Unravelling the effects of heterogeneity in space use on estimates of connectivity and population size: 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[1,7]

[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@gmail.com

Abstract

Spatial Capture-Recapture (SCR) models using least-cost path distance offer an unified framework to estimate landscape connectivity and population size from individual detection data, while accounting for individual and spatial heterogeneity in space use. In our case study on the Pyrenean brown bear population, two "outliers" individuals were detected more often over very large spatial extent compared to other individuals. The integration of such individuals made SCR modelling challenging, especially since it remains unclear how unmodelled heterogeneity in space use may bias connectivity and population size estimates. To address this gap, we used simulations reflecting the Pyrenean brown bear population, with two groups of individuals differing in their space use due to individual characteristics but also in their spatial responses to landscape structure. We compared six SCR model formulations that varied in whether individual and spatial heterogeneity in space use were (1) ignored, (2) explicitly modelled, or (3) handled by removing outliers. We then applied the same models to the empirical bear dataset.

By combining simulations and sensitivity analyses, we highlight the challenges of choosing the appropriate modelling approach when multiple sources of heterogeneity in space use

occur simultaneously. Our results show that the treatment of heterogeneity in space use should match the research objective: removing outliers supports accurate population size estimation, while explicitly modelling heterogeneity is essential for reliable connectivity assessment. Overall, we provide a practical framework for identifying and addressing heterogeneity in SCR applications, to guide practitioners in deciding when additional model complexity is warranted.

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 (e.g., resource availability, risk to encounter predators) (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. 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 estimated using 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 animal movement 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). Despite the importance of intraspecific variation, 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 spatially explicit individual detection histories to estimate both population size and distribution. A core assumption of SCR models is that detection probability declines with distance from an individual's activity centre (Royle et al., 2014). In other words, the detection function is designed to reflect individual space use, as it shapes where and how often an individual is detected. In the context of SCR models, we define space use as the spatial distribution of detection probability, with baseline detection probability set to 1 ($p_0 = 1$), and conditional on the estimated location of the activity centre. In its most commonly used form, a SCR model assumes a circular space use resulting in a symmetrical and stationary (invariant across space) space use extent. However, extensions of SCR models allow the estimation of nonstationary, asymmetrical individual space use extent caused by spatial heterogeneity in the landscape (Royle et al., 2013). In these models, distance in the detection function is not Euclidean but based on the least-cost path (LCP) distance, with landscape resistance directly inferred from the data. As a result, SCR models with LCP distance provide a unified framework for jointly estimating both population size and the effects of landscape resistance on space use (Royle et al., 2013; Sutherland et al., 2015), thereby quantifying population-level landscape connectivity. This approach has been applied to several species, including American mink (*Neogale vison*, Fuller et al., 2016), black bear (*Ursus americanus*, Morin et al., 2017), jaguar *(Panthera onca*, Tobler et al., 2018), snow leopard (*Panthera unica*, Pal et al., 2021), and brown bear (*Ursus arctos*, Kervellec et al., 2023). However, these studies focused on population-level connectivity and did not account for individual heterogeneity in space use, particularly in the resistance parameter that scales the spatial heterogeneity in space use.

Individual heterogeneity refers to variability in responses among individuals within a population, which may be linked to measurable or non-measurable individual characteristics (Gimenez et al., 2018). In non-spatial capture-recapture models, it is well established that accounting for individual heterogeneity in detection probability is crucial to prevent negative bias in population size estimates (Gimenez et al., 2018). In SCR models, heterogeneity in detection probability can arise from individual variation in the baseline detection probability as well as from individual and spatial heterogeneity in space use. Our focus here is on the latter. Non-Euclidean SCR models typically describe heterogeneity in space use with two parameters: (1) the scale parameter σ, which defines the extent to which an individual can be detected from its activity centre and is related to home range size (Royle et al., 2014), and (2) the resistance parameter α_2, which quantifies the influence of spatial covariates on space use across the landscape (i.e. spatial heterogeneity). Therefore, the scale parameter σ can vary among individuals to reflect individual variation in space use while the resistance parameter α_2 quantifies how individuals adjust their space use to spatial heterogeneity. Not all individuals may respond similarly to landscape structure, leaving the possibility for individual-specific α_2 parameters. Accounting for individual heterogeneity in the scale

parameter is important to prevent negative bias in population size estimates (Gardner et al., 2010; Sollmann et al., 2011). SCR models are generally robust to unmodelled spatial heterogeneity in space use (Royle et al., 2016; Theng et al., 2022) but strongly structured landscapes can nevertheless cause an underestimation of population size (Sutherland et al., 2015). In addition, unmodelled individual heterogeneity in the response of space use to landscape structure remains unknown. Although previous studies have shown that unmodelled heterogeneity in space use, whether arising from individual traits or from the location of individuals in the landscape, can bias population size estimates (Gardner et al., 2010; Sutherland et al., 2015), much less is known about the combined effects of individual and spatial heterogeneity on estimates of both population size and connectivity.

The motivation for this study stems from our work on the Pyrenean brown bear population (Kervellec et al., 2023; Sanz-Pérez et al., 2025). Nearly extinct in 1995, this population is now recolonising human-dominated landscapes (Kervellec et al., 2023; Sanz-Pérez et al., 2025). In this context, robust estimates of population size and a clear understanding of how landscape structure affects connectivity are critical. In a previous study, Kervellec et al. (2023) identified that two individuals were detected far more frequently and had a much larger apparent space use than the rest of the population. The common practice in this case is to remove these "outliers" individuals from the analyses, in order to avoid biased population size estimates due to unmodelled heterogeneity in the space use component of the detection probability (Kendall et al., 2019; Schmidt et al., 2022). However, such wide-ranging individuals may in fact play a key role when the goal of the study is to quantify population connectivity.

In this study, we investigate how heterogeneity in space use affects (1) population size estimate and (2) connectivity estimates, using non-Euclidean SCR models. For each objective, we built six different models to evaluate the consequences of (a) ignoring individual or/and spatial heterogeneity in space use or (b) removing the individuals with large space use ("outliers"). We first conducted a simulation study inspired by the Pyrenean brown bear population, with two groups of individuals differing in both space use extent and response to landscape structure. We then applied the same models to the empirical Pyrenean brown bear data. We expected that (1a) population size would be underestimated when individual heterogeneity in space use is not accounting for (Gimenez et al., 2018; Sutherland et al., 2015), (1b) removing outliers would yield unbiased population size estimates, and (2) removing outliers and not accounting for individual heterogeneity would undermine estimates of space use and connectivity.

2. Materials 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 spatial domain 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. For inhomogeneous point process, 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 space use extent. The larger the σ, the further from its activity centre an individual can be detected. These parameters can be constant or varying according to individual covariates (e.g., sex) to account for individual heterogeneity in the detection process.

By using the Euclidean distance, we assume that space use is symmetric and stationary (Royle et al., 2014), meaning that the presence of barriers or resources do not affect movements of individuals across the landscape. 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) $$d_{lcp}(\nu, \nu') = \min_{\mathcal{L}_1,\dots,\mathcal{L}_l} \sum_{g=1}^{m} \mathrm{cost}(\nu_g, \nu_{g+1}) \times d_{euc}(\nu_g, \nu_{g+1})$$

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 $\alpha_2$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 decreases ($\alpha_2 > 0$) or increases ($\alpha_2 < 0$) the space use extent 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. The resistance parameter therefore quantifies the extent to which spatial characteristics of the landscape influences individuals space use (spatial heterogeneity in space use). Similarly to σ, α_2 can vary according to individual characteristics according to a function of individual covariates (spatial and individual heterogeneity in space use). 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 (i.e. realized densities), also called density weighted connectivity (Morin et al., 2017).

2.2 Simulations

We conducted a simulation study to assess the consequences of not accounting for individual and spatial heterogeneity in space use on parameter estimation. We considered a population composed of two groups differing in their responses to landscape structure. The study area was a 25 x 25 distance unit (du) square domain, within which we centred an array of 100 traps spaced by a minimum of 2du. We added a 3.5du buffer around traps, following standard 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 distance to a road (see map Appendix S1).

The simulation study was designed to mimic issues encountered with the Pyrenean brown bear population (Kervellec et al., 2023). We simulated 100 individuals, assuming a homogeneous distribution of their activity centres within the spatial domain. The population was composed of two groups. The first encompassed 97% of the population (n = 97) and was called the "common group", while the second one gathered the outlier individuals (n=3) characterized by a larger space use and was called the "outlier group". We assumed that the larger space use of the "outlier" group was linked to individual variation in home range size (σ_{Com} = 7; σ_{Out} = 5), as well as a different response to the road. The common group

have smaller space use when the distance to road decreased ($\alpha_{2_{Com}}$= 3) while the outlier group was not affected ($\alpha_{2_{Out}}$= 0). To ensure realistic (i.e. number of detections per individual ≤ 30) detectability rates between the two groups despite their different space use, we also set a higher baseline detection probability for individuals of the common group ($p_{0_{Com}}$ = 0.2) compared to the outlier group ($p_{0_{Out}}$ = 0.1).

We simulated 25 detection histories. For each simulation, we fitted six models: four using the full data (denoted MOx) and two with all detections of all outlier individuals removed (denoted Mx). We considered models assuming an effect (MO3, MO4, M4) and no effect (MO1, MO2, M1) of landscape structure on space use. For datasets including outliers (MOx), we sequentially increased the number of model parameters (p_0, σ, α_2) that accounted for individual heterogeneity through a group effect (Table 1).

The simulations were performed using the R package *nimble* (version 1.01 de Valpine et al., 2017). We ran two chains of 50,000 iterations each and discarded the first 1,000 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). We used a data-augmentation approach, with an augmented population of 150 individuals (Royle et al., 2007), and priors are presented in the Appendix S2 (Table S2.A). For the six models, we compared parameter estimates with the true values and recorded whether 95% credible intervals encompassed the truth. We also illustrated differences in connectivity estimates by mapping potential connectivity for models MO3, MO4 and M4 with estimates from the simulated dataset 1. because the 25 simulated datasets were primarily intended to identify trends, we did not compute bias explicitly.

Table 1 – *Models formulas used for the simulation study and applied on each simulated datasets. It present the model name, the dataset, the baseline detection probability (p_0), the scale parameter (σ), the resistance parameter (α_2) when applicable. The empty cells means that the resistance parameter was not included in the model and therefore we used the Euclidean distance instead of the least-cost path distance.*

Model name	dataset	p_0	σ	α_2
MO1	all	~1	~1	
MO2	all	~group	~group	
MO3	all	~group	~group	~1
MO4	all	~group	~group	~group
M1	outlier removed	~1	~1	
M4	outlier removed	~1	~1	~1

2.3 Case study: Brown bears in the Pyrenees

2.3.1 Data

In this case study, we used the spatial capture-recapture dataset of the Pyrenean brown bear population monitored 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. An opportunistic trap is the centre of each grid cell where at least one depredation of a bear to livestock or beehives occurred in the cell between 2010 and 2020. The depredation events are reported by the owner. In order to get financial compensation, evidences of bears presence and genetic samples (hair or scat) are 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. The maximum distance moved per year, which is the maximum distance between the centroid of the detections and the detections, for Néré was between 39.8 and 66.5km and was between 27.6 and

51.1km for Goiat. The maximum distance moved per year per individual between detections was between 0 (i.e. individuals detected once) and 24.3km for the females (mean = 3.5km) and between 0 and 17.4km for the males (mean = 6.1km) (Figure 2). In average Néré was detected 9 times, Goiat 35.7 times, while males were detected 5.6 times and females 4.1 times per year.

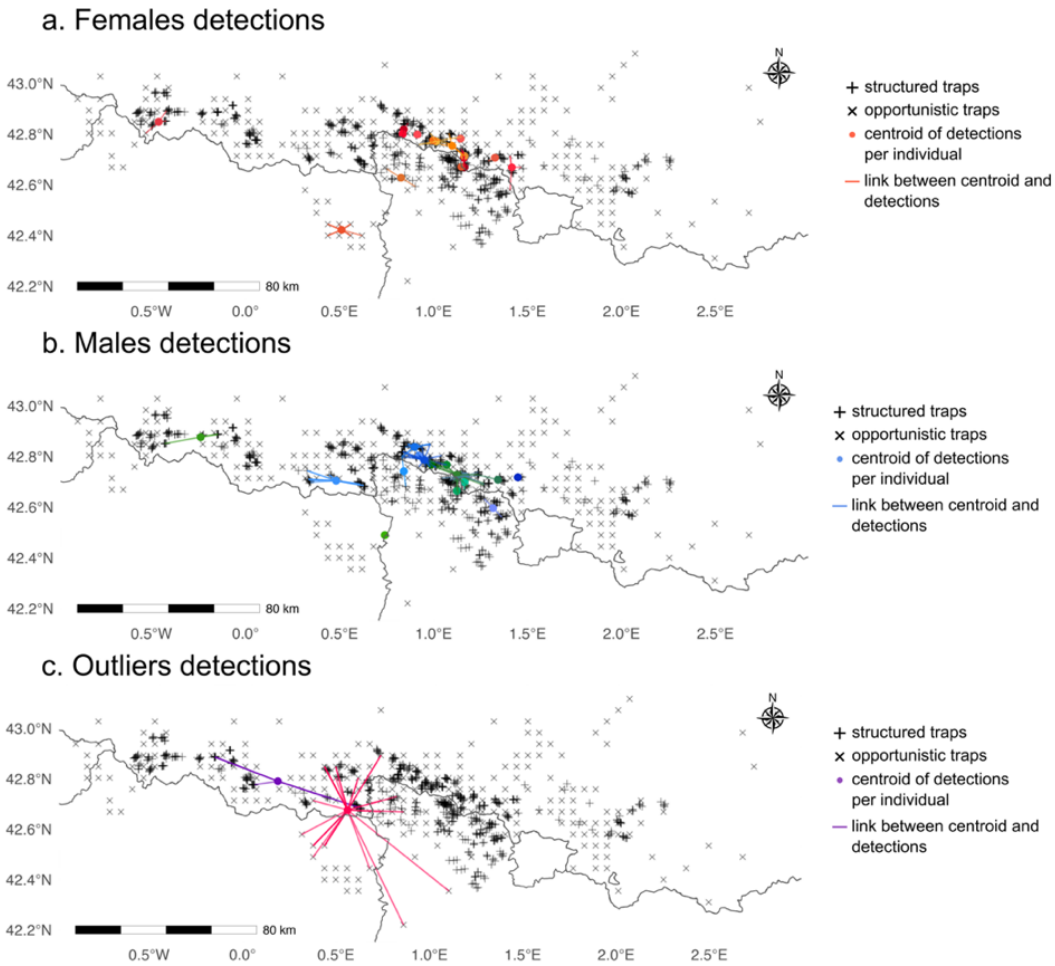


Figure 1 - Maps of traps distribution from the structured (+) and the opportunistic (x) monitoring across the Pyrenees mountains (France, Spain, Andorra). Points represents the centroid of detections per individuals and the lines links the centroid to the centroid to the traps where each individual was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

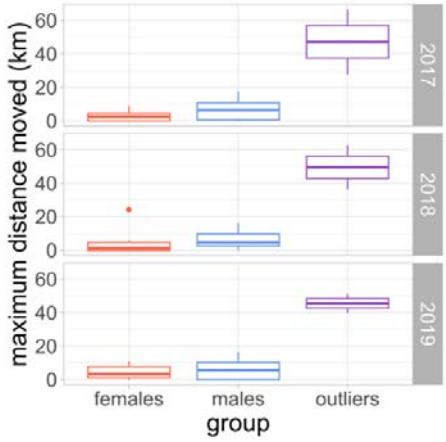


Figure 2 – Boxplot of the maximum distance moved in km for each group (females, males and outliers) between 2017 and 2019 for the Pyrenenan brown bear population. This distance is the maximum distance between each detection and the centroïd of its detections for each individual

Table 2 - *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 in the Pyrenean brown bear population.*

	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 monitoring 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 histories and $p_{Struct_{ijkt}}$ and $p_{Opp_{ijkt}}$ are the detection probabilities 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) using Equation 2.

(6) $$\text{logit}(p_{0_{Struct_{ijkt}}}) \sim \alpha_{Struct} + \beta_b \times \text{Behav}_{ijkt} + \beta_{t1} \times \text{Nvisits}_{jkt} + \beta_{t2} \times \text{Country}_j + \beta_{t3} \times \text{Trap}_j + \beta_{t4} \times m_k + \beta_{t5} \times m_k^2$$

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

The baseline detection probability (Equation 6) for the structured monitoring was assumed to be a function of a quadratic effect (β_4, β_5) of the sampling occasion (m) to account for higher detectability in the summer, the country (Country, β_2), the number of visits at the trap each month (Nvisits, β_1) as samples can be collected once or twice each month, and if the hair snag was combined or not with a camera trap (Trap, β_3). For the opportunistic monitoring (Equation 7), the baseline detection probability was a function of the country (Country, β_6) and the quadratic effect of the month (m, β_4, β_5). Both monitoring accounted for a behavioural effect (Behav, β_b), 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). We also estimated a different intercept for the structured (α_{Struct}) and opportunistic sampling (α_{Opp}).

Additionally, the scale parameter was estimated for each sex: $\sigma_i \sim \alpha_s + \beta_s \times \text{sex}_i$, for each group: $\sigma_i \sim \alpha_s + \beta_s \times \text{group}_i$, or was constant (Table 3).

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

(8) $$\log(\lambda(s)) = \beta_{d1} \times \text{Rug}_s + \beta_{d2} \times \text{HumDens}_s$$

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.

Spatial heterogeneity in space use was modelled using the LCP distance instead of the Euclidean distance using Equation 3. The resistance surface was based on road density, computed as the length of all roads within each 2.5x2.5km pixel extracted from Open Street Map (see Kervellec et al., 2023). We considered two models with no effect of spatial heterogeneity on space use, but considering individual variation with an effect of sex (MFO1) and group (MFO2) on σ. The group is a categorical variable that differentiate between females, males, and outliers. For the other models, we added an effect of spatial heterogeneity in space use assuming that all individuals responded similarly to landscape structure (MFO3) or that response to landscape structure was a function of the group (MFO4). We also added two models where the outliers were removed from the dataset, first modelling only individual heterogeneity in σ between males and females with no effect of spatial heterogeneity on space use (MF1) and, second modelling individual and spatial heterogeneity in space use (MF4).

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

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

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* (Bischof et al., 2020; 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)

Table 3 – *Models formulas used for the simulation study and applied on each simulated datasets. It present the model name, the dataset, the scale parameter (σ), the resistance parameter (α_2) when applicable. The empty cells means that the resistance parameter was not included in the model and therefore we used the Euclidean distance instead of the least-cost path distance.*

Model name	dataset	σ	α_2
MFO1	all	~sex	
MFO2	all	~group	
MFO3	all	~sex	~1
MFO4	all	~group	~group
MF1	outlier removed	~sex	
MF4	outlier removed	~sex	~sex

3. Results

3.1 Simulation

The simulated data included three outlier individuals that were always detected, while in the "common group" between 42 and 59 individuals were detected out of the 97 present in the population. As expected from the parameter settings, outliers were detected more frequently, with an average of 12.7 detections per individual compared to 2.7 in the common group. All models converged with $\hat{R} \leq 1.08$ and n.eff > 113.

3.1.1 Population size estimates

Estimates of population size were underestimated when we ignored individual and spatial hetereogenity in space use (Figure 3). When keeping all the individuals in the dataset, the model MO1, that was not accounting for individual and spatial heterogeneity in space use, consistently underestimated population size (95%BCI never cross the true value). Introducing individual heterogeneity in space use (model MO2) led to a mean estimate below the true value in all simulations, with the 95%BCI overlapping the true value for only 0.08% of the simulations. Accounting for spatial heterogeneity in space use (model MO3) resulted in 92% of 95%BCI overlapping the true value. When modelling individual heterogeneity on the resistance parameter (model MO4), 24% of simulations (8/25) were below the true value, and 92% of 95%BCI overlapped the true value (Figure 3). When the three outliers were removed from the dataset, the null model that ignored spatial and individual heterogeneity in space use (model M1) underestimated population size, with 95%BCI overlapping the true value in only 0.08% of the simulations. By contrast, model (M4) that integrated individual and spatial heterogeneity in space use achieved 92% overlap (23/25 simulations).

3.1.2 Connectivity estimates

Three models (MO3, MO4 and M4) allowed to look into connectivity estimates as they simultaneaously estimated individual and spatial heterogeneity in space use. Ignoring individual heterogeneity in the resistance parameter led to a resistance parameter close to the common group estimate that encompassed most of the population, with mean estimated α_2 between 2.2 and 3.3 (MO3) compared to between 2.6 and 3.5 for the common group and between -0.9 and 1.4 for the outlier group (MO4) (Figure 3). By contrast, removing outlier indiduals led to results comparable to when individual and spatial heteorgeneity in space use was modelled, with mean estimated α_2 between 2.6 and 3.5 (M4). To illustrate how these differences translate into the connectivity maps, we compared relative potential connectivity across these models (Figure 4), accounting for space use extent (σ) and its variability across the landscape (α_2). Using simulated dataset 1 as an example, the three model formulations produced similar overall patterns. However, ignoring individual variation in the response to spatial heterogeneity in the landscape led to an overestimation of connectivity, especially visible in areas far away from the barrier (MO3 relative bias between -15% and 85.3%). Conversely, removing outliers led to an underestimation of connectivity close the barrier (M4 relative bias between - 86.7% and 26.1%).

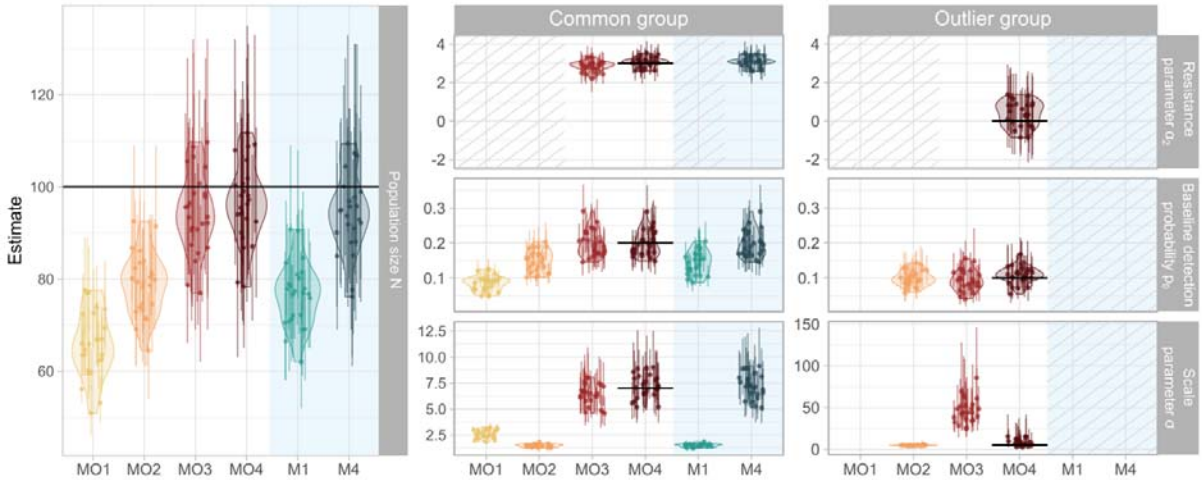


Figure 3 - Model parameter estimates across the six models compared (see Table 1) in the simulation study. The mean estimate of each simulation is represented by a point with errors bar displaying the 95% credible intervals. Violin plots illustrate the distribution of the mean estimated parameters for each model across the 25 simulations. Horizontal lines indicate the true parameter values from model MO4, which was used to generate the simulated datasets. The first column shows population size estimates for the entire population. The second and third columns display parameter estimates for the common group and the outlier group, respectively. Within these columns, rows correspond to specific model parameters: the resistance parameter α_2 (top row), baseline detection probability p_0 (middle row), and scale parameter σ (bottom row). Background colors indicate data inclusion, with a white background representing models using all the dataset and a blue background indicating models with outliers removed. Striped areas denote parameters that were not estimated in the model. Empty spaces indicate cases where a single parameter was estimated for the entire population, with the reference parameter displayed in the

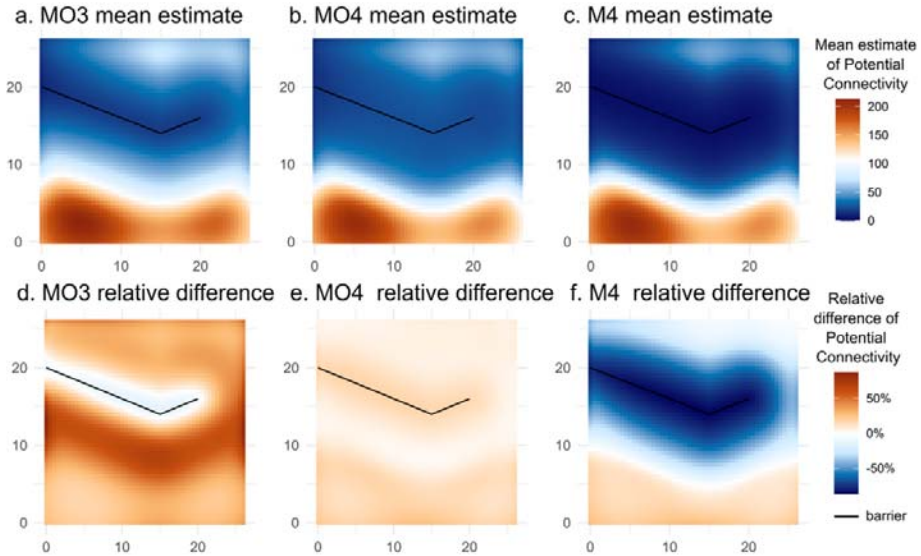


Figure 4 – Illustration, for simulated dataset 1, of mean estimate (top row: a, b, c) and relative difference (bottom row: d, e, f) maps of potential connectivity. The left panels (a, d) show results for model MO3, which estimated a single resistance parameter $\alpha_2$for the entire population. The central panels (b, e) show results for model MO4, which estimated a separate resistance parameter $\alpha_2$for each group. The left panels (c, f) show results for model M4, where outlier individuals were removed from the analyses.

3.2 Case study

The models MFO1, MFO2, MFO4, MF1 and MF4 did converge with $\hat{R} \leq 1.08$ and effective sample size > 200. On the contrary, the mixing for the resistance parameter in the model MFO3 was poor with an effective sample size of only 93 and $\hat{R}$ = 1.

3.2.1 Population size estimates

In our case study, modelling individual heterogeneity in space use only (i.e. accounting that males are made of two groups, model MFO2) led to an increase of the estimated number of males (from $N_{MFO1_{males}}$ = 23 [20; 28] to $N_{MFO2_{males}}$ = 29 [21; 42] and 2 outliers) and a decrease of the estimated number of females (from $N_{MFO1_{females}}$ = 130 [85; 173] to $N_{MFO2_{females}}$ = 79 [50; 122]), compared to the null model (MFO1) (Figure 5). However, including spatial heterogeneity in space use did not affect population size estimates whether only one resistance parameter was estimated ($N_{MFO3_{females}}$ = 124 [80; 170], $N_{MFO3_{males}}$ = 23 [20; 29] and 2 outliers) or one resistance parameter was estimated for each group ($N_{MFO4_{females}}$ = 74 [47; 119], $N_{MFO4_{males}}$ = 29 [21; 42] and 2 outliers).

Removing the outliers from the dataset led to similar estimated population size than modelling individual heterogeneity ($N_{MF1_{males}}$ = 33 [23; 47], $N_{MF1_{females}}$ = 60 [39; 92]). Similarly, modelling spatial heterogeneity in space use when the outliers were removed led to comparable population size estimates ($N_{MF4_{males}}$ = 33 [23; 50], $N_{MF4_{females}}$ = 60 [39; 91]) (Figure 5).

The two models where outliers were removed had comparable estimates of population size (N_{MF1} = 93 [67; 131] to N_{MF4} = 94 [67; 133] (Figure 5)).

3.2.2 Connectivity estimates

Integrating outlier individuals without modelling individual heterogeneity (model MFO1) led to a 2.7 times increase in the median estimated σ of males (σ_{MF01} = 0.15 [0.14; 0.17]), compared to when individual heterogeneity in space use was modelled ($\sigma_{Males_{MFO2}}$ = 0.06 [0.06; 0.07]) or when removing outliers ($\sigma_{Males_{MF1}}$ = 0.06 [0.05; 0.07]) (Figure 5).

Overall, the estimated resistance parameters was close to 0, whether we kept all the individuals in the dataset ($\alpha_{2_{Males_{MFO4}}}$ = 0.21 [0.03; 0.35], $\alpha_{2_{Outliers_{MFO4}}}$ = -0.16 [-0.64; 0.24] and $\alpha_{2_{Females_{MFO4}}}$ = 0.27 [0.01; 0.43]) or we removed the outliers ($\alpha_{2_{Males_{MF4}}}$ = 0.21 [0.02; 0.35] and $\alpha_{2_{Females_{MF4}}}$ = 0.27 [0.11; 0.42] (Figure 5)). However, ignoring individual heterogeneity in the resistance parameter when keeping all the individuals in the dataset led to an estimated resistance parameter that encompassed 0 and increased parameter uncertainty ($\alpha_{2_{MFO3}}$ = -0.26 [-0.45; 0.42]) (Figure 5).

The map of density weighted connectivity predicted from model MFO4 showed that females' movements were more restricted than males' in high road density areas (Figure 6ac). However, the uncertainty of $\alpha_{2_{Outlier}}$ was propagated to the potential connectivity map, which did not allow us to draw any conclusion about their space use extent (Figure 6ef).

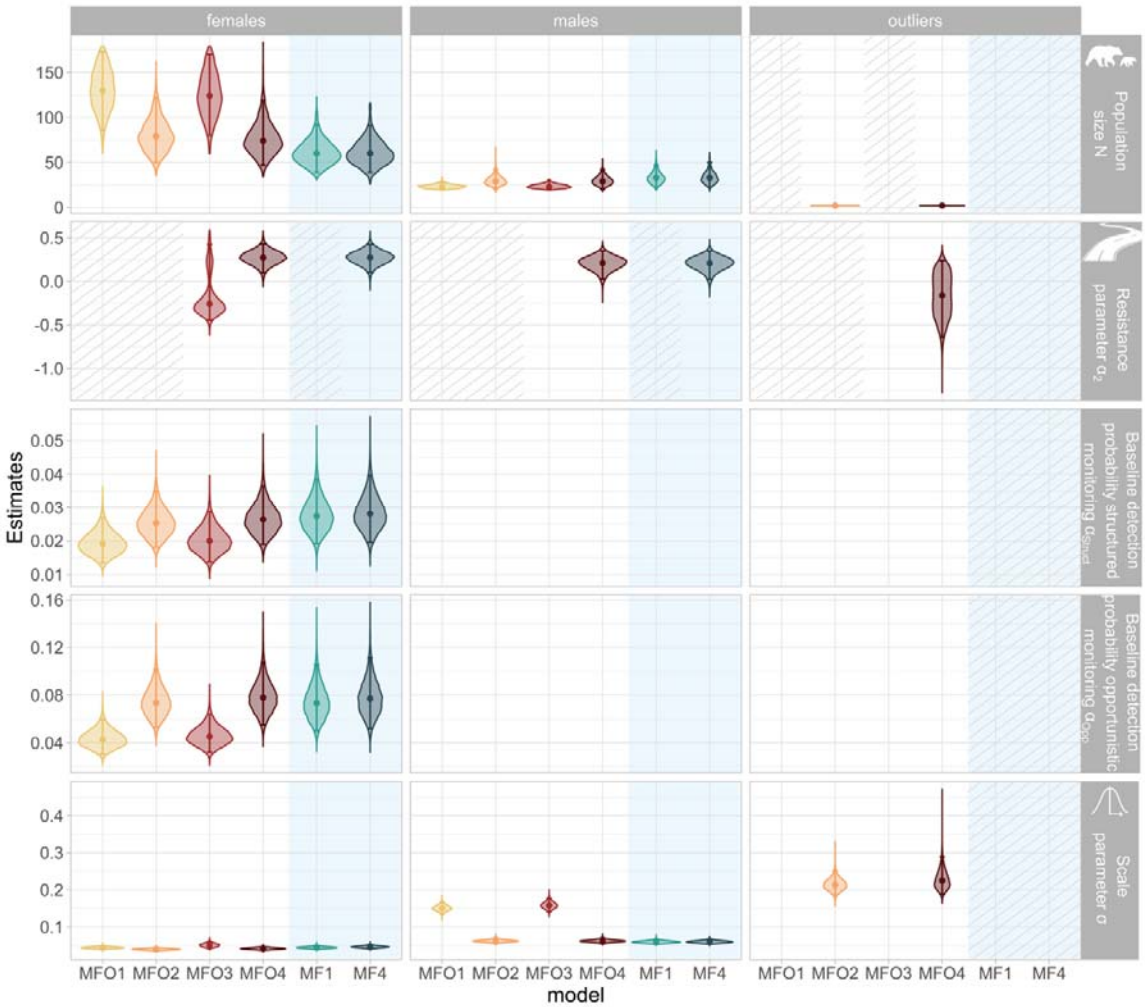


Figure 5 - Violin plot of estimated population size N (top row), resistance parameter α_2 (second row), intercept of baseline detection probability for the structured monitoring α_{Struct} (third row), intercept of baseline detection probability for the opportunistic monitoring α_{Opp} (fourth row), and scale parameter σ (bottom row) for the Pyrenean brown bear population. The median of each distribution is represented by a point with errors bar displaying 95% credible intervals. The panels in column display the estimated parameter's posterior distribution for the females (left column), males (middle column) and outliers' (right column) Pyrenean brown bear population in 2019 in function of the six models compared. Background colours indicate data inclusion, with a white background representing models using all the dataset and a blue background indicating models with outliers removed. Striped areas denote parameters that were not estimated in the model. Empty spaces indicate cases where a single parameter was estimated for the entire population, with the reference parameter displayed in the female panel, except for the estimated distribution of the scale parameter in model MFO1 and MFO3 that is common to the males and the outliers and this parameter is therefore display in the male panel.

407

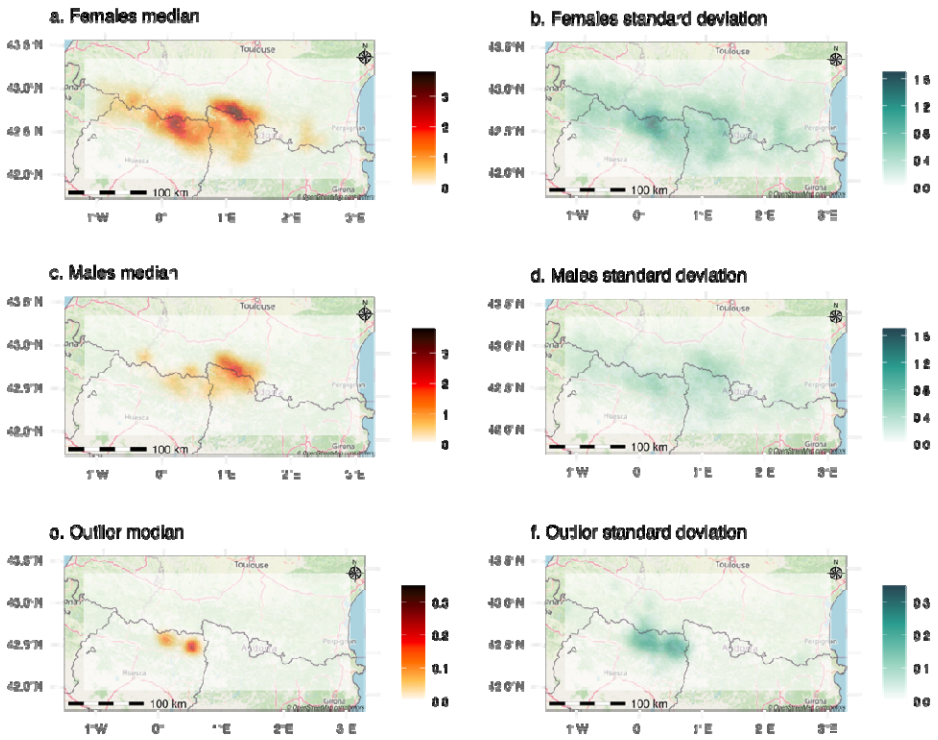


Figure 6 - Maps of median density weighted connectivity derived from estimates of model MFO4, for females (a), males (b) and the two outlier individuals (e) with respective standard deviation (b, d, f) for the Pyrenean brown bear population in 2019.

4. Discussion

4.1. Population size estimates

As predicted, the simulation study revealed a negative bias in population size estimates when individual and spatial heterogeneities in space use were ignored (Pledger, 2000; Royle et al., 2013). This bias was strongest when both types of heterogeneity remained unmodelled.

Although modelling individual heterogeneity is challenging, especially when part of the population remains undetected (Marrotte et al., 2022; Pledger, 2000), we demonstrated how latent state formulations in SCR models can distinguish and estimate individual and spatial heterogeneity in space use. Conversely, when the primary objective is to estimate population size, removing outlier individuals with larger space use extent provides a straightforward way to reduce individual heterogeneity in space use and avoid underestimating population size (Kendall et al., 2019; Schmidt et al., 2022). However, this strategy is only effective if individual heterogeneity in space use is driven by a small number of individuals. Moreover, removing outlier individuals does not address spatial heterogeneity in space use that may remain in the population and would still require explicit modelling.

In our case study, understanding the combined effect of individual and spatial heterogeneity in space use on population size estimates was not straightforward. Applying the models to the Pyrenean brown bear population showed that individual heterogeneity could be accounted for, even with the low number of outliers (n=2). As expected, the estimated number of males was larger after accounting for individual heterogeneity (from 23 [20; 28] to 29 [21; 42] and 2 outliers). Surprisingly, the estimated number of females was smaller after accounting for individual heterogeneity (from 130 [85; 173] to 79 [50; 122]). In this study, we have not been able to unambiguously identify the reason of this observed pattern. When modelling individual heterogeneity,

we noticed an increase in the intercept of the opportunistic baseline detection probability (from $\alpha_{Opp_{MFO1}}$ = 0.04 [0.03; 0.06] to $\alpha_{Opp_{MFO2}}$ = 0.07 [0.05; 0.10]), while intercept of the structured baseline detection probability remained unchanged (from $\alpha_{Strcut_{MFO1}}$ = 0.02 [0.01; 0.03] to $\alpha_{Struct_{MFO2}}$ = 0.03 [0.02; 0.03]), linked to the presence of outliers in the population.

The estimated number of females was also larger than the 52 individuals detected in 2019 (with 22 males, 23 females and 7 individuals whose sex remains unknown)(Vanpé et al., 2022). A hypothesis to explain this apparent overestimation is that females have a smaller space use extent compared to males ($\sigma_{MFO2_{females}}$ = 0.040 [0.036; 0.044] and $\sigma_{MFO2_{males}}$ = 0.060 [0.057; 0.070]). From the model point of view, the small space use of female makes possible to miss them between traps, especially in areas that are covered solely by opportunistic traps. Additionally, in the south west part of the study area (Ordesa and Monte Perdido National Park), the monitoring intensity was low and even the habitat was perceived favourable (i.e. low human density and high ruggedness). This led to the prediction of high density of females in that area, which is unlikely given the low number of observed individuals in that area by the additional observations not included in this study. Unbiased estimation of population size will require working towards alternative model formulation and study design approaches to deal with those undesirable issues.

Furthermore, including spatial heterogeneity in space use in the model led to similar population size estimates compared to when only individual heterogeneity in space use was included ($N_{MFO4_{females}}$ = 74 [47; 119], $N_{MFO4_{males}}$ = 29 [21; 42] and 2 outliers). The estimated resistance parameter was close to zero for each group (respectively $\alpha_{2_{MFO4_{females}}}$ = 0.27 [0.10; 0.43], $\alpha_{2_{MFO4_{males}}}$ = 0.21 [0.03; 0.35] and $\alpha_{2_{MFO4_{outliers}}}$ = -0.16 [-0.64; 0.24]) meaning that accounting for spatial heterogeneity in space use was not necessary in our case if the objective is to estimate population size (Sutherland et al., 2015).
Removing the outliers led to similar results that modelling individual heterogeneity ($N_{MF1_{males}}$ = 33 [23; 47], $N_{MF1_{females}}$ = 60 [39; 92]) demonstrating that this practice is a cost effective solution when the aim was to estimate population size (Kendall et al., 2019; Schmidt et al., 2022).

4.2 Connectivity estimates

The resistance parameter reflects whether space use extent increases ($\alpha_2 < 0$) or decreases ($\alpha_2 > 0$) as a response to a landscape covariate (Royle et al., 2013; Sutherland et al., 2015). In our simulations, the common group had smaller space use extent close to the barrier, while the outlier group had larger space use extent and was not affected by landscape structure. Under these conditions, the model that ignored the individual heterogeneity in the resistance parameter (model MFO3) estimated a resistance parameter skewed toward the common group value, suggesting that all individuals in the population reduced space use near the barrier. A comparable conclusion would be drawn when removing outlier individuals from the data set with a single estimated resistance parameter.
Potential connectivity maps enable to map population space use by assuming an activity centre in each cell of the spatial domain and summing the space use of each individual (Morin et al., 2017). Comparing maps from models MFO3, MFO4 and MF4 showed consistent reductions in connectivity near barriers, but values were sensitive to model specification (Figure 4). Ignoring individual heterogeneity in the resistance parameter (model MFO3) led to overestimating potential connectivity further away from the barrier, whereas removing outliers (model MF4) led to underestimating potential connectivity at the barrier. These results highlight the need for caution when comparing potential connectivity maps.

In the case study, the estimated resistance parameter demonstrated that both males and females exhibited smaller space use extents in areas with higher road density, and their responses were similar across sexes (Figure 5).
Density Weighted Connectivity (DWC) maps provide an integrated representation of population space use by combining realized densities, individual heterogeneity in space use extent (scale parameter σ), and heterogeneous responses to landscape features (resistance parameter α_2). These maps allow delineation of areas with high space use as well as those where accessibility is limited by the road network. In addition, the uncertainty maps associated with DWC showed that the outlier group map

uncertainty is too high to support reliable conclusions, while the map for females identifies in Ordesa and Monte Perdido National Park, elevated uncertainty that suggests the need for further investigation in this area to refine space use assessment (Figure 6).
DWC maps are based on realized densities which are known to be sensitive to the location of traps in the landscape (Durbach et al., 2024), contrary to predicted maps that are based on the estimated relationship between density and landscape variables. Given the difficulty to describe spatial variation in the Pyrenean brown bear density based on spatial covariates (Sanz-Pérez et al., 2025), and considering the intensive monitoring of the population, it is likely that the realized densities more accurately reflect the population distribution than predicted maps.

Overall, the effects of individual and spatial heterogeneity in space use on connectivity estimates are much more challenging to interpret than that on population size: they are sensitive to model formulation, trap location, and there is no clear parameter comparison to validate the outputs. Thus, interpreting SCR-based connectivity requires caution, even though jointly modelling spatial and individual heterogeneity remains conceptually appealing.

4.3 Challenges and perspectives

Our findings highlight the importance of tailoring the structure of SCR model to the research objectives. When the objective is to estimate population size, removing outlier individuals is an efficient solution that avoids unnecessary complexity. When the goal is to assess connectivity, however, disentangling individual and spatial heterogeneity is essential but more challenging.

Identifying outliers is a first challenge and itself subjective. In our case, we examined the distribution of the capture frequency and the maximum distance moved (MDM) as indicators, both of which can reveal unusual detectability patterns. While further research is necessary to identify outliers, these summary statistics are a good first approach to detect obvious outliers (Figure 2, Table 2). In addition, comparing population size estimates with the presence and absence of outliers may be a good test to decide whether accounting for individual heterogeneity is necessary. Large differences across models indicate that results are sensitive to heterogeneity assumptions, underscoring the importance of testing alternative formulations.

Accounting for individual heterogeneity also depends on data availability (Schmidt et al., 2022). In this study, we assumed known group membership for detected individuals and used latent state formulations to assign undetected individuals. When group membership is unknown, finite mixture models can be applied (Marrotte et al., 2022; Pledger, 2000), though interpretation of latent groups requires caution (Pledger and Phillpot, 2008). We provide the code for the finite mixture SCR model with ecological distance in *nimble* when groups are unknown (Appendix S1).

5. Conclusion

Our study underscores that SCR models must be tailored to research goals. When the objective is to estimate population size, removing outliers can provide efficient and unbiased estimates without increasing model complexity. When the objective is to assess connectivity, however, explicitly modelling both spatial and individual heterogeneity is necessary, despite the challenges. Ultimately, model complexity should be matched to both data availability and the ecological questions at hand to ensure robust and meaningful inference for management and conservation.

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. The simulations were performed on the Core Cluster of the Institut Français de Bioinformatique (IFB) (ANR-11-INBS-0013). We thank the two referee, for their helpful and constructive comments.

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. This research was partly funded by Biodiversa+, the European Biodiversity Partnership, in the context of the Big_Picture project under the 2022-2023 BiodivMon joint call. It is co-funded by the European Commission (GA No. 101052342) and the French Agence Nationale de la Recherche.

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.17259654). 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 35, 447–467. https://doi.org/10.1146/annurev.en.35.010190.002311

Bischof, R., Milleret, C., Dupont, P., Chipperfield, J., Tourani, M., Ordiz, A., de Valpine, P., Turek, D., Royle, J.A., Gimenez, O., Flagstad, Ø., Åkesson, M., Svensson, L., Brøseth, H., Kindberg, J., 2020. Estimating and forecasting spatial population dynamics of apex predators using transnational genetic monitoring. Proc Natl Acad Sci USA 117, 30531–30538. https://doi.org/10.1073/pnas.2011383117

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

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

Durbach, I., Chopara, R., Borchers, D.L., Phillip, R., Sharma, K., Stevenson, B.C., 2024. That's not the Mona Lisa! How to interpret spatial capture-recapture density surface estimates. Biometrics 80, ujad020. https://doi.org/10.1093/biomtc/ujad020

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.G., 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., Royle, J.A., Hare, M.P., 2016. Estimating population density and connectivity of American mink using spatial capture–recapture. Ecological Applications 26, 1125–1135. https://doi.org/10.1890/15-0315

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

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

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 56, 434–442. https://doi.org/10.1111/j.0006-341x.2000.00434.x.

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

Sanz-Pérez, A., Sazatornil, V., Palazón, S., Vanpé, C., Quenette, P.-Y., Sentilles, J., Kervellec, M., Milleret, C., López-Bao, J.V., Jordana, I.A., Jato, R., Solà de la Torre, J., Sollmann, R., 2025. Mapping sex- and age-structure reveals lonely males at the front in an expanding brown bear population. Biological Conservation 306. https://doi.org/10.1016/j.biocon.2025.111122

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

Sollmann, R., Furtado, M.M., Gardner, B., Hofer, H., Jácomo, A.T.A., Tôrres, N.M., Silveira, L., 2011. Improving density estimates for elusive carnivores: Accounting for sex-specific detection and movements using spatial capture–recapture models for jaguars in central Brazil. Biological Conservation 144, 1017–1024. https://doi.org/10.1016/j.biocon.2010.12.011

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

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

Appendix S1: Unravelling the effects of heterogeneity in space use on estimates of connectivity and population size: Insights from spatial capture-recapture modelling

```
source("SimulationsV2/functions/AccessCachedNim.R")
```

M. Kervellec, O. Gimenez, C. Vanpé, P.-Y. Quenette, J. Sentilles, S. Palazón, I. Afonso Jordana, R. Jato, M. M. Elósegui Irurtia, and C. Milleret

In the following appendix we provide the script to simulate spatial detection data for two groups and the nimble model. First we assume that groups are known as presented in the manuscript and then we provide an example to run the model when groups are unknown (mixture model)

1 Simulate data

1.1 Population parameters

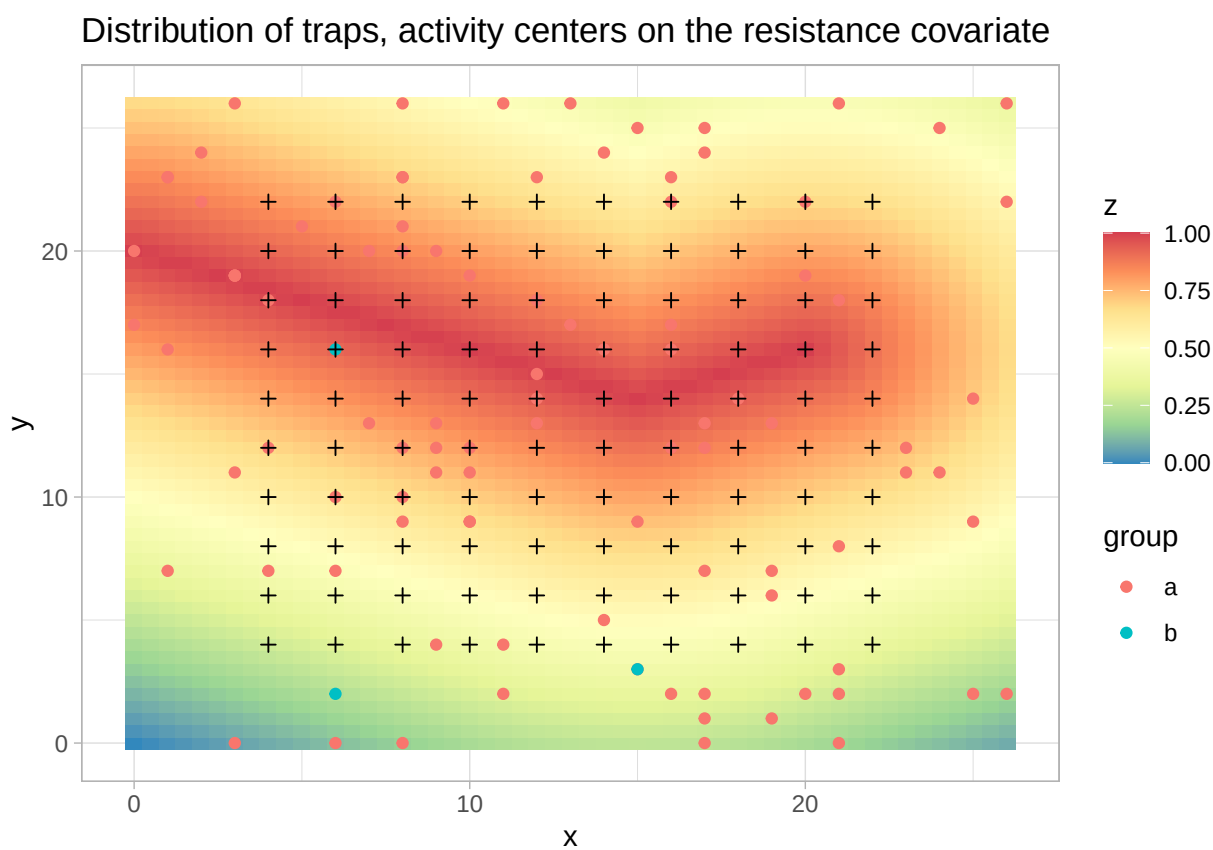
```
# Group a
N.a <- 97 # population size
p0.a <- 0.2 # baseline detection probability
sigma.a <- 7 # scale parameter
alpha1.a <- 1/(2*sigma.a*sigma.a)
alpha2.a <- 3 # resistance parameter

# Group b = outliers
N.b <- 3
DeltaP0 <- -0.1
p0.b <- p0.a + DeltaP0
DeltaSigma <- -2
sigma.b <- sigma.a + DeltaSigma
alpha1.b <- 1/(2*sigma.b*sigma.b)
DeltaAlpha <- -3
alpha2.b <- alpha2.a + DeltaAlpha

N <- N.a + N.b
```

1.2 Study Area and resistance covariate

```
# 1. Constant definition -------------------------------------------------------
K <- 7 # number of occasions
Array <- 10 # number of traps in one line: size of the trap array
```

```r
Resolution <- 0.5 # resolution of the resistance surface
trap_spacing <- 2 # 2*sigma.a

# 2. State space ------------------------------------------------------------
# Set a squared study area according to recommendation
buff <- 4*(trap_spacing/2) # size of the buffer -> 4*sigma
L <- 2*buff*(1/Resolution) + (Array-1)*trap_spacing*(1/Resolution) # total number of cell in a row

# 3. Trap locations ---------------------------------------------------------
traplocs <- expand.grid(X = seq(buff,
                                (L*Resolution)-buff,
                                trap_spacing),
                        Y = seq(buff,
                                (L*Resolution)-buff,
                                trap_spacing)) %>% # trap spacing = 2*sigma
  as.matrix()
ntraps <- nrow(traplocs)

# 4. Activity center location: Homogeneous point process ------------------------
spatialdomain <- expand.grid(X = seq(0, (L+1)*Resolution,1),
                             Y = seq(0, (L+1)*Resolution,1)) %>%
  as.matrix()
# sample across the locations in the spatial domain
S.a <-tibble(X = sample(spatialdomain[,1], size = N.a, replace = TRUE),
             Y =  sample(spatialdomain[,2], size = N.a, replace = TRUE)) %>%
  as.matrix()


S.b <-tibble(X = sample(spatialdomain[,1], size = N.b, replace = TRUE),
             Y =  sample(spatialdomain[,2], size = N.b, replace = TRUE)) %>%
  as.matrix()

S <- rbind(S.a,S.b) %>%
  as.data.frame() %>%
  mutate(group = c(rep("a",nrow(S.a)),rep("b",nrow(S.b))),
         ind = row_number())

# 5. Define resistant surface covariate ----------------------------------------
covariate <- expand_grid(seq(0,L*Resolution,Resolution),
                         seq(0,L*Resolution,Resolution)) %>%
  mutate(z  = runif((L+1)^2, -1, 1)) %>%
 rasterFromXYZ()

# set.seed(2025)
# covariate <- predict(gstat(formula = z ~ 1,
#                            locations = ~ x + y,
#                            dummy=T, beta=1,
#                            model=vgm(psill=1,model="Exp",range=20),
#                            nmax=20), # define the gstat object (spatial model)
#                        newdata = grid %>%
#                          raster::as.data.frame(xy = TRUE), # make one simulations based on the stat obj
#                        nsim = 1) %>%
#   mutate(sim1 =((sim1 - min(sim1))/(max(sim1)-min(sim1)))) %>% # scale between 0 & 1
```

```r
#   raster::rasterFromXYZ()

x_coords <- c(0,15,20)
y_coords <- c(20, 14, 16) # define the line points

points <- sp::SpatialPoints(cbind(x_coords, y_coords)) %>% # from spatial points
  sf::st_as_sf() %>%
  sf::st_combine()%>%
  sf::st_cast("LINESTRING") # convert to line

covariate[!is.na(covariate[])] <- sf::st_distance(SpatialPoints(coordinates(covariate)[!is.na(covariate
                                          points, byid = T) # from each resistance raster cell, we comput
covariate <- 1-((covariate - min(raster::values(covariate)))/(max(raster::values(covariate))-min(raster

# 6. look at what we have created so far ----------------------------------------
ss.plot <- ggplot() +
  geom_raster(covariate %>% raster::as.data.frame(xy=TRUE),
              mapping = aes(x = x, y = y, fill = z)) +
  scale_fill_distiller(palette = "Spectral", direction = -1) +
  geom_point(S, mapping = aes(x=X, y = Y, color = group))+
  geom_point(traplocs %>% as.data.frame(), mapping = aes(x=X, y = Y), pch = 3) +
  theme_light()

ss.plot +
  ggtitle("Distribution of traps, activity centers on the resistance covariate")
```



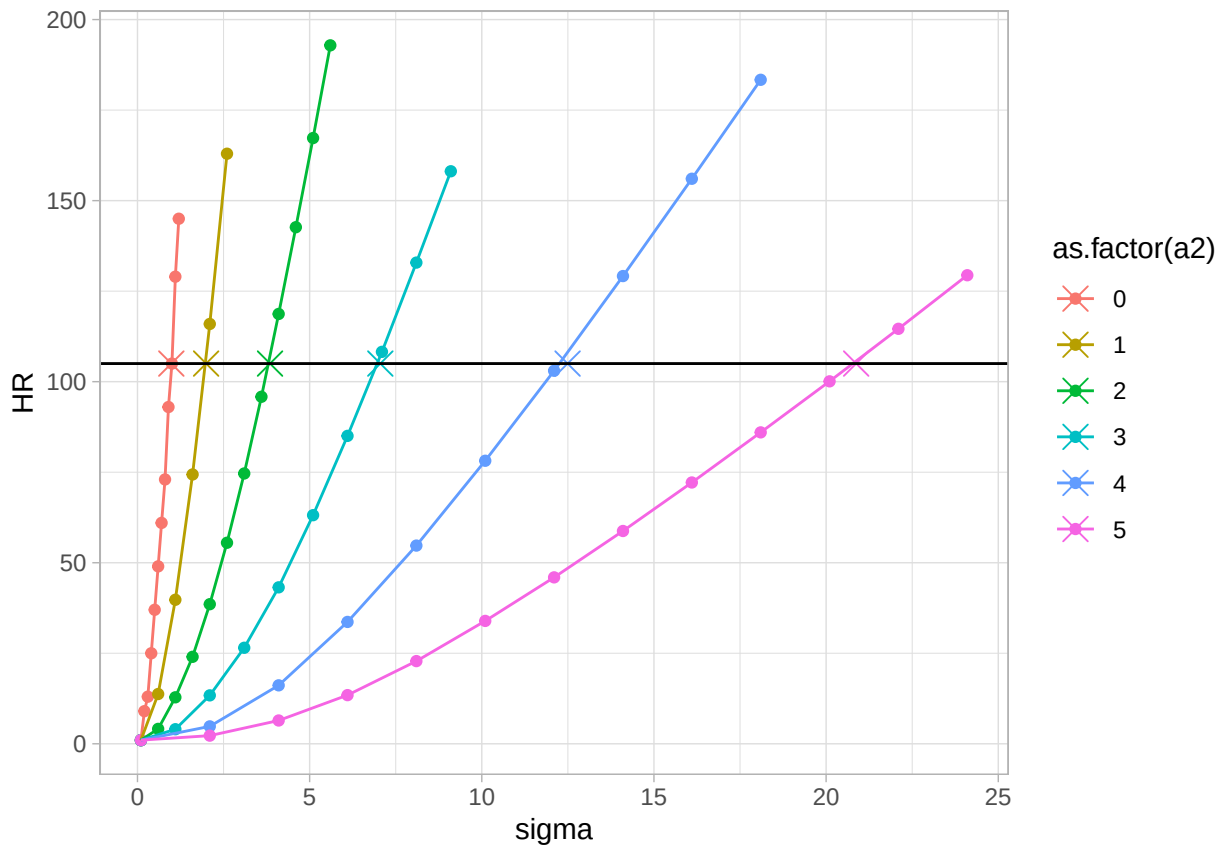
In this figure th underlying raster represents the length of roads per pixel. The black crosses represent traps and the points the activity centers simulated according to an homogeneous point process for group "a" in red and group "b" in blue.

1.3 Simulate Detections

Define resistance scenarios for landscape A Find a sigma for a given alpha2

```
source("SimulationsV2/functions/mean_homerange_size.R")
## extract coefficients  dont load this function it already exist in secr
coeff <- function(alpha2, data){
  data.a2 <- data %>%
    filter(a2 == alpha2) %>%
    mutate(sigma2 = sigma^2)
  lm <- lm(data.a2$HR ~ data.a2$sigma + data.a2$sigma2)
  return(c(coefficients(lm),alpha2))}

# solve polynomial, i.e. find the sigma given alpha and home range size
solve.poly <-function(alpha2, coef,HRsize){
  row <- which(coef[,4] == alpha2) # retrieve the coefficients according to a given alpha2
  Re(polyroot(c(coef[row,1]-HRsize, coef[row,2],coef[row,3])))[which(Re(polyroot(c(coef[row,1]-HRsize, c
}

# Define set of alpha2 and sigma by hand
SC <- data.frame(a2 = c(rep(0,12),
                        rep(1,6),
                        rep(2,12),
                        rep(3,10),
                        rep(4,10),
                        rep(5,13)),
                 sig = c(seq(0.1, 1.2, by=0.1),
                         seq(0.1, 3, by=0.5),
                         seq(0.1,6, by=0.5),
                         seq(0.1, 10, by=1),
                         seq(0.1, 20, by=2),
                         seq(0.1, 25, by=2)))

data <- data.frame(a2 = NA,
                   HR  = NA,
                   sigma = NA) # initialize output matrix

# Compute mean HR
for (i in 1:nrow(SC)){
  #print(i)
  out <- mean_homerange_size(sigma = SC[i,2],
                             a2 = SC[i,1],
                             covariate = covariate,
                             x.box = c(4,22), # where the HR are computed
                             y.box = c(4,22))
  data[i,] <- out
}

# Fit an exponential model to each of the curve
```

```r
Coefficients.poly <- coeff(0, data = data) %>%
  rbind(coeff(1, data = data)) %>%
  rbind(coeff(2, data = data)) %>%
  rbind(coeff(3, data = data)) %>%
  rbind(coeff(4, data = data)) %>%
  rbind(coeff(5, data = data))

# For each alpha for a set Home Range size find the corresponding sigma (choose HR such as sigma = 1 fo
HomeRangeSize <- 105
Coefficients.model <- data.frame(sigma = c(solve.poly(0, coef = Coefficients.poly, HRsize = HomeRangeSi
                                 solve.poly(1, coef = Coefficients.poly, HRsize = HomeRangeSize),
                                 solve.poly(2, coef = Coefficients.poly, HRsize = HomeRangeSize),
                                 solve.poly(3, coef = Coefficients.poly, HRsize = HomeRangeSize),
                                 solve.poly(4, coef = Coefficients.poly, HRsize = HomeRangeSize),
                                 solve.poly(5, coef = Coefficients.poly, HRsize = HomeRangeSize)),
                         a2 = c(0:5))

# Have a look to the results
row <- 5
sig.sets <- seq(0,1.2,0.5)
Check <- data.frame(sig = sig.sets,
                    HR = Coefficients.poly[row,1] + Coefficients.poly[row,2]*sig.sets + Coefficients.pol
HRsizeplot <- ggplot() +
  geom_point(data, mapping = aes(x = sigma, y = HR, col = as.factor(a2))) +
  geom_line(data, mapping = aes(x = sigma, y = HR, col = as.factor(a2))) +
  geom_point(Coefficients.model, mapping = aes(x = sigma, y  = HomeRangeSize, col = as.factor(a2)), pch
  geom_hline(yintercept = HomeRangeSize) +
  theme_light()
  #geom_line(Check, mapping = aes(x = sig, y = HR))
HRsizeplot
```




```
#ggplot2::ggsave(plot = HRsizeplot, dpi = 600,  width = 4.5, height = 4, filename = "plots/LandscapeAHR

# We define accordingly a sigma for each alpha2
Coefficients.model
```

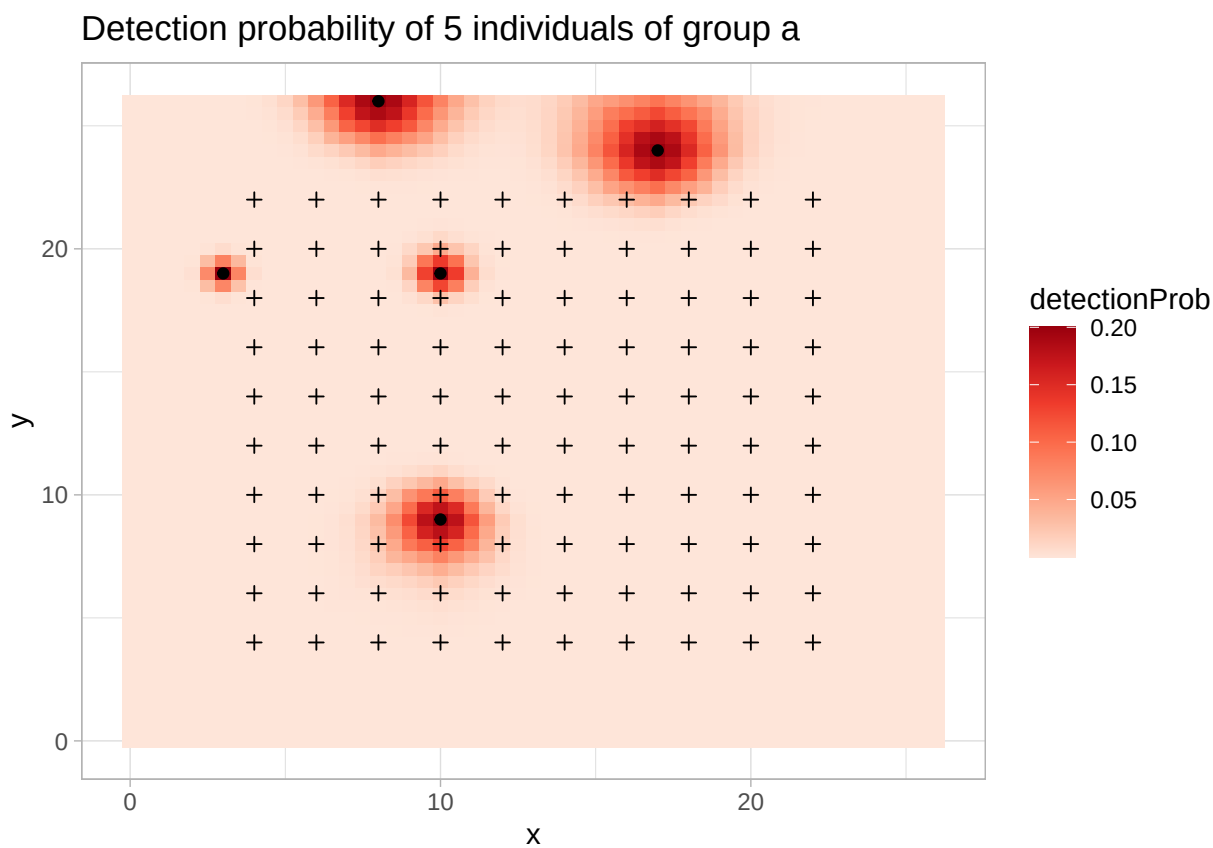
```
##        sigma a2
## 1  0.9829354  0
## 2  1.9931158  1
## 3  3.8496701  2
## 4  7.0527836  3
## 5 12.4913487  4
## 6 20.8730084  5
```

```
alpha2 <- c(0, 1, 2, 3, 4, 5)
sigma <- c(1, 2, 4, 7, 12, 21)
```

```
# 1. Compute detection probabilities for each individual -----------------------
## group a
cost.a <- exp(alpha2.a * covariate) # compute cost of each cell

D.a <- gdistance::transition(cost.a,
                             transitionFunction=function(x) 1/mean(x),
                             directions=8) %>%
  gdistance::geoCorrection(type="c",multpl=F,scl=F) %>%
  gdistance::costDistance(S.a, traplocs) # compute LCP distance between AC and traps, same as e2dist(S,
```

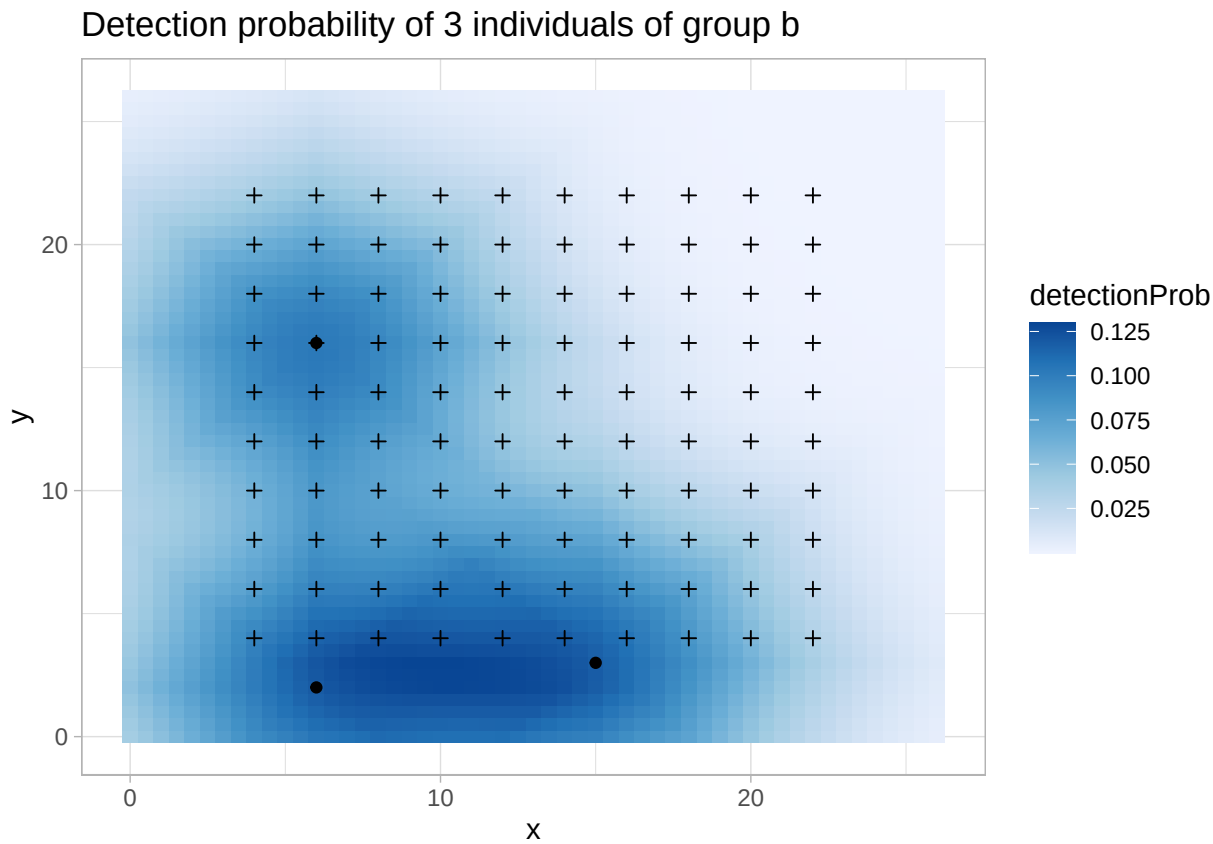
```r
probcap.a <- p0.a * exp(-alpha1.a * D.a * D.a) # compute detection probabilities

## group b
cost.b <- exp(alpha2.b * covariate)

D.b <- gdistance::transition(cost.b,
                             transitionFunction=function(x) 1/mean(x),
                             directions=8) %>%
  gdistance::geoCorrection(type="c",multpl=F,scl=F) %>%
  gdistance::costDistance(S.b, traplocs)

probcap.b <- p0.b * exp(-alpha1.b * D.b * D.b)

probcap <- rbind(probcap.a, probcap.b)

# 2. Simulate detection according to defined detection probabilities -----------
y <- array(0, dim=c(N, ntraps, K)) # empty 3D array (individuals by traps by sampling occasion)

# Simulate encounter histories
for(i in 1:N){                            # loop through each individual/activity center
  for(j in 1:ntraps){                     # loop through each trap
    y[i,j,1:K]<- rbinom(K, 1, probcap[i,j])  # y ~ binomial(p_ijk)
  }
}

Y <- apply(y, c(1,2), sum) # number of time an individual was captured at a trap in a session
#Y
caps.per.ind <- apply(y,1,sum) # shows # captures for each individual across all traps and all occasion
#caps.per.ind
N.det <- length(caps.per.ind[which(caps.per.ind>0)])
#N.det

# Keep detection history only for individuals detected
M <- N + 50 # augmented population size

eta_det <- c(rep(0, N.a),rep(1, N.b))[which(caps.per.ind>0)]
eta_data <- c(eta_det, rep(NA, M-N.det))

# Augmenting y
yy <- array(rep(0,M*ntraps*K), c(M,ntraps,K)) # augmented detection array
yy[1:N.det,,] <- y[which(caps.per.ind>0),,]

# 3. look at encounter probabilities of a few individual -----------------------
someGuys.a <- S.a[sample(1:N.a,5),]
D2.a <- gdistance::transition(cost.a,
                              transitionFunction=function(x) 1/mean(x),
                              directions=8) %>%
  gdistance::geoCorrection(type="c",multpl=F,scl=F) %>%
  gdistance::costDistance(someGuys.a, raster::coordinates(covariate))
probcap2.a <- p0.a * exp(-alpha1.a * D2.a * D2.a)

someGuys.b <- S.b[sample(1:N.b,3),]
D2.b <- gdistance::transition(cost.b,
```

```
                                transitionFunction=function(x) 1/mean(x),
                                directions=8) %>%
  gdistance::geoCorrection(type="c",multpl=F,scl=F) %>%
  gdistance::costDistance(someGuys.b, raster::coordinates(covariate))
probcap2.b <- p0.b * exp(-alpha1.b * D2.b * D2.b)

ggplot()+
  # group a
  geom_raster(data.frame(raster::coordinates(covariate),
                         detectionProb = apply(probcap2.a,2,sum)),
              mapping = aes(x=x, y = y, fill = detectionProb)) +
  scale_fill_distiller(palette = "Reds", direction = 1) +
  geom_point(someGuys.a %>% as.data.frame() , mapping = aes(x=X, y = Y)) +
  # traps
  geom_point(traplocs %>% as.data.frame(), mapping = aes(x=X, y = Y), pch = 3) +
  theme_light() +
  ggtitle("Detection probability of 5 individuals of group a ")
```




```
ggplot()+
  # group b
  geom_raster(data.frame(raster::coordinates(covariate),
                         detectionProb = apply(probcap2.b,2,sum)),
              mapping = aes(x = x, y = y, fill = detectionProb)) +
  scale_fill_distiller(palette = "Blues", direction = 1) +
  geom_point(someGuys.b %>% as.data.frame() , mapping = aes(x=X, y = Y)) +
```

```
  # traps
  geom_point(traplocs %>% as.data.frame(), mapping = aes(x=X, y = Y), pch = 3) +
  theme_light() +
  ggtitle("Detection probability of 3 individuals of group b")
```



These two maps represents the cumulative detection probability across the state space for the five randomly chosen activity center respectively for group "a" and group "b"

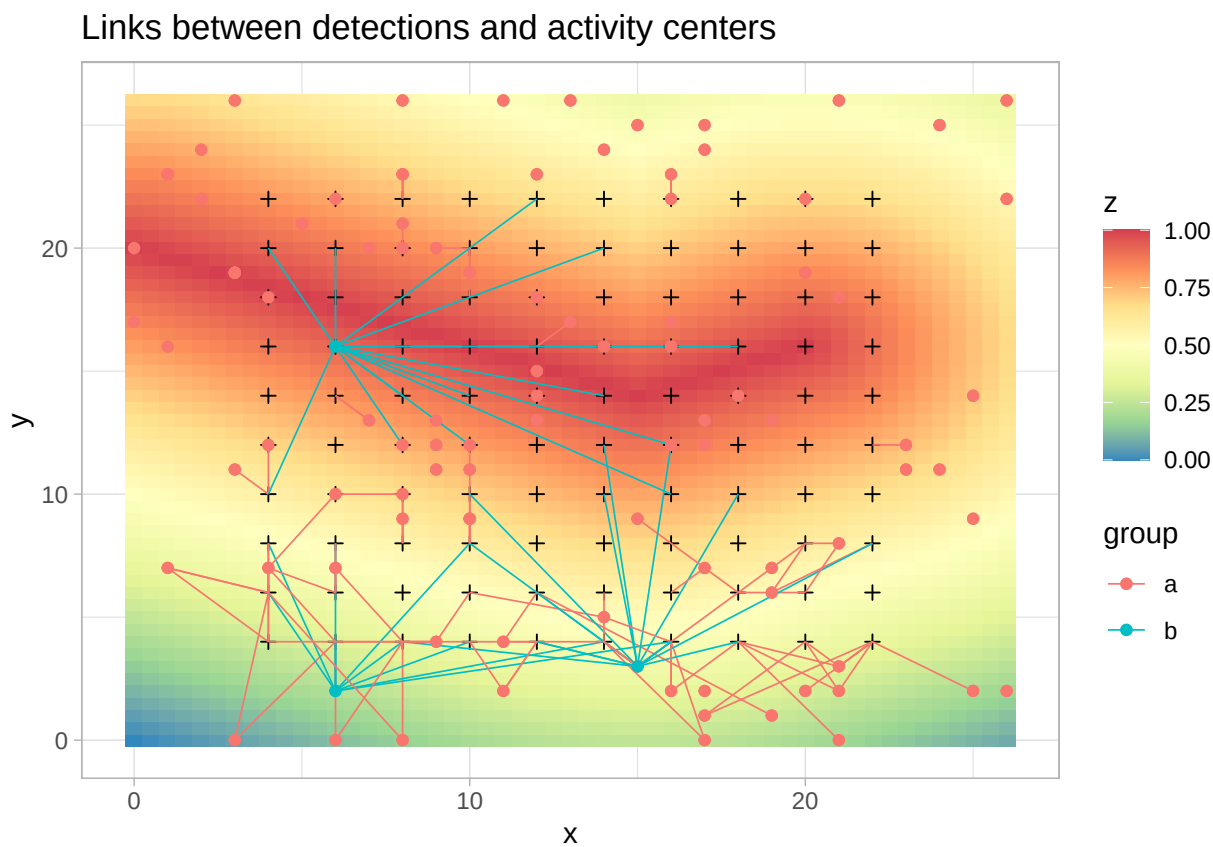
```
# 4. look at detection histories ----------------------------------------------
tdf <- traplocs %>%
  as.data.frame() %>%
  mutate(trap = paste0("V",c(1:ntraps)) %>%
           as.factor() %>%
           as.numeric())

for (k in 1:K){
  Det <- y[,,k] %>% # matrix of detections
    as_tibble(rownames = "ind", colnames = "trap") %>%
    pivot_longer(cols = -ind, names_to = "trap", values_to = "occ") %>% # transform to a longer datafra
    mutate(trap = as.numeric(as.factor(trap))) %>%
    filter(occ > 0) %>% # We only keep detections
    mutate(occ = occ*k) %>% # identify the occasion of capture
    left_join(tdf) %>% # retrieve trap location
    left_join(S %>% # retrieve AC location and group of each ind
                as.data.frame() %>%
                mutate(ind = as.character(1:nrow(S))) %>%
```

```r
                rename(sx=X,
                       sy=Y),
              by = "ind")
  # save for each occasion
  if(k==1){
    Dets <- Det}
  else{
    Dets <- rbind(Dets,Det)}
}
```

```
## Warning: The `x` argument of `as_tibble.matrix()` must have unique column names if
## `.name_repair` is omitted as of tibble 2.0.0.
## i Using compatibility `.name_repair`.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```

```
## Joining with `by = join_by(trap)`
## Joining with `by = join_by(trap)`
## Joining with `by = join_by(trap)`
## Joining with `by = join_by(trap)`
## Joining with `by = join_by(trap)`
## Joining with `by = join_by(trap)`
## Joining with `by = join_by(trap)`
```

```r
# Plot
ss.plot +
  geom_segment(data=Dets,
               aes(x=sx,y=sy,xend=X,yend=Y, color = group),lwd=0.3) +
  geom_point(S %>% as.data.frame() %>%  mutate(ind = as.character(1:nrow(S))) ,
             mapping = aes(x=X, y = Y, color = group)) +
  ggtitle("Links between detections and activity centers")
```



Then, this figure shows where each individual was detected by drawing lines between simulated activity centers and the traps where the individual was detected.

2. Fit model

2.1 Model with 2 known groups

```
# Least cost path function
leastcostpath <- function(alpha){
  # alpha is a scalar
  # rcov is a RasterLayer
  # xy is a matrix of dims nsites x 2
  # obtain resistance surface
  cost <- exp(alpha * covariate)

  D <- gdistance::transition(cost, transitionFunction=function(x) 1/mean(x), directions=8) %>%
    gdistance::geoCorrection(type="c",multpl=F,scl=F) %>%
    gdistance::costDistance(spatialdomain, traplocs)

  D^2
}

# Convert the function into NIMBLE
```

```r
LCP <- nimbleRcall(function(alpha = double(0)){},
                   Rfun = 'leastcostpath',
                   returnType = double(2))

A <- leastcostpath(-5) # test
```

Compute cached matrix

```r
# make correspondance between cell corrdinates and rownumber in the distance array
habitatgrid <- matrix(1:nrow(spatialdomain), nrow = max(spatialdomain[,1])+1) # identify AC location

cells = nrow(spatialdomain) # number of cells in the spatial domain
x.max = y.max = L*Resolution # extend of the spatial domain

# Alpha array
step_alpha <- 0.1
alphaValues <- round(seq(-5, 5, by=step_alpha), digits=1)

cached <- array(NA, c(cells, ntraps, length(alphaValues)))

for(a in 1:length(alphaValues)){
  cached[1:cells,1:ntraps,a] <- LCP(alphaValues[a]) # compute distances between habitat cells and traps
}

roundValue <- 1/step_alpha
indexValue <- which(alphaValues<=0) %>% length()
```

Model Code

```r
modelCode2groups <- nimbleCode({
  ## priors
  psi ~ dunif(0, 1)
  probB ~ dunif(0, 1)

  sigma[1] ~ dunif(0, 50)
  DeltaSigma ~ dunif(-50, 50)
  sigma[2] <- sigma[1] +  DeltaSigma
  three ~ dconstraint(sigma[2] > 0) # positive scale parameter

  p0[1] ~ dunif(0, 1)
  DeltaP0 ~ dunif(-1, 0)
  p0[2] <- p0[1] +  DeltaP0
  two ~ dconstraint(p0[2] > 0 & p0[2] < 1)

  alpha2[1] ~ dunif(-1, 5)
  DeltaAlpha ~ dunif(-5, 0)
  alpha2[2] <- alpha2[1] + DeltaAlpha
  one ~ dconstraint(alpha2[2] > -5 & alpha2[2] < 5) # be sure that alpha2[2] is not outside cached array

  AlphaID[1] <- trunc(alpha2[1]*roundValue+indexValue) # formula (round(alpha2,1)*10+51)  to find were
  AlphaID[2] <- trunc(alpha2[2]*roundValue+indexValue)

  distSq[1:cells, 1:ntraps,1] <- AccessCachedNim(AlphaID[1]) #cached[1:cells, 1:ntraps,AlphaID[1]]
```

```
  distSq[1:cells, 1:ntraps,2] <- AccessCachedNim(AlphaID[2]) #cached[1:cells, 1:ntraps,AlphaID[2]]

  alpha1[1] <- 1/(2 * sigma[1] * sigma[1])
  alpha1[2] <- 1/(2 * sigma[2] * sigma[2])

  ## loop over individuals
  for(i in 1:M) {
    ## AC coordinates
    sxy[i,1] ~ dunif(0, x.max)
    sxy[i,2] ~ dunif(0, y.max)

    sID[i] <- habitatgrid[trunc(sxy[i,1]) + 1, trunc(sxy[i,2]) + 1] # in which cell fall the AC
    eta[i] ~  dbern(probB)
    probcap[i,1:ntraps] <- p0[eta[i]+1] * exp(-alpha1[eta[i]+1] * distSq[sID[i],1:ntraps,eta[i]+1]) # d

    z[i] ~ dbern(psi) # latent dead/alive indicators

    for(j in 1:ntraps){
      muy[i,j] <-  probcap[i,j] * z[i]
      for(k in 1:K){
        y[i,j,k] ~  dbern(muy[i,j])}
    }
  }

  ## derived quantity: total population size
  N <- sum(z[1:M])
})
```

Compute the centroid of detection for each individual detected so his activity center is near it's detection

```
Sxy_inits <- tibble(X = rep(NA,M), Y = rep(NA,M))

# Centroid for detected individuals
for(i in 1:N.det){
  Tl <- traplocs[which(apply(yy[i,,],1, sum) >0),] # trap location were an individual was detected
  if (length(Tl)>2){
    Sxy_inits[i,"X"] <- apply(Tl,2,mean)["X"]
    Sxy_inits[i, "Y"] <- apply(Tl,2,mean)["Y"]} # centroid
  else {
    Sxy_inits[i,"X"] <- Tl["X"]
    Sxy_inits[i, "Y"] <- Tl["Y"]
  }
}

# random values for AC of undetected individuals

for(i in N.det:M){
    Sxy_inits[i,"X"] <- runif(1,0,x.max)
    Sxy_inits[i, "Y"] <- runif(1,0,y.max)
}
```

Define initial values

```r
# True values
# nimInits <- list(p0 = c(p0.a,p0.b),
#                  psi = N/M,
#                  sigma = c(sigma.a,sigma.b),
#                  alpha2 = c(alpha2.a,alpha2.b),
#                  probB  = c(N.b/(N.a+N.b)),
#                  DeltaAlpha = DeltaAlpha,
#                  DeltaSigma = DeltaSigma,
#                  DeltaP0 = DeltaP0,
#                  eta = c(rep(NA,N.det), rbinom(M-N.det, 1, prob=0.5)),
#                  sxy = Sxy_inits %>% as.matrix(),
#                  z = c(rep(1,N.det),rbinom(M-N.det, 1, prob=0.3)))

# Random values
p0Inits <- runif(1,0.1,0.4)
alphaInits <- runif(1, 1, 5)
DeltaAlphaInits <- runif(1, -1, -0.1)
sigmaInits <- runif(1, 1, 3)
DeltaSigmaInits <- runif(1, 0.5, 2)
DeltaP0Inits <- runif(1,-0.1,0)
psiInits <- runif(1,0.7,0.9)

nimInits <- list(p0 = c(p0Inits,p0Inits+DeltaP0),
                 psi = psiInits,
                 sigma = c(sigmaInits,sigmaInits+DeltaSigmaInits),
                 alpha2 = c(0,0),
                 probB  = runif(1,0.05,0.2),
                 DeltaAlpha = 0,
                 DeltaSigma = DeltaSigmaInits,
                 DeltaP0 = DeltaP0,
                 eta = c(rep(NA,N.det), rbinom(M-N.det, 1, prob=0.5)),
                 sxy = Sxy_inits %>% as.matrix(),
                 z = c(rep(1,N.det),rbinom(M-N.det, 1, prob=0.3)))
```

Define constants and data to run the model

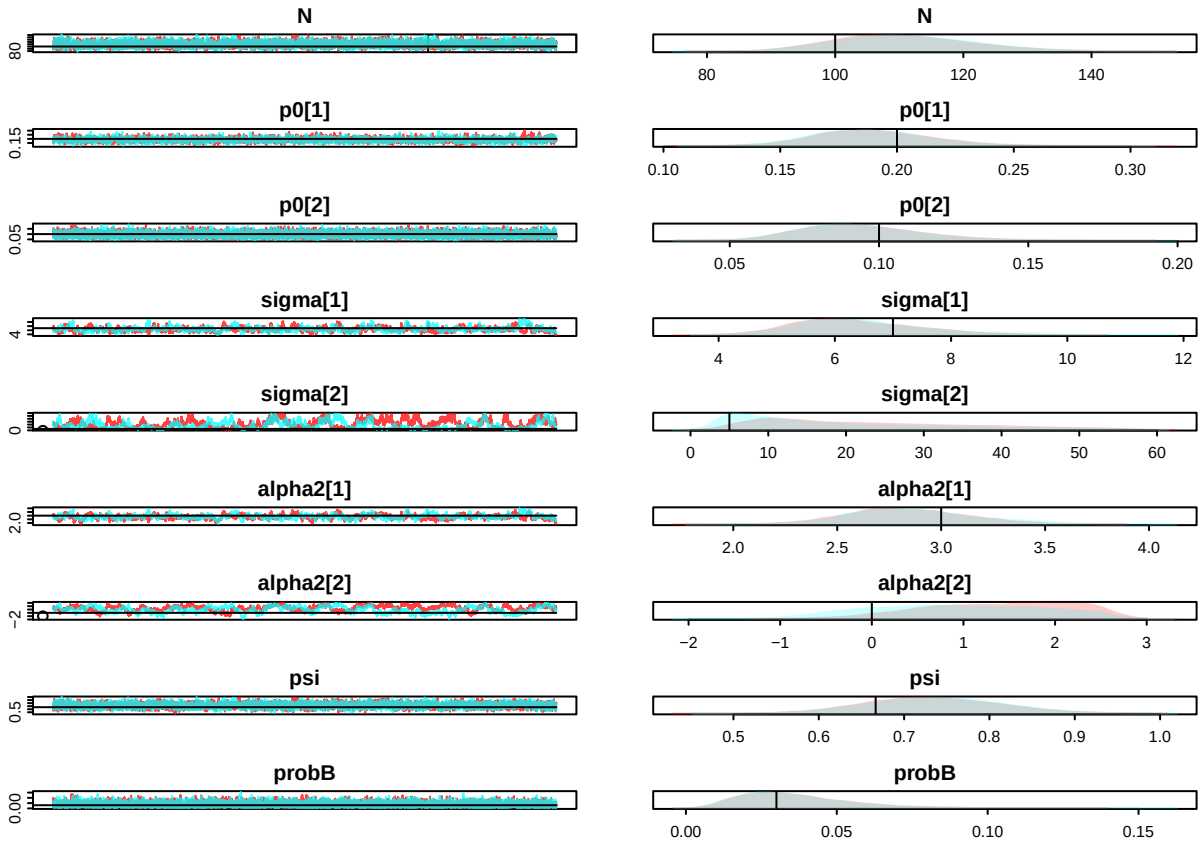
```r
nimConstants <- list(M = M,
                     ntraps = Array*Array,
                     K = K,
                     x.max = x.max,
                     y.max = y.max,
                     cells = nrow(spatialdomain))

nimData <- list(y = yy,
                habitatgrid = habitatgrid,
                eta = eta_data,
                indexValue = indexValue,
                roundValue = roundValue,
                #cached = cached,
                one = 1,
                two = 1,
                three = 1)
```

```r
model <- nimbleModel(code = modelCode2groups,
                     constants = nimConstants,
                     data = nimData,
                     inits = nimInits,
                     check = F,
                     calculate = F,
                     dimensions = list(habitatgrid = dim(habitatgrid),
                                       probcap = c(M, ntraps),
                                       sID = M,
                                       eta = M,
                                       AlphaID = 2))
```

```
## Defining model

## Building model

## Setting data and initial values

## Checking model sizes and dimensions

##   [Note] This model is not fully initialized. This is not an error.
##          To see which variables are not initialized, use model$initializeInfo().
##          For more information on model initialization, see help(modelInitialization).
```

```r
# model$calculate()
# cmodel <- compileNimble(model)
# cmodel$calculate()
# MCMCconf <- configureMCMC(model = model,
#                           monitors = c("N", "sigma", "p0","psi","alpha2","probB"),
#                           control = list(reflective = TRUE),
#                           thin = 1)
#
# MCMC <- buildMCMC(MCMCconf)
# cMCMC <- compileNimble(MCMC, project = model, resetFunctions = TRUE)
# # RUN THE MCMC
# MCMCRuntime <- system.time(myNimbleOutput <- runMCMC(mcmc = cMCMC,
#                                                      nburnin = 1000,
#                                                      niter = 50000,
#                                                      nchains = 2,
#                                                      samplesAsCodaMCMC = TRUE))
#
# save(myNimbleOutput, file = "myNimbleOutput_appendixS1.RData")
load("myNimbleOutput_appendixS1.RData")
```

```r
basicMCMCplots::chainsPlot(myNimbleOutput,var=c("N","p0","sigma","alpha2","psi","probB"), line = c(N,p0
```




```r
summary(myNimbleOutput)
```

```
##
## Iterations = 1:49000
## Thinning interval = 1
## Number of chains = 2
## Sample size per chain = 49000
##
## 1. Empirical mean and standard deviation for each variable,
##    plus standard error of the mean:
##
##                Mean       SD  Naive SE Time-series SE
## N         110.84580 10.51847 3.360e-02      0.2035925
## alpha2[1]   2.85215  0.29718 9.493e-04      0.0189300
## alpha2[2]   1.15945  0.89956 2.874e-03      0.0907727
## p0[1]       0.18902  0.02425 7.745e-05      0.0005747
## p0[2]       0.09353  0.01984 6.339e-05      0.0002192
## probB       0.03567  0.01765 5.639e-05      0.0001381
## psi         0.73578  0.07765 2.480e-04      0.0014527
## sigma[1]    6.45123  1.13536 3.627e-03      0.0693649
## sigma[2]   19.67454 12.93414 4.132e-02      1.1735708
##
## 2. Quantiles for each variable:
##
##                2.5%       25%       50%       75%     97.5%
```

```
## N         92.000000 103.00000 110.00000 118.0000 133.00000
## alpha2[1]  2.314221   2.64611   2.83128   3.0400   3.48523
## alpha2[2] -0.612783   0.50321   1.20869   1.8880   2.60534
## p0[1]      0.145310   0.17193   0.18762   0.2045   0.23980
## p0[2]      0.059751   0.07945   0.09167   0.1058   0.13745
## probB      0.009919   0.02276   0.03276   0.0454   0.07781
## psi        0.588951   0.68161   0.73380   0.7884   0.89239
## sigma[1]   4.595920   5.63898   6.30144   7.1189   9.10922
## sigma[2]   3.767919   9.21338  16.08018  28.0876  49.84539
```

```
MCMCvis::MCMCsummary(myNimbleOutput)
```

```
##                    mean          sd         2.5%          50%        97.5% Rhat
## N         110.84579592 10.51846981 92.000000000 110.00000000 133.00000000 1.00
## alpha2[1]   2.85214529  0.29717879  2.314221186   2.83128309   3.48522553 1.02
## alpha2[2]   1.15944955  0.89956153 -0.612782924   1.20868799   2.60533518 1.34
## p0[1]       0.18901802  0.02424580  0.145309536   0.18762341   0.23980142 1.01
## p0[2]       0.09352839  0.01984293  0.059751294   0.09167085   0.13744900 1.00
## probB       0.03566711  0.01765322  0.009919494   0.03276369   0.07780762 1.00
## psi         0.73578412  0.07764746  0.588950511   0.73379588   0.89239072 1.00
## sigma[1]    6.45122737  1.13536372  4.595919887   6.30144010   9.10922163 1.04
## sigma[2]   19.67453885 12.93413811  3.767919205  16.08017869  49.84539012 1.24
##           n.eff
## N          2666
## alpha2[1]   252
## alpha2[2]    95
## p0[1]      1776
## p0[2]      8197
## probB     16378
## psi        2858
## sigma[1]    274
## sigma[2]    116
```

2.2 Mixture Model

```
# # define parameters
# nimConstants <- list(M = M,
#                      ntraps = ntraps,
#                      K = K,
#                      x.max = x.max,
#                      y.max = y.max,
#                      cells = nrow(spatialdomain))
#
# nimData <- list(y = yy,
#                 habitatgrid = habitatgrid,
#                 indexValue = indexValue,
#                 roundValue = roundValue,
#                 cached = cached,
#                 one = 1,
#                 two = 1)
#
```

```
# probBInits <- runif(1,0.05,0.2)
# nimInits <- list(p0 = c(p0Inits,p0Inits+DeltaP0),
#                  psi = psiInits,
#                  sigma = c(sigmaInits,sigmaInits+DeltaSigmaInits),
#                  alpha2 = c(0,0),
#                  probB  = probBInits,
#                  DeltaAlpha = 0,
#                  DeltaSigma = DeltaSigmaInits,
#                  DeltaP0 = DeltaP0,
#                  eta = rbinom(M, 1, prob=probBInits),
#                  sxy = Sxy_inits %>% as.matrix(),
#                  z = c(rep(1,N.det),rbinom(M-N.det, 1, prob=0.3)))
#
# model <- nimbleModel(code = modelCode2groups,
#                      constants = nimConstants,
#                      data = nimData,
#                      inits = nimInits,
#                      check = F,
#                      calculate = T,
#                      dimensions = list(habitatgrid = dim(habitatgrid),
#                                        probcap = c(M, ntraps),
#                                        sID = M,
#                                        AlphaID = 2))
```

```
# calculate(model)
# cmodel <- compileNimble(model)
# MCMCconf <- configureMCMC(model = model,
#                           monitors = c("N", "sigma", "p0","psi","alpha2", "probB"),
#                           control = list(reflective = TRUE),
#                           thin = 1)
#
# MCMC <- buildMCMC(MCMCconf)
# cMCMC <- compileNimble(MCMC, project = model, resetFunctions = TRUE)
#
# # RUN THE MCMC
# myNimbleOutput <- runMCMC(mcmc = cMCMC,
#                           nburnin = 1000,
#                           niter = 10000,
#                           nchains = 2,
#                           samplesAsCodaMCMC = TRUE)
```

```
# basicMCMCplots::chainsPlot(myNimbleOutput,var=c("N","p0","sigma","alpha2","psi","probB"), line = c(N,
#
# summary(myNimbleOutput)
```

AppendixS2: Unravelling the effects of heterogeneity in space use on estimates of connectivity and population size: Insights from spatial capture-recapture modelling

M. Kervellec, O. Gimenez, C. Vanpé, P.-Y. Quenette, J. Sentilles, S. Palazón, I. Afonso Jordana, R. Jato, M. M. Elósegui Irurtia, and C. Milleret

Simulation Study

***Table S2.A**: Table with prior distribution for each model parameter*

Parameter	Distribution
sigma[1]	Uniform(0, 50)
DeltaSigma	Uniform(-50, 50) or Uniform(-50,150) for model MO3
sigma[2]	sigma[1]+DeltaSigma
alpha2[1]	Uniform(-1, 5)
DeltaAlpha	Uniform(-5, 0)
alpha2[2]	alpha2[1]+DeltaAlpha
p0[1]	Uniform(0,1)
DeltaP0	Uniform(-1,0)
p0[2]	P0[1]+DeltaP0
psi	Uniform(0,1)

AppendixS3: Unravelling the effects of heterogeneity in space use on estimates of connectivity and population size: Insights from spatial capture-recapture modelling

M. Kervellec, O. Gimenez, C. Vanpé, P.-Y. Quenette, J. Sentilles, S. Palazón, I. Afonso Jordana, R. Jato, M. M. Elósegui Irurtia, and C. Milleret

Case study:

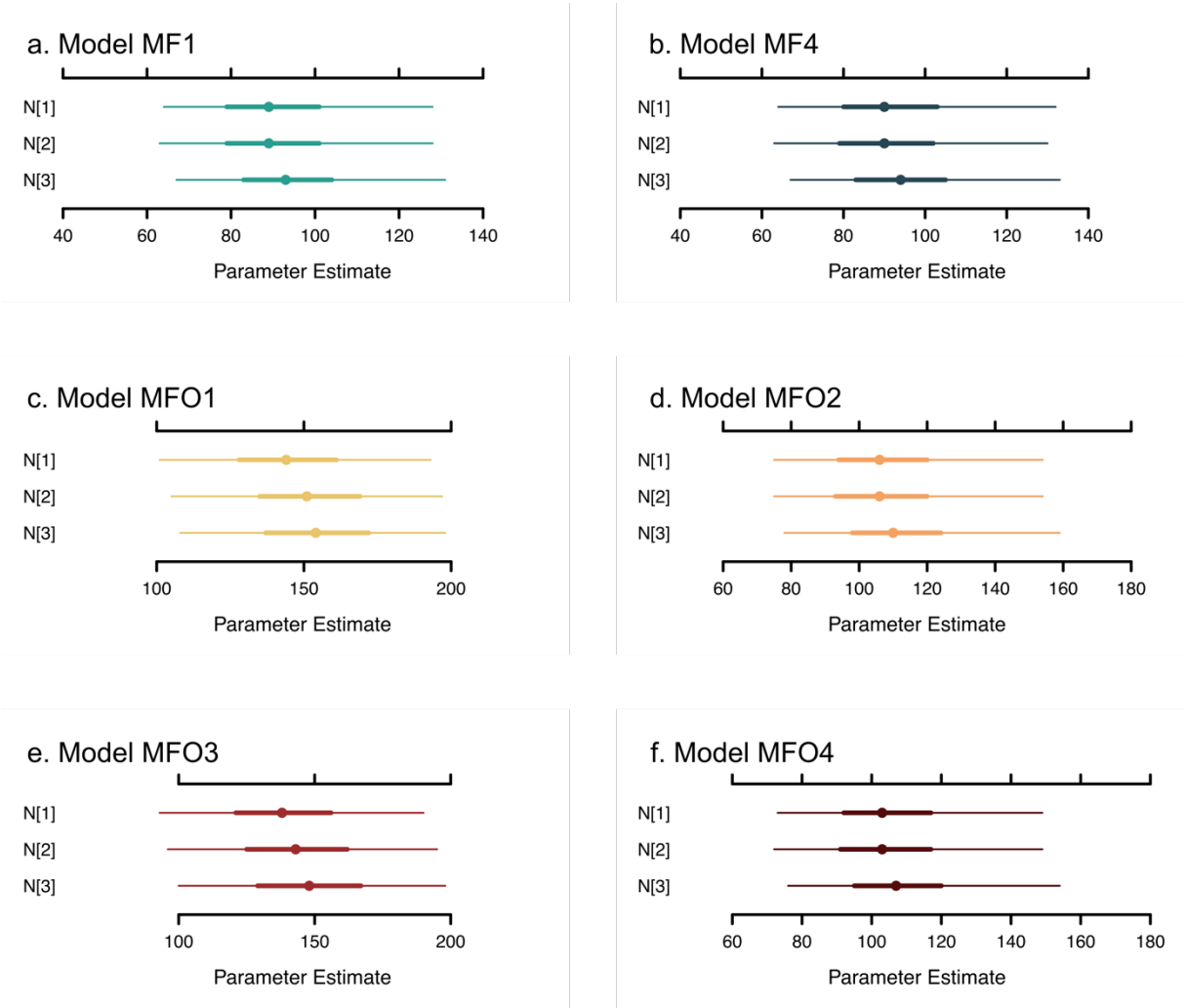


***Figure S3.A**: Estimated population size in 2017 (N[1]), 2018 (N[2]) and 2019 (N[3]) for the Pyrenean brown bear population. Caterpillar plots display the median of the posterior distribution and the 50% and 95% credible intervals for the 6 tested models. The first row shows the models with outliers removed i.e. MF1 (a) and MF4 (b) and the second and the third rows show the models with all the dataset i.e. models MFO1 (c), MFO2 (d), MFO3 (e) and MFO4 (f).*

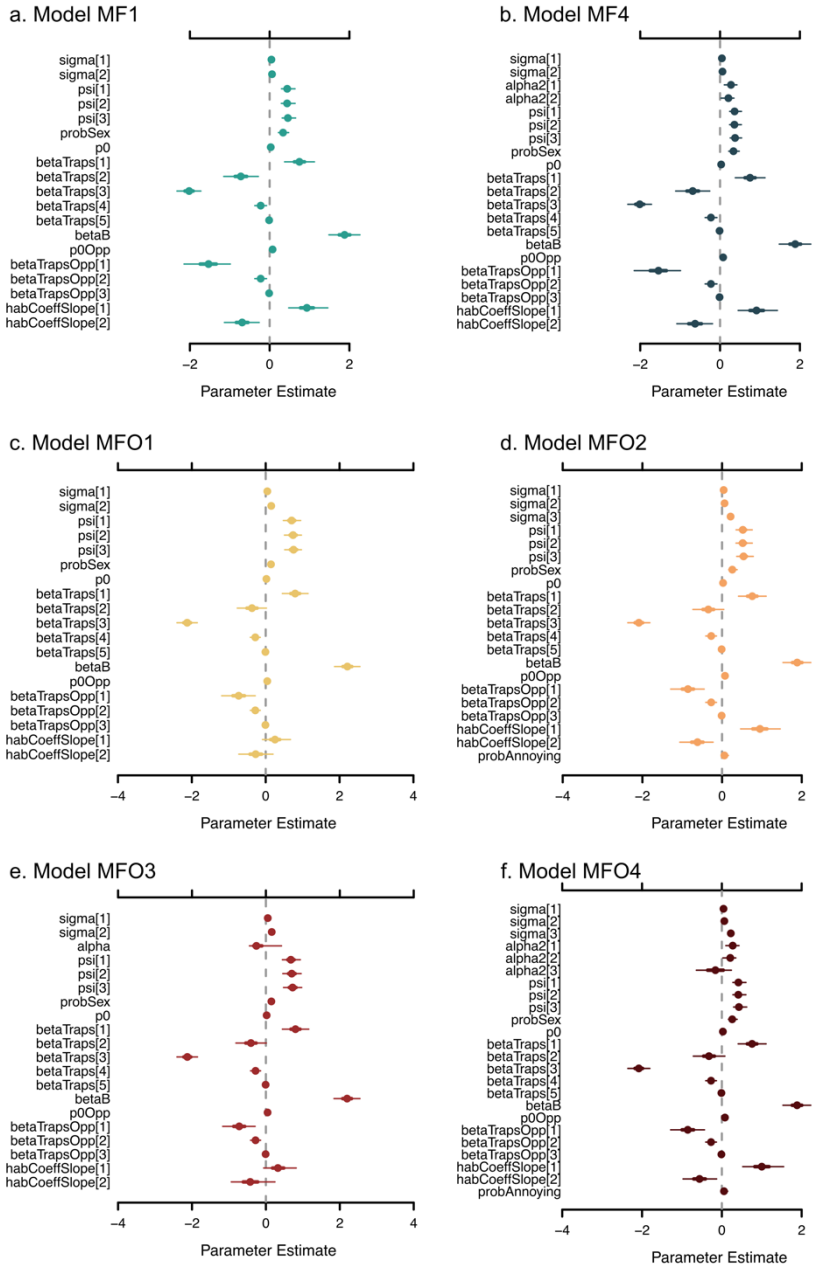


Figure S3.B*: Estimated parameters for the Pyrenean brown bear population. The parameters are the scale parameter for females (sigma[1]) and males (sigma[2]), the resistance parameters for females (alpha2[1]), males (alpha2[2]) and outliers (alpha2[3]), the inclusion parameters in 2017(psi[1]), 2018 (psi[2]) and 2019 (psi[3]), the probability to be a male (probSex), the intercept of the baseline detection probability for the structured monitoring (p0), the slope for the number of visits (betaTraps[1]), the slope for the country for the structured monitoring (betaTraps[2]), the slope for the type of trap (betaTraps[3]), the slope for effect of the month (betaTraps[4], betaTrapsOpp[2]) and the squared effect of the month (betaTraps[5], betaTrapsOpp[3]), the slope for the behavioral effect (betaB), the intercept of the baseline detection probability for the opportunistic monitoring (p0Opp), the slope for the country for the opportunistic monitoring (betaTrapsOpp[1]), the habitat slope for the ruggedness (habCoeffSlope[1]) and the habitat slope for the human density (habCoeffSlope[2]).*
Caterpillar plots display the median of the posterior distribution and the 50% and 95% credible intervals for the 6 tested models. The first row shows the models with outliers removed i.e. MF1 (a) and MF2 (b) and the second and the third rows show the models with all the dataset i.e. models MFO1 (c), MFO2 (d), MFO3 (e) and MFO4 (f).

***Table S3.A**: Table with prior distribution for each model parameter*

Parameter	Distribution
sigma	Uniform(0, 50)
alpha2[1]	Uniform(-1.5, 1.5)
DeltaAlpha	Uniform(-1.5, 0)
alpha2[2]	alpha2[1]+DeltaAlpha
DeltaAlpha2	Uniform(-1.5, 0)
alpha2[3]	alpha2[2]+DeltaAlpha2
p0	Uniform(0,1)
betaTraps	Uniform(-5,5)
betaB	Uniform(-5,5)
p0Opp	Uniform(0,1)
betaTrapsOpp	Uniform(-5,5)
habCoeffSlope	Normal(0,10)
probSex	Uniform(0,1)
probOutlier	Uniform(0,1)