

The importance of independence in unmarked spatial capture-recapture analysis

Catherine Sun¹ and Cole Burton²

¹Zambian Carnivore Programme

²The University of British Columbia

July 3, 2023

Abstract

Wildlife populations can be unmarked, meaning individuals lack visually distinguishing features for identification; populations may also exhibit non-independent movements, meaning individuals move together. For either unmarked or non-independent individuals, models based on spatial capture-recapture (SCR) approaches estimate abundance, density, and other population parameters critical for monitoring, management, and conservation. However, when individuals are both unmarked and non-independent, few model options exist. One approach has been to apply unmarked models and not address the non-independence despite unquantified impacts of overdispersion on bias, precision, and the ability to make robust ecological inferences. We conducted a simulation study to quantify the impact of non-independence on the performance of spatial count (SC) and spatial partial identity models (SPIM), two SCR-based unmarked modeling approaches, and used the performance of fully marked and independent SCR as a reference. We varied the levels of non-independence (aggregation and cohesion), detection probability, and the number of partial identity covariates used to resolve identities in SPIM estimation. We expected estimates of abundance and sigma (the spatial scale of individual movement) to be increasingly biased and less precise as aggregation and cohesion increased. Results showed that models indeed became less robust to increasing non-independence, especially for abundance, but importantly suggested that only SPIM could be reliably applied under low levels of cohesion when sufficient partial identity covariates are available. SC yielded consistently biased estimates with inflated precision that could not be corrected to nominal levels of coverage. SCR was the most robust across all combinations of aggregation and cohesion, as expected. We therefore advise against the use of SC models for estimating population parameters when individuals are known to be non-independent, caution that SPIM may be used under narrow ecological conditions, and encourage continued investigations into sampling design and methods development for populations of unmarked and non-independent individuals.

Introduction

Density estimation is a cornerstone of wildlife research and conservation. Spatially explicit capture-recapture methods (SCR), or capture-mark-recapture more broadly, are a well-established approach to estimating abundance and density (Efford, 2004; Royle & Young, 2008; Tourani, 2021). Traditionally in SCR, individuals are assumed to be identifiable and independent, but many species may not meet or adhere to these assumptions. Individuals may be unidentifiable (i.e., ‘unmarked’) for reasons including uninformative visual appearances from non-invasive sampling techniques such as camera traps, or the high costs of genetic approaches for obtaining identity (Cheng et al., 2017; Gilbert et al., 2021). The independence assumption may also be unmet because many species exhibit some degree of territoriality or sociality for at least some period of their lives (Prox & Farine, 2020). Accommodating common realities such as unmarked and non-independent individuals can be important for avoiding biases in density estimates. Approaches using the SCR framework have been developed that handle unidentifiability and non-independence separately (e.g., Augustine

et al., 2018; Chandler & Royle, 2013; Emmet et al., 2021; Royle, 2004), but there are as yet no SCR-based approaches that simultaneously accommodate both. Instead, models for unmarked populations are applied and the non-independence is acknowledged as a potential source of bias of unknown magnitude (e.g., Sun et al., 2022). However, the consequence of violating independence in SCR-based unmarked models should be quantified because precise and accurate density estimates are critical for determining population status and monitoring for conservation and management.

Identities enable the assignment of detections to individuals and the creation of individual-specific detection histories. These detection histories are informative about detection probability, and therefore help quantify the number of undetected individuals remaining in the population. When some detections cannot be assigned to individuals, they may be discarded (Tourani et al., 2020), incorporated along with detections of marked individuals (Sollmann et al. 2013), or modeled as having lost their identity with some probability (Jiménez et al., 2020). Occasional misidentifications may also be accounted for (McClintock et al., 2014; Morrison et al., 2011; Petersma et al., 2023; Rakhimberdiev et al., 2022; Yoshizaki et al., 2009). However, if all individuals are entirely unmarked or have few identifying marks such that their full identities are unknown, then spatial count (SC) and spatial partial identity models (SPIM) are two SCR-based modelling alternatives that can be used for density estimation. For wholly unmarked populations, SC models use the spatially correlated pattern of counts of detections across the sampling array, based on expectations of how single, independent individuals move on the landscape (Chandler & Royle, 2013). When partial identity marks are available, such as the spot or stripe pattern on one or both flanks of an animal, SPIM can probabilistically match detections to individuals and deterministically exclude non-matching detections from being from the same individual (Augustine et al., 2018, 2019). Precision with SPIM generally increases with more partial marks. SC and SPIM can be promising options for estimating populations with low density, but become less robust as density and individual movement increase space-use overlap among individuals and uncertainty in identities (Augustine et al., 2019; Chandler & Royle, 2013; Ruprecht et al., 2021).

In SCR-based models, individuals and their activity centers are modelled as outcomes of a spatial point process (Royle et al., 2013). Individuals may be distributed non-homogeneously due to habitat associations, but are usually assumed to not interact with each other. However, in nature, non-independence among individuals arises from interactions ranging from avoidance and territoriality to varying degrees of temporary or permanent grouping. We focus on grouping because it can also contribute to difficulty in determining identities. Grouping can provide safety from predators (Lehtonen & Jaatinen, 2016), increase foraging efficiency (McInnes et al., 2017), and improve chances of mating (Røed et al., 2002). Non-independence from grouping can be described with two components: aggregation and cohesion (Bischof et al., 2020). Aggregation describes the group size, so more individuals share the same activity center as aggregation increases. Cohesion is the degree to which group individuals move together, so detections of individuals become increasingly coordinated as members become more cohesive. Developing sampling designs that meet the independence assumption for group-living species such as wolves (*Canis lupis*), lions (*Panthera leo*), and some ungulates can be difficult, so there is growing interest in models that account for non-independence (e.g., see Emmet et al., 2021; Hickey & Sollmann, 2018; Reich & Gardner, 2014). Notably, Emmet et al. (2021) developed a SCR model using a cluster point process to estimate the size of cohesive groups as well as population abundance, while accounting for detection heterogeneity due to group size. By modeling non-independence among individuals, the cluster SCR model increases the understanding of a species' and population ecology, but still leaves a gap in the toolkit for density estimation because it cannot be used with unmarked individuals.

Due to the lack of SCR-based models for individuals that are both unmarked and non-independent, unmarked models are applied and non-independence is left unaddressed. Unmodeled correlation among detections can compromise model inferences, but so far has only been investigated for fully marked SCR models (Bischof et al., 2020; López-Bao et al., 2018; Moqanaki et al., 2021). Simulations show SCR to be relatively robust to low-to-moderate levels of aggregation and cohesion, but the overdispersion from correlated detections and resulting inflated precision and poor coverage of the confidence intervals around the true value may still lead to false inferences about population parameters. In unmarked density models, non-independence may contribute to the issues that SC and SPIM already face: higher local densities that increase overlap between

individual space-use and uncertainty about which detections originate from which individuals (Sun et al. 2022). Aggregation and especially cohesion may result in fewer apparent individuals with inflated detection probability and therefore underestimate density, while moderate levels of cohesion could inflate estimates of the spatial scale of individual movement (fewer individuals being detected over a larger spatial extent than otherwise expected). Thus, grouping would be expected to increase bias while reducing precision and coverage. Such biased estimates could be misleading about population density, insensitive to changes in population trajectory, and ultimately misinform conservation and management action.

We conducted a simulation study to assess the impacts of non-independence on SC and SPIM estimation and to identify any levels of tolerable aggregation and cohesion in order to inform further application of unmarked density models for ecological inference on species and populations with non-independent movements. For a template species, we considered boreal caribou (*Rangifer tarandus*) because it is an unmarked and grouping species of conservation concern in Canada (Festa-Bianchet et al., 2011; Hervieux et al., 2013). Many caribou populations are threatened and declining due to a combination of factors including habitat loss from natural resource exploration (Nagy-Reis et al., 2021), climate change (Barber et al., 2018; Bradley & Neufeld, 2012), and altered predator-prey dynamics (Burgar et al., 2019; Dickie et al., 2017; Hervieux et al., 2014). Tracking population responses to conservation interventions is critical to assessing populations across landscapes and reversing the fate of the species, but their wide distribution across remote areas makes it difficult to collect individual detection histories for SCR (McFarlane et al., 2020). As such, there has been interest using camera trapping and unmarked density models to estimate caribou density (Fisher et al., 2021; Sun et al., 2022).

A key challenge in monitoring caribou density with camera traps is that caribou violate requirements of independence and identifiability for traditional SCR approaches. Caribou form temporary multi-age herds or associations throughout the year (Body et al., 2015) and present challenges to visual identification, although they can be partially identified based on visible sex attributes and the number of antler points on both sexes. We conducted simulations using SCR, SC, and SPIM to estimate populations under varying levels of aggregation and cohesion. We also varied the number of partial identity marks to assess how non-independence interacts with the amount of available identity information in SPIM. Using SCR with known identities as a reference for comparison, we expected SCR to perform best and SPIM to outperform SC as aggregation and cohesion increase due to the availability of partial identities for assigning individuals to detections (Figure 1). Ultimately, the objective was to assess the reliability of SC and SPIM approaches for estimating densities of populations that are simultaneously unmarked and non-independent.

Materials and Methods

We simulated populations with varying degrees of grouping behavior, while setting $N = 140$ and movement scale of $\sigma = 3$ over a rectangular state-space of $60 \times 30 = 1800$ units² to mirror previous inferences and camera-trap sampling of the Algar herd of boreal woodland caribou in northern Alberta, Canada (part of the East Side Athabasca River population, Hervieux et al. 2013; for further details see (Beirne et al., 2021; Sun et al., 2022; Tattersall et al., 2020)). We generated different population scenarios by varying individual-level aggregation and cohesion, the number of available partial marks, and detection probability. Aggregation values included 1, 4, and 10 in order to range from single individuals (1) to the largest group size (10) observed in the Algar herd; cohesion values included 0, 0.3, 0.67, and 1 to range from none (0) to complete (1). This resulted in 9 grouping scenarios in which individuals could group together and move, ranging from a baseline independence scenario in which individuals are solitary and move independently (aggregation = 1; cohesion = 0) to a maximal non-independence scenario in which large groups of individuals always move together (aggregation = 10, cohesion=1) (Figure 1). When aggregation is 1, within-group cohesion is irrelevant.

For partial marks, we created 4 categorical independent marks or covariates: coat color (e.g., black, white, gray, brown, piebald; each with 0.20 population-level proportions), sex (equal 1:1 sex ratio), presence of GPS collar (4% yes / 96% no), and total number of points on left and right antlers (sequential integers from 0

to 17, each with a mean proportion of 0.049 and range from 0.004 – 0.180; Sun et al. 2022). Covariate values were randomly assigned to individuals according to these proportions. We considered 6 combinations in which identity covariates could be available, thereby spanning a range of expected probabilities of identity from 8% to 99% (Figure 2). We defined probability of identity as the probability that two randomly drawn individuals from the population would have the same set of partial identity covariate values. Finally, we considered populations to have either a low baseline detection probability ($p_0 = 0.05$) or higher ($p_0 = 0.20$). In total, we assessed 108 different population combinations of grouping scenarios, available identity information, and detection probability.

We simulated 100 datasets per population combination, each time subjecting individuals to 4 sampling occasions. Sampling occurred with a 5 x 15 grid of 75 sampling locations centered within the state-space to approximate camera trapping of the Algar caribou herd (Tattersall et al. 2020), with 3-unit spacing and a 9-unit buffer to the state-space edge (Figure 3). We adapted custom code from Bischof et al. 2020, which uses the *sim.pop* function in the ‘secr’ R package (Efford 2022) (Appendix 1). We summarized data by reporting the mean numbers of unique individuals detected, total detections, and traps with detections. To describe the potential consequence of non-independence, we used the counts of unique animals detected per sampling location to calculate Fletcher’s \hat{c} , a measure of over-dispersion that describes the degree to which the observed variance is greater than predicted by homogeneous density (Fletcher, 2012). The statistic is also a variance inflation factor as corrected variances may be obtained under low levels of overdispersion ($\hat{c} \lesssim 4$) multiplying variance by \hat{c} (Anderson et al. 1994).

Simulated detections were manipulated for density estimation in three ways. For SCR analysis, unique individual identities were retained and fit to a frequentist null SCR model using the ‘secr.fit’ function in the R package ‘secr’. For SC, identity information was removed to derive trap-and-occasion-specific counts of total detections that were fit to a Bayesian-formulated null SC model using the ‘nimble’ (de Valpine et al. 2022) package in R. For SPIM analyses, partial identity covariates were retained and associated with individual detections and then data were fit using the ‘SPIM’ package (Augustine et al. 2019) in R. For the Bayesian SC and SPIM, we used a slightly informative prior of gamma(24,8) on σ (Burgar et al., 2019), resulting in a mean (± 1 standard deviation, SD) $\sigma = 3 \pm 0.6$ units. Data augmentation was set to either $M = 400$ or $M = 600$ as necessary to prevent truncation of the posterior distribution in N as determined with visual inspection. We used an initial burn-in of 5,000 iterations and a subsequent 45,000 iterations split into 3 chains each with 15,000 iterations. Model results were assessed for convergence with the r-hat statistic and removed if r-hat > 1.1 (Gelman & Rubin, 1992). SC and SPIM were run using Microsoft Azure virtual machines [Standard D64s v4 (64 vcpus, 256 GiB memory)] to reduce computation time.

Parameters of interest were abundance (N) and the spatial scale parameter for individual movement (σ). We report the mean (for SCR) and median (for SC and SPIM) estimates of N and σ , and similar metrics as used in Bischof et al. (2020) to summarize model performance across the iterations of each scenario: relative bias, coefficient of variation, and coverage. Relative bias describes the extent of over- or under-estimation and is calculated as

$$RB = \frac{\hat{\theta} - \theta}{\theta}$$

where θ is the true value of the parameter of interest (in this case the value of N or σ used in simulations) and $\hat{\theta}$ is the point estimate. Coefficient of variation is a measure of precision and is calculated as

$$CV = \frac{SD(\hat{\theta})}{\hat{\theta}}$$

where SD is the standard deviation. For conservation and management purposes, $CV < 20\%$ is preferred (Morin et al., 2022). Coverage was calculated as the proportion of simulations that include θ within its 95% credible intervals.

$$Coverage = \frac{\sum_{i=1}^{100} I(95CI_{i|\lambda\omega\epsilon\rho} \leq \theta \leq 95CI_{i|\nu\pi\pi\epsilon\rho})}{100}$$

english100

Results

Aggregation and cohesion influenced the numbers of individuals detected, total detections, and traps with detections. The mean number of individuals detected and total detections changed little with aggregation and cohesion: when $p_0 = 0.05$, approximately one-third of the population was consistently detected (min – max: 43 – 47, of $N = 140$) a total of 65 – 71 times. However, the variance in these measures increased as aggregation and cohesion increased (Figure 4, Appendix 2 Table 1). For example, the standard deviation (SD) in individuals detected increased three-fold from $SD = 5$ to $SD = 18$ when changing from the independence (aggregation = 1, cohesion = 0) to maximal non-independence (aggregation = 10, cohesion = 1); SD in total detections similarly increased 3.5-fold from $SD = 9$ to $SD = 32$. In contrast, the number of traps with detections decreased with increasing aggregation and cohesion, from a maximum of 43 ± 5 (mean \pm 1 SD) under the independence scenario to a mean of only 7 ± 3 traps under maximal non-independence when $p_0 = 0.05$. The patterns of fewer traps with detections and greater variance in unique individuals and total detections with increasing non-independence also occurred with $p_0 = 0.20$. (Appendix 2 Table 1, Appendix 2 Figure 1).

Aggregation and cohesion increased over-dispersion in the number of unique individuals detected per sampling location (Figure 5). The overdispersion factor \hat{c} ranged from a minimum of 1.8 ± 0.2 to a maximum of 10.4 ± 0.8 when $p_0 = 0.05$. Generally, $\hat{c} < 4$ when cohesion was 0 and at low to intermediate values of cohesion and aggregation. Over-dispersion further increased with $p_0 = 0.20$ (Appendix 2 Figure 2).

Abundance

SCR estimates of abundance (\hat{N}) were consistently the least biased across all grouping scenarios compared to SC and SPIM (Figure 6, Appendix 2 Figure 3). Mean bias was negligible ($|\text{RB}| < 6\%$) under low cohesion (0, 0.3) and increased to a maximum of $\text{RB} = 16 - 22\%$ when animals exhibited maximal aggregation (10) and high cohesion (0.67, 1) with $p_0 = 0.05$ (Appendix 2 Table 2). Estimates were relatively precise with mean CV [?][?] 20% but suffered when both aggregation and cohesion were high, whereby the variance in CV increased to 23% (Appendix 2 Table 3). Coverage was nominal under the independence scenario (97%) and averaged $80 \pm 19\%$ across all grouping scenarios, but decreased to a minimum of 38% at maximal aggregation and cohesion when $p_0 = 0.05$ (Figure 7, Appendix 2 Table 4). After correcting the variance for overdispersion, coverage averaged $97 \pm 5\%$ and decreased to 82% only under maximal aggregation and cohesion (Appendix 2 Figure 4). Patterns with $p_0 = 0.20$ were similar but with less bias, more precision, and greater coverage (Appendix 2 Tables 2, 3, 4 and Appendix 2 Figure 4).

In comparison, SC consistently produced biased \hat{N} , as expected (Figure 6). Mean bias was negative, underestimating N with inflated precision when there was any aggregation or cohesion (mean RB ranged $-34 \pm 20\%$ to $-88 \pm 6\%$; Appendix 2 Table 2). The amount of bias ($|\text{RB}|$) increased with both aggregation and cohesion, although bias was less affected by increasing cohesion at greater levels of aggregations (i.e., 10). Precision increased with detection probability and with aggregation under intermediate cohesion (0.3, 0.67), with the greatest mean precision ($\text{CV} = 29 \pm 12\%$) under maximal non-independence when $p_0 = 0.20$ (Appendix 2 Table 3). Coverage with SC averaged $29 \pm 32\%$ across all grouping scenarios and detection probabilities (Figure 7). Coverage reached nominal levels only under the independence scenario, and decreased quickly with aggregation and cohesion to a minimum of 0 (Appendix 2 Table 4). Even after correction with \hat{c} , mean coverage improved to $44 \pm 33\%$ (Appendix 2 Figure 4).

SPIM estimation of N performed differently depending on the levels of aggregation and cohesion, and number of partial identity covariates. Estimates were generally overestimated, with the exception of high aggregation (10) and low cohesion (0, 0.3) in which median N was underestimated (Figure 6). Otherwise, as with SC, bias increased with aggregation and cohesion, reaching a maximum of $\text{RB} = 185 \pm 96\%$ under maximal non-independence with only the collar covariate when $p_0 = 0.20$ (Appendix 2 Table 2). Precision also increased with aggregation and cohesion; mean CV with only the collar covariate when $p_0 = 0.05$ decreased from $\text{CV} = 59 \pm 13\%$ under the independence scenario by 30% to $29 \pm 12\%$ under maximal non-independence

(Appendix 2 Table 3). Having more partial identity covariates, or more specifically a greater probability of identity, helped to moderate bias and inflated precision from aggregation and cohesion; estimates became less biased and more stable in precision with more partial identity covariates, such that use of all 4 partial ID covariates when $p_0 = 0.20$ consistently resulting in negligible mean bias (RB: 2 - 6%) and stable mean precision of CV = 9 - 10% across all combinations of aggregation and cohesion (Appendix 2 Tables 2, 3). Average coverage across all grouping scenarios was 62 +- 23% (Figure 7), with a minimum of 45% under maximal non-independence even when all 4 spatial identity covariates. Coverage only reached nominal or near-nominal levels when cohesion was low and probability of identity was greater than 54% (i.e., at least 2 partial identity covariates if collar was not one of the 2 covariates; Appendix 2 Table 4). Coverage improved when variance was corrected with \hat{c} , averaging 89 +- 11% across all scenarios (Appendix 2 Figure 4).

Sigma

SCR estimates of sigma ($\hat{\sigma}$) exhibited similar but less pronounced patterns of bias, precision, and coverage compared to \hat{N} (Figures 6, 7). SCR $\hat{\sigma}$ was generally unbiased (mean |RB| [?] 8%) and precise (mean CV [?] 15%) across all levels of aggregation and cohesion (Appendix 2 Tables 5, 6), although the variance in bias and precision increased with aggregation and cohesion. Coverage was nominal under the independence scenario and less than 90% only when cohesion was high (0.67, 1) (Figure 7, Appendix 2 Table 7). After correcting variance for overdispersion, coverage was an average 96 +- 7% and only decreased below 90% (to 82%) with maximal non-independence when $p_0 = 0.05$ (Appendix 2 Figure 4).

For SC, $\hat{\sigma}$ was less biased (range of mean |RB|: 0 +- 5% - 70 +- 4%) and more precise (CV < 20%) compared to \hat{N} , but followed similar patterns of more bias and inflated precision with increasing cohesion and aggregation (Figure 6, Appendix 2 Tables 5, 6). Coverage of SC $\hat{\sigma}$ was nominal or near-nominal when cohesion = 0 and decreased dramatically when cohesion and aggregation increased (Figure 7, Appendix 2 Table 7). Patterns in coverage did not improve with variance correction and reached an average of 62 +- 40% (Appendix 2 Figure 4).

Finally, SPIM $\hat{\sigma}$ had low bias (|RB|: 1-12%) and accordingly acceptable precision (CV [?] 20%) across all levels of aggregation, cohesion, and detection probabilities (Figure 6, Appendix 2 Tables 5, 6). As expected, estimates exhibited greater bias (or variance in bias), more precision, and less coverage with increasing cohesion and aggregation. However, as with SPIM \hat{N} , more partial identity covariates helped to moderate these effects. Notably, at high levels of cohesion (0.67, 1), SPIM $\hat{\sigma}$ had greater coverage than SCR $\hat{\sigma}$ (Figure 7). Once corrected for overdispersion, coverage for SPIM $\hat{\sigma}$ improved from an average of 83 +- 12% across grouping scenarios and detection probabilities to a mean of 93 +- 7% (Figure 7, Appendix 2 Figure 2, Table 7).

Discussion

Unmarked density estimation is a challenging but worthy endeavor due to the many species and populations that are unmarked and require information about population status for informed conservation. However, when individuals are both unmarked and non-independent, our simulations showed that two primary SCR-based unmarked approaches, SC and SPIM, perform poorly. Population abundance (N), and the spatial scale parameter (σ) to a lesser extent, were often estimated with increasing bias and imprecision as aggregation and cohesion increased. The reference SCR model remained the least biased and most precise despite non-independence (Bischof et al., 2020; López-Bao et al., 2018), while aggregation and cohesion affected SC and SPIM differently. In particular, we identified no conditions of non-independence under which SC could reliably estimate parameters. In contrast, SPIM was more robust to violating the independence assumption and could estimate N and σ under some conditions with low cohesion. Unmarked density models should be applied selectively and carefully to populations exhibiting non-independence among individuals.

Our simulations strongly suggest that SC models should not be used when individuals aggregate or co-

here, and are likely only appropriate for independent species or periods when individuals are independent. Abundance was severely underestimated with misleading precision at all levels of aggregation and cohesion, consistent with our expectations but opposite to the overestimation found in previous work with independent populations (Chandler & Royle, 2013; Ruprecht et al., 2021). Perhaps fortunately, the CV never reached 20% or lower to afford confident application of the biased estimates for informing conservation or management action (Morin et al., 2022). Estimates of σ were slightly better but still biased, rendering SC unfit for inferences about movement and space use when individuals exhibit grouping behaviors. These issues could not be resolved by applying \hat{c} as a variance inflation factor, as it failed to recover sufficient coverage for parameter estimates. Indeed, \hat{c} was ineffective because the relative variance (variance for a given grouping scenario divided by the variance for the independence scenario) far exceeded \hat{c} , in contrast to the SCR simulations of Bischof et al. (2020). Furthermore, overdispersion when $\hat{c} > 4$ is ideally addressed with additional model structure (Anderson et al., 1994; Fletcher, 2012). Realistically, other measures of overdispersion would be more appropriate for SC models because empirical or non-simulated SC data do not have the counts of unique individuals used for the formulation of \hat{c} in these simulations. We also recognize the potential to consider other distributions to model the SC data, such as the negative binomial, to allow for greater variance (Lindén & Mäntyniemi, 2011) rather than the standard Poisson distribution used here.

While SPIM estimation also suffered with increasing non-independence, SPIM may be acceptable when cohesion is minimal provided there are sufficient partial identity covariates. Applying SPIM under larger values of cohesion resulted in overestimated populations, which Sun et al. (2022) also found with empirical and simulated SPIM modeling under independence. Indeed, the variance in RB and CV (across the iterations for any single grouping scenario) were large even with high probability of identity (i.e., 99% from all 4 partial identity covariates) and $p_0 = 0.20$, despite small mean RB and CV values across grouping scenarios. The underestimated population observed with low cohesion and high aggregation was consistent with our expectations, while the otherwise mostly positive bias observed with SPIM was consistent with Augustine et al. (2019) – who also reported positive bias under independence due to high local densities that increased identity uncertainty and ultimately assigned detections to more individuals than actually detected. The change in pattern of bias with increasing cohesion, from under- to over-estimation, with few partial identity covariates is unexplained and motivates further examination. Broadly, the increase in bias with increasing non-independence is consistent with increasing uncertainty in individual identification due to overlapping space use.

The decreased bias and greater precision from using more SPIM partial identity covariates highlight the importance of maximizing identity resolution by increasing the numbers of partial identity covariates and their categorical values. More combinations of partial identity covariates could have been considered because probabilities of identity ranging 53 - 90% were not assessed. However, increasing the number partial identity covariates is not a panacea because in practice the number of partial identity covariates for a species or population will be limited (motivating the use of SPIM in the first place), and estimates were still biased under high aggregation and cohesion even when identities were completely known (SCR). Additional investigation of SPIM is nonetheless warranted to further explore the ecological and sampling conditions in which SPIM models could still be reliably used. Future simulations should explore more probabilities of identity, values of aggregation and cohesion, and different spacings of sampling locations. Density estimation for unmarked populations expected to have low levels of aggregation and cohesion may benefit from sampling spaced closer than generally recommended under independence in order to increase spatial variation among group members (Augustine et al., 2019; Chandler & Royle, 2013).

In addition to SCR-based unmarked approaches, other approaches instead rely on information about movement speed (or its inverse: staying time within a viewshed). Two such unmarked models, the random encounter model (Rowcliffe et al., 2008) and random encounter and staying time (Nakashima et al., 2018), have demonstrated potential to be robust to non-independence but also poor model fit that would hamper interpretation of spatial or temporal patterns (Hayashi & Iijima, 2022). Two other non-SCR approaches, space to event (Moeller et al., 2018) and time in front of camera models (Becker et al., 2022), have shown either bias or a lack of concordance with other approaches when non-independence is added (Fisher et al.,

2021; Hayashi & Iijima, 2022). Hayashi and Iijima (2022) also showed that bias in the non-SCR approaches is sensitive to the choice of statistical distribution for modeling count data, including Poisson, zero-inflated Poisson, exponential, and negative binomial. Comparing SCR-based and non-SCR-based approaches under identical simulating conditions would be informative for helping wildlife researchers and managers choose the appropriate sampling and modeling approaches for obtaining unbiased population estimates. Indeed, we recommend researchers run a small simulation study using expected population conditions and the anticipated sampling designs to confirm that robust density estimation is feasible before implementing in the field.

Continued efforts to identify optimal sampling designs and estimation approaches are essential for robust wildlife populations monitoring. If SCR-based unmarked approaches are to be tenable for non-independent populations, models must have procedures to account for – and ideally attempt to incorporate – cohesion and aggregation. In addition to using statistical distributions that account for overdispersion, some efforts to estimate density of group-living species have corrected group densities with average group size (e.g., Mattioli et al., 2018). However, including a mechanism for how aggregation and cohesion impact the detection process and lead to overdispersion would further advance the understanding of the focal species’ ecology. The clustered SCR model of Emmett et al. (2021) begins this work and presents some potential for modification to fit unmarked populations for ecological inference, but caution is necessary. The current formulation requires knowledge about group membership for detected individuals, which is difficult to obtain when identities are unknown unless groups are being monitored intensively, in which case detections with unknown identities are usually the exception and can be dealt with in ways previously described. Group membership might instead be estimated, informed by spatial patterns and supported with informative priors about range of movement and group size. A model that uses the spatial pattern of detections and any available partial identities to estimate the distribution of individuals and groups when identities and group membership remain unknown would be of great appeal and value. However, attempting to develop a model that weaves a tapestry of gold from relatively little information risks neglecting careful sampling design for statistical machismo (Gimenez et al., 2014). As a result, we encourage the continued development of unmarked models but recommend that applications of current SCR-based approaches to unmarked population estimation recognize that they are generally unsuitable for most cases of non-independence.

References

- Anderson, D., Burnham, K., & White, G. (1994). AIC Model Selection in Overdispersed Capture-Recapture Data. *Ecology* , 75 , 1780. <https://doi.org/10.2307/1939637>
- Augustine, B. C., Royle, J. A., Kelly, M. J., Satter, C. B., Alonso, R. S., Boydston, E. E., & Crooks, K. R. (2018). Spatial capture–recapture with partial identity: An application to camera traps. *Annals of Applied Statistics* , 12 (1), 67–95. <https://doi.org/10.1214/17-AOAS1091>
- Augustine, B. C., Royle, J. A., Murphy, S. M., Chandler, R. B., Cox, J. J., & Kelly, M. J. (2019). Spatial capture–recapture for categorically marked populations with an application to genetic capture–recapture. *Ecosphere* , 10 (4), e02627. <https://doi.org/10.1002/ecs2.2627>
- Barber, Q. E., Parisien, M.-A., Whitman, E., Stralberg, D., Johnson, C. J., St-Laurent, M.-H., DeLancey, E. R., Price, D. T., Arseneault, D., Wang, X., & Flannigan, M. D. (2018). Potential impacts of climate change on the habitat of boreal woodland caribou. *Ecosphere* , 9 (10), e02472. <https://doi.org/10.1002/ecs2.2472>
- Becker, M., Huggard, D. J., Dickie, M., Warbington, C., Schieck, J., Herdman, E., Serrouya, R., & Boutin, S. (2022). Applying and testing a novel method to estimate animal density from motion-triggered cameras. *Ecosphere* , 13 (4), e4005. <https://doi.org/10.1002/ecs2.4005>
- Beirne, C., Sun, C., Tattersall, E. R., Bugar, J. M., Fisher, J. T., & Burton, A. C. (2021). Multispecies modelling reveals potential for habitat restoration to re-establish boreal vertebrate community dynamics. *Journal of Applied Ecology* , 58 (12), 2821–2832. <https://doi.org/10.1111/1365-2664.14020>

- Bischof, R., Dupont, P., Milleret, C., Chipperfield, J., & Royle, J. A. (2020). Consequences of ignoring group association in spatial capture–recapture analysis. *Wildlife Biology* , 2020 (1). <https://doi.org/10.2981/wlb.00649>
- Body, G., Weladji, R. B., Holand, O., & Nieminen, M. (2015). Fission-fusion group dynamics in reindeer reveal an increase of cohesiveness at the beginning of the peak rut. *Acta Ethologica* ,18 (2), 101–110. <https://doi.org/10.1007/s10211-014-0190-8>
- Bradley, M., & Neufeld, L. (2012). Climate and management interact to explain the decline of woodland caribou (*Rangifer tarandus caribou*) in Jasper National Park. *Rangifer* , 183–191. <https://doi.org/10.7557/2.32.2.2268>
- Burgar, J. M., Burton, A. C., & Fisher, J. T. (2019). The importance of considering multiple interacting species for conservation of species at risk. *Conservation Biology* , 33 (3), 709–715. <https://doi.org/10.1111/cobi.13233>
- Chandler, R. B., & Royle, J. A. (2013). Spatially explicit models for inference about density in unmarked or partially marked populations. *The Annals of Applied Statistics* , 7 (2), 936–954. <https://doi.org/10.1214/12-AOAS610>
- Cheng, E., Hodges, K. E., Sollmann, R., & Mills, L. S. (2017). Genetic sampling for estimating density of common species. *Ecology and Evolution* , 7 (16), 6210–6219. <https://doi.org/10.1002/ece3.3137>
- de Valpine P, Paciorek C, Turek D, Michaud N, Anderson-Bergman C, Obermeyer F, Wehrhahn Cortes C, Rodriguez A, Temple Lang D, Paganin S (2022). NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling.. doi: 10.5281/zenodo.1211190(URL: <https://doi.org/10.5281/zenodo.1211190>), R package version 0.13.1, <URL:<https://cran.r-project.org/package=nimble>>.
- Dickie, M., Serrouya, R., McNay, R. S., & Boutin, S. (2017). Faster and farther: Wolf movement on linear features and implications for hunting behaviour. *Journal of Applied Ecology* , 54 (1), 253–263. <https://doi.org/10.1111/1365-2664.12732>
- Efford, M. (2004). Density Estimation in Live-Trapping Studies. *Oikos* , 106 (3), 598–610. <https://doi.org/10.2307/3548382>
- Efford, M. 2022. secr: Spatiay explicit capture-recapture models. R package version 4.5.8. url: <https://CRAN.R-project.org/package=scr>.
- Emmet, R. L., Augustine, B. C., Abrahms, B., Rich, L. N., & Gardner, B. (2021). A spatial capture–recapture model for group-living species. *Ecology* , 00 (00), e03576. <https://doi.org/10.1002/ecy.3576>
- Festa-Bianchet, M., Ray, J. C., Boutin, S., Cote, S. D., & Gunn, A. (2011). Conservation of caribou (*Rangifer tarandus*) in Canada: An uncertain future. *Canadian Journal of Zoology* , 89 (5), 419–434. <https://doi.org/10.1139/z11-025>
- Fisher, J., Burgar, J., Dickie, M., Burton, C., & Serrouya, R. (2021). *Density Estimates of Unmarked Large Mammals at Camera Traps Vary among Models, Species, and Years, Signalling Importance of Model Assumptions* [Preprint]. Authorea. <https://doi.org/10.22541/au.163768980.02146968/v1>
- Fletcher, D. J. (2012). Estimating overdispersion when fitting a generalized linear model to sparse data. *Biometrika* ,99 (1), 230–237. <https://doi.org/10.1093/biomet/asr083>
- Gelman, A., & Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science* , 7 (4), 457–472. <https://doi.org/10.1214/ss/1177011136>
- Gilbert, N. A., Clare, J. D. J., Stenglein, J. L., & Zuckerberg, B. (2021). Abundance estimation of unmarked animals based on camera-trap data. *Conservation Biology* , 35 (1), 88–100. <https://doi.org/10.1111/cobi.13517>

- Jimenez, O., Buckland, S. T., Morgan, B. J. T., Bez, N., Bertrand, S., Choquet, R., Dray, S., Etienne, M.-P., Fewster, R., Gosselin, F., Merigot, B., Monestiez, P., Morales, J. M., Mortier, F., Munoz, F., Ovaskainen, O., Pavoine, S., Pradel, R., Schurr, F. M., ... Rexstad, E. (2014). Statistical ecology comes of age. *Biology Letters* . <https://doi.org/10.1098/rsbl.2014.0698>
- Hayashi, K., & Iijima, H. (2022). Density estimation of non-independent unmarked animals from camera traps. *Ecological Modelling* , *472* , 110100. <https://doi.org/10.1016/j.ecolmodel.2022.110100>
- Hervieux, D., Hebblewhite, M., DeCesare, N. J., Russell, M., Smith, K., Robertson, S., & Boutin, S. (2013). Widespread declines in woodland caribou (*Rangifer tarandus caribou*) continue in Alberta. *Canadian Journal of Zoology* , *91* (12), 872–882. <https://doi.org/10.1139/cjz-2013-0123>
- Hervieux, D., Hebblewhite, M., Stepnisky, D., Bacon, M., & Boutin, S. (2014). Managing wolves (*Canis lupus*) to recover threatened woodland caribou (*Rangifer tarandus caribou*) in Alberta. *Canadian Journal of Zoology* , *92* (12), 1029–1037. <https://doi.org/10.1139/cjz-2014-0142>
- Hickey, J. R., & Sollmann, R. (2018). A new mark-recapture approach for abundance estimation of social species. *PLOS ONE* , *13* (12), e0208726. <https://doi.org/10.1371/journal.pone.0208726>
- Jimenez, J., C. Augustine, B., Linden, D. W., B. Chandler, R., & Royle, J. A. (2020). Spatial capture–recapture with random thinning for unidentified encounters. *Ecology and Evolution* , *11* (3), 1187–1198. <https://doi.org/10.1002/ece3.7091>
- Lehtonen, J., & Jaatinen, K. (2016). Safety in numbers: The dilution effect and other drivers of group life in the face of danger. *Behavioral Ecology and Sociobiology* , *70* (4), 449–458. <https://doi.org/10.1007/s00265-016-2075-5>
- Linden, A., & Mantyniemi, S. (2011). Using the negative binomial distribution to model overdispersion in ecological count data. *Ecology* , *92* (7), 1414–1421. <https://doi.org/10.1890/10-1831.1>
- Lopez-Bao, J. V., Godinho, R., Pacheco, C., Lema, F. J., Garcia, E., Llana, L., Palacios, V., & Jimenez, J. (2018). Toward reliable population estimates of wolves by combining spatial capture-recapture models and non-invasive DNA monitoring. *Scientific Reports* , *8* (1), Article 1. <https://doi.org/10.1038/s41598-018-20675-9>
- Mattioli, L., Canu, A., Passilongo, D., Scandura, M., & Apollonio, M. (2018). Estimation of pack density in grey wolf (*Canis lupus*) by applying spatially explicit capture-recapture models to camera trap data supported by genetic monitoring. *Frontiers in Zoology* , *15* (1), 38. <https://doi.org/10.1186/s12983-018-0281-x>
- McClintock, B. T., Hill, J. M., Fritz, L., Chumbley, K., Luxa, K., & Diefenbach, D. R. (2014). Mark-resight abundance estimation under incomplete identification of marked individuals. *Methods in Ecology and Evolution* , *5* (12), 1294–1304. <https://doi.org/10.1111/2041-210X.12140>
- McFarlane, S., Manseau, M., Steenweg, R., Hervieux, D., Hegel, T., Slater, S., & Wilson, P. J. (2020). An assessment of sampling designs using SCR analyses to estimate abundance of boreal caribou. *Ecology and Evolution* , *10* (20), 11631–11642. <https://doi.org/10.1002/ece3.6797>
- McInnes, A. M., McGeorge, C., Ginsberg, S., Pichegru, L., & Pistorius, P. A. (2017). Group foraging increases foraging efficiency in a piscivorous diver, the African penguin. *Royal Society Open Science* , *4* (9), 170918. <https://doi.org/10.1098/rsos.170918>
- Moeller, A. K., Lukacs, P. M., & Horne, J. S. (2018). Three novel methods to estimate abundance of unmarked animals using remote cameras. *Ecosphere* , *9* (8), e02331. <https://doi.org/10.1002/ecs2.2331>
- Moqanaki, E. M., Milleret, C., Tourani, M., Dupont, P., & Bischof, R. (2021). Consequences of ignoring variable and spatially autocorrelated detection probability in spatial capture-recapture. *Landscape Ecology* . <https://doi.org/10.1007/s10980-021-01283-x>

- Morin, D. J., Boulanger, J., Bischof, R., Lee, D. C., Ngoprasert, D., Fuller, A. K., McLellan, B., Steinmetz, R., Sharma, S., Garshelis, D., Gopalswamy, A., Nawaz, M. A., & Karanth, U. (2022). Comparison of methods for estimating density and population trends for low-density Asian bears. *Global Ecology and Conservation* , 35 , e02058. <https://doi.org/10.1016/j.gecco.2022.e02058>
- Morrison, T. A., Yoshizaki, J., Nichols, J. D., & Bolger, D. T. (2011). Estimating survival in photographic capture–recapture studies: Overcoming misidentification error. *Methods in Ecology and Evolution* , 2 (5), 454–463. <https://doi.org/10.1111/j.2041-210X.2011.00106.x>
- Nagy-Reis, M., Dickie, M., Calvert, A. M., Hebblewhite, M., Hervieux, D., Seip, D. R., Gilbert, S. L., Venter, O., DeMars, C., Boutin, S., & Serrouya, R. (2021). Habitat loss accelerates for the endangered woodland caribou in western Canada. *Conservation Science and Practice* , 3 (7), e437. <https://doi.org/10.1111/csp2.437>
- Nakashima, Y., Fukasawa, K., & Samejima, H. (2018). Estimating animal density without individual recognition using information derivable exclusively from camera traps. *Journal of Applied Ecology* , 55 (2), 735–744. <https://doi.org/10.1111/1365-2664.13059>
- Petersma, F. T., Thomas, L., Thode, A. M., Harris, D., Marques, T. A., Cheoo, G. V., & Kim, K. H. (2023). *Accommodating false positives within acoustic spatial capture-recapture, with variable source levels, noisy bearings and an inhomogeneous spatial density* (arXiv:2207.09343). arXiv. <https://doi.org/10.48550/arXiv.2207.09343>
- Prox, L., & Farine, D. (2020). A framework for conceptualizing dimensions of social organization in mammals. *Ecology and Evolution* , 10 (2), 791–807. <https://doi.org/10.1002/ece3.5936>
- Rakhimberdiev, E., Karagicheva, J., Saveliev, A., Loonstra, A. H. J., Verhoeven, M. A., Hooijmeijer, J. C. E. W., Schaub, M., & Piersma, T. (2022). Misidentification errors in reencounters result in biased estimates of survival probability from CJS models: Evidence and a solution using the robust design. *Methods in Ecology and Evolution* , 13 (5), 1106–1118. <https://doi.org/10.1111/2041-210X.13825>
- Reich, B. J., & Gardner, B. (2014). A spatial capture-recapture model for territorial species. *Environmetrics* , 25 (8), 630–637. <https://doi.org/10.1002/env.2317>
- Roed, K. H., Holand, O., Smith, M. E., Gjostein, H., Kumpula, J., & Nieminen, M. (2002). Reproductive success in reindeer males in a herd with varying sex ratio. *Molecular Ecology* , 11 (7), 1239–1243. <https://doi.org/10.1046/j.1365-294X.2002.01509.x>
- Rowcliffe, J. M., Field, J., Turvey, S. T., & Carbone, C. (2008). Estimating animal density using camera traps without the need for individual recognition. *Journal of Applied Ecology* , 45 (4), 1228–1236. <https://doi.org/10.1111/j.1365-2664.2008.01473.x>
- Royle, J. A. (2004). N-Mixture Models for Estimating Population Size from Spatially Replicated Counts. *Biometrics* , 60 (1), 108–115. <https://doi.org/10.1111/j.0006-341X.2004.00142.x>
- Royle, J. A., Chandler, R. B., Sollmann, R., & Gardner, B. (2013). *Spatial Capture-Recapture* (1 edition). Academic Press.
- Royle, J. A., & Young, K. V. (2008). A Hierarchical Model for Spatial Capture-Recapture Data. *Ecology* , 89 (8), 2281–2289. <https://doi.org/10.2307/27650753>
- Ruprecht, J. S., Eriksson, C. E., Forrester, T. D., Clark, D. A., Wisdom, M. J., Rowland, M. M., Johnson, B. K., & Levi, T. (2021). Evaluating and integrating spatial capture–recapture models with data of variable individual identifiability. *Ecological Applications* , 31 (7), e02405. <https://doi.org/10.1002/eap.2405>
- Sun, C., Burgar, J. M., Fisher, J. T., & Burton, A. C. (2022). A cautionary tale comparing spatial count and partial identity models for estimating densities of threatened and unmarked populations. *Global Ecology and Conservation* , 38 , e02268. <https://doi.org/10.1016/j.gecco.2022.e02268>

Tattersall, E. R., Burgar, J. M., Fisher, J. T., & Burton, A. C. (2020). Mammal seismic line use varies with restoration: Applying habitat restoration to species at risk conservation in a working landscape. *Biological Conservation* , 241 , 108295. <https://doi.org/10.1016/j.biocon.2019.108295>

Tourani, M. (2021). A review of spatial capture–recapture: Ecological insights, limitations, and prospects. *Ecology and Evolution* , 12 (1), e8468. <https://doi.org/10.1002/ece3.8468>

Tourani, M., Dupont, P., Nawaz, M. A., & Bischof, R. (2020). Multiple observation processes in spatial capture–recapture models: How much do we gain? *Ecology* , 101 (7), e03030. <https://doi.org/10.1002/ecy.3030>

Yoshizaki, J., Pollock, K. H., Brownie, C., & Webster, R. A. (2009). Modeling misidentification errors in capture–recapture studies using photographic identification of evolving marks. *Ecology* , 90 (1), 3–9. <https://doi.org/10.1890/08-0304.1>

Figure captions

Figure 1. For spatial count (SC) and spatial partial identity models (SPIM), yellow cells show expectations of model performance across different values of individual-level aggregation and cohesion (i.e., non-independence). Red crosses and green check marks indicate patterns identified in simulations. Values of aggregation and cohesion considered in simulations are shown in parentheses.

Figure 2. Expected probabilities that two draws from a population with 4 partial identity covariates (coat color, sex, GPS collar, antler points, in various population proportions) would be distinct from each other, given an increasing number of those covariates used for SPIM estimation. Closed circles show all combinations, while open squares highlight those evaluated in simulations. Colored lines indicate the mean probability of identity for a given partial identity covariate for each specified number of total covariates used.

Figure 3. Example patterns of detections for $N = 140$ individuals (colored circles) for select combinations of aggregation and cohesion with detection probabilities of $p_0 = 0.05$ (top) and $p_0 = 0.20$ (bottom), illustrating how non-independence affects the distribution of individuals and their detections across the statespace (rectangle). Lines connect individuals to sampling locations (crosses) where detected. The sampling array has 3-unit spacing with a 9-unit buffer to the statespace edge.

Figure . Numbers of unique individuals detected (top), total detections (middle), and traps with detections (bottom) from simulating detection data for populations of $N = 140$ under 9 different non-independence scenarios with detection probability $p_0 = 0.05$, generated by varying aggregation (group size) and cohesion.

Figure . Calculated overdispersion factor (\hat{c}) of simulated detection data with detection probability $p_0 = 0.05$ when individuals exhibit non-independence through varying levels of aggregation and cohesion.

Figure . Point estimates of abundance, N (top row), and spatial scale of individual movement, σ (bottom row), with SCR, SC, and SPIM across a range of aggregation and cohesion values and combinations of partial identity covariates, with detection probability $p_0 = 0.05$. Dotted lines indicate truth: $N = 140$, and $\sigma = 3$.

Figure 7. Coverage of estimates of abundance, N (top row), and spatial scale of individual movement, σ (bottom row), for SCR, SC, and SPIM, across a range of aggregation and cohesion values and combinations of partial identity covariates, with detection probability $p_0 = 0.05$. Dashed line indicates 95% coverage.

Cohesion	High (0.67, 1)	Small groups of coordinated individuals		Large groups of coordinated individuals		
		<i>More bias</i>	SC ✓	SPIM ✓	<i>Most bias</i>	SC ✓
	<i>More precision</i>	SC ✗	SPIM ✓	<i>Most precision</i>	SC ✓	SPIM ✓
	Low (0, 0.3)	Small groups of mostly independent individuals		Large groups of mostly independent individuals		
<i>Less biased</i>		SC ✓	SPIM ✓	<i>Some bias</i>	SC ✓	SPIM ✓
	<i>Less precision</i>	SC ✗	SPIM ✓	<i>Some im/precision</i>	SC ✗	SPIM ✓
	Low (4)			High (10)		
	Aggregation					







