Spatial capture–recapture model performance with known small-mammal densities

Brian D. Gerber[1,3] and Robert R. Parmenter[2]

[1]*Colorado Cooperative Fish and Wildlife Research Unit, Department of Fish, Wildlife and Conservation Biology, Colorado State University, Fort Collins, Colorado 80523 USA*
[2]*Valles Caldera National Preserve, P.O. Box 359, 090 Villa Louis Martin, Jemez Springs, New Mexico 87025 USA*

Abstract. Abundance and density of wild animals are important ecological metrics. However, estimating either is fraught with challenges; spatial capture–recapture (SCR) models are a relatively new class of models that attempt to ameliorate common challenges, providing a statistically coherent framework to estimate abundance and density. SCR models are increasingly being used in ecological and conservation studies of mammals worldwide, but have received little testing with empirical field data. We use data collected via a web and grid sampling design to evaluate the basic SCR model where small-mammal abundance (N) and density (D) are known (via exhaustive sampling). We fit the basic SCR model with and without a behavioral effect to 11 small-mammal populations for each sampling design using a Bayesian and likelihood SCR modeling approach. We compare SCR and ad hoc density estimators using frequentist performance measures. We found Bayesian and likelihood SCR estimates of density ($\hat{D}$) and abundance ($\hat{N}$) to be similar. We also found SCR models to have moderately poor frequentist coverage of D and N (45–73%), high deviation from truth (i.e., accuracy; $\hat{D}$, 17–29%; $\hat{N}$, 16–29%), and consistent negative bias across inferential paradigms, sampling designs, and models. With the trapping grid data, the basic SCR model generally performed more poorly than the best ad hoc estimator (behavior CR super-population estimate divided by the full mean maximum distance moved estimate of the effective trapping area), whereas with the trapping web data, the best-performing SCR model (null) was comparable to the best distance model. Relatively poor frequentist SCR coverage resulted from higher precision (SCR coefficients of variation [CVs] < ad hoc CVs); however $\hat{D}$ and D were fairly well correlated (r^2 range of 0.77–0.96). SCR's negative relative bias (i.e., average underestimation of the true density) suggests additional heterogeneity in detection and/or that small mammals maintained asymmetric home ranges. We suggest caution in the use of the basic SCR model when trapping animals in a sampling grid and more generally when small sample sizes necessitate the spatial scale parameter (σ) apply to all individuals. When possible, researchers should consider variation in detection and incorporate individual biological and/or ecological variation at the trap level when modeling σ.

Key words: abundance; accuracy; Bayesian inference; bias; coverage; estimation; evaluation; field test; likelihood; population; spatially explicit capture–recapture models.

INTRODUCTION

Animal abundance and density are primary measures of interest in the study of wild animals. They are critically important in the conservation of threatened and endangered species, are useful in making management decisions of harvested populations, and are often used in the study of ecological relationships (e.g., species interactions). It has long been recognized that obtaining a census (i.e., a total enumeration) of wild animals is rare in ecological studies, necessitating the incorporation of the probability of detection to correct simple animal counts (Nichols 1992). However, dealing with variation in detection probability among individuals and characterizing the effective area sampled to estimate animal

abundance ($\hat{N}$) or density ($\hat{D}$) is challenging for most species.

A common approach for estimating abundance is to set up an array of detectors or traps to mark and subsequently recapture individuals over two or more sampling occasions (e.g., days) and use closed capture–recapture (CR) models (Otis et al. 1978, White et al. 1982). A frequent issue that arises is the lack of geographic closure around the trapping array, such that individuals will move on and off the array throughout the sampling period. This type of movement is especially pervasive with individuals that primarily range beyond the trapping array but may encounter a trap at the periphery. As such, animal abundance estimates using closed CR models reflect a super-population of animals that could have used the trapping array area during the total sampling period (Kendall 1999). It is challenging to then estimate animal density, as the super-population

belongs to an unknown area that extends beyond the area delineated around the trapping array (Dice 1938). Density estimates that are not corrected will be positively biased, with the magnitude generally depending on the ratio of home range to the detector array size (Bondrup-Nielsen 1983).

Researchers and managers often seek to estimate animal density because it is easily scalable and thus comparable across studies. As such, a variety of strategies have been considered to estimate the area of the super-population, often referred to as the effective trapping area (ETA), and subsequently calculate animal density as $\hat{D} = \hat{N}/\widehat{ETA}$, where the $\widehat{var}(\hat{D})$ is approximated by the delta method. Two common ad hoc estimators include adding area around each trap or the minimum convex polygon of all traps based on the mean maximum distance moved of individuals detected (full MMDM [mean maximum distance moved]) and half this value (half-MMDM; Wilson and Anderson 1985). There is, however, no theoretical support of one estimator of the ETA, leading to many estimators with varying reliability (Wilson and Anderson 1985, Parmenter et al. 2003, Balme et al. 2009, Efford and Fewster 2013, Ivan et al. 2013).

A relatively new class of statistical models that provides a unified and statistically coherent framework to estimate animal density and abundance is spatial capture–recapture (SCR) models (Borchers and Efford 2008, Royle et al. 2013a). These models are generally applicable to any species and many different types of detection devices; they have been used to study the ecology and conservation of rare and threatened carnivores (Sharma et al. 2010, Gerber et al. 2012), apes (Moore and Vigilant 2013), whales (Marques et al. 2012), small mammals (Krebs et al. 2011), and much more. SCR models can be defined hierarchically where a spatial process model defines the number of individual activity or home-range centers over an area of interest (an area that includes the trapping array) and an observation model describes the process of how individuals are observed. The spatial model is either assumed to follow a binomial point process, where the number of individuals in the area of interest is finite, or considered to be a realization of a Poisson point process (Efford and Fewster 2013). The observation model depends on the CR sampling method, such as whether individuals can be observed at one or multiple detectors on a single occasion (e.g., live trapping vs. photographic sampling; Efford et al. 2009). The basic SCR model assumes the proximity of a trap to an activity center is a significant source of variability in detections of individuals (Borchers and Efford 2008), where a detection function is used to define how detection probability declines the farther a trap is from an activity center.

A significant advantage of SCR models is that the detection function allows for temporary movement on and off the trapping array, leading to SCR models being robust to variation in trapping array size and spacing relative to animal movement (Sollmann et al. 2012). These models have developed along two parallel lines, where parameter estimation is done by Bayesian inference using Markov chain Monte Carlo (MCMC; Royle et al. 2013b) methods or classical inference by maximizing the integrated likelihood (Borchers and Efford 2008). MCMC methods require higher costs in computation time and user knowledge compared to likelihood maximization (Efford 2011, Noss et al. 2012). However, Bayesian inference has been argued to be more appropriate in the case of small data sets (e.g., few unique individuals detected; Noss et al. 2012, Thompson et al. 2012, Royle et al. 2013b), which is common in CR studies, as likelihood inference predicates on asymptotic conditions that are unlikely to be met with small samples, while Bayesian inference does not depend on this condition.

Several studies have investigated properties of maximum likelihood (ML) and Bayesian SCR models and their performance under a variety of scenarios via simulation (Blanc et al. 2012, Sollmann et al. 2012, Efford and Fewster 2013, Ivan et al. 2013, Royle et al. 2013a, c, Tobler and Powell 2013). In addition, several studies have reported empirical differences between SCR models and ad hoc density estimators (Obbard et al. 2010, Gerber et al. 2011, O'Brien and Kinnaird 2011, Sollmann et al. 2011, Noss et al. 2012, Howe et al. 2013). However, to our knowledge, only a single study has evaluated SCR estimates from empirical data with a known population (Sharma et al. 2010). SCR empirical performance with field data is thus largely untested.

The development of SCR models has added greatly to the field of ecological modeling, but the statistical development has largely outpaced any real-world evaluation of these models using real and messy data. Both simulation and empirical model evaluation are critical in understanding the appropriate use of a statistical model and its expected performance. Simulations are useful to evaluate whether a model behaves as intended under ideal conditions (e.g., unbiased and nominal confidence coverage of truth) and to evaluate the robustness of the model to violation of specific assumptions. However, simulations are inherently limited and cannot consider the breadth of ecological and sampling processes (which are never known in their entirety) that lead to observational data. Empirical studies are required to field test a priori models with common ecological data (e.g., sparse data) to gain an understanding of realistic model performance. When models perform poorly with empirical data, it is challenging to recognize how model assumptions were violated or why the model failed to produce satisfactory results. Ultimately, however, it is accurate and precise estimates from field data that are desired by ecologists, which hinge on the theoretical properties of statistical models and the appropriateness of model assumptions to ecological and sampling processes.

Sampling grids

Complete enumeration

~50 m

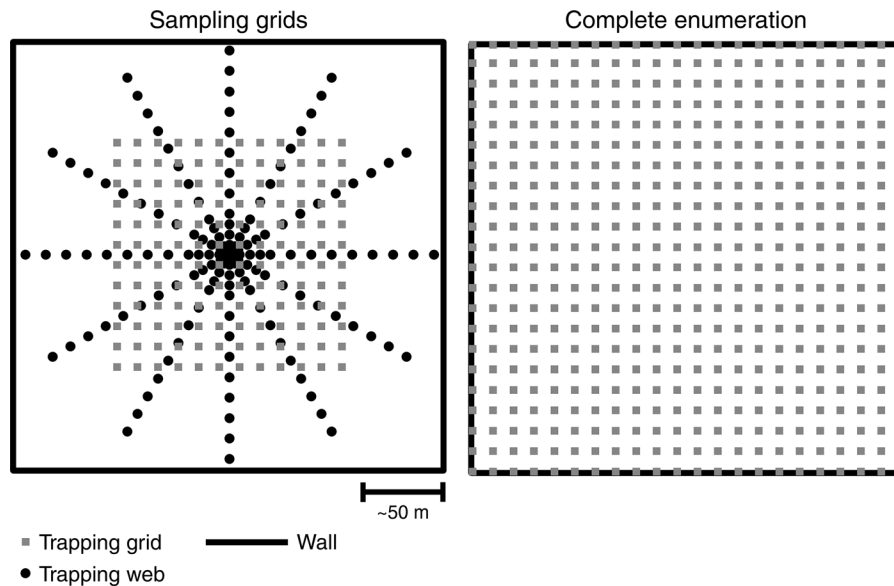■ Trapping grid ▬▬ Wall
● Trapping web

Fig. 1. Sampling designs to evaluate small-mammal density and abundance estimates with known population size using complete enumeration. All sampling occurred in four terrestrial predator-free enclosures of ~4.2 ha each, arranged in a 2 × 2 array (Table 1) over a relatively homogenous shrub-steppe vegetation community in the Sevilleta National Wildlife Refuge, New Mexico, USA. Grid and web layout (left panel) consisted of a trapping grid with 144 traps (gray squares) arranged 12 × 12 with a 10-m inter-trap distance for a trapping area of 1.21 ha. The trapping web was composed of 12 radial lines that extended 100 m from the center; 12 traps (black circles) were placed along each line with the first four at 5-m intervals and the following eight at 10-m intervals. An additional four traps were placed at the center of the web, for a total of 148 traps. The circular area of each trapping web was 3.14 ha. Saturation trapping (right panel) was done to completely enumerate small-mammal populations, following the collection of the test data via grid sampling. Each previously sampled enclosure was trapped for an additional four days, and then until no unmarked individuals were captured (13 July–18 July 1998 and 20 July–24 July 1998); the saturation trapping grid was 22 × 22 with a 10-m inter-trap interval (484 traps [gray squares], ~115 traps/ha).

We use the small-mammal CR data of Parmenter et al. (2003) and empirically evaluate the performance of the basic SCR model by estimating small-mammal abundance and density. Parmenter et al. (2003) describe 11 small-mammal populations trapped using a grid and web sampling design within four replicated enclosures (Fig. 1) where the rodent populations were enumerated (i.e., known abundance [N] and density [D]). They estimated density from the trapping grid data using CR super-population estimates with ETA estimators, while the trapping web data was used in a distance-sampling approach (Buckland et al. 2001). Frequentist performance measures of estimator error were computed for both approaches using true abundance and density values. The number of unique individuals captured per population ranged from five to 46 with the trapping grid layout and from nine to 51 for the trapping web (Appendix A: Table A1). Sample sizes were realistic for small-mammal trapping, but were generally smaller than recommended for CR modeling (White et al. 1982). Our objectives are to compare (1) Bayesian and likelihood SCR modeling approaches to estimating density and abundance from small sample data, (2) performance measures between the grid and web sampling designs, and (3) SCR performance with density estimation performance from Parmenter et al. (2003). Our study

is the first to evaluate the basic SCR model using multiple known animal populations.

Material and Methods

We describe relevant small-mammal trapping procedures, while additional details can be found in Parmenter et al. (2003).

Small-mammal trapping

Trapping occurred in four terrestrial predator-free enclosures of ~4.2 ha each, arranged in a 2 × 2 array (Table 1) over a relatively homogenous shrub-steppe vegetation community in the Sevilleta National Wildlife Refuge, New Mexico, USA (see Plate 1). Enclosure walls were constructed in winter/spring, 1997; construction disturbances were minimal and confined to within 0.5 m of the inside wall. All rodent species within the enclosure were native residents, except for some of the deer mice (*Peromyscus maniculatus*), which had been introduced to the enclosures in summer 1997 as part of an ongoing hantavirus study. Avian predation was possible, but not observed. Sherman live traps (7.62 × 8.89 × 22.86 cm folding traps; H. B. Sherman Traps, Tallahassee, Florida, USA) were used to capture and recapture small mammals; traps were baited with seeds coated in molasses, and were opened each evening and checked in early morning. All individuals were double-tagged

TABLE 1. Small-mammal population characteristics where abundance and density were determined by complete enumeration.

Population number	Species or groups of species	Grid and area (ha)	True abundance	True density (individuals/ha)
1	*Perognathus flavus*	NE (4.28)	88	20.57
2	*P. flavus*	SW (4.13)	39	9.44
3	*P. flavus*	SE (4.19)	58	13.83
4	*P. flavus*	NW (4.25)	81	19.06
5	Cricetines	SW (4.13)	65	15.73
6	Cricetines	SE (4.19)	22	5.24
7	Cricetines	NW (4.25)	33	7.76
8	Cricetines	NE (4.28)	33	7.71
9	*Peromyscus maniculatus*	SW (4.13)	26	6.29
10	*P. maniculatus*	NE (4.28)	14	3.27
11	*Dipodomys* spp.	SE (4.19)	18	4.29

Note: In grid and area column, area is shown in parentheses. NE, NW, SE, and SW refer to the different enclosure areas, referenced by their east–west, north–south position.

with both PIT tags (MUSICC Chip Identification System, Avid Identification Systems, Norco, California, USA) and ear tags (Gey Band and Tag, Norristown, Pennsylvania, USA; except for *Perognathus flavus*, for which unique toe clips were used for identification because their ears were too small for tags). The trapping grid consisted of 144 traps arranged 12×12 with a 10-m inter-trap distance, for a trapping area of 1.21 ha (Fig. 1). The trapping web was composed of 12 radial lines that extended 100 m from the center (Fig. 1); 12 traps were placed along each line with the first four at 5-m intervals and the following eight at 10-m intervals. An additional four traps were placed at the center of the web, for a total of 148 traps. The circular area of each trapping web was 3.14 ha.

Test data were collected by trapping each enclosure for five days. Two enclosures were each trapped using the grid and web layout (28 June–2 July 1998, nearly a year after enclosure construction), and then the trapping layout was switched (5 July–9 July 1998). Following the collection of the test data, saturation trapping was done to completely enumerate small-mammal populations. Each enclosure was trapped for an additional four days, and then until no unmarked individuals were captured (13 July–18 July 1998 and 20 July–24 July 1998); the saturation trapping grid was 22×22 with a 10-m inter-trap interval (Fig. 1; 484 traps, ~115 traps/ha). Parmenter et al. (2003) identified 11 populations by species or groups of species (Heteromyidae [*Perognathus* and *Dipodomys*], Cricetidae [*Peromyscus*]); groups of species were aggregated to a population due to small sample size and taxonomic similarity (e.g., *Peromyscus* spp.). True abundance (N) was calculated as the number of unique individuals by species or group of species caught per enclosure and true density (D) as abundance divided by the exact area of the enclosure (Table 1).

Population estimation

We estimated small-mammal density ($\hat{D}$) and abundance ($\hat{N}$) by population using two basic SCR models;

the first model only considers space use to affect detection (null), while the second model also includes a behavioral effect (behavior; e.g., trap-happy or trap-shy). It is expected that when animals are trapped with a food reward, there will likely be a behavioral response; however, due to small sample size the behavioral effect may be undetectable and a null model more appropriate (Otis et al. 1978). We fit both models to each population's test data collected from the trapping grid and web using available Bayesian and likelihood methods.

The Bayesian SCR model was first defined with a state-space (S) with the same dimensions as each enclosure. The number of activity centers ($s_i = s_{xi},\ s_{yi}$) were finite and uniformly distributed within S ($s_i \sim$ Uniform[S]) for individual $i = 1, 2, \ldots, N$. We defined observations by the encounter histories y_{ijk} for individual $i = 1, 2, \ldots, N$, trap $j = 1, 2, \ldots, J$, and occasion $k = 1, 2, \ldots, K$. Each trap is indexed by the coordinate x_j and we restricted individuals to only be allowed to be captured at a single trap per occasion; the encounter history $y_{ijk} \sim$ Categorical(π_{ijk}) indicates the trap each individual was encountered at from $1, 2, \ldots, J$, or $J + 1$ if it was not encountered (Royle et al. 2013b). For the null model, we defined variation in detection using the multinomial-logit function (mlogit), such that mlogit(π_{ijk}) $= \alpha_0 - d_{ij}/2\sigma^2$, where d_{ij} is the squared Euclidian distance between s_i and x_j ($d_{ij} = ||s_i - x_j||^2$). The parameter α_0 represents the probability of detecting an individual when a trap is located at the activity center and σ represents how quickly detection probability declines as the distance between activity center and trap increases. This detection function is equivalent to assuming individuals have a strict circular bivariate normal home range. The categorical observation model is technically misspecified for single-animal traps, as it allows multiple individuals to be caught in a single trap. Trap competition among individuals with single-animal traps violates the independence assumption of the observational model and creates dependency among

the capture histories; this issue has yet to be formally resolved (Royle et al. 2013b). However, we expect density estimation bias to be negligible, given that mean trap saturation (the proportion of traps occupied across occasions) was much lower (<0.31; Appendix A: Table A2) than needed to incur noticeable bias (Efford et al. 2009). The behavior model is the same as the null model, but also includes an individual-level, non-trap-specific binary covariate (behav) in the detection function, where $behav_{ik} = 1$ if individual i was captured prior to occasion k and 0 if not. We assume that after an individual is initially trapped, it will respond similarly to all traps on the grid as a learned response of recognizing the shape and material of the traps and the reward within.

As is common to fit Bayesian SCR models, we augment the test data with all zero capture histories from n to M, where n is the number of unique individuals detected and M is an arbitrarily large value that is much greater than N (e.g., Royle and Dorazio 2012). Further, we define a set of latent variables, $z_i \sim$ Bernoulli(ψ) from $i = 1, 2, \ldots, M$, where $z_i = 1$ indicates individual i is a member of the population N, or otherwise is not. We chose relatively uninformative prior distributions for unknown parameters: $\alpha_0 \sim$ Normal(0, 200), $\sigma \sim$ Uniform(0, 200), including the behavioral effect parameter as $\beta \sim$ Normal(0, 200). Instead of using the common discrete uniform prior on N (implied by $\psi \sim$ Beta(1,1)), which can lead to improper posterior distributions, we used the scale prior ($\psi \sim$ Beta(0.000001,1); Link 2013). The null SCR model using data augmentation is completely specified in Appendix B: Fig. B1.

We fit models using the R package rjags, which interfaces with software JAGS (Plummer 2013), in which MCMC is used to simulate samples from the full conditional distributions of unknown parameters of our statistical model (see Supplement). We initialized three chains with over-dispersed starting values that were each run for 60 000 iterations, discarding the first 10 000 as a burn-in, and computed posterior summaries from the remaining 50 000 samples. We assessed convergence graphically and using the Gelman-Rubin statistic, $\hat{R}$ (Gelman and Rubin 1992); an $\hat{R}$ below 1.1 indicates convergence (Gelman 2007). For all models and all test data, $\hat{R} \leq 1.02$. To summarize posterior distributions, we used the posterior mode (i.e., the most probable value) as our measure of central tendency and calculated 95% highest posterior density (HPD) intervals, because we expect asymmetric posteriors because of small sample sizes (Chen et al. 2000). The posterior mode has been shown to be unbiased in SCR simulation studies with small samples and is favored over the positively biased posterior mean (Marques et al. 2012, Thompson et al. 2012), especially when interested in frequentist performance (Kery et al. 2011). When the posterior distribution is symmetric and unimodal, the mode is equal to the mean and HPD intervals are equal

to the more commonly used equal-tailed credible intervals (Chen et al. 2000, Wilks 2011).

We estimated small-mammal population size similarly using the SCR likelihood approach available in the R package secr (Efford 2013). We considered the population size as a finite number of individual activity centers, rather than a realization from a Poisson process, which is common in likelihood SCR modeling. The observation process is defined differently than the Bayesian model, where an additive hazard model is used to restrict individuals to a single trap per occasion (Borchers and Efford 2008); the half-normal detection function is the same. As with the Bayesian observation model, there is still competition among individuals for traps, such that the model is still misspecified. We fit the null and behavior models to each population by maximizing the unconditional integrated likelihood using the Nelder-Mead optimization algorithm (Nelder and Mead 1965); secr uses a discretized state-space, where the likelihood function was evaluated over a grid the size of each enclosure ($A[S]$) with equally spaced points at 1-m intervals, representing potential activity centers. We report the maximum likelihood estimate (MLE) of D and 95% asymmetric confidence intervals. While D is explicitly in the likelihood, abundance is derived as $\hat{N} = A(S)\hat{D}$ and is referred to as fixed-N (Efford and Fewster 2013). We estimated the variance of fixed-N as $\widehat{var}(\hat{N}) = A(S)^2 \widehat{var}(\hat{D})$ and calculated lognormal confidence intervals, as they have performed well in simulation and naturally satisfy that the lower interval is greater than n (Efford and Fewster 2013).

SCR model performance

To evaluate model performance, we computed five attributes of frequentist error using the estimates from each population and the true values (defined by exhaustive sampling), following Parmenter et al. (2003): (1) Confidence interval coverage was calculated as the percentage of times the true value (x) lay within the estimated 95% lower and upper confidence or 95% HPD limits for each population. Under many repeat surveys, we would hope for 95% coverage, but for our 11 populations, coverage could be considered good when it was relatively high, such as >90 % (10/11 or 11/11). (2) Mean relative mean squared error ($\overline{\text{RMSE}}$) was computed as

$$\overline{\text{RMSE}} = \frac{1}{11} \sum_{i=1}^{11} \frac{(\hat{x}_i - x_i)^2}{x_i} \times 100$$

where $i = 1, 2, \ldots, 11$ estimates. Smaller values indicate a lower bias, variance, or both. (3) Mean relative accuracy ($\overline{\text{RA}}$) was computed as

$$\overline{\text{RA}} = \frac{1}{11} \sum_{i=1}^{11} \frac{|\hat{x}_i - x_i|}{x_i} \times 100$$

where $i = 1, 2, \ldots, 11$ estimates. This is a measure of absolute deviation of estimates from truth across the
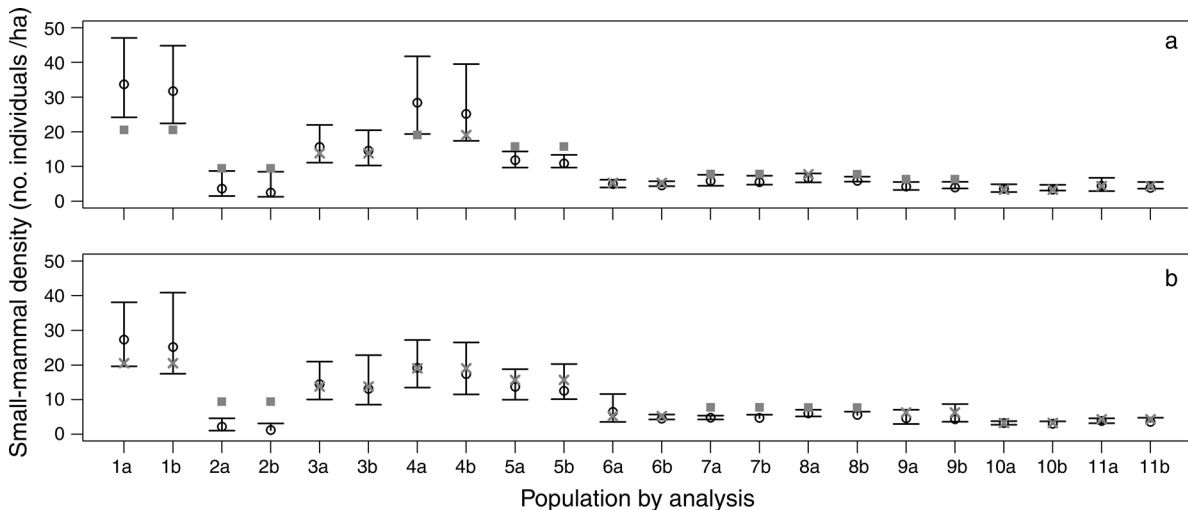
Fɪɢ. 2. Small-mammal density estimates (open circles) with 95% highest posterior density (HPD) intervals from two spatial capture–recapture (SCR) models where traps were arranged in a grid: (a) null SCR, (b) behavior effect SCR. The *x*-axis indicates the population (1–11; see Table 1) and whether the estimate is from a maximum-likelihood (designated with a) or Bayesian analysis (designated with b). The true density is indicated as an "x" when included in the HPD interval and a square when not.

populations; smaller values indicate estimates were closer to truth. (4) Mean relative precision or mean coefficient of variation ($\overline{\text{CV}}$), was computed as

$$\overline{\text{CV}} = \frac{1}{11} \sum_{i=1}^{11} \frac{\text{SE}(\hat{x}_i)}{\hat{x}_i} \times 100$$

where $i = 1, 2, \ldots, 11$ estimates. This is a measure of uncertainty around the point estimate. (5) Mean relative bias ($\overline{\text{RB}}$) was computed as

$$\overline{\text{RB}} = \frac{1}{11} \sum_{i=1}^{11} \frac{\hat{x}_i - x_i}{x_i} \times 100$$

where $i = 1, 2, \ldots, 11$ estimates. Values close to zero indicate better model performance across the populations. Lastly, we also evaluated the accuracy of each models point estimate using linear regression analyses of $\hat{D}$ vs. D; the most accurate model would have a slope nearest to 1 and the highest coefficient of determination (r^2).

Ad hoc density estimators

Parmenter et al. (2003) estimated small-mammal density from the trapping grid data using four estimators of the ETA with 10 closed CR models. The trapping web data were analyzed using 12 distance-sampling models. We compare SCR estimates with the best-overall-performing ETA estimators (half-MMDM and full MMDM, calculated using individuals detected at >1 trap locations) using capture–recapture models that had similar a priori assumptions as SCR models on detection: constant (M_0), heterogeneity (M_h), behavior (M_b), and behavior with heterogeneity (M_{bh}; Otis et al. 1978). We also compare SCR estimates to the best-

performing distance model with the trapping web data, where the detection function is uniform-cosine, and with a model using the half-normal function due to its similarity to the basic SCR model.

Rᴇsᴜʟᴛs

Comparison of Bayesian and ML SCR modeling

We found that $\hat{D}$ and $\hat{N}$ from the ML and Bayesian SCR modeling approaches were similar overall for a given model and sampling design (Figs. 2 and 3, Appendix B: Tables B1–B4). However, the posterior mode was mostly smaller than the ML SCR estimates. As predicted, all posterior distributions of population size were either lightly or heavily positively skewed (e.g., Appendix B: Fig. B2). We found the frequentist coverage of the true value of population size (D and N) was similar and below the nominal level, regardless of sampling design, inferential paradigm, and model (45–73%; Table 2). The likelihood modeling approach consistently had lower $\overline{\text{CV}}$ and $\overline{\text{RB}}$. However, $\overline{\text{RA}}$ was mostly similar between the two approaches, but the r^2 of the point estimates was consistently slightly better under the Bayesian paradigm (Appendix B: Fig. B3). In other words, the likelihood approach was more precise and had a better balance of estimates that were both larger and smaller than the true values, but both methods had a similar level of overall inaccuracy. When truth was not covered by density or abundance intervals, point and interval estimates tended to be lower than truth.

Comparing SCR performance between the trapping web and grid

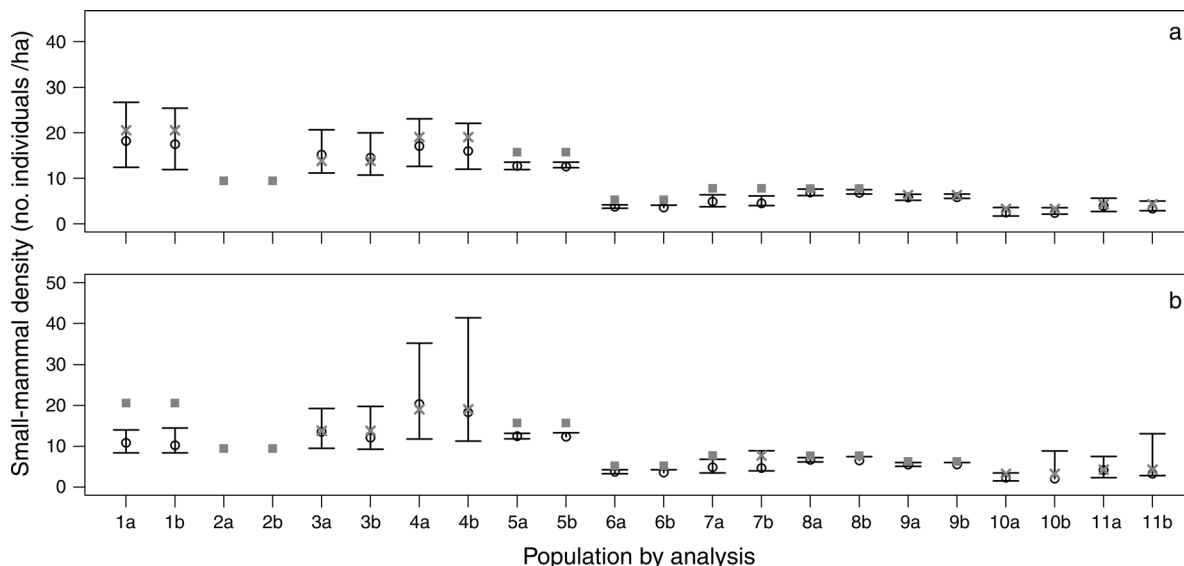Neither data from the trapping grid or web were consistently better in terms of number of unique

Fig. 3. Small-mammal density estimates (open circles) with 95% HPD intervals from two SCR models where traps were arranged in a web: (a) null SCR, (b) behavior effect SCR. The *x*-axis indicates the population (1–11; see Table 1) and whether the estimate is from a maximum-likelihood (designated with a) or Bayesian analysis (designated with b). The true density is indicated as an "x" when included in the interval and a square when not.

individuals detected or frequency of recaptures (Appendix A: Table A1). However, there were distinct differences between estimates from the sampling designs; the null model performed better with the trapping web, while the behavior model performed better with the trapping grid, except in terms of $\overline{RB}$. Overall, the trapping web had a small $\overline{RMSE}$ and $\overline{RA}$, but a higher $\overline{RB}$ (Table 2, Appendix B: Fig. B3). The trapping web data under the null model had a slope closest to 1.0 (ML had 0.933 and Bayesian had 0.905) and the highest r^2 (ML had 0.956 and Bayesian had 0.964); the most inaccurate estimates were at larger

densities, which got worse using the behavior model (Appendix B: Fig. B3).

Performance of SCR models compared to ad hoc estimators

For the trapping grid, we found SCR model performance was worse than the best-performing ad hoc estimator (M_b and M_{bh} super-population estimates with the full MMDM estimator). However, for the trapping web data, the best-performing SCR model (null) had comparable performance with the best-performing distance model (uniform-cosine), except in

TABLE 2. Model performance attributes of the trapping grid and web density ($\hat{D}$) and abundance ($\hat{N}$) estimation using a null spatial capture–recapture (SCR) model and a SCR model with a behavioral effect.

Model and summary statistics	ML SCR estimation				Bayesian SCR estimation			
	Trapping grid		Trapping web		Trapping grid		Trapping web	
	$\hat{D}$	$\hat{N}$	$\hat{D}$	$\hat{N}$	$\hat{D}$	$\hat{N}$	$\hat{D}$	$\hat{N}$
Null SCR								
CI coverage (%)	45.45	45.45	60.00	60.00	45.45	45.45	60.00	60.00
$\overline{RMSE}$ (%)	176.82	724.06	30.70	121.26	154.94	653.16	41.66	175.99
$\overline{RA}$ (%)	27.74	27.29	17.01	16.12	28.59	28.61	20.08	20.08
$\overline{CV}$ (%)	18.25	18.01	12.28	10.93	23.38	23.10	13.03	13.03
$\overline{RB}$ (%)	−2.85	−2.06	−15.03	−14.21	−11.94	−11.92	−19.07	−19.05
Behavior SCR								
CI coverage (%)	72.73	72.73	40.00	40.00	72.73	72.73	50.00	50.00
$\overline{RMSE}$ (%)	94.70	379.38	72.91	300.87	105.88	441.49	87.34	370.98
$\overline{RA}$ (%)	22.75	22.27	19.68	18.80	25.33	25.33	24.33	24.31
$\overline{CV}$ (%)	17.72	17.49	14.54	14.31	22.40	22.40	42.00	42.00
$\overline{RB}$ (%)	−11.49	−11.13	−18.28	−17.38	−21.22	−21.20	−24.33	−24.31

Note: Models are fitted using a maximum likelihood (ML) and a Bayesian approach. Summary statistics are mean relative mean squared error ($\overline{RMSE}$), accuracy ($\overline{RA}$), precision ($\overline{CV}$), and bias ($\overline{RB}$).

TABLE 3. Density estimator summary attributes from Parmenter et al. (2003).

| | Trapping grid | | | | | | | | Trapping web distance estimation | |
| | (M_0) | | (M_h) | | (M_b) | | (M_{bh}) | | | |
Summary statistic	Half-MMDM	Full MMDM	Half-MMDM	Full MMDM	Half-MMDM	Full MMDM	Half-MMDM	Full MMDM	Uniform cosine	Half-normal cosine
CI coverage (%)	80.0	100.0	70.0	90.0	90.0	100.0	90.0	100.0	100.0	100.0
$\overline{RMSE}$ (%)	500.5	165.0	1457.1	531.6	206.1	40.5	206.1	40.5	37.2	834.7
$\overline{RA}$ (%)	44.2	26.3	80.6	38.6	38.2	16.4	38.2	16.4	16.9	73.4
$\overline{CV}$ (%)	24.0	35.0	26.8	37.0	24.9	35.6	24.9	35.6	24.5	48.1
$\overline{RB}$ (%)	44.2	−3.3	80.6	22.3	38.2	−8.5	38.2	−8.5	−6.5	67.3

Notes: Model notation for traditional capture–recapture models: constant detection probability (M_0), heterogeneous detection probability (M_h), behavioral effect on detection probability (M_b), and behavior and heterogeneity effect on detection probability (M_{bh}). Summary statistic abbreviations are as in Table 1. For comparison with spatial capture–recapture models, we selected density estimators that consider similar variation in detection and performed relatively well. MMDM stands for mean maximum distance moved of individuals studied.

terms of coverage and $\overline{RB}$, which were worse (Tables 2 and 3). Coverage was worse in all SCR analyses compared with the ad hoc density estimators and distance-sampling models (Tables 2 and 3). We did find that SCR models had a lower $\overline{RMSE}$, $\overline{RA}$, and $\overline{RB}$ compared to the half-MMDM estimator and the half-normal distance-sampling model. SCR model $\overline{CV}$ was similar or smaller than ad hoc estimators and distance-sampling models, but the smaller variance estimates led to poorer confidence interval coverage of the true densities. Point estimates for the trapping web data under the SCR null model were as accurate as the best-performing ad hoc density estimators and distance-sampling models (Appendix B: Fig. B3; Parmenter et al. 2003).

DISCUSSION

SCR models have advanced since the development of the basic model and no longer have to assume individuals distribute their activity according to a circular home range. Two such models incorporate landscape attributes through least-cost-path and resource selection (Royle et al. 2013*a*, *c*). However, many CR studies will commonly preclude the use of these more complex models due to relatively few individuals detected and few spatial locations of individuals. These models also require that landscape variables important to animal movement can be identified and measured, which can be difficult; this is especially true when sampling occurs in relatively homogenous vegetation, such as in this study. We believe the basic SCR model will likely continue to be used more often than the advanced models, as well as preferred over ad hoc density estimators, given the appeal of SCR models' coherent statistical framework. However, a certain amount of caution should be considered based on our analyses, given that the basic SCR model was often overly precise, leading to only moderate frequentist coverage, and on average underestimated density. This is especially problematic given that the extent of sampling, the number of traps deployed, and sampling duration, coupled with double-tagging of individuals in

this study represents a best case scenario; small-mammal capture–recapture studies commonly use less sampling effort (fewer traps, smaller grids, shorter sampling time).

Given the empirical nature of this study, it is difficult to determine how or if SCR model assumptions were unmet and thus why SCR performance was sometimes underwhelming. However, we offer some insights about possible mismatches between SCR model assumptions and basic animal ecology and published SCR simulations. First, it is unlikely that the trapping layout or number of traps were a limiting factor, as trap density was high, the sampling layout proved to be larger than the extent of individual movement, and inter-trap distances were short enough to detect individuals at multiple traps. It is possible that two of the core assumptions of the basic model may be overly simplifying; (1) that individual activity centers are randomly distributed (i.e., independent) and (2) animal space-use can be approximated with a simple detection function where the spatial scalar (σ) is assumed to be the same for all individuals. Individuals are unlikely to be independent of conspecifics due to territoriality or sociality. For small mammals that share burrows or nests, their activity centers may be completely dependent. In addition, animal movement is a highly complex process (Nathan et al. 2008); the necessary simplification of this process due to limited spatial information leads to an unrealistic assumption of a stationary circular home range (White and Garrott 1990). Interspecific interactions may also cause individuals to exhibit very different patterns of space use, depending on their proximity and the effect of these interactions (e.g., attraction or repulsion). Another possibility is that the enclosure walls may unnaturally impede dispersing individuals and inflate density along the enclosure edges. If so, we would expect to observe a peak of individuals in the outermost ring of the sampling web, which was not the case. Similarly, the outermost grid traps were 45 m from the enclosure wall, and the means of full MMDM ranged from 16.18 to 54.54 m, indicating that the effects of enclosure walls at this distance were likely minimal. Lastly, because we often found point and interval

estimates underestimated density and abundance, this suggests additional heterogeneity in detection that was not accommodated. We attempted to address this heterogeneity through a mixture-modeling approach (Pledger 2000), but limited sample size led to generally unreliable results, and models were often unable to converge. Empirical studies have suggested that un-modeled variation in SCR models can lead to overly precise and biased estimates (Howe et al. 2013), as seen in our results; individual biological variation in σ, such as due to sex differences in ranging patterns, is likely to negatively bias density estimates (Tobler and Powell 2013). The basic SCR model estimates also have been shown to be negatively biased when not accounting for movement due to resource use on the landscape (Royle et al. 2013a), as well as when home ranges are asymmetric (Ivan et al. 2013). Heterogeneity in detection among individuals will continue to be a challenge in population estimation because of the many influencing factors. Incorporating the effects of space use is logical, but not the only and perhaps sometimes not the dominating factor influencing detection; the best re-course is to remove the effects of heterogeneity by detecting most of the population many times (Gerber et al. 2014).

Compared to passive detectors (e.g., camera traps), live-trapping will commonly lead to less spatial infor-mation about individual movement due to trapped individuals being confined. This is especially true for this study, in which the sampling duration was much shorter compared to studies of larger mammals sampled using passive detectors (Gerber et al. 2012). The potential for more spatial information from alternative trapping methods may lead to more reliable estimates than found in this study. However, modeling additional individual heterogeneity beyond that accommodated by SCR models, such as using finite mixtures, will also require a reasonable number of unique individuals to be detected; this also will be affected by sampling methods.

We found that data from the trapping web configu-ration led to more accuracy but consistently negative bias, whereas the trapping grid data led to reduced accuracy but also reduced relative bias. With regard to relative accuracy, the null SCR model performed similarly to the best-performing distance-sampling model. Using the trapping grid data, the SCR models performed better than using half-MMDM, but worse than using full MMDM. Among the comparative empirical studies, researchers have generally found SCR models produce lower estimates than CR models with ETA estimators, of which the SCR estimates are taken as superior (Obbard et al. 2010, Gerber et al. 2011, Noss et al. 2012), perhaps due to a higher risk of making poor conservation/management decisions when overes-timating compared to underestimating population size. However, our findings do not support this claim. We found SCR models are on average lower than true values, but not as bad as the higher estimates of half-

PLATE 1.   (Top) the enclosure fence, (middle) the trapping web center from the north, and (bottom) two researchers processing small mammals. Photo credits: R. R. Parmenter.

MMDM, while full MMDM was on average only slightly lower than the true values. Full MMDM has often been evaluated in simulation studies, producing mixed results, and should not be considered generally applicable to all study designs. In comparing a known tiger population to density estimates, Sharma et al. (2010) found the full MMDM and the ML SCR density estimates to be comparable and close to the true value. In a simulation study of jaguar density estimates, full MMDM was found to be similar to SCR estimates across a wide range of trapping array sizes (Tobler and Powell 2013). In contrast, another simulation study that examined alternative space-use patterns by individuals

found SCR estimates to generally outperform the full MMDM estimator (Ivan et al. 2013).

Moving forward, we suggest two developments are needed. First, more SCR simulations are needed to evaluate model assumptions explicitly; future simulation work should aim to explore realistic animal movement as the basis for generating data when applied to SCR models (e.g., Ivan et al. 2013) and consider dependency among individuals' movements. Second, development of formal goodness-of-fit (GOF) tests are needed. There has been some preliminary work in this direction (Borchers and Efford 2008, Royle et al. 2013b), but no assessment as to their expected performance. Reliable GOF tests are especially needed for small sample sizes, in which low-quality data may often lead to a failure to detect a departure from model assumptions.

The parallel development of likelihood and Bayesian SCR models has led many ecologists to analyze their data using both inferential paradigms, without clear reasoning. From a practical perspective, the similarity of estimates and performance between the ML and Bayesian approaches does not support the costs and challenges to practitioners not familiar with Bayesian methods in fitting the basic Bayesian SCR model (Efford 2011). There is of course the philosophical difference between the paradigms, which is independent of the practical argument and will depend solely on the practitioner's inferential preference. Bayesian inference produces valid posterior distributions for any sample size and provides a complete summary of the state of belief about unknown quantities, which is why some have suggested Bayesian SCR models are more valid for small sample data than likelihood models (Noss et al. 2012, Thompson et al. 2012, Royle et al. 2013b). However, it is important for ecologists less familiar with Bayesian methods to be aware that Bayesian inference is conditional on the available data, priors, and the model fitted, as well as assuming proper posterior convergence has been reached. Bayesian models do not guarantee frequentist coverage of truth or unbiasedness. Ecologists interested in Bayesian model fitting should be aware of convergence assessment, especially when dealing with small sample sizes and relatively uninformative priors; even still, improper posterior distributions are of concern and often difficult to assess (Link 2013). If researchers are interested in point and interval estimates, our results, along with simulation support of small sample size (Efford 2011), suggest the asymptotic argument against likelihood methods for SCR is moot.

LITERATURE CITED

Balme, G. A., L. T. B. Hunter, and R. O. B. Slotow. 2009. Evaluating methods for counting cryptic carnivores. Journal of Wildlife Management 73:433–441.

Blanc, L., E. Marboutin, S. Gatti, and O. Gimenez. 2012. Abundance of rare and elusive species: empirical investigation of closed versus spatially explicit capture–recapture models with lynx as a case study. Journal of Wildlife Management 77:372–378.

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

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

Buckland, S. T., D. R. Anderson, K. P. Burnham, and J. L. Laake. 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford, UK.

Chen, M. H., Q. M. Shao, and J. G. Ibrahim. 2000. Monte Carlo methods in Bayesian computation. Springer, New York, New York, USA.

Dice, L. R. 1938. Some census methods for mammals. Journal of Wildlife Management 2:119–130.

Efford, M. 2013. secr: spatially explicit capture–recapture models. R package version 2.6.1. http://cran.r-project.org/web/packages/secr/index.html

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

Efford, M. G., D. L. Borchers, and A. E. Byrom. 2009. Density estimation by spatially explicit capture–recapture: likelihood-based methods. Pages 255–269 in D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. Modeling demographic processes in marked populations. Springer, New York, New York, USA.

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

Gelman, A. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, Cambridge, UK.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457–472.

Gerber, B., J. Ivan, and K. Burnham. 2014. Estimating the abundance of rare and elusive carnivores from photographic-sampling data when the population size is very small. Population Ecology 56:463–470.

Gerber, B., S. Karpanty, and M. Kelly. 2011. Evaluating the potential biases in carnivore capture–recapture studies associated with the use of lure and varying density estimation techniques using photographic-sampling data of the Malagasy civet. Population Ecology 54:1–12.

Gerber, B. D., S. M. Karpanty, and J. Randrianantenaina. 2012. The impact of forest logging and fragmentation on carnivore species composition, density and occupancy in Madagascar's rainforests. Oryx 46:414–422.

Howe, E. J., M. E. Obbard, and C. J. Kyle. 2013. Combining data from 43 standardized surveys to estimate densities of female American black bears by spatially explicit capture–recapture. Population Ecology 55:595–607.

Ivan, J., G. C. White, and T. M. Shenk. 2013. Using simulation to compare methods for estimating density from capture–recapture data. Ecology 94:817–826.

Kendall, W. L. 1999. Robustness of closed capture–recapture methods to violations of the closure assumption. Ecology 80:2517–2525.

Kery, M., B. Gardner, T. Stoeckle, D. Weber, and J. A. Royle. 2011. Use of spatial capture-recapture modeling and DNA

data to estimate densities of elusive animals. Conservation Biology 25:356–364.

Krebs, C. J., D. Reid, A. J. Kenney, and S. Gilbert. 2011. Fluctuations in lemming populations in north Yukon, Canada, 2007–2010. Canadian Journal of Zoology 89:297–306.

Link, W. A. 2013. A cautionary note on the discrete uniform prior for the binomial N. Ecology 94:2173–2179.

Marques, T. A., L. Thomas, S. W. Martin, D. K. Mellinger, S. Jarvis, R. P. Morrissey, C. A. Ciminello, and N. DiMarzio. 2012. Spatially explicit capture–recapture methods to estimate minke whale density from data collected at bottom-mounted hydrophones. Journal of Ornithology 152:445–455.

Moore, D. L., and L. Vigilant. 2013. A population estimate of chimpanzees (*Pan troglodytes schweinfurthii*) in the Ugalla region using standard and spatially explicit genetic capture–recapture methods. American Journal of Primatology 76:335–346.

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

Nelder, J. A., and R. Mead. 1965. A simplex method for function minimization. Computer Journal 7:308–313.

Nichols, J. D. 1992. Capture-recapture models. BioScience 42:94–102.

Noss, A., B. Gardner, L. Maffei, E. Cuéllar, R. Montaño, A. Romero-Muñoz, R. Sollman, and A. O'Connell. 2012. Comparison of density estimation methods for mammal populations with camera traps in the Kaa-Iya del Gran Chaco landscape. Animal Conservation 15:527–535.

Obbard, M. E., E. J. Howe, and C. J. Kyle. 2010. Empirical comparison of density estimators for large carnivores. Journal of Applied Ecology 47:76–84.

O'Brien, T. G., and M. F. Kinnaird. 2011. Density estimation of sympatric carnivores using spatially explicit capture–recapture methods and standard trapping grid. Ecological Applications 21:2908–2916.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs 62:3–135.

Parmenter, R. R., et al. 2003. Small-mammal density estimation: a field comparison of grid-based vs. web-based density estimators. Ecological Monographs 73:1–26.

Pledger, S. 2000. Unified maximum likelihood estimates for closed capture–recapture models using mixtures. Biometrics 56:434–442.

Plummer, M. 2013. rjags: Bayesian graphical models using mcmc. R package version 3-11. http://cran.r-project.org/web/packages/rjags/index.html

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

Royle, J. A., R. B. Chandler, R. Sollmann, and B. Gardner. 2013b. Spatial capture-recapture. Academic Press, Salt Lake City, Utah, USA.

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

Royle, J. A., and R. M. Dorazio. 2012. Parameter-expanded data augmentation for Bayesian analysis of capture–recapture models. Journal of Ornithology 152:521–537.

Sharma, R., Y. Jhala, Q. Qureshi, J. Vattakaven, R. Gopal, and K. Nayak. 2010. Evaluating capture–recapture population and density estimation of tigers in a population with known parameters. Animal Conservation 13:94–103.

Sollmann, R., M. M. Furtado, B. Gardner, H. Hofer, A. T. Jácomo, N. M. Tôrres, and L. Silveira. 2011. Improving density estimates for elusive carnivores: accounting for sex-specific detection and movements using spatial capture–recapture models for jaguars in central Brazil. Biological Conservation 144:1017–1024.

Sollmann, R., B. Gardner, and J. L. Belant. 2012. How does spatial study design influence density estimates from spatial capture-recapture models? PLoS ONE 7:e34575.

Thompson, C. M., J. A. Royle, and J. D. Garner. 2012. A framework for inference about carnivore density from unstructured spatial sampling of scat using detector dogs. Journal of Wildlife Management 76:863–871.

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

White, G. C., D. R. Anderson, K. P. Burnham, and D. L. Otis. 1982. Capture–recapture and removal methods for sampling closed populations. Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

White, G. C., and R. A. Garrott. 1990. Analysis of wildlife radio-tracking data. Academic Press, New York, New York, USA.

Wilks, D. S. 2011. Statistical methods in the atmospheric sciences. Academic Press, Waltham, Massachusetts, USA.

Wilson, K. R., and D. R. Anderson. 1985. Evaluation of two density estimators of small mammal population size. Journal of Mammalogy 66:13–21.

SUPPLEMENTAL MATERIAL

Ecological Archives

Appendices A and B and the Supplement are available online: http://dx.doi.org/10.1890/14-0960.1.sm