Unravelling the effects of individual heterogeneity on connectivity estimates: Insights from Spatial Capture-Recapture modelling

Maëlis Kervellec*[1], O. Gimenez[1], C. Vanpé[2], P.-Y. Quenette[2], J. Sentilles[2], S. Palazón[3], I. Afonso Jordana[4], R. Jato[5], M. M. Elósegui Irurtia[6], and C. Milleret[7]

[1] CEFE, University of Montpellier, CNRS, EPHE, IRD, Montpellier, France
[2] Équipe Ours, Office Français de la Biodiversité, Impasse de la Chapelle, 31800 Villeneuve-de Rivière, France
[3] Fauna and Flora Service, Department of Climate Action, Food and Rural Agenda, Carrer del Foc, 57. 08038 Barcelona, Spain
[4] Departament de territòri, paisatge e gestion ambientau, Conselh Generau d'Aran, Plaça d'Aran 1-2, 25530 Vielha, Spain
[5] Departamento de Producción y Asistencias Técnicas, Sociedad Aragonesa de Getión Agroambiental, S. L. U., Ctra HUV-Fornillos, 22195 Fornillos de Apiés, Huesca, Spain
[6] Servicio de Biodiversidad, Sección de Espacios Naturales y Especies Amenazadas, Gobierno de Navarra-Nafarroako Gobernua, Pamplona-Iruñea, Spain
[7] Faculty of Environmental Sciences and Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway

*Corresponding author
Correspondence: maelis.kervellec@cefe.cnrs.fr

Abstract

Spatial Capture-Recapture (SCR) models, parametrized with least cost path distance, provide a unifying framework for explicitly estimating landscape connectivity and population size from individual detection data. However, we frequently encounter individuals with larger apparent space use than the rest of the population. To avoid biased population size estimates, it is common practice to remove these outliers from the analysis. Yet such individuals are likely to be very important when the aim is to estimate connectivity. We therefore investigated how unmodelled individual heterogeneity in space use affects population size and connectivity estimates from SCR models.

We first conducted a simulation study. We varied the proportion of the population whose space use is most constrained by landscape structure. We fit four SCR models, increasing the degree of modelled individual heterogeneity to assess its effect on parameter estimates. Secondly, we applied these models to the Pyrenean brown bear population, examining the effects of removing outliers or modelling explicitly heterogeneity on parameter estimates.

Our simulation study showed that population size was underestimated when no individual heterogeneity was modelled and increased with increasing the level of individual

heterogeneity modelled. Moreover, unmodeled individual heterogeneity led to biased connectivity estimates towards the group, which is the most detected. In our case study, modelling individual heterogeneity was key to unravel the patterns in population space use. Female brown bear movements were restricted by road density (estimated resistance: α_2 = 0.27 [0.01, 0.43]), whereas a small group of males were not affected by road density (α_2 = -0.16 [-0.64, 0.24]).
Our study provides valuable insights for optimizing the application of SCR models. When the primary objective is to accurately estimate population size, we recommend removing outlier individuals with larger home range sizes to avoid the risk of underestimating population size. Conversely, when the aim is to assess landscape population-level connectivity, modelling individual heterogeneity is essential. Our findings underscore the importance of tailoring the SCR model to the specific research goal, ensuring more precise and meaningful ecological inferences.

Keywords: Hierarchical model, Least-cost path, Nimble, Resistance surface, Spatially-explicit capture-recapture model, Species distribution model, *Ursus arctos*

1. Introduction

Animals move through the landscape according to intrinsic (e.g., locomotion abilities, need to feed or to find mates) and extrinsic factors (Bell, 1990). This means that animals do not move randomly through the landscape, but constantly have to adjust their movement to the distribution of opportunities and pressures. Identifying the spatial features affecting animal movement can be done by quantifying connectivity. In other words, connectivity aims at describing "the degree to which the landscape facilitates or impedes movement among resource patches" (Taylor et al., 1993). Population-level connectivity can be based on a map called a resistance surface that quantifies the cost of moving through each pixel of the landscape. This surface is estimated from spatial environmental variables and information on the cost of movement or space use derived from expert opinion or empirical data (Dutta et al., 2022; Zeller et al., 2012). In landscape management, resistance surfaces are used to map corridors, detect barriers or predict population range shift due to climate change (Dutta et al., 2022). However, connectivity does not only have a spatial dimension, but also varies with time (Zeller et al., 2019), or with individual characteristics: e.g. life stages (Thorsen et al., 2022), or sex (García-Sánchez et al., 2022; Zeller et al., 2023). Even if it could be key to unravel intraspecific connectivity patterns, group specific connectivity is rarely considered (Garcia-Sanchez et al., 2022; Zeller et al. 2023).

Spatial capture-recapture (SCR) models, which have become standard in population ecology (Efford, 2004; Tourani, 2021), are hierarchical models that use spatialised individual detection histories to estimate population size and distribution. At the core of the SCR model lies the assumption that individual detection probability is a function of distance from the activity centre (Royle et al., 2014). In its simplest form, a SCR model assume a circular space use by individuals. However, SCR models have been extended to allow the estimation of irregularly shaped space use caused by spatial heterogeneity of the landscape. In such models, the distance used in the detection function is not Euclidean but based on the least cost-path (LCP) distance, with the resistance of landscape features directly inferred by the model. Consequently, SCR models with LCP provide a unified framework for explicitly quantifying both population size and the resistance of landscape features on space use (Royle et al., 2013; Sutherland et al., 2015) that is population-level landscape connectivity. This model has been applied on different species: e.g. American mink (Neogale vison, Fuller et al., 2016), black bear (Ursus americanus, Morin et al., 2017), jaguars (Panthera onca, Tobler et al., 2018), snow leopard (Panthera unica, Pal et al., 2021), brown bear (Ursus arctos, Kervellec et al., 2023). However, these studies were limited to connectivity estimation at the population level and did not take into account individual heterogeneity in the resistance parameter scaling the resistance surface.

Individual heterogeneity can be defined as the variability of responses among individuals within a population that can be associated with measurable or non-measurable individual characteristics. In non-spatial capture-recapture models, properly modelling individual heterogeneity by including covariates (age, sex, body size, breeding state, etc.) is crucial to prevent creating negative bias in abundance estimates (Gimenez et al., 2018). While abundance estimates from SCR models are generally robust to unmodeled heterogeneity in movement (Royle et al., 2016; Theng et al., 2022), highly structured landscapes can cause an underestimation of population size (Royle et al., 2013). Individuals within a population can also display heterogenous levels of resistance to landscape structure. For example, in a brown bear population, some individuals can avoid human settlement to avoid the risk of encountering humans, whereas others might be prone to go close to cities to find food (Bombieri et al., 2021). Such heterogeneity, that cannot be explained by any particular characteristics of the individual, can be challenging to deal with in SCR models and often remain unaccounted for. In some cases, individuals that exhibit a space usage not representative of the whole population can even be considered as outliers and be removed prior from the analysis to avoid biases in parameter estimation (Kendall et al., 2019; Kervellec et al., 2023; Schmidt et al., 2022), However, quantifying and accounting for individual heterogeneity are key steps in the analysis when the objective is to estimate population-level landscape connectivity.

In this study, we investigated the consequences of not accounting for individual heterogeneity in space use response to landscape structure on population size and connectivity estimates. First, we conducted a

simulation study with two scenarios varying the proportion of individuals exhibiting heterogeneous responses to landscape features. The first scenario reflected a population with 10% of individuals exhibiting larger space use than others. The second scenario represented a population with two distinct behavioural responses to landscape structure. We assumed that population size will be underestimated when connectivity was not modelled (Royle et al., 2013) and that connectivity will be overestimated for the part of the population the most restricted by the landscape structure. Secondly, as a case study, we applied the same models on the Pyrenean brown bear population. We examined the effects of removing outliers (as we did in Kervellec et al., 2023), versus modelling explicitly heterogeneity on parameter estimates. We assumed that removing outliers can be an efficient way of obtaining reliable population size estimates when quantifying connectivity is not the main goal of the study. However, we predict that modelling individual heterogeneity will be essential to understand and map population-level connectivity.

2. Material and Methods

2.1 Spatial Capture Recapture models to estimate connectivity

Spatial capture-recapture (SCR) models are hierarchical models that estimate the location of each individual activity centre denoted s_i in the state space S according to a spatial point process. This process can be homogeneous, in this case $s_i \sim \mathrm{Uniform}(S)$ or can be driven by spatial variables. In the later, an intensity function $\lambda(s) = e^{\beta X(s)}$ describes the placement of activity centres, according to the vector of regression coefficients β and the vector of spatial covariate $X(s)$ evaluated at cell s (Zhang et al., 2023). SCR models combine an ecological sub-model describing the distribution of the population's activity centres and observational sub-model. The observation sub-model is based on the detection and non-detection of individuals through an array of traps and account for the imperfect detection of individuals by explicitly estimating detection probability (Efford, 2004; Royle et al., 2014). Therefore, the detection history corresponds to detecting or not the individual i at trap j at the sampling occasion k in session t, and is noted y_{ijkt}:

(1) $$y_{ijkt} \sim \mathrm{Bernoulli}(p_{ijkt})$$

Where p_{ijkt} is the detection probability usually modelled by a half-normal function:

(2) $$p_{ijkt} = p_{0ijkt} \times \exp\left(\frac{-d_{euc}(x_j, s_i)^2}{2\sigma^2}\right)$$

In this relation, the detection probability decreases with the Euclidean distance, noted $d_{euc}(x_j, s_i)$, between the trap x_j and the activity centre s_i. The baseline detection probability p_{0ijkt} is the probability of detecting an individual if its activity centre is located exactly on the trap. The scale parameter σ provides the extent to which an individual can be detected and is correlated with home range size. The larger the σ, the further from its activity centre an individual can be detected. These parameters can be constant or varying according to covariates or groups to account for heterogeneity in the detection process. By using the Euclidean distance, we assume that space use is symmetric (Royle et al., 2014), meaning that the presence of barriers or resources do not affect movements of individuals. We account for the effect of landscape structure by assuming that detection probability is a function of the ecological distance instead of Euclidean distance. The ecological distance can be represented by the least-cost path distance, which is based on a discrete resistance surface, where each pixel, noted ν_g, is given a cost. The distance between two pixels in the landscape is the succession of m steps, noted $\nu_1, \nu_2, \dots, \nu_m$, forming a path. We calculate the cost of all possible paths $(\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_l)$ connecting these two pixels and keep the one with the minimum value:

(3) $$d_{lcp}(\nu, \nu') = \min_{\mathcal{L}_1,\dots,\mathcal{L}_l} \sum_{g=1}^{m-1} \mathrm{cost}(\nu_g, \nu_{g+1}) \times d_{euc}(\nu_g, \nu_{g+1})$$

where, $\mathrm{cost}(\nu_g, \nu_{g+1}) = \frac{\exp(\alpha_2 W(\nu_g)) + \exp(\alpha_2 W(\nu_{g+1}))}{2}$

The resistance parameter α_2 is explicitly estimated in the SCR model from the encounter data and $W(\nu)$ is the resistance variable considered to build the resistance surface at cell ν. The estimated resistance parameter reflects the extent to which W facilitates ($\alpha_2 > 0$) or impedes ($\alpha_2 < 0$) the movement of individuals across the landscape (Royle et al., 2013; Sutherland et al., 2015). When $\alpha_2 = 0$, the resistance is null and the distance is exactly Euclidean. Here the least-cost path distance is computed by the Dijkstra algorithm implemented in the *gdistance* R package (version 1.6.4, van Etten, 2012).
A map of the potential landscape connectivity can be derived from these estimates. Given the location of an activity centre, we can compute its space use from equation 2, by setting $p_0 = 1$. We then assume that an activity centre is present in each pixel of the landscape and compute the sum of the space use of each individual at each pixel (Morin et al., 2017) to obtain a measure of potential connectivity. Therefore, potential connectivity goes from 0, meaning that not a single individual can reach this pixel, to the total number of pixels, where all individuals can reach this pixel. A realized measure of landscape connectivity is computed by weighted the potential connectivity map by the estimated number of activity centres in each pixel, also called density weighted connectivity.

2.2 Simulations

We conducted a simulation study to assess the consequences of not accounting for individual heterogeneity in space use on parameter estimation. We considered a population composed of two groups with different responses to landscape structure. We set a 25 x 25 distance unit (du) squared spatial domain in which we centred an array of 100 traps, each spaced by a minimum of 2du. We kept a 3.5du buffer around traps, following usual recommendations (Royle et al., 2014). To represent the landscape structure, we defined the resistance covariate by a raster of 0.5x0.5du resolution representing landscape fragmentation with the density of roads (see map AppendixS2 FigureS2A).

We simulated 200 individuals assuming homogeneous distribution of their activity centres within the spatial domain. We assumed that the population was composed of two groups (A and B), where space used of group A was more affected by the density of roads (α_{2_A}= 1) than group B (α_{2_B}= 0.5). In addition, we also assumed that individuals of group A had a smaller home range (σ_A = 1) than individuals of group B (σ_B = 2), regardless of the constraints imposed by roads. In order to obtain comparable detectability rates between the two groups despite their different space use, we also set a higher baseline detection probability for individuals of the group A (p_{0_A} = 0.2) compared to group B (p_{0_B} = 0.1).

We then considered two scenarios. In the first scenario, 10% of the population (n = 20 individuals) was assigned to group B and in the second scenario 50% of the population (n = 100 individuals) was assigned to group B. The first scenario is intended to represent a population with only few individuals with contrasted response to landscape structure like dispersers. The second scenario illustrates a population with two even groups of individuals showing a heterogeneous response to landscape as it can be the case for males and females (Bischof et al., 2017). For each scenario, we simulated 25 detection histories, representing a total of 50 different simulated data sets.

For each simulation, we fitted four models with increasing levels of modelled individual heterogeneity in the detection process. In other words, for each of the three parameters in the detection model (p_0 , σ or α_2), we estimated either one value for the whole population or a value for each group. For the first three models, we used the R package *oSCR* (version 0.42.0, Sutherland et al., 2019). Two models used the Euclidean distance and did not include a resistance parameter ($M1 = p_0 \sim \alpha_p, \sigma \sim \alpha_s$ and $M2 = p_{0_i} \sim \alpha_p + \beta_p \times group_i,\ \sigma_i \sim \alpha_s + \beta_s \times group_i$). The third model included the LCP distance based on the resistance surface derived from road density ($M3 = p_{0_i} \sim \alpha_p + \beta_p \times group_i,\ \sigma_i \sim \alpha_s + \beta_s \times group_i, \alpha_2 \sim \beta_{connect} \times road$). For the last model estimating group-specific resistance parameters ($M4 = p_{0_i} \sim \alpha_p + \beta_p \times group_i,\ \sigma_i \sim \alpha_s + \beta_s \times group_i, \alpha_{2_i} \sim \beta_{connect} \times road \times group_i$), we could not use oSCR. We therefore developed a Bayesian SCR model with LCP distance by using the R package nimble (version 1.01) and nimbleSCR (version 0.2.1, de Valpine et al., 2017). We ran two chains of 50,000 iterations each and discarded the first 10,000 iterations as burn-in. For the simulation study, we discarded models that did not converge, based on visual inspection of the trace plots, the Gelman–Rubin statistic ($\widehat{R}$ > 1.1)

and effective sample size (n_{eff} < 100) (Brooks and Gelman, 1998). For the Bayesian models, we used a data-augmentation approach, with an augmented population of 250 individuals (Royle et al., 2007), and priors are presented in the AppendixS2 (TableS2.C).

2.3 Case study: Brown bears in the Pyrenees

2.3.1 Data

In this case study, we analysed the spatial capture-recapture dataset of the Pyrenean brown bear population conducted across the Pyrenees mountains in France, Spain and Andorra, presented in Kervellec et al. (2023). In this transboundary study, we collected data from three sampling sessions from 2017 to 2019, in which we defined seven sampling occasions i.e., the months from May to November. We used the data from both the structured and opportunistic monitoring of the population (Vanpé et al., 2022). The structured monitoring consisted of a total of 708 DNA hair snags set on walking transects, around 10% of which are combined with a camera trap (Figure 1). The opportunistic monitoring consisted of 246 traps, defined according to a 25km^2 grid over the Pyrenees. A cell is considered active if at least one depredation of a bear to livestock or beehives occurred in the cell between 2010 and 2020. In these cells, genetic samples (hair or scat) could be collected after the depredation (Figure 1). Each sample was genetically analysed following details available in Vanpé et al. (2022), to obtain the identity and sex of the bear detected at a given trap in a given sampling occasion. To account for individual heterogeneity, we divided the population into three groups, females, males and outliers. We identified two outlier individuals (named "Goiat" and "Néré"), both males were detected more frequently than the others throughout a large part of Pyrenees mountains (Table 1, Figure 1). The maximum distance per year between detections for Néré was between 68 and 108.6km and between 49.2 and 78.3km for Goiat. Whereas the maximum distance per year per individual between detections was between 0 (i.e. individuals detected once) and 39.5km for the females (mean = 6.1km) and between 0 and 28.5km for the males (mean = 10.7km). In average outliers were detected 14 times, males 4.1 times and females 2.9 times per year.

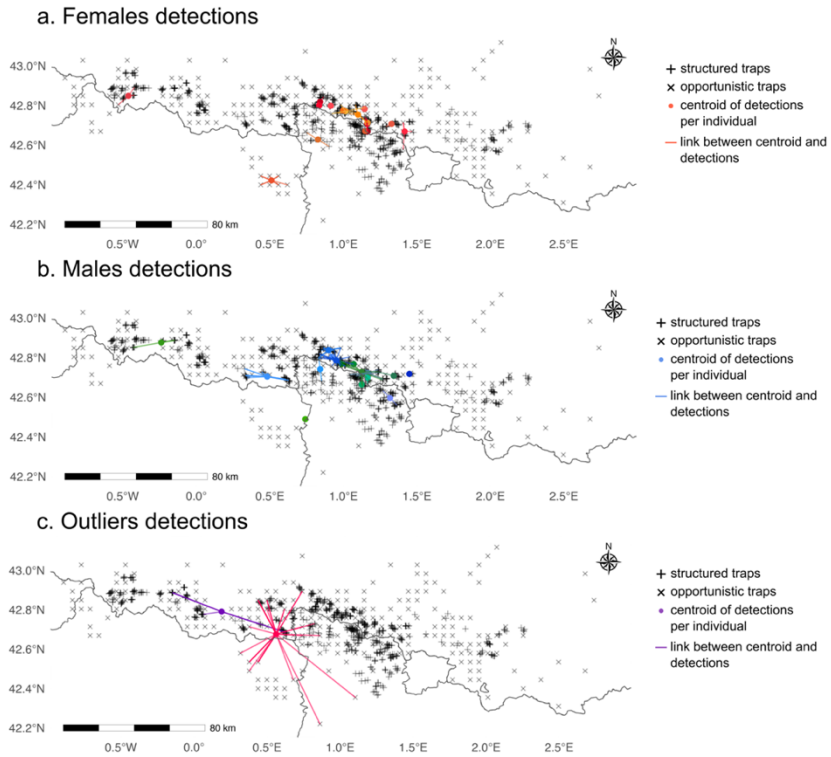

Figure 1 - Maps of traps distribution from the structured (+) and the opportunistic (x) monitoring across the Pyrenees mountains (France, Spain, Andorra). Points represents the centroid of detections per individuals and the lines links the centroid to the centroid to the traps where each individual was detected in 2019 for females (a), males (b) and outliers (c).

Table 1 - *Number of females, males and outliers detected and number of detections across traps and sampling occasions each year (2017, 2018, 2019) by the structured and the opportunistic monitoring.*

	females	males	outliers
Number of individuals detected			
2017	17	14	2
2018	14	17	2
2019	16	17	2
Number of detections			
2017	47	53	30
2018	32	57	32
2019	58	89	22

2.3.2 Models

We applied a SCR model similar to the one used in Kervellec et al. (2023). We assumed that the structured and opportunistic sampling followed different detection processes such as:

(4) $$y_{Struct_{ijkt}} \sim \text{Bernoulli}(p_{Struct_{ijkt}})$$

(5) $$y_{Opp_{ijkt}} \sim \text{Bernoulli}(p_{Opp_{ijkt}})$$

Where $y_{Struct_{ijkt}}$ and $y_{Opp_{ijkt}}$ are the encounter history and $p_{Struct_{ijkt}}$ and $p_{Opp_{ijkt}}$ are the detection probability for the structured and opportunistic monitoring, respectively. We modelled heterogeneity in the detection probability arising from differences in the sampling protocols or according to brown bear ecology on the baseline detection probability (p_0) and the scale parameter (σ).

(6) $$\text{logit}(p_{0_{Struct_{ijkt}}}) \sim \alpha_{Struct} + \beta_b \times Behav_{ijkt} + \beta_{t1} \times Nvisits_{jkt} + \beta_{t2} \times Country_j + \beta_{t3} \times Trap_j + \beta_{t4} \times m_k + \beta_{t5} \times m_k^2$$

(7) $$\text{logit}(p_{0_{Opp_{ijkt}}}) \sim \alpha_{Opp} + \beta_b \times Behav_{ijkt} + \beta_{t6} \times Country_j + \beta_{t4} \times m_k + \beta_{t5} \times m_k^2$$

The baseline detection probability (Equation 6) for the structured monitoring was assumed to be a function of a quadratic effect of the sampling occasion (m), the country (Country), the number of visits at the trap each month (Nvisits), and if the hair snag was combined or not with a camera trap (Trap). For the opportunistic monitoring (Equation 7), the baseline detection probability was a function of the country (Country) and the quadratic effect of the month (m). Both monitoring accounted for a behavioural effect (Behav), meaning that an individual detected once at a trap within a session was more likely to be detected at this trap in the following sampling occasions (Kervellec et al., 2023). Additionally, the scale parameter was estimated for each sex: $\sigma_i \sim \alpha_s + \beta_s \times sex_i$.

The brown bear densities were assumed to vary according to an inhomogeneous binomial point process:

(8) $$\log(\lambda(s)) = \beta_{d1} \times Rug_s + \beta_{d2} \times HumDens_s$$

The intensity function (Equation 8) was a function of the linear and additive effect of ruggedness (Rug), defined as average of the absolute values of elevation differences between the focal cell and the eight surrounding cells, and human population density (HumDens) (see Kervellec et al., 2023).

In this Bayesian formulation of the model, we used a data-augmentation approach. The augmented population M has to be large enough to encompass all individuals that were never detected (Royle et al., 2007). Here we used M = 204 individuals, which is four times the number of detected individuals. The latent state variable z_i model the inclusion of an individual from the augmented population to the population with an inclusion probability Ψ:

(9) $$z_i \sim \text{Bernoulli}\ (\Psi)$$

From this equation we can derive the total population size $N = \sum_{i=1}^{M} z_i$, which is the sum of the individuals included in the population.

We tested four models which differed in terms of the data considered and the number of resistance parameters estimated. In the first two models, outlier individuals were excluded. Model MF1 estimated a single resistance parameter for the entire population (α_2), while model MF2 estimated two resistance parameters ($\alpha_{2_{\text{Female}}}; \alpha_{2_{\text{Male}}}$), one for females and one for males. The remaining two models included all detections. Model MFO1 estimates a single resistance parameter for the entire population (α_2), while model MFO3 estimated three resistance parameters independently ($\alpha_{2_{\text{Female}}}$; $\alpha_{2_{\text{Male}}}$; $\alpha_{2_{\text{Outlier}}}$), for females, males, and outliers.

We assumed that we know the group of each detected individuals. However, for undetected individuals their group η_i is estimated according to a categorical distribution:

(10) $$\eta_i \sim \text{Categorical}\ (p_{\text{Female}}, p_{\text{Male}}, p_{\text{Outlier}})$$

Where p_{Female}, p_{Male} and $p_{Outlier}$ are respectively the probability to belong to the group female, male or outlier.

We used the R package *nimbleSCR* (Turek et al., 2021). We run three chains of 30,000 iterations each and discarded the first 1,000 iterations as burn-in for each model, and priors are presented in Appendix S3 (Table S3.A). We checked model convergence, based on visual inspection of the trace plots, the Gelman–Rubin statistic ($\hat{R}$ < 1.1) and effective sample size (n_{eff} > 100) (Brooks and Gelman, 1998). We derived potential connectivity maps and density weighted connectivity maps from the distributions of model MFO3 and propagated associated uncertainty. All the analysis were conducted in R (R Core Team, 2023, version 4.3.1)

3. Results

3.1 Simulation

In the scenario 1, all model converged, with $\hat{R} \leq 1.02$ (Appendix S2, TableS2.A). Whereas in scenario 2, five simulations (1, 6, 7, 12 and 18) did not converge (Appendix S2, TableS2.B), likely because the number of simulated detections for one group was too scarce (see Appendix S2, FigureS2.A). These simulations were removed for the following results.

Mean estimates of population size showed lower apparent bias when increasing the amount of individual heterogeneity modelled (Figure 2, and Appendix S2). In model M1, without modelling individual heterogeneity, population size was consistently underestimated (mean estimate of 89% of simulations (40/45) was below the true value). However, all confidence intervals overlapped the true value. Introducing individual heterogeneity in the scale parameter (model M2) led to a mean estimate below the true value in 71% of simulations (32/45), with all confidence intervals overlapping the true value. Modelling Least Cost Path (LCP) distance in the detection function (model M3) resulted to a mean estimate below the true value in 36% of simulations (16/45), with 98% of confidence intervals overlapping the true value. When modelling individual heterogeneity on the resistance parameter (model M4), 24% of simulations (11/45) were below the true value, and 89% of credible intervals overlapped the true value (Figure 2).

In both scenarios, when we did not account for individual heterogeneity (model M1), the detection estimates (p_0 and σ) were intermediate between the two true values (Figure 2). The detection estimates were driven toward the group that was the most detected, that is the group A in scenario 1 (Figure 2a) and the group B in scenario 2 (Figure 2b). We observed the same pattern in the model M3, when no

heterogeneity was modelled on the resistance parameter. In the scenario 1, the resistance parameter was biased towards the group A and in the scenario 2 the estimated resistance parameter is smaller and biased toward the group B. When heterogeneity was modelled (M2 and M4), all estimates of p_0, σ and α_2 were estimated without apparent bias.

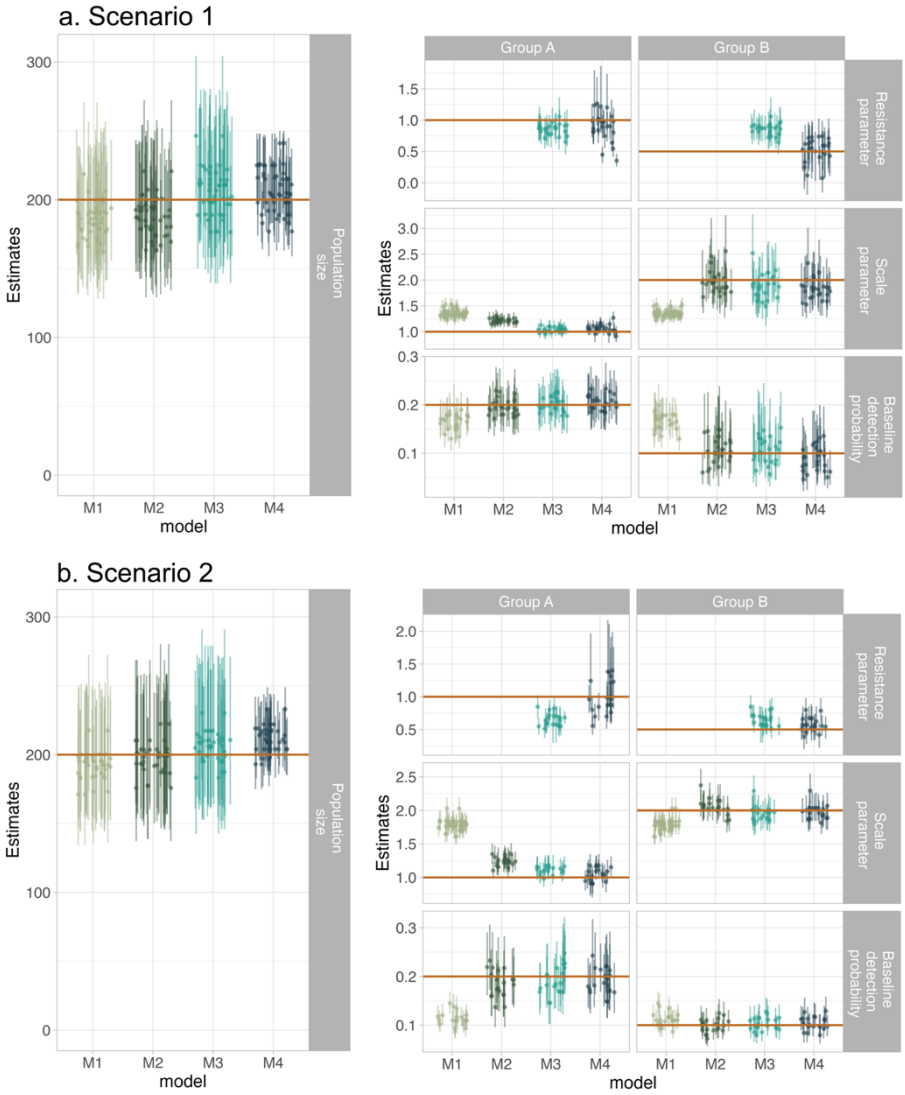

Figure 2 - *For each simulation, estimated values for scenario 1 (a, top panels) where 10% of the population is in group B and scenario 2 (b, bottom panels) where 50% of the population is in group B. In the left panels, we display estimates of population size (N) and in the right panels, we display estimates for detection parameters (resistance parameter (α_2), scale parameter (σ) and baseline detection probability (p_0)) of group A and B. We compare four models: M1 = D ~ 1 p_0 ~ 1 σ ~ 1, M2 = D ~ 1 p_0 ~ group σ ~ group, M3 = D ~ 1 p_0 ~ group σ ~ group α_2~ road and M4 = D ~ 1 p_0 ~ group σ ~ group α_2~ group*road. Horizontal lines are true values. The median of each simulation is represented by a point with errors bar displaying 95% confidence or credible intervals. Simulations 1, 6, 7, 12, 18 for the scenario 2 did not converge and were removed.*

3.2 Case study

The models MF1, MF2 and MFO3 did converge with $\hat{R} \leq 1.02$ and effective sample size > 200. On the contrary, the mixing for the resistance parameter in the model MFO1 was poor with an effective sample size of only 93 (Appendix S2, Table S2A, B).

When outliers were removed, modelling individual heterogeneity in the resistance parameter did not affect the population size estimates when the two groups considered had close α_2 values from $N_{MF1} = 93$ [66; 132] to $N_{MF2} = 94$ [67; 133] (Figure 3). Integrating outliers without modelling heterogeneity in σ or

α_2 (model MFO1) led to 1.6 times increase in population size estimates ($N_{MFO1} = 148$ [100; 198]) compared to previous estimates (MF1 and MF2), especially for females. Whereas, modelling heterogeneity in the resistance parameter (model MFO3) tended to increase the mean population size estimate, but with 95% credible intervals overlapping the mean estimated population size obtained without integrating outliers (Figure 3; $N_{MFO3} = 107$ [76; 154]).

Similarly, the estimated scale parameter and resistance parameter for males and females slightly differed when we modelled individual heterogeneity in the resistance parameter from $\alpha_{2_{MF1}} = 0.25$ [0.09; 0.37] to $\alpha_{2_{Males_{MF2}}} = 0.21$ [0.02; 0.35] and $\alpha_{2_{Females_{MF2}}} = 0.27$ [0.11; 0.42] (Figure 3). However, integrating outlier individuals without modelling individual heterogeneity (model MFO1) led to a 2.7 times increase in the median estimated σ of males (from $\sigma_{Males_{MF03}} = 0.06$ [0.06; 0.07] to $\sigma_{MF01} = 0.16$ [0.14; 0.18]), as well as an increase in the uncertainty of α_2 (from $\alpha_{2_{Males_{MF03}}} = 0.21$ [0.03; 0.35] to $\alpha_{2_{Males_{MF01}}} = -0.26$ [-0.45; 0.42]) (Figure 3). Then, when modelling individual heterogeneity (model MFO3), integrating outliers did not affect males' and females' resistance estimates but showed that road density did not affect outliers' space use as α_2 95% credible interval overlapped 0 ($\alpha_{2_{Outliers_{MFO3}}} = -0.16$ [-0.64, 0.24]). Parameter estimates for the models are available in the appendices (Appendix S3).

The map of potential connectivity derived from model MFO3 showed that females' movements were more restricted than males' in high road density areas (Figure 4ac). However, the uncertainty of $\alpha_{2_{Outlier}}$ propagated in the potential connectivity map, which did not allow us to identify barriers or corridors to their movements (Figure 4ef). Density weighted connectivity was the highest at the border between Catalonia (Val d'Aran and Pallars Sobirà) and Ariège and Haute-Garonne (Figure 5ac). The activity centres of outliers were estimated to be located in Aragon (Figure 5e).

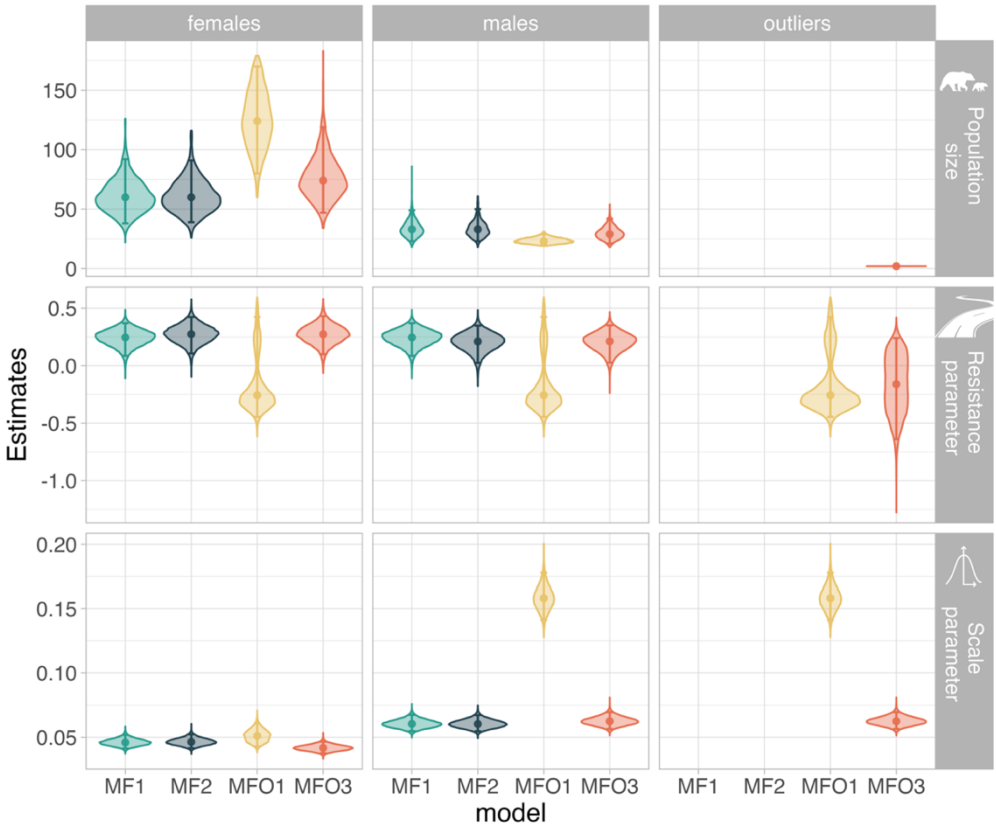

Figure 3 - Violin plot of estimated population size (N), resistance parameter (α_2), scale parameter (σ) for females, males and outliers' Pyrenean brown bears in 2019 in function of four models. Models MF1 and MF2 use only data from males and females and model MFO1 and MFO3 include data from the two outlier individuals. Models MF2 and MFO3 estimate respectively two and three resistance parameters corresponding to the groups in the data, whereas models MF1 and MFO1 estimate only 1 resistance parameter for all groups. The median of each distribution is represented by a point with errors bar displaying 95% credible intervals.

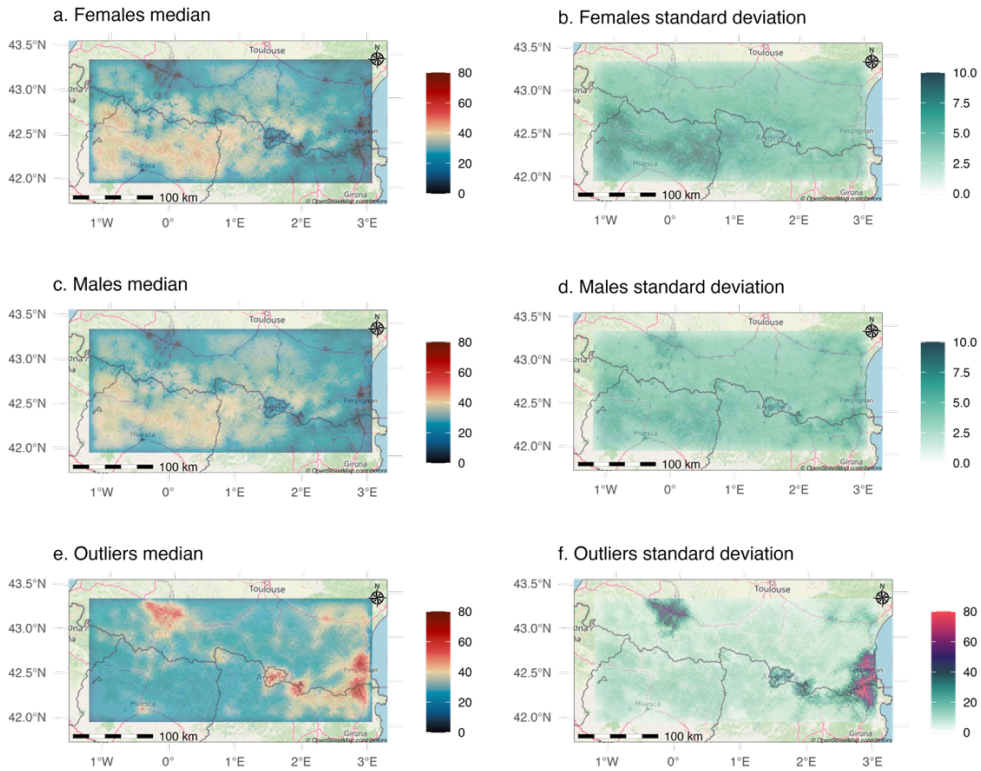

Figure 4 - Maps of median potential connectivity derived from estimates of model MFO3, for females (a), males (b) and outlier individuals (e) with respective standard deviation (b, d, f) for the Pyrenean brown bear population in 2019.

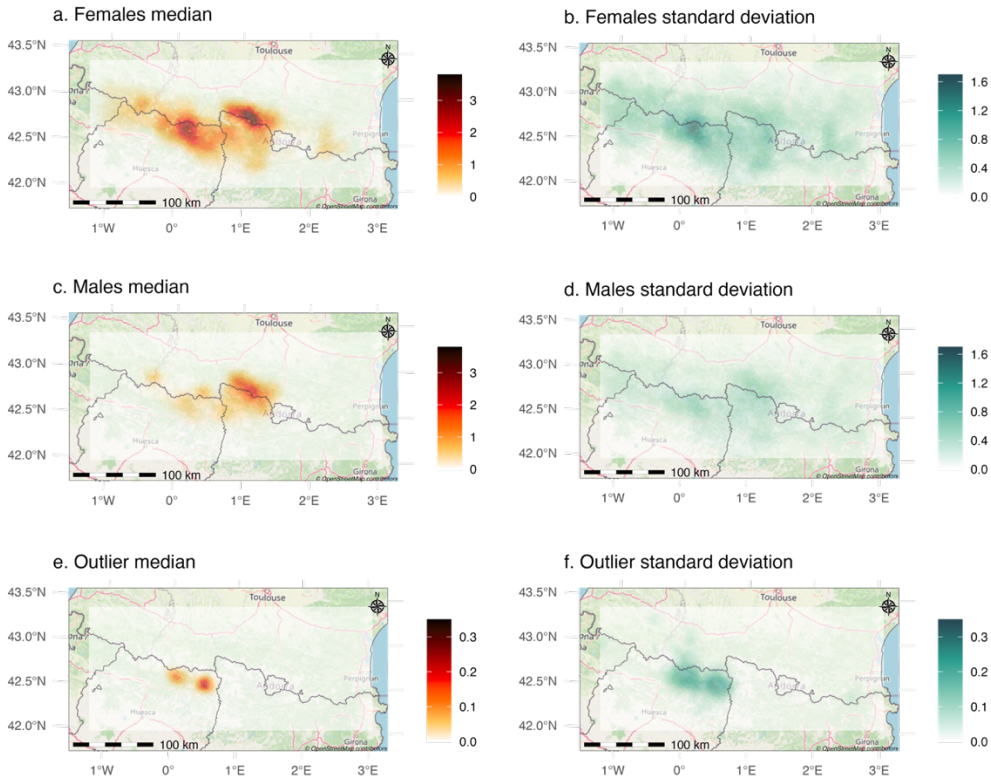

Figure 5 - Maps of median density weighted connectivity derived from estimates of model MFO3, for females (a), males (b) and the two outlier individuals (e) with respective standard deviation (b, d, f) for the Pyrenean brown bear population in 2019.

4. Discussion

As predicted, a negative bias in population size estimates occurs when individual heterogeneity in space use is unaccounted for (Pledger, 2000; Royle et al., 2013). The combination of the simulation and Pyrenean brown bear case study also revealed that accounting for individual heterogeneity in space use is key to obtain reliable population-level inferences in connectivity (García-Sánchez et al., 2022; Zeller et al., 2023). Although individual heterogeneity can be challenging to model, especially when a part of the population remains undetected (Marrotte et al., 2022; Pledger, 2000), we also demonstrated how the latent state formulation of SCR models can distinguish and estimate the space use of different types of individuals. Conversely, when the primary objective is to accurately estimate population size, removing outlier individuals with larger home range sizes is a straightforward option to avoid the risk of underestimating population size (Kendall et al., 2019; Schmidt et al., 2022). Our findings underscore the importance of tailoring the SCR model to the specific research goal, ensuring more precise and meaningful ecological inferences.

Introducing connectivity into a SCR model serves as a mechanism for implicitly capturing individual heterogeneity in space use among individuals based on the location of their activity centres across the landscape (Royle et al., 2013). However, a common modelling assumption is that all individuals share the same response to landscape features. Whether it is linked to age, sex or other intrinsic characteristics of the individual, different responses across individuals are expected (Bombieri et al., 2021; García-Sánchez et al., 2022; Zeller et al., 2023). When a single parameter is estimated for the entire population, the estimates will tend to reflect the group response with the largest number of individuals. For example, when only individuals with the highest connectivity were detected, ignoring individual heterogeneity led to underestimating the landscape resistance for individuals that were the most affected by landscape features.

The application of the model to the Pyrenean brown bear population showed that it was possible to account for individual heterogeneity, even if the number of outliers was small (n=2). However, keeping outliers without modelling individual heterogeneity in connectivity and scale parameter led to increase the estimated number of females, and to slightly decrease the estimated number of males. We were expecting a negative bias in population size, when individual heterogeneity remained unmodelled (Gardner et al., 2010; Pledger, 1994). However, we observed a decrease in the probability to be a male ($p_{Male_{MF1}} = 0.34\,[0.22, 0.49]$ to $p_{Male_{MFO1}} = 0.15\,[0.01, 0.22]$) and an increase in detection probability at opportunistic traps in Spain ($\beta_{t6_{MF1}} = -1.54\,[-2.15, -0.99]$ to $\beta_{t6_{MFO1}} = -0.72\,[-1.17, -0.29]$). This ultimately led to an increase in female density in Ordesa and Monte Perdido National Park likely because monitoring intensity was low in this area, and the habitat was predicted to be favourable according to the set of habitat covariate chosen. Overall, removing outliers or keeping them and modelling individual heterogeneity in connectivity gave us comparable population size and detection estimates ($p_0, \sigma_{Males}, \sigma_{Females}, \alpha_{2_{Males}}, \alpha_{2_{Females}}$). This means that removing outliers did not modify the conclusion of the study when the aim was to estimate population size (Kendall et al., 2019; Schmidt et al., 2022).

SCR models are a flexible framework to study group-based connectivity from non-invasive detection data. Although studies already found sex differences in parameters such as dispersal, or home range size (Dahle et al., 2006; Greenwood, 1980), sex or group differences in connectivity are rarely considered. Modelling individual heterogeneity in connectivity arising from behavioural differences is a step towards a better understanding of the differences in space use across a population (García-Sánchez et al., 2022; Zeller et al., 2023). In the Pyrenean brown bear population, outliers were not restricted by road density, contrasting to the more constrained movements of both females and males, as observed in our previous study (Kervellec et al., 2023).

However, the ability to account for individual heterogeneity in connectivity will depend on the amount of data available (Schmidt et al., 2022), because accurate inference in a more complex model requires more data. In this study, we assumed that we already knew to which group each individual detected belonged.

We used a latent state formulation to assign undetected individuals to each group. When it is not the case, finite mixture models allow to account for two or more latent groups and estimate the proportion of individuals in each group (Marrotte et al., 2022; Pledger, 2000). We provide the code for the finite mixture SCR model with ecological distance in *NIMBLE* when groups are unknown (Appendix S1). Yet, the interpretation of the groups must be done with care as they may not have biological meaning (Pledger and Phillpot, 2008).

We assumed here a dichotomy in movement variability by modelling group-based connectivity. However, other studies showed a continuous heterogeneity in connectivity (Bell, 1990; Bombieri et al., 2021). This behavioural gradient to landscape structure could be investigated with SCR models. Future extensions of SCR models could integrate a random effect on the scale parameter or the resistance parameter to avoid the need of grouping individuals into classes and investigate these ecological questions. More generally, the development of SCR models to address movement ecology questions is broad (McClintock et al., 2021). As we stated in the introduction, connectivity is unlikely to only vary spatially but also temporally. Continuous-time SCR models (Distiller et al., 2020) are another step towards a better understanding of the role of time in connectivity. Not only this would allow us to map where, but also when connectivity is the highest or the lowest, which would have direct implications for management and conservation

5. Conclusion

In summary, Spatial Capture-Recapture (SCR) models prove flexible for studying movement ecology, particularly in assessing connectivity. The study underscores the importance of explicitly modelling group-based connectivity, often overlooked in traditional analyses, in order to understand population pressures and informing effective management decisions. However, when the aim of the study is only to provide density estimates, removing outliers can be an efficient way to provide unbiased estimates without having to model individual heterogeneity. We recommend to match model complexity with the data availability and the research question, ensuring accurate ecological inferences.

Acknowledgements

We gratefully acknowledge all brown bear field co-workers from the Brown Bear Network (ROB) and from Spanish and Andorran teams. We thank Perry de Valpine for his help on the implementation of the model in Nimble.

Funding

The work was partially supported by the French National Research Agency grant ANR-16-CE02-0007. Financial and logistical supports for this study were provided by the ONCFS / OFB in France. Genetic analyses were performed by the ANTAGENE Company in France and the Servei Veterinari de Genètica Molecular, Universitat Autònoma de Barcelona in Spain.

Conflict of interest disclosure

The authors declare no conflict of interest.

Data, scripts and supplementary information availability

Appendices, scripts and simulations data are available on the online Zenodo repository (https://doi.org/10.5281/zenodo.10718643). Pyrenean brown bear detection data as well as trap locations are sensitive data and are therefore not provided in this repository.

References

Bell, W.J., 1990. Searching Behavior Patterns in Insects. Annual Reviews of Entomology 23, 447–467.

Bischof, R., Steyaert, S.M.J.G., Kindberg, J., 2017. Caught in the mesh: roads and their network-scale impediment to animal movement. Ecography 1369–1380. https://doi.org/10.1111/ecog.02801

Bombieri, G., Penteriani, V., Delgado, M. del M., Groff, C., Pedrotti, L., Jerina, K., 2021. Towards understanding bold behaviour of large carnivores: the case of brown bears in human-modified landscapes. Anim. Conserv. 24, 783–797. https://doi.org/10.1111/acv.12680

Brooks, S.P., Gelman, A., 1998. General Methods for Monitoring Convergence of Iterative Simulations. Journal of Computational and Graphical Statistics 7, 434–455. https://doi.org/10.1080/10618600.1998.10474787

Dahle, B., Støen, O.-G., Swenson, J.E., 2006. Factors influencing home-range size in subadult brown bears. Journal of Mammalogy 87, 859–865. https://doi.org/10.1644/05-MAMM-A-352R1.1

de Valpine, P., Turek, D., Paciorek, C.J., Anderson-Bergman, C., Lang, D.T., Bodik, R., 2017. Programming with models: writing statistical algorithms for general model structures with NIMBLE. Journal of Computational and Graphical Statistics 26, 403–413. https://doi.org/10.1080/10618600.2016.1172487

Distiller, G.B., Borchers, D.L., Foster, R.J., Harmsen, B.J., 2020. Using Continuous-Time Spatial Capture–Recapture models to make inference about animal activity patterns. Ecology and Evolution 10, 11826–11837. https://doi.org/10.1002/ece3.6822

Dutta, T., Sharma, S., Meyer, N.F.V., Larroque, J., Balkenhol, N., 2022. An overview of computational tools for preparing, constructing and using resistance surfaces in connectivity research. Landsc Ecol. https://doi.org/10.1007/s10980-022-01469-x

Efford, M., 2004. Density estimation in live-trapping studies. Oikos 106, 598–610. https://doi.org/10.1111/j.0030-1299.2004.13043.x

Fuller, A.K., Sutherland, C.S., Royle, J.A., Hare, M.P., 2016. Estimating population density and connectivity of American mink using spatial capture–recapture. Ecological Applications 26, 11.

García-Sánchez, M.P., González-Ávila, S., Solana-Gutiérrez, J., Popa, M., Jurj, R., Ionescu, G., Ionescu, O., Fedorca, M., Fedorca, A., 2022. Sex-specific connectivity modelling for brown bear conservation in the Carpathian Mountains. Landsc Ecol 37, 1311–1329. https://doi.org/10.1007/s10980-021-01367-8

Gardner, B., Royle, J.A., Wegan, M.T., Rainbolt, R.E., Curtis, P.D., 2010. Estimating Black Bear Density Using DNA Data From Hair Snares. J Wildl Manag 74, 318–325. https://doi.org/10.2193/2009-101

Gimenez, O., Cam, E., Gaillard, J.-M., 2018. Individual heterogeneity and capture-recapture models: what, why and how? Oikos 127, 664–686. https://doi.org/10.1111/oik.04532

Greenwood, P.J., 1980. Mating systems, philopatry and dispersal in birds and mammals. Animal Behaviour 28, 1140–1162. https://doi.org/10.1016/S0003-3472(80)80103-5

Kendall, K.C., Graves, T.A., Royle, J.A., Macleod, A.C., McKelvey, K.S., Boulanger, J., Waller, J.S., 2019. Using bear rub data and spatial capture-recapture models to estimate trend in a brown bear population. Scientific Reports 9, 16804. https://doi.org/10.1038/s41598-019-52783-5

Kervellec, M., Milleret, C., Vanpé, C., Quenette, P.-Y., Sentilles, J., Palazón, S., Jordana, I.A., Jato, R., Elósegui Irurtia, M.M., Gimenez, O., 2023. Integrating opportunistic and structured non-invasive surveys with spatial capture-recapture models to map connectivity of the Pyrenean brown bear population. Biological Conservation 278, 109875. https://doi.org/10.1016/j.biocon.2022.109875

Marrotte, R.R., Howe, E.J., Beauclerc, K.B., Potter, D., Northrup, J.M., 2022. Explaining detection heterogeneity with finite mixture and non-Euclidean movement in spatially explicit capture-recapture models. PeerJ 10, e13490. https://doi.org/10.7717/peerj.13490

McClintock, B.T., Abrahms, B., Chandler, R.B., Conn, P.B., Converse, S.J., Emmet, R.L., Gardner, B., Hostetter, N.J., Johnson, D.S., 2021. An integrated path for spatial capture–recapture and animal movement modeling. Ecology. https://doi.org/10.1002/ecy.3473

Morin, D.J., Fuller, A.K., Royle, J.A., Sutherland, C., 2017. Model-based estimators of density and connectivity to inform conservation of spatially structured populations. Ecosphere 8, 1–16. https://doi.org/10.1002/ecs2.1623

Pal, R., Sutherland, C., Qureshi, Q., Sathyakumar, S., 2021. Landscape connectivity and population density of snow leopards across a multi-use landscape in Western Himalaya. Anim. Conserv. acv.12754. https://doi.org/10.1111/acv.12754

Pledger, S., 2000. Unified Maximum Likelihood Estimates for Closed Capture–Recapture Models Using Mixtures. Biometrics 50, 494–500.

Pledger, S., 1994. Unified Maximum Likelihood Estimates for Closed Capture–Recapture Models Using Mixtures. Biometrics 50, 494–500.

Pledger, S., Phillpot, P., 2008. Using Mixtures to Model Heterogeneity in Ecological Capture-Recapture Studies. Biometrical J 50, 1022–1034. https://doi.org/10.1002/bimj.200810446

R Core Team, 2023. R : A Language and Environment for Statistical Computing.

Royle, J.A., Chandler, R.B., Gazenski, K.D., Graves, T.A., 2013. Spatial capture–recapture models for jointly estimating population density and landscape connectivity. Ecology 94, 287–294. https://doi.org/10.1890/12-0413.1

Royle, J.A., Chandler, R.B., Sollmann, R., Gardner, B., 2014. Spatial capture-recapture. Elsevier, Amsterdam.

Royle, J.A., Dorazio, R.M., Link, W.A., 2007. Analysis of Multinomial Models With Unknown Index Using Data Augmentation. Journal of Computational and Graphical Statistics 16, 67–85. https://doi.org/10.1198/106186007X181425

Royle, J.A., Fuller, A.K., Sutherland, C., 2016. Spatial capture–recapture models allowing Markovian transience or dispersal. Population Ecology 58, 53–62. https://doi.org/10.1007/s10144-015-0524-z

Schmidt, G.M., Graves, T.A., Pederson, J.C., Carroll, S.L., 2022. Precision and bias of spatial capture–recapture estimates: A multi-site, multi-year Utah black bear case study. Ecological Applications 32. https://doi.org/10.1002/eap.2618

Sutherland, C., Fuller, A.K., Royle, J.A., 2015. Modelling non-Euclidean movement and landscape connectivity in highly structured ecological networks. Methods in Ecology and Evolution 6, 169–177. https://doi.org/10.1111/2041-210X.12316

Sutherland, C., Royle, J.A., Linden, D.W., 2019. oSCR: a spatial capture–recapture R package for inference about spatial ecological processes. Ecography 42, 1459–1469. https://doi.org/10.1111/ecog.04551

Taylor, P.D., Fahrig, L., Henein, K., Merriam, G., 1993. Connectivity Is a Vital Element of Landscape Structure. Oikos 68, 571–573. https://doi.org/10.2307/3544927

Theng, M., Milleret, C., Bracis, C., Cassey, P., Delean, S., 2022. Confronting spatial capture–recapture models with realistic animal movement simulations. Ecology. https://doi.org/10.1002/ecy.3676

Thorsen, N.H., Hansen, J.E., Støen, O.-G., Kindberg, J., Zedrosser, A., Frank, S.C., 2022. Movement and habitat selection of a large carnivore in response to human infrastructure differs by life stage. Mov Ecol 10, 52. https://doi.org/10.1186/s40462-022-00349-y

Tobler, M.W., Garcia Anleu, R., Carrillo-Percastegui, S.E., Ponce Santizo, G., Polisar, J., Zuñiga Hartley, A., Goldstein, I., 2018. Do responsibly managed logging concessions adequately protect jaguars and other large and medium-sized mammals? Two case studies from Guatemala and Peru. Biological Conservation 220, 245–253. https://doi.org/10.1016/j.biocon.2018.02.015

Tourani, M., 2021. A review of spatial capture–recapture: Ecological insights, limitations, and prospects. Ecology and Evolution ece3.8468. https://doi.org/10.1002/ece3.8468

Turek, D., Milleret, C., Ergon, T., Brøseth, H., Dupont, P., Bischof, R., Valpine, P., 2021. Efficient estimation of large-scale spatial capture–recapture models. Ecosphere 12. https://doi.org/10.1002/ecs2.3385

van Etten, J., 2012. R package gdistance: distances and routes on geographical grids.

Vanpé, C., Piédallu, B., Quenette, P.-Y., Sentilles, J., Queney, G., Palazón, S., Jordana, I.A., Jato, R., Irurtia, M.M.E., de la Torre, J.S., Gimenez, O., 2022. Estimating abundance of a recovering transboundary brown bear population with capture-recapture models. Peer Comunity Journal 2. https://doi.org/doi.org/10.24072/pcjournal.199

Zeller, K.A., McGarigal, K., Whiteley, A.R., 2012. Estimating landscape resistance to movement: a review. Landscape Ecol 27, 777–797. https://doi.org/10.1007/s10980-012-9737-0

Zeller, K.A., Wattles, D.W., Conlee, L., DeStefano, S., 2019. Black bears alter movements in response to anthropogenic features with time of day and season. Mov Ecol 7, 19. https://doi.org/10.1186/s40462-019-0166-4

Zeller, K.A., Wultsch, C., Welfelt, L.S., Beausoleil, R.A., Landguth, E.L., 2023. Accounting for sex-specific differences in gene flow and functional connectivity for cougars and implications for management. Landsc Ecol 38, 223–237. https://doi.org/10.1007/s10980-022-01556-z

Zhang, W., Chipperfield, J.D., Illian, J.B., Dupont, P., Milleret, C., De Valpine, P., Bischof, R., 2023. A flexible and efficient Bayesian implementation of point process models for spatial capture–recapture data. Ecology 104, e3887. https://doi.org/10.1002/ecy.3887