a number of factors that might have limited recovery, including breeding-season habitat quality, non-breeding-season habitat quality, and the effects of severe winter weather on survival (Rusch et al. 1989; Conroy, Costanzo, and Stotts 1989; Conroy, Miller, and Hines 2002). Modeling how these factors relate to changes in demographic rates over space and time can increase understanding of black duck population limitation and improve management outcomes (Robinson et al. 2016).

Severe winter weather has been investigated as another potential limiting factor on black duck survival, but previous analyses have proved inconclusive, and none have been conducted range-wide. Low winter temperatures, and particularly freezing conditions, affect wintering black duck habitat use (Morton et al. 1989b) and home range size (Ringelman et al. 2015). Studies of the effects of severe winter weather on survival have anecdotally reported mass die-offs following severe winter weather, and some recorded deaths due to hypothermia and starvation (Conroy, Costanzo, and Stotts 1989), but no link between severe winter weather and survival during late winter or subsequent seasons has been established (Conroy, Costanzo, and Stotts 1989; Newcomb et al. 2016), though these studies were small in terms of their spatial and temporal extents. This limits power to detect an effect, fails to capture potentially large variation in winter conditions across the black duck range, and makes it difficult to conduct range-wide inference. In

general, it has been difficult to study drivers of survival in black ducks at both broad spatial and temporal extents and fine scales.

Band recovery models (Brownie models; Brownie et al. 1985) use data from birds that are banded and subsequently recovered dead to estimate survival probabilities. Specifically, these models estimate band recovery probabilities (the probability that a bird is killed, retrieved, and its band is reported) and either annual survival probability (single-season Brownie models) or survival probability over hunting and non-hunting seasons (two-season Brownie models). These models can be used to increase knowledge of survival and harvest processes, including whether mortality is additive or compensatory (Arnold et al. 2016), or to determine drivers of variation in harvest vulnerability (Devers et al. In press). Single-season Brownie models have been developed to account for spatial variation and interannual variation in survival (Royle and Dubovsky 2001; Rice et al. 2010; Zhao, Boomer, and Kendall 2018), but none of these models account for variation in survival throughout the annual cycle (i.e., differences between hunting-season and non-hunting-season survival). Two-season Brownie models can estimate differing survival throughout the annual cycle (Arnold et al. 2016; Devers et al. In press), but incorporating spatial variation into survival and recovery probabilities has been limited in these models.

Extension of existing spatially and temporally explicit band recovery models to include both spatial and temporal variation throughout the annual cycle is feasible, but difficult for several reasons. Current spatially explicit band recovery models can require the delineation of pre-defined regions (Rice et al. 2010; Zhao, Boomer, and Kendall 2018), which limits the spatial scale of inference and prediction. Models using spatially correlated random effects may be able to estimate survival and recovery probabilities at arbitrarily fine scales (Royle and Dubovsky

2001), but these and other band recovery models often require the use of computationally costly Markov chain Monte Carlo (MCMC) methods (Royle and Dubovsky 2001; Arnold et al. 2016; Zhao, Boomer, and Kendall 2018). In addition, cost-effective bird banding requires banding in areas where the target species is likely to occur, possibly resulting in biased survival or recovery probabilities due to preferential sampling unless the process of banding site selection is accounted for in the band recovery model (Pollock and Raveling 1982).

Integrated nested Laplace approximation (INLA) is a computationally efficient Bayesian approach that can be used to estimate a variety of spatial and temporal random effects and correlation structures (Rue et al. 2017) and account for site selection bias (Pennino et al. 2019; Watson, Zidek, and Shaddick 2019; Sicacha-Parada et al. 2020). INLA has been used for a variety of ecological models, including species distribution models (Pennino et al. 2019), disease models (Jousimo et al. 2014), and other applications (Rue et al. 2017). Within this framework, using banding locations to approximate exposure to severe winter weather (as black ducks display high winter site fidelity, rarely migrating farther south during inclement weather; Diefenbach, Nichols, and Hines 1988), we can use a range-wide, spatially explicit model to investigate the effects of severe winter weather on survival and recovery probabilities of black ducks across large spatial and temporal extents.

Severe winter weather is measured using metrics of average or extreme values of variables like temperature (Schummer et al. 2010) and may be considered extreme from either a physical or biological perspective (Latimer and Zuckerberg 2019). Severe winter weather events and their effects may be acute or chronic (Latimer and Zuckerberg 2019). There are known relationships between extended periods of severe winter weather and habitat use, movements, and bio-energetics of black ducks (Morton et al. 1989a; Cramer 2009; Cramer et al. 2012; Jones et al.

2012; Ringelman et al. 2015) and reports of mass die-offs following winter storms (Conroy, Costanzo, and Stotts 1989). Lower body mass, which could be a result of severe winter weather, is also associated with decreases in black duck winter survival (Conroy, Costanzo, and Stotts 1989). Thus, we hypothesized that measures of chronic exposure to extreme (i.e., freezing) winter conditions would be related to non-hunting-season survival, such that survival would decrease as chronic exposure to freezing weather increased. However, we did not expect to observe any relationship between severe winter weather and recovery probability in the subsequent hunting season, if severe winter weather kills black ducks directly and relatively quickly via hypothermia and/or starvation before the hunting season (i.e., lowers non-hunting-season survival).

To test these hypotheses, we developed a novel approach to estimating non-hunting-season survival and recovery probabilities using a spatially explicit band recovery model built in INLA. We applied this model to banding and recovery data collected on black ducks over two seasons of each year 2010-2017. After accounting for spatial correlation, site selection bias, temporal effects, age and sex, and relative abundance, we investigated the effects of severe winter weather on non-hunting-season survival and recovery of black ducks.

5.2 Methods

5.2.1 Focal species and study area

Black ducks are partially migratory; those individuals who migrate generally complete spring migration February-May (Longcore et al. 2000; Figure 5.1). Nesting occurs March-May, and fall migration generally occurs September-January, which is also when black duck hunting seasons occur (Longcore et al. 2000). While black ducks are on their wintering grounds (roughly November-March, depending on the population), black ducks use a variety of freshwater and

saltwater habitats, selecting more for forested wetlands and subtidal habitats during periods of freezing weather (Morton et al. 1989b; Ringelman et al. 2015). Our study area includes the Mississippi and Atlantic Flyways east of 100.8° W longitude and south of 60° N latitude, which contain the majority of the global black duck population (Longcore et al. 2000; Figure 5.2).

5.2.2 Band recovery data

We considered banding and recovery data from a pilot two-season banding program for black ducks. From July 2009 to March 2017, black ducks were banded during two banding occasions: before the hunting season (July 1st-September 30th; hereafter pre-season), and after the hunting season (February 1st-March 10th; hereafter post-season). These banding seasons were designed to occur on either side of fall migration, which is generally when waterfowl hunting seasons occur in North America (Figure 5.1). Banding was conducted by federal, provincial, and state agencies and non-governmental organizations throughout the annual range of the black duck, though specific banding locations and effort in each region of the black duck range did vary by season (Devers et al. In press). All banding records were submitted to the Bird Banding Office in Canada and the USGS Bird Banding Lab (BBL); banding and subsequent recovery data were obtained from these sources. Further details on banding sampling design and field protocols are described in Devers et al. (In press).

We limited band recovery data to black ducks (BBL species code 1330) and black-duck-dominant mallard hybrids (BBL species code 1337). We also included adult (after hatch year or after second year, BBL age codes 1, 6, 7, or 8) males and females only, due to the difficulties in comparing post-season-banded juveniles (second year, BBL age code 5, which have survived one hunting season) and pre-season-banded juveniles (hatch year, BBL age codes 2, 3, or 4, which have yet to encounter a hunting season). For the purposes of summarizing data, we created

a grid of approximately 32 by 32 km based on modeled weather data from the North American Regional Reanalysis (NARR), provided by the Physical Sciences Laboratory, one of the Earth Systems Research Laboratories of NOAA's Office of Oceanic and Atmospheric Research, Boulder, Colorado, USA, from their Web site at https://psl.noaa.gov/data/gridded/data.narr.html. Using the NARR data grid as a reference, we summed, for each year, season, and cohort, the number of black ducks banded in each grid cell, and the number of black ducks banded in that grid cell that were recovered in the first hunting season following banding.

5.2.3 Weather and relative abundance data

We calculated multiple metrics of exposure to severe winter weather using North American Regional Reanalysis (NARR) modeled weather data. We used NARR daily mean surface air temperature data from February 1st-March 10th, 2010-2017 to calculate four candidate metrics of severe winter weather, based on both physical and biological thresholds for what might be considered extreme or severe weather (Latimer and Zuckerberg 2019). We calculated the mean temperature ($MEANTEMP$) and minimum temperature ($MINTEMP$) in each grid cell during each year. We also calculated a winter severity index, the 4-day freeze index ($FROZEN4$), which was shown to be correlated with changes in movement of black ducks (Ringelman et al. 2015). This index is the number of periods of 4 consecutive days in which the mean daily temperature was below 0° C (Celsius). Finally, we summed the number of days on which the mean daily temperature was below 0° C (Celsius) in each grid cell to create a "freezing days" index ($FREEZEDAYS$). These covariates were chosen to capture a range of categories of severe winter weather. $MEANTEMP$ measures chronic exposure to average conditions from a physical perspective, while $MINTEMP$ measures acute exposure to extreme conditions from a physical

perspective. $FROZEN4$ and $FREEZEDAYS$ measure chronic exposure to extreme conditions from biological and physical perspectives, respectively.

We estimated relative abundance of black ducks using observations from eBird, a community science platform allowing for collection of semi-structured data on bird occurrence and relative abundance (Sullivan et al. 2009). We used observations from complete checklists recorded in the Mississippi and Atlantic Flyway states during February 1st-March 10th, 2010-2017. To reduce variation between checklists, we used only stationary and traveling checklists with fewer than 10 observers per checklist, fewer than 5 hours' duration, and less than 5 km traveled. We fit a generalized additive model (GAM) using mgcv version 1.8-31 (Wood 2017) in R version 4.0.2 (R Core Team 2020), containing a linear effect of protocol type (stationary or traveling), linear and quadratic effects of the observation start time, and cubic spline smooths of duration (minutes), distance traveled (km), day of year, and number of observers, and a thin-plate spline of Easting and Northing. Full descriptions of the model and the data used can be found in Appendix S3. For each winter location included in our final band recovery model, we predicted the expected number of black ducks that would be observed on February 20th at that banding location, by 1 observer traveling 1 kilometer over 1 hour, beginning at 1000 hours local time. We used this predicted number as a relative abundance index ($RELABUND$) in our models.

For the purposes of extracting covariate values, we approximated locations of black ducks during February 1st-March 10th of their banding year as follows. For post-season-banded black ducks, we assumed that their banding location was a reasonable approximation of their exposure to severe winter weather during the period February 1st-March 10th of their banding year, given the high winter site fidelity of black ducks (Diefenbach, Nichols, and Hines 1988) and the relatively coarse resolution (~32 km) of the NARR data used to calculate winter severity metrics. We

assumed that pre-season-banded black ducks wintered at locations that were a weighted average of their recovery locations, and were exposed to winter conditions that were a weighted average of conditions at their recovery locations. Specifically, we assumed that pre-season-banded black ducks that were banded in the same grid cell also wintered in the same grid cell, and that the coordinates and winter conditions of this grid cell would be a weighted average of the coordinates and winter conditions of the recovery locations' grid cells, weighted by the number of black ducks recovered in each of the recovery locations' grid cells. Though we modeled recoveries of black ducks recovered throughout the hunting season (September-January), to impute winter locations, we specifically used recovery locations from all hunting seasons after banding, November-January. In imputing wintering locations of pre-season-banded black ducks in this way, we assumed that recovery locations during the later part of the hunting season (November-January) were a reasonable approximation to winter distribution (Diefenbach, Nichols, and Hines 1988; Zimpfer and Conroy 2006). We used a weighted average of winter conditions, rather than the winter conditions present at the weighted average of recovery location coordinates, to account for variation in winter conditions across recovery locations. In cells where no black ducks banded during pre-season banding were subsequently recovered, we used the recovery locations of the nearest neighboring pre-season-banded cell instead.

5.2.4 Spatially explicit band recovery model

5.2.4.1 Band recovery model development

To allow for estimation of non-hunting-season survival and recovery probabilities without having non-hunting-season survival be a non-linear function of covariates, we only model recoveries from the first hunting season after banding. This is essentially subsetting the data typically used in a two-season Brownie model. For instance, we only use band recoveries from

September 2010-January 2011 for black ducks banded February-March 2010 and July-September 2010. This is different from how years are recorded under a standard two-season Brownie model. Under a two-season Brownie model, "hunting season 2010" would cover September 2010-January 2011, "pre-season 2010" would cover July-September 2010, and "post-season 2010" would cover February 1st-March 10th, 2011. Under our model "post-season 2010" would instead be the post-season banding occasion falling within calendar year 2010: February 1st-March 10th, 2010. Thus, for a given year of data under our model, the timeline moves from post-season banding in February-March, to pre-season banding during July-September, to the hunting season beginning in September of that same calendar year and ending in January of the next calendar year. While this approach does not allow for the separate estimation of recovery probability and hunting-season survival probability, it does allow for efficient estimation of non-hunting-season survival and recovery probabilities.

Traditional two-season Brownie models use a multinomial likelihood to model the number of birds in a cohort banded in pre-season (N_t) and post-season (M_t) of year t, and the number of band recoveries during the hunting season beginning in year t (N_t and M_t for pre- and post-season, respectively; Table 5.1). Under a standard two-season Brownie model, this likelihood allows for estimation of non-hunting-season survival (S_t^{nh}), hunting-season survival (S_t^h), and recovery (f_t) probabilities during year t. By restricting the analysis to only modeling recoveries during the calendar year of banding, we can use the Poisson distribution as an approximation to estimate survival and recovery probabilities. Specifically, using the expected values shown in Table 5.1, we can derive the following Poisson approximations:

$$n_t \sim Poisson(N_t f_t)$$

and

$$m_t \sim Poisson(M_t S_t^{nh} f_t)$$

If N_t and M_t are treated as fixed, known quantities (offsets), a Poisson regression with a log link function can be used to model the effects of covariates on S_t^{nh} and f_t.

We consider a grid of cells $i = 1,2, \dots, I$ over the study area, and a sampling situation in which $N_{i,t,c}$ black ducks are banded in grid cell i in year $t = 1,2, \dots, T$, during pre-season from cohort $c = 1,2$ (where 1 = adult male and 2 = adult female), and $M_{i,t,c}$ are banded during post-season. $n_{i,t,c}$ and $m_{i,t,c}$ the number of black ducks banded in grid cell i, year t, and cohort c that were recovered in the hunting season beginning in year t during pre- and post-season, respectively (e.g., $m_{i,t,c}$ for 2010 would be the number of black ducks banded February-March 2010 that were recovered during September 2010-January 2011). We then use the following Poisson approximations:

$$n_{i,t,c} \sim Poisson(N_{i,t,c} f_{i,t,c})$$

and

$$m_{i,t,c} \sim Poisson(M_{i,t,c} S_{i,t,c}^{nh} f_{i,t,c})$$

We considered four competing models; each model included a different winter severity covariate. For the purposes of illustration, we describe the model including the covariate $FREEZEDAYS$; this covariate would be replaced by $MEANTEMP$, $MINTEMP$, or $FROZEN4$ in each of the competing models. Band recoveries of pre-season- and post-season-banded black

ducks were modeled, using $I(s = 2)$ to denote parameters assigned to post-season-banded black ducks, as:

$$log(b_{i,t,c}) = \beta_{0,t,c} + \beta_1 FREEZEDAYS_{i,t} + \beta_2 RELABUND_{i,t} + \beta_{3,t,c} I(s = 2) + \beta_4 I(s = 2) FREEZEDAYS_{i,t} + \beta_5 I(s = 2) RELABUND_i + \eta_i.$$

$b_{i,t,c}$ is the probability that a black duck is recovered during the first hunting season after it is banded, which is $f_{i,t,c}$ for pre-season-banded black ducks and $S_{i,t,c}^{nh} f_{i,t,c}$ for post-season-banded black ducks. The post-season-specific ($I(s = 2)$) parameters describe relationships between covariates and non-hunting-season survival. Here $\beta_{0,t,c}$ and $\beta_{3,t,c}$ are cohort- and year-specific random effects for pre-season- and post-season-banded black ducks, respectively, with

$$\beta_{0,t,c} \sim Normal(\overline{\beta}_0, \sigma_{\beta_0}^2)$$

and

$$\beta_{3,t,c} \sim Normal(\overline{\beta}_3, \sigma_{\beta_3}^2).$$

We use the INLA approach to estimate a spatial field component of recovery probability, denoted by η_i. Specifically, η_i is a spatially varying but temporally constant random effect:

$$\eta_i \sim Normal(0, Q).$$

with precision matrix Q. We used the stochastic partial differential equation (SPDE) approach, where the solution to an SPDE is estimated which approximates a Matérn covariance function (Lindgren, Rue, and Lindström 2011). This covariance function can be parametrized in terms of the range (the approximate distance at which correlation between points falls to 0.1; Lindgren,

Rue, and Lindström 2011) and the marginal standard deviation (the standard deviation of the spatial field at points at 0 distance from each other; Fuglstad et al. 2019).

In this model, the spatial random effect η_i represents some combination of unmodeled spatial covariates that account for spatial variation in recovery probability across different wintering locations. These covariates are assumed to be relatively constant over time (e.g., habitat availability). Because η_i is being used to model spatial factors affecting recovery of both pre-season-banded and post-season-banded black ducks, it represents factors that affect some combination of non-hunting-season survival and recovery probabilities for black ducks wintering at a spatial location.

Using a Poisson distribution and log link function does allow for probabilities that are greater than 1 in theory, because the log link function only restricts responses to be greater than 0. However, given the relatively low recovery rates of black ducks (Devers et al. In press), we did not anticipate observing estimated probabilities greater than 1. Common link functions for the binomial distribution, the logit, probit, and complementary log-log, would allow for restriction to the probability scale but result in non-hunting-season survival being a non-linear function of potential covariates, making it difficult to interpret parameters and separately interpret effects of covariates on survival and recovery probabilities.

Based on our hypotheses, if recovery probability has no relationship to severe winter weather covariates such as $FREEZEDAYS$, we predict that β_1 would have a Bayesian credible interval overlapping 0 (owing to the Bayesian approach we take within INLA). If non-hunting-season survival has a relationship to severe winter weather covariates, we predict that β_4 would have a Bayesian credible interval that does not overlap 0, and would be positive for $MEANTEMP$ and

$MINTEMP$ (i.e., increasing mean or minimum temperature is associated with an increase in non-hunting-season survival probability) and negative for $FROZEN4$ and $FREEZEDAYS$ (i.e., increasing the number of 4-day frozen periods or total freezing days is associated with a decrease in non-hunting-season survival probability).

5.2.4.2 Modeling site-selection bias

Our method of estimating late winter (February 1st-March 10th) locations of black ducks to measure exposure to winter weather may lead to bias that must be accounted for. Specifically, the winter locations or "sites" selected are a combination of post-season banding locations and (for pre-season-banded black ducks) a weighted average of recovery locations, either of which may be located in areas that are subject to higher proportions of bands recovered. Because banding effort was focused in areas of high black duck density to ensure sufficiently large sample sizes (Devers et al. In press), and because hunting effort may be higher in areas of high black duck density as well, it is possible that these sites provide a biased view of black duck survival and recovery probabilities. For instance, if sites were selected to occur in higher-quality habitat or habitats with higher population density, the same factors that affected site selection may also have influenced survival or recovery by way of density dependence or habitat quality effects. Thus, spatial random effect estimates may be biased if the process by which sites are chosen is not taken into account (Watson, Zidek, and Shaddick 2019).

We address this potential source of bias by jointly modeling the band recovery and site selection processes, using a shared spatial field to account for dependence between the two processes (Watson, Zidek, and Shaddick 2019). We model a site selection indicator variable R_i, which equals 1 if grid cell i is included in the sample (i.e., as a winter location or site where the band

recovery process was observed in the model) and 0 otherwise, as the outcome of a Bernoulli trial:

$$R_i \sim Bernoulli(p_i)$$

where p_i is a site selection probability. This probability is a function of relative abundance and the spatial field η_i:

$$logit(p_i) = \gamma_0 + \gamma_1 RELABUND_i + \gamma_2 \eta_i.$$

The spatial field η_i is the same one that occurs in the likelihood of the band recovery model, but scaled by γ_2 to account for the different scales of the responses in the band recovery and site selection models. Under this framework, the observation process of band recoveries is conditional on site selection (Watson, Zidek, and Shaddick 2019).

We selected 790 pseudosites systematically, choosing 1 out of every 4 NARR grid cells east of ~100.8° W longitude (-200,000 meters for UTM zone 16 north) and south of ~60° N latitude (7,000,000 meters for UTM zone 16 north), where banding did not occur and which were on land.

5.2.5 Data analysis with INLA

All data analyses were conducted in R version 4.0.1 (R Core Team 2020). We created a mesh based on the locations of both observed locations and pseudosites. To increase computational efficiency and minimize the effect of study area boundaries on spatial random effect predictions, we chose to not assign unique points in the mesh to points within 384 km (approximately 8-12 grid cells) of each other, and we chose a boundary around the study area of 4,000 km (Lindgren and Rue 2015). We used INLA's default priors for all fixed effects, which are mean 0 for both intercept terms and coefficients of covariates, and precision of 0 and 0.001 for intercept terms

and coefficients of covariates, respectively. We used $Gamma(0.01, 0.01)$ priors for the precision of the cohort-year random effects. We used penalized complexity (PC) priors for the range (ρ) and marginal standard deviation (σ) parameters of the spatial field (Fuglstad et al. 2019). PC priors for a spatial field with Matern covariance in INLA are designed to penalize deviation from a base, non-spatial model with infinite range and 0 marginal variance (Fuglstad et al. 2019). Following a sensitivity analysis (Supplementary Material Figures S1, S2), we chose PC priors such that $P(\rho < 128000) = 0.05$ and $P(\sigma > 5) = 0.05$, implying that a range less than 128 km or a marginal standard deviation greater than 5 was unlikely.

All models included the spatial field, interaction effects of cohort, year, and banding season, and$RELABUND$. Because winter severity covariates were highly correlated, we only included one covariate in each model. We scaled all covariates to improve model convergence and compare effect sizes. Specifically, we scaled winter severity covariates at all banding sites, and eBird relative abundance at all banding sites and pseudosites. We used the Wakanabe-Akaike information criterion (WAIC, Watanabe 2013) and the negative log sum of conditional predictive ordinates (CPO) to compare models; smaller WAIC and CPO both indicate better model fit. We assessed covariate statistical significance by noting whether 95% credible intervals included 0.

5.3 Results

We included 24,794 banded black ducks in the analysis; of those, 841 were recovered in the first hunting season after banding and thus included in the model as recoveries. There were 286 unique wintering locations (post-season banding locations or weighted averages of recovery locations for post-season-banded and pre-season-banded black ducks, respectively) included in the final model. Of these locations, 208 had birds subsequently recovered from them.

Based on the WAIC and CPO rankings (Table 5.2), the model with $FREEZEDAYS$ was considered the top model in our candidate set; the full results of the parameter estimates from this model are shown in Table 5.3. We found no association between $FREEZEDAYS$ (or any of the other winter severity covariates) and non-hunting-season survival probability (see Table 5.3). We did, however, find that $MINTEMP$, $FROZEN4$, and $FREEZEDAYS$ were all significantly associated with recovery probability. For example, for the $FREEZEDAYS$ model, an increase of 1 freezing day was associated with a 1.8% increase in recovery probability in the subsequent hunting season.

We fit the model to predict eBird relative abundance using 43,531 checklists from 16,047 observers. In the eBird relative abundance model, all covariates except for day of year were significant predictors of eBird relative abundance (Appendix S3: Tables S1, S2). Predicted relative abundance was highest in many regions that had relatively high predicted recovery probabilities, including the northeastern US states and Nova Scotia (Figure 5.2). $RELABUND$ was negatively associated with both non-hunting-season survival and recovery probabilities in the band recovery model (Table 5.3); none of these relationships were significant, though the posterior probability that the effect of $RELABUND$ on recovery was less than 0 (implying that higher relative abundance was associated with lower recovery probability) was 0.92. However, $RELABUND$ was significantly positively related to site selection probability, implying that banders are more likely to select sites for banding where there are more black ducks (Table 5.3).

The range and standard deviation parameters of the spatial random effect on recovery probability were similar across all four competing models; as an example, for the $FREEZEDAYS$ model, the mean range was about 1,860 km and a mean marginal standard deviation was about 3.85 (Table 5.3)). Site selection probability also increased as the spatial random effect increased, implying

that proportions of bands recovered were higher from locations where black ducks were banded during the post-season banding period and where pre-season-banded black ducks were being recovered (i.e., where they were estimated to have spent the previous winter; Table 5.3). Taking the average survival and recovery probabilities across cohorts and years, and under average conditions of severe winter weather and relative abundance, the predicted proportion of bands recovered was highest in Nova Scotia, the Northeast US, southern Ontario and Michigan, and some regions of the Midwest US (Figure 5.3). Thus, the spatial field indicates areas of relatively high recovery and spatial correlation in proportion of bands recovered in the North Atlantic region, a core area of the black duck winter range (Robinson, McGowan, and Devers 2016; Devers et al. In press).

The inclusion of a site selection sub-model to account for preferential sampling altered model estimates substantially. This can be demonstrated with the results of the model containing the $FREEZEDAYS$ variable. Without a site selection component, the spatial field moved closer to a null model of infinite range and zero variance; specifically, the estimated range increased and the estimated variance decreased relative to the spatial field when the site selection component was included (Table 5.4). When a site selection component was added, the variance of the spatial model increased substantially, the fixed-effect intercept (representing average recovery probability across all sites, years, and cohorts) increased, and the fixed effect representing the effect of severe winter weather on recovery probability increased in effect size and moved from being non-significant (i.e., the 95% credible interval overlapped 0) to significant (Table 5.3).

5.4 Discussion

We developed a novel implementation of a band recovery model that allowed us to estimate spatial and temporal patterns of non-hunting-season survival and recovery probabilities. Our

results did not support our hypothesis that measures of chronic exposure to extreme winter weather conditions would be related to non-hunting-season survival. In all four models, the credible intervals for the coefficient describing the relationship between severe winter weather and survival included 0. However, the signs of the coefficients matched what we predicted, with non-hunting-season survival decreasing as severe winter weather increased in duration or intensity (Table 5.3). Thus, it is possible that there is a relationship between severe winter weather and non-hunting-season survival that we did not have sufficient power to detect. There are several plausible explanations for why severe winter weather may not be related to non-hunting-season survival, or why we may not have detected a relationship if it were present. The timing of banding relative to occurrence of severe winter weather may be a factor. Conroy, Costanzo, and Stotts (1989) observed that much of the non-hunting mortality that they recorded occurred during a prolonged cold period in December-January, before post-season banding would occur. Two-season Brownie models condition on birds being alive during the occasion in which they are banded. Thus, our model could be failing to detect increased mortality during late hunting season (early winter). Changes in habitat use during severe winter weather episodes could explain this pattern; increasing use of sheltered or subtidal habitats during severe winter weather episodes may buffer some black ducks against direct impacts of severe winter weather that kill other black ducks in early winter, such as starvation or hypothermia (Morton et al. 1989b; Ringelman et al. 2015). If there is no period of high early winter mortality for which our model fails to account, it may be that severe winter weather simply does not affect survival during the non-hunting-season as described in our model (late winter, spring migration, and breeding season). Previous studies have found that body mass is related to survival, but winter weather is not (Conroy, Costanzo, and Stotts 1989; Newcomb et al. 2016). Body mass itself

could be related to severe winter weather, and may be a better direct measure of black duck vulnerability during the non-hunting season than severe winter weather is.

We found that both physically ($MINTEMP$) and biologically ($FROZEN4, FREEZEDAYS$) significant measures of extreme conditions were related to recovery probability of black ducks. The single measure of average conditions that we used, $MEANTEMP$, did not have any relationship with either non-hunting-season survival or recovery probability. This suggests that either acute or chronic exposure to extreme winter weather conditions, or some combination, may result in a carryover effect on vulnerability to hunting. One possible mechanism for such a carryover effect might be reduction in body condition via starvation. In response to prolonged (greater than 4 days) freezing temperatures, black ducks increase their home range sizes, possibly to seek out new, local, non-frozen habitat when they are close to depleting their energy reserves (about 3-4 days for adult females; Ringelman et al. 2015; Cramer 2009). Increased energy expenditure due to foraging movements and low temperatures (Cramer 2009; Cramer et al. 2012; Jones 2012; Jones, Williams, and Castelli 2014), combined with decreased foraging success if non-frozen habitat is not found or is inadequate, could lead to decreased body condition, which has been observed in black ducks during late winter (Morton, Kirkpatrick, and Vaughan 1990). If lower body condition persists through spring migration and breeding season into fall migration, it could be a factor in increased hunting vulnerability (Haramis et al. 1986; Hepp et al. 1986; Robb 2002). Future research on survival involving bands or satellite tags might further investigate this possibility by comparing body condition of captured and recovered black ducks banded both before and after hunting season.

We found that site selection probability and proportion of band recovered ($b_{i,t,c}$) were correlated. Specifically, sites with higher site selection probability also had higher proportions of bands

recovered for both pre-season-banded and post-season-banded black ducks. This suggests that winter locations where banders banded black ducks (for post-season-banded black ducks) and where black ducks were recovered in subsequent hunting seasons (for pre-season-banded black ducks) were also locations for which the ducks which had wintered there had higher recovery probabilities in the subsequent hunting season. This pattern may occur if waterfowl hunters and banders select the same sites (or types of sites) due to the abundance of black ducks in the area. It is possible that these results were somewhat influenced by our decision to impute winter locations of pre-season-banded birds using recovery locations. Imputed winter locations would be more informed by the pre-season banding areas from which more black ducks were recovered in the subsequent hunting season; thus, these winter locations might be associated with higher recovery probabilities. This pattern may be somewhat mitigated by the inclusion of post-season banding locations, although the joint modeling of pre-season and post-season banding locations using the same spatial random effect means that it is impossible to separately estimate spatial patterns in non-hunting-season survival and recovery probability. Even if the spatial random effect and the correlation between site selection and recovery probabilities is partially an artefact of how we imputed winter locations, it is probably still capturing important drivers of spatial variation in recovery probability, such as hunter effort.

Unfortunately, it was impossible to estimate hunting-season survival under our model. Recovery probability is a product of the probability of not surviving the hunting season and being retrieved and reported by hunters (Brownie et al. 1985). Future modeling efforts could model the relationship between hunting-season survival and severe winter weather directly by including band reporting rates as an offset. Alternatively, the model could be expanded to include recoveries from multiple years after banding using multinomial logistic regression. This would

require careful specification of the design matrix, including the addition of recovery year as a covariate, and specific seasonal survival and recovery probabilities could be retrieved using combinations of log-odds ratios. This model might provide increased information on non-hunting-season survival and recovery probabilities, in addition to allowing for the estimation of hunting-season survival probability. However, at least two major limitations would need to be overcome under this model. First, wintering locations of both pre-season- and post-season-banded black ducks are often unknown after banding. These locations would have to be imputed using recovery locations or live recapture data. Second, numbers of recovered black ducks may decline steeply after the first hunting season following banding. Achieving sufficient sample size to estimate spatially explicit hunting-season survival rates may be difficult, though it may be possible to estimate annual survival rates and their relationship with severe winter weather with an appropriately specified model.

A major limitation of this model is the inability to determine pre-season-banded black ducks' winter locations and exposure to severe winter weather in the same way that we determined these quantities for post-season-banded black ducks. Imputing winter locations of pre-season-banded black ducks is difficult, as black duck migration strategies differ throughout their range (Zimpfer and Conroy 2006; Robinson, McGowan, and Devers 2016); thus, not all pre-season-banded black ducks can be expected to winter in the same areas, nor would black ducks likely migrate completely at random to areas of known high winter population density. Imputing locations of pre-season-banded black ducks based on locations of post-season-banded black ducks would probably not adequately account for variation in winter locations across pre-season-banded black ducks. Using $RELABUND$ to randomly assign pre-season-banded black ducks to winter locations would not account for high winter site fidelity in black ducks unless the imputation method were

designed with this site fidelity in mind. Coarse-scale information on transitions to winter locations can be gleaned from previous band recovery studies (Diefenbach, Nichols, and Hines 1988; Zimpfer and Conroy 2006; Robinson, McGowan, and Devers 2016), though this approach for imputing winter locations would lack fine-scale information on exposure to severe winter weather. Satellite tagging studies, recapture data from banding efforts, and stable isotope analysis all hold promise as direct, fine-scale sources of information on winter locations of pre-season-banded black ducks. While recapture data on black ducks are available and we did consider using them to impute wintering locations of pre-season-banded black ducks, the number of cross-season recaptures (e.g., pre-season-banded birds being recaptured during post-season banding) was sufficiently low that it may have been difficult to reliably impute winter locations of large numbers of pre-season-banded black ducks. Until data are available to understand broad-extent, fine-scale patterns in migratory connectivity of black ducks, we believe that recovery locations give a reasonable approximation to winter locations of pre-season-banded black ducks (Diefenbach, Nichols, and Hines 1988).

Given that we used recovery locations of pre-season-banded black ducks to estimate winter locations, it is entirely possible that the coefficients measuring the relationship between severe winter weather and recovery probability (estimated using pre-season data) could be measuring relationships of hunting effort or other covariates to recovery probability, to a certain extent. To check this, we ran a model without pre-season-banded black ducks, estimating the effects of severe winter weather on the joint probability of non-hunting-season survival and recovery for post-season-banded black ducks in each year. We found that this joint probability was not significantly associated with any severe winter weather covariates. However, there was still some evidence of variable importance for these covariates; for instance, the posterior probability that

the effect of $FREEZEDAYS$ on the joint probability was greater than 0 (implying that increasing the number of freezing days increased the joint probability of surviving the non-hunting-season and being recovered) was 0.897. Because severe winter weather was still a possible predictor of recoveries for post-season-banded black ducks for at least some covariates, and because non-hunting-season survival was not significantly associated with winter severity covariates when pre-season-banded black ducks were included in the model, we conclude that the relationship between severe winter weather and recovery that we observed is most likely measuring a response to severe winter weather, rather than hunting effort.

If black duck harvest vulnerability (as measured by recovery probability) is affected by severe winter weather during the winter prior to hunting, there may be management implications. We echo Ringelman et al. (2015) in suggesting that winter habitat planning for black ducks may need to include more thermal refuges and reliable food sources such as forested wetlands and subtidal habitats, in order to mitigate any potential effects of severe winter weather on subsequent harvest vulnerability. If instead recovery probabilities are some function of hunting effort that is correlated with severe winter weather, winter weather should still be taken into account when predicting hunter effort.

More broadly, our study takes a novel approach to analyzing band recovery data. Existing spatially explicit band recovery models, even those that model spatial variation at coarse scales and broad extents, have often required computationally intensive Bayesian methods (Royle and Dubovsky 2001; Zhao, Boomer, and Kendall 2018) and tracking cohorts of banded birds over multiple years. We have demonstrated that spatially explicit band recovery models can be fit quickly and efficiently using INLA, relatively simple log-linear regression equations, and as few as two seasons of banding data and one recovery occasion, and can use covariate information

collected from banding locations to test ecological hypotheses concerning drivers of seasonal survival and recovery probabilities. This approach to band recovery modeling offers the opportunity to estimate fine-scale, spatially explicit survival and recovery probabilities using relatively efficient spatial models for migratory bird species.

5.5 Tables and Figures

Table 5.1. Example band recovery matrix with expected band recoveries as functions of recovery (f) and survival (S) probabilities. In each year (denoted by the subscripts, e.g., M_3 is recorded during year 3), N is the number of birds banded during the pre-season banding period, M is the number of birds banded during the post-season banding period, S^h is the probability of survival through the hunting season, S^{nh} is the probability of survival through the non-hunting season, and f is the probability of a bird being killed, retrieved, and reported during the hunting season.

	Year 1 Recovery	Year 2 Recovery	Year 3 Recovery
N_1	$N_1 f_1$	$N_1 S_1^h S_2^{nh} f_2$	$N_1 S_1^h S_2^{nh} S_2^h S_3^{nh} f_3$
M_2	-	$M_2 S_2^{nh} f_2$	$M_2 S_2^{nh} S_2^h S_3^{nh} f_3$
N_2	-	$N_2 f_2$	$N_2 S_2^h S_3^{nh} f_3$
M_3	-	-	$M_3 S_3^{nh} f_3$
N_3	-	-	$N_3 f_3$

Table 5.2. Watanabe-Akaike information criterion (WAIC) and conditional predictive ordinate (CPO) values for band recovery models including each of four severe winter weather covariates. All four covariates are used to model variation in both non-hunting-season survival and recovery probabilities. MEANTEMP represents the mean daily average air temperature during post-season banding (February 1st-March 10th) at each location and year, MINTEMP represents the minimum daily average air temperature, FREEZEDAYS represents the number of days during post-season banding during which the daily average air temperature was below freezing, and FROZEN4 represents the number of consecutive 4-day periods with daily average air temperatures below freezing during post-season banding.

	WAIC	CPO
MEANTEMP	3038.562	3355.698
MINTEMP	3025.800	3264.929
FROZEN4	3037.032	3286.463
FREEZEDAYS	3025.714	3232.093

Table 5.3. Posterior mean, standard deviation and 95% credible interval estimates for fixed and random effects of the model containing the FREEZEDAYS winter severity covariate with site selection.

Response Parameter	Explanation	Mean	SD	Lower CI Limit	Upper CI Limit
Recovery	Intercept	-8.352	2.222	-13.574	-4.662
Recovery	Effect of FREEZEDAYS	0.255	0.093	0.073	0.437
Recovery	Effect of RELABUND	-0.105	0.074	-0.254	0.036
Non-hunting-season survival	Intercept	-0.523	0.110	-0.741	-0.310
Non-hunting-season survival	Effect of FREEZEDAYS	-0.029	0.095	-0.213	0.160
Non-hunting-season survival	Effect of RELABUND	0.066	0.089	-0.106	0.243
Site selection	Intercept	-8.112	4.640	-18.902	-0.266
Site selection	Effect of RELABUND	1.962	0.257	1.479	2.489
Site selection	Spatial field multiplier	2.163	0.213	1.747	2.586
Recovery	Precision for recovery cohort-year random effect	21.090	13.858	5.318	57.363
Survival	Precision for survival cohort-year random effect	43.556	52.188	5.879	176.895

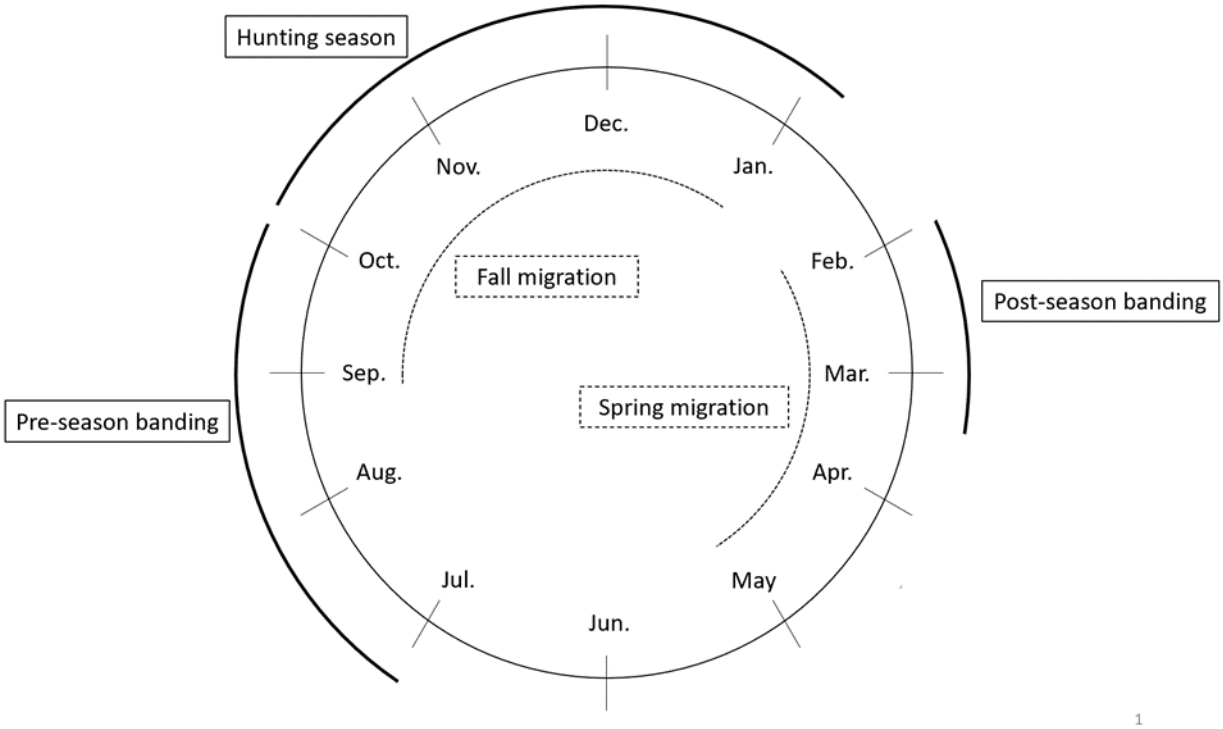
-	Range for spatial field	1858465	683451	932619	3562858
-	Marginal standard deviation for spatial field	3.849	1.231	2.109	6.870

Table 5.4. Posterior mean, standard deviation and 95% credible interval estimates for fixed and random effect parameters of the model containing the FREEZEDAYS winter severity covariate without the site selection component.

Response Parameter	Explanation	Mean	SD	Lower CI Limit	Upper CI Limit
Recovery	Intercept	-3.070	0.900	-4.654	-1.531
Recovery	Effect of FREEZEDAYS	0.095	0.097	-0.096	0.284
Recovery	Effect of RELABUND	-0.087	0.070	-0.228	0.046
Non-hunting-season survival	Intercept	-0.548	0.102	-0.750	-0.348
Non-hunting-season survival	Effect of FREEZEDAYS	-0.006	0.092	-0.184	0.176
Non-hunting-season survival	Effect of RELABUND	0.056	0.083	-0.104	0.220
Recovery	Precision for recovery cohort-year random effect	29.924	19.477	7.054	80.566
Survival	Precision for survival cohort-year random effect	57.737	68.373	7.661	233.536
-	Range for spatial field	4693487	8598620	477594	23884957

-	Marginal standard deviation for spatial field	0.512	0.261	0.222	1.196

Figure 5.1. Black duck life cycle shown inside the calendar and banding and primary hunting seasons shown outside the calendar (hunting seasons for black ducks differ by state, flyway, and year, with some starting in September and others extending through January,).


Hunting season
Dec.
Nov.
Jan.
Oct.
Fall migration
Feb.
Post-season banding
Sep.
Mar.
Spring migration
Pre-season banding
Aug.
Apr.
Jul.
May
Jun.

1

Figure 5.2. Predictions of relative abundance of black ducks across the study area from the generalized additive model fit to the eBird data (see Appendix S3). Predicted relative abundances greater than 20 (n=6) were set to 20 for improved visualization.

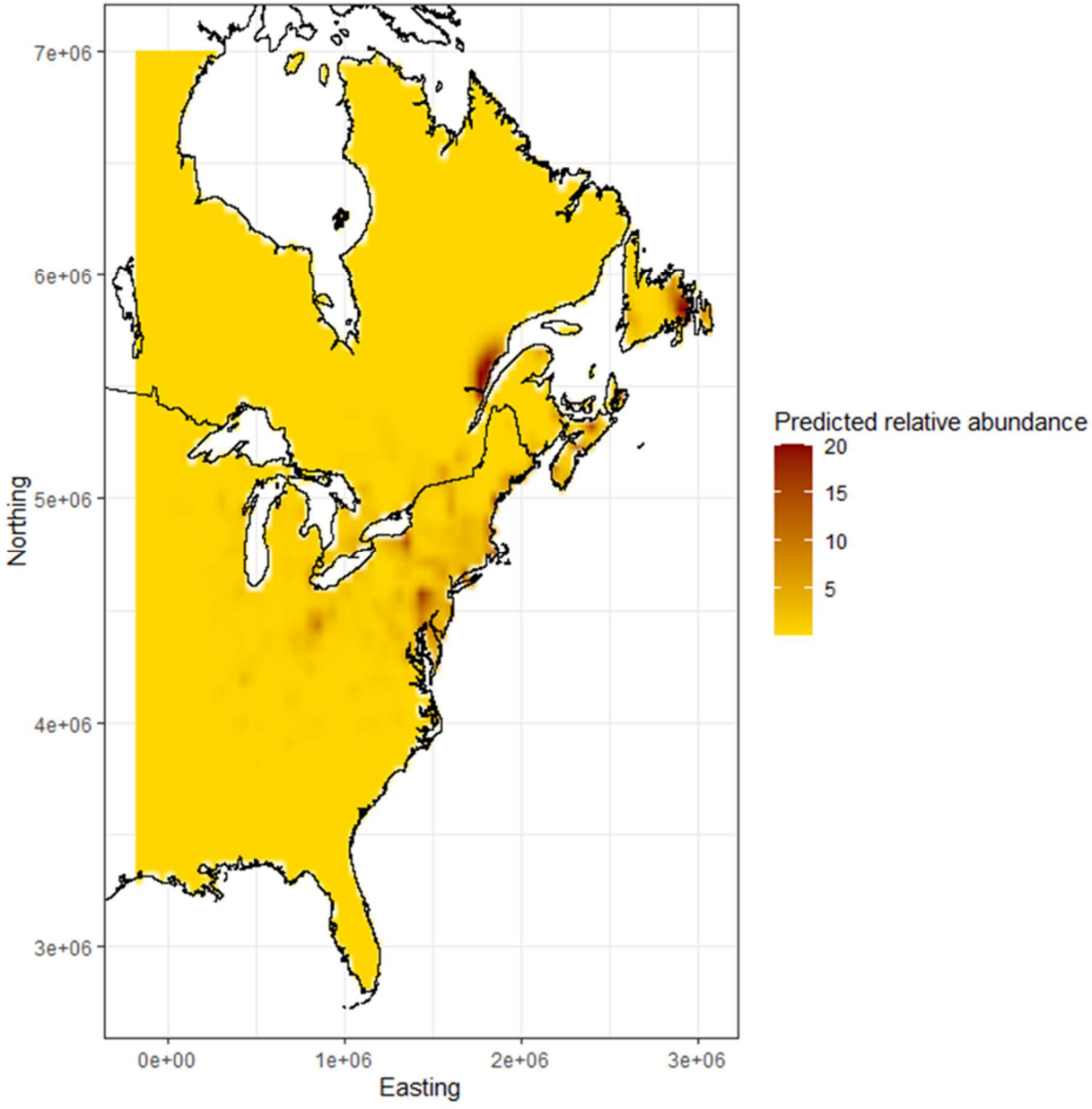
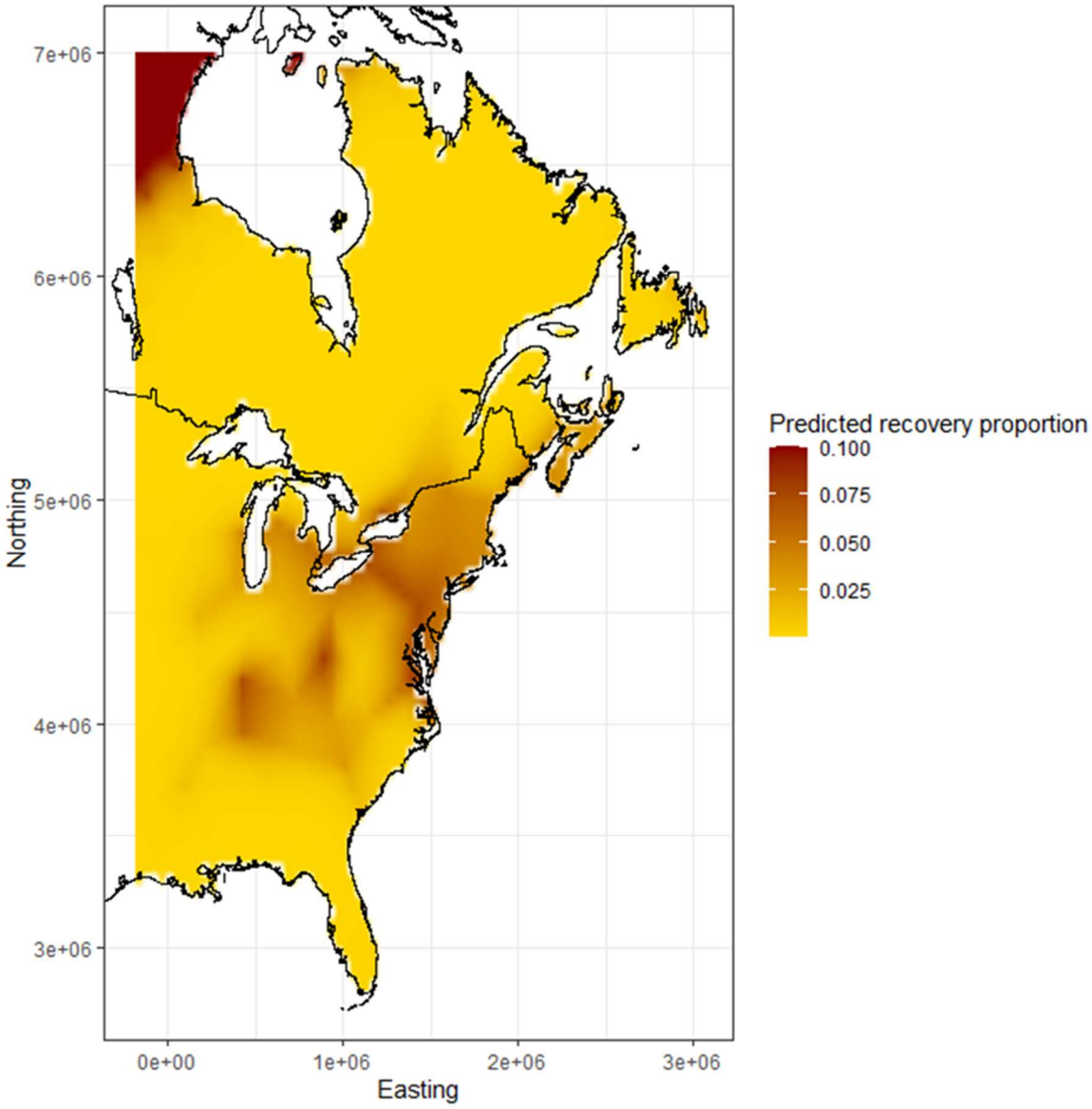


Figure 5.3. Predictions of the proportion of banded black ducks recovered in the first hunting season after banding, of black ducks that experienced average winter severity and relative abundance in the previous winter. This map specifically shows variation in the proportion of banded black ducks recovered $b_{i,t,c}$, and thus represents variation in factors affecting both non-hunting-season survival and recovery probability. Predicted proportions greater than 0.1 (n=76) were set to 0.1; these primarily occurred in the northwest corner of the study area due to INLA boundary effects (Lindgren and Rue 2015).



References

Arnold, T.W., A.D. Afton, M.J. Anteau, D.N. Koons, and C.A. Nicolai. 2016. "Temporal Variation in Survival and Recovery Rates of Lesser Scaup." *Journal of Wildlife Management* 80 (5): 850–61.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.S. Robson. 1985. "Statistical inference from band recovery data: a handbook." Washington, D.C.: U.S. Fish; Wildlife Service.

Conroy, M.J., G.R. Costanzo, and D.B. Stotts. 1989. "Winter Survival of Female American Black Ducks on the Atlantic Coast." *Journal of Wildlife Management* 53 (1): 99–109.

Conroy, M.J., M.W. Miller, and J.E. Hines. 2002. "Identification and synthetic modeling of factors affecting American Black Duck populations." *Wildlife Monographs* 150: 1–64.

Cramer, D.M. 2009. "Estimating habitat carrying capacity for American black ducks wintering in southern New Jersey." M.S. Thesis, University of Delaware.

Cramer, D.M., P.M. Castelli, T.Yerkes, and C.K. Williams. 2012. "Food resource availability for American black ducks wintering in southern New Jersey." *Journal of Wildlife Management* 76 (1): 214–19.

Devers, P.K., R.L. Emmet, G.S. Boomer, G.S. Zimmerman, and J. Andrew Royle. 2021. "Evaluation of a two-season banding program to estimate and model migratory bird survival." *Ecological Applications* in press.

Diefenbach, D.R., J.D. Nichols, and J.E. Hines. 1988. "Distribution patterns of American black duck and mallard winter band recoveries." *Journal of Wildlife Management* 52 (4): 704–10.

Duriez, O., S.A. Sæther, B.J. Ens, R. Choquet, R. Pradel, R.H.D. Lambeck, and M. Klaassen. 2009. “Estimating survival and movements using both live and dead recoveries: A case study of oystercatchers confronted with habitat change.” *Journal of Applied Ecology* 46 (1): 144–53.

Fuglstad, G.A., D. Simpson, F. Lindgren, and H. Rue. 2019. “Constructing priors that penalize the complexity of Gaussian random fields.” *Journal of the American Statistical Association* 114 (525): 445–52.

Haramis, G.M., J.D. Nichols, K.H. Pollock, and J.E. Hines. 1986. “The relationship between body mass and survival of wintering canvasbacks.” *Auk* 103: 506–24.

Hepp, G.R., R.J. Blohm, R.E. Reynolds, J.E. Hines, and J.D. Nichols. 1986. “Physiological condition of autumn-banded mallards and its relationship to hunting vulnerability.” *Journal of Wildlife Management* 50 (2): 177–83.

Jones, O.E. 2012. “Constructing a 24 hour time-energy budget for American black ducks wintering in coastal New Jersey.” M.S. Thesis, University of Delaware.

Jousimo, J., A.J.M. Tack, O. Ovaskainen, T. Mononen, H. Susi, C. Tollenaere, and A.L. Laine. 2014. “Ecological and evolutionary effects of fragmentation on infectious disease dynamics.” *Science* 344 (6189): 1289–93.

Krementz, D.G., M.J. Conroy, J.E. Hines, and H.F. Percival. 1987. “Sources of variation in survival and recovery rates of American black ducks.” *Journal of Wildlife Management* 51 (4): 689–700.

Latimer, C.E., and B. Zuckerberg. 2019. “How extreme is extreme? Demographic approaches inform the occurrence and ecological relevance of extreme events.” *Ecological Monographs* 89 (4): 1–15.

Lindgren, F., and H. Rue. 2015. “Bayesian spatial modelling with R-INLA.” *Journal of Statistical Software* 63 (19): 1–25.

Lindgren, F., H. Rue, and J. Lindström. 2011. “An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach.” *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 73 (4): 423–98.

Longcore, J.R., D.G. McAuley, G.R. Hepp, and J.M. Rhymer. 2000. “American Black Duck (Anas rubripes).” In *The Birds of North America*, https://doi.org/10.2173/bna.481.

Marra, P.P., E.B. Cohen, S.R. Loss, J.E. Rutter, and C.M. Tonra. 2015. “A call for full annual cycle research in animal ecology.” *Biology Letters* 11 (8).

Morton, J.M., R.L. Kirkpatrick, and M.R. Vaughan. 1990. “Changes in body composition of American black ducks wintering at Chincoteague, Virginia.” The Condor 92 (3): 598–605.

Morton, J.M., A.C. Fowler, and R.L. Kirkpatrick. 1989a. “Time and Energy Budgets of American Black Ducks in Winter.” *Journal of Wildlife Management* 53 (2): 401–10.

Morton, J.M., R.L. Kirkpatrick, M.R. Vaughan, and D.F. Stauffer. 1989b. “Habitat use and movements of American black ducks in winter.” *Journal of Wildlife Management* 53 (2): 390–400.

Newcomb, K.C., J.B. Davis, R.M. Kaminski, and M.J. Gray. 2016. “Winter survival of female American black ducks in Tennessee , USA.” *The Condor* 118 (1): 33–45.

Pennino, M.G., I. Paradinas, J.B. Illian, F. Muñoz, J.M. Bellido, A. López-Quílez, and D. Conesa. 2019. “Accounting for preferential sampling in species distribution models.” *Ecology and Evolution* 9 (1): 653–63.

Pollock, K.H., and D.G. Raveling. 1982. “Assumptions of modern band-recovery models, with emphasis on heterogeneous survival rates.” *Journal of Wildlife Management* 46 (1): 88–98.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rice, M.B., D.A. Haukos, J.A. Dubovsky, and M.C. Runge. 2010. “Continental Survival and Recovery Rates of Northern Pintails Using Band-Recovery Data.” *Journal of Wildlife Management* 74 (4): 778–87.

Ringelman, K.M., C.K. Williams, P.K. Devers, J.M. Coluccy, P.M. Castelli, K.A. Anderson, J.L. Bowman, G.R. Costanzo, D.M. Cramer, M.T. Dibona, M.W. Eichholz, M. Huang, B. Lewis Jr., D.M. Plattner, and T. Yerkes. 2015. “A Meta-Analysis of American Black Duck Winter Habitat Use Along the Atlantic Coast.” *Journal of Wildlife Management* 79 (8): 1298–1307.

Robb, J.R. 2002. “Band recovery and recapture rates of American black ducks and mallards.” *Journal of Wildlife Management* 66 (1): 153–61.

Robinson, O.J., C.P. McGowan, and P.K. Devers. 2016. “Updating movement estimates for American black ducks (Anas rubripes).” *PeerJ* 4: e1787.

Robinson, O.J., C.P. McGowan, P.K. Devers, R.W. Brook, M. Huang, M. Jones, D.G. McAuley, and G.S. Zimmerman. 2016. “A full annual cycle modeling framework for American black ducks.” *Natural Resource Modeling* 29 (1): 159–74.

Rockwell, S.M., J.M. Wunderle, T.S. Sillett, C.I. Bocetti, D.N. Ewert, D. Currie, J.D. White, and P.P. Marra. 2017. “Seasonal survival estimation for a long-distance migratory bird and the influence of winter precipitation.” *Oecologia* 183 (3): 715–26.

Roy, C., E.J.B. McIntire, and S.G. Cumming. 2016. “Assessing the spatial variability of density dependence in waterfowl populations.” *Ecography* 39 (10): 942–53.

Royle, J.A., and J.A. Dubovsky. 2001. “Modeling Spatial Variation in Waterfowl Band-Recovery Data.” *Journal of Wildlife Management* 65 (4): 726–37.

Rue, H., A. Riebler, S.H. Sorbye, J.B. Illian, D.P. Simpson, and F.K. Lindgren. 2017. “Bayesian Computing with INLA: a review.” *Annual Review of Statistics and Its Application* 4: 395–421.

Runge, C.A., T.G. Martin, H.P. Possingham, S.G. Willis, and R.A. Fuller. 2014. “Conserving mobile species.” *Frontiers in Ecology and the Environment* 12 (7): 395–402.

Rusch, D.H., C.D. Ankney, H. Boyd, J.R. Longcore, F. Montalbano III, J.K. Ringelman, and V.D. Stotts. 1989. “Population Ecology and Harvest of the American Black Duck: A Review.” *Wildlife Society Bulletin* 17 (4): 379–406.

Rushing, C.S., T.B. Ryder, and P.P. Marra. 2016. “Quantifying drivers of population dynamics for a migratory bird throughout the annual cycle.” *Proceedings of the Royal Society B: Biological Sciences* 283 (1823).

Schummer, M.L., R.M. Kaminski, A.H. Raedeke, and D.A. Graber. 2010. “Weather-Related Indices of Autumn–Winter Dabbling Duck Abundance in Middle North America.” *Journal of Wildlife Management* 74 (1): 94–101.

Sedinger, J.S., J.L. Schamber, D.H. Ward, C.A. Nicolai, and B. Conant. 2011. “Carryover effects associated with winter location affect fitness, social status, and population dynamics in a long-distance migrant.” *American Naturalist* 178 (5): 110–23.

Sicacha-Parada, J., I. Steinsland, B. Cretois, and J. Borgelt. 2020. “Accounting for spatial varying sampling effort due to accessibility in Citizen Science data: a case study of moose in Norway.” *Spatial Statistics*, 100446.

Sullivan, B.L., C.L. Wood, M.J. Iliff, R.E. Bonney, D. Fink, and S.T. Kelling. 2009. “eBird : A citizen-based bird observation network in the biological sciences.” *Biological Conservation* 142 (10): 2282–92.

Watanabe, S. 2013. “A widely applicable bayesian information criterion.” *Journal of Machine Learning Research* 14 (1): 867–97.

Watson, J., J.V. Zidek, and G. Shaddick. 2019. “A general theory for preferential sampling in environmental networks.” *The Annals of Applied Statistics* 13 (4): 2662–2700.

Wood, S. 2017. *Generalized additive models: an introduction with R.* 2nd ed. Chapman; Hall/CRC.

Zhao, Q, G.S. Boomer, and W.L. Kendall. 2018. “The non-linear, interactive effects of population density and climate drive the geographical patterns of waterfowl survival.” *Biological Conservation* 221: 1–9.

Zhao, Q., E. Silverman, K. Fleming, and G.S. Boomer. 2016. “Forecasting waterfowl population dynamics under climate change- does the spatial variation of density dependence and environmental effects matter?” *Biological Conservation* 194: 80–88.

Zimpfer, N.L., and M.J. Conroy. 2006. “Modeling Movement and Fidelity of American Black Ducks.” *The Journal of Wildlife Management* 70 (6): 1770–7.

6 Chapter 4: A spatial capture-recapture model for group-living species

Publication history: This chapter is currently in review at Ecology as: Emmet, R.L., B. Augustine, B. Abrahms, L.N. Rich, and B. Gardner. "A spatial capture-recapture model for group-living species."

6.1 Introduction

Group-living, social species play an important role in ecological systems and include species ranging from small taxa (e.g., meerkats or social insects; Doolan and Macdonald 1996, Elizalde et al. 2020) to large carnivores (e.g., canids; Gittleman 1989) and megaherbivores (e.g., elephants; Coverdale et al. 2016). Within a species, the size and structure of groups can have enormous implications for individual fitness, population dynamics, and even ecosystem processes. Living in large groups can benefit individuals, for example, by increasing foraging success or decreasing the need for anti-predator vigilance (Lima 1995; Wrona and Dixon 1991). Living in large groups can also harm individuals by increasing competition for resources such as food (Holekamp, Sakai, and Lundrigan 2007; Creel and Creel 2015), or increasing the risk of exposure to disease (Nunn et al. 2015). At the population level, density-dependent survival and recruitment can lead to both group-level Allee effects (i.e., positive density dependence of within-group population growth rate) and population-level Allee effects (i.e., positive density dependence of overall population growth rate) (Angulo et al. 2018). Group size and structure also affect species interactions and ecosystem processes, such as predator-prey dynamics (Fryxell et al. 2007) and parasitism (Patterson and Ruckstuhl 2013). Thus, understanding group dynamics, the number of groups, and average size of those groups, is crucial for learning about the ecology of the species, as well as long-term monitoring and management.

While a variety of methods may be used to estimate abundance, number of groups, or group size (Hickey and Sollmann 2018; Schmidt et al. 2012; Belant et al. 2016), spatial capture-recapture (SCR) models offer many advantages. Relative to non-spatial capture-recapture models (Hickey and Sollmann 2018), SCR models account for heterogeneity in detection rates of animals due to individual locations and movement (Royle and Young 2008; Borchers and Efford 2008). SCR models combine a point process model of activity centers with a detection function allowing individuals' detection rates to vary as a function of distance between activity centers and detectors. SCR models have been applied to group-living species, using either individual activity centers and detection histories (individual-level SCR models; Russell et al. 2012; Lopez-Bao et al. 2018) or group activity centers and detection histories (group-level SCR models; Mattioli et al. 2018). The choice of individual versus group activity centers is crucial for estimating abundance of group-living species, as these models can have varying levels of ecological realism depending on the group movement behavior of the study species. From an SCR perspective, group-living species can vary widely in their level of aggregation (the degree of spatial clustering of individuals' activity centers in the population) and in their level of cohesion (the degree to which individuals in the population move and are detected with their groups; Bischof et al. 2020). Individual-level and group-level SCR models represent the extreme ends of aggregation and cohesion, with individual-level SCR models assuming solitary individuals that move completely independently (i.e., no aggregation or cohesion) and group-level SCR models assuming that all individuals in the population share an activity center with one or more other individuals, and that these individuals have a completely shared detection history (i.e., high aggregation and perfect cohesion).

The choice between the current individual-level and group-level SCR models involves a tradeoff in assumptions of independent individual detections and group-level homogeneity in detection, which is governed in part by the level of cohesion within the study population. Species vary widely in their cohesion within groups, which we specifically define as a tendency to move and be detected with members of their groups (Aureli et al. 2008). Aureli et al. (2008) describe intra-group cohesion as one of two measures of behavioral fission-fusion dynamics in a species, the other measure being inter-group stability (the stability of the membership of groups). In terms of cohesion, species can vary from moving in solitude (low cohesion), to moving constantly with groups (high cohesion) (Aureli et al. 2008), to cohesion varying over time as a group splits into subgroups based on food availability (e.g., Asensio, Korstjens, and Aureli (2009)). Low cohesion, where animals are moving relatively independently of their groups, is less likely to violate the assumption that detections of individuals are independent in individual-level SCR models (Bischof et al. 2020). High cohesion, however, likely violates the assumption of independent individual detections, making group-level SCR models more appropriate as they assume individuals in a group share a common activity center and detections are independent only at the group level (Mattioli et al. 2018; Bischof et al. 2020).

One shortcoming of group-level SCR models is that they assume that groups' detection probabilities do not vary with group size, which is considered unobserved and not modeled in group-level SCR models; that is, two groups of different sizes would be expected to have equal encounter rates or probabilities with a given detector if their activity centers were the same distance from the detector. Larger groups, however, may be detected more easily, as they contain more members who could be exposed to detection (Hickey and Sollmann 2018). Group-level SCR models do not include a parameter for group size and consequently, cannot account for this

extra variation in detection, which may yield biased density estimates. Moreover, group-level SCR models do not include a model component for individual-level detection, even though accounting for detection histories of individuals can be informative of group size and can allow for modeling the dependence in detections among individuals within groups.

Existing SCR methods have estimated group size from auxiliary data and multiplied an ad-hoc estimate of group size by group density to estimate individual density (Russell et al. 2012; Mattioli et al. 2018). This approach limits the ability to model predictors of group size and abundance in a unified framework, and does not account for the effects of group size on group-level detection. Bischof et al. (2020) developed an overdispersion correction for individual-level SCR models applied to group-living species, which can improve coverage of individual density interval estimates when cohesion is high. To date, however, no solution has been applied to account for heterogeneity in detection of group-living species while estimating both group abundance and group size.

Thus, what is needed in the SCR detection process is 1) the flexibility to have shared expected group detection rates while allowing for individual-level detection, thus accounting for dependence in detections of individuals within groups (i.e., individuals moving together) and 2) an additional model component that accounts for different group sizes. A potential framework for jointly estimating group abundance and size is cluster capture-recapture in which clustered point processes allow the estimation of cluster density and size (Stevenson et al. 2020). In the model presented by Stevenson et al. (2020), cluster members are calls associated with a particular individual, and clusters are individuals. Each individual's calls are members of a group of calls, and that group is associated with the individual via a clustered point process. We can

apply a similar framework to groups, modeling groups as clusters and group members as cluster members to simultaneously estimate group abundance and size.

Failure to account for dependence between individuals within groups could lead to biased assessments of population status, and ultimately lead to adverse conservation outcomes. In addition, average group size can be an important quantity for conservation decision-making. An ideal case study are African wild dogs (*Lycaon pictus*, hereafter wild dogs or African wild dogs), a group-living species of international conservation concern (Woodroffe and Sillero-Zubiri 2012). African wild dogs are categorized as Endangered on the IUCN red list, with a declining population trend and an estimated 1,500 mature adults remaining in the wild (Woodroffe and Sillero-Zubiri 2012). Group size in African wild dogs varies between 2 and 40 individuals and individuals within a group often move together (Courchamp and Macdonald 2001; Woodroffe et al. 2020). Group size can affect measures of individual fitness, as larger groups can have increased litter size (Courchamp, Clutton-Brock, and Grenfell 2000), may experience higher hunting success, and may have greater ability to drive off kleptoparasites (Courchamp, Clutton-Brock, and Grenfell 2000). Group size may also play a role in regulating population growth in wild dogs (Courchamp, Clutton-Brock, and Grenfell 2000; Lerch, Nolting, and Abbott 2018; Woodroffe, O'Neill, and Rabaiotti 2019).

Our research aims to enhance our understanding of the complex group- and population-level effects that may limit the recovery of wild dogs and other group-living species by advancing our ability to accurately and simultaneously model individual abundance, group abundance, and group size. We developed a cluster SCR model to estimate group size and group and individual abundance while accounting for dependence in detection between individuals within groups, using a model based on group activity centers. We applied the cluster SCR model for group-

living species to detections of African wild dogs on camera traps in 2014 and 2015 in the Okavango Delta region of Botswana. This well-studied population provided a unique test case to evaluate the model, as the true group size was known for the packs. To determine how the model performed, we developed simulation studies using different levels of aggregation (i.e., group size) and density of groups to examine the effects of these quantities on bias and precision of parameter estimates. We compared the model to a group-level SCR model when using data generated under the cluster SCR model assumptions, and we fit the model to simulations of dynamic group movement to investigate its robustness to realistic variations in group movement.

6.2 Methods

6.2.1 Model

SCR models include an ecological process, a spatial point process describing density and distribution of activity centers (typically denoted s), and an observation process, which is a detection process describing encounter rates with detectors as a function of distance between detectors and activity centers (Borchers and Efford 2008; Royle and Young 2008). We extend SCR models to account for group-structured detection and abundance processes of group-living species. The cluster SCR model's ecological process model includes the density and distribution of group activity centers, group size, and individual density. The observation process model includes two stages. In the first stage, groups encounter detectors in a manner similar to the detection process in standard SCR models (i.e., detectors farther away from the group activity center are expected to be encountered less often). This process is not perfectly observed; instead, in the second stage, individuals within the group may be detected conditional on a group encountering a detector and thus being detectable. In the terminology of Bischof et al. (2020), our model assumes some degree of aggregation for all individuals (i.e., individuals within groups

have the same activity center, but there are multiple groups in the study area) and high cohesion (i.e., all individuals follow the movement patterns of the group around their activity center). Group membership of individuals is assumed to be known, though this assumption can be relaxed with future model extensions. While the model parameters and structure are described here, the MCMC algorithm we use to implement the model is described in Appendix S4.

We start by developing the ecological process portion of the model. To estimate abundance, we used a data augmentation approach for capture-recapture models (Royle et al 2013). The cluster model described below includes components to estimate both the number of groups and the number of individuals; thus, we augmented a large number of all-zero encounter histories to both the observed group- and individual-level encounter histories. Each group and individual in the augmented dataset is assigned a latent inclusion variable which indicates whether it is part of the actual population of interest. Observed groups and individuals are automatically included in the population, as detection is conditional on being in the population of interest. The sums of the latent inclusion variables yield estimates separately of the realized group and individual abundances.

Groups $g = 1,2, \dots, Z$ have activity centers, s_g, uniformly distributed over a state space M:

$$\boldsymbol{s_g} \sim Uniform(M)$$

where Z is number of observed and augmented groups in the dataset. . Each group has a latent population inclusion variable, z_g, which is equal to 1 when the group is included in the population of groups and 0 otherwise, with inclusion probability ψ:

$$z_g \sim Bernoulli(\psi)$$

Each group has a latent group size n_g, with parameter ζ such that

$$n_g \sim ZeroTruncatedPoisson(\zeta)$$

This means that groups cannot have size zero, though groups can have size one, which implies that the population can include "floaters" who do not have any other group members.

Each individual $i = 1,2, \dots, W$ (where W is the size of the augmented data set of individuals) must be a member of a group g. The group membership of individuals, $\gamma_i \in 1,2, \dots, Z$, is assumed to be known for detected individuals. Individual abundance $N = \sum_{i=1}^{W} w_i = \sum_{g=1}^{Z} z_g n_g$, like group abundance $G = \sum_{g=1}^{Z} z_g$, is estimated using data augmentation; individuals can be included in the population using inclusion variable w_i, which can only equal 1 if the group to which individual belongs is also included in the population ($z_{\gamma_i} = 1$). When $z_{\gamma_i} = 1$, the model allows exactly n_g individuals to be included in the size of group g, such that $n_g = \sum_{i:\gamma_i=g} w_i$.

The distance-based encounter (detection) process used in standard SCR occurs at the group level in this model. That is, it is assumed that groups are less likely to encounter detectors that are farther away from their activity centers. For example, using a half-normal detection function, detection rate $\lambda_{g,j,t}$ for group g at detector j on occasion t is calculated as:

$$\lambda_{g,j,t} = \lambda_0 \exp(-\frac{d_{g,j}^2}{2\sigma^2})$$

where $d_{g,j} = ||\boldsymbol{s_g} - \boldsymbol{x_j}||$ is the distance between the location of activity center of group g ($\boldsymbol{s_g}$) and the location of detector j $(\boldsymbol{x_j})$. λ_0 is the baseline group encounter rate and σ is a detection scale parameter.

This rate, $\lambda_{g,j,t}$, is the expected number of group encounters at a specific detector for a group g. The number of group encounters is the number of times, in a sampling occasion *t*, that individuals in a group are able to be detected at a detector (i.e., the group is close enough to the detector that all of its individual members could potentially be detected). Group encounters are not the same as groups being detected, or all members of a group being detected; instead, group encounters are times when all individuals within a group are able to be detected, when the group encounters a detector. The number of latent group encounters with detectors (hereafter group encounters), $v_{g,j,t}$, is generated according to encounter rate $\lambda_{g,j,t}$:

$$v_{g,j,t} \sim Poisson(\lambda_{g,j,t})$$

Under our model, groups are assumed to encounter detectors in a process that is not directly observed; in particular, if an individual is detected at a given detector during a given occasion, it is assumed that the whole group must be able to be detected at that detector and occasion. This does not necessarily imply that individuals in a group always move together, or that all individuals in a group must be detected or even immediately present at a detector during a group encounter. Group encounters only require that if any individual of a group is detected at a given detector and occasion, others could be as well, each at an equal rate, though this assumption is most likely to be met when individuals always travel with their group.

Finally, the number of individual detections is conditional on the latent group encounters:

$$y_{i,j,t}|v_{\gamma_i,j,t} \sim Poisson(w_i v_{\gamma_i,j,t}\theta)$$

where θ is the individual detection rate. This rate is the expected number of detections per individual per latent group encounter. Note that individuals cannot be observed at a detector

unless their assigned group has one or more latent encounters at that detector, which is governed in part by the group activity center.

The individual detection process accounts for group-level heterogeneity in detection due to varying group size. The individual is the unit of detection in this model, and groups are only indirectly observed through the group membership vector $\boldsymbol{\gamma}$. More individuals in a group represent more opportunities for a group to be detected; thus, including more individuals in a group indirectly raises the overall probability that the group is detected.

6.2.2 Relationship to individual-level and group-level SCR models

The cluster SCR model is a hybrid of individual-level and group-level SCR models. Individual- and group-level SCR models include data augmentation with a single inclusion variable, as the models only estimate individual abundance and group abundance, respectively, but not both. The cluster SCR model, on the other hand, includes data augmentation via inclusion variables for both individuals and groups. Individual-level and group-level SCR models contain detection data and activity centers at a single level (e.g., group-level SCR model consider activity centers and detection data on groups). The cluster SCR model has group activity centers and a detection process based on distance from group activity center to detectors, but the detection data are modeled at an individual level.

6.2.2.1 Individual-level SCR model

Under an individual-level SCR model, individuals $i = 1,2, \dots, W$ would have activity centers s_i and group membership γ_i would be ignored:

$$s_i \sim Uniform(M)$$

The detection model would not include group encounters; these would be replaced by individual encounters, and the individual detection rate parameter θ is removed:

$$\lambda_{i,j,t} = \lambda_0 \exp(-\frac{d_{i,j}^2}{2\sigma^2})$$

$$y_{i,j,t} \sim Poisson(w_i \lambda_{i,j,t})$$

6.2.2.2 Group-level SCR model

Under a group-level model, groups $g = 1,2, \dots, Z$ would have activity centers s_i and group membership γ_i would be known:

$$s_g \sim Uniform(M)$$

Individual identity may not be known under this model, as individuals might be identifiable as members of a group but not as individuals, depending on sampling and identification methods.

The detection model would include group encounters, but these would not be latent. Individual detection rate would not be estimated under this model; instead, the group detection rate would be some function of unknown individual detection rate and group size. Relative to the cluster SCR model, the detection histories for the group-level SCR model would be summed over an unknown number of individuals, since the group-level SCR model does not include a model for group size:

$$y_{g,j,t} = \sum_{i:\gamma_i = g} y_{i,j,t}$$

$$\lambda_{g,j,t} = \lambda_0 \exp(-\frac{d_{g,j}^2}{2\sigma^2})$$

$$y_{g,j,t} \sim Poisson(z_g \lambda_{g,j,t})$$

Though the group-level SCR model includes no model for group size, the average group size may be estimated from auxiliary data and used to calculate individual density or abundance (Russell et al. 2012; Mattioli et al. 2018).

6.2.3 Case study

6.2.3.1 Study area and data collection

We used detections of wild dogs on camera traps collected between August and November 2014 and February and August 2015 in and around the Moremi Game Reserve in northern Botswana's Okavango Delta. Cameras were deployed for approximately 30 days each, and a total of 134 and 221 locations were monitored in 2014 and 2015, respectively (Figure 6.1). Only two locations monitored in 2014 were not monitored in 2015. Cameras were placed along roads to maximize detections of carnivores. Rich et al. (2017) contains a full description of the study area and data collection process. This is a well-studied population where individuals can be identified using pelage patterns from photographs (Rich et al. 2017) and the group membership of each individual is known from long-term monitoring (Behr et al. 2020); thus, we have more information available than in a typical camera trapping study where only individuals may be identified.

6.2.3.2 Data analysis

For the analysis, we used a single survey occasion ($T = 1$) and adjusted for varying effort at each camera by multiplying $\lambda_{g,j,t}$ at each camera j by the number of days each camera was operational. Thus, $\lambda_{g,j,t}$ can be interpreted in this case as a daily group detection rate. To create a state space for the model, we created a buffer of 15 km around the camera arrays, resulting in a total area of 5387 km^2 for the 2014 state space and 6924 km^2 for the 2015 state space. We used

only detections of wild dogs whose individual and group identities were known from previous studies.

We fit separate cluster SCR models to the wild dog 2014 and 2015 data using a custom MCMC algorithm in R; the algorithm is described in detail in Appendix S4. We also fit individual- and group-level SCR models to the 2014 and 2015 data using the jagsUI package (Kellner 2016). We used vague priors for all parameters (Appendix S4). For all models fit to the wild dog data, we ran 10 chains of 100,000 iterations each, with a thinning rate of 10. After thinning, we discarded burn-in iterations based on visual inspection of the thinned chains. Specifically, we discarded 500 of the 10,000 thinned iterations available for the models for both years. We calculated Gelman-Rubin diagnostics and visually inspected chains for all models to ensure adequate convergence and mixing.

We recorded posterior distribution summaries for all parameters, including posterior modes and 95% credible intervals (calculated from the 2.5% and 97.5% percentiles of the posterior samples). To calculate individual density for group-level SCR models, we multiplied the average observed group size (i.e., average number of individuals detected in each *detected* group) by estimated group density (Russell et al. 2012). This estimation method does not account for variation in group sizes in the population, and thus does not account for variation in detection due to differences in group size.

6.2.4 Simulation

We conducted a simulation study to assess the ability of the cluster SCR model to estimate group and individual density and group membership using data generated from the model. We simulated data directly from the cluster SCR model. We simulated populations of either 10 or 20 groups in a 57- by 60-unit state space with a regularly spaced camera grid of 132 cameras

(spaced at intervals of 2/3 times σ, fixed at about 3 units) with a buffer of 13.5 units. The area of this state space is about half the size of the state space used in the 2015 wild dog data set, where each unit is a kilometer. We varied the number of groups, with either 10 or 20 groups, and group size, with average group size parameters of either 5 or 12.5 (leading to average group sizes of close to 5 and 12.5, respectively). All other parameters were fixed at constant values.

To assess how well the cluster SCR model performs under different group movement dynamics, we conducted a simulation study using a biased correlated random walk model formulated as a hidden Markov model (Langrock et al. 2014). We used a version of this model with a single behavioral state, in which dependence in individuals' movements was induced via attraction to a shared group centroid that moves over time. This results in individuals within groups having loosely similar movement trajectories, where individuals within groups do not share exactly the same movements but their movements are correlated (relatively high but not perfect cohesion).

Elaborating on the hidden Markov model, for each group, we simulated movement of a group centroid using a correlated random walk; this centroid was not observed or modeled in the cluster SCR model but was used as a center of attraction for individuals within the group. Then, individual movements were simulated using a biased correlated random walk, where at each time step, individuals' movements were influenced both by their movement in the previous time step and by attraction to their group centroid. This movement model is described more fully in Langrock et al. (2014). We simulated the data with the R package momentuHMM version 1.5.2 (McClintock and Michelot 2018). To explore the effect of increasing independence of movement of individuals, we varied the parameter governing attraction of individuals to their time-referenced group centroid. This resulted in two levels of attraction to group centroids, which we term relatively "high" and "low", though neither scenario results in very high attraction to group

centroids. These do not correspond directly to scenarios of high and low cohesion, because cohesion requires that individuals move and are detected together often, and the latter cannot be guaranteed in our simulations. However, under the lower attraction scenario, individuals are allowed to range more widely and thus would more likely be detected without their group members. We simulated detections of individuals using a detection function that varies by distance between individual locations and detectors in each time step (Royle et al. 2016). We varied the number of groups, with either 10 or 20 groups, and group size, with average group size parameters of either 5 or 12.5.

For each unique combination of parameters in the simulation studies (hereafter scenario), we simulated 100 data sets from the model. These data sets included detection histories and group identities of detected individuals. We fit the cluster SCR models and a group-level SCR model to each data set simulated from the cluster SCR model, and fit only the cluster SCR model to each data set simulated from biased correlated random walks, using the same custom MCMC algorithm and JAGS code used in the data analysis. We did not fit individual-level SCR models in either simulation study, because these data were simulated assuming high cohesion (or at least highly correlated movement) between individuals within groups and the effects of cohesion on individual-level SCR models have been previously described by Bischof et al. (2020). We also did not fit the group-level SCR model in the biased correlated random walk study, because our aim was to understand the robustness of the cluster SCR model to violations of the assumptions of stationary activity centers and perfect cohesion of individual movement within groups.

We used a single occasion (T=1) for all models, and ran 3 chains for 30,000 iterations with 5,000 iterations of burn-in and thinning the results at a rate of 150, resulting in 500 samples. We used vague priors for all parameters. We recorded the posterior modes of group abundance and group

size for each simulated data set. We also calculated, using only models that had converged (as measured by Gelman-Rubin diagnostics ($\hat{R}$) that were non-missing and less than 1.1, which resulted in some scenarios with <100 data sets; see Appendix S5: Tables S4 and S6), mean relative bias and coverage of group abundance and the average group size parameter for each scenario. In the simulations from the cluster SCR model, we also calculated mean relative bias and coverage for all other cluster SCR model parameters, and the detection parameters λ_0 and σ, and the group data augmentation parameter ψ from the group-level SCR model, and included these in Appendix S5.

6.3 Results

6.3.1 Case study

There was a total of 356 detections in the 2014 data set, from 70 individually identifiable wild dogs in 6 groups. In the 6 groups detected, the average number of individuals detected was 11.67, and ranged from 3 to 21 individuals. Group overlap at cameras was generally low: only 5 of the 53 cameras that detected at least one wild dog detected individuals from different groups. Individuals were detected at an average of 4.26 cameras, while groups (i.e., groups for which at least one member of the group was detected) were detected at an average of 9.67 cameras.

There was a total of 267 detections in the 2015 data set, from 46 individually identifiable wild dogs in 6 groups. In the 6 groups detected, the average number of individuals detected was 7.67 individuals, and ranged from 2 to 11 individuals, a much smaller range than 2014. Group overlap at cameras was low in 2015 as well as 2014: only 6 of the 63 cameras that detected at least one wild dog detected individuals from different groups. Similar to 2014, individuals were detected at an average of 4.94 cameras, while groups were detected at an average of 11.5 cameras.

Gelman-Rubin diagnostics and visual inspection of chains indicated that the cluster SCR models converged (all $\hat{R} < 1.1$). The group-level SCR model for the 2014 data indicated convergence for all parameters except for deviance. The 95% credible intervals for all parameters of the cluster SCR model overlapped previous estimates of individual density and group size for 2014, but not group density for 2014 or any previous estimates for 2015 (Table 6.1). Group density, individual density, and group size were all estimated to be smaller in 2015 than in 2014 (Table 6.1). The cluster SCR model estimated higher individual and group density than the group-level SCR model. The cluster SCR model generally yielded density estimates that were closer to previous estimates than those from the individual-level and group-level SCR models, except for 2014 individual density, where the individual density estimated from the group-level SCR model was closer to the previous estimate (Table 6.1). Results for all parameters in the case study and simulations, except for individual density/abundance in the simulations, are available in Appendix S5.

6.3.2 Simulation

The cluster SCR model returned estimates with relatively little bias when fit to data generated from the model (Table 6.2); bias was only greater than 3-4% for the estimated number of groups when the true number of groups was 10. The cluster SCR model also gave close to nominal coverage of 95% credible intervals when fit to simulations from the model (Table 6.2). The group-level SCR model fit to data simulated from our cluster SCR model returned extremely negatively biased estimates of the number of groups (Table 6.2). The average number of recaptures of groups increased with both the average group size parameter and the number of groups (Table 6.3). The average number of individual recaptures stayed roughly the same regardless of scenario (Table 6.3).

When fit to detection data generated from hidden Markov movement models, the cluster SCR model returned average group size parameter estimates that were consistently biased low and group abundance estimates that were consistently biased high, though bias never reached >25% (Table 3). When attraction to group centroids was high, bias was relatively low and coverage was close to nominal; when attraction to group centroids was low, bias increased and coverage decreased substantially (Table 6.3). The average number of individual and group recaptures was relatively similar across all scenarios, but the average number of individual recaptures generally decreased and the number of group recaptures generally increased when attraction to group centroids was low (Appendix S5: Table S5).

6.4 Discussion

We developed a cluster SCR model that included several extensions to standard SCR models: joint estimation of group and individual abundance using data augmentation; an individual detection process that is conditional on group encounters to account for dependence in individual-level detections; and the ability to estimate average group size without auxiliary data. By simultaneously estimating individual abundance, group abundance, and group size, and providing relatively accurate estimates, the cluster SCR model offers opportunities to investigate aspects of the ecology of group-living species, including drivers of group size and group density.

The results of the simulations and analysis of the African wild dog data match the behavior expected of individual- and group-level SCR models when they are applied to group-living animals with high cohesion and group-level heterogeneity in detection. Though we did not apply individual-level models to simulated data, Bischof et al. (2020) showed that individual-level SCR models applied to data with high levels of cohesion had positive bias in individual density estimates. High cohesion was present in the data simulated from the cluster SCR model so we

would expect to find that individual density estimates from individual-level SCR models would be higher than those from the cluster SCR model. Individual density estimates for the wild dog population met this expectation (Table 6.1) In group-level SCR models, our simulations would lead us to expect that group density estimates from group-level SCR models would be lower than estimates from cluster SCR models, due to negative bias in group density estimates from group-level SCR models in simulations (Table 6.2). The results of our data analysis met this expectation (Table 6.1).

There is a possibility that wild dog group density estimates from the cluster SCR model were biased, as the density of groups in the population was lower than previous estimates in both years and the simulations showed bias in group density when the number of groups in the population was small (Table 6.2). Estimates of group density from the cluster SCR and group SCR models were closer to estimates from the study area from Creel, Mills, and McNutt (2004), which range from 0.23 to 0.5 groups per 100 km^2. Group density estimates based on recent observations of the population are substantially higher (0.70-0.81 groups per 100 km^2 ;Table 6.1). The discrepancy in group density estimates could be explained by the relatively small number of groups detected; of the 22 groups present in 2014 and the 19 groups present in 2015, only 6 were detected in the camera trapping data set in each year (the same 6 were not detected in both years). In spite of this, generally, the cluster SCR model returned estimates of wild dog individual density, group density, and group size that were not far from previous estimates from the study area, although only 2014 individual density and group size parameter credible intervals captured previous estimates (Table 6.1). The cluster SCR model did have relatively unbiased estimates and high coverage in simulations when fit to data simulated from the model, and retained high coverage but moderate bias (<20%) particularly when fit to hidden Markov

movement model simulations generated with high levels of dependence in individuals' movement (i.e., attraction to group centroids). Thus, the cluster SCR model seems to be capable of overcoming multiple sources of bias in individual-level and group-level SCR models, including group-level heterogeneity in detection due to varying group size and non-independence of individuals' detections due to shared group activity centers and high cohesion.

Unmodeled heterogeneity in detection due to variation in group size may lead to negative bias in abundance in capture-recapture and SCR models (Hickey and Sollmann 2018). Both the simulations from the cluster SCR model and data analysis display this pattern, as the group-level SCR model yielded lower group abundance and density estimates than the cluster SCR model in both analyses (Table 6.1; Table 6.2). There is also reason to believe that ad-hoc estimates of group size based on the number of detected individuals per group, which have been used with group-level SCR models (Russell et al. 2012), may lead to bias. While estimated group sizes were similar between cluster SCR and group-level SCR models fit to the wild dog data (Table 6.1), this method of calculating group size may yield biased individual density estimates in other cases, because group size may not be perfectly observed (i.e., the estimate is based only on the number of detected individuals), and because larger groups may be easier to detect and thus average group size may be overestimated under this method. While both the cluster SCR and group-level SCR model require that group membership of detected individuals is known, our results suggest that the cluster SCR model can yield less biased group density and group size estimates than the group-level SCR model, particular due to its ability to account for group-level heterogeneity in detection.

The cluster SCR model makes several assumptions that could be relaxed for different studies. One of the most restrictive assumptions is that group membership of detected individuals is

known. This assumption can be relaxed by treating group membership as a latent variable, similar to how individual identity is treated as a latent variable in unmarked SCR (Chandler and Royle 2013) and similar models. In order to fit this model, we must either marginalize over all possible group membership assignments of individuals to groups or sample them directly in the MCMC algorithm. The zero-truncation of the Poisson model for group size prevents marginalization; however, this may be a fruitful strategy for species living in larger groups where a Poisson group size assumption would never or rarely lead to groups of size zero being proposed.

If an assumption of known group membership is required, auxiliary data collection or group assignment based on proximity or other factors may become necessary. Strategies for identifying group membership of detected individuals will vary by species. For species with high inter-group stability (i.e., species in which individuals rarely or never change group membership) and strongly defended territories, group membership may be inferred from co-detection of individuals. However, this approach will likely be prone to error if territories overlap or co-detections of individuals are rare. If group membership cannot be inferred from detection histories themselves, auxiliary data may have to be collected. This can often take the form of genetic data to infer both individual identity and group membership. Sites that groups use collectively, such as rendezvous sites (Bassing et al. 2020) and nests (Hickey and Sollmann 2018), can be used to collect data on both individual identity and group membership. Another approach that has been used is the aggregation of individuals into groups based on spatial proximity of their detections (Russell et al. 2012), though again, this approach may be prone to error if territories overlap or the spatial scale used to assign groups often results in multiple groups being combined or groups being split apart incorrectly. Given that assigning groups by

proximity is likely to lead to biased parameter estimates, we recommend that this model be used only in well-studied populations where group membership of detected individuals can be definitively assigned. While this does limit the types of data sets on which this model can be used, the cluster SCR model can still provide more information (i.e., group size) and account for dependence in detections to a greater extent than other SCR models, and has the potential to be extended to allow group membership to be a latent variable.

The cluster SCR model assumes a closed population, such that individual abundance, group abundance, and group membership do not change over the course of the sampling period. As in all closed-population capture-recapture models, careful consideration must be given to the length and timing of sampling. As an example, when movement dynamics such as dispersal can impact group membership, sampling should not occur during dispersal periods, if possible, to avoid changes in group membership. In the case of African wild dogs, individuals or same-sex groups disperse at any time of year (Woodroffe et al. 2020), so it is possible that dispersers or "floaters" with movement behavior different from that of established groups may be present in the population at any time. In situations like this, where sampling cannot be timed to avoid dispersal periods, the cluster SCR model would require an extension to model some individuals independently of their groups. While under the current cluster SCR model, it might be possible to treat dispersers as groups of size one, dispersal could cause substantial heterogeneity in detection parameters (particularly the detection scale parameter σ), and a modeling solution would need to be developed to account for that heterogeneity. This might be accomplished using an open-population SCR framework to let individual and group abundance and group membership change over time, although letting group membership change over time would add considerable complexity to the model.

As with all SCR models, animal movement plays a key role in the underlying assumptions of the detection process. The cluster SCR model makes assumptions about not only individual movements, but also group movements. The model assumes groups move around an activity center and that the activity center is stationary during the study, as would be expected with territorial or stable home range behavior. Individuals are assumed to move with a high level of group cohesion, such that if one individual is detected, all individuals within its group are assumed to be detectable. Our simulations of group movement with imperfect cohesion and dynamic (i.e., nonstationary) group activity centers showed that the cluster SCR model is not very robust to the combined violations of the assumptions of individual detections being conditional on group encounters, stationary activity centers, and perfect cohesion of individual movement within groups. When attraction to the group activity center is relatively high, the cluster SCR model returns group abundance and average group size parameter estimates with moderate bias but high coverage, in spite of the presence of individual trajectories and dynamic group activity centers. Lower attraction to group centroids may result in unmodeled heterogeneity in group encounters (because fewer individuals may be available for detection at each group encounter than if attraction were higher) and biased estimates of detection rates, group abundance, and group size estimates. This seems to become more likely under our low-attraction scenario: individual recapture rates are fairly similar across all scenarios, but group recaptures increase under low-attraction scenarios, suggesting that individuals are ranging more widely and being detected without other group members more often (Appendix S5: Table S5). Indeed, in our movement simulations, bias increases and coverage decreases under low-attraction scenarios (Table 6.3). Thus, when individuals move very independently of their groups (i.e.,

cohesion is low), individual-level SCR models may be more appropriate than the cluster SCR model (Bischof et al. 2020).

In reality, animal movement, including group movement patterns, are complex, and we have tried to capture some of this complexity in simulating different aspects of cohesion and attraction to dynamic group activity centers. However, the realism of our simulations varies depending on the study area size, how quickly animals move relative to home range size, and the length of the study. The degree of cohesion in group movement can vary over time and space (Aureli et al. 2008); for example, some species live relatively consistently in groups, such as caribou (Langrock et al. 2014), while others like the Mongolian gazelle may cluster during calving but otherwise range independently (Mueller et al. 2011). African wild dogs, being cooperative breeders and hunters, exhibit dynamic activity centers and other behaviors that lead to positive bias, but they likely exhibit relatively high cohesion, especially during periods when pups are being cared for and home ranges shrink (Pomilia et al. 2015). Thus, our movement simulation study is not fully capturing the group dynamics of the African wild dogs; however, the simulation study does highlight the importance of applying the cluster SCR model to species that exhibit high cohesion in group movement.

Many patterns of group movement beyond cohesion may violate more assumptions of the cluster SCR model. Activity centers and movement patterns of groups may change seasonally in some group-living species. For example, in African wild dogs, denning behavior can impact movement as group home ranges contract during denning periods (Pomilia, McNutt, and Jordan 2015). Sampling could be timed to avoid denning periods for populations that reproduce seasonally (McNutt, Groom, and Woodroffe 2019) and change their movement behavior as a result. Extending the model to include dispersal processes (described above) or modeling group activity

centers as a random walk (Royle, Fuller, and Sutherland 2016; Gardner et al. 2018) would result in more flexible and realistic group movement models. Additionally, with sufficient spatial recaptures and information on captures over time, and the possible integration of telemetry data to improve inference, a more explicit group movement model such as a biased correlated random walk could be used (Langrock et al. 2014; McClintock et al., this issue). Incorporating more realistic movement and population change processes into the cluster SCR model would allow researchers to link individual and group movements to population dynamics, further integrating the fields of movement and population ecology.

By extending SCR models to account for dependence in detections and explicitly model group size, the cluster SCR model can provide unbiased estimates of group abundance, individual abundance, and group size in populations of group-living species. The model can be extended to include covariates on group size, such that it can be used to test hypotheses about ecological predictors of group size (e.g., food availability). Particularly if it is extended to an open-population framework, and individual and group population growth rates can be calculated and correlated with group sizes, the model can provide evidence of density dependence at the group and population levels. Given the prevalence of group-living behavior across species and taxa, developing models that accurately capture group dynamics and population estimates is important for ecological research and biodiversity conservation.

6.5 Tables and Figures

Table 6.1. Posterior modes and 95% credible intervals of African wild dog individual density (dogs/100 km^2), group density (groups/100 km^2), and group size for the cluster SCR model and both individual-level and group-level SCR models applied to the wild dog data for 2014 and 2015. The previous estimate column shows previously published estimates from the Okavango Delta region of Botswana. For the previous estimate of individual density, we report estimates from individual-level SCR model applied to the 2015 data set, averaged across habitat types (Rich et al. 2019). For the previous estimates of group density and group size we report observations derived from long-term monitoring of African wild dog populations (Creel, Mills, and McNutt 2004). The estimates marked by an asterisk are derived using the average number of individuals detected in each group that was detected.

Parameter	Year	Cluster Mode (CI)	Individual-level Mode (CI)	Group-level Mode (CI)	Previous Estimate
Group density	2014	0.28 (0.15-0.48)	-	0.17 (0.10-0.30)	0.81
Group density	2015	0.13 (0.09-0.25)	-	0.12 (0.09-0.22)	0.70
Average group size (μ)	2014	11.97 (9.60-15.33)	-	11.67*	12.32

Average group size (μ)	2015	8.05 (5.91-10.65)	-	7.67*	11.89
Individual density	2014	2.75 (1.73-6.03)	3.79 (2.93-4.42)	1.95* (1.30-3.25*)	2.10
Individual density	2015	1.20 (0.71-2.05)	1.47 (1.13-1.73)	0.89* (0.66-1.66*)	2.10

Table 6.2. Mean relative bias (%) and coverage (%) of the estimated number of groups (G) and average group size parameter (ζ) based on posterior modes for all simulation scenarios. The models considered here are the cluster SCR model and the standard group-level SCR model. Bias and coverage are calculated using replicate data sets which resulted in converged models and models with non-missing Gelman-Rubin diagnostics (some Gelman-Rubin diagnostics could not be calculated).

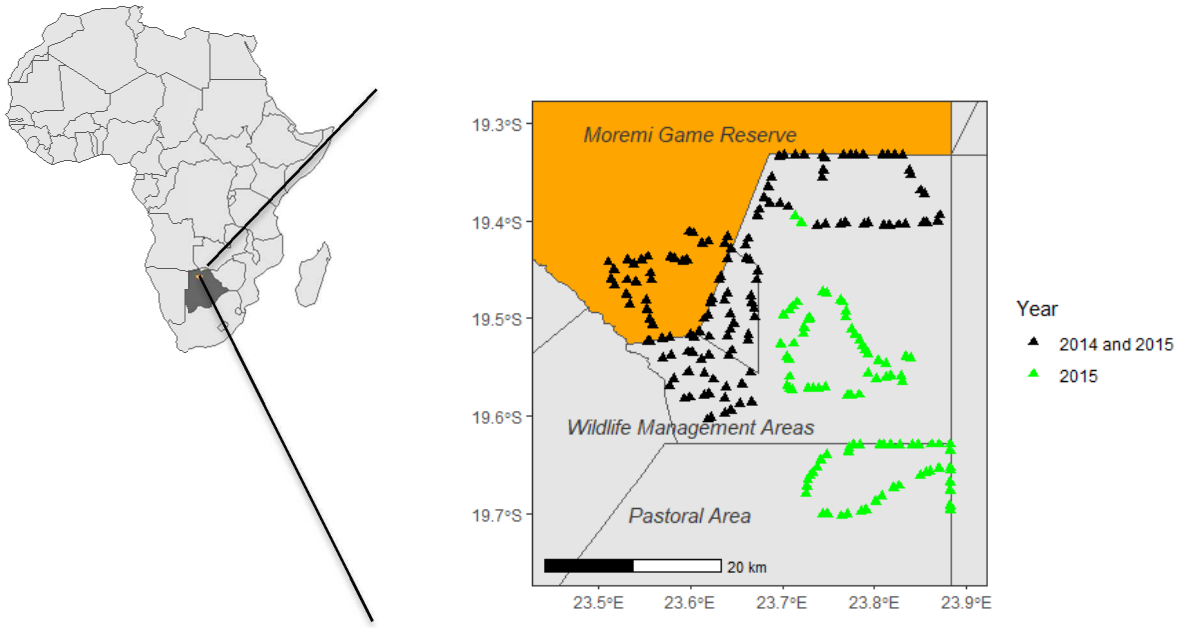
Parameter	True number of groups	True average group size parameter	Cluster SCR Bias	Group-level SCR Bias	Cluster SCR Coverage	Group-level SCR coverage
G	10	5.0	-7.28	-42.65	97.73	29.11
G	10	12.5	-7.77	-42.77	96.94	9.64
G	20	5.0	-1.71	-47.95	96.97	1.08
G	20	12.5	-1.84	-47.74	96.00	0
ζ	10	5.0	3.79	-	94.32	-
ζ	10	12.5	3.40	-	94.90	-
ζ	20	5.0	-1.27	-	95.96	-
ζ	20	12.5	-0.55	-	93.00	-

Table 6.3. Mean relative bias (%) and coverage (%) of the estimated number of groups (G) and average group size parameter (ζ) based on posterior modes for all simulation scenarios using hidden Markov movement models. The model considered here is the cluster SCR model only. Bias and coverage are calculated using replicate data sets which resulted in converged models and models with non-missing Gelman-Rubin diagnostics (some Gelman-Rubin diagnostics could not be calculated).

Parameter	True number of groups	True average group size parameter	Attraction to group centroid	Cluster SCR Bias	Cluster SCR Coverage
G	10	5.0	High	16.40	87.95
G	10	5.0	Low	22.24	89.41
G	10	12.5	High	14.96	88.89
G	10	12.5	Low	22.61	86.90
G	20	5.0	High	15.16	88.89
G	20	5.0	Low	18.39	80.65
G	20	12.5	High	13.36	93.00
G	20	12.5	Low	21.25	68.19
ζ	10	5.0	High	-2.70	93.98
ζ	10	5.0	Low	-12.72	88.24
ζ	10	12.5	High	-7.34	92.93
ζ	10	12.5	Low	-15.71	72.62

ζ	20	5.0	High	-6.81	93.93
ζ	20	5.0	Low	-14.38	76.49
ζ	20	12.5	High	-7.12	85.00
ζ	20	12.5	Low	-19.81	36.36

Figure 6.1. Location of Moremi Game Reserve (orange) in the context of Botswana (dark gray) and Africa (light gray). Inset: plot of camera locations in northern Botswana's Okavango Delta region in 2014 (black) and 2015 (black + green). Two camera locations monitored in 2014 but not 2015 are not plotted here. Boundaries of management units are overlaid.


Moremi Game Reserve
Wildlife Management Areas
Pastoral Area
20 km
19.3°S
19.4°S
19.5°S
19.6°S
19.7°S
23.5°E
23.6°E
23.7°E
23.8°E
23.9°E
Year
2014 and 2015
2015

References

Angulo, E., G.M. Luque, S.D. Gregory, J.W. Wenzel, C. Bessa-Gomes, L. Berec, and F. Courchamp. 2018. "Allee effects in social species." *Journal of Animal Ecology* 87 (1): 47–58.

Asensio, N., A.H. Korstjens, and F. Aureli. 2009. "Fissioning minimizes ranging costs in spider monkeys: a multiple-level approach." *Behavioral Ecology and Sociobiology* 63 (5): 649–59.

Aureli, F., C.M. Schaffner, C. Boesch, S.K. Bearder, J. Call, C.A. Chapman, R. Connor, A. Di Fiore, Robin I. M. Dunbar, S.P. Henzi, K. Holekamp, A.H. Korstjens, R.Layton, P. Lee, J. Lehmann, J.H. Manson, G. Ramos□Fernandez, K.B. Strier, and C.P. van Schaik. 2008. "Fission-fusion dynamics: new research frameworks." *Current Anthropology* 49 (4): 627–54.

Bassing, S.B., D.E. Ausband, M.S. Mitchell, M.K. Schwartz, J.J. Nowak, G.C. Hale, and L.P. Waits. 2020. "Immigration does not offset harvest mortality in groups of a cooperatively breeding carnivore." *Animal Conservation* 23 (6): 750–61.

Behr, D.M., J.W. McNutt, A. Ozgul, and G. Cozzi. 2020. "When to stay and when to leave? Proximate causes of dispersal in an endangered social carnivore." *Journal of Animal Ecology* 89 (10): 2356–66.

Belant, J.L., F. Bled, C.M. Wilton, R. Fyumagwa, S.B. Mwampeta, and D.E. Beyer Jr. 2016. "Estimating Lion Abundance using N-mixture Models for Social Species." *Scientific Reports* 6: 35920.

Bischof, R., P. Dupont, C. Milleret, J. Chipperfield, and J.A. Royle. 2020. "Consequences of ignoring group association in spatial capture-recapture analysis." Wildlife Biology 2020 (1).

Borchers, D.L., and M.G. Efford. 2008. “Spatially explicit maximum likelihood methods for capture-recapture studies.” *Biometrics* 64: 377–85.

Courchamp, F., T. Clutton-Brock, and B. Grenfell. 2000. “Multipack dynamics and the Allee effect in the African wild dog, Lycaon pictus.” *Animal Conservation* 3 (4): 277–85.

Courchamp, F., and D.W. Macdonald. 2001. “Crucial importance of pack size in the African wild dog Lycaon pictus.” *Animal Conservation* 4 (2): 169–74.

Coverdale, T.C., T. R. Kartzinel, K.L. Grabowski, R.K. Shriver, A.A. Hassan, J.R. Goheen, T.M. Palmer, and R.M. Pringle. 2016. “Elephants in the understory: opposing direct and indirect effects of consumption and ecosystem engineering by megaherbivores.” *Ecology* 97 (11): 3219-3230.

Creel, S., and N.M. Creel. 2015. “Opposing effects of group size on reproduction and survival in African wild dogs.” *Behavioral Ecology* 26 (5): 1414–22.

Creel, S., M.G.L. Mills, and J.W. McNutt. 2004. “Demography and population dynamics of African wild dogs in three critical populations.” In *The Biology and Conservation of Wild Canids*, edited by D.W. Macdonald and C. Sillero-Zubiri, 337–50. Oxford: Oxford University Press.

Doolan, S.P., and D.W. Macdonald. 1996. “Diet and foraging behaviour of group-living meerkats, Suricata suricatta, in the southern Kalahari.” *Journal of Zoology* 239 (4): 697–716.

Elizalde, L., M. Arbetman, X. Arnan, P. Eggleton, I.R. Leal, M.N. Lescano, A. Saez, V. Werenkraut, and G.I. Pirk. 2020. “The ecosystem services provided by social insects: traits, management tools and knowledge gaps.” *Biological Reviews* 95 (5): 1418-1441.

Fryxell, J.M., A. Mosser, A.R.E. Sinclair, and C. Packer. 2007. “Group formation stabilizes predator-prey dynamics.” *Nature* 449 (7165): 1041–3.

Gardner, B., R. Sollmann, N.S. Kumar, D. Jathanna, and K.U. Karanth. 2018. “State space and movement specification in open population spatial capture – recapture models.” *Ecology and Evolution* 8 (July): 10336–44.

Gittleman, J.L. 1989. “Carnivore group living: comparative trends.” In *Carnivore Behavior, Ecology, and Evolution*, edited by John L. Gittleman, 183–207. Boston, MA: Springer.

Hickey, J.R., and R. Sollmann. 2018. “A new mark-recapture approach for abundance estimation of social species.” *PLoS ONE* 13 (12): e0208726.

Holekamp, K.E., S.T. Sakai, and B.L. Lundrigan. 2007. “Social intelligence in the spotted hyena (Crocuta crocuta).” *Philosophical Transactions of the Royal Society B: Biological Sciences* 362 (1480): 523–38.

Kellner, K. 2016. “jagsUI: a wrapper around ’rjags’ to streamline ’JAGS’ analyses.” *https://cran.r-project.org/package=jagsUI*.

Langrock, R., J.G.C. Hopcraft, P.G. Blackwell, V. Goodall, R. King, M. Niu, T.A. Patterson, M.W. Pedersen, A. Skarin, and R.S. Schick. 2014. “Modelling group dynamic animal movement.” *Methods in Ecology and Evolution* 5 (2): 190–99.

Lerch, B.A., B.C. Nolting, and K.C. Abbott. 2018. “Why are demographic Allee effects so rarely seen in social animals?” *Journal of Animal Ecology* 87 (6): 1547–59.

Lima, S.L. 1995. “Back to the basics of anti-predatory vigilance: the group-size effect.” *Animal Behaviour* 49 (1): 11–20.

Lopez-Bao, J.V., R. Godinho, C. Pacheco, F.J. Lema, E. Garcia, L. Llaneza, V. Palacios, and J. Jimenez. 2018. “Toward reliable population estimates of wolves by combining spatial capture-recapture models and non-invasive DNA monitoring.” *Scientific Reports* 8: 2177.

Mattioli, L., A. Canu, D. Passilongo, M. Scandura, and M. Apollonio. 2018. “Estimation of pack density in grey wolf (Canis lupus) by applying spatially explicit capture-recapture models to camera trap data supported by genetic monitoring.” *Frontiers in Zoology* 15: 38.

McClanahan, K., F. Rosell, and M. Mayer. 2020. “Minding your own business: low pair cohesion in a territorial, monogamous mammal.” *Animal Behaviour* 166: 119–28.

McClintock, B.T., B. Abrahms, R.B. Chandler, P.B. Conn, S.J. Converse, R.L. Emmet, B. Gardner, N.J. Hostetter, and D.S. Johnson. 2021. “An integrated path for spatial capture recapture and animal movement modeling.” *Ecology* this issue.

McClintock, B.T., and T. Michelot. 2018. “momentuHMM: R package for generalized hidden Markov models of animal movement.” *Methods in Ecology and Evolution* 9 (6): 1518-1530. http://dx.doi.org/10.1111/2041-210X.12995.

McNutt, J.W., R. Groom, and R.Woodroffe. 2019. “Ambient temperature provides an adaptive explanation for seasonal reproduction in a tropical mammal.” *Journal of Zoology* 309 (3): 153–60.

Mueller, T., K.A. Olson, G. Dressler, P. Leimgruber, T.K. Fuller, C. Nicolson, A.J. Novaro, M.J. Bolgeri, D. Wattles, S. DeStefano, J.M. Calabrese, and W.F. Fagan. 2011. “How landscape dynamics link individual- to population-level movement patterns: a multispecies comparison of ungulate relocation data.” *Global Ecology and Biogeography* 20(5): 683-694.

Nunn, C.L., F. Jordan, C.M. McCabe, J.L. Verdolin, and J.H. Fewell. 2015. “Infectious disease and group size: More than just a numbers game.” *Philosophical Transactions of the Royal Society B: Biological Sciences* 370: 20140111.

Patterson, J.E.H., and K.E. Ruckstuhl. 2013. “Parasite infection and host group size: A meta-analytical review.” *Parasitology* 140 (7): 803–13.

Pomilia, M.A., J.W. McNutt, and N.R. Jordan. 2015. “Ecological predictors of African wild dog ranging patterns in Northern Botswana.” *Journal of Mammalogy* 96 (6): 1214–23.

Rich, L.N., D.A.W. Miller, D.J. Muñoz, H.S. Robinson, J.W. McNutt, and M.J. Kelly. 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.

Rich, L.N., D.A.W. Miller, H.S. Robinson, J.W. McNutt, and M.J. Kelly. 2017. “Carnivore distributions in Botswana are shaped by resource availability and intraguild species.” *Journal of Zoology* 303 (2): 90–98.

Royle, J.A., A.K. Fuller, and C. Sutherland. 2016. “Spatial capture–recapture models allowing Markovian transience or dispersal.” *Population Ecology* 58 (1): 53–62.

Royle, J. Andrew, and Kevin V. Young. 2008. “A hierarchical model for spatial capture-recapture data.” *Ecology* 89 (8): 2281–9.

Royle, J.A., R.B. Chandler, R. Sollmann, and B. Gardner. 2013. *Spatial Capture-Recapture*. Cambridge, Massachusetts: Academic Press.

Russell, R.E., J.A. Royle, R. Desimone, M.K. Schwartz, V.L. Edwards, K.P. Pilgrim, and K.S. McKelvey. 2012. “Estimating abundance of mountain lions from unstructured spatial sampling.” *Journal of Wildlife Management* 76 (8): 1551–61.

Schmidt, J.H., K.L. Rattenbury, J.P. Lawler, and M.C. MacCluskie. 2012. “Using distance sampling and hierarchical models to improve estimates of Dall’s sheep abundance.” *Journal of Wildlife Management* 76 (2): 317–27.

Stevenson, B.C., P. van Dam-Bates, C.K.Y. Young, and J. Measey. 2020. “A spatial capture–recapture model to estimate call rate and population density from passive acoustic surveys.” *Methods in Ecology and Evolution* 2020 (October): 1–11.

Sun, C.C., A.K. Fuller, and J.A. Royle. 2014. “Trap configuration and spacing influences parameter estimates in spatial capture-recapture models.” *PLoS ONE* 10 (10): e0141634.

Woodroffe, R., H.M.K. O’Neill, and D. Rabaiotti. 2019. “Within- and between-group dynamics in an obligate cooperative breeder.” *Journal of Animal Ecology* 00: 1–11.

Woodroffe, R., D. Rabaiotti, D.K. Ngatia, T.R.C. Smallwood, S. Strebel, and H.M.K. O’Neill. 2020. “Dispersal behaviour of African wild dogs in Kenya.” *African Journal of Ecology* 58 (1): 46–57.

Woodroffe, R., and C. Sillero-Zubiri. 2012. “Lycaon pictus.” The IUCN Red List of Threatened Species.

Wrona, F.J., and R.W.J. Dixon. 1991. “Group size and predation risk: a field analysis of encounter and dilution effects.” *The American Naturalist* 137 (2): 186–201.

7 Appendix S1: Chapter 1 power analysis methods

For each simulation replicate, we started with a population of 25 individuals in the year prior to the start of surveys. We simulated population change over 10 years, simulating either a 25% or 50% decline or increase in abundance over that time frame (annual growth rates of $R = 0.933, 0.972, 1.022, 1.041$). We simulated adult male and female home ranges using bivariate normal distributions, with the home range radius being approximately two times the standard deviation of the bivariate normal distribution (which was the same for the two normal random variables in each sex-specific bivariate normal distribution). We used estimated 95% fixed-kernel home range estimates of male and female wolverines from Inman et al. (2012) to back-calculate home range radius (assuming roughly circular home ranges), and subsequently the standard deviations we used for adult males and females, respectively. Subadult dispersing individuals moved on average 7.2 km per 12 hours in a correlated random walk, with parameters estimated from telemetry data on 2 dispersing wolverines in the Cascades using R package momentuHMM (C. Raley, unpubl. data; McClintock and Michelot, 2018). We simulated movements of each individual over 240 12-hour time steps (120 days).

Once we simulated population changes and individuals' movements, we simulated a remote camera network and individuals' encounters with each camera. We overlaid a hexagonal grid of cell sizes 225 km^2. We simulated a fixed number of cameras in each simulation, with the number of cameras depending on the proportion of cells sampled (20%, 50%, or 80% of cells). We used only one camera per sampled cell. These cells remained the same throughout the 10 years of each simulation; that is, cameras were never moved. An individual was detected perfectly if it moved within 1.5 km of a camera.

For each of the 10 years in a given simulation replicate, we created encounter histories for each survey site consisting of the number of days with detections in a given 30-day secondary occasion (each primary occasion had 4). Using these encounter histories, we fit both our continuous-time multi-scale model in which occupancy is parametrized using colonization and extinction probabilities (the explicitly dynamic model), and our model in which occupancy is estimated independently each year (the implicitly dynamic model). We fit each model in JAGS using R package jagsUI (Kellner, 2016). We ran all models in parallel using 3 cores. We used 5,000 iterations after a burn-in period of 2,000 iterations for 3 chains, with a thin rate of 2. We then ran linear models of both occupancy and snapshot occupancy over time, i.e., $\psi_t = \beta_0 + \beta_1 t + w_t$, where ψ_t is an annual derived occupancy (or snapshot occupancy) estimate for year t, β_0 and β_1 are intercept and slope parameters, respectively, and $w_t \sim N(0, \sigma^2)$, with σ estimated. If a significant trend was detected (the p-value from the slope parameter of the linear model was less than 0.05) and the sign of the trend was correct (e.g., the slope parameter was negative when there was a population decline), we recorded the trend as detected.

We repeated these steps for 100 replicates of each combination of population decline and proportion of cells sampled. For each scenario and model type (explicitly vs. implicitly dynamic), we calculated the proportion of replicates where a trend in occupancy was detected, and the proportion where a trend in snapshot occupancy was detected, which gave simulated power of occupancy and snapshot occupancy, respectively, for each scenario.

Supplemental References

Agostinelli, C., and U. Lund. 2018. CircStats: circular statistics, from "Topics in Circular Statistics" (2001).

Bivand, R.S., E.J. Pebesma, and V. Gomez-Rubio. 2013. Applied spatial data analysis with R. 2nd edition. Springer, New York, USA.

Bivand, R.S. and C. Rundel. 2019. rgeos: Interface to Geometry Engine - Open Source (’GEOS’).

Hijmans, R.J. 2019. Raster: Geographic data analysis and modeling.

Inman, R.M., Packila, M.L., Inman, K.H., McCue, A.J., White, G.C., Persson, J., et al. (2012) Spatial ecology of wolverines at the southern periphery of distribution. *Journal of Wildlife Management*, **76**, 778–792.

Kellner, K. (2016) jagsUI: a wrapper around ’rjags’ to streamline ’JAGS’ analyses.

McClintock, B.T., and T. Michelot. 2018. momentuHMM: R package for generalized hidden Markov models of animal movement. Methods in Ecology and Evolution 9: 1518–1530.

Pebesma, E.J., and R.S. Bivand. 2005. Classes and methods for spatial data in R. R News 5.

Wickham, H. 2016. Ggplot2: Elegant graphics for data analysis.

8 Appendix S2: JAGS code and simulation results tables for Chapter 1

JAGS code for continuous-time multi-scale occupancy model:

```
model{

   # continuous-time multi-scale occupancy model with one primary occasion

   #Priors

   int_psi ~ dunif(0,1) # occupancy parameter

   int_q ~ dunif(0,1) # use parameter, called theta in paper

   int_lambda ~ dunif(0,1)    # detection rate parameter



   #Likelihood

   for(i in 1:nsites){

   psi[i] <- int_psi

   z[i] ~ dbin(psi[i],1)

   for(j in 1:ntimes){ # ntimes is number of secondary occasions

   q[i,j] <- int_q

   u[i,j] ~ dbin(z[i]*q[i,j],1) # use is conditional on the site being occupied

   y[i,j] ~ dpois(u[i,j]*int_lambda*L[i,j]) # detection is conditional on use, and includes an effort
variable, L, read in as data
```

```
    }

    }

    psiq <- int_psi*int_q # this is instantaneous occupancy; we monitor it as a separate parameter rather than calculating post-hoc summaries from occupancy and use parameters


    }
```

Figure S1. Posterior mean estimates of detection probability from all simulation scenarios. True occupancy probability increases from the left panel (probability 0.2) to the right panel (probability 0.8). Detection probability increases from the top panel (probability 0.2) to the bottom panel (probability 0.8). The dark, medium, and light gray boxplots represent scenarios with true use probability of 0.2, 0.5, and 0.8, respectively. Axis limits vary for each row of panels.

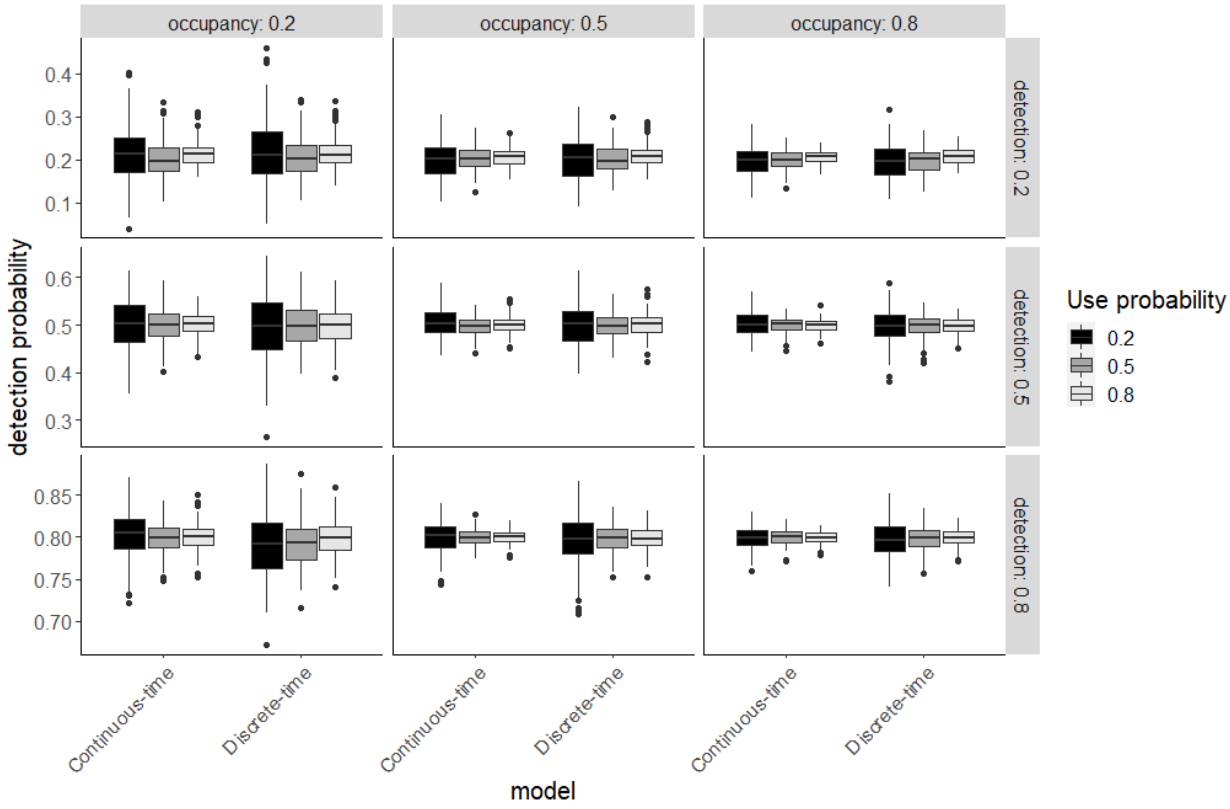


Figure S2. Posterior mean estimates of use probability from all simulation scenarios. True occupancy probability increases from the left panel (probability 0.2) to the right panel (probability 0.8). Detection probability increases from the top panel (probability 0.2) to the bottom panel (probability 0.8). The dark, medium, and light gray boxplots represent scenarios with true use probability of 0.2, 0.5, and 0.8, respectively.

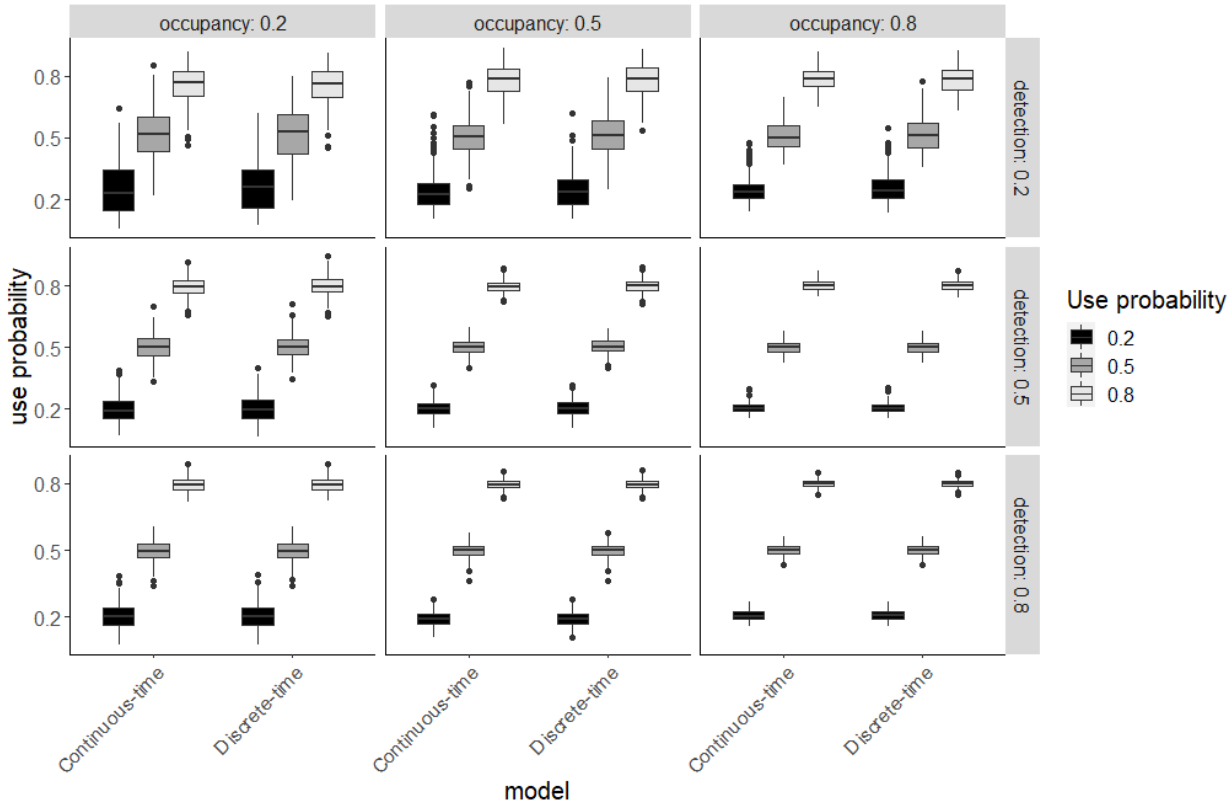


Table S1. Mean relative bias of detection probabilities (percent) from continuous-time and discrete-time multi-scale occupancy models for all detection, occupancy and use scenarios.

Detection	Use	Occupancy	Continuous-time	Discrete-time
0.2	0.2	0.2	7.454	9.227
0.5	0.2	0.2	0.409	-1.065
0.8	0.2	0.2	0.244	-1.187
0.2	0.5	0.2	0.472	1.839
0.5	0.5	0.2	0.008	-0.457
0.8	0.5	0.2	-0.209	-0.914
0.2	0.8	0.2	6.270	8.522
0.5	0.8	0.2	0.179	-0.323
0.8	0.8	0.2	-0.038	-0.292
0.2	0.2	0.5	-0.596	0.479
0.5	0.2	0.5	0.817	0.104
0.8	0.2	0.5	-0.049	-0.476
0.2	0.5	0.5	1.270	0.089
0.5	0.5	0.5	-0.416	-0.413
0.8	0.5	0.5	-0.105	-0.287
0.2	0.8	0.5	3.318	4.520

0.5	0.8	0.5	0.123	0.010
0.8	0.8	0.5	0.001	-0.222
0.2	0.2	0.8	-1.567	-2.850
0.5	0.2	0.8	0.385	-0.436
0.8	0.2	0.8	-0.102	-0.416
0.2	0.5	0.8	0.237	-0.880
0.5	0.5	0.8	0.081	-0.300
0.8	0.5	0.8	-0.048	-0.196
0.2	0.8	0.8	3.437	3.925
0.5	0.8	0.8	-0.171	-0.280
0.8	0.8	0.8	-0.049	-0.114

Table S2. Mean relative bias of occupancy probabilities (percent) from continuous-time and discrete-time multi-scale occupancy models for all detection, occupancy and use scenarios.

Detection	Use	Occupancy	Continuous-time	Discrete-time
0.2	0.2	0.2	55.068	51.315
0.5	0.2	0.2	30.627	28.963
0.8	0.2	0.2	17.621	17.457
0.2	0.5	0.2	13.214	12.005
0.5	0.5	0.2	5.876	5.941
0.8	0.5	0.2	3.853	3.838
0.2	0.8	0.2	4.359	4.298
0.5	0.8	0.2	2.814	2.792
0.8	0.8	0.2	1.121	1.123
0.2	0.2	0.5	10.822	10.065
0.5	0.2	0.5	7.041	6.971
0.8	0.2	0.5	8.113	8.215
0.2	0.5	0.5	5.299	5.738
0.5	0.5	0.5	1.579	1.590
0.8	0.5	0.5	1.191	1.190
0.2	0.8	0.5	1.871	1.915
0.5	0.8	0.5	0.199	0.194

0.8	0.8	0.5	0.717	0.708
0.2	0.2	0.8	-8.762	-8.594
0.5	0.2	0.8	-0.151	-0.104
0.8	0.2	0.8	-1.296	-1.296
0.2	0.5	0.8	0.770	0.998
0.5	0.5	0.8	0.156	0.186
0.8	0.5	0.8	-0.098	-0.114
0.2	0.8	0.8	-0.273	-0.200
0.5	0.8	0.8	-0.381	-0.380
0.8	0.8	0.8	-0.472	-0.473

Table S3. Mean relative bias of use probabilities (percent) from continuous-time and discrete-time multi-scale occupancy models for all detection, occupancy and use scenarios.

Detection	Use	Occupancy	Continuous-time	Discrete-time
0.2	0.2	0.2	28.173	33.104
0.5	0.2	0.2	-0.906	1.419
0.8	0.2	0.2	1.244	1.428
0.2	0.5	0.2	3.128	4.258
0.5	0.5	0.2	0.140	0.526
0.8	0.5	0.2	-0.882	-0.747
0.2	0.8	0.2	-5.130	-5.866
0.5	0.8	0.2	-0.390	0.015
0.8	0.8	0.2	-0.628	-0.565
0.2	0.2	0.5	19.859	22.911
0.5	0.2	0.5	0.934	1.704
0.8	0.2	0.5	-2.697	-2.643
0.2	0.5	0.5	1.053	3.006
0.5	0.5	0.5	0.589	0.708
0.8	0.5	0.5	-0.298	-0.263
0.2	0.8	0.5	-2.601	-2.918
0.5	0.8	0.5	-0.163	-0.040

0.8	0.8	0.5	-0.548	-0.513
0.2	0.2	0.8	23.430	28.541
0.5	0.2	0.8	2.499	3.039
0.8	0.2	0.8	3.517	3.592
0.2	0.5	0.8	1.391	3.224
0.5	0.5	0.8	-0.028	0.206
0.8	0.5	0.8	0.247	0.269
0.2	0.8	0.8	-2.072	-2.377
0.5	0.8	0.8	0.023	0.117
0.8	0.8	0.8	-0.154	-0.142

Table S4. Mean relative bias of instantaneous occupancy probabilities (percent) from continuous-time and discrete-time multi-scale occupancy models for all detection, occupancy, and use scenarios.

Detection	Use	Occupancy	Continuous-time	Discrete-time
0.2	0.2	0.2	35.576	43.335
0.5	0.2	0.2	4.23	6.135
0.8	0.2	0.2	1.888	2.053
0.2	0.5	0.2	8.206	8.909
0.5	0.5	0.2	5.029	5.534
0.8	0.5	0.2	2.294	2.417
0.2	0.8	0.2	-2.078	-2.719
0.5	0.8	0.2	2.482	2.879
0.8	0.8	0.2	0.443	0.507
0.2	0.2	0.5	15.34	18.255
0.5	0.2	0.5	1.788	2.484
0.8	0.2	0.5	0.36	0.453
0.2	0.5	0.5	3.484	6.056
0.5	0.5	0.5	1.856	1.99
0.8	0.5	0.5	0.686	0.722
0.2	0.8	0.5	-1.201	-1.433

0.5	0.8	0.5	-0.02	0.086
0.8	0.8	0.5	0.126	0.152
0.2	0.2	0.8	6.984	11.823
0.5	0.2	0.8	0.396	0.965
0.8	0.2	0.8	0.405	0.485
0.2	0.5	0.8	0.886	2.981
0.5	0.5	0.8	-0.101	0.156
0.8	0.5	0.8	-0.024	0.028
0.2	0.8	0.8	-2.584	-2.763
0.5	0.8	0.8	-0.38	-0.28
0.8	0.8	0.8	-0.624	-0.613

Table S5. RMSE of detection probabilities from continuous-time and discrete-time multi-scale occupancy models for all detection, occupancy, and use scenarios.

Detection	Use	Occupancy	Continuous-time	Discrete-time
0.2	0.2	0.2	0.069	0.079
0.5	0.2	0.2	0.054	0.070
0.8	0.2	0.2	0.028	0.040
0.2	0.5	0.2	0.043	0.046
0.5	0.5	0.2	0.035	0.044
0.8	0.5	0.2	0.017	0.029
0.2	0.8	0.2	0.029	0.038
0.5	0.8	0.2	0.024	0.034
0.8	0.8	0.2	0.015	0.021
0.2	0.2	0.5	0.043	0.049
0.5	0.2	0.5	0.032	0.045
0.8	0.2	0.5	0.018	0.029
0.2	0.5	0.5	0.028	0.033
0.5	0.5	0.5	0.019	0.026
0.8	0.5	0.5	0.010	0.017
0.2	0.8	0.5	0.021	0.025
0.5	0.8	0.5	0.017	0.023

0.8	0.8	0.5	0.008	0.014
0.2	0.2	0.8	0.033	0.040
0.5	0.2	0.8	0.026	0.034
0.8	0.2	0.8	0.012	0.022
0.2	0.5	0.8	0.022	0.028
0.5	0.5	0.8	0.016	0.022
0.8	0.5	0.8	0.009	0.014
0.2	0.8	0.8	0.016	0.020
0.5	0.8	0.8	0.013	0.017
0.8	0.8	0.8	0.006	0.010

Table S6. RMSE of occupancy probabilities from continuous-time and discrete-time multi-scale occupancy models for all detection, occupancy, and use scenarios.

Detection	Use	Occupancy	Continuous-time	Discrete-time
0.2	0.2	0.2	0.169	0.166
0.5	0.2	0.2	0.114	0.109
0.8	0.2	0.2	0.087	0.086
0.2	0.5	0.2	0.061	0.06
0.5	0.5	0.2	0.035	0.035
0.8	0.5	0.2	0.03	0.03
0.2	0.8	0.2	0.035	0.035
0.5	0.8	0.2	0.03	0.029
0.8	0.8	0.2	0.03	0.03
0.2	0.2	0.5	0.145	0.146
0.5	0.2	0.5	0.104	0.105
0.8	0.2	0.5	0.102	0.103
0.2	0.5	0.5	0.078	0.079
0.5	0.5	0.5	0.039	0.039
0.8	0.5	0.5	0.038	0.038
0.2	0.8	0.5	0.044	0.044
0.5	0.8	0.5	0.033	0.033

0.8	0.8	0.5	0.033	0.033
0.2	0.2	0.8	0.126	0.116
0.5	0.2	0.8	0.068	0.068
0.8	0.2	0.8	0.073	0.072
0.2	0.5	0.8	0.068	0.069
0.5	0.5	0.8	0.036	0.036
0.8	0.5	0.8	0.037	0.037
0.2	0.8	0.8	0.043	0.042
0.5	0.8	0.8	0.028	0.028
0.8	0.8	0.8	0.026	0.026

Table S7. RMSE of use probabilities from continuous-time and discrete-time multi-scale occupancy models for all detection, occupancy, and use scenarios.

Detection	Use	Occupancy	Continuous-time	Discrete-time
0.2	0.2	0.2	0.144	0.148
0.5	0.2	0.2	0.069	0.071
0.8	0.2	0.2	0.058	0.059
0.2	0.5	0.2	0.130	0.131
0.5	0.5	0.2	0.057	0.057
0.8	0.5	0.2	0.048	0.048
0.2	0.8	0.2	0.095	0.100
0.5	0.8	0.2	0.048	0.051
0.8	0.8	0.2	0.031	0.031
0.2	0.2	0.5	0.100	0.100
0.5	0.2	0.5	0.038	0.040
0.8	0.2	0.5	0.034	0.034
0.2	0.5	0.5	0.094	0.102
0.5	0.5	0.5	0.037	0.038
0.8	0.5	0.5	0.033	0.033
0.2	0.8	0.5	0.078	0.082
0.5	0.8	0.5	0.029	0.031

0.8	0.8	0.5	0.022	0.022
0.2	0.2	0.8	0.079	0.091
0.5	0.2	0.8	0.025	0.026
0.8	0.2	0.8	0.025	0.025
0.2	0.5	0.8	0.071	0.084
0.5	0.5	0.8	0.030	0.030
0.8	0.5	0.8	0.023	0.024
0.2	0.8	0.8	0.056	0.063
0.5	0.8	0.8	0.025	0.026
0.8	0.8	0.8	0.017	0.017

Table S8. RMSE of instantaneous occupancy probabilities from continuous-time and discrete-time multi-scale occupancy models for all detection, occupancy, and use scenarios.

Detection	Use	Occupancy	Continuous-time	Discrete-time
0.2	0.2	0.2	0.029	0.035
0.5	0.2	0.2	0.009	0.009
0.8	0.2	0.2	0.008	0.008
0.2	0.5	0.2	0.027	0.029
0.5	0.5	0.2	0.018	0.019
0.8	0.5	0.2	0.015	0.015
0.2	0.8	0.2	0.028	0.029
0.5	0.8	0.2	0.026	0.027
0.8	0.8	0.2	0.024	0.024
0.2	0.2	0.5	0.035	0.039
0.5	0.2	0.5	0.013	0.013
0.8	0.2	0.5	0.013	0.013
0.2	0.5	0.5	0.044	0.051
0.5	0.5	0.5	0.024	0.024
0.8	0.5	0.5	0.023	0.023
0.2	0.8	0.5	0.045	0.048
0.5	0.8	0.5	0.028	0.028

0.8	0.8	0.5	0.027	0.027
0.2	0.2	0.8	0.039	0.048
0.5	0.2	0.8	0.015	0.015
0.8	0.2	0.8	0.013	0.013
0.2	0.5	0.8	0.05	0.062
0.5	0.5	0.8	0.023	0.023
0.8	0.5	0.8	0.022	0.022
0.2	0.8	0.8	0.05	0.058
0.5	0.8	0.8	0.028	0.029
0.8	0.8	0.8	0.025	0.025

9 Appendix S3: Prior and mesh sensitivity analyses and eBird GAM methods

Prior and mesh sensitivity analyses

We varied parameters of the prior distribution and mesh we used in INLA to assess the sensitivity of our results to our parameter choices. We varied both range and marginal standard deviation threshold choices from the defaults of 128 km (such that $P(\rho < 128000) = 0.05$) and 5 (such that $P(\sigma > 5) = 0.05$), respectively. For the prior sensitivity analysis, we used the same mesh and other model settings as for our original analysis. For each combination of prior parameters, we fit a model using $FREEZEDAYS$ and stored the hyperparameter estimates (random effect precision estimates, range and standard deviation for the spatial field, and spatial field multiplier) for each model run. Similarly, for the mesh sensitivity analysis, we varied values of the inner maximum edge argument (the maximum edge length of mesh triangles in the inner boundaries of the mesh), the cutoff value (the maximum distance between points at which multiple points are assigned the same vertex), and the offset value (the distance between the inner mesh boundary and the points).

For the prior sensitivity analysis, we found that the mean range estimate and the mean marginal standard deviation increased as the range parameter increased, but that neither was especially affected by the standard deviation parameter (Figure S1, S2). For the mesh sensitivity analysis, results were especially sensitive to cutoff and offset values, but neither mean range nor mean marginal standard deviation varied greatly over different mesh parameters given the same PC prior distribution (Figure S3, S4).

Figure S1. Mean range values for all combinations of range and marginal standard deviation parameters for the prior sensitivity analysis.

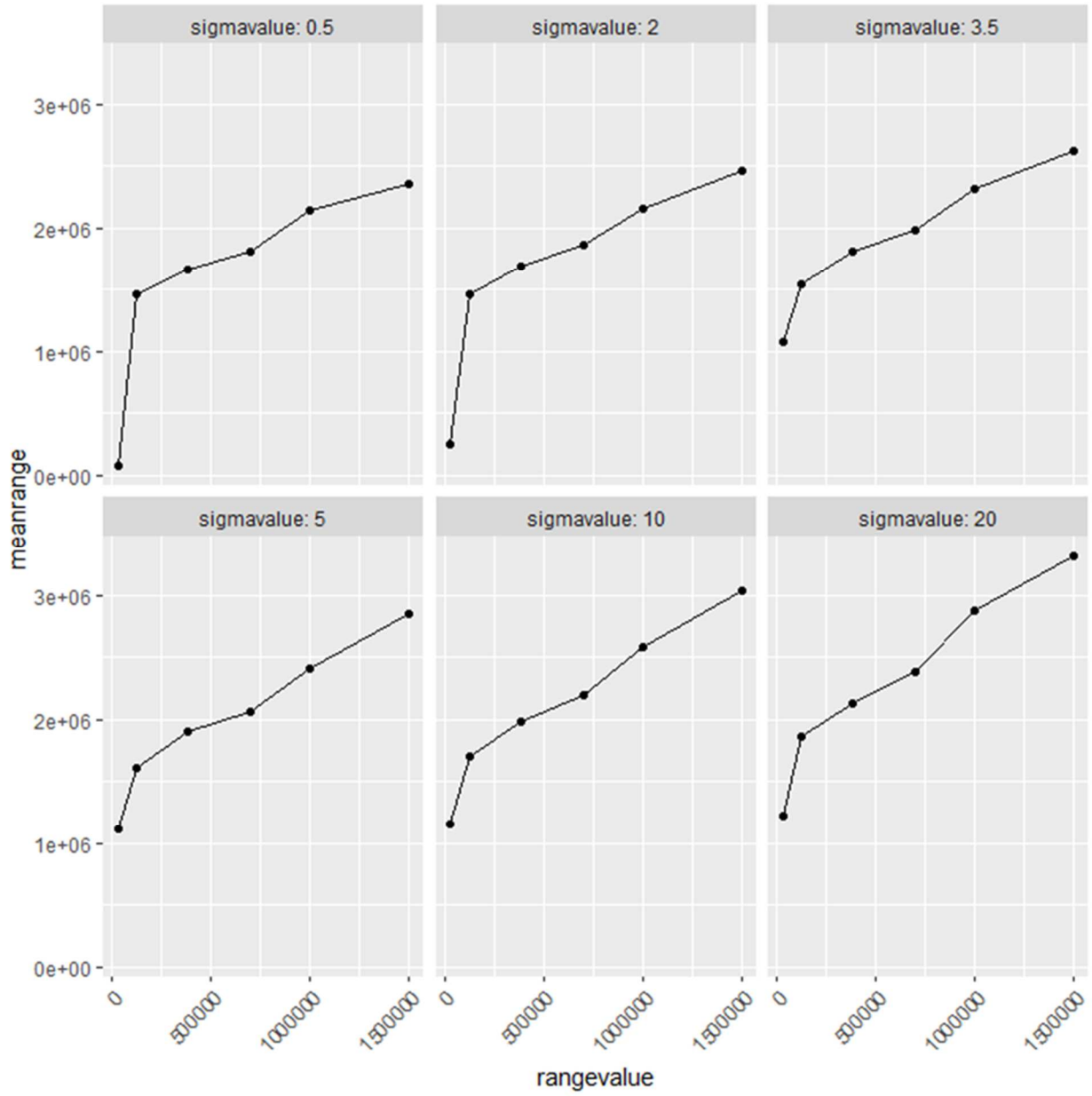


Figure S2. Mean marginal standard deviation values for all combinations of range and marginal standard deviation parameters for the prior sensitivity analysis.

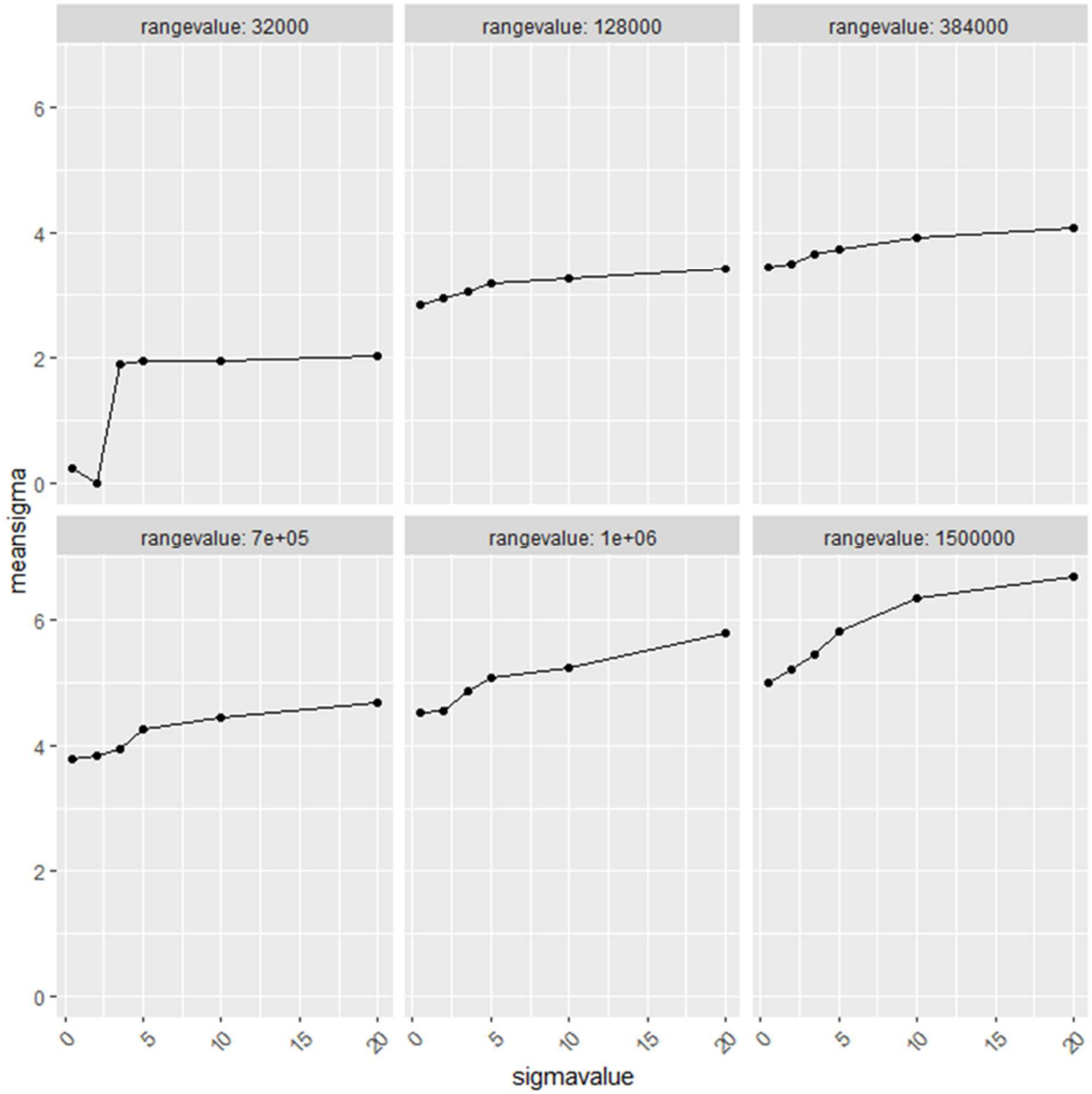


Figure S3. Mean range values for all combinations of maximum edge, cutoff, and offset parameters for the mesh sensitivity analysis.

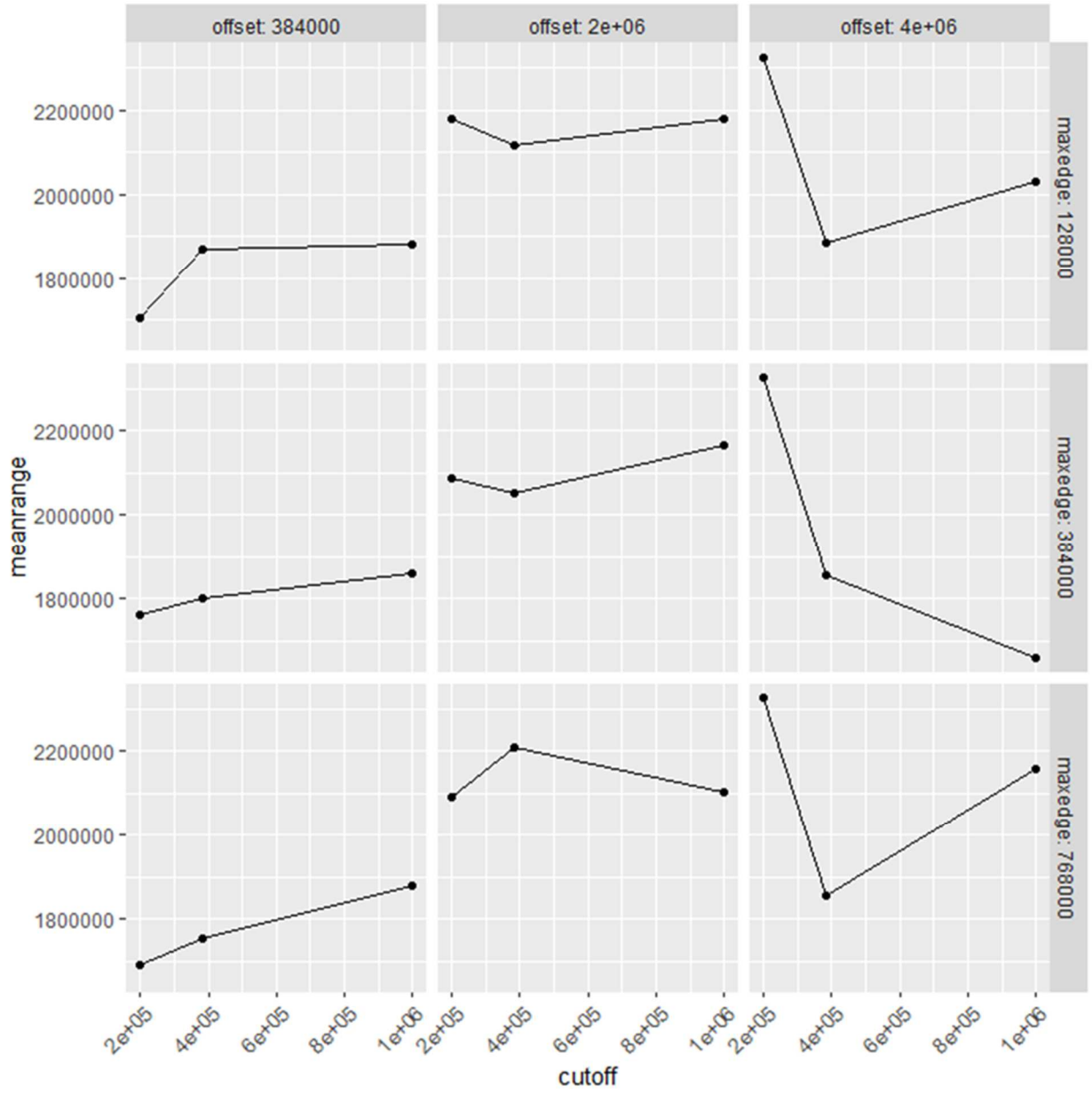
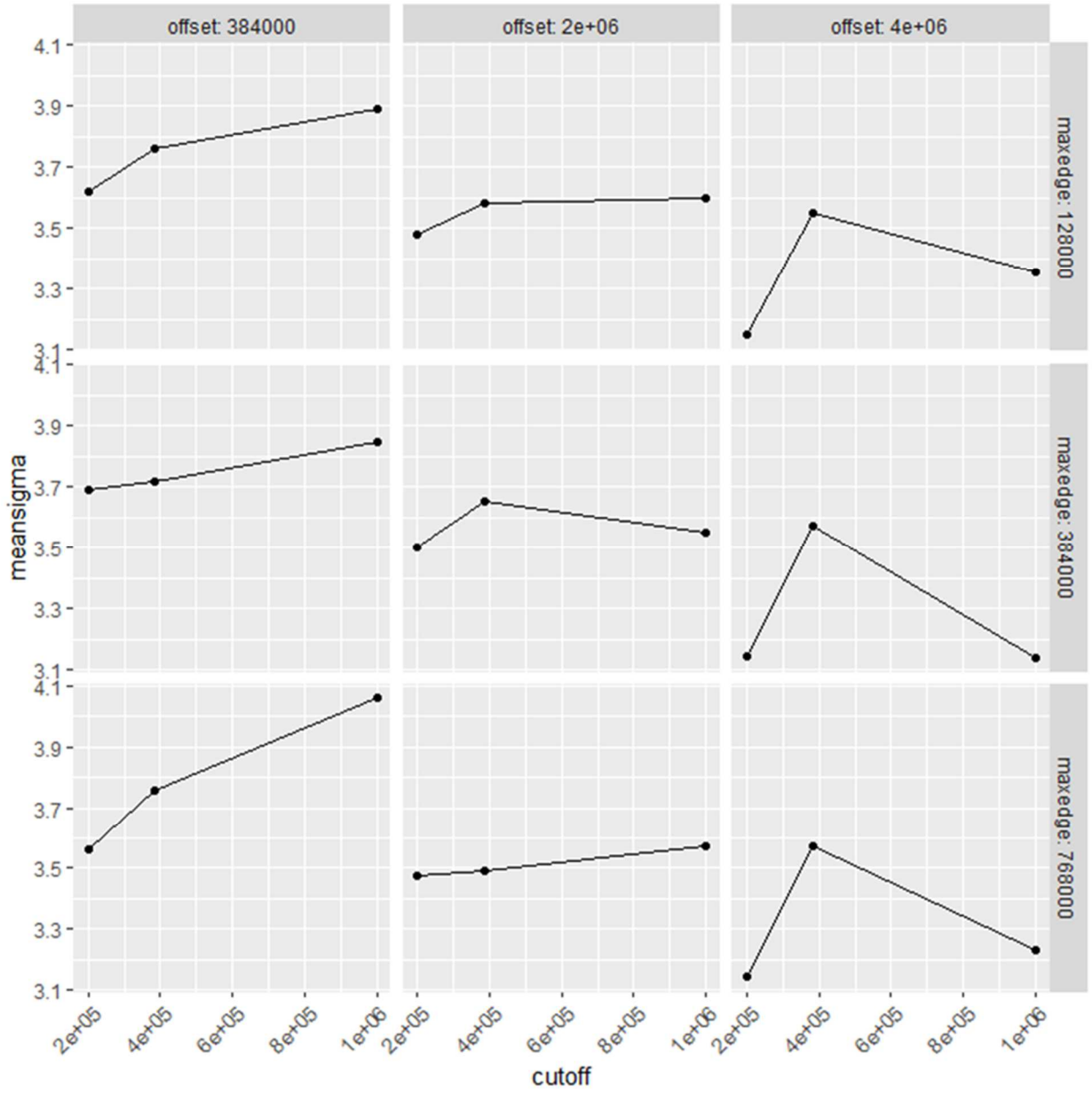


Figure S4. Mean marginal standard deviation values for all combinations of maximum edge, cutoff, and offset parameters for the mesh sensitivity analysis.



eBird GAM methods

To estimate relative abundance of black ducks across the study area, we used observations from eBird and fit a spatial generalized additive model (GAM) to the data. There are many factors that an influence the number of birds that an eBird observer records including if they are traveling, how far they travel, how many people are observing, how long they were counting birds, etc. Because of this, we used only stationary and traveling checklists with fewer than 10 observers per checklist, fewer than 5 hours' duration, and less than 5 km traveled (for traveling checklists; distance traveled was recorded as 0 for stationary checklists; Strimas-Mackey et al. 2020). We did not consider observations west of 100˚ W longitude, and subsampled all checklists down to one checklist in each NARR grid cell, week, and year to reduce spatial and temporal sampling bias (Strimas-Mackey et al. 2020). Because we were interested in relative abundance when post-season banding occurred, we used observations from the Mississippi and Atlantic Flyway states during February 1st-March 10th, 2010-2017. In order to infer when black ducks were not observed, we used only complete checklists. We fit a GAM using mgcv version 1.8-31 (Wood 2017) in R version 4.0.2 (R Core Team 2020). The resulting GAM contained a linear effect of protocol type (stationary or traveling), linear and quadratic effects of the observation start time, and cubic spline smooths of duration (minutes), distance traveled (km), day of year, and number of observers, and a thin-plate spline of Easting and Northing using a UTM projection. We chose these covariates to account for variation in counts due to changes in effort, black duck activity at different times of day and year (Morton, Fowler, and Kirkpatrick 1989; Jones, Williams, and Castelli 2014), and spatial variation, all of which can be present in community science data (Johnston et al. 2019). We used a negative binomial distribution to model observed counts, and used efficient matrix computation techniques and discretization of covariates provided by the command bam() in R package mgcv to reduce memory usage (Wood 2017). For each winter location included in our final band recovery model, we predicted the expected number of black ducks that would be observed on February 20th at that banding location, by 1 observer traveling 1 kilometer over 1 hour, beginning at 1000 hours local time.

Table S1. Parameter estimates, standard deviations, t values, and p values for linear and quadratic effects in GAM fit to eBird data. The t and p values are used for a 2-tailed t-test to test whether each estimate is significantly different from 0.

Variable	Estimate	SD	t value	p value
Intercept	-4.712	0.650	-7.240	4.550×10^{-13}
Protocol type (traveling)	0.435	0.050	8.691	$< 2.000 \times 10^{-16}$
Time observations started	0.130	0.032	4.115	3.880×10^{-5}
(Time observations started)^2	-0.003	0.001	-2.366	0.018

Table S2. Estimated degrees of freedom, reference degrees of freedom, F values and p values for smooth functions in GAM fit to eBird data. The F values and approximate p values are used to test whether each smooth term is significantly different from 0 (Wood 2017).

Variable	Estimated degrees of freedom	Reference degrees of freedom	F value	p value
Day of year	1.001	1.002	1.217	0.270
Duration (minutes)	14.048	15.808	11.418	$< 2.000 \times 10^{-16}$
Distance traveled (km)	1.000	1.000	33.262	8.050×10^{-9}
Number of observers	1.862	2.280	15.095	1.480×10^{-7}

Longitude + latitude	684.113	839.180	11.206	$< 2.000 \times 10^{-16}$

Supplementary Literature Cited

Johnston, A., W.M. Hochachka, M.E. Strimas-Mackey, V. Ruiz Gutierrez, O.J. Robinson, E.T. Miller, T. Auer, S.T. Kelling, and D. Fink. 2019. "Analytical guidelines to increase the value of citizen science data: using eBird data to estimate species occurrence." bioRxiv, 574392.

Jones, O.E., C.K. Williams, and P.M. Castelli. 2014. "A 24-Hour Time-Energy Budget for Wintering American Black Ducks (Anas rubripes) and its Comparison to Allometric Estimations." Waterbirds 37 (3): 264–73.

Strimas-Mackey, M., W.M. Hochachka, V. Ruiz Guterriez, O.J. Robinson, E.T. Miller, T. Auer, S.T. Kelling, D. Fink, and A. Johnston. 2020. Best practices for using eBird data. Version 1. Ithaca, NY: Cornell Lab of Ornithology. *https://doi.org/https://doi.org/10.5281/zenodo.3620739*.

Morton, J. M., A. C. Fowler, and R. L. Kirkpatrick. 1989. "Time and Energy Budgets of American Black Ducks in Winter." Journal of Wildlife Management 53:401–10.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Wood, S. 2017. Generalized additive models: an introduction with R. 2nd ed. Chapman; Hall/CRC.

10 Appendix S4: MCMC Algorithm for Chapter 4

The joint posterior distribution is:

$$[\psi, \zeta, \lambda_0, \sigma, \theta, \mathbf{n_g}, \mathbf{S}, \mathbf{z}, \mathbf{w}|\mathbf{Y}]$$

$$\propto \left\{ \prod_{g=1}^{Z} [z_g|\psi][n_g|\zeta][\mathbf{s}_h] \left\{ \prod_{j=1}^{J} \left\{ \prod_{t=1}^{T} [v_{gjt}|\lambda_0, \sigma] \left\{ \prod_{i \ni \gamma_i = g} [y_{ijt}|v_{gjt}, \theta, w_i] \right\} \right\} \right\} \right\}$$

$$\times [\psi][\zeta][\lambda_0][\sigma][\theta]$$

We assume the following prior distributions:

$$\pi(\psi) \sim \text{Uniform}(0,1)$$

$$\pi(\mathbf{s}_g) \sim \text{Uniform}(\mathcal{M})$$

$$\pi(\zeta) \sim \text{Uniform}(0, \infty)$$

$$\pi(\lambda_0) \sim \text{Uniform}(0, \infty)$$

$$\pi(\sigma) \sim \text{Uniform}(0, \infty)$$

$$\pi(\theta) \sim \text{Uniform}(0, \infty)$$

Data Augmentation Structure and Initialization

In order to describe several steps for the MCMC sampling, we must first describe the data augmentation structure and how it is initialized. The observed data consists of detections of individuals, with the number of detections for individual i at site j and occasion t being y_{ijt}. We also require that the group membership of every detected individual, γ_i, is known, though this can be relaxed (not considered here). Then, the latent variables for group encounters consists of the number of group encounters by each group across sites and occasions with v_{gjt} being the number of detections for group g at site j on occasion t. Group encounters are fully latent,

though if at least one individual in group g is observed at site j on occasion t, v_{gjt} is constrained to be greater than zero (individuals cannot be detected if their group does not visit a site).

Because both groups and individuals are augmented, we require two data augmentation level parameters, Z and W. We augment $\mathbf{V}$ up to dimension $Z \times J \times K$ and we augment $\mathbf{Y}$ up to dimension $W \times J \times K$. The Z-length indicator vector $\mathbf{z}$ is associated with groups and the W-length indicator vector $\mathbf{w}$ is associated with individuals. Finally, because $\boldsymbol{\gamma}$ links individual i to their group number, g, we must augment this vector up to length W. As with typical data augmentation, we require that $Z >> G$ and $W >> N$, though we require more considerations for setting W, discussed below.

A key requirement of this model is that when an undetected group is included in the population, switching a z_g from 0 to 1, we are able to jointly include an appropriate number of individuals belonging to this group in the population. This appropriate number is n_g, the size of group g. When $z_g = 1$, $n_g = \sum_{i \ni \gamma_i = g} w_i$, the number of individuals i assigned to group g who are included in the population (for which $w_i = 1$). When $z_g = 0$, n_g is a latent count updated from ζ. Therefore, when proposing to turn on an undetected group, we require there to always be n_g or greater augmented individuals with group identity g (individuals with $\gamma_i = g$). Because the latent indices of n_g are continuously updated, we require each group to be augmented with more individuals than the maximum value n_g will ever take during sampling, n_g^{max}. Further, when proposing to add group members to detected groups we also require there be up to n_g^{max} augmented individuals that can be turned on.

Given these considerations, when augmenting $\mathbf{Y}$ up to size W, we ensure that each group has at least n_g^{max} individuals allocated to it by assigning at least n_g^{max} elements of $\boldsymbol{\gamma}$ to each group

($W \geq Z \times n_g^{max}$). The approximate size of the required n_g^{max} can be obtained from the PMF of the Zero-Truncated Poisson group size distribution using the expected value of ζ (e.g., how large must a count be in order for it to be very improbable given the rate parameter) and n_g^{max} can be raised from there if n_g is ever sampled to be larger than n_g^{max}.

MCMC Sampling

1. Update the group encounter parameters, λ_0 and σ. We use α with a subscript to denote tuning parameters, e.g., α_ζ would be the tuning parameter for ζ, the standard deviation of the symmetric, Normal proposal distribution (see below). For both parameters, we use a symmetric Metropolis-Hastings step drawn from Normal($\lambda_0^{curr}, \alpha_{\lambda_0}$) and Normal($\sigma^{curr}, \alpha_\sigma$), respectively, with proposals falling below 0 automatically rejected. For λ_0, $[\lambda_0|\mathbf{V}, \sigma, \mathbf{s}] \propto \prod_g \prod_j \prod_t \text{Poisson}(\lambda_{gjt})$ where $\lambda_{gtk} = \lambda_0 \exp\left(\frac{-d_{gj}^2}{2\sigma^2}\right)$. Similarly, for σ, $[\sigma|\mathbf{V}, \lambda_0, \mathbf{s}] \propto \prod_g \prod_j \prod_t \text{Poisson}(\lambda_{gjt})$.

2. Update the individual detection parameter, θ. We use a symmetric Metropolis-Hastings update drawn from Normal($\theta^{curr}, \alpha_\theta$) and $[\theta|\mathbf{Y}, \mathbf{V}, \mathbf{w}] \propto \prod_i \prod_j \prod_t \text{Poisson}(\theta \times v_{gjt} \times w_i)$ for $g \ni \gamma_i = g$ (mapping individual i to the number of group encounters for the group it is assigned to).

3. Update the group size parameter, ζ. We use a symmetric Metropolis-Hastings step drawn from Normal($\zeta^{curr}, \alpha_\zeta$) and $[\zeta|\mathbf{n}] \propto \prod_g \text{ZTPoisson}(\zeta)$. We reject all proposals that fall below zero.

4. Update the group size, n_g, for groups with $z_g = 0$. Because these group size counts are fully latent, we update them using a Gibbs step and the full conditional for n_g, $[n_g] \propto$ ZTPoisson(ζ).

5. Update group encounters, $\mathbf{V}$. We update all elements of $\mathbf{V}$ using a symmetric Metropolis-Hastings step that proposes to add or subtract 1 site visit each with probability 0.5. For elements of v_{gjt} for which the members of group g have at least one detection at site j on occasion t, we automatically reject proposals of $v_{gjt} = 0$ because a group must visit a site in order for one of its members to be detected. We also automatically reject proposals of $v_{gjt} = -1$ because values below 0 are not in the support of the Poisson group encounter distribution. For all indices, $[v_{gjt}|\lambda_0, \sigma, s_g, \theta, y_{ijt}] \propto$ Poisson$(\lambda_{gjt}) \times \prod_{i \ni \gamma_i = g}$ Poisson $(\theta \times v_{gjt} \times w_i)$

6. Update the group indicator vector $\mathbf{z}$. We sequentially update all members of $\mathbf{z}$ using a Metropolis-Hastings step with a symmetric proposal.

- Case 1: $z_g^{curr} = 1$. If group g is currently included in the population (set to 1), we jointly propose to exclude it (set it to 0) , identify the individuals associated with that group and included in the population, and propose to set their w_i elements to 0. This proposal for both z_g and w_i for $i \ni \gamma_i = g$ is either jointly accepted or rejected. If accepted, the count n_g associated with group g is not changed, but transitions from representing a count of the number of group members included in the population to a fully latent count. Also, because group encounter is fully latent, the group encounter for this group, $v_{g,1:J,1:K}$, also is not changed. The distribution of z_g is $[z_g|\theta, y_{i \ni \gamma_i = g,1:J,1:K}, v_{g,1:J,1:K}, w_{i \ni \gamma_i = g}] \propto$ Bernoulli$(\psi) \times$

$\prod_{i \ni \gamma_i = g} \prod_j \prod_t \text{Poisson}(\theta \times v_{gjt} \times w_i)$. If group g has detected individuals associated with it, it cannot be excluded from the population and this index of g is skipped.

- Case 2: $z_g^{curr} = 0$. If group g is currently not included in the population, we jointly propose to include it, randomly choose n_g individuals associated with it to include, and propose to set their w_i elements to 1. As in Case 1, n_g and $v_{g,1:J,1:K}$ are not changed and the distribution of z_g is the same as in Case 1.

7. Update ψ. We update ψ from the full conditional $[\psi] \propto \text{Beta}(1 + \sum_g z_g, 1 + Z - \sum_g z_g)$, as is typical in standard spatial capture-recapture using data augmentation.

8. Jointly update the individual indicator vector $\mathbf{w}$, and the group size parameter, n_g for the group g they are associated with. We sequentially update all members of $\mathbf{w}$ using a Metropolis-Hastings step with an asymmetric proposal with respect to n_g.

- Case 1: $w_i^{curr} = 1$. If individual l associated with group $\gamma_i = g$ is currently included in the population, we propose to exclude it and subtract 1 from n_g, unless this involves setting $n_g = 0$ in which case we skip this i index (group size must be larger than 0). Then, $[w_i^{curr} | \zeta, \theta, y_{i,1:J,1:K}, v_{g,1:J,1:K}] \propto \text{Poisson}(\zeta) \times \prod_j \prod_t \text{Poisson}(\theta \times v_{gjt} \times w_i)$ for $g \ni \gamma_i = g$. The detection component of this distribution evaluates to 1 when proposing $w_i = 1$. This proposal is asymmetric with respect to n_g with the forward proposal probability being the probability of excluding a w_i associated with this group and thus decrementing n_g by one. This probability is $\sum_{i \gamma_i = g} w_i^{curr} / \sum_{i \ni \gamma_i = g} 1$. The backwards proposal probability is then $\sum_{i \ni \gamma_i = g} w_i^{cand} / \sum_{i \ni \gamma_i = g} 1$

- Case 2: $w_i^{curr} = 0$. If individual i associated with group $\gamma_i = g$ is currently not included in the population, we propose to include it and add 1 to n_g. Then, we proceed exactly as in Case 1, except the detection component of $[w_i^{curr}|\zeta, \theta, y_{i,1:J,1:K}, v_{g,1:J,1:K}]$ does not evaluate to 1 and the forward and backwards proposal probabilities are reversed.

9. Update the group activity centers, **s**. We use a typical SCR Metropolis-Hastings activity center update with symmetric proposal distribution Normal$(s_{g1}^{curr}, \alpha_{s_1})$ and Normal$(s_{g2}^{curr}, \alpha_{s_2})$ for the x and y dimensions, respectively. We propose a candidate s_i, automatically rejecting if the proposal falls outside the state space S. For proposals that fall inside the state space, $[s_g, |v_{g,1:J,1:K}, \lambda_0, \sigma] \propto \prod_j \prod_t$ Poisson $(\lambda_{g,1:J,1:T})$. Note, because group site use is fully latent, its distribution is not conditioned on z_g as it is in typical SCR.

10. Update the number of groups and individuals. These are derived parameters. On each iteration, $G = \sum_g z_g$ and $N = \sum_i w_i$, or equivalently, $N^I = \sum_g n_g \times z_g$. Similarly, the total number of site visits can be derived by summing all elements of **V**.

11. Repeat.

11 Appendix S5: Further data analysis and simulation results for Chapter 4

Data analysis results

Table S1. Full model results for all parameters of cluster SCR, individual-level SCR, and group-level SCR models fit to African wild dog camera data in 2014 and 2015. V is the total number of group encounters, such that $V = \sum_{g,j,k} v_{g,j,k}$.

Parameter	Year	Cluster Mode	Cluster CI	Individual-level Mode	Individual-level CI	Group-level Mode	Group-level CI	Previous Estimate
Group density	2014	0.28	0.15-0.48	-	-	0.17	0.1-0.3	0.81
Group density	2015	0.13	0.09-0.25	-	-	0.12	0.09-0.22	0.70
Average group size (μ)	2014	11.97	9.60-15.33	-	-	11.67*	-	12.32
Average group size (μ)	2015	8.05	5.91-10.65	-	-	7.67*	-	11.89
Individual density	2014	2.75	1.73-6.03	3.79	2.93-4.42	1.95*	1.30-3.25*	2.10

Individual density	2015	1.20	0.71-2.05	1.47	1.13-1.73	0.89*	0.66-1.66*	2.10
λ_0	2014	0.04	0.03-0.05	0.01	0.01-0.02	0.20	0.17-0.26	-
λ_0	2015	0.02	0.01-0.03	0.01	0.01-0.0104	0.07	0.06-0.09	-
σ	2014	4.74	4.12-5.75	4.77	4.46-5.22	5.57	5.11-6.07	-
σ	2015	6.39	5.66-7.44	5.71	5.28-6.24	5.74	5.34-6.19	-
θ	2014	0.29	0.25-0.34	-	-	-	-	-
θ	2015	0.33	0.28-0.39	-	-	-	-	-
ψ	2014	0.19	0.09-0.40	0.65	0.52-0.80	0.09	0.04-0.19	-
ψ	2015	0.14	0.07-0.29	0.32	0.24-0.42	0.09	0.04-0.19	-
V	2014	91.00	84.00-114.00	-	-	-	-	-
V	2015	97.98	90.00-119.00	-	-	-	-	-

Simulation results

Table S2. Bias and coverage of all parameters for cluster SCR and group-level SCR models fit to data simulated from the group-level SCR model. V is the total number of group encounters, such that $V = \sum_{g,j,k} v_{g,j,k}$.

Parameter	True number of groups	True average group size parameter	Cluster SCR Bias	Group-level SCR Bias	Cluster SCR Coverage	Group-level SCR coverage
Number of groups (*G)*	10	5.0	-7.28	-42.65	97.73	29.11
Number of groups (*G*)	10	12.5	-7.77	-42.77	96.94	9.64
Number of groups (*G*)	20	5.0	-1.71	-47.95	96.97	1.08
Number of groups (*G*)	20	12.5	-1.84	-47.74	96.00	0
Average group size (μ)	10	5.0	3.79	-	94.32	-
Average group size (μ)	10	12.5	3.40	-	94.90	-

Average group size (μ)	20	5.0	-1.27	-	95.96	-
Average group size (μ)	20	12.5	-0.55	-	93.00	-
N	10	5.0	-8.81	-	96.59	-
N	10	12.5	-5.33	-	97.96	-
N	20	5.0	-4.78	-	96.97	-
N	20	12.5	-1.84	-	93.00	-
λ_0	10	5.0	5.93	-11.03	93.18	56.96
λ_0	10	12.5	-0.14	91.54	92.86	16.87
λ_0	20	5.0	2.94	-34.58	94.95	23.66
λ_0	20	12.5	4.06	48.59	93.00	21.43
σ	10	5.0	-2.38	91.24	90.91	3.80
σ	10	12.5	0.14	88.96	92.86	2.41
σ	20	5.0	-1.93	116.95	92.93	0
σ	20	12.5	-1.13	115.40	95.00	0
θ	10	5.0	-1.03	-	93.18	-
θ	10	12.5	0.40	-	97.96	-

θ	20	5.0	-4.23	-	91.92	-
θ	20	12.5	-1.97	-	92.00	-
ψ	10	5.0	-1.10	-39.86	100	84.81
ψ	10	12.5	-1.99	-38.55	100	87.95
ψ	20	5.0	-0.36	-46.12	100	27.96
ψ	20	12.5	0.43	-45.26	100	15.48
V	10	5.0	0.23	-	93.18	-
V	10	12.5	-1.46	-	93.88	-
V	20	5.0	1.66	-	91.92	-
V	20	12.5	0.49	-	95.00	-

Table S3. Average individual recaptures (traps/detected individual) and group recaptures (traps/detected group) for all scenarios of the simulation of data from the cluster SCR model. Averages are calculated using all simulated data sets, including those which resulted in models that did not converge.

Number of groups	Average group size parameter	Average individual recaptures	Average group recaptures
10	5.0	2.50	7.23
10	12.5	2.50	10.00
20	5.0	2.46	7.84
20	12.5	2.53	11.04

Table S4. Number of replicate data sets used to create simulation bias and coverage results for simulations from the cluster SCR model. Datasets were excluded due to lack of evidence of convergence (any Gelman-Rubin diagnostics > 1.1) or model failures.

True number of groups	True average group size parameter	Cluster SCR sample size	Group-level SCR sample size
10	5.0	88	79
10	12.5	98	83
20	5.0	99	93
20	12.5	100	84

Table S5. Average individual recaptures (traps/detected individual) and group recaptures (traps/detected group) for all hidden Markov movement simulation scenarios. Averages are calculated using all simulated data sets, including those which resulted in models that did not converge.

Number of groups	Average group size parameter	Attraction to group centroid	Average individual recaptures	Average group recaptures
10	5.0	High	2.96	6.08
10	5.0	Low	2.77	6.26
10	12.5	High	3.02	7.62
10	12.5	Low	2.68	8.35
20	5.0	High	2.99	5.83
20	5.0	Low	2.69	6.12
20	12.5	High	3.03	7.65
20	12.5	Low	2.73	8.33

Table S6. Number of replicate data sets used to create simulation bias and coverage results for hidden Markov movement model simulations. Datasets were excluded due to lack of evidence of convergence (any Gelman-Rubin diagnostics > 1.1) or model failures.

True number of groups	True average group size parameter	Attraction to group centroid	Cluster SCR sample size
10	5.0	High	83
10	5.0	Low	85
10	12.5	High	99
10	12.5	Low	84
20	5.0	High	99
20	5.0	Low	93
20	12.5	High	100
20	12.5	Low	88