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

Guillermo Gomez-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

guillermo.gomez@ebd.csic.es

[†] Shared first authorship

Running headline

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. Computer resources were provided by the Deputy General Secretariat for Informatics SGAI-CSIC (Drago HPC) and the Estación Biológica de Doñana EBD-CSIC.

Supplementary material and the scripts used for making the simulation study and case study have been deposited in a ZIP file and in Figshare (Supplementary material S12 link: https://figshare.com/s/7c6f92b973683e154b91, Appendix A link: https://figshare.com/s/e6e4d1e80b4dae5dc434, Appendix B link: https://figshare.com/s/760610fb6997e596f8c7, and Appendix C link: https://figshare.com/s/2b8aa6942ae0ff7125df). These links are also included in the ZIP file in a text document named "Supplementary_material_S16_AppendixA_AppendixB_AppendixC_links.txt". If the manuscript is accepted, all data will be archived in Figshare and all scripts will be uploaded to Zenodo.

Abstract

1. Human infrastructures are among the most impactful wildlife threats. Although estimates of animal mortality by these structures exist over a given period, they typically do not account for several detection biases (i.e., difference between recorded and true mortality). 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 processes that produce biases 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 apply our framework on a case study to estimate the total number of roadkills per km in Mediterranean ecosystems while evaluating the cross-efficiency of different sampling methods.

3. Our framework accurately recovers 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 was high and Bayesian priors were highly diffuse. In the case study, our results show notably high roadkill numbers

(e.g., estimating 48.92 per km passerines based on 8.04 observed counts), along with substantial variation across different vertebrate groups. Furthermore, our case study confirms that walking and cycling surveys outperform driving surveys in carcass observation rate and provide complementary information between them, observing partially distinct sets of species and carcass sizes.

4. Our modelling framework offers an efficient approach to estimate mortality rates for a wide range of taxa. Optimizing application requires extensive fieldwork for bias estimation and integration. We provide a checklist to help managers to assess when infrastructure-related mortality can be assessed most robustly to prioritize conservation efforts.

We have made all data and code available in a ZIP file and on Figshare (links also in ZIP)

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 lead to multiple nested sources of bias, i.e., mismatch among recorded roadkills and true road mortality due to lack of detection, when inferring infrastructure-induced mortality from carcass surveys (Barrientos et al., 2018; Román et al., 2024). The first of these biases is carcass location bias and concerns missing 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 bias affecting standard mortality surveys along infrastructures is carcass persistence bias, which occurs when carcasses disappear from the survey area over time before a given survey (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), as well as to scavenger activity (DeVault et al., 2017; Dhiab et al., 2023). On 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 bias 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

We first describe 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) (Fig. 1)). We then evaluate the model's performance through a simulation study, testing different roadkill scenarios across different vertebrate groups and conducting a prior sensitivity analysis. Finally, we apply our model to data from a field case study to estimate the total number of roadkills based on empirical census datasets.

173 *2.2 Modelling framework*

174 We introduce a hierarchical latent-state model to estimate the total number of
175 roadkills, explicitly accounting for the three nested levels of bias: carcass location,
176 persistence, and observation. The model structure is based on the widely used N-
177 mixture models, which estimate abundances from count data while accounting for
178 imperfect detection (Hostetter et al., 2019; Kery & Royle, 2021; Royle, 2004).

179 We model the total number of roadkills $N_{i,t,D}$ across *i* = 1…*I* road transects, within *t* =
180 1…*T* survey periods (with months used as periods in our model, as more frequent
181 surveys are rarely performed), considering a retrospective carcass accumulation
182 period of *D* days (representing the maximum number of days a carcass remains on
183 the survey area before disappearing). We define $N_{i,t,D}$ as a random Poisson variable
184 sampled from $\lambda_{t,D}$ = mean number of roadkills for the period *t* over *D* days:

185 $$N_{i,t,D} \sim Poisson\,(\lambda_{t,D})$$ eqn. 1

186 As $\lambda_{t,d}$, and consequently $N_{i,t,D}$, can vary across months *t*, our model accounts for
187 seasonal changes in roadkill numbers throughout the year. Note, however, that we
188 do not model variability in $\lambda_{t,d}$ at the transect level *i*.

189 We assume that $N_{i,t,D} = \sum_{d=1}^{D} N_{i,t,d}$, where each daily total number of roadkills $N_{i,t,d}$
190 can fluctuate across the days within period *D*, following the daily $\lambda_{t,d}$ in month *t*.
191 However, our framework assumes that $N_{i,t,d}$ and $\lambda_{t,d}$ cannot be modelled directly and
192 instead need to be estimated over the maximum persistence time *D*, as conducting
193 daily road monitoring is too resource-demanding to be feasible.

We then define $N2_{i,t,D}$ as the subset 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 subset 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., a Cox-hazard model as in Santos et al. 2011), then $p_P = \int_{d=1}^{D} S(d)d(d)$, being the average persistence probability from *d* = 1 to *D* (for details, see Supplementary Material S1). For example, if daily $N2_{i,t,d}$ values are known and, for illustration, we set *D* = 3 days (with *d1*, *d2* and *d3* denoting days 1, 2 and 3 since roadkills occur), our framework can in theory model: $N3_{i,t,d1} \sim Binomial\left(p_{Pd1}p_{Pd2}p_{Pd3},\ N2_{i,t,d1}\right)$, $N3_{i,t,d2} \sim Binomial\left(p_{Pd2}p_{Pd3},\ N2_{i,t,d2}\right)$, $N3_{i,t,d3} \sim Binomial\left(p_{Pd3},\ N2_{i,t,d3}\right)$, where p_{Pd1}, p_{Pd2} and p_{Pd3} are the daily carcass persistence probability on first, second and third days since roadkills occur, respectively.

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 implemented our Bayesian hierarchical latent-state modelling framework to run 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 using the method of moments relative to the standard parameterization of the beta distribution. We adopted a non-informative prior for p_{Om} ($p_{Om} \sim \text{Uniform}(0,1)$), and weakly informative priors for $\lambda_{t,D}$ setting the upper limit sufficiently wide to accommodate the expected number of roadkills ($\lambda_{t,D} \sim \text{Uniform}(0,300)$, $\sim \text{Uniform}(0,600)$ or $\sim \text{Uniform}(0,800)$, as detailed in Supplementary Material S2, S3 and S6), this specific upper limit was selected to ensure values remained biologically reasonable, while increasing the computational stability and convergence of the Bayesian models.

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 v. 1.6.2 (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 evaluated the effective sample size (ESS) and visually inspected the traceplots of the posterior distributions among the different MCMC chains to check convergence

or mixing issues (see Appendix C). 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. Furthermore, given this structural reliance on prior information, we conducted a specific prior sensitivity analysis to evaluate the robustness of our estimates under different prior specifications (see Section 2.4.3).

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 groups were delineated based on species traits (body size and mobility); consequently, these groups determined the characteristics of observed roadkill numbers and seasonal 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 groups (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 2.2 and Supplementary Material S1, see also Supplementary Material S2, S3, S6, S8 and S9 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 group, 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)

310

311 *2.4 Simulation study*

312 We used a simulation study to stress-test under which scenarios our modelling
313 framework accurately estimated the total number of roadkills $N_{t,D}$ and recovered p_L,
314 p_P and p_{Om} as latent states. The simulation study generated census data $C_{i,j,t,D,m}$
315 based on different biological and observation processes described below.

316 *2.4.1 Principles of census data generation*

317 To generate different census datasets, we followed the nested levels of data as
318 described in the modelling framework section 2.2 (N → $N2$ → $N3$ → C). The
319 progression through these levels was carried out considering the values of p_L, p_P
320 and p_{Om} specific to each vertebrate group (Table 1), in order to create a range of
321 biologically realistic data.
322
323 We first sampled $N_{i,t,d}$ for each road transect (*i*), month (*t*), and day (*d* = 1-*D*, where
324 *D* was the maximum carcass persistence period for a given vertebrate group) as a
325 random Poisson variable based on their mean total number of roadkills $\lambda_{t,d}$ (using
326 equation 1). We used expert knowledge to assign variation in month t dimension
327 based on data collected in 2021 and 2022 in southern Spain, which
328 incorporated known seasonal trends for each vertebrate group (Supplementary
329 Material S5).
330 As an example, considering *D* = 3 days, $\lambda_{t,d1}$, $\lambda_{t,d2}$, and $\lambda_{t,d3}$ would be generated
331 (i.e., the mean number of roadkills occurring three days, two days, or the day before
332 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 $N_{i,t,d1}$, $N_{i,t,d2}$ and $N_{i,t,d3}$, 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 on the 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 $N_{i,t,D} = N_{i,t,d}$.

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 (equation 2).

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 (equation 3). For each vertebrate group with their respective p_{Pd} value (see Supplementary material S1), 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 (equation 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 observation 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 Simulation scenarios of variable data

We generated different simulated census dataset for each of the vertebrate groups considered here (Table 1) using scenarios that introduced different levels of parameters variability (Table 2) to assess when our model can recover simulated parameters and when there is a risk of under- or overestimating total roadkills.

First, we introduced variability in the daily number of roadkills and persistence rates across simulations of *i* = 10 or 100 road transects. Specifically, we introduced variability in the mean daily number of roadkills $\lambda_{t,d}$ by multiplying it by a value sampled from a truncated random Normal distribution (mean = 1; SD = 0, 0.5, or 1.5), and similarly applied variability to the daily persistence probability p_{Pd} by sampling from a truncated random Normal distribution (mean = p_{Pd}; SD = 0, 0.05, or 0.15). We didn't add variability to p_L, as we assume p_L doesn't vary among days. We then analysed these datasets using p_L and p_P priors fitted with both a strong informative prior distribution (Standard Error (SE) = 0.05) and a weaker prior distribution (SE = 0.1) to test how uncertainty in the priors affects the recovery of

parameters in the presence of variability, as p_L and p_P values are prior-driven. We couldn't simulate $\lambda_{t,d}$ and p_{Pd} variability for Mammals G5 as we simulated a single $N_{i,t,d}$ value for the entire month and this group is not affected by carcass persistence bias, respectively.

We also simulated vertebrate groups, i.e., "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 are 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.

We simulated 20 datasets for each vertebrate group and scenario combination, resulting in 720 simulated datasets per vertebrate group (dataset simulation code is detailed in Supplementary Material S2, S3, S4 and S6).

2.4.3 Prior sensitivity

In our Bayesian models, information on p_L and p_P comes entirely from prior knowledge, as the count data alone ($C_{i,j,t,D,m}$) does not contain sufficient information to independently disentangle these intermediate latent processes without external information. Therefore, we conducted a sensitivity analysis to stress-test whether key

simulated parameters can be recovered and p_L and p_P are identifiable by using informative vs. uninformative prior knowledge. We again simulated different datasets for each vertebrate group (as described in 2.4.2), but we fixed the following parameters: i = 10 road transects, $\lambda_{t,d}$ and p_{Om} values, and SD = 0 for both $\lambda_{t,d}$ and p_{Pd}. We then simulated datasets using pairwise combinations of p_L and p_{Pd} as shown in Supplementary Material S7. For each simulated dataset, we then used three prior specifications in Bayesian models: accurate informative priors centered on the p_L and p_P values associated with each simulated data set (Supplementary Material S7); inaccurate informative priors (0.7 if p_L or p_P < 0.5 and 0.3 if p_L or p_P > 0.5); and finally, uninformative priors using a uniform distribution from 0 to 1. Prior precision was set to SE = 0.05 (see 2.4.2).

We generated 27 scenario combinations per vertebrate group affected by carcass location and persistence bias, leading to 540 simulated datasets per vertebrate group, while 9 scenario combinations were generated for vertebrate groups only affected by carcass location or persistence bias, i.e., 180 simulated datasets (dataset simulation code is detailed in Supplementary Material S8, S9 and S10).

All datasets were generated and analysed in R v. 4.2.2 (R Core Team, 2022) .

2.4.4 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 $\theta_{s,v,sim,t}$. Here, the subscripts denoted *s* = 1…*S* simulation scenario, *v* = 1…*V* vertebrate groups, *sim* = 1…*Sim* specific simulation iteration and *t* = 1…*T* months. 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 model performance across simulation scenarios and vertebrate groups. $RRMSE$ standardizes the error relative to the magnitude of the true parameter, allowing for comparisons between parameters with different scales (e.g., total number of roadkills vs. carcass bias probabilities). It was calculated 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. Intuitively, $RRMSE$ represents the size of the error as a proportion of the true simulated value. An $RRMSE$ of 0 indicates perfect accuracy, while a value of 1 implies that the magnitude of the error is equal to the true simulated value itself. Consequently, $RRMSE$ values can exceed 1 (or >0 on a logarithmic scale used for graphical clarity) when the estimation error is larger than the parameter being estimated, indicating high uncertainty. In the case of $N_{t,D}$ $\theta_{s,v,sim,t}$ values, we added 1 to all values as $RRMSE(\hat{\theta}_{s,v,sim,t})$ could not be calculated when $N_{t,D}$ = 0 (Supplementary Material S11).

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; Supplementary Material S12):

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

To ensure that the true known simulated values were recovered, we assessed whether the average true values $N_{t,D}\ \theta_{s,v,sim,t}$ lay within the 95% credible interval of Bayesian posterior estimates of average $N_{t,D}\ \hat{\theta}_{s,v,sim,t}$. In contrast for carcass bias probabilities, we assessed whether the true values p_L, p_P, $p_{Om}\ \theta_{s,v,sim}$ lay within the full range of their pooled respective posterior distributions p_L, p_P, and $p_{Om}\ \hat{\theta}_{s,v,sim}$.

In the prior sensitivity analysis, we also evaluated across the different simulation scenarios whether the true values $\theta_{s,v,sim,t}$ correctly overlapped with the posterior distributions $\hat{\theta}_{s,v,sim,t}$. Specifically, we assessed the accuracy of estimates for the mean number of roadkills for period t over D days $\lambda_{t,D}$ the total number of roadkills $N_{t,D}$, and the carcass observation probabilities per survey method p_{Om}. Additionally, we verified the successful recovery of the p_L and p_P values integrated into the model as priors.

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 S13).

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. However, unlike the Simulation study (Section 2.4.2) where broad theoretical categories were used (e.g., “Birds and Bats G1”), here we adapted the group nomenclature to strictly reflect the specific taxa actually observed during fieldwork (e.g., “Passerines”).

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 A; B; C).

3.1 Simulation scenarios of variable data

Our 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 of variability in parameters (Fig. 2; see also Supplementary Material S14 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 mean number of roadkills (SD $\lambda_{t,d}$) was also high (Fig. 2). The number of road transects simulated (10 or 100 transects) had minimal impact on $RRMSE(\hat{\theta}_{s,v,sim,t})$, except for Reptiles G1, where the $RRMSE(\hat{\theta}_{s,v,sim,t})$ decreased notably in the case of full-year datasets, including extended periods of low number of roadkills (Supplementary Material S14). For Amphibians and Reptiles G1, $RRMSE(\hat{\theta}_{s,v,sim,t})$ were lower when limiting the analysis to months when animals are active (abundance peak), compared to when extended periods of low number of roadkills were included in the datasets (Supplementary Material S14).

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, i.e., was withing the 95% credible interval of $N_{t,D}\,\hat{\theta}_{s,v,sim,t}$ (Fig. 3a; see Supplementary Material S15 for more detailed plots). However, Reptiles G1 was an

exception: when considering full-year datasets with 10 road transects, estimates for this vertebrate group were overestimated in all scenarios, whereas with 100 road transects they were well recovered. Accounting only for the abundance peak led to better fits overall, even though estimates for this vertebrate group were always underestimated. (Supplementary Material S15). On the other hand, when p_L and p_P prior SE was high, the 95% credible interval overlap widened, leading to overestimations across most vertebrate groups (Supplementary Material S15). The only exception were Amphibians and Reptiles G1 accounting only for the abundance peak, where estimates were typically underestimated, resulting in increased uncertainty but reduced underestimation in most scenarios (Supplementary Material S15).

$\overline{RRMSE}(\hat{\theta}_{s,v,sim})$ 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 G1, Birds G2 and Mammals G3 under high variability in daily carcass persistence probabilities (SD p_{Pd}) (Fig. 4). 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. 3b; see Supplementary Material S16 for more detailed plots). The only exception was the Reptiles G1 full-year dataset with 10 road transects, where p_{Om} was not recovered although recovery was successful with 100 transects (Supplementary Material S16).

3.2 Prior sensitivity analyses

Regarding our Prior Sensitivity Analysis, we observed that when priors are accurate, the posterior distributions correctly overlapped with the true values, accurately estimating the mean number of roadkills for the period *t* over *D* days $\lambda_{t,D}$, the total number of roadkills for the period *t* over *D* days $N_{t,D}$, and the carcass observation probabilities per survey method p_{Om}, in addition to successfully recovering the p_L and p_P values integrated in the model as priors (Supplementary material S17). In the case of inaccurate and uninformative prior scenarios, the estimation of p_{Om} values remained robust, although the remaining parameters were affected (Supplementary material S17). Specifically, in the case of inaccurate priors, the p_L and p_P values were never recovered (although the model occasionally recovered estimates of $N_{t,D}$ and $\lambda_{t,D}$, but with lower accuracy than in accurate prior scenarios (Supplementary material S17). On the other hand, uninformative priors produced extremely wide posterior distributions of p_L and p_P, which resulted in an extreme overestimation of $N_{t,D}$ and $\lambda_{t,D}$. Furthermore, uninformative priors often resulted in model run errors, as impossible values during the Bayesian estimation process were generated (Supplementary material S17).

3.3 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 taxa: 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 taxa. 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 S18).

4. Discussion

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

In the present study, we 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. Consequently, we were able to 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 thanks 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

4.2.1 Impact of variability in daily parameters

Our simulation scenarios of variable data indicate that, under low variability in daily number of roadkills and daily carcass persistence probabilities, the N-mixture model

provides reliable estimates of total number of roadkills. However, when we simulated scenarios with high variability in parameters, total number of roadkills were both over- and underestimated, which was reflected in increased Relative Root Mean Square Error ($RRMSE$) values. These reliability results are consistent with findings by Dennis et al. (2015); Duarte et al. (2018); Link et al. (2018) and Monroe et al. (2019), who emphasize that N-mixture models are highly sensitive to excessive variation in model parameters. Consequently, estimates of total number of roadkill should be interpreted with caution in datasets characterized by high variability in parameters (see Table 3). However, despite these inaccuracies in estimating total roadkill numbers, the models consistently yielded reliable estimates of relative numbers across months. This finding supports the "can't lose" proposition described by Kéry & Royle (2021): even when violations of parametric assumptions compromise the precision of the absolute population size, the N-mixture framework remains a robust tool for inferring relative dynamics. As noted by Barker et al. (2018); Knape & Korner-Nievergelt (2015) and Martijn et al. (2023), such models effectively characterize relative abundance even when data are sparse or variable.

4.2.2 *Impact of priors information for carcass location and persistence probability*

In most current road survey designs, data on carcass location (p_L) and persistence (p_P) probability are not explicitly collected, and these two parameters are not identifiable from count data alone. Thus, the estimation of these specific biases becomes prior-driven; that is, the posterior distributions are dominated by the prior assumptions rather than by the data itself (Banner et al., 2020; Northrup & Gerber, 2018). Our simulation scenarios of variable data and prior sensitivity show that weak informative, uninformative or inaccurate prior distributions (i.e., wide distributions

with high standard errors, uniform distributions or informative priors intentionally biased away from the true values) propagated uncertainty directly to estimates total number of roadkill or resulted in low recovery and identifiability of simulated carcass location and persistence probability priors by the Bayesian model, an issue highlighted by (Fidino, 2021). This behaviour is consistent with literature on Bayesian mixture models, which warns that inference can become unstable when data are sparse and priors are uninformative (Depaoli, 2013; Depaoli et al., 2017).

Our results demonstrate that to obtain reliable estimates of the total number of roadkills, future roadkill studies cannot rely on vague priors for carcass bias probabilities; they require informative priors derived from independent empirical data; or, ideally, would incorporate independent data to facilitate the estimation of posterior distributions of these parameters. Thus, research efforts must prioritize collecting auxiliary data to quantify carcass location probability (Román et al., 2024) and persistence rates (Ruiz-Capillas et al., 2015; Santos et al., 2018), as these independent constraints are necessary to anchor the model parameters.

4.2.3 Differences among vertebrate groups

When simulated data are highly variable or priors were uninformative, the total number of roadkills can be overestimated, depending on the vertebrate group. The fact that lizards were overestimated when considering full-year datasets with 10 road transects and informative priors is likely due to extremely low persistence and observation probabilities in this group, which resulted in a zero-inflated simulated dataset for analysis. In such cases, a higher sampling efforts (Guillera-Arroita et al., 2010; MacKenzie et al., 2002) and employing a zero-inflated Poisson version of the

671 N-mixture model can yield more accurate results (Joseph et al., 2009; Wenger &
672 Freeman, 2008).
673
674 *4.2.4 Reliability checklist to assess robustness of Bayesian framework in future*
675 *applications*
676 Based on the model's performance across our simulation scenarios, we provide a
677 reliability checklist to guide researchers in ensuring the reliability of total mortality
678 estimates in infrastructure mortality surveys (Table 3). The model generally provides
679 robust estimates of all target parameters when accurate priors are available for the
680 carcass location and carcass persistence probabilities. However, as outlined in the
681 checklist (Table 3), the precision of these estimates may be compromised in
682 scenarios where prior uncertainty or data complexity affect the reliability and
683 performance of the model. In such instances, while carcass observation probability
684 estimates tend to remain robust, the estimates for the total number of roadkills
685 should be approached with caution.
686
687 Future applications must cautiously evaluate and document prior knowledge on
688 carcass location and persistence probabilities and ideally incorporate, auxiliary field
689 experiments to estimate posterior distributions. In addition, we strongly recommend
690 restricting the analysis to the biologically active season for taxa with marked
691 seasonal dynamics (e.g., amphibians, reptiles). Lastly, results must be interpreted
692 with caution when applying these models to datasets where daily roadkill numbers or
693 persistence rates fluctuate drastically. For such highly variable data, estimates
694 should be interpreted as robust indices of relative abundance (Kéry & Royle, 2021).
695

Crucially, beyond adhering to this checklist, we strongly recommend that researchers using our Bayesian models to estimate infrastructure mortality perform their own simulations to validate the model's suitability for their specific study system and data. By utilizing the simulation R scripts provided in this study (Supplementary Material S2, S3, S4 and S6), users can easily adapt our specifications to generate simulated datasets that mimic their specific study conditions (e.g., number of survey transects; carcass location, persistence and observation per method probabilities values).

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

721 underfoot, the elevated perspective provided by cycling allows for a broader field of
722 view, making it easier to observe carcasses on the roadside.
723
724 These results highlight that using the driving method in surveys not only reduces the
725 proportion of carcasses observed on the road but can also lead to an overestimation
726 of the total number of collisions. In N-mixture models, lower observation probabilities
727 result in larger extrapolations in the estimated values. Since observation probabilities
728 while driving are extremely low, the estimated total number of roadkills ultimately
729 would be much higher than the real one (Dennis et al., 2015; Hostetter et al., 2019).
730
731 Regarding our case study survey methodology, one important consideration is that,
732 typically, roadkill studies alternate the direction of search and the side of the road
733 randomly in order to cover the area as thoroughly as possible along the
734 infrastructure (D'Amico et al., 2015). However, in our case, as our study was an
735 initial phase of a citizen science project with volunteers, we had to employ a simple
736 and easy sampling method, conducting surveys on only one side of the road and
737 always in the same direction. Although we recognize that this may decrease the
738 carcass observation probability, it would be interesting to investigate in the future
739 whether randomizing the direction and side of the road would actually reduce
740 carcass observation bias.
741
742 *4.4 Limitations and future perspectives*
743 A limitation of our methodology is that it requires extensive knowledge of carcass
744 location, persistence, and observation biases specific to each infrastructure,
745 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 number of roadkills is that they are limited by the maximum number of days a carcass from a specific vertebrate group remains on the road before disappearing (*D*-day period). This means that to estimate the number of roadkill for periods larger than the D-day period (e.g., one month or a specific season), we currently simply extrapolate our estimates for the D-day period over a larger time window (e.g., 30 days / *D*-day period). Thus, for vertebrate groups with shorter persistence times (such as amphibians and lizards), the extrapolation gap required to cover the unobserved temporal window is significantly larger than for groups 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.

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, such as roads (Barrientos et al., 2021), power lines (Biasotto & Kindel, 2018; D'Amico et al., 2019) and 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

Abra, F. D., Huijser, M. P., Pereira, C. S., & Ferraz, K. M. P. M. B. (2018). How reliable are your data? Verifying species identification of road-killed mammals recorded by road maintenance personnel in São Paulo State, Brazil. *Biological Conservation*, *225*, 42–52. https://doi.org/10.1016/j.biocon.2018.06.019

Ascensão, F., D'Amico, M., & Barrientos, R. (2022). No Planet for Apes? Assessing Global Priority Areas and Species Affected by Linear Infrastructures. *International Journal of Primatology*, *43*(1), 57–73. https://doi.org/10.1007/s10764-021-00207-5

Banner, K. M., Irvine, K. M., & Rodhouse, T. J. (2020). The use of Bayesian priors in Ecology: The good, the bad and the not great. *Methods in Ecology and Evolution*, *11*(8), 882–889. https://doi.org/10.1111/2041-210X.13407

Barker, R. J., Schofield, M. R., Link, W. A., & Sauer, J. R. (2018). On the reliability of N-mixture models for count data. *Biometrics*, *74*(1), 369–377. https://doi.org/10.1111/biom.12734

Barrientos, R., Ascensão, F., D'Amico, M., Grilo, C., & Pereira, H. M. (2021). The lost road: Do transportation networks imperil wildlife population persistence? *Perspectives in Ecology and Conservation*, *19*(4), 411–416. https://doi.org/10.1016/j.pecon.2021.07.004

Barrientos, R., Martins, R. C., Ascensão, F., D'Amico, M., Moreira, F., & Borda-de-Água, L. (2018). A review of searcher efficiency and carcass persistence in infrastructure-driven mortality assessment studies. *Biological Conservation*, *222*, 146–153. https://doi.org/10.1016/j.biocon.2018.04.014

796 Bech, N., Beltran, S., Boissier, J., Allienne, J.-F., Resseguier, J., & Novoa, C. (2012). Bird
797 mortality related to collisions with ski-lift cables: Do we estimate just the tip of the
798 iceberg? *Animal Biodiversity and Conservation*, *35*(1), 95–98.
799 Bernardino, J., Bevanger, K., Barrientos, R., Dwyer, J. F., Marques, A. T., Martins, R. C.,
800 Shaw, J. M., Silva, J. P., & Moreira, F. (2018). Bird collisions with power lines: State
801 of the art and priority areas for research. *Biological Conservation*, *222*, 1–13.
802 https://doi.org/10.1016/j.biocon.2018.02.029
803 Bernardino, J., Bispo, R., Martins, R. C., Santos, S., & Moreira, F. (2020). Response of
804 vertebrate scavengers to power line and road rights-of-way and its implications for
805 bird fatality estimates. *Scientific Reports*, *10*(1), 15014.
806 https://doi.org/10.1038/s41598-020-72059-7
807 Biasotto, L. D., & Kindel, A. (2018). Power lines and impacts on biodiversity: A systematic
808 review. *Environmental Impact Assessment Review*, *71*, 110–119.
809 https://doi.org/10.1016/j.eiar.2018.04.010
810 Borner, L., Duriez, O., Besnard, A., Robert, A., Carrere, V., & Jiguet, F. (2017). Bird collision
811 with power lines: Estimating carcass persistence and detection associated with
812 ground search surveys. *Ecosphere*, *8*(11), e01966. https://doi.org/10.1002/ecs2.1966
813 Brooks, S. P., & Gelman, A. (1998). General Methods for Monitoring Convergence of
814 Iterative Simulations. *Journal of Computational and Graphical Statistics*, *7*(4), 434–
815 455. https://doi.org/10.1080/10618600.1998.10474787
816 D'Amico, M., Ascensão, F., Fabrizio, M., Barrientos, R., & Gortázar, C. (2018). Twenty years
817 of road ecology: A topical collection looking forward for new perspectives. *European*
818 *Journal of Wildlife Research*, *64*, 1–2.
819 D'Amico, M., Catry, I., Martins, R. C., Ascensão, F., Barrientos, R., & Moreira, F. (2018). Bird
820 on the wire: Landscape planning considering costs and benefits for bird populations
821 coexisting with power lines. *Ambio*, *47*(6), 650–656. https://doi.org/10.1007/s13280-
822 018-1025-z

D'Amico, M., Martins, R. C., Álvarez-Martínez, J. M., Porto, M., Barrientos, R., & Moreira, F. (2019). Bird collisions with power lines: Prioritizing species and areas by estimating potential population-level impacts. *Diversity and Distributions*, *25*(6), 975–982. https://doi.org/10.1111/ddi.12903

D'Amico, M., Román, J., De los Reyes, L., & Revilla, E. (2015). Vertebrate road-kill patterns in Mediterranean habitats: Who, when and where. *Biological Conservation*, *191*, 234–242.

Delgado, J. D., Durán Humia, J., Rodríguez Pereiras, A., Rosal, A., del Valle Palenzuela, M., 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

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

Duarte, A., Adams, M. J., & Peterson, J. T. (2018). Fitting N-mixture models to count data with unmodeled heterogeneity: Bias, diagnostics, and alternative approaches. *Ecological Modelling*, *374*, 51–59. https://doi.org/10.1016/j.ecolmodel.2018.02.007

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

Fidino, M. (2021, August 25). *A gentle introduction to an integrated occupancy model that combines presence-only and detection/non-detection data, and how to fit it in `JAGS`*. https://masonfidino.com/bayesian_integrated_model/

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

Guillera-Arroita, G., Ridout, M. S., & Morgan, B. J. T. (2010). Design of occupancy studies with imperfect detection. *Methods in Ecology and Evolution*, *1*(2), 131–139. https://doi.org/10.1111/j.2041-210X.2010.00017.x

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

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. (2021). *Applied Hierarchical Modeling in Ecology: Analysis of Distribution, Abundance and Species Richness in R and BUGS. Academic Press.*

Kéry, M., & Royle, J. A. (2021). *Applied Hierarchical Modeling in Ecology: Analysis of Distribution, Abundance and Species Richness in R and BUGS: Volume 2: Dynamic and Advanced Models*. 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

Knape, J., & Korner-Nievergelt, F. (2015). Estimates from non-replicated population surveys rely on critical assumptions. *Methods in Ecology and Evolution*, *6*(3), 298–306. https://doi.org/10.1111/2041-210X.12329

Link, W. A., Schofield, M. R., Barker, R. J., & Sauer, J. R. (2018). On the robustness of N-mixture models. *Ecology*, *99*(7), 1547–1551. https://doi.org/10.1002/ecy.2362

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Andrew Royle, J., & Langtimm, C. A. (2002). Estimating Site Occupancy Rates When Detection Probabilities Are Less Than One. *Ecology*, *83*(8), 2248–2255. https://doi.org/10.1890/0012-9658(2002)083%255B2248:ESORWD%255D2.0.CO;2

Martijn, B., Jim, C., Natalie, B., & Thomas, N. (2023). Simulation-based assessment of the performance of hierarchical abundance estimators for camera trap surveys of

905 unmarked species. *Scientific Reports*, *13*(1), 16169. https://doi.org/10.1038/s41598-
906 023-43184-w
907 Meijer, J. R., Huijbregts, M. A. J., Schotten, K. C. G. J., & Schipper, A. M. (2018). Global
908 patterns of current and future road infrastructure. *Environmental Research Letters*,
909 *13*(6), 064006. https://doi.org/10.1088/1748-9326/aabd42
910 Monroe, A. P., Wann, G. T., Aldridge, C. L., & Coates, P. S. (2019). The importance of
911 simulation assumptions when evaluating detectability in population models.
912 *Ecosphere*, *10*(7), e02791. https://doi.org/10.1002/ecs2.2791
913 Nazir, M. S., Ali, N., Bilal, M., & Iqbal, H. M. N. (2020). Potential environmental impacts of
914 wind energy development: A global perspective. *Current Opinion in Environmental*
915 *Science & Health*, *13*, 85–90. https://doi.org/10.1016/j.coesh.2020.01.002
916 Northrup, J. M., & Gerber, B. D. (2018). A comment on priors for Bayesian occupancy
917 models. *PLOS ONE*, *13*(2), e0192819. https://doi.org/10.1371/journal.pone.0192819
918 Ogletree, K., & Mead, A. (2020). What Roadkills Did We Miss in a Driving Survey? A
919 Comparison of Driving and Walking Surveys in Baldwin County, Georgia. *Georgia*
920 *Journal of Science*, *78*(2). https://digitalcommons.gaacademy.org/gjs/vol78/iss2/8
921 Pearce-Higgins, J. W., Stephen, L., Douse, A., & Langston, R. H. W. (2012). Greater
922 impacts of wind farms on bird populations during construction than subsequent
923 operation: Results of a multi-site and multi-species analysis. *Journal of Applied*
924 *Ecology*, *49*(2), 386–394. https://doi.org/10.1111/j.1365-2664.2012.02110.x
925 Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using
926 Gibbs sampling. *Working Papers*.
927 R Core Team. (2022). *R: A language and environment for statistical computing.* [R
928 Foundation for Statistical Computing].
929 Ravache, A., Barré, K., Normand, B., Goislot, C., Besnard, A., & Kerbiriou, C. (2024).
930 Monitoring carcass persistence in windfarms: Recommendations for estimating
931 mortality. *Biological Conservation*, *292*, 110509.
932 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

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., 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

959 Tabassum-Abbasi, Premalatha, M., Abbasi, T., & Abbasi, S. A. (2014). Wind energy:
960 Increasing deployment, rising environmental concerns. *Renewable and Sustainable*
961 *Energy Reviews*, *31*, 270–288. https://doi.org/10.1016/j.rser.2013.11.019
962 Teixeira, F. Z., Coelho, A. V. P., Esperandio, I. B., & Kindel, A. (2013). Vertebrate road
963 mortality estimates: Effects of sampling methods and carcass removal. *Biological*
964 *Conservation*, *157*, 317–323. https://doi.org/10.1016/j.biocon.2012.09.006
965 Wenger, S. J., & Freeman, M. C. (2008). Estimating Species Occurrence, Abundance, and
966 Detection Probability Using Zero-Inflated Distributions. *Ecology*, *89*(10), 2953–2959.
967 https://doi.org/10.1890/07-1127.1
968 Winton, S. A., Taylor, R., Bishop, C. A., & Larsen, K. W. (2018). Estimating actual versus
969 detected road mortality rates for a northern viper. *Global Ecology and Conservation*,
970 *16*, e00476. https://doi.org/10.1016/j.gecco.2018.e00476
971

972 Figures and tables

973

974 Table 1. Descriptive characteristics of the different vertebrate groups used to
975 simulate roadkill numbers, including examples of species, their features of observed
976 roadkill numbers and their seasonal variation, maximum days their carcass remains
977 on the road without disappearing (*D*), probability of their carcass being located on
978 the road (p_L), average probability across *D* of their carcass persisting on the road
979 (p_P) and carcass observation probability (p_{Om}) by walking (p_{Ow}), cycling (p_{Oc}) and
980 driving (p_{Odr}) 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	Frequently	High	2	1	0.36	0.5	0.3	0.02

	amphibians such as *Bufo spinosus* or *Epidalea calamita*	observed							
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*	Frequently observed	Low	4	1	0.35	0.8	0.6	0.2

	granatensis								
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

981

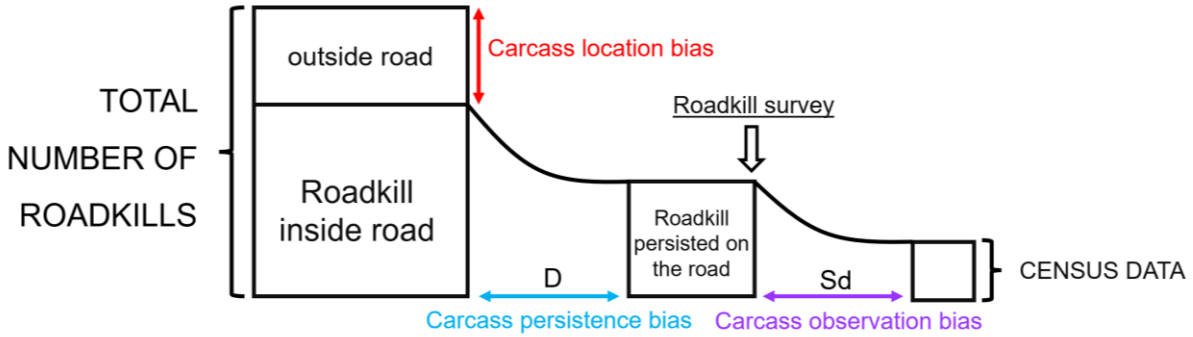
982 Table 2. Simulation scenarios to generate roadkill census data, including levels of
983 variation and justification for the scenario choice. $\lambda_{t,d}$= daily mean number of roadkills
984 in month t for each specific day *d* across *D* (maximum persistence), p_{Pd} = daily
985 carcass persistence 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, 2021). 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

Table 3: Checklist to assess reliability of our model's absolute infrastructure mortality estimates. The matrix classifies reliability into three levels: Robust (green), Caution (yellow), and High risk of erroneous outputs (red), based on three critical modelling constraints: Priors information (specifically for carcass location and persistence probabilities), variability of parameters in input data, and seasonality of road mortality events.

Key factors affecting model robustness	Robust	Caution	High risk of erroneous outputs
Prior information (carcass location and persistence probabilities)	Accurate: Based on literature from same ecosystems and vertebrate group or derived from independent field experiments.	Inaccurate: Based on literature but from different ecosystems or vertebrate groups. Result: Non-identifiability. Better use only as relative index.	Weak informative or uninformative: wide distributions with high SE or Uniform distributions (0-1). Result: Non-identifiability, model run errors, and mortality overestimation. Use only as relative index.
Variability of parameters in input data	Low variability: Daily carcass persistence and daily mortality numbers are relatively constant.	Moderate variability: Variability is present but remain within a moderate order of magnitude. Result: Increased uncertainty in estimates.	High variability: Extreme variability in daily carcass persistence or daily mortality numbers. Result: Use only as relative index.
Seasonality of road mortality events	Active season only: Dataset restricted to biologically active months of specific vertebrate groups.	Full year with low seasonality: Vertebrate groups with continuous activity along the year (e.g., some mammals). Result: Increased uncertainty in estimates.	Full year with high seasonality: vertebrate groups with inactive periods (e.g. amphibians/lizards). Result: Increased uncertainty and reduced accuracy in estimates.



994

995 Figure 1. Roadkill survey bias framework. This diagram illustrates how three types of
996 survey bias (carcass location bias, carcass persistence bias, and carcass
997 observation bias) impact the census data of roadkill within the surveyed road. These
998 theoretical different sizes of the squares in the diagram symbolize the quantity of
999 roadkill that would be available at each nested level of the framework. Additionally, *D*
1000 represents the time elapsed between the roadkill event and the maximum days a
1001 carcass remains on the road without disappearing until survey day, where carcass
1002 persistence bias occurs, while Sd represents the survey duration, during which
1003 observational bias occurs.

1004

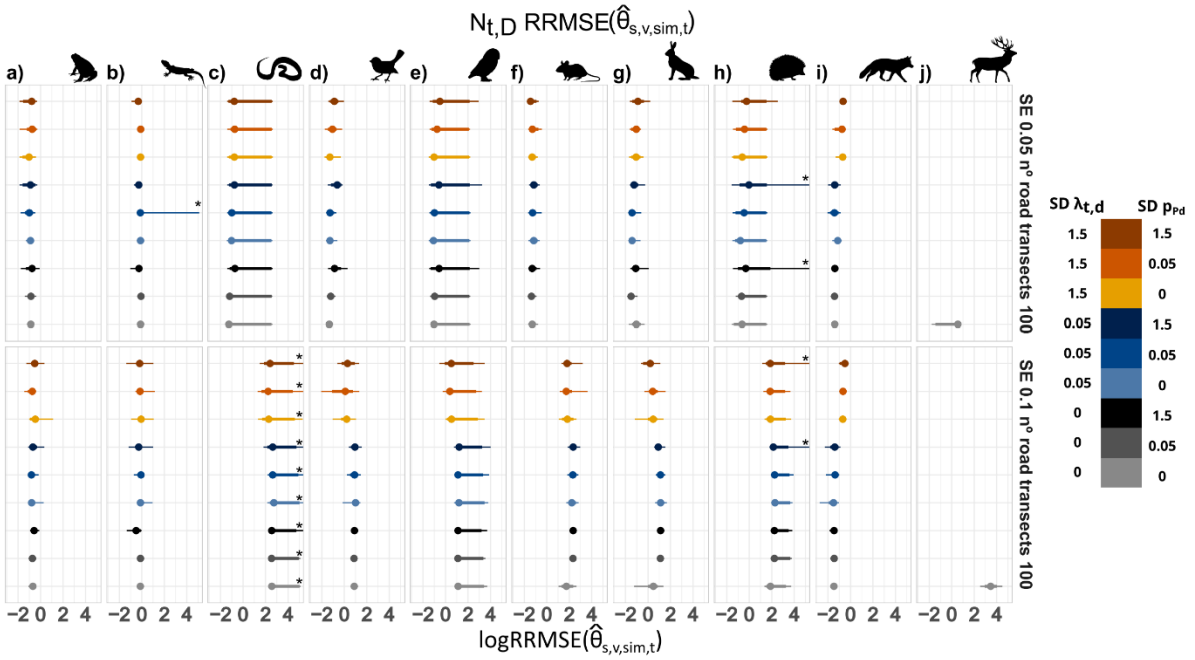


Figure 2: $N_{t,D}\ RRMSE(\hat{\theta}_{s,v,sim,t})$ values (Equation 5), where log values <0 indicate high accuracy of the model, log values =0 indicate that the magnitude of the error equals the true simulated value, and log values >0 indicates low accuracy of the model. This is evaluated across *s* = 9 different scenario combinations of mean daily number of roadkills and daily carcass persistence variability ($SD\ \lambda_{t,d}$ and $SD\ p_{Pd}$), *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. Coloured circles represent the mean, bold lines for 66% intervals, and thin lines 95% intervals. An asterisk (*) in the distributions indicates values exceeding 5 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.

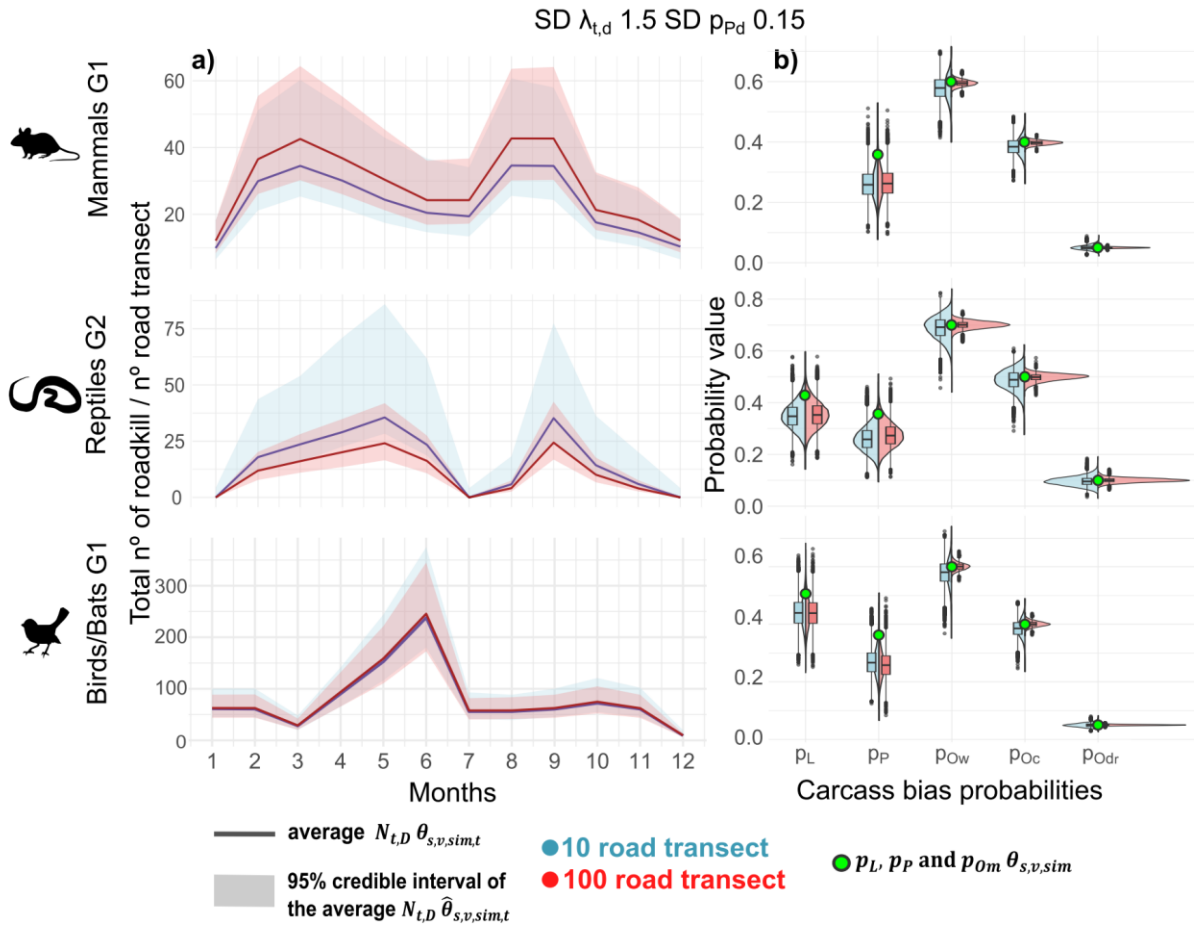


Figure 3: Comparison between the Bayesian estimation distribution of $N_{t,D}$ given $\hat{\theta}_{s,v,sim,t}$ and of p_L, p_P and p_{Om} given $\hat{\theta}_{s,v,sim}$, and their true simulated values, $N_{t,D}$ given $\theta_{s,v,sim,t}$ and p_L, p_P and p_{Om} given $\theta_{s,v,sim}$, for Mammals G1, Reptiles G2, and Birds/Bats G1. Here census data were simulated under high variability scenario for daily mean number of roadkills ($\lambda_{t,d}$) and daily carcass persistence probability (p_{Pd}), considering a SE = 0.05 in p_L and p_P priors. a) Comparison of total number of roadkills per transect. Lines represent the average $N_{t,D}\ \theta_{s,v,sim,t}$ over 20 simulations, while the shaded areas show the average 95% credible interval of Bayesian posterior estimates $N_{t,D}\ \hat{\theta}_{s,v,sim,t}$ over the 20 simulated census data. b) Comparison of carcass location, persistence and observation probability per method. Green dots represent the p_L, p_P and $p_{Om}\ \theta_{s,v,sim}$ values for *m* = walking (p_{Ow}), cycling (p_{Oc}) or

driving (p_{Odr}) survey methods. Violin plots combined with boxplots (representing the same underlying data) show the pooled p_L, p_P and p_{Om} $\hat{\theta}_{s,v,sim}$ over 20 simulated census data.

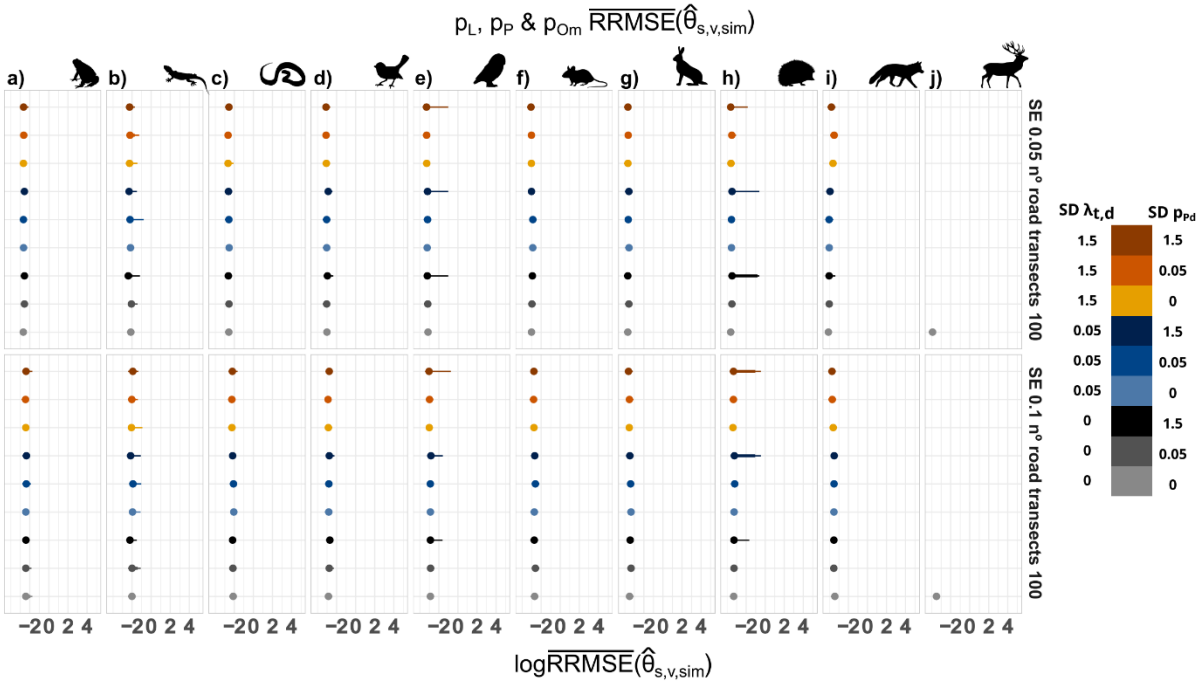


Figure 4: p_L, p_P & p_{Om} $\overline{RRMSE}(\hat{\theta}_{s,v,sim})$ (Equation 6), where log values <0 indicate high accuracy of the model, log values =0 indicate that the magnitude of the error equals the true simulated value, and log values >0 indicates low accuracy of the model. This is evaluated across *s* = 9 different scenario combinations of daily mean number of roadkills and daily carcass persistence variability ($SD\ \lambda_{t,d}$ and $SD\ p_{Pd}$), *v* = 10 vertebrate groups, *sim* = 20 simulations and *m* = walking, cycling and driving survey methods. Each distribution represents p_L, p_P & p_{Om} $\overline{RRMSE}(\hat{\theta}_{s,v,sim})$ values derived from each *sim* level 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. Coloured circles represent the mean, bold lines for 66% intervals, and thin lines 95% intervals. Note: Amphibians and Reptiles G1 vertebrate groups models only account for peak abundance months, excluding periods of typical absence.

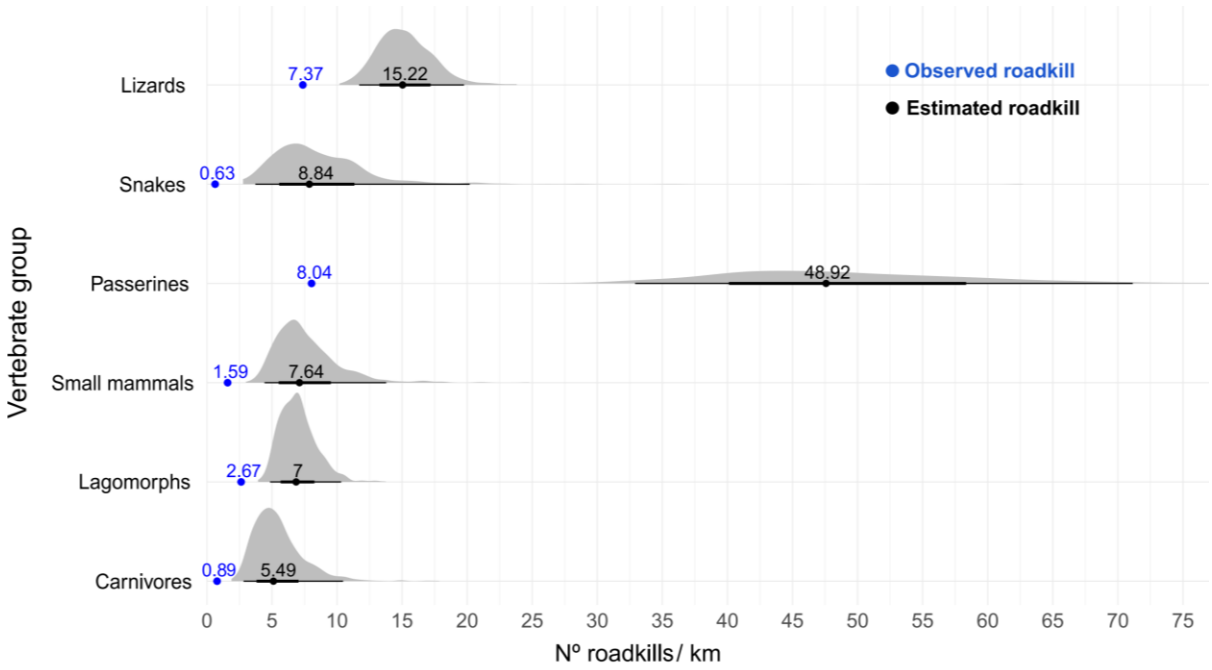


Figure 5. Observed roadkill rates per kilometer in road surveys (blue) and Bayesian posterior estimates of total roadkill rates per kilometer (black), derived from aggregating four monthly census data of the case study, for each vertebrate group. These estimated roadkill rates are limited to those that 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.

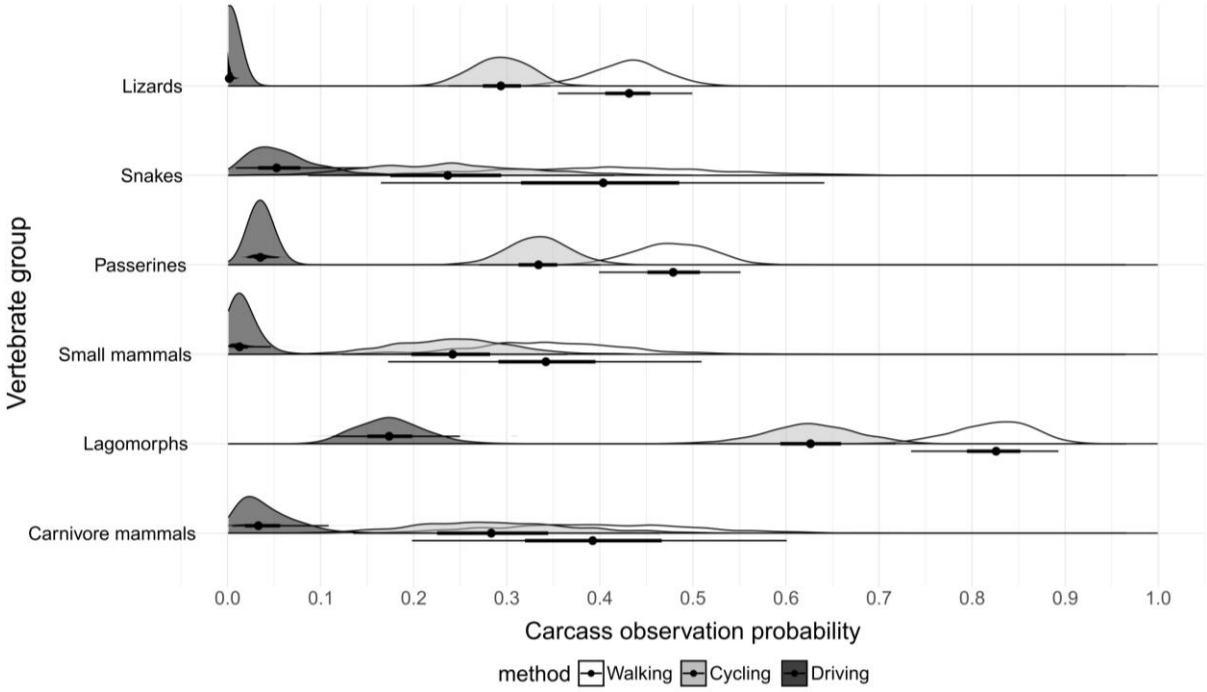


Figure 6. Bayesian posterior distribution of the carcass observation probabilities from case study, for each considered vertebrate groups. “Walking” means the estimation of carcass observation probability by walking survey method, “Cycling” by cycling survey method and “Driving” by driving survey method.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