A novel method to estimate actual infrastructure-induced mortality by integrating sampling biases

Guillermo Gómez-Peña[1*†], Marcello D'Amico[1†], Carlos Rodríguez[1], Jacinto Román[1], Alberto García-Rodríguez[1], Eloy Revilla[1], Maria Paniw[1]

[1] Estación Biológica de Doñana CSIC – Calle Américo Vespucio 26, 41092 Seville, Spain

[*] Corresponding author

[†] Shared first authorship

* A running headline of not more than 45 characters*

Bayesian latent-state models for roadkill estimates

* Acknowledgments, data availability, conflict of interest and author contribution statements*

This study has been funded by the Spanish Ministry for the Ecological Transition and the Demographic Challenge MITECO, through the SAFE Project – Stop Atropellos de Fauna en España to the Estación Biológica de Doñana EBD-CSIC. GGP has been supported by the JAE Intro 2022 (JAEINT_22_02345) and JAE Pre 2023 (SOLAUT_00055805) scholarship programs from the Spanish National Research Council CSIC. MD has been supported by Juan de la Cierva - Incorporación (IJC2019-039662-I) from the Spanish Ministry of Science, Innovation and Universities MICIU, and the project “Plan Complementario de I+D+i en el área de Biodiversidad (PCBIO)”, funded by the European Union within the framework of the Recovery, Transformation and Resilience Plan - NextGenerationEU and by the Regional Government of Andalusia. CR was supported by the SAFE Project. AGR was supported by the Spanish Ministry of Science and Innovation (MICINN) through the SUMHAL Project (LIFEWATCH2019-09-CSIC-13. MP was supported by was supported by the Spanish Ministry of Economy and Competitiveness (MINECO) and by the European Social Fund through the Ramón y Cajal Program (RYC2021-033192-I). Miguel Clavero contributed to the conceptualization of this study. Juan Carlos Rivilla and dozens of volunteers contributed to the field samplings. Land Rover España kindly lent a vehicle.

Supplementary material and the scripts used for making the simulation study and case study have been deposited in Github https://github.com/Guillermo247/A-novel-method-to-estimate-actual-infrastructure-induced-mortality-by-integrating-sampling-biases

ABSTRACT

1. Human infrastructures are among the most impactful threads to wildlife. While estimates exist on the number of animals killed by these structures over a given period, such estimates typically do no account for several detection biases. Consequently, true mortality rates may be severely underestimated, as well as their impact on populations and species.

2. We present a hierarchical Bayesian latent-state modelling framework that sequentially accounts for three main biases probabilities in estimating mortality abundance: the probability that a hit animal dies on the surveyed area (carcass location probability), the probability that the carcass remains on the surveyed area until the survey is conducted (carcass persistence probability), and the probability that the carcass is observed during the survey process (carcass observation probability). We employ a comprehensive simulation study where we test the effects of variability in species characteristics, sampling design, latent-state parameters, and prior information on the ability of our model to estimate mortality abundance on roads as total number of roadkills. We then demonstrate the applicability of our framework on a case study to estimate the total number of roadkills per km in different Mediterranean ecosystems while evaluating the cross-efficiency of different sampling methods.

3. Our framework is able to accurately recover the total number of roadkills from simulated census data for most simulation scenarios. We detected the highest disagreement between modelling outcomes and simulated data when variability in simulated carcass persistence probability, as well as related prior information in the Bayesian model, were high. In the case study, our results showed notably high roadkill numbers (e.g., for passerines, we estimate a total of 48.92 roadkills per km rate based on 8.04 observed rate during the road survey), along with substantial variation across different vertebrate groups. Furthermore, our case study confirms that walking and cycling surveys are more effective than driving surveys in detecting carcasses.

4. Our modelling framework offers an efficient approach to estimate mortality rates for a wide range of taxa. To optimize its application, extensive fieldwork for bias estimation and integration in analysis is needed. The accuracy of our framework may help managers to assess the impact of infrastructure-related mortality and prioritize conservation efforts to mitigate it.

1. Introduction

Linear infrastructures such as roads, power lines and wind turbines have become extremely widespread and are expected to increase substantially in the next decades, particularly in developing countries that host rich biodiversity (D'Amico, Catry, et al., 2018; Meijer et al., 2018; Tabassum-Abbasi et al., 2014). This is worrying because linear infrastructures contribute to the decline and even extinction of wildlife populations, and ultimately to biodiversity loss (Barrientos et al., 2021; D'Amico et al., 2019; Pearce-Higgins et al., 2012). In the last decades, this ecological impact has been extensively studied, with the majority of research focusing on infrastructure-induced mortality (Barrientos et al., 2021; D'Amico, Ascensão, et al., 2018; Nazir et al., 2020). Most research has primarily aimed at investigating the spatiotemporal patterns of such mortality (D'Amico et al., 2015; Guil et al., 2015), although a growing body of studies has more recently sought to quantify the magnitude of this threat.

However, when estimating infrastructure-induced mortality, standard carcass counts may not accurately reflect the total number of individuals affected. This is because the recorded carcasses are the result of a series of sequential processes, including the affected animal remaining near the infrastructure after the mortality event, the carcass persisting until the survey, and finally the observer detecting it (Barrientos et al., 2018; Bech et al., 2012; Román et al., 2024). Not accounting for these three hierarchical processes may result in several nested levels of biases in carcass surveys (Barrientos et al., 2018; Román et al., 2024). The first of these process is carcass location bias and concerns animals injured by collisions with power lines, wind turbines, or vehicles on roads that die outside the survey area (Bernardino et al., 2018; Román et al., 2024; Smallwood, 2007). The second process affecting standard mortality surveys along infrastructures is carcass persistence bias, which occurs when carcasses disappear from the survey area over time (Barrientos et al., 2018; Borner et al., 2017; Ravache et al., 2024). This is typically due to natural decomposition and environmental factors influencing it (such as weather conditions; Barrientos et al., 2018; Borner et al., 2017), but also to scavenger activity (DeVault et al., 2017; Dhiab et al., 2023). Regarding roads, carcass persistence can also be impacted by repeated crushing by vehicles and road maintenance (Abra et al., 2018; Barrientos et al., 2018; Santos et al., 2011). Finally, the third process affecting standard mortality surveys along infrastructures is carcass observation bias, which occurs when carcasses within the survey area are not detected by observers, typically due to the sampling method used and the observers' level of experience (Barrientos et al., 2018; Borner et al., 2017; Domínguez del Valle et al., 2020). On roads, this bias tends to be particularly pronounced when roadkill surveys are conducted from vehicles compared to those conducted cycling or walking (Delgado et al., 2019; Guinard et al., 2012; Teixeira et al., 2013).

Although the hierarchical nature of biases in carcass surveys along infrastructures may appear evident, this aspect has received relatively little attention in the scientific literature. While carcass location bias has been largely neglected in mortality estimates (Barrientos et al., 2018; Román et al., 2024), several authors have highlighted the significant underestimation of carcass records due to both persistence and observation bias (Barrientos et al., 2018; Kitano et al., 2023; Teixeira et al., 2013). Nonetheless, not even the hierarchical nature of these two biases has been sufficiently disclosed in the scientific literature. Some notable exceptions relate to road-mortality research, where recent studies have implemented hierarchical statistical models to account for carcass persistence and observation bias combined as latent states when estimating roadkill numbers (Santos et al., 2018), or even extrapolating such estimates to assess the population abundance of the affected species (Fernández-López et al., 2022). However, despite these recent advances, methods that integrate the varying magnitudes of all three biases in carcass surveys are still lacking, hindering the estimation of the total number of killed animals.

In this study, we developed a Bayesian latent-state modelling framework that can effectively integrate location, persistence, and observation biases into a reliable estimate of actual infrastructure-induced mortality across different vertebrate groups. More specifically, we focused on road mortality and roadkill surveys, as the scientific literature on this topic is more extensive than that available for other infrastructures. Our framework is an extension of Bayesian N-mixture models, which estimate abundances from repeated counts (Royle, 2004). We conducted a simulation study to assess the framework's accuracy in recovering the simulated total number of roadkills for different vertebrate groups and survey methods (walking, cycling, and driving). In this study, we implemented multiple scenarios in which we varied the number of road transects surveyed, the daily variability in roadkill numbers and carcass persistence rate, and finally the certainty of prior expert knowledge on location and persistence bias probabilities, which we integrated into our model. We then applied our model to a case study with real data collected by road surveys in southern Spain.

2. Material and methods

2.1 General overview

In this study, we first described our Bayesian hierarchical latent-state modelling framework, which quantifies the total number of roadkills by sequentially assessing how carcass location, persistence and observation biases cause deviations in roadkill census data from actual roadkill (i.e., similar to detection biases in abundance estimation from count data (e.g. Barrientos et al., 2018; Smallwood, 2007) (Figure 1)). We then evaluated the model's performance through a simulation

study, testing different roadkill scenarios across different vertebrate groups. Finally, we applied our model to data from a field case study to estimate the total number of roadkills based on empirical census datasets.

2.2 Modelling framework

We introduce a hierarchical latent-state model to estimate the total number of roadkills, explicitly accounting for the three nested levels of bias: carcass location, persistence, and observation. The model structure is based on the widely used N-mixture models, which estimate abundances from count data while accounting for imperfect detection (Hostetter et al., 2019; Kery & Royle, 2020; Royle, 2004).

We assume that the total number of roadkills $N_{i,t,D}$ varies across *i* = 1…*I* road transects, within *t* = 1…*T* survey periods (with months used as periods in our model, as more frequent surveys are rarely performed), and for a given *D* = the maximum number of days a carcass remains on the survey area before disappearing. This timeframe serves as the window during which we can estimate the total number of roadkills based on the number of observed carcasses in the survey. For instance, if *D* = 3 days, we assume that when carcasses are observed during a road survey (in the morning of day *d*), the roadkill events could have occurred on any day between *d* and *d* – 3, contributing to the total count of carcasses observed in the survey. We define $N_{i,t,D}$ as a random Poisson variable sampled from an average number of roadkills over *D* days $\lambda_{t,D}$:

$$N_{i,t,D} \sim Poisson\,(\lambda_{t,D}) \qquad \text{eqn. 1}$$

As $\lambda_{t,d}$, and consequently $N_{i,t,D}$, can vary across months *t*, our model accounts for seasonal changes in roadkill numbers throughout the year. We assume that $N_{i,t,D} = \sum_{d=1}^{D} N_{i,t,d}$, where each daily total number of roadkills $N_{i,t,d}$ can fluctuate across the days within period *D*, following the daily $\lambda_{t,d}$ in month *t*. However, our framework assumes that $N_{i,t,d}$ and $\lambda_{t,d}$ cannot not be modelled directly and instead needs to be estimated over the maximum persistence time *D*, as conducting daily road monitoring is too resource-demanding to be feasible.

We then define $N2_{i,t,D}$ as the proportion of the total number of roadkills ($N_{i,t,D}$) whose carcasses were located on the road survey area after the collision, determined by the probability of a carcass being located on the road (p_L, carcass location probability):

$$N2_{i,t,D} \sim Binomial\,(p_L, N_{i,t,D}) \qquad \text{eqn. 2}$$

As for $N_{i,t,D}$, we assume that $N2_{i,t,D} = \sum_{d=1}^{D} N2_{i,t,d}$. Based on previous studies (Román et al., 2024), we assume that p_L does not vary among days *d* = 1 – *D*.

Subsequently, we define $N3_{i,t,D} = \sum_{d=1}^{D} N3_{i,t,d}$ as the proportion of roadkills located on the road ($N2_{i,t,D}$) that remain on it until the day of the road survey, determined by the cumulative probability of a carcass persisting on the road survey area, weighted by *D* (p_P, carcass persistence probability)

$$N3_{i,t,D} \sim Binomial\,(p_P, N2_{i,t,D}) \qquad \text{eqn. 3}$$

More precisely, if we assume that the carcass persistence probability could be modelled using a survival function d (e.g., Cox-hazard model as in Santos et al. 2011), then $p_P = \int_{d=1}^{D} S(d)d(d)$, i.e, the average persistence probability from *d* = 1 to *D* (for details, see Supplementary Material S1). If daily $N2_{i,t,d}$ values are known, and assuming *D* = 3, our framework can in theory model: $N3_{i,t,d0} \sim Binomial\left(p_{Pd0}p_{Pd1}p_{Pd2},\ N2_{i,t,d0}\right)$, $N3_{i,t,d1} \sim Binomial\left(p_{Pd1}p_{Pd2},\ N2_{i,t,d1}\right)$, $N3_{i,t,d2} \sim Binomial\left(p_{Pd2},\ N2_{i,t,d2}\right)$, , being p_{Pd} daily carcass persistence probability.

Finally, we define $C_{i,j,t,D,m}$ as the census data, representing the proportion of the total number of roadkills that have persisted in the road survey area during *D* and are recorded in a given road survey, which depend on the carcass observation probability p_{Om}:

$$C_{i,j,t,D,m} \sim Binomial\,(p_{Om}, N3_{i,t,D}) \qquad \text{eqn. 4}$$

We assume a robust-design road survey (Royle, 2004), and thus $C_{i,j,t,D,m}$ varies by road transect *i*, by *m* = 1…*M* methods used for surveying (here: walking, bike, or vehicle), as well as by month *t* (the primary sampling occasion), with *j* = 1…*J* independent sampling replicates each month (secondary sampling occasion). In turn, p_{Om} differs depending on the sampling method *m* used.

Equation 4 builds upon the N-mixture model introduced by Royle (2004), where the estimation of p_{Om} comes from the variability among the independent sampling replicates for each method from a robust design census dataset. That is, we assume that independent observers sampled a given road transect repeatedly during a given road survey. This allows us to make an independent estimation of the observation probability per method p_{Om}.

2.3 Implementation of the model

We used a Bayesian framework to implement our model, using Markov chain Monte Carlo (MCMC) to estimate the parameters (Hobbs & Hooten, 2015). Carcass location probability p_L and carcass persistence probability p_P parameters are not typically estimated directly in roadkill census data, and therefore we assumed them to be latent parameters. We employed beta-distributed informative priors for p_L and p_P, with different parameter estimates for different vertebrate groups. The beta distribution is ideal for modelling probabilities like p_L and p_P because it is defined on

the interval [0,1] and its probability density distribution can take on various shapes, allowing us to represent different levels of prior belief and uncertainty. For each vertebrate group, we defined the α and β parameters of the beta distribution based on a mean estimate for p_L and p_P, reflecting our prior knowledge, and a standard error (SE) that captured our uncertainty around this knowledge (see sections 2.3.1 and 2.4.2 for more information).

Using the respective mean and SE values for p_L and p_P, we calculated the α and β parameters for their prior beta distributions as follows:

$$\alpha = \left(\frac{(1-mean)}{SE^2} - \frac{1}{mean}\right) * mean^2 \quad \text{eqn. 5}$$

$$\beta = \alpha * \left(\frac{1}{mean} - 1\right) \quad \text{eqn. 6}$$

We adopted non-informative priors for the $\lambda_{t,D}$ and p_{Om} parameters, as detailed in Supplementary Material S2.

The MCMC sampling process was conducted in JAGS (Plummer, 2003), operated within the R statistical framework v. 4.2.2 (R Core Team, 2022) through the jagsUI package (Kellner, 2015). To determine model convergence, we used the Gelman–Rubin $\bar{R}$ diagnostic criterion, considering models to have converged when $\bar{R}$ was less than 1.1, following the guidelines by Brooks and Gelman (1998). We also visually inspected the posterior distributions among the different MCMC chains. For each model run, we used three chains of 400,000 iterations with a burn-in period of 100,000 iterations, an adaptive period of 100,000 iterations, and a thinning rate of 1,000.

2.3.1 *Prior information on carcass location (p_L) and carcass persistence (p_P) probability*

We assumed that in most roadkill estimation studies carcass location and persistence probabilities estimations were not available and could not be easily estimated from the census data ($C_{i,j,t,D,m}$). They would have to be entirely modelled as latent states based on prior information. Therefore, in both our simulation and case studies, we integrated such priors based on literature data for these two probabilities.

2.3.1.1 *Carcass location probability (p_L)*

We obtained information on p_L from a recent publication, in which authors determined the probability of a carcass being located on the road after the collision from direct and indirect first-hand observations of vehicle-animal collisions (Román et al., 2024). Based on their data, we reorganized their 150 observations into 10

vertebrate groups (G) using their supplementary material (Amphibians, Reptiles G1, Reptiles G2, Birds/Bats G1, Birds G2, Mammals G1, Mammals G2, Mammals G3, Mammals G4 and Mammals G5; see Table 1). These categories were delineated based on species traits (body size and mobility) which then determined the characteristics of observed roadkill numbers and annual trends (differences in abundances across months), as well as the maximum days their carcasses remain on the road without disappearing (D), and the average p_L, p_P, and p_{Om} values, as shown in Table 1.

We used the observations in Román et al. (2024) to designate a carcass that was located inside the road as success (1) and outside the road as failure (0), and then calculated the mean of successes over each vertebrate group in order to estimate their p_L. In groups where the value of p_L was 1, we assumed the absence of carcass location bias and hence an extremely low probability of being displaced by the collision or being capable of moving after the impact. For this reason, we excluded equation 2 when modelling such categories (i.e., Amphibians, Reptiles G1, Mammals G1, Mammals G2 and Mammals G3 in Table 1), in such cases $N3_{i,t,D}$ being directly dependent on $N_{i,t,D}$ (Supplementary Material S3).

2.3.1.2 Carcass persistence probability (p_P)

Santos, Carvalho, and Mira (2011) was, to our knowledge, the only study providing estimates of mean daily carcass persistence probability (p_{Pd}) for a diverse array of vertebrate groups from Mediterranean habitats, which we were able to adapt to our classification. We used these values to derive p_P (as discussed in section 2.1., see also Supplementary Material S2 and S3 for R code). Santos, Carvalho, and Mira (2011) did not provide information for "Mammals G5". Nonetheless, based on available scientific literature, we contended that this group likely does not demonstrate carcass persistence bias within a monthly time period between successive roadkill surveys (Barrientos et al. 2018). For this reason, we excluded equation 3 when modelling this category, in these cases $C_{i,j,t,D,m}$ being directly dependent on $N2_{i,t,D} = \sum_{d=1}^{D} N2_{i,t,d}$ (Supplementary Material S4)

2.4 Simulation study

To evaluate how well our modelling framework estimated the total number of roadkills and recovered p_L, p_P and p_{Om} as latent states, we simulated different datasets. These datasets were census data $C_{i,j,t,D,m}$, generated based on variations on the number of road transects surveyed, the mean total number of roadkills , $\lambda_{t,d}$, for each specific day *d* in month *t* over the period of *D* days, the daily carcass persistence probability p_{Pd}, and the standard error values that defined the range of

the prior distributions for the latent parameters p_L and p_P (Table 2). Values for p_{Om} were fixed in the simulations depending on the vertebrate group (Table 1).

2.4.1 Census data simulation

For census data simulation, we followed the nested levels of data as described in the modelling framework. We started from the daily total number of roadkills for each respective $N_{i,t,d}$, then the daily proportion of each $N_{i,t,d}$ that were located in the road survey area after the collision $N2_{i,t,d}$, then the proportion of each $N2_{i,t,d}$ that persisted on the road survey area until the road survey day, $N3_{i,t,d}$, and the proportion of the total number of $N3_{i,t,d}$ that had persisted on the road for the maximum persistence period *D*, $N3_{i,t,D}$ and were finally recorded in a given road survey on the census data $C_{i,j,t,D,m}$. The progression through the different nested levels in the simulations was carried out considering the values of p_L, p_P, and p_{Om} which were specific to each vertebrate group (Table 1) and the different scenarios of simulation (Table 2).

We first sampled $N_{i,t,d}$ for each road transect (*i* = 1-10 or 1-100, depending on the simulation scenario, Table 2), month (*t* = 1-12), and day (*d* = 1-*D*, where *D* was the maximum carcass persistence period for a given vertebrate group) as a random Poisson variable based on their mean total number of roadkills in month *t* for each specific day *d* along the *D*-day period $\lambda_{t,d}$ (using eqn. 1). We used expert knowledge to assign variation in month *t* dimension based on data collected by the authors in 2021 and 2022 in southern Spain, which incorporated known monthly fluctuation trends for each vertebrate group, making the census data more realistic (Supplementary Material S5).

As $N_{i,t,d}$ values could vary across the maximum persistence period *D*, along with their corresponding $\lambda_{t,d}$ values, we introduced variability by multiplying each $\lambda_{t,d}$ by a value sampled from a truncated random Normal distribution (mean = 1; *SD* = 0, 0.5, or 1.5, depending on the simulation scenario; Table 2). We used these $\lambda_{t,d}$ to generate $N_{i,t,d}$ values through the Poisson sampling process. For example, considering *D* = 3 days, $\lambda_{t,d0}$, $\lambda_{t,d1}$, and $\lambda_{t,d2}$ would be generated (i.e., the mean number of roadkills occurring three days, two days, or the day before the road survey day, respectively). Roadkills on the survey day itself were not considered as surveys typically occur in the first half of the day. Through the Poisson sampling process, we then obtained the respective $N_{i,t,d0}$, $N_{i,t,d1}$ and $N_{i,t,d2}$, being the total number of roadkills three days, two days and one day before the road survey day respectively. From these values, we could obtain the simulated total number of roadkills $N_{i,t,D} = \sum_{d=1}^{D} N_{i,t,d}$, which we wanted to recover by applying our modelling framework. For Mammals G5 such as ungulates, since we assumed that their carcasses remain onthe road survey area all month and their roadkill numbers were low, simulating $N_{i,t,d}$ values along a D = 30 days period led to an unrealistically high value for $N_{i,t,D}$.

Therefore, here, we simulated a single $N_{i,t,d}$ value for the entire month, such that $N_{i,t,D} = N_{i,t,d}$. As a result, we did not include daily variation in the daily number of roadkills within the month for the simulations.

Next, for vertebrate groups affected by carcass location bias (Table 1), we sampled $N2_{i,t,d}$ values from their respective $N_{i,t,d}$, from a random binomial distribution with p_L as the probability of success (eqn. 2). The p_L value was constant for each respective $N_{i,t,d}$, as we assume that p_L did not vary among days 1 – *D*, using our specific prior mean values on p_L for vertebrate group (Table 1).

Lastly, we sampled $N3_{i,t,d}$ values from their respective $N2_{i,t,d}$, from a random binomial distribution with the daily persistence probabilities p_{Pd} as the probability of success (eqn. 3). For each vertebrate group with their respective p_{Pd} value (see Supplementary material S1), we introduced variability by sampling the p_{Pd} from a truncated Normal distribution (mean = p_{Pd}; *SD* = 0, 0.05, or 0.15, depending on the simulation scenario; Table 2). From these $N3_{i,t,d}$ values we obtained the simulated total number of roadkills that are available to be observed in the survey day $N3_{i,t,D} = \sum_{d=1}^{D} N3_{i,t,d}$. For Mammals G5, which were not affected by carcass persistence bias (i.e., p_P = 1), $N3_{i,t,D}$ was directly dependent on $N_{i,t,D}$ (Supplementary Material S4).

Finally, we sampled census data $C_{i,j,t,D,m}$ from $N3_{i,t,D}$, from a random binomial distribution with the carcass observation probability, p_{Om}, as the probability of success (eqn .4) using *m* = 3 survey methods (i.e. walking, cycling and driving), with *j* = 3 independent sampling replicates per method. We considered the following evidence when assigning p_{Om} values for the different vertebrate groups (in the absence of more concrete data and based on our expert knowledge): (a) we assumed that detection was highest when walking, followed by cycling, and then driving (Guinard et al., 2012; Winton et al., 2018); (b) we assumed that observation for any of the three methods would be low for small vertebrate groups and high for the large, more visible groups (Gerow et al., 2010; Teixeira et al., 2013) (Table 1).

2.4.2 Analysis of simulated data

We aimed to test how well the modelling framework could recover latent parameters assuming different levels of uncertainty in the p_L and p_P prior distributions (eqn. 2 and 3). We analysed the simulated datasets using a low standard error value (*SE* p_L and p_P=0.05), which created a narrow prior distribution, and a high standard error value (*SE* p_L and p_P=0.1), which resulted in a wider prior distribution.

We used "Amphibians" and "Reptiles G1", characterized by a significant peak in roadkill numbers over just a few months, as examples of high seasonal roadkill numbers due to presumed absence or low numbers of roadkills in certain months

where animals were not active (monthly abundance from the 2021 and 2022 data collected <5; see Supplementary Material S5). Here, as we assumed that active and inactive periods were independent, we fitted additional models that only included months where monthly abundance from the 2021 and 2022 data collected was > 5 (see Supplementary Material S6). The aim was to see if model performance improved without accounting for the extended periods with very low roadkill counts, compared to the peak abundance months.

Models were run for all possible scenarios for each vertebrate group, aiming to recover the value of $N_{t,D} = \sum_I^i \sum_D^d N_{i,t,d}$ in our posterior distributions, as well as the values of p_L, p_P, and p_{Om}. In total, we simulated 20 datasets for each of the 36 scenario combinations per vertebrate group, resulting in the analysis of 720 simulated datasets per vertebrate group.

The dataset simulation code is detailed in Supplementary Material S2, S3, S4 and S6. All datasets were generated in R v. 4.2.2 (R Core Team, 2022).

2.4.3 Model evaluation

To evaluate the ability of the modelling framework to recover the simulated parameters we compared their Bayesian posterior distribution of parameters $\hat{\theta}_{s,v,sim,t}$ with the real known simulated parameter value $\theta_{s,v,sim,t}$. In our simulation study we focused on the recovery of $N_{t,D}$ $\theta_{s,v,sim,t}$, and p_L, p_P, and p_{Om} $\theta_{s,v,sim}$ (note that there was no dimension *t* as p_L, p_P, and p_{Om} values did not change across *t* months).

We used the Relative Root Mean Squared Error ($RRMSE$) to compare the model development along every simulation scenario and vertebrate group estimations, as shown in the following equation (Rosenbaum et al., 2024):

$$\mathrm{RRMSE}(\hat{\theta}_{s,v,sim,t}) = \frac{1}{\theta_{s,v,sim,t}} \sqrt{(E(\hat{\theta}_{s,v,sim,t}) - \theta_{s,v,sim,t})^2 + \mathrm{Var}(\hat{\theta}_{s,v,sim,t})} \qquad \text{eqn. 5}$$

where $E(\hat{\theta}_{s,v,sim,t})$ was the mean of the Bayesian posterior distribution for *s* = 1…*S* scenario combination, *v* = 1…*V* vertebrate groups, *sim* = 1…*Sim* simulations and *t* = 1…*T* months. In the case of $N_{t,D}$ $\theta_{s,v,sim,t}$ values, we added 1 to all values because some of them were equal to 0, and so we could not obtain the $RRMSE(\hat{\theta}_{s,v,sim,t})$.

In the case of p_L, p_P, and p_{Om}, we generalized their $RRMSE$ ($\hat{\theta}_{s,v,sim}$) values as the geometric mean of all probability estimates for each vertebrate group ($\overline{RRMSE}(\hat{\theta}_{s,v,sim})$), as shown in the following equation (Rosenbaum et al. 2024):

$$\overline{\mathrm{RRMSE}}(\hat{\theta}_{s,v,sim}) = \prod_{x \in \{p_L, p_P, p_{Om}\}} \mathrm{RRMSE}(\hat{\theta}_{s,v,sim})^{1/|\{p_L, p_P, p_{Om}\}|} \qquad \text{eqn. 6}$$

We also tested whether $N_{t,D}$ $\theta_{s,v,sim,t}$, and p_L, p_P, and p_{Om} $\theta_{s,v,sim}$ overlapped with the 95% credible interval of the $N_{t,D}$ $\hat{\theta}_{s,v,sim,t}$ and p_L, p_P, and p_{Om} $\hat{\theta}_{s,v,sim}$ distributions respectively, ensuring that the true known value was correctly estimated with 95% credibility.

2.5 Case study

We applied our modelling framework to estimate the total number of roadkills across *i* = 9 road transects of 3 km each, in three different Mediterranean ecosystems in south-western Spain (Supplementary Material S5). The first ecosystem was located in a hilly landscape belonging to the Natural Park of *Sierra Norte de Sevilla*, characterized by mountainous Mediterranean dehesa, composed mainly of holm oaks *Quercus ilex*, bushes and grasslands. We surveyed the SE-5405 road from Castilblanco de los Arroyos to Almadén de la Plata (37°45'08.9"N, 6°02'38.7"W). The second ecosystem was mostly plain farmland characterized by sunflower, wheat and olive tree plantations. We selected three roads: the SE-6103 from Carmona to La Campana (37°31'41.5"N, 5°29'37.5"W), the A-456 from La Campana to Lora del Río, and the A-457 from Lora del Rio to Carmona (37°36'31.0"N, 5°28'46.7"W). Finally, the third ecosystem was the mainly flat agroforestry matrix surrounding Doñana National Park, characterized by Mediterranean dehesa, pine trees and interspersed orchards, fruit trees and other minor crops. We selected the A-481 road from Hinojos to Villamanrique de la Condesa (37°14'01.7"N, 6°19'38.0"W) and the A-494 road from Mazagón to Matalascañas (37°07'05.9"N 6°46'06.5"W).

We collected data on these road sections using *M* = 3 different methods carried simultaneously (walking, cycling, and driving), with *J* = 2 independent sampling repetitions per method (thereby guaranteeing a robust sampling design) for each method and *T* = 4 monthly surveys from February to May in 2023. For each 3 km transect, one observer conducted an initial survey, followed by a second observer after a 10-minute break, considering this interval short enough to assume that the roadkill population was closed. Due to administrative and legal requirements, during the initial phase of the driving surveys, the first observer was solely responsible for roadkill sampling while the second focused entirely on driving. In the subsequent transect sampling repetition, the roles were reversed, allowing the driver to also take on the task of searching for roadkill, ensuring both observers made independent samplings. The survey velocity while driving was the minimum allowed on the road.

For each roadkill detected, we noted the observer's identity, the surveyed transect, sampling method, observation month and the exact georeferenced location of the roadkill (with less than 10 m error). Roadkills were documented with zenithal photographs and identified to the lowest feasible taxonomic level, although the ultimate goal was to group them into functional groups.

3. Results

In our Bayesian model analysis, the $\bar{R}$ statistic consistently showed values below 1.1, indicating good convergence and precise parameter estimations from the MCMC chains (Appendix A1; A2).

3.1 Simulation study

Our model outputs demonstrated overall low $RRMSE(\hat{\theta}_{s,v,sim,t})$ values recovering the simulated total number of roadkills $N_{t,D}(\theta_{s,v,sim,t})$ across nearly all scenarios (Fig. 2; see also Supplementary Material S7 for more detailed plots for each of the vertebrate groups). Nevertheless, the vertebrate groups Reptiles G2, Birds G2, and Mammals G3 showed very high variation in their distributions, ranging from log $RRMSE(\hat{\theta}_{s,v,sim,t})$ values below -1 to over 4 (Fig. 2). Across vertebrate groups, the highest $RRMSE(\hat{\theta}_{s,v,sim,t})$ scores, indicating relatively worse performance of the model in recovering simulated parameters, corresponded to scenarios with high variability in daily persistence probabilities (*SD* p_{Pd}) (Fig. 2). Additionally, $RRMSE(\hat{\theta}_{s,v,sim,t})$ increased when the *SE* was high for the prior distributions for p_L and p_P compared with low *SE*, and increased further when variability in daily roadkill numbers (*SD* $\lambda_{t,d}$) was also high (Fig. 2). The number of road transects simulated (10 or 100 transects) had minimal impact on $RRMSE(\hat{\theta}_{s,v,sim,t})$ (Supplementary Material S7). For Amphibians and Reptiles G1, $RRMSE(\hat{\theta}_{s,v,sim,t})$ were lower when considering only those months in the analyses when animals are active, compared to when extended periods of low number of roadkills were included in the datasets (Supplementary Material S7). Regarding the overlap of the 95% credible interval for $N_{t,D}\,\hat{\theta}_{s,v,sim,t}$ across all scenarios, when the p_L and p_P prior *SE* was low, the simulated total number of roadkills $N_{t,D}(\theta_{s,v,sim,t})$ was generally well recovered for all vertebrate groups (Fig. 4a; see Supplementary Material S8 for more detailed plots). However, there were some exceptions: Amphibians and Mammals G4 were underestimated in some scenarios, while Reptiles G1 were underestimated in all scenarios (Supplementary Material S8). On the other hand, when p_L and p_P prior *SE* was high, the 95% credible interval overlap widened, leading to overestimations across all vertebrate groups (Supplementary Material S8). The only exceptions were Amphibians and Reptiles G1, which were typically underestimated, resulting in increased uncertainty but reduced underestimation (Supplementary Material S8).

$\overline{RRMSE}(\hat{\theta}_{s,v,sim})$ scores for p_L, p_P and p_{Om} $(\theta_{s,v,sim})$ showed the same relative differences as $RRMSE(\hat{\theta}_{s,v,sim,t})$, being the highest for Reptiles G2, Birds G2, and Mammals G3 under high variability in daily carcass persistence probabilities (*SD* p_{Pd}) (Fig. 3). This was largely due to the fact that the Bayesian hierarchical models could not recover well p_P under high *SD* p_{Pd} and a high p_P prior *SE* , although p_{Om} values where always well recovered, being much more precise in 100 survey sites scenario (Fig. 4b; see Supplementary Material S9 for more detailed plots).

3.2 *Case study*

During the sampling period, we recorded a total of 650 different carcasses of 45 identified species (386 of these carcasses could only be classified into higher taxonomic groups). For further modelling, we classified these carcasses into the following functional groups: 199 lizards, 17 snakes, 217 passerines, 43 small mammals, 72 lagomorphs, and 24 carnivores. Although we observed 40 amphibians, 22 medium-sized birds, 12 hedgehogs and 4 big-sized mammals we were unable to estimate the total number of roadkills for these groups. Standardizing observations across the 27 km surveyed (9 transects × 3 km), the roadkill rates per kilometre were highest for passerines (8.04/km) and lizards (7.37/km), followed by lagomorphs (2.67/km), small mammals (1.59/km), carnivores (0.89/km), and snakes (0.63/km).

Our model generated estimates for the total number of roadkills over the 4 months of sampling on our study roads, taking into account prior distributions of p_L and p_P, alongside the estimated values of p_{Om} for each sampling method used. The estimated roadkill rates per kilometre were 15.22 for lizards (2.07 times higher than observed), 8.84 for snakes (14.03 times higher), 48.92 for passerines (6.08 times higher), 7.64 for small mammals (4.81 times higher), 7 for lagomorphs (2.62 times higher), and 5.49 for carnivores (6.16 times higher) (see Fig. 5). For each vertebrate, p_{Om} estimation is highest for walking survey method p_{Ow}, followed by cycling p_{Oc}, and is considerably lower for driving p_{Od}. This was particularly evident in lizards, passerines, and lagomorphs, where p_{Ow} was markedly higher compared to the other methods. For lizards and small mammals, the probability of observation was generally low, with values concentrated close to zero when using the driving method (Fig. 6).

Finally, our data revealed that some carcasses were observed exclusively by one survey method and not by the others: 294 carcasses were only observed using the walking method, 134 by the cycling method, and 1 by the driving method (Supplementary material 11).

4. Discussion

4.1 *Integrating biases in surveys of infrastructure-induced mortality*

In the present study, we successfully integrated the three intrinsic survey biases of infrastructure-induced mortality (i.e., carcass location, persistence, and observation bias) within the predefined conceptual framework of our modelling approach. As a consequence, we were able to robustly infer the actual mortality from carcass census data, which represents a significant step forward in methodological research on this type of impact, with potentially important implications for the conservation of

threatened species as well as for taxa providing ecosystem services. Unlike earlier studies that implemented similar statistical approaches, which provided abundance indices (e.g., Fernández-López et al., 2022) or roadkill risk metrics (e.g., Santos et al., 2018), the application of modified Bayesian N-mixture models in our study allowed us to derive actual roadkill estimates while propagating uncertainty throughout the model due to the Bayesian approach (Schmelter et al., 2012). Our roadkill estimates were between 2.07 and 14.03 times higher than the observed records in the case study (depending on the species group considered), highlighting that road mortality is a far greater threat than previously recognized, especially for species more affected by sampling biases, such as small birds and bats (Barrientos et al., 2018; Román et al., 2024). Since the biases analyzed in this study are very similar to those affecting other infrastructure-induced mortality surveys (Barrientos et al., 2018; Bernardino et al., 2020), it is reasonable to assume that this threat is also underestimated along power lines, wind farms and other linear developments.

4.2 *Model performance in simulation scenarios*

The worse performance of the model in scenarios with high standard error in the Bayesian prior distributions of carcass location and persistence probabilities could be due to the fact that a higher prior variation is translated to weakly informative priors. Since we lacked empirical data to refine these weakly informative priors, this uncertainty propagated through the model, resulting in more uncertain posterior distribution estimates. Bayesian mixture models have been noted before to run into performance issues when data are scarce and prior information for latent parameters is uninformative (Depaoli (2013) or Depaoli, Yang, and Felt (2017)). Thus, improving knowledge of carcass location and persistence probabilities is crucial for future roadkill studies. Regarding carcass location bias, this can be addressed through collaborative research efforts, such as those already being conducted by some authors (Román et al. 2024). On the other hand, our knowledge of persistence bias can be improved through targeted experiments of persistence times (Ruiz-Capillas et al., 2015; Santos & Ascensão, 2019) and more intensive censuses, although this is often associated with high logistical and economic costs (Costa et al., 2015; Henry et al., 2021). It is important to emphasize that the study of these biases needs to be further explored in other infrastructures as well, and in fact, even more than in roads, about which comparatively more studies have been conducted (Barrientos et al., 2018).

4.2.2 *Impact of daily parameters variability*

Not surprisingly, the combination of a high variability in daily roadkill numbers and daily persistence probabilities negatively affected the agreement between modelled and simulated total roadkill numbers in some simulations, thereby increasing the variation in Relative Root Mean Square Error ($RRMSE$). This occurs because total roadkill numbers are a non-linear function of the combination of the parameters. By averaging daily values of potentially highly variable parameters (e.g., high roadkill numbers on certain days, combined with low persistence probabilities and vice versa), we introduce non-linear averaging that may skew estimates upwards or downwards (Denny, 2017), causing the $RRMSE$ value to vary.

Our results also suggest that true average persistence probabilities may be underestimated when the simulated daily persistence shows high variation and carcasses may persist more than five days. This can be explained by the fact that the prior for the average persistence probability is obtained using a persistence curve that follows a convex function (see Supplementary Material S1). Under Jensen's inequality (Ruel & Ayres, 1999) given our convex function, the average value of our persistence curve across its range is typically lower than the value of the mean of the different daily persistence probabilities values that generate the convex curve (see Supplementary Material S11).

4.2.3 *Differences among vertebrate groups*

The fact that we underestimated simulated roadkill numbers for amphibians and lizards is likely due to the fact that extremely low persistence and observation probabilities in these groups resulted in a zero-inflated simulated dataset for analysis. In such cases, employing a zero-inflated Poisson version of the N-mixture model could potentially yield more accurate results (Joseph et al., 2009; Wenger & Freeman, 2008). Regarding carnivores, in our models, this vertebrate group has particularly high carcass location, persistence and observation probabilities, in agreement with previous knowledge (Barrientos et al., 2018), leading to posterior distributions that are very precise and narrow (Veech et al., 2016). However, the underestimation of this group in certain scenarios remains unexplained, requiring further investigation.

4.3 *Case study application*

Applying the hierarchical modelling framework to empirical data in our case study showed an important increase in the estimated number of roadkills compared to those observed, aligning with the findings of other studies (e.g Teixeira et al. (2013); Winton et al. (2018)). Also, our estimates for carcass observation probabilities align with previous findings in the literature, as it is highest for walking surveys, followed by cycling, and lowest for driving (Guinard et al., 2012; Ogletree & Mead, 2020; Winton et al., 2018), and it is also lower for smaller vertebrate groups and higher for larger, more visible species (Gerow et al., 2010; Teixeira et al., 2013). Our study is the first to compare all three survey methods simultaneously within the same study. We not only demonstrate that walking surveys—while the most effective method—are not perfect and should not be assumed to observe all roadkill events, as was done in Teixeira et al. (2013), but we also show that a significant number of carcasses were missed by walking surveys but observed by cycling. This suggests that walking, cycling, and driving surveys should not be seen as a ranking from best to worst but rather as complementary methods, each with its own advantages and limitations. For example, while walking likely helps observe carcasses directly underfoot, the elevated perspective provided by cycling allows for a broader field of view, making it easier to observe carcasses on the roadside.

These results highlight that using the driving method in surveys not only reduces the proportion of carcasses observed on the road but can also lead to an overestimation of the total number of collisions. In N-mixture models, lower observation probabilities result in larger extrapolations in the estimated values. Since observation probabilities

while driving are extremely low, the estimated number of roadkills of collisions ultimately be much higher than the real number of roadkills (Dennis et al., 2015; Hostetter et al., 2019).

4.4 *Limitations and future perspectives*

A limitation of our methodology is that it requires extensive knowledge of carcass location, persistence, and observation biases specific to each infrastructure, vertebrate group, and study environment. The bias values for each of these contexts may vary, which is crucial for making accurate estimates in each case.

Another limitation of our estimates of total roadkill numbers is that they are limited by the maximum number of days a carcass from a specific vertebrate group remains on the road before disappearing. This means that to estimate the number of roadkill within specific time units (day, month, etc.), further methodological development of the model would be needed to ensure that the estimates are properly linked to the chosen time unit. Consequently, when extrapolating monthly roadkill levels, vertebrate groups with shorter persistence times (such as amphibians and lizards) may show significantly higher roadkill estimates compared to those with longer persistence times (such as large birds and carnivores). To address the accuracy of monthly extrapolations, roadkill survey frequency should take into account the persistence period of the target vertebrate group. This approach would be particularly useful in studies focused on endangered or high-interest species, due to most studies do not typically follow this method, as they generally assess overall vertebrate mortality (e.g. D'Amico et al., 2015). For species with short persistence times, such as lizards, surveys should be done every day throughout the study season to avoid extrapolation and rely on actual observed data.

In the case study, one important consideration is that, typically, roadkill studies alternate the direction of search and the side of the road randomly in order to cover the area as thoroughly as possible along the infrastructure (D'Amico et al., 2015). However, in our case, as our study was an initial phase of a citizen science project with volunteers, we had to employ a simple and easy sampling method, conducting surveys on only one side of the road and always in the same direction. Although we recognize that this may decrease the carcass observation probability, it would be interesting to investigate in the future whether randomizing the direction and side of the road would actually reduce this bias.

Finally, our modelling framework could be used for animal conservation issues by combining it with population abundance estimation models near to infrastructure, offering a valuable tool to assess what proportion of the study population may succumb to infrastructure-related mortality (e.g. roads: Barrientos et al. (2021); power lines: Biasotto & Kindel (2018) and D'Amico et al. (2019); multiple linear infrastructures: Ascensão et al. (2022)) This information would facilitate the identification of species or populations more significantly affected by infrastructure-related mortality (e.g. species with very low population sizes and highly susceptible to roadkill), thereby prioritizing conservation efforts.

Bibliography

693 Abra, F. D., Huijser, M. P., Pereira, C. S., & Ferraz, K. M. P. M. B. (2018). How reliable are
694 your data? Verifying species identification of road-killed mammals recorded by road
695 maintenance personnel in São Paulo State, Brazil. *Biological Conservation*, *225*, 42–
696 52. https://doi.org/10.1016/j.biocon.2018.06.019
697 Ascensão, F., D'Amico, M., & Barrientos, R. (2022). No Planet for Apes? Assessing Global
698 Priority Areas and Species Affected by Linear Infrastructures. *International Journal of*
699 *Primatology*, *43*(1), 57–73. https://doi.org/10.1007/s10764-021-00207-5
700 Barrientos, R., Ascensão, F., D'Amico, M., Grilo, C., & Pereira, H. M. (2021). The lost road:
701 Do transportation networks imperil wildlife population persistence? *Perspectives in*
702 *Ecology and Conservation*, *19*(4), 411–416.
703 https://doi.org/10.1016/j.pecon.2021.07.004
704 Barrientos, R., Martins, R. C., Ascensão, F., D'Amico, M., Moreira, F., & Borda-de-Água, L.
705 (2018). A review of searcher efficiency and carcass persistence in infrastructure-
706 driven mortality assessment studies. *Biological Conservation*, *222*, 146–153.
707 https://doi.org/10.1016/j.biocon.2018.04.014
708 Bech, N., Beltran, S., Boissier, J., Allienne, J.-F., Resseguier, J., & Novoa, C. (2012). Bird
709 mortality related to collisions with ski-lift cables: Do we estimate just the tip of the
710 iceberg? *Animal Biodiversity and Conservation*, *35*(1), 95–98.
711 Bernardino, J., Bevanger, K., Barrientos, R., Dwyer, J. F., Marques, A. T., Martins, R. C.,
712 Shaw, J. M., Silva, J. P., & Moreira, F. (2018). Bird collisions with power lines: State
713 of the art and priority areas for research. *Biological Conservation*, *222*, 1–13.
714 https://doi.org/10.1016/j.biocon.2018.02.029
715 Bernardino, J., Bispo, R., Martins, R. C., Santos, S., & Moreira, F. (2020). Response of
716 vertebrate scavengers to power line and road rights-of-way and its implications for
717 bird fatality estimates. *Scientific Reports*, *10*(1), 15014.
718 https://doi.org/10.1038/s41598-020-72059-7

719 Biasotto, L. D., & Kindel, A. (2018). Power lines and impacts on biodiversity: A systematic
720 review. *Environmental Impact Assessment Review*, *71*, 110–119.
721 https://doi.org/10.1016/j.eiar.2018.04.010
722 Borner, L., Duriez, O., Besnard, A., Robert, A., Carrere, V., & Jiguet, F. (2017). Bird collision
723 with power lines: Estimating carcass persistence and detection associated with
724 ground search surveys. *Ecosphere*, *8*(11), e01966. https://doi.org/10.1002/ecs2.1966
725 Brooks, S. P., & Gelman, A. (1998). General Methods for Monitoring Convergence of
726 Iterative Simulations. *Journal of Computational and Graphical Statistics*, *7*(4), 434–
727 455. https://doi.org/10.1080/10618600.1998.10474787
728 Costa, A. S., Ascensão, F., & Bager, A. (2015). Mixed sampling protocols improve the cost-
729 effectiveness of roadkill surveys. *Biodiversity and Conservation*, *24*(12), 2953–2965.
730 https://doi.org/10.1007/s10531-015-0988-3
731 D'Amico, M., Ascensão, F., Fabrizio, M., Barrientos, R., & Gortázar, C. (2018). Twenty years
732 of road ecology: A topical collection looking forward for new perspectives. *European*
733 *Journal of Wildlife Research*, *64*, 1–2.
734 D'Amico, M., Catry, I., Martins, R. C., Ascensão, F., Barrientos, R., & Moreira, F. (2018). Bird
735 on the wire: Landscape planning considering costs and benefits for bird populations
736 coexisting with power lines. *Ambio*, *47*(6), 650–656. https://doi.org/10.1007/s13280-
737 018-1025-z
738 D'Amico, M., Martins, R. C., Álvarez-Martínez, J. M., Porto, M., Barrientos, R., & Moreira, F.
739 (2019). Bird collisions with power lines: Prioritizing species and areas by estimating
740 potential population-level impacts. *Diversity and Distributions*, *25*(6), 975–982.
741 https://doi.org/10.1111/ddi.12903
742 D'Amico, M., Román, J., De los Reyes, L., & Revilla, E. (2015). Vertebrate road-kill patterns
743 in Mediterranean habitats: Who, when and where. *Biological Conservation*, *191*,
744 234–242.
745 Delgado, J. D., Durán Humia, J., Rodríguez Pereiras, A., Rosal, A., del Valle Palenzuela, M.,
746 Morelli, F., Arroyo Hernández, N. L., & Rodríguez Sánchez, J. (2019). The spatial

distribution of animal casualties within a road corridor: Implications for roadkill monitoring in the southern Iberian rangelands. *Transportation Research Part D: Transport and Environment*, *67*, 119–130. https://doi.org/10.1016/j.trd.2018.11.017

Dennis, E. B., Morgan, B. J. T., & Ridout, M. S. (2015). Computational Aspects of N-Mixture Models. *Biometrics*, *71*(1), 237–246. https://doi.org/10.1111/biom.12246

Denny, M. (2017). The fallacy of the average: On the ubiquity, utility and continuing novelty of Jensen's inequality. *Journal of Experimental Biology*, *220*(2), 139–146. https://doi.org/10.1242/jeb.140368

Depaoli, S. (2013). Mixture class recovery in GMM under varying degrees of class separation: Frequentist versus Bayesian estimation. *Psychological Methods*, *18*(2), 186–219. https://doi.org/10.1037/a0031609

Depaoli, S., Yang, Y., & Felt, J. (2017). Using Bayesian Statistics to Model Uncertainty in Mixture Models: A Sensitivity Analysis of Priors. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(2), 198–215. https://doi.org/10.1080/10705511.2016.1250640

DeVault, T. L., Seamans, T. W., Linnell, K. E., Sparks, D. W., & Beasley, J. C. (2017). Scavenger removal of bird carcasses at simulated wind turbines: Does carcass type matter? *Ecosphere*, *8*(11), e01994. https://doi.org/10.1002/ecs2.1994

Dhiab, O., D'Amico, M., & Selmi, S. (2023). Experimental evidence of increased carcass removal along roads by facultative scavengers. *Environmental Monitoring and Assessment*, *195*(1), 216.

Domínguez del Valle, J., Cervantes Peralta, F., & Jaquero Arjona, M. I. (2020). Factors affecting carcass detection at wind farms using dogs and human searchers. *Journal of Applied Ecology*, *57*(10), 1926–1935. https://doi.org/10.1111/1365-2664.13714

Fernández-López, J., Blanco-Aguiar, J. A., Vicente, J., & Acevedo, P. (2022). Can we model distribution of population abundance from wildlife–vehicles collision data? *Ecography*, *2022*(5), e06113. https://doi.org/10.1111/ecog.06113

Gerow, K., Kline, N. C., Swann, D. E., & Pokorny, M. (2010). Estimating annual vertebrate mortality on roads at Saguaro National Park, Arizona. *Human-Wildlife Interactions*, *4*(2), 283–292.

Guil, F., Àngels Colomer, M., Moreno-Opo, R., & Margalida, A. (2015). Space–time trends in Spanish bird electrocution rates from alternative information sources. *Global Ecology and Conservation*, *3*, 379–388. https://doi.org/10.1016/j.gecco.2015.01.005

Guinard, É., Julliard, R., & Barbraud, C. (2012). Motorways and bird traffic casualties: Carcasses surveys and scavenging bias. *Biological Conservation*, *147*(1), 40–51. https://doi.org/10.1016/j.biocon.2012.01.019

Henry, D. A. W., Collinson-Jonker, W. J., Davies-Mostert, H. T., Nicholson, S. K., Roxburgh, L., & Parker, D. M. (2021). Optimising the cost of roadkill surveys based on an analysis of carcass persistence. *Journal of Environmental Management*, *291*, 112664. https://doi.org/10.1016/j.jenvman.2021.112664

Hobbs, N. T., & Hooten, M. B. (2015). *Bayesian Models: A Statistical Primer for Ecologists* (STU-Student edition). Princeton University Press. https://www.jstor.org/stable/j.ctt1dr36kz

Hostetter, N. J., Gardner, B., Sillett, T. S., Pollock, K. H., & Simons, T. R. (2019). An integrated model decomposing the components of detection probability and abundance in unmarked populations. *Ecosphere*, *10*(3), e02586. https://doi.org/10.1002/ecs2.2586

Joseph, L. N., Elkin, C., Martin, T. G., & Possingham, H. P. (2009). Modeling abundance using N-mixture models: The importance of considering ecological mechanisms. *Ecological Applications*, *19*(3), 631–642. https://doi.org/10.1890/07-2107.1

Kellner, K. (2015). jagsUI: a wrapper around rjags to streamline JAGS analyses. *R Package Version*, *1*(1).

Kery, M., & Royle, J. A. (2020). *Applied Hierarchical Modeling in Ecology: Analysis of Distribution, Abundance and Species Richness in R and BUGS. Academic Press.*

Kitano, M., Smallwood, K. S., & Fukaya, K. (2023). Bird carcass detection from integrated trials at multiple wind farms. *The Journal of Wildlife Management*, *87*(1), e22326. https://doi.org/10.1002/jwmg.22326

Meijer, J. R., Huijbregts, M. A. J., Schotten, K. C. G. J., & Schipper, A. M. (2018). Global patterns of current and future road infrastructure. *Environmental Research Letters*, *13*(6), 064006. https://doi.org/10.1088/1748-9326/aabd42

Nazir, M. S., Ali, N., Bilal, M., & Iqbal, H. M. N. (2020). Potential environmental impacts of wind energy development: A global perspective. *Current Opinion in Environmental Science & Health*, *13*, 85–90. https://doi.org/10.1016/j.coesh.2020.01.002

Ogletree, K., & Mead, A. (2020). What Roadkills Did We Miss in a Driving Survey? A Comparison of Driving and Walking Surveys in Baldwin County, Georgia. *Georgia Journal of Science*, *78*(2). https://digitalcommons.gaacademy.org/gjs/vol78/iss2/8

Pearce-Higgins, J. W., Stephen, L., Douse, A., & Langston, R. H. W. (2012). Greater impacts of wind farms on bird populations during construction than subsequent operation: Results of a multi-site and multi-species analysis. *Journal of Applied Ecology*, *49*(2), 386–394. https://doi.org/10.1111/j.1365-2664.2012.02110.x

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Working Papers*.

R Core Team. (2022). *R: A language and environment for statistical computing.* [R Foundation for Statistical Computing].

Ravache, A., Barré, K., Normand, B., Goislot, C., Besnard, A., & Kerbiriou, C. (2024). Monitoring carcass persistence in windfarms: Recommendations for estimating mortality. *Biological Conservation*, *292*, 110509. https://doi.org/10.1016/j.biocon.2024.110509

Román, J., Rodríguez, C., García-Rodríguez, A., Diez-Virto, I., Gutiérrez-Expósito, C., Jubete, F., Paniw, M., Clavero, M., Revilla, E., & D'Amico, M. (2024). Beyond crippling bias: Carcass-location bias in roadkill studies. *Conservation Science and Practice*, *6*(4), e13103. https://doi.org/10.1111/csp2.13103

Rosenbaum, B., Li, J., Hirt, M. R., Ryser, R., & Brose, U. (2024). Towards understanding interactions in a complex world: Design and analysis of multi-species functional response experiments. *Methods in Ecology and Evolution*, *15*(9), 1704–1719. https://doi.org/10.1111/2041-210X.14372

Royle, J. A. (2004). N-Mixture Models for Estimating Population Size from Spatially Replicated Counts. *Biometrics*, *60*(1), 108–115. https://doi.org/10.1111/j.0006-341X.2004.00142.x

Ruel, J. J., & Ayres, M. P. (1999). Jensen's inequality predicts effects of environmental variation. *Trends in Ecology & Evolution*, *14*(9), 361–366. https://doi.org/10.1016/S0169-5347(99)01664-X

Ruiz-Capillas, P., Mata, C., & Malo, J. E. (2015). How many rodents die on the road? Biological and methodological implications from a small mammals' roadkill assessment on a Spanish motorway. *Ecological Research*, *30*, 417–427.

Santos, R. A. L., & Ascensão, F. (2019). Assessing the effects of road type and position on the road on small mammal carcass persistence time. *European Journal of Wildlife Research*, *65*, 1–5.

Santos, R. A. L., Mota-Ferreira, M., Aguiar, L. M. S., & Ascensão, F. (2018). Predicting wildlife road-crossing probability from roadkill data using occupancy-detection models. *Science of The Total Environment*, *642*, 629–637. https://doi.org/10.1016/j.scitotenv.2018.06.107

Santos, S. M., Carvalho, F., & Mira, A. (2011). How Long Do the Dead Survive on the Road? Carcass Persistence Probability and Implications for Road-Kill Monitoring Surveys. *PLOS ONE*, *6*(9), e25383. https://doi.org/10.1371/journal.pone.0025383

Schmelter, M. L., Erwin, S. O., & Wilcock, P. R. (2012). Accounting for uncertainty in cumulative sediment transport using Bayesian statistics. *Geomorphology*, *175–176*, 1–13. https://doi.org/10.1016/j.geomorph.2012.06.012

Smallwood, K. S. (2007). Estimating Wind Turbine-Caused Bird Mortality. *The Journal of Wildlife Management*, *71*(8), 2781–2791. https://doi.org/10.2193/2007-006

857 Tabassum-Abbasi, Premalatha, M., Abbasi, T., & Abbasi, S. A. (2014). Wind energy:
858 Increasing deployment, rising environmental concerns. *Renewable and Sustainable*
859 *Energy Reviews*, *31*, 270–288. https://doi.org/10.1016/j.rser.2013.11.019
860 Teixeira, F. Z., Coelho, A. V. P., Esperandio, I. B., & Kindel, A. (2013). Vertebrate road
861 mortality estimates: Effects of sampling methods and carcass removal. *Biological*
862 *Conservation*, *157*, 317–323. https://doi.org/10.1016/j.biocon.2012.09.006
863 Veech, J. A., Ott, J. R., & Troy, J. R. (2016). Intrinsic heterogeneity in detection probability
864 and its effect on -mixture models. *Methods in Ecology and Evolution*, *7*(9), 1019–
865 1028. https://doi.org/10.1111/2041-210X.12566
866 Wenger, S. J., & Freeman, M. C. (2008). Estimating Species Occurrence, Abundance, and
867 Detection Probability Using Zero-Inflated Distributions. *Ecology*, *89*(10), 2953–2959.
868 https://doi.org/10.1890/07-1127.1
869 Winton, S. A., Taylor, R., Bishop, C. A., & Larsen, K. W. (2018). Estimating actual versus
870 detected road mortality rates for a northern viper. *Global Ecology and Conservation*,
871 *16*, e00476. https://doi.org/10.1016/j.gecco.2018.e00476
872

873 **Figures and tables**
874
875 **Table 1**. Descriptive characteristics of the different vertebrate groups used to
876 simulate roadkill numbers, including examples of species, their features of observed
877 roadkill numbers and their annual trend, maximum days their carcass remains on the
878 road without disappearing (D), probability of their carcass being located on the road
879 (p_L), average probability across D of their carcass persisting on the road (p_P) and
880 carcass observation probability (p_{Om}) by walking (p_{Ow}), cycling (p_{Oc}) and driving (p_{Odr})
881 survey method.

Vertebrate groups	Example	Observed roadkill abundance	Seasonal variation	D (days)	p_L	p_P	p_{Om}		
							p_{Ow}	p_{Oc}	p_{Odr}
Amphibians	Small amphibians such as *Bufo spinosus* or *Epidalea calamita*	Frequently observed	High	2	1	0.36	0.5	0.3	0.02

Reptiles G1	Small reptiles such as *Psammodromus algirus* or *Timon lepidus*	Frequently observed	High	1	1	0.54	0.5	0.3	0.02
Reptiles G2	Medium-sized ophidians such as *Malpolon monspessulanus* or *Zamenis scalaris*	Frequently observed	High	3	0.43	0.36	0.7	0.5	0.1
Birds/Bats G1	Small birds such as *Carduelis carduelis* or bats	Frequently observed	Low	3	0.36	0.36	0.6	0.4	0.05
Birds G2	Medium-sized birds such as *Alectoris rufa* or large birds as *Asio otus*	Rarely observed	Low	10	0.69	0.34	0.8	0.6	0.2
Mammals G1	Small mammals such as *Mus spretus* or *Rattus rattus*	Frequently observed	Low	3	1	0.36	0.6	0.4	0.05
Mammals G2	Medium-sized mammals such as *Oryctolagus cuniculus* or *Lepus granatensis*	Frequently observed	Low	4	1	0.35	0.8	0.6	0.2
Mammals G3	Mammals with keratinous structures such as *Erinaceus europaeus*	Rarely observed	Low	12	1	0.34	0.8	0.6	0.2
Mammals G4	Medium-sized carnivores as *Felis catus* or *Vulpes vulpes*	Frequently observed	Low	14	0.65	0.34	0.9	0.7	0.3
Mammals G5	Big mammals as *Sus scrofa* or *Cervus elaphus*	Rarely observed	Low	30	0.5	1	1	0.9	0.8

Table 2. Simulation scenarios to generate roadkill census data, including levels of variation and justification for the scenario choice. $\lambda_{t,d}$= mean total number of roadkills in month t for each specific day *d* across *D* (maximum persistence), p_{Pd} = daily carcass persistence probability, *SD* = Standard Deviation and *SE* = Standard Error.

Parameter	Levels	Justification
Nº road transect	10/100	N-mixture models can be sensitive to the spatial replication of count surveys (Kery & Royle, 2020). Increasing the number of transects can enhance the precision of estimates by improving the spatial representativeness of the data
SD in $\lambda_{t,d}$	0/0.5/1.5	Since we model the total number of roadkills as the sum over the maximum persistence period (D), we aim to know how this modelling approach impacts our estimates when daily values show no variation, moderate variation, or high variation

SD in p_{Pd}	0/0.05/0.15	Since we model the carcass persistence probability as the average of carcass persistence probabilities over the maximum persistence period (D), we aim to know how this modelling approach impacts our estimates when daily values show no variation, moderate variation, or high variation
SE in priors p_L and p_P	0.05/0.1	Since we model our prior beta-distribution α and β parameters for a p_L and p_P from their mean values and a SE that captures our uncertainty around this knowledge, we aim to know how low and high uncertainty impacts our estimates

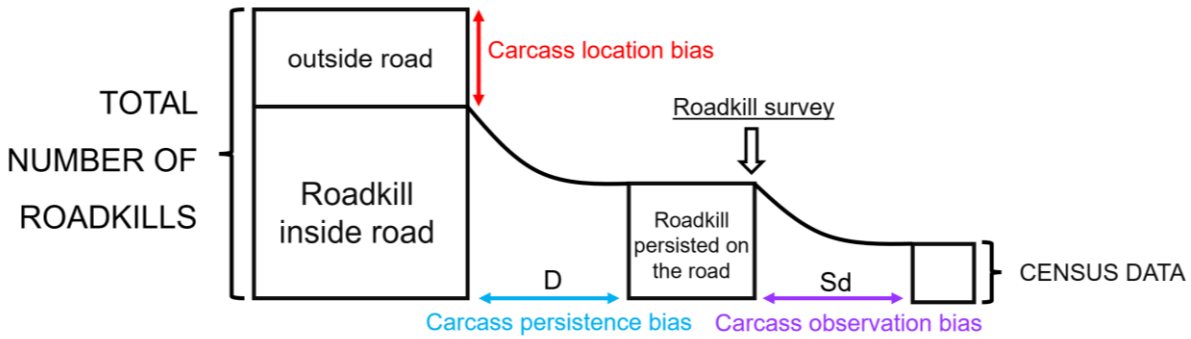


Figure 1. Roadkill survey bias framework. This diagram illustrates how three types of survey bias (carcass location bias, carcass persistence bias, and carcass observation bias) impact the census data of roadkill within the surveyed road. These theoretical different sizes of the squares in the diagram symbolize the quantity of roadkill that would be available at each nested level of the framework. Additionally, D represents the time elapsed between the roadkill event and the maximum days a carcass remains on the road without disappearing until survey day, where carcass persistence bias occurs, while Sd represents the survey duration, during which observational bias occurs.

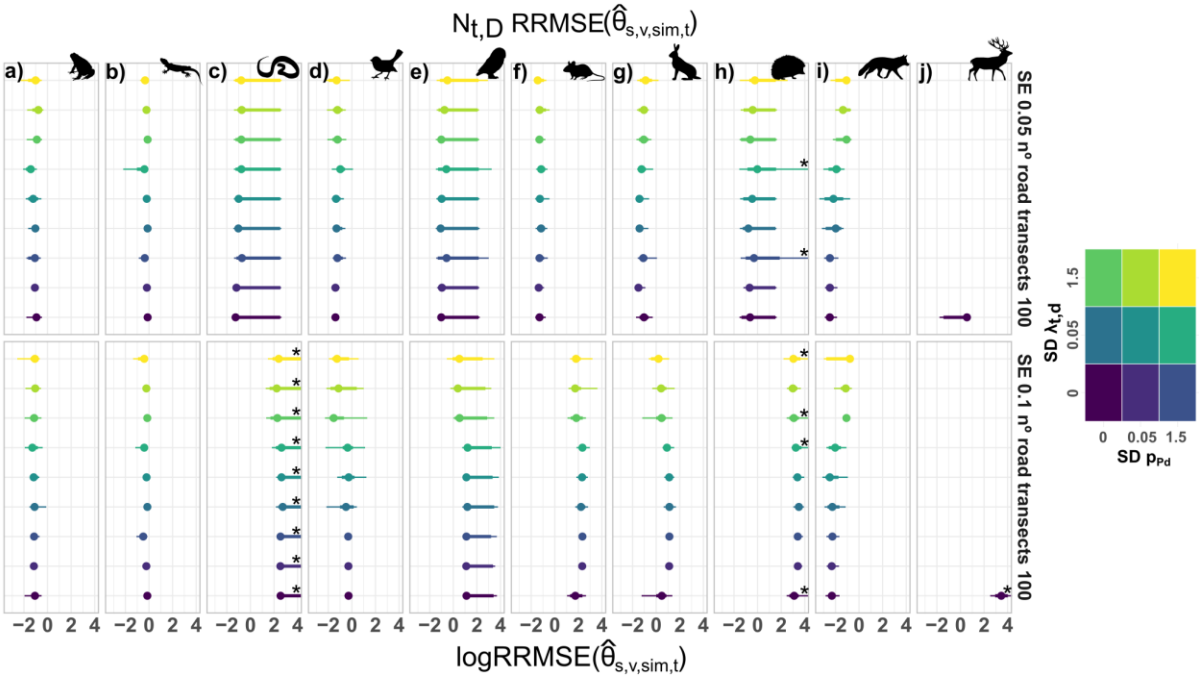
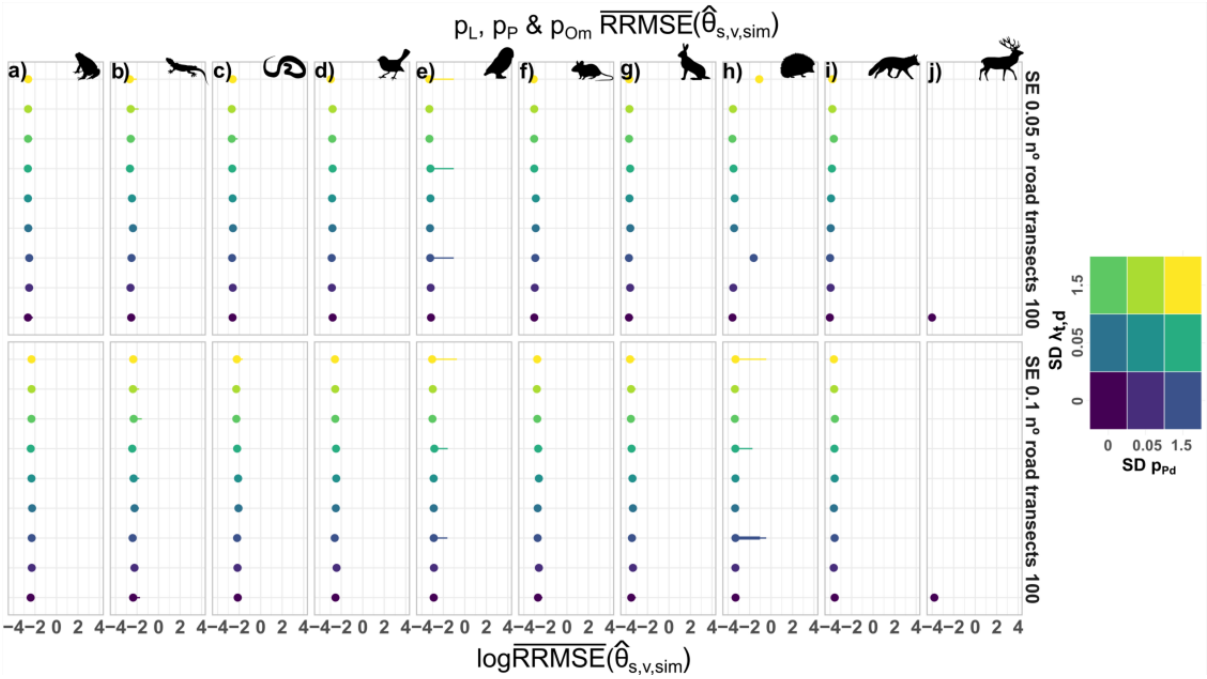


Figure 2: $N_{t,D}\ RRMSE(\hat{\theta}_{s,v,sim,t})$ values (Equation 5), where lower values indicate better model performance in recovering the simulated value. This is evaluated across s = 9 different scenario combinations of daily roadkill numbers and daily carcass persistence variability ($SD\ \lambda_{t,d}$ and SDp_{Pd}), v = 10 vertebrate groups, sim = 20 simulations, t = 12 months and D = maximum days a carcass remains on the road without disappearing. Each distribution represents $N_{t,D}\ RRMSE(\hat{\theta}_{s,v,sim,t})$ values derived from each sim and t levels described above for a) Amphibians, b) Reptiles G1, c) Reptiles G2, d) Birds/Bats G1, e) Birds G2, f) Mammals G1, g) Mammals G2, h) Mammals G3, i) Mammals G4 and j) Mammals G5. The results are shown for 2 levels of standard error (0.05 or 0.1) for the p_L and p_P prior distributions, and for 100 road transects surveyed. Dot represent the mean, bold lines for 66% intervals, and thin lines 95% intervals. An asterisk (*) in the distributions indicates values exceeding 4 that are part of the distribution. Note: Amphibians and Reptiles G1 vertebrate groups models only account for peak abundance months, excluding periods of typical absence, therefore t = 4 months were considered.



915

916 **Figure 3**: p_L, p_P & p_{Om} $\overline{RRMSE}(\hat{\theta}_{s,v,sim})$ (Equation 6), where lower values indicate
917 better model performance in recovering the simulated value. This is evaluated
918 across s = 9 different scenario combinations of daily roadkill numbers and daily
919 carcass persistence variability ($SD\ \lambda_{t,d}$ and SDp_{Pd}), v = 10 vertebrate groups, sim =
920 20 simulations and m = walking, cycling and driving survey methods. Each
921 distribution represents p_L, p_P & p_{Om} $\overline{RRMSE}(\hat{\theta}_{s,v,sim})$ values derived from each sim
922 level described above for a) Amphibians, b) Reptiles G1, c) Reptiles G2, d)
923 Birds/Bats G1, e) Birds G2, f) Mammals G1, g) Mammals G2, h) Mammals G3, i)
924 Mammals G4 and j) Mammals G5.
925 The results are shown for 2 levels of standard error (0.05 or 0.1) for the p_L and p_P
926 prior distributions, and for 100 road transects surveyed. Design: Log-spaced grid
927 with dots for means, bold lines for 66% intervals, and thin lines for 95% intervals.
928 Note: Amphibians and Reptiles G1 vertebrate groups models only account for peak
929 abundance months, excluding periods of typical absence.

930

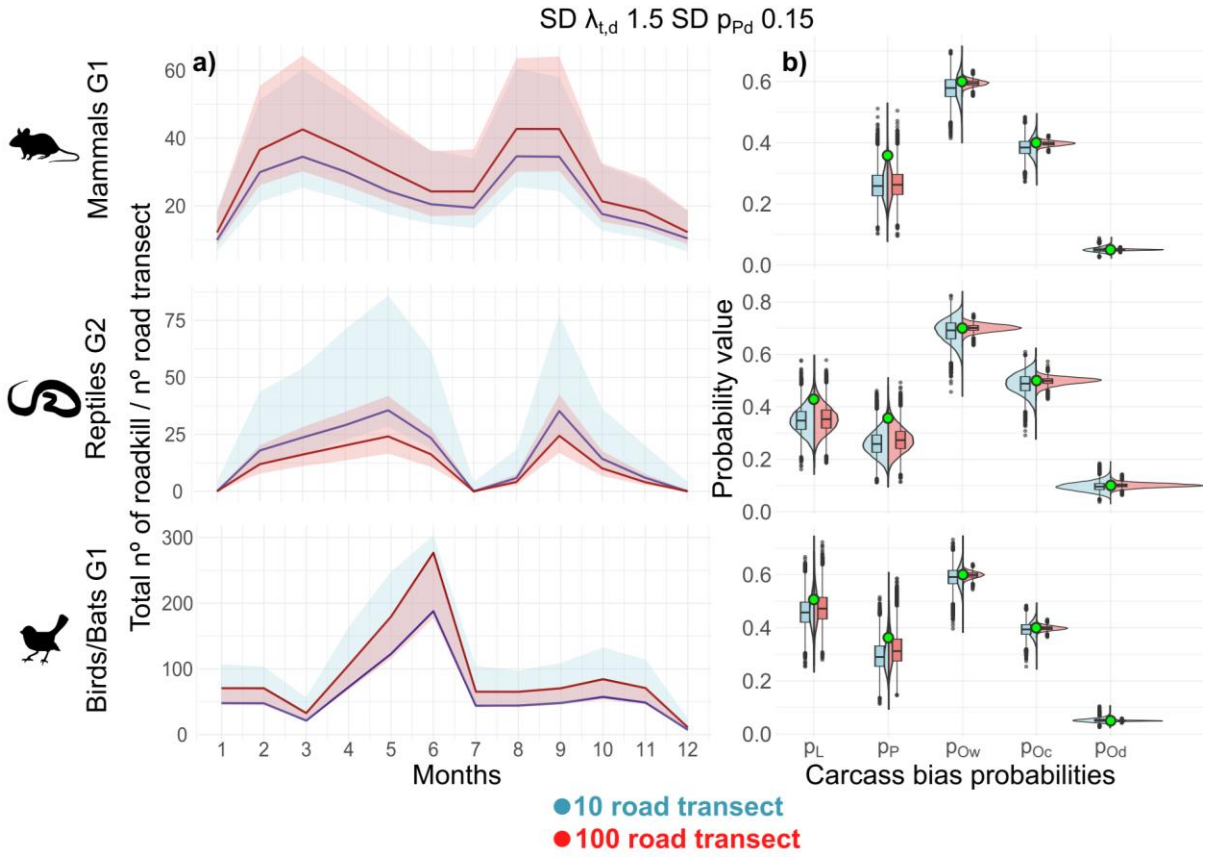


Figure 4: Overlap of total number of roadkills simulated values $N_{t,D}\ \theta_{s,v,sim,t}$ per road transect (a) and carcass location, persistence and observation per method probabilities simulated data p_L, p_P and $p_{Om}\ \theta_{s,v,sim}$ (b) between their Bayesian estimation distribution, $N_{t,D}\ \hat{\theta}_{s,v,sim,t}$ and p_L, p_P and $p_{Om}\ \hat{\theta}_{s,v,sim}$, for Mammals G1, Reptiles G2, and and Birds/Bats G1, when census data were simulated under variability scenario for daily roadkill numbers ($\lambda_{t,d}$) and daily carcass persistence probability (p_{Pd}), considering a SE = 0.05 in p_L and p_P priors. a) Lines represent the averaged $N_{t,D}\ \theta_{s,v,sim,t}$ over 20 simulations, while the shaded areas show the 95% credible interval of the average $N_{t,D}\ \hat{\theta}_{s,v,sim,t}$ over each 20 simulated census data. b) Green dots represent the p_L, p_P and $p_{Om}\ \theta_{s,v,sim}$ for m = walking (p_{Ow}), cycling (p_{Oc}) or driving (p_{Od}) survey methods, while the boxplots with violin plots show the credible interval of the pooled p_L, p_P and $p_{Om}\ \hat{\theta}_{s,v,sim}$ over 20 simulated census data.

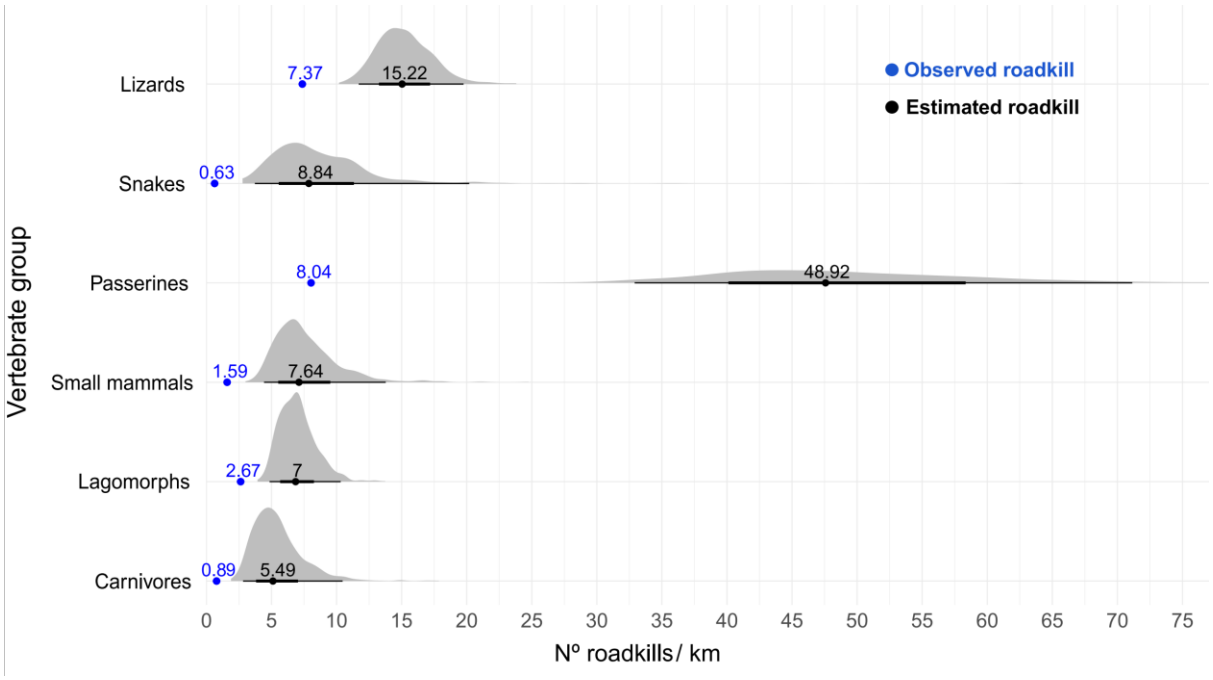


Figure 5. Observed roadkill rates per kilometer in road surveys (blue) and Bayesian posterior estimates of total roadkill rates per kilometer (black), derived from aggregating four monthly census data of the case study, for each vertebrate group. These estimated roadkill rates are limited to those that occurred within the time interval where each vertebrate group remains visible on the road without disappearing. Dots for means, bold lines for 66% credible intervals, and thin lines for 95% credible intervals.

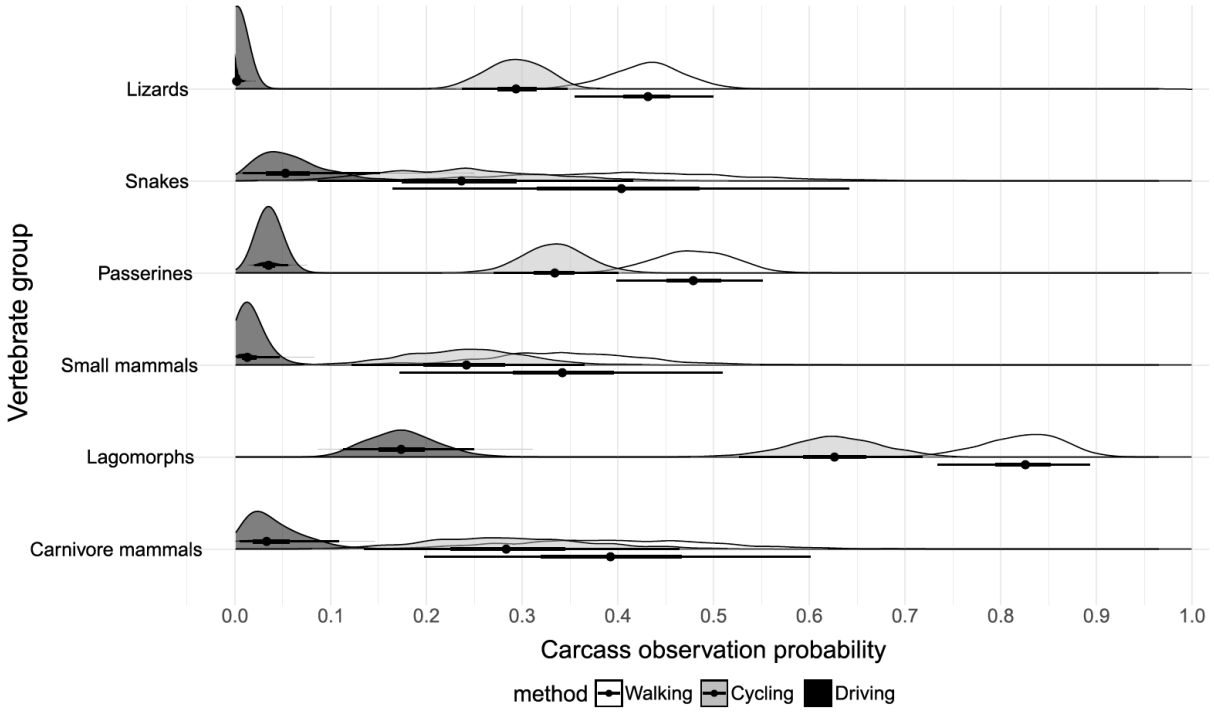


Figure 6. Bayesian posterior distribution of the carcass observation probabilities from case study, for each considered vertebrate groups. “Walking” means the estimation of carcass observation probability by walking survey method, “Cycling” by cycling survey method and “Driving” by driving survey method. Dots for means, bold lines for 66% credible intervals, and thin lines for 95% credible intervals.