Ghostbusting - Reducing bias due to identification errors in spatial capture-recapture histories

Abinand Reddy Kodi[1], Jasmin S. Howard[1], David L. Borchers[1], Hannah Worthington[1], Justine Shanti Alexander[2], Purevjav Lkhagvajav[3], Gantulga Bayandonoi[4], Munkhtogtokh Ochirjav[4], Sergelen Erdenebaatar[4], Choidogjamts Byambasuren[3], Nyamzav Battulga[4], Örjan Johansson[5,2], and Koustubh Sharma[2]

[1]Centre for Research into Ecological and Environmental Modelling, School of Mathematics and Statistics, University of St Andrews, The Observatory, St Andrews, Fife, KY16 9LZ, Scotland

[2]Snow Leopard Trust, 4649 Sunnyside Avenue North, Suite 325, Seattle, WA 98103, USA

[3]Snow Leopard Conservation Foundation, 602, Khan uul tower, 3th khoroo, Khan Uul district, Mongolia

[4]World Wide Fund for Nature - Mongolia, Amar Str. 4, Internom office, 400, Ulaanbaatar 14200, Mongolia

[5]Grimsö Wildlife Research Station, Department of Ecology, Swedish University of Agricultural Sciences, SE-73091, Riddarhyttan, Sweden.

Abstract

1. Identifying individuals is key to estimating population sizes by spatial capture-recapture, but identification errors are sometimes made. The most common identification error is the failure to recognise a previously detected individual, thus creating a "ghost" (Johansson et al., 2020). This results in positively biased abundance estimates.

2. Ghosts typically manifest as single detection individuals ("singletons") in the capture history. To deal with ghosts, we develop a spatial capture-recapture method conditioned on at least K detections. The standard spatial capture-recapture (SCR) model is the special case of $K = 1$. Ghosts can mostly be excluded by fitting a model with $K = 2$ (SCR-2).

3. We investigated the effect of 'singleton' ghosts on the estimates by simulation. The SCR method increas-

ingly over-estimated abundance with increasing percentage of ghosts, with positive bias even when only 10% of the detected individuals were ghosts, and bias between 30% and 58% when 30% were ghosts. Estimates from the SCR-2 method remained unbiased in the presence of ghosts, at the cost of some precision. The mean squared error of the estimated abundance from the SCR-2 method was lower than that from the SCR method even with only 10% ghosts, except in the lowest density and lowest encounter rate scenario, and was lower in all scenarios with 20% or more ghosts. We also applied our method to capture histories from camera trap surveys of snow leopards (*Panthera uncia*) at 2 sites from Mongolia and find that the SCR method produced higher abundance estimates at both sites.

4. Capture histories are susceptible to errors when generated from passive detectors such as camera traps and genetic samples. The SCR-2 method can remove bias from ghost capture histories, at the cost of some loss in precision. At identification error rates that result in at least 10% of the detected individuals as ghost capture histories, we recommend using the SCR-2 method, except perhaps when the sample size is very low. We recommend using it in all cases when there may be more than 10% ghosts or surveys with a large number of single detection capture histories.

Keywords: camera-trapping, misidentification, population estimation, singletons, spatial capture-recapture

1 INTRODUCTION

Estimating abundance and tracking changes in population sizes are imperative to inform conservation efforts for wildlife species. Spatial capture-recapture (SCR) methods (M. Efford, 2004; Borchers & Efford, 2008; Royle & Young, 2008) are widely used to estimate the abundance and distribution of species that can be individually identified. These methods rely on correctly identifying re-detections of animals and may give biased estimates if errors are made in animal identification. We focus here with non-invasive surveys in which animals are not marked by surveyors but are uniquely identified from some innate features such as body markings in photographs in the case of a camera trap survey, or genetic fingerprinting from scat samples in the case of a scat survey.

Identification errors on camera trap surveys can occur due to poor observation conditions, poor quality photographs of individuals, or variation in markings over time (Stevick et al., 2001). The level of experience of the people who perform the identification may also be a factor. For example, observers Gibbon et al. (2015) found that those experienced in working with mountain bongos (*Tragelaphus eurycerus isaaci*) made identification errors in around one-sixth of cases, while inexperienced observers made errors in about one-fifth of cases when asked to determine whether pairs of photographs were of the same individual.

Like camera trapping, Genotyping methods are known to be prone to errors. This led to the development of $M_{t,\alpha}$ models for non-spatial capture-recapture (CR) surveys that explicitly model error rates (Lukacs & Burnham, 2005). However, in this model, identification errors could only lead to artificially introduced individuals which Yoshizaki (2007) termed "ghosts". Models to handle ghosts were further developed by Link et al. (2010) and extended by Bonner et al. (2016) to incorporate more complex misidentification error processes. However, it has been difficult to apply these models to real data, and the resulting population size estimates suffer from poor precision unless there are high capture probabilities or many capture occasions (Vale et al., 2014). Wright et al. (2009) proposed that the identification error rate may be estimated directly by genotyping all samples at least twice, but this is not always cost-effective, so manual error correction prior to modelling may be preferable (Fewster, 2017). Yoshizaki (2007) proposed and developed a model dropping all single captures and conditioning on at least 2 captures in the CR likelihood. Conn et al. (2011) generalise this method to account for transient animals in a survey.

Identification errors from photographs can be more complex than simply creating ghosts. Different types of misidentification errors can have different impacts on the structure of a capture history (the record of detections for each individual), and hence on abundance estimates. As outlined in Johansson et al. (2020), *splitting errors* occur when a photographic capture of an individual is falsely not matched to that individual and is not matched to any other individual, splitting the capture history of that individual in two and creating a "ghost" individual. A *combination error* occurs when all captures of an individual are falsely matched to a different individual, removing its entire capture history and reducing the (recorded) number of individuals detected. A *shifting error* is when a splitting error, reducing the capture history of one individual, is combined with a combination error from another individual; the capture event moves between individuals without changing the number of individuals detected. An *exclusion error* occurs when photographs are discarded despite containing sufficient information to make an identification, reducing the number of captures in their capture history. Splitting errors will lead to positively biased abundance estimates, combination errors will lead to negative bias. Johansson et al. (2020) found abundance estimates to be largely unaffected by shifting errors, but may be biased due to exclusion errors if these do not also cause a reduction in the number of animals within capture histories, or if there is heterogeneity in the probability of exclusions among individuals.

If marking patterns do not change considerably over the survey period and vary sufficiently between individuals, falsely identifying captures of two different individuals as being from the same individual (as in combination errors and some shifting errors) should be rare (Morrison et al., 2011). Van Horn et al. (2014) found that observers who continued to make errors after training were twice as likely to falsely identify two photos of the same bear as two different Andean bears (*Tremarctos ornatus*) than falsely match two different individuals. Stevick et al. (2001) found that observers did not falsely match different individuals, they only produced false negative errors in which a photograph and a biopsy were falsely identified as being of two different humpback whales (*Megaptera novaeangliae*), mainly due to poor photographic quality. Johansson et al. (2020) found that the impact of splitting errors far outweighed the impact of combination errors in snow leopard (*Panthera uncia*) CR surveys, resulting in systematically positively biased abundance estimates - by one-third on average. They further noted that splitting errors, typically resulting in singleton ghosts, were the most common of the misidentification errors.

SCR methods have been developed to deal with situations in which no identities are known (Chandler & Royle, 2013), in which only a fraction of the population is marked and these animals are identifiable without errors (Sollmann et al., 2013; Whittington et al., 2018), and in which single flanks of individuals are observed without errors, but the two flanks of the same individual cannot be confidently matched (Augustine et al., 2018). Jiménez et al. (2021) implemented a random thinning model to account for exclusion errors - detections that are discarded as a result of not being able to identify them either as an existing or as a new individual. There currently is not a method that can effectively account for ghosts in SCR.

Here we extend the method proposed by Yoshizaki (2007) for non-spatial CR, to deal with ghosts in SCR. We do this by developing a likelihood and maximum likelihood estimator for capture histories that involve at least K detections of individuals. We fit the SCR model conditioned on at least 2 captures (SCR-2) to exclude ghosts, the rationale being that ghosts are predominantly singletons and so can be excluded by using $K = 2$.

We first investigate the properties of the SCR-2 method by simulations. We investigate the bias and mean squared error of SCR and SCR-2 abundance estimates caused by ghost capture histories, and use the simulation results to identify when using the SCR-2 method would be advisable. We also implement the SCR-2 method on 2 camera trap surveys of snow leopards (*Panthera uncia*) and compare estimates from the SCR and SCR-2 models.

2 MATERIALS AND METHODS

2.1 Model formulation

Consider a camera trapping survey where detectors record counts of individuals encountered over a survey period. We assume the individuals $i = 1, 2, \ldots, N$ move around their activity centres and model the spatial distribution of these activity centres as a realisation of a Poisson point process over the survey region $\boldsymbol{X} \in \mathbb{R}^2$. Let $\boldsymbol{s} \in \boldsymbol{X}$ denote the location of any activity centre and $\boldsymbol{s}_i$ denote the activity centre of individual i. The intensity of the point process that generates activity centres is $D(\boldsymbol{z_s}, \boldsymbol{\phi})$, or $D(\boldsymbol{s}; \boldsymbol{\phi})$ for brevity, where $\boldsymbol{z_s}$ is a vector of covariate values at $\boldsymbol{s}$ and $\boldsymbol{\phi}$ is the parameter vector of the point process.

Suppose J traps are placed across the survey region. Let $\boldsymbol{x}_j \in \boldsymbol{X}$ denote the location of the j^{th} trap. Let $\omega_{ij} \in \mathbb{N}$ denote the number of times individual i was detected at trap j, $\boldsymbol{\omega}_i = (\omega_{i1}, \ldots, \omega_{iJ})$ be the vector of the i^{th} individual's capture history across all J traps, and $\omega_{i\cdot} = \sum_{j=1}^{J} \omega_{ij}$ be the total number of times individual i was detected. Let $n^{(K)}$ be the number of individuals detected at least K times in the survey (so that $n^{(1)}$ is the total number of individuals detected). $\boldsymbol{\Omega}^{(K)} = (\boldsymbol{\omega}_i;\ i = 1, \ldots, n^{(K)})$ denotes the set of capture histories for the $n^{(K)}$ individuals with $\omega_{i\cdot} \geq K$, i.e. the capture histories with at least K detections. The survey can be further split into multiple days or occasions if encounter rates varied across occasions, but we consider a single-occasion model for simplicity.

We define the expected number of encounters at trap j located at $\boldsymbol{x}_j$ for an individual i with an activity centre at $\boldsymbol{s}_i$ to be $\lambda_j(\boldsymbol{s}_i; \boldsymbol{\theta})$. $\boldsymbol{\theta}$ is the vector of encounter function parameters; for readability, we usually omit the $\boldsymbol{\theta}$ term below. The number of encounters at trap j of animal i with activity centre at $\boldsymbol{s}_i$ is $\omega_{ij} | \boldsymbol{s}_i \sim \text{Po}(\lambda_j(\boldsymbol{s}_i))$. The expected number of encounters $\lambda_j(\boldsymbol{s}_i)$ can be modelled by various functions based on how we expect the animal to move around its activity centre $\boldsymbol{s}_i$. Most commonly, $\lambda_j(\boldsymbol{s}_i)$ is modelled as a function of $d_{i,j}$, the distance between individual i's activity centre $\boldsymbol{s}_i$ and trap j, and the encounter parameters $(\lambda_0, \sigma) \in \boldsymbol{\theta}$, where λ_0 is the encounter rate at the individual's activity centre and σ is the range parameter. The expected number of encounters across all J traps of an individual with an activity centre at $\boldsymbol{s}_i$ is $\lambda_\cdot(\boldsymbol{s}_i) = \sum_{j=1}^{J} \lambda_j(\boldsymbol{s}_i)$. Via the additive property of the Poisson distribution $\omega_{i\cdot}|\boldsymbol{s}_i \sim \text{Po}(\lambda_\cdot(\boldsymbol{s}_i))$.

The probability of detecting animal i with an activity centre at $\boldsymbol{s}_i$ at least K times is $p_\cdot^{(K)}(\boldsymbol{s}_i; \boldsymbol{\theta}) = 1 - \sum_{k=0}^{K-1} p(\omega_{i\cdot} = k|\boldsymbol{s}_i)$. Based on our assumption that the number of detections follows a Poisson distribution:

$$p_\cdot^{(K)}(\boldsymbol{s}_i; \boldsymbol{\theta}) = 1 - \sum_{k=0}^{K-1} \frac{\lambda_\cdot(\boldsymbol{s}_i)^k \exp[-\lambda_\cdot(\boldsymbol{s}_i)]}{k!}. \tag{1}$$

This allows us to compute the probability of the capture history of an individual conditioned on at least K encounters.

Under our assumptions, the number of individuals detected at least K times, $n^{(K)}$, is a Poisson random variable with parameter:

$$\Lambda^{(K)} = \int_{\boldsymbol{X}} D(\boldsymbol{s}; \boldsymbol{\phi}) p_\cdot^{(K)}(\boldsymbol{s}; \boldsymbol{\theta}) d\boldsymbol{s}. \tag{2}$$

Adapting the likelihood function of Borchers and Efford (2008) to deal only with animals that were detected at least K times, we have the likelihood function:

$$L(\boldsymbol{\theta}, \boldsymbol{\phi}|\boldsymbol{\Omega}^{(K)}, n^{(K)}) = \frac{e^{-\Lambda^{(K)}}}{n^{(K)}!} \prod_{i=1}^{n^{(K)}} \int_{\boldsymbol{X}} \prod_{j=1}^{J} p(\omega_{ij}|\boldsymbol{s}; \boldsymbol{\theta}) D(\boldsymbol{s}; \boldsymbol{\phi}) d\boldsymbol{s}, \tag{3}$$

The likelihood function of Borchers and Efford (2008) is recovered by setting $K = 1$. From the maximum likelihood estimate of the density, the expected abundance, $E[N]$ can be derived by

$$E[N] = \int_{\boldsymbol{X}} D(\boldsymbol{s}; \widehat{\boldsymbol{\phi}}) d\boldsymbol{s}. \tag{4}$$

In common with other SCR inference methods, we discretise $\boldsymbol{X}$ into a mesh. Likelihood maximisation was performed using the R programming language (R Core Team, 2022). Standard SCR models were fit using the secr package (M. Efford, 2022).

2.2 Simulations

The performance of the methods was assessed using a simulation study. In particular, we investigated the impact of ghosts in the capture histories on the bias of the parameter estimates for both models. Reducing sample size by dropping single detections in the SCR-2 method is bound to increase the variance of the estimated parameters. To understand the trade-off between bias and variance between the two models, we check the mean squared errors of the parameter estimates to identify when using the SCR-2 method would be advisable over using standard SCR.

We simulated 100 survey regions as a 56 x 56 square grid with each grid cell having an area of $2500m^2$.

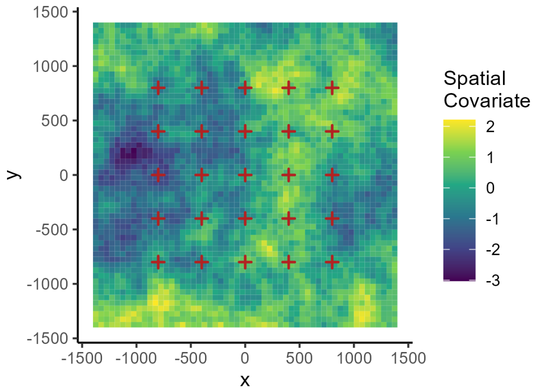

Figure 1: One realisation of the simulated survey landscape. The colour in each grid cell represents the value of the spatial covariate at that cell. Traps are shown by the red crosses.

The surveys consisted of 25 traps deployed in a grid layout with a spacing of $400m$. Each of the grid cells in a landscape was assigned a spatial covariate value drawn from a standard normal distribution and the entire grid was smoothed to introduce spatial correlation. To reduce any site-specific effects, each of the 100 landscapes was simulated with different realisations of the spatial covariate. One realisation of a landscape with the trap layout is shown in Figure 1.

Populations within each survey were simulated as an inhomogeneous point process, modelled with a log-linear relationship to the simulated spatial covariate: $D(\boldsymbol{s}) = \exp(\beta_0 + \beta_1 z_{\boldsymbol{s}})$; where $z_{\boldsymbol{s}}$ is the value of the spatial covariate at point $\boldsymbol{s}$ and β_0, β_1 are the intercept and slope of the log-linear model. A high abundance($\beta_0 = 0.1$) and a low abundance ($\beta_0 = 0.03$) scenario were simulated for each of the landscapes. The slope was fixed at $\beta_1 = 0.1$ in both cases. Sets of capture histories were simulated from these populations using a half normal encounter rate, $\lambda_j(\boldsymbol{s_i}) = \lambda_0 e^{-d_{i,j}^2/2\sigma^2}$. $\lambda_0 = 1$ and $\lambda_0 = 2$ were used to generate capture histories with low and high encounter rates respectively. σ was fixed at 300.

Ghosts were then introduced to each set of the simulated capture histories by randomly splitting detections of individuals encountered more than once and creating new single-detection ghost capture histories. More than one detection of any individual could become a ghost as long as there remained at least one detection in the original individual's capture history. The number of ghosts introduced was proportional to the total number of individuals detected in each simulation. For each simulated set of capture histories, 3 new sets of capture histories were created with 10%, 20% and 30% of the total number of individuals truly detected introduced as ghosts.

This resulted in a total of 1600 sets of capture histories. The SCR and the SCR-2 models were fit to each dataset. We investigated relative bias (RB) and the mean squared error (MSE) = Bias^2 + Variance on the estimated parameters λ_0, σ and $E[N]$ for both models.

2.3 Case Study: Snow Leopards in Mongolia

The population assessment of the World's snow leopards (PAWS) is a multinational survey programme that aims to produce a robust estimate of the distribution and abundance of the world's snow leopards. Mongolia has

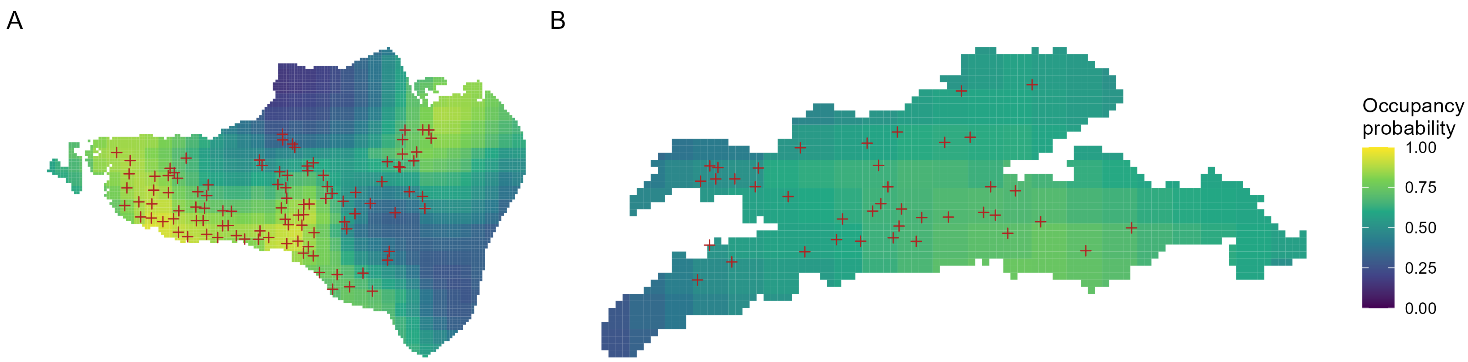

Figure 2: Survey regions of Munkhkhairkhan(A) and Tost(B). Background colours show the occupancy probabilities estimated by Bayandonoi et al. (2021), used to model densities. Configurations of traps across the two survey sites are represented by red crosses. The survey regions are not to scale.

been one of the leaders in conducting surveys for this programme. Bayandonoi et al. (2021) obtained estimates of the distribution of snow leopards across the country using occupancy methods (MacKenzie et al., 2017) to bolster abundance estimates from existing camera trap surveys. These occupancy estimates were used to design further surveys that spanned the estimated distribution range of snow leopards in Mongolia. For this case study, we verify data from 2 of these sites: Munkhkhairkhan and Tost. Tost has been a long-term monitoring site for snow leopards, with annual camera trap surveys being conducted for over a decade. Snow leopard individuals in Tost have been followed across years, with several cats that have been captured and collared, and we expect fewer misidentification errors in the capture histories as compared to Munkhkhairkhan which was surveyed for the first time.

Munkhkhairkhan was surveyed between May and September 2017 where 102 traps were deployed over an area of $9,841km^2$. The Tost survey for PAWS was conducted between September 2019 and January 2020. 40 traps were set up over an area of $1,902km^2$ in Tost. Survey regions and trap layouts can be seen in Figure 2.

We estimated abundance and related parameters in both the sites using standard SCR methods (including singletons) and the SCR-2 model (removing the singletons). Density was modelled as a log-linear function of the estimated occupancy probability that was obtained from Bayandonoi et al. (2021). The half normal encounter rate was used to model the counts of detections at the traps conditional on the activity centres.

3 RESULTS

3.1 Simulations

Summaries of the populations and capture histories simulated can be seen in Table 1. The average number of ghost capture histories introduced for every 10% increase in the proportion of ghosts introduced under the 4 scenarios were: high abundance, high encounters: 6.5; high abundance, low encounters: 5.5; low abundance, high encounters: 2.1; and low abundance, low encounters: 1.6. This, however, resulted in an average increase of around 7 single detection capture histories under the high abundance scenarios and 2 single detection capture histories under the low abundance scenarios irrespective of the encounter rates. Lower encounter rates result in capture histories with fewer recaptures, which are more prone to being turned to single detection capture histories by misidentification errors.

For scenarios where no ghosts were introduced into the simulated dataset, both methods were unbiased across all estimated parameters Figure 3. The SCR model however, consistently overestimated abundance when ghosts were introduced. The RB of abundance increased with the proportion of detected individuals introduced as ghosts. The RB was higher under each scenario with ghosts when the encounter rates were lower, under both scenarios of total abundance. As expected, the variance of the parameter estimates from the SCR-2 model was higher than those from the SCR model due to lower sample sizes. MSE was therefore higher for all parameter estimates for the SCR-2 models when no ghosts were introduced. With respect to the expected abundance, SCR-2 models performed similarly to the SCR model in terms of MSE at just 10% of the detected individuals being introduced as ghosts and outperformed SCR models at higher proportions of ghosts introduced.

Estimates of σ remained unbiased for both models even when ghost capture histories were introduced. Furthermore, there was little loss in precision for σ estimates using the SCR-2 model as the MSE were similar to that of SCR models under all simulation scenarios. The SCR model underestimated λ_0 when ghosts were introduced and bias increased with the proportion of individuals detected as ghosts. Even in the SCR-2 model, the increasing trend in bias of λ_0 with increasing number of ghosts is observable. Fewer recaptured individuals imply low encounter rates. In the SCR model, ghosts misinform the model by falsely suggesting lower encounter rates. Ghosts are recaptures falsely made into singletons; in the SCR-2 model, dropping all singletons translates to dropping all true single-capture individuals and some misidentified recaptures. Dropping misidentified recaptures is analogous to an exclusion error, which results in lower encounter rates. While they may increase the uncertainty of estimates due to fewer detections, random exclusions do not bias the abundance estimate.

3.2 Case Study: Snow Leopards in Mongolia

54 detections of snow leopards were made across the 102 traps in Munkhkhairkhan. 19 unique individuals were identified from the detections of which 8 were detected just once. 177 detections of 22 uniquely identified individuals were made at the 40 traps in Tost. Three of the 22 were singletons. Estimates from both models are given in Table 2

$E[N]$, the expected number of individuals, estimated by the SCR-2 method was lower in both sites as compared to SCR. The SCR estimate was around 8% higher in Tost and over 24% higher in Munkhkhairkhan. The higher estimates in abundance is reflected inversely in lower encounter rates (λ_0) estimated by the SCR method. Estimated σ from both models were similar, with the absolute difference between both models being < 5. Except for in the case of the σ estimate in Tost, dropping single individuals resulted in higher coefficients of variation for the estimates in the SCR-2 model. There was no loss in the precision on σ in Tost; having had

Table 1: Summaries of the simulated populations and sets of capture histories under the various scenarios. β_1 and σ were both fixed under all scenarios at 0.1 and 300 respectively.

Density Intercept (β_0)	Encounter Rate (λ_0)	Proportion of Ghosts Introduced	Mean Values Across Simulations			
			Population Size	Total Detections	Individuals Detected	Single Detections
High (0.1)	High (2)	0%	78.8	280.1	65.2	12.2
		10%	78.8	280.1	71.7	19.1
		20%	78.8	280.1	78.2	26.1
		30%	78.8	280.1	84.6	33.1
	Low (1)	0%	78.8	140.5	55.3	18.5
		10%	78.8	140.5	60.8	25.4
		20%	78.8	140.5	66.2	31.8
		30%	78.8	140.5	71.6	39.0
Low (0.03)	High (2)	0%	22.7	81.3	20.8	3.5
		10%	22.7	81.3	20.8	5.5
		20%	22.7	81.3	22.6	7.5
		30%	22.7	81.3	24.4	9.5
	Low (1)	0%	22.7	40.7	15.7	5.2
		10%	22.7	40.7	17.3	7.2
		20%	22.7	40.7	18.9	9.1
		30%	22.7	40.7	20.3	11.1

a large number of recaptures, dropping single captures may not have led to a loss in precision of the σ estimate.

Table 2: Estimates and standard errors (SE) from the SCR and SCR-2 models for both the sites. CV, the coefficient of variation, is SE/Estimate. ΔEst and ΔCV are the relative differences in the parameter estimates and CVs of the SCR and SCR-2 models with respect to the SCR-2 model.

Site	Parameter	Estimate (SE)		CV		ΔEst(%)	ΔCV(%)
		SCR	SCR-2	SCR	SCR-2		
Munkhkhairkhan	$E[N]$	33.06 (8.76)	26.58 (10.49)	0.26	0.39	24.38	-32.86
	λ_0	0.011 (0.0024)	0.013 (0.0031)	0.23	0.24	-16.67	-6.56
	σ	5491.24 (575.79)	5758.48 (657.8)	0.10	0.11	-4.64	-8.21
Tost	$E[N]$	23.99 (5.13)	22.15 (5.11)	0.21	0.23	8.34	-7.47
	λ_0	0.016 (0.0018)	0.017 (0.002)	0.12	0.12	-7.44	-0.91
	σ	7826.52 (431.34)	7684.95 (421.48)	0.06	0.05	1.84	0.49

4 DISCUSSION

Non-invasive methods used in surveys for abundance estimation by SCR generate data such as photographs or genetic samples that have to be converted into capture histories for the model. The step of converting the raw data to capture histories is seldom error-free, irrespective of whether it is done manually or is automated. Errors in data processing leading to single detection individuals or ghosts is the most common in capture histories derived from both genetic as well as camera trap data (Yoshizaki, 2007; Johansson et al., 2020). We propose the SCR-2 method to address the bias resulting from this common misidentification error. The SCR-2 likelihood disregards all single captures (and therefore, it is assumed, all ghosts) and is parameterised in terms of the probability of being detected at least twice, rather than at least once.

Our simulation study shows the effect of ghosts in the capture history on parameter estimates. Johansson

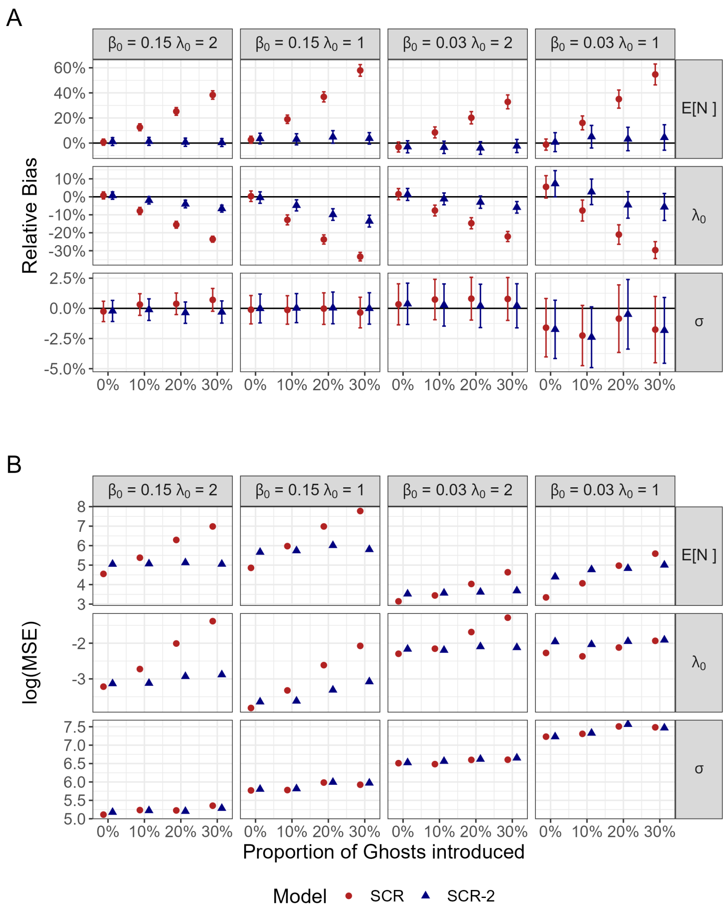

Figure 3: Results from the simulation study. Panel A shows relative bias. Red circles represent the estimated mean relative bias of the parameters from the SCR model and the blue triangles represent the estimated mean relative bias of the parameters from the SCR-2 model. The bars around the estimates show 95% error bars for the mean of the simulated estimates. Relative bias was computed as the ratio between the deviation of the estimate from the true value, and the true value. Panel B shows the log of the mean sqaured errors of the parameter estimates from the true value. Red circles show log(MSE) of the SCR model and the blue triangles show the log(MSE) of the SCR-2 model. Each panel column represents one of the 4 scenarios for combinations of high and low abundance and encounter rates. Each row of a panel shows results for the estimates of expected abundance: $E[N]$, expected encounter rate at the activity centre: λ_0, and the range parameter: σ.

et al. (2020) reported a splitting error probability of 0.11 for each detection that led to up to 4 ghosts (25% of detected individuals) by experts in their experiment. However, they did not investigate whether this probability is affected by the total number of detections or individuals detected. More detections could lead to more ghosts or alternatively provide more reference images to reduce splitting errors. We therefore chose a conservative approach of introducing ghosts as a proportion of the number of individuals detected rather than a proportion of the total number of detections. Even at a 10% proportion of detection individuals introduced as ghosts, there was a significant bias in the abundance estimate by the SCR method. Estimates from the SCR-2 method remained unbiased even at 30% of the detected individuals introduced as ghost individuals.

Dropping single detections and fitting the SCR-2 model lead to a loss in precision of the parameter estimates but even at 10% of the detected individuals being introduced as ghosts outperformed the SCR method with regards to MSE, except for the lowest sample size scenario. If error rates are expected to generate 10% or more ghost capture histories, we recommend using the SCR-2 method over SCR for abundance estimation, except perhaps when sample size is around the lowest in our simulations. With real data where the rates and types of misidentifications are not known, we show that there is a relative difference of 24% in abundance estimates in the site we suspected to be prone to identification errors. The simulations also showed that splitting errors could lead to more single detections than just the number of ghosts created. Dropping single detections in certain cases could therefore lead to bias even in the SCR-2 method, especially in surveys with low encounter rates, due to poor sample sizes.

Apart from camera trap surveys and genetic sampling, ghost recaptures have been shown to be the most common misidentification error in other noninvasive sampling methods (Stevick et al., 2001), making the SCR-2 method applicable in a larger range of surveys. Furthermore, SCR conditioned on at least K detections can also be used to deal with transient animals in a similar way to the CR method used by Conn et al. (2011). The specific SCR-2 model, where $K = 2$, has been independently developed and used to deal with potential false positive detections of bowhead whale (*Baleana mysticetus*) calls in an acoustic spatial capture-recapture study (Petersma et al., 2023) and our method can be used to extend such work to surveys requiring at least K detections.

It is also straightforward to address heterogeneity in how ghosts are introduced. If individuals that are difficult to identify are more likely to generate ghosts, this heterogeneity can be modelled by adding individual-based covariates to λ_0 to model encounter rates. This could be the case in species where identifying individuals of one sex is harder than the other, or if juveniles have less distinct markings than adults.

We also develop a hypothesis test that can be used to diagnose identification errors within capture histories (supplementary material). The test, however, proved to have low power, and not being able to reject the null hypothesis may not mean that there are no ghosts in the capture histories. Additionally, while ghosts are the result of the most common misidentification errors, other misidentification errors also occur and further investigation is required to understand how they affect SCR and SCR-2 estimates. Choo et al. (2020) review common errors in identifying camera trap images and outline best practices for teams to avoid them. Meanwhile, machine learning and deep learning methods for individual identification are areas of active research (Petso et al., 2022) and are becoming more accessible to practitioners. Alongside these advances, there is a need to develop methods that are able to incorporate the uncertainty of individual identities to develop an overall integrated workflow for abundance estimation.

Conversely, a large number of single detection individuals need not indicate large number of identification errors or the presence of ghosts. The expected number of single captures is a function of the encounter function parameters, the trap layout and the state space. Low encounter rates yield lower detections and fewer recaptures. Small home ranges relative to the distance between traps can lead to greater number of single detections and traps placed further away from suitable habitats are likely to have fewer detections. The trade-off between maximising the number of individuals detected or the number of recaptures is an active area of research to design SCR surveys. For more rigorous treatment of study design for SCR surveys and the effect of the trap array on the number of animals detected and the number of recaptures readers can refer to M. G. Efford and

Boulanger (2019) and Durbach et al. (2021).

In summary, we extend the standard SCR model to condition on at least K detections. The new likelihood can be used to estimate abundance and detection parameters on a subset of the capture histories if capture histories that contain fewer than K detections are expected to bias estimates and our work shows that despite the loss in precision, it may be preferable to use over the SCR method when misidentifications are expected. The standard SCR model is a special case of the more general method developed, the spatial capture-recapture conditioned on at least K detections, where $K = 1$. The method lends itself to other extensions of SCR, is easily implementable in existing software and adds no additional computation time, making it accessible to both researchers and practitioners.

AUTHOR'S CONTRIBUTIONS

AKR, JSH, DLB and KS conceived the idea. AKR and JSH conducted the analysis and wrote the manuscript. DLB, HW, OJ and KS reviewed and provided critical feedback for the manuscript. JSA, PL, GB and KS designed and organised camera trap surveys, and managed the data. JSA, PL, CB, GB, MO, SE and NB conducted the camera trap surveys and classified individuals in the camera trap images.

ACKNOWLEDGEMENTS

The camera trap survey work is a part of the nation-wide snow leopard population assessment initiated and implemented by WWF-Mongolia and SLCF-Mongolia, along with partner organizations aiming to plan effective conservation activities in the country under the umbrella species, the snow leopard. The funds for the nation-wide snow leopard population assessment in Mongolia to WWF Mongolia were provided by WWF Netherlands, its generous individual donors, Bram Linnartz, Patrick Munsters, Rob ten Heggeler, Olivier Gorter and Frank Thuis, and WWF Germany. Zoo Basel and David Shepherd Wildlife Foundation supported the Snow Leopard Trust and Snow Leopard Conservation Foundation's contribution to the assessment. We are also grateful to Global Environment Facility, United Nations Development Program and Snow Leopard Trust for supporting the Global Snow Leopard and Ecosystem Protection Program and development of tools and methods for Population Assessment of the World's Snow Leopards (PAWS). We would like to thank Dr. Charu Mishra for crucial advice at various stages of planning and implementing this body of work. Lastly, we thank the Ministry of Environment and Tourism of Mongolia and State protected area Administration of Mongolia for providing leadership to the nation-wide snow leopard survey in Mongolia.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

DATA AVAILABILITY STATEMENT

The data and code used in this study are available at `https://github.com/abinandkr/Ghostbusting`.

References

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. *The Annals of Applied Statistics*, *12*(1), 67–95.

Bayandonoi, G., Sharma, K., Shanti Alexander, J., Lkhagvajav, P., Durbach, I., Buyanaa, C., . . . others (2021). Mapping the ghost: Estimating probabilistic snow leopard distribution across mongolia. *Diversity and Distributions*, *27*(12), 2441–2453.

Bonner, S. J., Schofield, M. R., Noren, P., & Price, S. J. (2016). Extending the latent multinomial model with complex error processes and dynamic markov bases. *The Annals of Applied Statistics*, *10*(1), 246–263.

Borchers, D. L., & Efford, M. (2008). Spatially explicit maximum likelihood methods for capture–recapture studies. *Biometrics*, *64*(2), 377–385.

Chandler, R. B., & Royle, J. A. (2013). Spatially explicit models for inference about density in unmarked or partially marked populatioins. *The Annals of Applied Statistics*, 936–954.

Choo, Y. R., Kudavidanage, E. P., Amarasinghe, T. R., Nimalrathna, T., Chua, M. A., & Webb, E. L. (2020). Best practices for reporting individual identification using camera trap photographs. *Global Ecology and Conservation*, *24*, e01294.

Conn, P. B., Gorgone, A. M., Jugovich, A. R., Byrd, B. L., & Hansen, L. J. (2011). Accounting for transients when estimating abundance of bottlenose dolphins in choctawhatchee bay, florida. *The Journal of Wildlife Management*, *75*(3), 569–579.

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

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

Efford, M. (2022). secr: Spatially explicit capture-recapture models [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=secr` (R package version 4.5.3)

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

Fewster, R. M. (2017). Some applications of genetics in statistical ecology. *AStA Advances in Statistical Analysis*, *101*, 349-379. https://doi.org/10.1007/S10182-016-0273-0

Gibbon, G. E. M., Bindemann, M., & Roberts, D. L. (2015). Factors affecting the identification of individual mountain bongo antelope. *PeerJ*, *3:e1303*. https://doi.org/10.7717/peerj.1303

Jiménez, J., C Augustine, B., Linden, D. W., B Chandler, R., & Royle, J. A. (2021). Spatial capture–recapture with random thinning for unidentified encounters. *Ecology and evolution*, *11*(3), 1187–1198.

Johansson, Ö., Samelius, G., Wikberg, E., Chapron, G. Mishra, C., & Low, M. (2020). Identifcation errors in camera-trap studies result in systematic population overestimation. *Scientific Reports*, *10:6393*. https://doi.org/10.1038/s41598-020-63367-z

Link, W. A., Yoshizaki, J., Bailey, L. L., & Pollock, K. H. (2010). Uncovering a latent multinomial: Analysis of mark–recapture data with misidentification. *Biometrics*, *66*(1), 178-185. https://doi.org/10.1111/j.1541-0420.2009.01244.x

Lukacs, P. M., & Burnham, K. P. (2005). Estimating population size from DNA-based closed capture-recapture data incorporating genotyping error. *The Journal of Wildlife Management*, *69*(1), 396–403. https://doi.org/10.2193/0022-541X(2005)069⟨0396:EPSFDC⟩2.0.CO;2

MacKenzie, D. I., Nichols, J. D., Royle, J. A., Pollock, K. H., Bailey, L. L., & Hines, J. E. (2017). *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*. Elsevier.

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

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. *Journal of Agricultural, Biological and Environmental Statistics*, 1–20.

Petso, T., Jamisola, R. S., & Mpoeleng, D. (2022). Review on methods used for wildlife species and individual identification. *European Journal of Wildlife Research*, *68*(1), 1–18.

R Core Team. (2022). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

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

Sollmann, R., Gardner, B., Parsons, A. W., Stocking, J. J., McClintock, B. T., Simons, T. R., ... O'Connell, A. F. (2013). A spatial mark–resight model augmented with telemetry data. *Ecology*, *94*(3), 553-559. Retrieved from `https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1890/12-1256.1` https://doi.org/10.1890/12-1256.1

Stevick, P. T., Palsbøll, P. J., Smith, T. D., Bravington, M. V., & Hammond, P. S. (2001). Errors in identification using natural markings: rates, sources and effects on capture-recapture estimates of abundance. *Canadian Journal of Fisheries and Aquatic Sciences*, *58*, 1861–1870. https://doi.org/10.1139/f01-131

Vale, R. T. R., Fewster, R. M., Carroll, E. L., & Patenaude, N. J. (2014). Maximum likelihood estimation for model mt,α for capture–recapture data with misidentification. *Biometrics*, *70*(4), 962-971. https://doi.org/10.1111/biom.12195

Van Horn, R. C., Zug, B., LaCombe, C., Velez-Liendo, X., & Paisley, S. (2014). Human visual identification of individual andean bears *Tremarctos ornatus*. *Wildlife Biology*, *20*(5), 291–299. https://doi.org/10.2981/wlb.00023

Whittington, J., Hebblewhite, M., & Chandler, R. B. (2018). Generalized spatial mark–resight models with an application to grizzly bears. *Journal of Applied Ecology*, *55*(1), 157–168.

Wright, J. A., Barker, R. J., Schofield, M. R., Frantz, A. C., Byrom, A. E., & Gleeson, D. M. (2009). Incorporating genotype uncertainty into mark—recapture-type models for estimating abundance using dna samples. *Biometrics*, *65*(3), 833–840. Retrieved from `http://www.jstor.org/stable/20640581` https://doi.org/10.1111/j.1541-0420.2008.01165.x.

Yoshizaki, J. (2007). *Use of natural tags in closed population capture-recapture studies: Modeling misidentification* (Unpublished doctoral dissertation). Biomathematics Graduate Program and Department of Zoology, North Carolina State University, Raleigh.