



Spatial capture–recapture for categorically marked populations with an application to genetic capture–recapture

BEN C. AUGUSTINE,^{1,†} J. ANDREW ROYLE,² SEAN M. MURPHY¹ ,³ RICHARD B. CHANDLER,⁴ JOHN J. COX,³ AND MARCELLA J. KELLY⁵

¹Atkinson Center for a Sustainable Future and Department of Natural Resources, Cornell University, Ithaca, New York 14843 USA

²Patuxent Wildlife Research Center, U.S. Geological Survey, Laurel, Maryland 20708 USA

³Department of Forestry, University of Kentucky, Lexington, Kentucky 40546 USA

⁴Department of Forestry and Natural Resources, University of Georgia, Athens, Georgia 30602 USA

⁵Department of Fish and Wildlife Conservation, Virginia Tech, Blacksburg, Virginia 24061 USA

Citation: Augustine, B. C., J. A. Royle, S. M. Murphy, R. B. Chandler, J. J. Cox, and M. J. Kelly. 2019. Spatial capture–recapture for categorically marked populations with an application to genetic capture–recapture. *Ecosphere* 10(4): e02627. 10.1002/ecs2.2627

Abstract. Recently introduced unmarked spatial capture–recapture (SCR), spatial mark–resight (SMR), and 2-flank spatial partial identity models (SPIMs) extend the domain of SCR to populations or observation systems that do not always allow for individual identity to be determined with certainty. For example, some species do not have natural marks that can reliably produce individual identities from photographs, and some methods of observation produce partial identity samples as is the case with remote cameras that sometimes produce single-flank photographs. Unmarked SCR, SMR, and SPIM share the feature that they probabilistically resolve the uncertainty in individual identity using the spatial location where samples were collected. Spatial location is informative of individual identity in spatially structured populations because a sample is more likely to have been produced by an individual living near the trap where it was recorded than an individual living further away from the trap. Further, the level of information about individual identity that a spatial location contains is related to two key ecological concepts, population density and home range size, which we quantify using a proposed Identity Diversity Index (IDI). We show that latent and partial identity SCR models produce imprecise and biased density estimates in many high IDI scenarios when data are sparse. We then extend the unmarked SCR model to incorporate categorical, partially identifying covariates, which reduce the level of uncertainty in individual identity, increasing the reliability and precision of density estimates, and allowing reliable density estimation in scenarios with higher IDI values and with more sparse data. We illustrate the performance of this “categorical SPIM” via simulations and by applying it to a black bear data set using microsatellite loci as categorical covariates, where we reproduce the full data set estimates with only slightly less precision using fewer loci than necessary for confident individual identification. We then discuss how the categorical SPIM can be applied to other wildlife sampling scenarios such as remote camera surveys, where natural or researcher-applied partial marks can be observed in photographs. Finally, we discuss how the categorical SPIM can be added to SMR, 2-flank SPIM, or other latent identity SCR models.

Key words: genetic mark–recapture; Identity Diversity Index; microsatellites; partial identity; spatial capture–recapture; spatial mark–resight; unmarked spatial capture–recapture.

Received 28 July 2018; revised 13 December 2018; accepted 9 January 2019; final version received 31 January 2019. Corresponding Editor: Lucas N. Joppa.

Copyright: © 2019 The Authors. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

† **E-mail:** ben.augustine@cornell.edu

INTRODUCTION

Animal population density is a fundamental quantity in wildlife ecology, and therefore, estimating population density is a primary challenge for ecologists (Laake et al. 1993, Efford 2004). Mark–recapture and spatial capture–recapture (SCR) methods are among the most reliable methods for estimating population abundance and density; however, they generally require that individual identities of captured animals are determined with certainty (e.g., marks are recorded correctly and not lost; Otis et al. 1978). Recently, several classes of SCR models that utilize latent or partially latent individual identities have been introduced, extending the utility of SCR models to populations that are unmarked (unmarked SCR, alternatively known as a spatial counts model; Chandler and Royle 2013), populations for which only a subset of individuals are marked (spatial mark–resight [SMR]; Sollmann et al. 2013), populations for which individuals are identified only in a subset of the years in a multiyear study (spatially explicit Integrated Population Models; Chandler and Clark 2014), and populations for which some or all samples carry only partial identifications (spatial partial identity models [SPIMs]; Augustine et al. 2018). Similar to recently developed classical partial identity models (Bonner and Holmberg 2013, McClintock et al. 2013), this class of SCR models with latent individual identities shares the feature that the true capture histories for some or all observed individuals are latent and must be probabilistically reconstructed using Markov Chain Monte Carlo (MCMC; or possibly marginalized out of the likelihood if computationally feasible). Unlike the classical partial identity models, however, these SCR models with latent individual identities use the spatial locations where samples were collected, together with a spatially explicit model of sample deposition, to probabilistically associate latent or partial identity samples, thereby reducing uncertainty in individual identity. This reduced uncertainty in individual identity is then propagated to a reduced uncertainty in other model parameter estimates, including population size and density. Therefore, the spatial location where a sample was collected constitutes a *continuous partial identity* and unmarked SCR and SMR can be

considered special cases of a SPIM, along with the “2-flank” SPIM that probabilistically associates left-flank and right-flank photographs for species with bilateral identification (Augustine et al. 2018).

The key feature of a SPIM is that the magnitude of uncertainty in individual identity and the model used to resolve it stems directly from two key aspects of population ecology—population density and home range size. In a SPIM, there will be more uncertainty about sample identity when more individuals are exposed to capture at the same traps, which happens when animals occur at higher densities, and when their home ranges are larger. Within the context of an SCR model, this is equivalent to scenarios in which there are higher densities of individual activity centers within the state space and larger detection function spatial scale parameters, for example, σ for the common half-normal detection function model. Note, however, that the key feature that both population density and home range size determine, which is directly responsible for the magnitude of uncertainty in individual identity, is the *magnitude of home range overlap*, or more conceptually, the magnitude of overlap in individual utilization distributions (e.g., as quantified by Fieberg and Kochanny 2005), which increases independently by increasing either population density or home range size.

We propose a metric to quantify the degree of home range overlap for a given population density and σ , and thus, the expected magnitude of uncertainty in individual identity. The Simpson’s Diversity Index can be applied to the spatially explicit, individual detection probabilities and averaged over many points on the landscape and realizations of the SCR process model (configurations of activity centers) to produce an Identity Diversity Index (IDI), which is conceptually similar to a metric of utilization distribution overlap for all individuals in the population. The IDI quantifies the expected diversity in individual identity of collected samples in one location, averaged over the landscape, described in detail in Appendix S1. Fig. 1 provides a visualization of how the magnitude of uncertainty in individual identity conceptually scales with population density and home range size as quantified by the SCR σ parameter. This relationship between the magnitude of uncertainty in individual identity

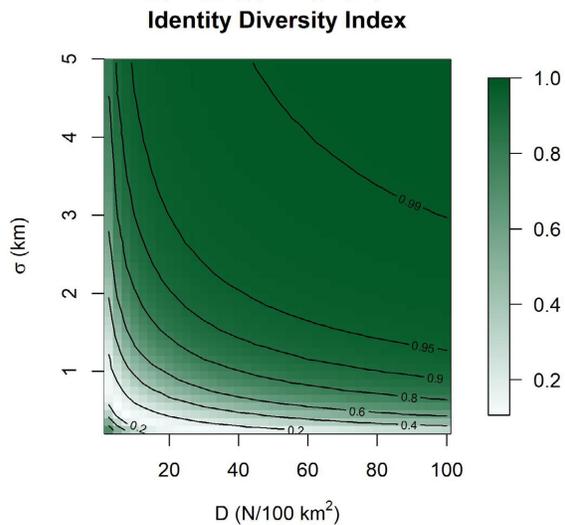


Fig. 1. The Identity Diversity Index, quantifying the magnitude of uncertainty in individual identity, as a function of population density (D) and the spatial capture–recapture spatial scale parameter (σ). Increasing D and/or σ increases the magnitude of uncertainty in individual identity and with N fixed, the expected precision of N and D estimates.

and the spatial features of animal populations is what set a SPIM apart from the classical partial identity models (Bonner and Holmberg 2013, McClintock et al. 2013, Knapp et al. 2009), where the magnitude of uncertainty in individual identity scales with population abundance alone.

Another key feature of the currently available SPIMs is that partial identity information can reduce the uncertainty in individual identity through three mechanisms: (1) by adding deterministic identity associations, (2) by adding deterministic identity exclusions, and (3) by improving probabilistic identity associations. To our knowledge, the mechanics of these models have not previously been framed in this manner, which we believe is important for understanding how they use partial identity information to improve density estimation, and for extending the types of partial identity information used by future SPIMs. Here, we define a deterministic identity association as a connection between samples from the same individual that also implies that the samples are excluded from being connected with samples from other individuals. This is distinguished from a deterministic

identity exclusion, which can only prevent certain samples from being combined together. A probabilistic identity association occurs when two samples have a positive posterior probability of belonging to the same individual, and as this probability increases, the probability they belong to another individual necessarily decreases. Probabilistic identity associations can be improved with partial identity information, effectively converging to deterministic identity associations as partial identity information increases, as we later demonstrate using microsatellite loci.

All SPIMs use the spatial location where samples were collected to improve probabilistic identity associations. The unmarked SCR model (Chandler and Royle 2013) and the model of Chandler and Clark (2014) that relates occupancy data to a latent SCR model use spatial information alone to inform individual identity. Typical SCR, on the other hand, uses all possible deterministic identity associations. Spatial mark–resight (Sollmann et al. 2013) and the “2-flank” SPIM (Augustine et al. 2018) represent two intermediate cases that both utilize some deterministic identity associations and exclusions. Spatial mark–resight makes deterministic identity associations between the samples of marked and identifiable individuals, which simultaneously excludes them from being connected to samples from other individuals. Deterministic identity exclusions are then made between the samples of unidentifiable individuals whose mark status can be observed (e.g., an unmarked sample cannot belong to a marked individual; Royle et al. 2013). The 2-flank SPIM makes deterministic identity associations across the same flank of the same individual, enforcing exclusions with non-matching samples from the same flank. Finally, deterministic identity exclusions arise in the 2-flank SPIM from the fact that an individual can only have one left and right flank.

Augustine et al. (2018) demonstrated that further deterministic identity exclusions are possible in SPIMs by using individual sex to split a data set into two population subgroups whose identities could not logically match, reducing the uncertainty in individual identity and thus abundance and density estimates. Splitting the population into identity subgroups of increasingly smaller size is conceptually similar to applying

unmarked SCR to separate populations, each with increasingly lower population densities, moving the population under study to more favorable regions of the IDI along the density axis (Fig. 1). However, rather than splitting data sets into increasingly smaller subsets, it is desirable to have a model that incorporates these categorical identity exclusions, allows for imperfect observation of the category levels, and allows parameters to be shared across identity subgroups. Further, when all categories are combined into a single analysis, the distribution of individuals across the category levels provides further information that can improve probabilistic identity associations. For example, if a population is 75% female, it is more likely that two nearby male samples came from a single individual than if the population is 75% male. Finally, the proportion of the population in each identity subgroup may be of ecological interest, such as individual sex or age class. Therefore, we are introducing a new class of SCR model, the “categorical SPIM”, which uses partially identifying categorical covariates to add both deterministic identity exclusions and reduce the uncertainty in probabilistic identity associations.

Partially identifying categorical covariates exist in many types of invasive and noninvasive wildlife sampling; for example, in studies using remote cameras, features such as sex, age class, and color morph may be observable in at least some photographs and similar categorical features can be extracted from bioacoustic sensors for some species (Reby et al. 1999, McIntosh et al. 2015). In more invasive wildlife sampling involving live capture, many more features are measurable and researchers may apply categorical marks whose combination do not provide full identities (e.g., colored collars or ear tags), or categorical marks may be fully identifying (Lewis et al. 2015), but imperfectly observed, or fully identifying categorical marks may be partially lost over time, similar to the problem of complete tag loss typically resolved by double tagging (Cowen and Schwarz 2006). Perhaps the most informative source of categorical identity covariates in wildlife sampling, however, comes from microsatellite genotypes.

Microsatellite loci are not typically thought of as categorical identity covariates because studies have traditionally aimed to ensure that unique

multilocus genotypes correspond to unique individuals with high probability through the use of many highly variable loci. This determination is typically made using P(ID) and/or P(sib) criteria, both of which estimate the probability that any two randomly selected individuals (or full siblings) in a population would share the same multilocus genotype, given the observed allele diversities and frequencies (Waits et al. 2001). The possibility that multiple individuals have the same multilocus genotype in a capture–recapture data set has been referred to as the “shadow effect” (Mills et al. 2000) and is considered a type of low-frequency error in assigning individual identities that introduces minimal bias into parameter estimates in capture–recapture studies if P(ID) or P(sib) criteria are strictly enforced. Using the categorical SPIM to model genotype data is especially appealing when genotypes are not variable enough to be considered unique because it does not make deterministic connections between samples with matching genotypes; thus, in the categorical SPIM we propose, multiple individuals in the population may have the same genotype, shifting the “shadow effect” from a source of bias to an additional source of uncertainty. Further, microsatellite data are ideal for investigating the performance of the categorical SPIM across the full range of uncertainty in individual identity—from unmarked SCR to SCR—due to the near perfect individual information content in the full genotype. Therefore, we will use genotypes as the main application to demonstrate how categorical identity and spatial information combine to determine the magnitude of uncertainty in individual identity and the resulting density estimates, but other types of categorical covariates could be interchanged without loss of generality.

Here, we generalize the unmarked SCR model to develop the “categorical SPIM.” We show via simulation that in scenarios with more sparse data than previously considered and/or scenarios with larger σ s and higher densities, the unmarked SCR density estimator is biased, very imprecise, and the parameters are frequently not identifiable, demonstrating the importance of population density and home range size to the application of latent and partial identity SCR models. We then show that adding categorical identity covariates removes this bias and removes parameter non-

identifiability, increases precision, allowing for reliable density estimation across a wider range of values of density and σ for a given capture process scenario. We also demonstrate that the uncertainty in the posterior for n^{cap} , the latent number of individuals captured during a survey, correlates well with the uncertainty in the posterior of N , suggesting it is a good single metric to quantify the observed magnitude of uncertainty in individual identity for a given data set. Finally, we apply the categorical SPIM to a previously published black bear data set in which we demonstrate how well the proposed model can reproduce the original density estimate using fewer loci than originally genotyped. Using this data set, we demonstrate that all uncertainty in individual identity can be removed with enough categorical identity covariates, producing equivalent estimates to an SCR model where all identities are known with certainty.

METHODS—DATA AND MODEL DESCRIPTION

Methods—unmarked SCR Foundation

First, we introduce the version of the unmarked SCR model that we will expand to allow categorical identity covariates. The unmarked SCR model is a typical hierarchical SCR model except that information about individual identity is not retained during the observation process. Formal inference is achieved by relating the spatial pattern of observed counts or detections at each of the J traps to the latent structure of the SCR process model. For the process model, we assume the N individuals in the population have activity centers that are distributed uniformly across a two-dimensional state space \mathcal{S} of arbitrary size (A) and shape, that is, $s_i \sim \text{Uniform}(\mathcal{S})$, $i = 1, \dots, N$ (see Borchers and Efford 2008, Reich and Gardner 2014, Royle et al. 2016 for alternative specifications). The activity centers are organized in the $N \times 2$ matrix \mathbf{S} .

For the observation model, we introduce the $N \times J$ fully latent capture history \mathbf{Y}^{true} , recording the number of detections or counts for each individual at each trap summed across the K occasions. The locations of the J traps are stored in the $J \times 2$ matrix \mathbf{X} . We assume that the number of counts or detections for each individual at each trap is a decreasing function of distance between the activity centers and traps. If using a

count model, we assume the latent counts are Poisson: $y_{ij}^{\text{true}} \sim \text{Pois}(K\lambda(s_i, \mathbf{x}_j))$, where $\lambda(s_i, \mathbf{x}_j) = \lambda_0 \exp\left(-\frac{\|s_i - \mathbf{x}_j\|^2}{2\sigma^2}\right)$, \mathbf{x}_j is the location of trap j , λ_0 is the expected number of counts for a trap located at distance 0 from an activity center, and σ is the spatial scale parameter determining how quickly the expected counts decline with distance from the activity center. We also consider an alternative Bernoulli observation model for which $y_{ij}^{\text{true}} \sim \text{Bin}(p(s_i, \mathbf{x}_j), K)$, where $p(s_i, \mathbf{x}_j) = 1 - \exp(-\lambda(s_i, \mathbf{x}_j))$.

During the observation process, the true, latent capture history, \mathbf{Y}^{true} , is disaggregated into the observed capture history, \mathbf{Y}^{obs} , discarding information about individual identity by storing one observation per row in \mathbf{Y}^{obs} (e.g., no samples are deterministically connected to the same individual). More specifically, \mathbf{Y}^{obs} is the $n^{\text{obs}} \times J$ matrix with entries 1 if sample m was recorded in trap j and 0 otherwise. Note that if we assume a Bernoulli observation model, each detection event will constitute a single observation, while if we assume a Poisson observation model, counts are disaggregated into observations of single counts, because counts from the same individuals cannot be deterministically connected without certain and unique identities. To visualize this, below is an example of true and disaggregated observed data set where $N = 2$ and $J = 3$:

$$\mathbf{Y}^{\text{true}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \quad \mathbf{Y}^{\text{obs}} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Methods—categorical SPIM

We propose a class-structured version of the unmarked SCR model in which class membership is determined by each individual's full categorical identity (e.g., full genotype). Here, we define a *full categorical identity* to be an individual's set of true values for n_{cat} categorical covariates, where n^{cat} is the maximum number of categorical covariates considered, and multiple individuals in the population can share the same full categorical identity. We will modify the unmarked SCR model such that single or multiple categorical covariates, potentially partially or even fully unobserved, are recorded with each observed sample at each trap. Further, continuous covariates could be

discretized into categories if it is safe to assume there is no measurement error. The linked density and categorical covariate models (joint process model) are fully latent and we use the n^{obs} trap-referenced, observed categorical covariates to make inference about this latent structure. The observed data then consist of two linked data structures: \mathbf{Y}^{obs} , an $n^{\text{obs}} \times J$ capture history indicating the trap at which each sample was recorded and \mathbf{G}^{obs} , an $n^{\text{obs}} \times n^{\text{cat}}$ identity history recording the observed categorical covariate(s) of each sample with category level enumerated sequentially as described below, or recorded as a 0 if not observed.

For the joint process model, we assume that each individual has a full categorical identity associated with its activity center. Following Wright et al. (2009), we assume that all possible category levels for each categorical covariate are known, with the number of categories for each covariate l being n_l^{levels} , $l = 1, \dots, n^{\text{cat}}$. Next, we introduce the population category level probabilities for covariate l , γ_l , of length n_l^{levels} and corresponding to the enumerated category levels $(1, \dots, n_l^{\text{levels}})$ for covariate l . Then, we introduce the $N \times n^{\text{cat}}$ matrix \mathbf{G}^{true} , where g_i^{true} is the full categorical identity of the individual with activity center s_i . Finally, we assume the categorical identity of each individual for each covariate is distributed following the covariate-specific category level probabilities according to $g_{il}^{\text{true}} \sim \text{Categorical}(\gamma_l)$, implying that category levels are independent across covariates (e.g., linkage equilibrium in the genetic context) and individuals. Using the example true and observed capture histories above, potential true and observed structures for the categorical identities assuming 3 categorical identity covariates with 4 levels each are as follows:

$$\mathbf{G}^{\text{true}} = \begin{bmatrix} 1 & 4 & 3 \\ 4 & 2 & 3 \end{bmatrix} \quad \mathbf{G}^{\text{obs}} = \begin{bmatrix} 1 & 4 & 3 \\ 1 & 4 & 3 \\ 4 & 0 & 3 \\ 0 & 0 & 3 \end{bmatrix}$$

In this case, the first two observed samples could possibly have come from the same individual as could the third and fourth sample; however, the third sample could not have come from the same individual as the first two samples. The fourth sample with two unobserved categories

could possibly belong to the same individual as that which produced the first three samples, with only the third sample being a correct match. In this example, potential values of l for any of the γ_l for any of the 3 categories could be (0.3, 0.1, 0.2, 0.4), corresponding to category levels (1, 2, 3, 4), with the only requirement on these category level probabilities being that they sum to 1 for each identity covariate. Also, note that each covariate may have a different number of category levels.

The observation process is the same as unmarked SCR except that a categorical identity, potentially partially or fully latent, is associated with each trap-referenced observation. The missing data process could be a simple binomial model for identification success, perhaps with covariate-specific identification probabilities; however, if we assume that covariate observation values do not vary by individual or by category level, the likelihood for the missing data process does not change when updating latent identities or latent categorical covariate values and can be ignored in the MCMC algorithm. Therefore, we make the assumption that covariate identification probabilities do not vary by individual or category level, but this assumption would be easy to relax.

The unmarked SCR model and categorical SPIM use a process similar to data augmentation to estimate population abundance and density (Royle et al. 2007) and to model the uncertainty in individual identity by providing latent structure to allow for different configurations of the observed samples across the individuals in the population (Chandler and Royle 2013, Augustine et al. 2018). Unlike typical data augmentation, the number of captured individuals, n^{cap} , and their capture histories are unknown in unmarked SCR, so rather than augmenting the observed capture history, we augment a possibly true capture history constructed by assigning individual identities to the observed samples in \mathbf{Y}^{obs} based on the spatial proximity of samples and the compatibility of their observed categorical identities.

We then augment \mathbf{G}^{true} to size $M \times n^{\text{cat}}$, which is initialized using the minimally implied categorical identity of the samples each individual is initialized with, and the remaining elements of \mathbf{G}^{true} that are not determined by these samples are simulated from the category level probabilities. Similar to typical data augmentation, M is

chosen by the analyst to be much larger than N and we use \mathbf{z} , a latent indicator vector of length M , to indicate which individuals are in the population; however, unlike typical data augmentation, this vector is fully, rather than partially latent. We assume $z_i \sim \text{Bernoulli}(\phi)$, inducing the relationship $N \sim \text{Binomial}(M, \phi)$. Then, population abundance is $N = \sum_{i=1}^M z_i$ and population density, D , is $\frac{N}{A}$. See Appendix S2 for a full description of the MCMC algorithm.

Note that n^{cap} , the total number of individuals captured, typically denoted by n and an observed statistic in capture–recapture models, is a derived, random variable in unmarked SCR with a posterior distribution that quantifies the magnitude of uncertainty in individual identity. Specifically, the posterior sample for n^{cap} on each MCMC iteration is calculated following $n^{\text{cap}} = \sum_{i=1}^M (\sum_{j=1}^J Y_{ij}^{\text{true}}) > 0$ —simply the number of individuals currently allocated at least one sample. As more categorical identity information is added, the posterior distribution of n^{cap} should converge to the single, true value. Finally, we introduce the derived vector, \mathbf{ID} , of size n^{obs} , that records the latent individual, $1, \dots, M$, each sample is assigned to. This vector is updated on each MCMC iteration, producing a posterior for true identity for each sample, which can be post-processed to obtain pairwise posterior probabilities that any two samples originated from the same individual. The posterior distribution of the true covariate values of samples with missing values can also be recorded.

SIMULATIONS

We conducted two simulation studies (A and B) to explore the performance of the categorical SPIM.

Simulation A specifications

First, we conducted a simulation study to demonstrate the utility of introducing an increasing number of categorical identity covariates for improving density estimation over the baseline case of unmarked SCR. Further, we sought to demonstrate that the effectiveness of adding categorical identity covariates for improving density estimation depends on the two axes of the IDI, population density, and σ . Here, the number of identity categories is defined to be the total

number of unique categories implied by the n_{cat} covariates with n_1^{levels} each. We start with sampling scenarios more challenging for unmarked SCR than considered by Chandler and Royle (2013) by considering scenarios with more sparse detection data achieved by using a smaller trapping array, a lower λ_0 , and scenarios with higher D given N . Specifically, our trapping array consisted of 81 traps in a 9×9 grid with unit spacing, buffered by 3 units to define the state space of 225 units and we chose densities $D \in \{0.17, 0.35\}$, corresponding to $N \in \{39, 78\}$. We considered that populations were sampled for $K = 5$ occasions for all scenarios.

We conducted simulations across 4 scenarios with a 2×2 factorial design using low and high values of σ and D . Scenarios A1 and A2 were the low σ scenarios with $\sigma = 0.5$, and scenarios A3 and A4 doubled σ to 1. To account for compensation in the detection function parameters (Efford and Mowat 2014) and maintain similar levels of data sparsity with the larger σ , we lowered λ_0 from 0.25 to 0.061 to approximately match the expected number of captures for each individual to that of the scenarios with $\sigma = 0.5$ ($E[\text{caps}] \sim 1.65$, achieved by trial and error). On this unit spacing grid, with $\sigma = 0.5$, the majority of an individual's captures fell within a 4-trap area, whereas with $\sigma = 1$ the majority of an individual's captures fell within a 16-trap area. Scenarios A1 and A3 were the high abundance scenarios with $N = 78$, and scenarios A2 and A4 were the low abundance scenarios with $N = 39$. The approximate Identity Diversity Indices (interpolated from Fig. 1) for scenarios A1–A4 were 0.38, 0.23, 0.76, and 0.58. Within each scenario, we explored 9 subscenarios with differing numbers of identity categories, with 1 identity category corresponding to unmarked SCR. Following the unmarked SCR subscenario, we sequentially added identity covariates with 2 category levels each, leading to the number of unique identity categories increasing exponentially with base 2 (2, 4, 8, 16, 32, 64, 128, 256). As noted previously, these categorical identities are not unique—they may be represented multiple times by different individuals in a population, but with decreasing frequency as the number of unique identity categories increase.

Further, scenario A2 was modified to better disentangle the effects of increasing D from increasing N . Increasing D by increasing N

simultaneously increases uncertainty in individual identity and reduces data sparsity, which have opposite effects on estimator precision. Therefore, to better explore the effect of increasing D on the uncertainty in individual identity, D must be increased by constraining a fixed N into a smaller state space. The state space can be reduced in two ways; the number of traps can be reduced, keeping the same state space buffer, or the number of traps can be fixed while reducing the state space buffer. In the first scenario, data sparsity is increased since a lower proportion of individuals will be located on the interior of the trapping array and we suspect the absolute number of traps is important for unmarked SCR density estimation. In the second scenario, data sparsity is decreased because a larger proportion of individuals live on the interior of the trapping array. In order to retain the same number of traps, we chose to increase the density of scenario A2 by reducing the state space buffer from 3 to 1 units, thus constraining the N individuals into a smaller state space area (Scenario A2b). The reduction in the state space area increased D from 0.17 to 0.32, and raising the approximate IDI value from 0.23 to 0.37. The factorial layout for simulation scenario A can be found in Table 1.

Simulation B specifications

We conducted a second simulation study to demonstrate that the categorical SPIM can accommodate partially observed categorical identities (missing identity covariate values) and provide a proof of concept for using partial genotypes that are the result of failed DNA amplification, rather than as part of the study design as would be the case in the first set of simulations if identity categories were genotype loci. We used

Table 1. Factorial design for the simulation scenarios A for two levels of abundance (N) and detection function spatial scale parameter (σ).

		σ	
		0.5	1.0
N	78	A1 – $D = 0.35$	A3 – $D = 0.35$
	39	A2 – $D = 0.17$	A4 – $D = 0.17$
		A2b – $D = 0.32$	

Note: Subscenarios A2 and A2b vary individual density (D) for a fixed abundance by varying the state space extent.

the parameter values from scenario A3 above, but introduced imperfect detection to the observed genotypes. We simulated data sets with 7 categorical identity covariates, each with 5 equally common category levels, and the category value for each categorical covariate was then observed with probability 0.5, leading to the average categorical identity being observed at 3.5 of the categorical identity covariates. With 5 equally probable levels per identity covariate, these identity covariates are more informative than the 2-level covariates used in simulation scenario A and closer in information content to a multilocus genotype. While each genotype loci usually has more than 5 levels, they are usually not equally distributed, so limiting the number of equal probability levels per identity covariate to 5 likely represents a more fair comparison to genotypes in practice. We fit the categorical SPIM to these data sets, assuming all partial categorical identities were usable (Scenario B1) or 75% of the partial categorical identities were usable as might be the case when using partial genotypes if a subset was deemed to be unreliable due to the likelihood of containing genotyping errors (Scenario B2). We then fit the null SCR model to the perfectly observed data for comparison (Scenario B3).

Simulation MCMC specifications

For simulation scenario A, we simulated and fit our model to 144 data sets within each subscenario, and for simulation scenario B, we simulated and fit our model to 128 data sets (differing due to cluster computing availability). Within each subscenario of simulation scenario A, we ran 3 chains for 100,000 iterations for the unmarked SCR estimator. We calculated the Gelman-Rubin statistic, R_c (Gelman and Rubin 1992) and only computed inferential quantities (listed below) for sets of chains that indicated convergence, which we define as those with $R_c < 1.1$ for parameter N . If fewer than 95% of the sets of chains indicated convergence, we ran 3 chains for 100,000 iterations for the next scenario with a categorical identity covariate (or additional covariate), until >95% indicated convergence. After that, we reverted to a single chain to save computation time for the remaining subscenarios. For simulation scenario B, there were no convergence problems, so we ran a

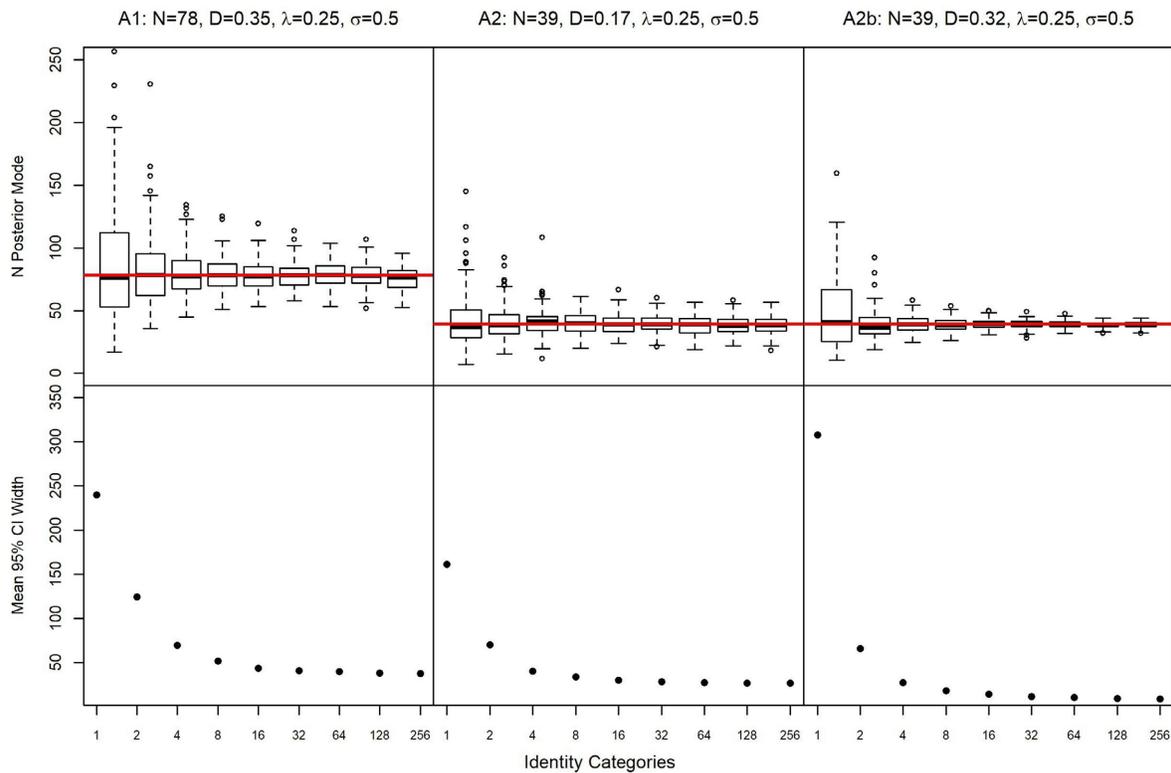


Fig. 2. Histograms of the posterior modes and mean 95% credible interval widths of abundance (N) estimates plotted against the number of identity categories in scenarios A1, A2, and A2b with baseline detection rate $\lambda_0 = 0.25$ and detection function spatial scale parameter $\sigma = 0.5$. Population density, D varies across scenarios, and A2b is the scenario that increases D while N remains fixed. Note the number of identity categories increase exponentially as 2-level identity covariates are added sequentially. Also, note that the histograms depict the median posterior mode, while the mean posterior mode is used to calculate the expected values and bias in Appendix S3: Table S1. The red line depicts the simulated value of N .

single chain for 60,000 iterations. We calculated point estimates using the posterior mode and interval estimates using the highest posterior density (HPD) interval. We were interested in the frequentist bias and coverage of the categorical SPIM estimator, the accuracy of the estimator depicted visually by the variance and right skew of the sampling distribution and quantitatively by the mean squared error and coefficient of variation ($CV: 100 \times \text{posterior sd}/\text{posterior mode}$), and the precision, quantified by the mean 95% HPD interval width. Also of interest was the use of the precision of n^{cap} , also quantified by the mean 95% HPD interval width, as a metric of uncertainty in the individual identity of observed samples that can predict the uncertainty in N .

Simulation A results

No unmarked SCR scenarios led to reliable convergence, with the percentage of simulated data sets for which the Gelman-Rubin statistic, R_c , indicated convergence ranging from 0.52 to 0.90 (Appendix S3: Table S1). The percentage of simulated data sets for which R_c indicated convergence increased as more categorical identity covariates were added, with >95% convergence achieved in all scenarios with 2 categorical identity covariates. Inspection of the MCMC chains (not shown) indicated that the main cause of lack of convergence was data realizations for which σ was not identifiable. The unmarked SCR abundance estimator was right-skewed with high variance (Figs. 2 and 3), except in Scenario A2

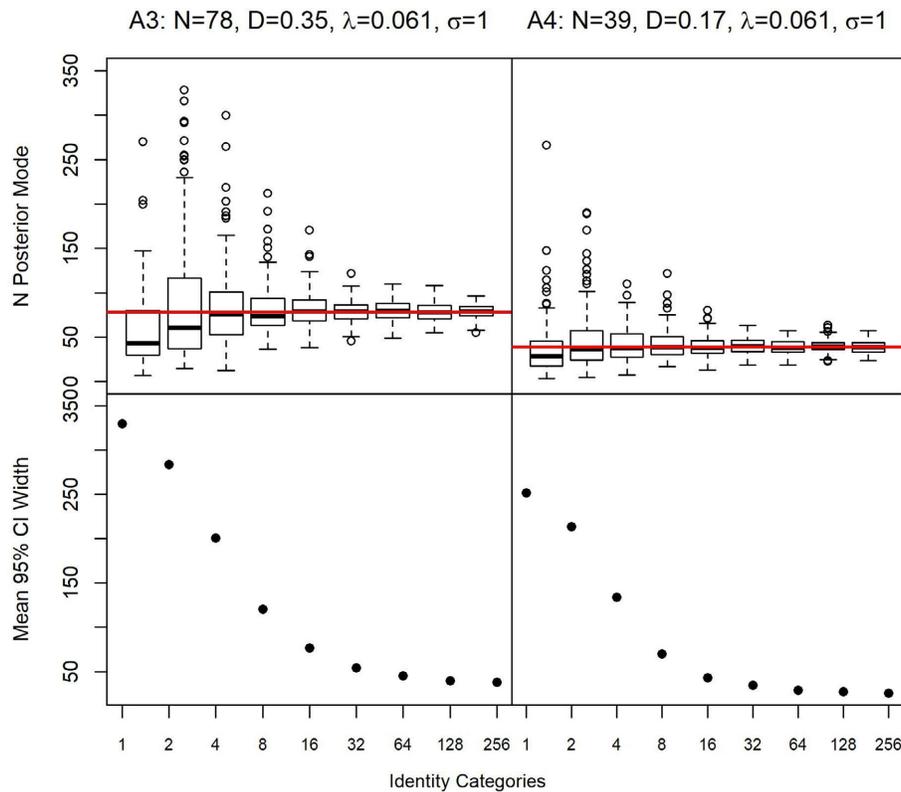


Fig. 3. Histograms of the posterior modes and mean 95% credible interval width of abundance (N) estimates plotted against the number of identity categories in scenarios A3 and A4 with baseline detection rate $\lambda_0 = 0.061$ and detection function spatial scale parameter $\sigma = 1$. Population density, D , varies across scenarios. Note the number of identity categories increases exponentially as 2-level identity covariates are added sequentially. Also, note that the histograms depict the median posterior mode, while the mean posterior mode is used to calculate the expected values and bias in Appendix S3: Table S1. The red line depicts the simulated value of N .

where both σ was small and abundance was low. The unmarked SCR estimator had a large mean 95% credible interval [CI] width relative to abundance in all scenarios. The unmarked SCR abundance estimates were generally positively biased (Appendix S3: Table S1), although the positive bias in Scenario A4 was only 5%. The only scenario where the unmarked SCR estimator was negatively biased was A3; however, because the A3 scenario with 1 categorical identity covariate was positively biased, we attribute the observed negative bias in the A3 unmarked SCR scenario to the 44% of simulations that were discarded due to lack of convergence. Data sets that did not lead to convergence likely disproportionately correspond to those producing N estimates in the upper half of the sampling distribution because

they disproportionately had more capture events (data not shown).

Adding and increasing the number of identity categories (from 2 to 256 using 8 2-level categories) reduced bias and increased precision in all scenarios, but with diminishing returns as more identity categories were added. The reduction in mean 95% CI width for n^{cap} by the introduction of identity categories was closely related to the reduction in mean 95% CI width for abundance; however, the relationship was not linear and varied by scenario (Fig. 4). More identity categories were required to reach maximum precision when abundance was higher and σ larger. The largest improvement in precision and abundance with the addition of identity categories was seen in the low abundance, high density,

low σ scenario (A2b), where the majority of uncertainty in abundance was removed with the addition of one 2-level categorical covariate. Note that the precision of estimates in scenario A2b converged to a lower value than scenario A2 because the same $N = 39$ individuals were constrained to be located within 1 unit of the trapping array, rather than 3 units, decreasing data sparsity. Positive bias in \hat{N} was $<5\%$ in scenarios A1, A2, A2b, and A4 when 2 identity categories were available, and $<5\%$ in Scenario A3 when 8

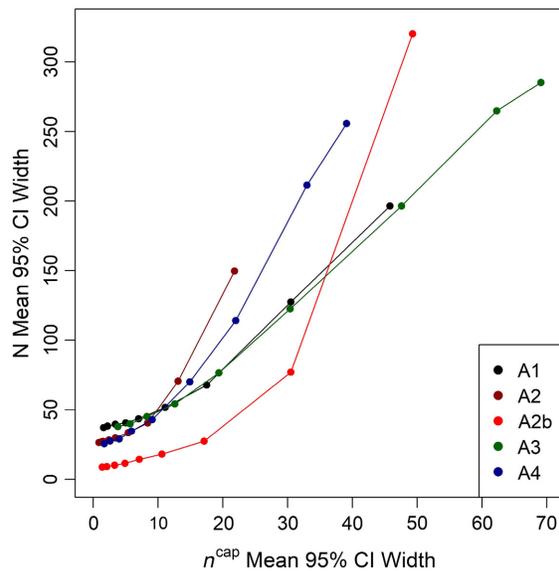


Fig. 4. The relationship between uncertainty in the number of individuals captured, n^{cap} and uncertainty in abundance, N for each of the 5 A scenarios. For each of the 5 scenarios plotted, the scenario with maximal uncertainty in n^{cap} is the unmarked spatial capture–recapture estimator and the uncertainty in n^{cap} converges toward zero as identity categories are added one 2-level identity covariate at a time.

identity categories were available. With no or very few identity categories, the latent identity samples tended to be allocated to more individuals than were actually captured, resulting in positive bias in n^{cap} , with more bias in the large σ scenarios, and we attribute this as a cause for the positive bias in \hat{N} .

Increasing σ decreased the precision and accuracy of the unmarked SCR and categorical SPIM n^{cap} and N estimates (A1 vs. A3 and A2 vs. A4; Appendix S3). Increasing D by increasing N decreased the precision and accuracy of the n^{cap} estimates, but increased the precision and accuracy of the N estimates demonstrating that the additional uncertainty in individual identity was more than offset by the lower data sparsity (A1 vs. A2 and A3 vs. A4). Increasing D without increasing N decreased the precision and accuracy of both n^{cap} and N estimates. The IDI did not perfectly predict the metrics of precision and accuracy across scenarios that changed both σ and D for all levels of identity categories, but the IDI generally correlated negatively with precision and accuracy (Appendix S3).

Simulation B results

The partially observed categorical identity covariate simulations (scenarios B1-3) produced minimally biased abundance point estimates (1 and 2.5% positive bias) and interval estimates that were nearly as precise as the scenario in which the identities were perfectly observed (Table 2). In these scenarios, categorical identities were observed at half of the 7 identity covariates, on average, producing data sets with an average of <1 sample observed at all category levels—data sets that would be unusable if full categorical identities were required, as might be the case with genetic capture–recapture requiring

Table 2. Simulation results for the partial genotype analyses.

Scenario	$\hat{\lambda}_0$	$\hat{\sigma}$	\hat{N}	\hat{n}^{cap}	Cov	N Wid	n^{cap} Wid	n^{cap}	# complete
True	0.250	0.500	38.0
B1	0.253	0.490	39.4	18.4	0.953	27.6	1.8	18.5	0.48
B2	0.189	0.481	40.0	17.1	0.953	32.7	2.4	17.0	0.31
B3	0.256	0.492	38.7	.	0.953	26.1	.	18.5	64.31

Notes: Scenarios B1 and B2 had loci amplification probabilities of 0.5. Scenario B1 used all partial genotype samples, while Scenario B2 used 75% of the partial genotype samples. Scenario B3 is a typical spatial capture–recapture model using all full genotype samples for comparison. Parameter mean point estimates are presented, along with coverage of N (Cov), the mean 95% credible interval widths for N (N Wid) and n^{cap} (n wid), and the mean number of samples with complete genotypes (# complete). Dots indicate non-applicable cells.

complete genotypes. In Scenario B1, where all partial categorical identities were used, the interval estimate was 95% as precise as the complete data analysis, and in Scenario B2, where 25% of the partial categorical identities were unusable, the interval estimate was 80% as precise than the complete data analysis. In both scenarios, an average of 3.5 categorical covariates provided enough information that the uncertainty in n^{cap} was small (mean credible interval widths of 1.8 and 2.4 relative to mean n^{cap} of 18.5 and 17.0 in B1 and B2, respectively).

APPLICATION—CENTRAL APPALACHIAN BLACK BEARS

We applied the categorical SPIM to a hair trapping data set that used 7 microsatellite loci for individual identification of American black bears (*Ursus americanus*) in the Central Appalachians, USA. This data set comes from a study conducted along the Kentucky-Virginia, USA, border across 2 study areas during 2012 and 2013 to estimate the population density and abundance of a recently reintroduced population that was in the process of recolonizing vacant range (Murphy et al. 2016). We chose to use the data set from the larger study area in 2013 because our model should perform better on the larger trapping array and more samples were collected in 2013 than in 2012 at this site. The specifics of the data collection methods are described by Murphy et al. (2016); of particular relevance is that eighty-one hair traps were deployed across the 215-km² study area with an average trap spacing of 1.6 km, and all traps were checked weekly for 8 consecutive weeks, with a week constituting a capture occasion. Similar to most bear hair trapping studies, hair samples were subsampled for genotyping because of the prohibitive costs of genotyping thousands of samples, such that at most 1 hair sample per trap per occasion produced an individual identity. The capture and subsampling processes resulted in 95 samples from 45 females and 87 samples from 37 males, determined using the P(sib) criterion. The spatial distribution of traps and individually identified hair sample observations are depicted in Fig. 5. The microsatellites used were G10H, G10L, G10M, MU23, G10J, G10B, and G10P, which had genotype frequencies of 19, 22, 19, 17, 12, 15, and

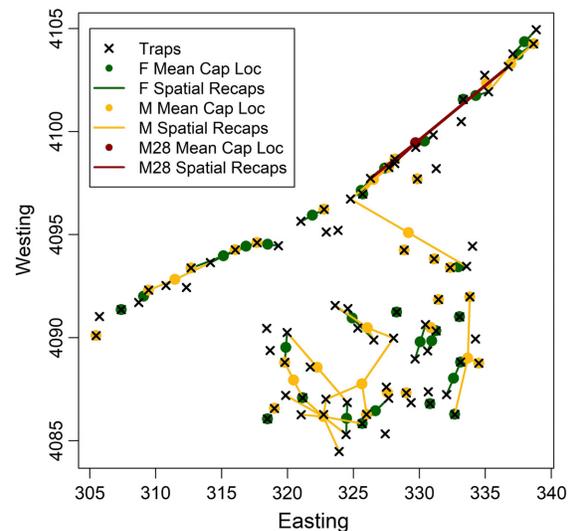


Fig. 5. Mean male and female capture locations with spatial recaptures for the Kentucky black bear data set. Male 28 is highlighted in yellow because it had a large spatial recapture and was not detected at several traps in between the two traps where it was detected, requiring more loci for the categorical spatial partial identity models to link these two samples together.

10 for females and 21, 18, 18, 22, 14, 13, and 10 for males. Despite the large number of genotypes at each locus, the majority of individuals shared just 2–4 genotypes at each locus, making them less informative than if the loci-specific genotypes were equally distributed as they were in our simulation studies.

The goal of this analysis was to fit the categorical SPIM using from 1 to 7 loci, added in the order listed above, and to compare the estimates to the null SCR estimate that does not allow for any uncertainty in individual identity. Further, we also consider a scenario adding partial genotype samples (2 for females, 4 for males) into the analysis that were originally discarded. For all genotype scenarios, the trapping array was buffered in the X and Y dimension by 3 km for females and 6 km for males, leading to state space sizes of 1042.5 km² and 1473.4 km² for females and males, respectively. For each sex-specific, 1–7 loci data set, we ran 32 Markov chains for 250,000 iterations each, thinned by 50, and discarded the first 25,000 iterations as burn in, leaving 1.4 million samples from the

Table 3. Sex-specific estimates of the detection function baseline encounter rate, λ_0 , the detection function spatial scale parameter, σ , and abundance from regular spatial capture–recapture (SCR, Full), categorical spatial partial identity models (SPIM) using 1–7 loci (1L–7L), and categorical SPIM with 7 loci, adding partial identity samples not included in the SCR estimate (7L+).

Genotype data	$\hat{\lambda}_0$	$\hat{\sigma}$	\hat{N}	\hat{n}^{cap}	\hat{N} SE	\hat{N} CV
Full	0.156	0.92	182.9	45	28.7	15.7
1L	0.204	0.79	198.7	44.66	37.4	18.8
2L	0.175	0.82	210.0	46.53	36.9	17.6
3L	0.161	0.90	188.4	45.06	31.8	16.9
4L	0.157	0.92	183.4	45.00	29.6	16.1
5L	0.156	0.92	183.0	45.00	28.8	15.8
6L	0.156	0.93	182.5	45.00	28.9	15.8
7L	0.157	0.93	182.2	45.00	28.9	15.9
7L+	0.163	0.92	180.2	45.00	28.1	15.6
Full	0.066	2.04	117.5	37.00	19.0	16.2
1L	0.113	0.74	436.9	52.55	165.0	37.8
2L	0.081	1.58	159.6	40.38	28.8	18.0
3L	0.079	1.73	137.7	38.13	23.2	16.9
4L	0.073	1.85	127.2	37.89	22.3	17.5
5L	0.067	2.02	118.8	37.00	19.9	16.7
6L	0.066	2.04	118.1	37.00	19.3	16.3
7L	0.066	2.04	117.8	37.00	19.1	16.2
7L+	0.064	2.03	130.0	40.00	20.7	15.9

Notes: Point estimates are the posterior mode, uncertainty is quantified by the posterior standard deviation (\hat{N} SE) and the coefficient of variation (\hat{N} CV). For the categorical SPIMs, \hat{n}^{cap} is the estimated number of captured individuals. Female estimates are listed first, followed by males.

posterior. This large number of posterior samples was likely far more than necessary; however, it allowed us to explore the behavior of the MCMC chains as the last uncertainty in n^{cap} was removed by adding genotype information (see Appendix S4). Because the hair sample subsampling process allowed for at most 1 sample per individual/trap/occasion, we used a Bernoulli observation model. The metrics of comparison were the point estimates (posterior modes), posterior standard deviations, and coefficients of variation ($100 \times$ posterior sd/posterior mode) for abundance as well as the posterior distributions of n^{cap} . Note, however, that the analyses adding the partial genotypes should not be expected to reproduce the null SCR point estimates, standard deviations, and number of individuals captured because they included additional data not used by the null SCR estimate.

Application—results

The categorical SPIM estimates (Table 3) for both sexes generally demonstrated the same patterns seen in the simulations. Abundance estimates with few identity categories were positively biased (relative to the SCR estimate) because the estimates of σ were negatively biased and/or the estimates of n^{cap} were positively biased. The magnitude of bias was larger for males, either partially or fully the result of a larger σ , estimated to be 2 times larger than females. As more loci were added, the categorical SPIM abundance estimates and their posterior standard deviations converged toward those of the SCR model, the posterior modes of n^{cap} converged toward the true number captured, and the posterior variance of n^{cap} converged toward 0 (Fig. 6). The coefficient of variation for all categorical SPIM estimates was lower than 0.20, except for the 1 locus male estimate.

The 1 locus female estimate was only 20% percent less precise than the full SCR estimate, as judged by the CV, with a positive bias (relative to the complete data estimate) of 8.6%. The 3 loci female estimate was substantially better—7.6% less precise than the full SCR estimate and positively biased by only 2.9%. The 4 loci female estimate was effectively equivalent to the full SCR estimate, with 3% less precision and minimal bias. The 5–7 loci female estimates were negligibly improved. Adding 2 partial genotype samples reduced the posterior standard deviation by 2.7% and coefficient of variation by 1.9%. One partial genotype sample was consistent with two different individuals in the full genotype data set, matching one with posterior probability of 0.452 and the other with posterior probability 0.544, leaving just a 0.004 posterior probability that this sample was from a new individual. The higher posterior match probability of 0.544 corresponded to the complete genotype individual that was captured at the same trap as the partial genotype sample. The other partial genotype matched an individual with 2 captures in the full genotype data set with posterior probability 0.997, leaving a probability of 0.003 that this was a new individual, and adding a high probability spatial recapture for this individual. These genotypes can be found in Table 4.

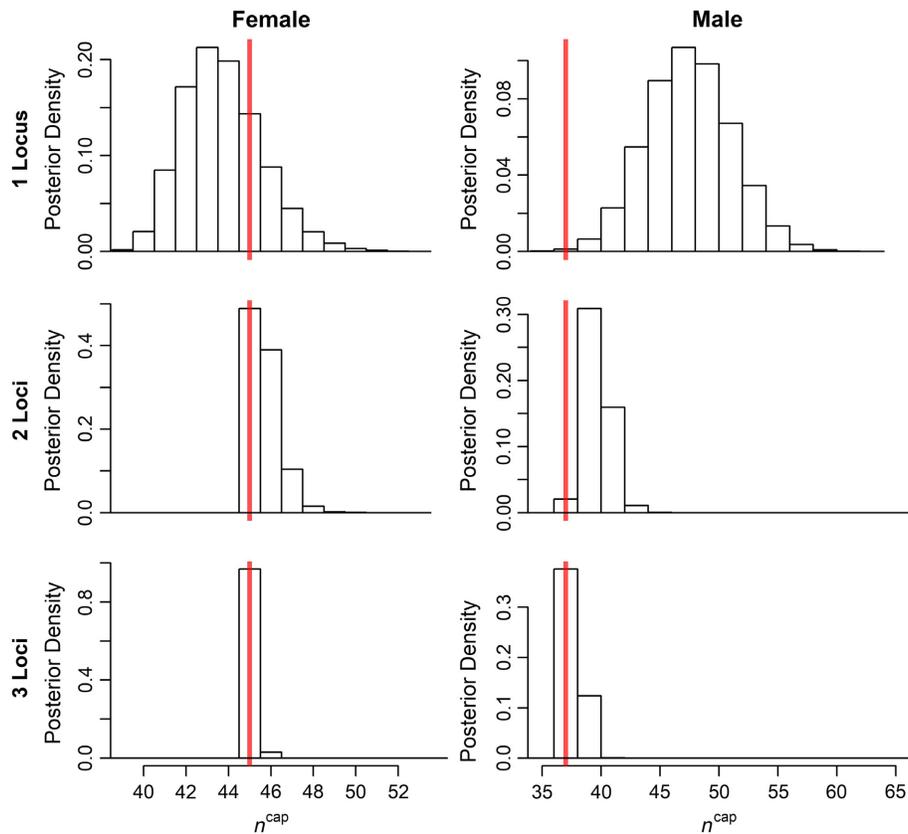


Fig. 6. Posterior distributions for the number of unique individuals captured (n^{cap}) for the Kentucky black bear female and male data set using 1, 2, and 3 loci. True values from full data sets are marked in red (45F, 37M).

The 1 locus male estimate was too biased (relative to the full SCR analysis) and imprecise to be of use, whereas the 2 and 3 loci male estimates had reasonable precision but perhaps too much positive bias to be useful. The 4 loci male estimate was 7.4% less precise than the full SCR estimate, with a positive bias of 7.6%. The 5 loci estimate was only negligibly less precise than the full SCR estimate, and adding the 6th and 7th loci improved precision negligibly. Adding 4 partial genotype samples modestly increased the abundance estimate, increased the posterior standard deviation (due to the larger point estimate), and decreased the coefficient of variation by 1.9%. The posterior probability that three of the partial genotype samples each came from separate individuals not represented in the full genotype data set was 1, while the fourth partial genotype sample matched with 8 other samples from 1 individual, each with posterior probability 1.

DISCUSSION

We developed a spatial capture–recapture model for categorically marked populations that uses any number of partially identifying categorical covariates to reduce the uncertainty in the individual identity of latent identity samples via three mechanisms. First, any samples that are inconsistent at any observed covariates are deterministically excluded from matching. Second, as the number of identity categories created by covariates increases and as the category level probabilities for each covariate become more equal, it is increasingly unlikely that more than one individual locally, and in the population, will have the same full categorical identity. Third, the spatial location of the latent identity samples and the estimated detection function scale parameter, σ , spatially restrict which samples matching at all observed covariates could have been

Table 4. Two partial genotype black bear samples and the full 7-loci genotype samples they had a positive posterior probability of matching with.

Sample no.	Genotype							X	Y
1	6	7	6	7	4	7	7	330461	4090636
2	6	7	6	7	4	7	7	329667	4088966
3	9	12	1	7	7	2	7	331260	4090333
4	9	12	1	7	7	2	7	330620	4089362
5	.	.	.	7	.	.	.	320620	4089362
6	7	8	1	1	2	8	2	333045	4091015
7	7	8	1	1	2	8	2	333045	4091015
8	.	.	1	1	2	.	.	331455	4091851

Notes: The first five samples correspond to one partial genotype (sample 5) and its matches (samples 1–4) and the final three samples correspond to a second partial genotype (sample 8) and its matches (samples 6 and 7). Each loci-level genotype is uniquely numbered, for example, the value 6 at loci 1 corresponds to a genotype with 158 repeats for allele 1 and 160 repeats for allele 2. X and Y are the coordinates where the samples were collected and missing loci-level genotypes are indicated with dots. The posterior probabilities that full genotype sample pairs 1–2, 3–4, and 6–7 each came from the same individual were 1. The posterior probabilities that partial genotype sample 5 came from the same individual as samples 1 and 2, the same individual as samples 3 and 4, or a new individual were 0.452, 0.544, and 0.004, respectively. The posterior probability that partial genotype sample 8 came from the same individual as samples 6 and 7 was 0.997, with a 0.003 probability this sample came from a new individual.

produced by the same individual. Thus, the categorical SPIM reduces uncertainty in the individual identity of latent identity samples by providing deterministic identity exclusions and reducing the uncertainty in probabilistic identity associations using both spatial and categorical covariate information. The categorical SPIM simulation and MCMC functions are maintained in the SPIM R package (Augustine 2018) and can also be found in Supplement S1.

A specific case of categorically marked populations that is of practical importance is that in which individuals are marked by multilocus genotypes. In this case, each locus of a genotype is a single categorical covariate and the categorical SPIM provides an incremental model for genotype uniqueness as the information about individual identity in the genotypes increases. Thus, the categorical SPIM is an alternative to using the P(ID) and P(sib) criteria currently used that allows for uncertainty in individual identity as might be the case when fewer loci are amplified than necessary to meet probability of identity criteria, which might occur in populations

with very low genetic diversity (McCarthy et al. 2009). The categorical SPIM also introduces the possibility of using fewer loci than necessary to meet probability of identity criteria by design, trading some certainty in individual identity for lower genotyping costs. Genotyping costs do not increase linearly with the number of loci and Puckett (2017) found that variability in the number of loci used in microsatellite studies explains very little variability in total project costs. However, there may be some cost savings if using fewer loci allows for the use of fewer multiplex panels, which explain a moderate amount of variability in total project costs (Puckett 2017).

Simulation scenarios A1–4 (Figs. 2 and 3; Appendix S3: Table S1) show the importance of population abundance, density given abundance, and σ for the accuracy and precision of the unmarked SCR and categorical SPIM estimators. Unmarked SCR estimates from populations with lower density given abundance and smaller σ s showed less bias (Appendix S3: Table S1), and unmarked SCR estimates were more precise for populations with higher abundances, lower density given abundance, and smaller σ s. More categorical identity groups were necessary to maximize precision when abundance was higher and σ larger; however, for the scenario that raised D without raising N , the majority of precision gains came from the addition of the first 4 identity categories. This scenario raising D without raising N demonstrates the importance of disentangling the relationship between N and D when assessing the performance of unmarked SCR and SCR models with latent or partial individual identities, more generally. Further, it suggests that categorical identity covariates are more effective in populations where uncertainty in individual identity is due to a large D relative to N , rather than a large σ . We propose that the uncertainty in n^{cap} , the number of individuals actually captured, is a good measure of the overall magnitude of uncertainty in individual identity for the unmarked SCR and categorical SPIM estimators. Figure 4 demonstrates that the precision of N increases as the precision of n^{cap} increases through the introduction of categorical identity covariates, though with diminishing returns (with precision quantified by the mean 95% credible interval width). The rate at which precision in N increases with increasing precision

in n^{cap} depends on the parameter values of each scenario (λ_0 , σ , D , and N given D), with greater precision gains in N to be had using categorical identity covariates in scenarios where there is greater uncertainty in individual identity in the unmarked SCR estimator.

The IDI correlated negatively with the precision and accuracy of n^{cap} estimates. For the scenarios that increased the IDI holding N fixed (A1 vs. A3, A2 vs. A4, A2 vs. A2b), the IDI also correlated negatively with the accuracy and precision of N estimates. In scenarios that increased D by increasing N (A1 vs. A2, A3 vs. A4), the increase in uncertainty in n^{cap} was outweighed by the decrease in data sparsity and N estimates were more precise and accurate. Our exploration of IDI values here is very limited and a larger simulation study is necessary to determine how well this index correlates with estimator performance and to what degree do scenarios with differing population density and σ values producing the same index value share the same estimator performance. Our results suggest that scenarios with differing D and σ values that produce the same IDI values will not necessarily produce the same precision in n^{cap} . We speculate that this is related to how the latent identity samples interact with different densities of activity centers in the SCR process model. Specifically, when D is large, there are necessarily more nearby individuals that a latent identity sample can be allocated to, which increases the uncertainty in individual identity. Finally, note that some of the largest values of the IDI may represent ecologically implausible or even impossible scenarios since home range size generally varies inversely with density (Efford et al. 2016).

We demonstrated that the unmarked SCR abundance estimator can be biased (Appendix S3: Table S1) when data are sparse and that this bias is magnified when density is higher for a fixed abundance and/or σ is larger. Further, none of the unmarked SCR scenarios we considered led to convergence rates $>90\%$, with only 52% of the MCMC chains converging for one scenario. Inspection of the MCMC chains indicated that σ was not identifiable for many realizations of the data for which convergence was not indicated. These patterns should also be seen in other SCR models with latent individual identities. For example, these patterns should be seen in SMR

when using few marked individuals and/or when detection data for the marked individuals are sparse as can occur when using natural marks and/or surveying low density populations. In fact, even though our simulation specifications were more challenging than those of Chandler and Royle (2013), many studies in practice use fewer than 81 traps that we considered and an expected 1.65 captures per individual (including the individuals captured 0 times) is likely optimistic for many sampling scenarios. The addition of marked individuals should allow for more reliable estimation for the unmarked population component because the marked individuals provide more information about the detection function parameters and reduce the number of samples with latent or partial individual identifications. Further, the use of individual-linked telemetry data and/or a marking process capture history in conjunction with generalized SMR (Whittington et al. 2018) improves the estimation of model parameters and thus, the reliability of density estimates. Still, until some practical guidelines can be established, relating estimator performance to abundance, density given abundance, and metrics of data sparsity (e.g., λ , K , number/spacing/extent of traps), we recommend researchers conduct simulations with unmarked SCR or SMR parameter values appropriate to their study design to determine if their study designs are sufficient to produce reliable estimates using these models. If not, the categorical SPIM offers a second route to remove bias and increase precision via reducing the uncertainty in individual identity, with the first route being the reduction of uncertainty in σ using telemetry data (Sollmann et al. 2013) and/or informative priors (Chandler and Royle 2013, Ramsey et al. 2015).

Simulation scenarios B1–3 (Table 2) demonstrate a proof of concept for using data sets where some or most of the full categorical identities are partial (missing values for the identity covariates) such as partial genotypes. With a 0.5 probability of successful amplification at each locus, no data sets produced enough full genotype samples that would have been usable in a model that required certain individual identification, but the categorical SPIM produced an estimate that was 95% as precise as the estimate where all loci were amplified with probability 1. Even assuming only 75% of the samples were usable, the estimate was 80% as precise.

Therefore, the categorical SPIM provides a way to use the partial identity information, such as partial genotype samples that are currently being discarded. We caution, however, that if partial genotype samples are more likely to have genotyping errors, the categorical SPIM needs to be extended to accommodate those errors, or perhaps a subset of the partial genotype samples could be deemed reliable through consultation with a wildlife geneticist. For example, allelic dropout could be ruled out by discarding any loci that were homozygous or using appropriate tests (e.g., available in MicroChecker, Van Oosterhout et al. 2004), or partial genotypes could perhaps be deemed reliable if they repeatedly produced the same partial genotype after a sufficient number of amplification attempts in a multi-tubes approach. Alternatively, adapting the allelic dropout model of Wright et al. (2009) to the categorical SPIM would be straightforward and Wang (2017) present a model for false alleles that could be adopted if sufficiently general. Both of these error processes would require the modeling of replicate amplification attempts rather than the consensus genotypes used here. We do note, however, that the density estimate precision that is possible for any set of partial genotypes will depend on the number of loci available, the underlying genetic diversity, and the percentage of genotype information remaining in the data set (in addition to the typical SCR model parameters). We expect qualitatively similar results to the simulations in Scenario B in typical data sets, but the expected results for a given population should be determined via simulation using the known or best-guess numbers and frequencies for the loci-level genotypes. Better results will be obtained for the same level of missing genotype information in populations with higher genetic diversity.

The bear genotype analysis demonstrates that the categorical SPIM estimator performed similarly on a real-world data set as it did for simulated data; however, the positive bias in the male estimates with few loci was larger than seen in the simulated data sets. We suspect individual heterogeneity in detection function parameters, particularly σ , may have been present in the male bear data set. If so, this could have led to poorer performance with few loci/identity categories, and the requirement of more loci/identity categories to remove bias and increase precision

than if there were no individual heterogeneity. The distribution of observed spatial recaptures in Fig. 5 does seem to suggest individual heterogeneity in σ for males, with one particular individual having a very long-distance spatial recapture and many individuals having no spatial recaptures. The samples for the individual with a long-distance spatial recapture were rarely combined into one individual until 3 loci were used and as $n^{\text{cap}} = 37$ (the correct number of captured males) became increasingly probable with the addition of more loci at which point, the estimate of σ converged upwards to the full SCR estimate. This behavior is consistent with the simulations where σ is large, but is more pronounced in this data set, which could be explained by individual heterogeneity in the detection function parameters. A second factor that tends to split the samples from this potentially large σ individual apart is that there were several traps between the two traps where this individual was captured and the categorical SPIM found it unlikely that this individual would not have been captured at these traps closer to its estimated activity center until enough categorical covariate information was available to make it even more unlikely that two individuals in the population had the same multilocus genotype. The second longest spatial recapture in the male data set spans a gap with no traps and required fewer loci to reliably link its samples together.

The posterior distributions of n^{cap} in Fig. 6 demonstrate what we believe is a source of the positive bias in the categorical SPIM estimator. With the addition of just 2 loci for females and 1 locus for males, all incorrect identity associations were ruled out by the genotype information and the spatial distribution of the samples. However, a 1 or 2 loci genotype is not sufficient to guarantee the local uniqueness of a genotype, leading to a situation in which samples cannot be erroneously combined into fewer individuals than produced them, but they can be erroneously split apart into more individuals than produced them. Thus, in these scenarios, n^{cap} can never take a value lower than the true value, but rather must always be equal to or larger than the true value. We believe the identity exclusions are removing the lower tail of the posterior distribution of n^{cap} that would be present in the unmarked SCR

estimator and introducing some positive bias, which can be removed by adding more categorical identity information and reducing the upper tail of the posterior of n^{cap} . In the simulations, this occurs with the addition of just a few identity categories; however, individual heterogeneity in detection function parameters as argued above may require more categorical identity information to remove the positive bias in n^{cap} and thus N .

This black bear genotype analysis also demonstrated the use of genotypes that are partial as a consequence of DNA amplification failure, with two caveats. First, there were very few usable partial genotypes because of the DNA amplification protocol used, in which samples at the same trap/occasion were subsequently genotyped until a full genotype was obtained. This process led to the partial genotypes matching the complete genotype individual at a particular trap/occasion with high probability because bears usually leave multiple hair samples in a hair snare, violating the Bernoulli observation process. Second, we assumed the partial genotype samples did not contain any genotyping errors. Three of the 6 partial genotype samples used matched other individuals in the population with high probability, but 3 partial genotypes had posterior probabilities of 1 that they were new individuals. These may have indeed been new individuals, or perhaps they did not match any other individuals because the partial genotypes were corrupted. Including partial genotypes in this manner needs to be done with caution and in consultation with a wildlife geneticist, or the categorical SPIM could be extended to accommodate genotyping errors (Wright et al. 2009). If partial genotypes, or even a subsample of the partial genotypes, can be deemed reliable, including them in the analysis can increase the precision of abundance and density estimates, especially if high probability spatial recaptures can be added, as was the case in the female bear data set.

We now will discuss four assumptions of particular importance that we make for the categorical SPIM. First, and perhaps most consequential, is that there is no individual heterogeneity in detection function parameters, or similarly, there is no transience in the activity centers during the time of the survey (Royle et al. 2016). Like unmarked SCR and spatial mark-resight (SMR),

the categorical SPIM uses the detection function likelihood to determine how likely it is that each latent identity sample came from each individual. As demonstrated in Appendix S1, of the detection function parameters, σ largely determines the degree to which samples from different individuals overlap in space. If there is individual heterogeneity in detection function parameters, especially σ , and the SCR model does not include this individual heterogeneity, an average λ_0 and/or σ will be estimated, except this average will be a biased estimate since the detected samples will disproportionately belong to the more detectable individuals. Still, the latent identity samples from the most detectable individuals will tend to be incorrectly split across two or more latent individuals because their true spatial distribution of samples will be deemed unlikely by the model based on the averaged λ_0 and/or σ values. Therefore, n^{cap} will tend to be biased high, introducing positive bias into N . The black bear example suggests that individual heterogeneity in detection function parameters (if it was present) can be overcome with increasing identity category information to correctly reproduce the true n^{cap} , but adding individual heterogeneity to the categorical SPIM detection model (as well as unmarked SCR and SMR) would be required to obtain appropriate abundance estimates. This should be investigated in the future, although we expect the introduction of individual heterogeneity in detection function parameters to drastically increase the uncertainty in individual identity and, thus, the utility of the categorical SPIM, unmarked SCR, and SMR for density estimation. A better strategy would be the use of covariate-specific detection function parameters for specific covariates such as the sex, if the subsets of the population that have different detection function parameters can be at least partially identified (Royle 2009, Sollmann et al. 2011).

The second assumption of note is that all possible category levels are known, which may not be the case for genotypes. There are two ways in which genotypes can be identified and enumerated for the categorical SPIM: identifying all observed genotypes at each locus, or identifying all implied genotypes at each locus based on the observed alleles at each locus. The latter method is certainly more thorough, but given sufficiently large data sets, any unobserved genotypes will

occur in the population very rarely. Regardless, there will always be the possibility that very low probability genotypes exist in the population that were not observed in the data set, and not accounting for these genotypes will introduce negative bias in abundance estimates (Wright et al. 2009), though the magnitude of bias will be small if the majority of genotypes with non-negligible frequencies are identified. Wright et al. (2009) raise the possibility of using independent, reference genotypes from the same population to improve abundance estimates. This information could aid in identifying all possible genotypes for each locus and could also be used to aid the estimation of genotype frequencies, which would be especially helpful for sparse data sets. The latter use would require that the reference genotypes were representative of the population subject to capture but did not contain any of the same individuals, which would violate independence (Wright et al. 2009).

The third assumption of note is that the full categorical identities are independent among individuals. When using genotypes, this assumption will be violated if genetic structure exists at the home range level as a result of relatedness, natural or anthropogenic impediments to movement, or other factors. Due to the importance of the spatial proximity of samples in the categorical SPIM, the combination of relatedness and philopatry (e.g., female black bears) can lead to a spatial correlation in genotypes across the landscape (Moyer et al. 2006). This could introduce negative bias if sufficiently strong because latent identity samples of nearby related individuals will be erroneously combined into one individual too often when using few loci data sets. We suspect spatial correlation in genotypes at the home range level will be weak in most spatially structured, non-isolated populations, and the categorical SPIM to be robust to this effect, although this might not be the case for species such as canids (*Canis sp.*) that travel in packs of highly related individuals. As the spatial uniformity of activity centers can be regarded as a weak prior for the distributions of individuals across the landscape (Royle et al. 2013), the spatial uniformity of category levels across the landscape is likely a similarly weak prior with the posterior able to deviate substantially from spatial uniformity, given sufficient data. Regardless, the

positive bias from individual heterogeneity in detection function parameters will likely outweigh any negative bias from the non-independence of genotypes, but this effect should be further investigated. The robustness of the categorical SPIM to assumption violations in typical genetic data sets can be further established by investigating how well it can reproduce estimates from full data sets, as we did in the black bear application, across many species and study areas.

The fourth assumption we will discuss is that the categorical covariates are independent of one another. When using genotypes, this is equivalent to linkage equilibrium, which can be tested for (Rousset and Raymond 1995), and results from a linkage disequilibrium test for the black bear data set used in our application did not detect nonrandom association of alleles (Murphy et al. 2016). When present, linkage disequilibrium causes pseudoreplication in genetic data sets (Selkoe and Toonen 2006), which we expect to introduce negative bias into the categorical SPIM estimator because there will be less variability than expected in the full categorical identities. Non-independence of the categorical covariates may occur for other types of data; for example, when using natural marks, body size and sex may not be independent. In this case, a composite covariate that combined body size and sex could be constructed, or the body size distribution could be estimated independently for each sex. In this latter case, body size would be informative about any missing sex covariates.

We added categorical covariates to the unmarked SCR model, but categorical identity covariates can also be added to SMR, 2-flank SPIM, and likely other types of specialized SPIMs that may not have been developed yet. Combining the categorical SPIM with SMR is especially appealing because it would allow all of the features currently available to improve density estimates over unmarked SCR to be combined into a single model that can accommodate a subset of marked individuals, individual-linked telemetry data, a marking process capture history, and categorical identity covariates. Further, this development will be required to address researcher-deployed categorical marks, such as colored ear tags and collars, because only the “marked” subset of individuals are categorically marked and the number of deployed marks of each type may

be known or perhaps the maximum number may be known if there is tag loss, requiring a more constrained MCMC algorithm for updating the latent individual identities. Another possible extension would be to allow continuous identity covariates, such as body size, perhaps measured with error, in a more general “covariate SPIM.” One continuous covariate of particular ecological relevance that is informative of identity is the time a sample was recorded, with samples collected closer together in time at the same location being more likely to have been produced by the same individual. However, making use of the time of sample deposition would require a model for animal movement and is probably less informative than spatial location, unless σ is large relative to the trapping array and movement speed relatively slow.

While our application and much of the discussion focuses on the use of microsatellite loci as categorical identity covariates, other observation systems provide categorical identity covariates that can be used by the categorical SPIM. Remote cameras, for example, sometimes provide individual sex, age class, color morph, natural markings, and/or other morphological features (Villafañe-Trujillo et al. 2018). Some species such as mustelids (Royle et al. 2011, Sirén et al. 2016, Villafañe-Trujillo et al. 2018) and Andean bears (*Tremarctos ornatus*; Molina et al. 2017) have markings on the chest, head, and neck, that can be used to classify unique individuals. We suspect categorical and/or continuous covariates can be extracted from these types of observations that would allow this identifying information to be used without having to assign a unique identity with certainty, which could be erroneous if two individuals shared the same markings, or to include the photographs that were not distinct enough to provide a unique identity. A second observation method with potentially abundant partially identifying information is bioacoustic monitoring where covariates such as note duration, bandwidth, and frequency can be extracted from the spectrograms of individual calls (Reby et al. 1999, McIntosh et al. 2015, Clink et al. 2017), which could improve density estimation for a large number of species that are difficult to detect by methods other than bioacoustics. Finally, in addition to the possible ways we envision the categorical SPIM may be used outlined in this paper, it may allow for better density estimation for

currently used observation systems we are not aware of, or may spur the adoption of new ways of categorically marking species for which unique marks are not currently available once researchers are aware that density can be estimated using the information contained in categorical marks. Unique marks will always be the most informative, but the categorical SPIM offers a middle ground between uniquely marked and unmarked populations.

ACKNOWLEDGMENTS

We thank the editor, associate editor, two anonymous reviewers, Lissette Waits, and Dana Morin for suggestions that led to the improvement of this paper. B.A. would also like to thank Chris Satter, Chris Rowe, and Robert Alonso for discussions that helped solidify the ideas developed in this paper. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

LITERATURE CITED

- Augustine, B. C. 2018. SPIM R package. <https://github.com/benaug/SPIM>
- Augustine, B. C., J. A. Royle, M. J. Kelly, C. B. Satter, R. S. Alonso, E. E. Boydston, and K. R. Crooks. 2018. Spatial capture-recapture with partial identity: an application to camera traps. *Annals of Applied Statistics* 11:67–95.
- Bonner, S. J., and J. Holmberg. 2013. Mark-recapture with multiple, non-invasive marks. *Biometrics* 69:766–775.
- Borchers, D. L., and M. Efford. 2008. Spatially explicit maximum likelihood methods for capture–recapture studies. *Biometrics* 64:377–385.
- Chandler, R. B., and J. D. Clark. 2014. Spatially explicit integrated population models. *Methods in Ecology and Evolution* 5:1351–1360.
- Chandler, R. B., and J. A. Royle. 2013. Spatially explicit models for inference about density in unmarked or partially marked populations. *Annals of Applied Statistics* 7:936–954.
- Clink, D. J., H. Bernard, M. C. Crofoot, and A. J. Marshall. 2017. Investigating individual vocal signatures and small-scale patterns of geographic variation in female bornean gibbon (*Hylobates muelleri*) great calls. *International Journal of Primatology* 38:656–671.
- Cowen, L., and C. J. Schwarz. 2006. The Jolly-Seber model with tag loss. *Biometrics* 62:699–705.
- Efford, M. 2004. Density estimation in live-trapping studies. *Oikos* 106:598–610.

- Efford, M., D. K. Dawson, Y. Jhala, and Q. Qureshi. 2016. Density-dependent home-range size revealed by spatially explicit capture–recapture. *Ecography* 39:676–688.
- Efford, M., and G. Mowat. 2014. Compensatory heterogeneity in spatially explicit capture–recapture data. *Ecology* 95:1341–1348.
- Fieberg, J., and C. O. Kochanny. 2005. Quantifying home-range overlap: the importance of the utilization distribution. *Journal of Wildlife Management* 69:1346–1359.
- Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. *Statistical science* 7:457–472.
- Knapp, S. M., B. A. Craig, and L. P. Waits. 2009. Incorporating genotyping error into non-invasive DNA-based mark–recapture population estimates. *Journal of Wildlife Management* 73:598–604.
- Laake, J., K. Burnham, and D. Anderson. 1993. Distance sampling: estimating abundance of biological populations. Chapman & Hall, London, UK.
- Lewis, J. S., K. A. Logan, M. W. Alldredge, L. L. Bailey, S. VandeWoude, and K. R. Crooks. 2015. The effects of urbanization on population density, occupancy, and detection probability of wild felids. *Ecological Applications* 25:1880–1895.
- McCarthy, T. M., L. P. Waits, and B. Mijiddorj. 2009. Status of the Gobi bear in Mongolia as determined by noninvasive genetic methods. *Ursus* 20: 30–38.
- McClintock, B. T., P. B. Conn, R. S. Alonso, and K. R. Crooks. 2013. Integrated modeling of bilateral photo-identification data in mark-recapture analyses. *Ecology* 94:1464–1471.
- McIntosh, B., K. M. Dudzinski, and E. Mercado III. 2015. Do dolphins' whistles reveal their age and sex. *Animal Behavior and Cognition* 2:313–333.
- Mills, L. S., J. J. Citta, K. P. Lair, M. K. Schwartz, and D. A. Tallmon. 2000. Estimating animal abundance using noninvasive DNA sampling: promise and pitfalls. *Ecological Applications* 10:283–294.
- Molina, S., A. K. Fuller, D. J. Morin, and J. A. Royle. 2017. Use of spatial capture–recapture to estimate density of Andean bears in northern Ecuador. *Ursus* 28:117–126.
- Moyer, M. A., J. W. McCown, T. H. Eason, and M. K. Oli. 2006. Does genetic relatedness influence space use pattern? A test on Florida black bears. *Journal of Mammalogy* 87:255–261.
- Murphy, S. M., J. J. Cox, B. C. Augustine, J. T. Hast, J. M. Guthrie, J. Wright, J. McDermott, S. C. Maehr, and J. H. Plaxico. 2016. Characterizing recolonization by a reintroduced bear population using genetic spatial capture–recapture. *Journal of Wildlife Management* 80:1390–1407.
- 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.
- Puckett, E. E. 2017. Variability in total project and per sample genotyping costs under varying study designs including with microsatellites or snps to answer conservation genetic questions. *Conservation Genetics Resources* 9:289–304.
- Ramsey, D. S., P. A. Caley, and A. Robley. 2015. Estimating population density from presence–absence data using a spatially explicit model. *Journal of Wildlife Management* 79:491–499.
- Reby, D., B. Cargnelutti, J. Joachim, and S. Aulagnier. 1999. Spectral acoustic structure of barking in roe deer (*capreolus capreolus*). sex-, age- and individual-related variations. *Comptes Rendus de l'Académie des Sciences-Series III-Sciences de la Vie* 322:271–279.
- Reich, B. J., and B. Gardner. 2014. A spatial capture–recapture model for territorial species. *Environmetrics* 25:630–637.
- Rousset, F., and M. Raymond. 1995. Testing heterozygote excess and deficiency. *Genetics* 140:1413–1419.
- Royle, J. A. 2009. Analysis of capture–recapture models with individual covariates using data augmentation. *Biometrics* 65:267–274.
- Royle, J. A., R. B. Chandler, R. Sollmann, and B. Gardner. 2013. Spatial capture–recapture. Academic Press, Cambridge, Massachusetts, USA.
- Royle, J. A., R. M. Dorazio, and W. A. Link. 2007. Analysis of multinomial models with unknown index using data augmentation. *Journal of Computational and Graphical Statistics* 16: 67–85.
- Royle, J. A., A. K. Fuller, and C. Sutherland. 2016. Spatial capture–recapture models allowing Markovian transience or dispersal. *Population Ecology* 58:53–62.
- Royle, J. A., A. J. Magoun, B. Gardner, P. Valkenburg, and R. E. Lowell. 2011. Density estimation in a wolverine population using spatial capture–recapture models. *Journal of wildlife management* 75:604–611.
- Selkoe, K. A., and R. J. Toonen. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9:615–629.
- Sirén, A. P., P. J. Pekins, P. L. Abdu, and M. J. Ducey. 2016. Identification and density estimation of American martens (*Martes americana*) using a novel camera-trap method. *Diversity* 8:3.

- Sollmann, R., M. M. Furtado, B. Gardner, H. Hofer, A. T. Jácomo, N. M. Tôres, 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, A. W. Parsons, J. J. Stocking, B. T. McClintock, T. R. Simons, K. H. Pollock, and A. F. O’Connell. 2013. A spatial mark–resight model augmented with telemetry data. *Ecology* 94:553–559.
- Van Oosterhout, C., W. F. Hutchinson, D. P. Wills, and P. Shipley. 2004. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Resources* 4:535–538.
- Villafañe-Trujillo, Á. J., C. A. López-González, and J. M. Kolowski. 2018. Throat patch variation in tayra (*Eira barbara*) and the potential for individual identification in the field. *Diversity* 10:7.
- Waits, L. P., G. Luikart, and P. Taberlet. 2001. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Molecular Ecology* 10:249–256.
- Wang, J. 2017. Estimating genotyping errors from genotype and reconstructed pedigree data. *Methods in Ecology and Evolution* 9:109–120.
- Whittington, J., M. Hebblewhite, and R. B. Chandler. 2018. Generalized spatial mark-resight models with an application to grizzly bears. *Journal of Applied Ecology* 55:157–168.
- Wright, J. A., R. J. Barker, M. R. Schofield, A. C. Frantz, A. E. Byrom, and D. M. Gleeson. 2009. Incorporating genotype uncertainty into mark–recapture-type models for estimating abundance using DNA samples. *Biometrics* 65:833–840.

SUPPORTING INFORMATION

Additional Supporting Information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/ecs2.2627/full>