

Estimating invasive rodent abundance using removal data and hierarchical models

Olivier Gimenez^{1*}

2024-12-26

¹ CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

* Corresponding author: olivier.gimenez@cefe.cnrs.fr

blabla.

Keywords: Invasive species, Multinomial N-mixture, Population size, Statistical ecology

Introduction

Invasive species are a significant global issue, with wide-ranging impacts on ecosystems, economies, and public health (Pyšek et al. 2020, Roy et al. 2024). Among these, the financial, epidemiological, social, and ecological costs associated with invasive rodents are substantial, as they damage infrastructures, degrade agricultural systems, and act as reservoirs for zoonotic diseases (Diagne et al. 2023).

Effective management of invasive species requires the estimation of population abundance for guiding control efforts and evaluating the success of eradication or regulation programs (Williams et al. 2002, Thompson et al. 2021). However, the challenge in estimating animal abundance is that individuals are not always observed even when present due to imperfect detection (Borchers et al. 2002, Seber and Schofield 2023). Ignoring imperfect detection leads to biased estimates of population abundance (Kéry and Schmidt 2008). To account for imperfect detection, capture-recapture methods are usually used to correct observed counts (McCrea and Morgan 2015). Yet, for invasive species, capture-recapture is often impractical, as ethical and management concerns typically prevent the release of captured animals.

An alternative approach involves the use of removal methods (Rodriguez de Rivera and McCrea 2021) in which individuals are captured and permanently removed from the study area during successive sampling occasions. This process leads to a decrease in the expected number of captures by a consistent proportion over time (rather than by a fixed amount decline), which informs on the total abundance as the initial population determines how quickly the number of individuals available for capture diminishes.

While standard removal methods are well-established (Moran 1951, Zippin 1956, 1958, Rodriguez de Rivera and McCrea 2021), recent advances in population ecology remain underutilized in the context of invasive species. Hierarchical models, in particular, have gained traction (Royle and Dorazio 2008, Kéry and Royle 2015) due to their ability to: (i) explicitly

separate biological processes of interest (e.g., population dynamics) from observation processes (e.g., imperfect detection), thus enabling more accurate modeling; (ii) incorporate environmental, spatial, or temporal covariates at multiple levels, allowing exploration of how various factors influence ecological processes; and (iii) share information across groups by modeling parameters hierarchically with random effects, which improves estimates for groups with fewer data.

In this paper, I showcase the application of a hierarchical formulation of removal models, the multinomial N-mixture model ([Dorazio et al. 2005](#)), to estimate the abundance of rodents in Europe. In this study, I focus on the coypu (*Myocastor coypus*) in France and the muskrat (*Ondatra zibethicus*) in the Netherlands. Both species are semi-aquatic rodents introduced to Europe in the early 20th century following escapes or releases from fur farms. The coypu, native to South America, has formed widespread invasive populations in France ([Bonnet et al. 2023](#)), where it causes significant damage to infrastructure and crops. Additionally, it serves as a healthy carrier of leptospirosis, a zoonotic disease with potentially serious consequences. Similarly, the muskrat, native to North America, has established extensive populations in the Netherlands. By burrowing into riverbanks, dykes, and dams, muskrats compromise the integrity of these structures, posing a threat to public safety ([Loon et al. 2017](#)).

Using removal data, I demonstrate the application of the multinomial N-mixture model to estimate the abundance of coypu and muskrat populations. First, I conduct a simulation study to evaluate the model's performance under varying numbers of sampling sites and sampling occasions. Second, I present a case study on a coypu population in France to illustrate the hierarchical structure of the multinomial N-mixture model, demonstrating how covariates can be incorporated to account for variations in abundance and capture probabilities. Third, I use a case study on muskrats in the Netherlands to demonstrate the integration of random effects within the model. To facilitate reproducibility, I provide the accompanying code and data,

aiming to promote the broader adoption of removal models in the study of biological invasions.

Methods

Multinomial N-mixture model

Think of a dice with six sides. The dice has a 1 in 6 chance of landing on face 1, the same for face 2, and so on. If I roll the dice 30 times, I expect to get face 1 five times, face 2 five times, and so on, on average. In this experiment, y_1 , the vector made of the number of 1s, y_2 , the number of 2s, \dots , and y_6 , the number of 6s, follows a multinomial distribution with parameters the number of rolls (30) and probabilities $(1/6, 1/6, \dots, 1/6)$.

Now think of a removal campaign conducted over 3 months. We record the number of rodents y_1 captured in month 1, y_2 in month 2, y_3 in month 3, and let y_4 represent the number of rodents never captured. Let p be the probability of capturing a rodent in a given month. The probability of capturing a rodent in the first month is $\pi_1 = p$. The probability of capturing a rodent in the second month is $\pi_2 = (1 - p)p$ the probability of not capturing it in the first month $(1 - p)$ multiplied by the probability of capturing it in the second month p . The probability of capturing a rodent in the third month is $\pi_3 = (1 - p)(1 - p)p$, the probability of not capturing it in the first and second months, $(1 - p)(1 - p)$, multiplied by the probability of capturing it in the third month, p . Finally, the probability of never being captured is $\pi_4 = 1 - (p + (1 - p)p + (1 - p)(1 - p)p)$ the complement of the probability of being captured in the first, second, or third month. If we assume that N represents the abundance, then we have that the vector of counts (y_1, y_2, y_3, y_4) follows a multinomial distribution with parameters N and probabilities $(\pi_1, \pi_2, \pi_3, \pi_4)$. In general, we assume that N follows a Poisson distribution with parameter the expected number of rodents denoted λ . And there you have it, the multinomial N-mixture model for a removal experiment, which is similar to throwing a

dice N times and the π 's give the probabilities that I get a given face of that dice. In general, we monitor rodents on several sites $i = 1, \dots, S$ and we need to estimate local population size N_i for site i . AJOUTER L'INTERET D'AVOIR PLUSIEURS SITES (cf ABSTRACT PAPIER DORAZIO) AND POLISH TEXT.

Parameters N , p , and λ are unknown and need to be estimated. In a frequentist framework, marginalization is performed by summing over all possible values of N (Dorazio et al. 2005). In a Bayesian framework, all these parameters are estimated directly, which simplifies the process (Royle and Dorazio 2006). Both parameters, λ and p , can be modeled as functions of explanatory spatial and temporal variables, in the spirit of generalized linear models, and Poisson or logistic regressions for example. I warmly recommend chapter 7 in Kéry and Royle (2015) for a detailed description of the multinomial mixture model.

Simulations

I conducted a simulation study to evaluate the model's performance by examining parameter bias under varying numbers of sampling sites and sampling occasions. I simulated removal data over 1, 5, 10 and 50 sites using a Poisson distribution with expected number of animals λ between 10 and 100 (20 values) for the ecological process. I simulated the observation process with a capture probability p varying between 0.3 and 0.9 (20 values) across 3, 5 and 10 occasions per site. In total, I considered 4800 scenarios. I fitted the multinomial N-mixture model to the simulated data within the frequentist framework using function `multinomPois()` in the R package `unmarked` (Kellner et al. 2023), and I repeated this procedure 500 times. Eventually, I calculated the relative bias for each parameter.

Case studies

Coypus in France

Tableau avec les données. Et les covariables.

To illustrate the approach, I investigated the impact of human disturbance on the occupancy of European otter (*Lutra lutra*) in France. The otter, a semi-aquatic mammal, faced near extinction in the 20th century in France due to extensive hunting for its fur. With hunting bans and protection efforts, the species is now recolonizing the country, and the ecological question is assessing its current distribution. Data on otter detection and non-detection were collected in 2003-2005 in the Midi-Pyrénées region. Observers searched for signs of otter presence at a small river catchment scale, which was used as the spatial sampling unit.

Muskrats in the Netherlands

Pour muskrats: https://ipt.nlbif.nl/resource?r=hwh_muskrat_1987-2014#anchor-versions and <https://www.gbif.org/dataset/7d75109d-a6cb-4e90-89d0-79d08577c580>.

Implementation

Temperature. Citer krigR (Kusch and Davy 2022). Month effect. First case study for hierarchical and covariates. Second case study for random effect.

Bayesian and MCMC easier to implement random effects.

For all analyses, I used the statistical language R (R Core Team 2024). I used the tidyverse (Wickham et al. 2019) suite of packages for data manipulation and visualization and sf (Pebesma and Bivand 2023) for dealing with spatial data. Bayesian framework using the ubms (Kellner et al. 2021) package, also NIMBLE (de Valpine et al. 2017). I specified weakly informative priors for all parameters, specifically normal distributions with mean 0 and

standard deviation 1.5 for regression parameters, and a uniform distribution for the standard deviation of the random effects. I ran two chains for a total of 15,000 iterations with a burn-in of 5,000 iterations. I summarized posterior distributions with posterior mean and 95% credible intervals. I assessed model convergence using R-hat values (< 1.1), effective sample size (> 100), and visual inspection of the trace plots.

Results and discussion

The European Union Regulation No. 1143/2014 was established to mitigate the negative impacts of invasive alien species on biodiversity. This regulation outlines measures to prevent the introduction of invasive alien species and manage their established populations. Two species listed as species of Union Concern under the regulation, which obligates member states to implement appropriate management actions.

The results of the simulation study underscored the importance of accounting for spatial autocorrelation - see Fig. 1. Ignoring spatial autocorrelation led to a relative bias of -26% in the slope of the covariate, compared to just 0.7% when spatial autocorrelation was included. Although the range parameter exhibited substantial bias (330%), this outcome was expected. In spatial models, there is typically a strong positive relationship between the range parameter and the partial sill (Zhang 2004, Ver Hoef et al. 2024), which showed a bias of 12%. Importantly, the ratio of the range parameter to the partial sill can be reliably estimated (Zhang 2004), as evidenced by its negligible bias (0.9%). Overall, the root mean square prediction error (RMSPE) was lower for the model with spatial correlation (RMSPE = 0.18) compared to the model without it (RMSPE = 0.23).

Coypus case study. I provide the parameter estimates from the new model accommodating spatial autocorrelation, unless otherwise specified. Detection probability was less than one, estimated at 0.71 (0.59, 0.80), which justified the use of occupancy models. The proportion of

cultivated area had no effect, with a slope estimated at 0.60 (-0.67, 1.96). Population density also had no effect on occupancy probability, with a slope estimated at -0.96 (-2.24, 0.17). However, I did find a negative effect when spatial autocorrelation was ignored, with a slope estimated at -1.10 (-1.99, -0.34). This latter result aligns with a previous analysis of a more extensive dataset (Couturier et al. 2023) that also ignored spatial autocorrelation.

Muskrats case study. - see Fig. 2

Two short-term perspectives arise from this work. From a methodological perspective, the new approach could be extended to multi-season occupancy models, enabling the modeling of colonization probability as a function of distance to habitat features that may impede species movement (e.g., Kervellec et al. 2024). This would facilitate the quantification of landscape connectivity in freshwater ecosystems. Such development requires moving to spatio-temporal models for stream and river data, which have recently become available (Santos-Fernandez et al. 2022). From an ecological perspective, the new approach presents significant potential for the analysis of environmental DNA (eDNA). The eDNA methodology offers substantial promise for the non-invasive monitoring of biodiversity in freshwater ecosystems (Carraro et al. 2020). While spatial stream network models have been employed to analyze eDNA data (Winkowski et al. 2024), these models have overlooked the issue of imperfect detection. Previous studies have recognized occupancy models as effective tools for eDNA data analysis (Burian et al. 2021), with some considering spatial autocorrelation (Chen and Ficetola 2019), however they have yet to integrate spatial stream networks. The new approach addresses this gap by incorporating both imperfect detection and spatial stream networks.

Recommendation. Record effort, 0s from non-sampling and 0s for non-detections. Close-kin as alternative. Use for projections. Also multispecies (cf N-mixture multispecies). Also immigration et papers par d'autres sur multinomial N-mixture extensions. Here, just AR(1) as in Outhwaite et al. (2018) random walk, mais mechanisms et demog parameters possible.

175 Effectifs en eux-mêmes un peu meaningless, boucle de gestion adaptative.

176 Two main assumptions: constant capture, closed pop.

177 Pb lack of fit. Main three are overdispersion, ok with negative binomial (give results); closure

178 ok w/ open models w/ survival/recruitment or simple multisession (papier census aérien);

179 spatial autocorrelation, ok en classique (GAM ou autre), et aussi possible w/ stream networks

180 pour coypus; iv) Hierarchical models are well-suited for representing dependencies in space,

181 time, or species interactions, which are common in ecological systems.

182 Pour ouvert, papier Eleni open (mais permanent emigration) ([Matechou et al. 2016](#)), puis Zhou

183 pour temporary emigration ([Zhou et al. 2019](#)). Ajouter aussi le papier de Link pour dispersal

184 ([Link et al. 2018](#)).

185 Un peu chiant unmarked. Avantage bayésien. No boundary estimates, easy random effects.

186 Difference entre population et species.

187 Les hypothèses du modèle. Ici ou en discussion. Population fermée: pas de naissances/morts

188 ou émigration/immigration pendant la période de piégeage. Les sites de piégeage sont

189 indépendants: les prélèvements sur un site n'affectent pas ceux faits sur un autre site.

190 Lack of fit, overdispersion and negative binomial.

191 Si random effect site pas pertinent, on peut faire du spatial autocorrelation avec RSR. Easy avec

192 ubms, possible aussi avec NIMBLE

193 (<https://wildlife.onlinelibrary.wiley.com/doi/10.1002/jwmg.22296>). Pas possible avec

194 unmarked.

195 Pour l'ouvert, possible avec unmarked, pas ubms. OK avec NIMBLE.

196 Pour NB, pas possible avec ubms. OK avec NIMBLE et unmarked.

Acknowledgments

I would like to warmly thank X and Y for data.

Data availability statement

Data and code are available at <https://github.com/oliviergimenez/counting-rodents>.

Funding statement

Ici mettre financement UM.

Conflict of interest disclosure

The author has no conflicts of interest to declare.

References

- Bonnet, M., G. Guédon, S. Bertolino, C. Harmange, A. Pagano, D. Picard, and O. Pays. 2023. Improving the management of aquatic invasive alien rodents in france: Appraisal and recommended actions. *Management of Biological Invasions* 14:625–640.
- Borchers, D. L., S. T. Buckland, and Z. W. 2002. Estimating animal abundance: Closed populations. Springer-Verlag, London.
- Burian, A., Q. Mauvisseau, M. Bulling, S. Domisch, S. Qian, and M. Sweet. 2021. Improving the reliability of eDNA data interpretation. *Molecular Ecology Resources* 21:1422–1433.
- Carraro, L., E. Mächler, R. Wüthrich, and F. Altermatt. 2020. Environmental DNA allows upscaling spatial patterns of biodiversity in freshwater ecosystems. *Nature Communications* 11:3585.

216 Chen, W., and G. F. Ficetola. 2019. Conditionally autoregressive models improve occupancy
 217 analyses of autocorrelated data: An example with environmental DNA. *Molecular Ecology*
 218 *Resources* 19:163–175.

219 Couturier, T., J. Steinmetz, P. Defos du Rau, D. Marc, E. Trichet, R. Gomes, and A. Besnard. 2023.
 220 Intensive agriculture as the main limiting factor of the otter's return in southwest france.
 221 *Biological Conservation* 279:109927.

222 de Valpine, P., D. Turek, C. Paciorek, C. Anderson-Bergman, D. Temple Lang, and R. Bodik.
 223 2017. [Programming with models: Writing statistical algorithms for general model structures](#)
 224 [with NIMBLE](#). *Journal of Computational and Graphical Statistics* 26:403–413.

225 Diagne, C., L. Ballesteros-Mejia, R. Cuthbert, T. Bodey, J. Fantle-Lepczyk, E. Angulo, A. Bang, G.
 226 D. G., and F. Courchamp. 2023. Economic costs of invasive rodents worldwide: The tip of
 227 the iceberg. *PeerJ* 11:e14935.

228 Dorazio, R. M., H. L. Jelks, and F. Jordan. 2005. Improving removal-based estimates of
 229 abundance by sampling a population of spatially distinct subpopulations. *Biometrics*
 230 61:1093–1101.

231 Kellner, K. F., N. L. Fowler, T. R. Petroelje, T. M. Kautz, D. E. Beyer, and J. L. Belant. 2021. [ubms:](#)
 232 [An R package for fitting hierarchical occupancy and n-mixture abundance models in a](#)
 233 [bayesian framework](#). *Methods in Ecology and Evolution* 13:577–584.

234 Kellner, K. F., A. D. Smith, J. A. Royle, M. Kery, J. L. Belant, and R. B. Chandler. 2023. [The](#)
 235 [unmarked R package: Twelve years of advances in occurrence and abundance modelling in](#)
 236 [ecology](#). *Methods in Ecology and Evolution* 14:1408–1415.

237 Kervellec, M., T. Couturier, S. Bauduin, D. Chenesseau, P. D. du Rau, N. Drouet-Hoguet, C.
 238 Duchamp, J. Steinmetz, J.-M. Vandel, and O. Gimenez. 2024. Bringing circuit theory into
 239 spatial occupancy models to assess landscape connectivity. *Methods in Ecology and*
 240 *Evolution*.

241 Kéry, M., and J. A. Royle. 2015. Applied hierarchical modeling in ecology: Analysis of

distribution, abundance and species richness in r and BUGS. Academic Press, London, UK.

Kéry, M., and B. Schmidt. 2008. Imperfect detection and its consequences for monitoring in conservation. *Community Ecology* 9:207–216.

Kusch, E., and R. Davy. 2022. [KrigR-a tool for downloading and statistically downscaling climate reanalysis data](#). *Environmental Research Letters* 17.

Link, W. A., S. J. Converse, A. A. Yackel Adams, and N. J. Hostetter. 2018. Analysis of population change and movement using robust design removal data. *Journal of Agricultural, Biological and Environmental Statistics* 23:463–477.

Loon, E. E. van, D. Bos, C. J. van Hellenberg Hubar, and R. C. Ydenberg. 2017. A historical perspective on the effects of trapping and controlling the muskrat (*ondatra zibethicus*) in the netherlands. *Pest Management Science* 73:305–312.

Matechou, E., R. S. McCrea, B. J. T. Morgan, D. J. Nash, and R. A. Griffiths. 2016. Open models for removal data. *The Annals of Applied Statistics* 10:1572–1589.

McCrea, R. S., and B. J. T. Morgan. 2015. Analysis of capture-recapture data. Chapman & Hall.

Moran, P. A. P. 1951. A mathematical theory of animal trapping. *Biometrika* 38:307–311.

Outhwaite, C. L., R. E. Chandler, G. D. Powney, B. Collen, R. D. Gregory, and N. J. B. Isaac. 2018. Prior specification in bayesian occupancy modelling improves analysis of species occurrence data. *Ecological Indicators* 93:333–343.

Pebesma, E., and R. Bivand. 2023. [Spatial Data Science: With applications in R](#). Chapman and Hall/CRC.

Pyšek, P., P. E. Hulme, D. Simberloff, S. Bacher, T. M. Blackburn, J. T. Carlton, W. Dawson, F. Essl, L. C. Foxcroft, P. Genovesi, J. M. Jeschke, I. Kühn, A. M. Liebhold, N. E. Mandrak, L. A. Meyerson, A. Pauchard, J. Pergl, H. E. Roy, H. Seebens, M. van Kleunen, M. Vilà, M. J. Wingfield, and D. M. Richardson. 2020. Scientists’ warning on invasive alien species. *Biological Reviews* 95:1511–1534.

R Core Team. 2024. [R: A language and environment for statistical computing](#). R Foundation for

Statistical Computing, Vienna, Austria.

Rodriguez de Rivera, O., and R. McCrea. 2021. Removal modelling in ecology: A systematic review. *Plos One* 3:e0229965.

Roy, H. E., A. Pauchard, P. J. Stoett, T. Renard Truong, L. A. Meyerson, S. Bacher, B. S. Galil, P. E. Hulme, T. Ikeda, S. Kavileveettil, M. A. McGeoch, M. A. Nuñez, A. Ordonez, S. J. Rahlao, E. Schwindt, H. Seebens, A. W. Sheppard, V. Vandvik, A. Aleksanyan, M. Ansong, T. August, R. Blanchard, E. Brugnoli, J. K. Bukombe, B. Bwalya, C. Byun, M. Camacho-Cervantes, P. Cassey, M. L. Castillo, F. Courchamp, K. Dehnen-Schmutz, R. D. Zenni, C. Egawa, F. Essl, G. Fayvush, R. D. Fernandez, M. Fernandez, L. C. Foxcroft, P. Genovesi, Q. J. Groom, A. I. González, A. Helm, I. Herrera, A. J. Hiremath, P. L. Howard, C. Hui, M. Ikegami, E. Keskin, A. Koyama, S. Ksenofontov, B. Lenzner, T. Lipinskaya, J. L. Lockwood, D. C. Mangwa, A. F. Martinou, S. M. McDermott, C. L. Morales, J. Müllerová, N. A. Mungi, L. K. Munishi, H. Ojaveer, S. N. Pagad, N. P. K. T. S. Pallewatta, L. R. Peacock, E. Per, J. Pergl, C. Preda, P. Pyšek, R. K. Rai, A. Ricciardi, D. M. Richardson, S. Riley, B. J. Rono, E. Ryan-Colton, H. Saeedi, B. B. Shrestha, D. Simberloff, A. Tawake, E. Tricarico, S. Vanderhoeven, J. Vicente, M. Vilà, W. Wanzala, V. Werenkraut, O. L. F. Weyl, J. R. U. Wilson, R. O. Xavier, and S. R. Ziller. 2024. Curbing the major and growing threats from invasive alien species is urgent and achievable. *Nature Ecology & Evolution* 8:1216–1223.

Royle, J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology: The analysis of data from populations, metapopulations and communities. Academic Press. San Diego, California.

Royle, J. andrew, and R. Dorazio. 2006. Hierarchical models of animal abundance and occurrence. *Journal of Agricultural, Biological, and Environmental Statistics* 11:249–263.

Santos-Fernandez, E., J. M. Ver Hoef, E. E. Peterson, J. McGree, D. J. Isaak, and K. Mengersen. 2022. Bayesian spatio-temporal models for stream networks. *Computational Statistics & Data Analysis* 170:107446.

- Seber, G. A. F., and M. R. Schofield. 2023. Estimating presence and abundance of closed populations. Springer.
- Thompson, B. K., J. D. Olden, and S. J. Converse. 2021. Mechanistic invasive species management models and their application in conservation. *Conservation Science and Practice* 3:e533.
- Ver Hoef, J. M., E. Blagg, M. Dumelle, P. M. Dixon, D. L. Zimmerman, and P. B. Conn. 2024. Marginal inference for hierarchical generalized linear mixed models with patterned covariance matrices using the laplace approximation. *Environmetrics* n/a:e2872.
- Wickham, H., M. Averick, J. Bryan, W. Chang, L. D. McGowan, R. François, G. Grolemund, A. Hayes, L. Henry, J. Hester, M. Kuhn, T. L. Pedersen, E. Miller, S. M. Bache, K. Müller, J. Ooms, D. Robinson, D. P. Seidel, V. Spinu, K. Takahashi, D. Vaughan, C. Wilke, K. Woo, and H. Yutani. 2019. [Welcome to the tidyverse](#). *Journal of Open Source Software* 4:1686.
- Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, San Diego, California, USA.
- Winkowski, J. J., J. D. Olden, and S. Brown. 2024. Integrating spatial stream network models and environmental DNA to estimate current and future distributions of nonnative smallmouth bass. *Transactions of the American Fisheries Society* 153:180–199.
- Zhang, H. 2004. Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *Journal of the American Statistical Association* 99:250–261.
- Zhou, M., R. S. McCrea, E. Matechou, D. J. Cole, and R. A. Griffiths. 2019. Removal models accounting for temporary emigration. *Biometrics* 75:24–35.
- Zippin, C. 1956. An evaluation of the removal method of estimating animal populations. *Biometrics* 12:163–189.
- Zippin, C. 1958. The removal method of population estimation. *The Journal of Wildlife Management* 22:82–90.

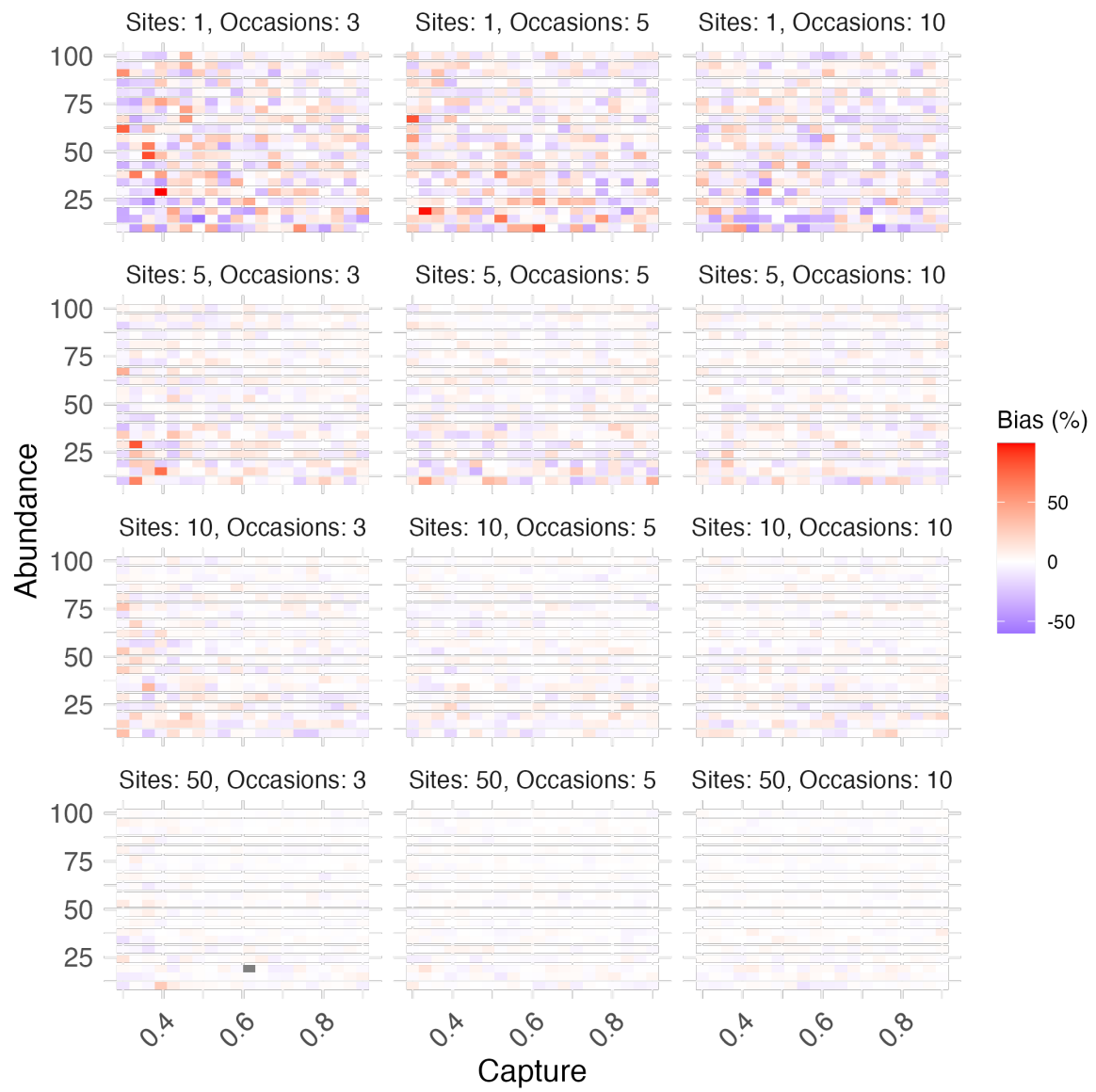


Figure 1: Relative bias.

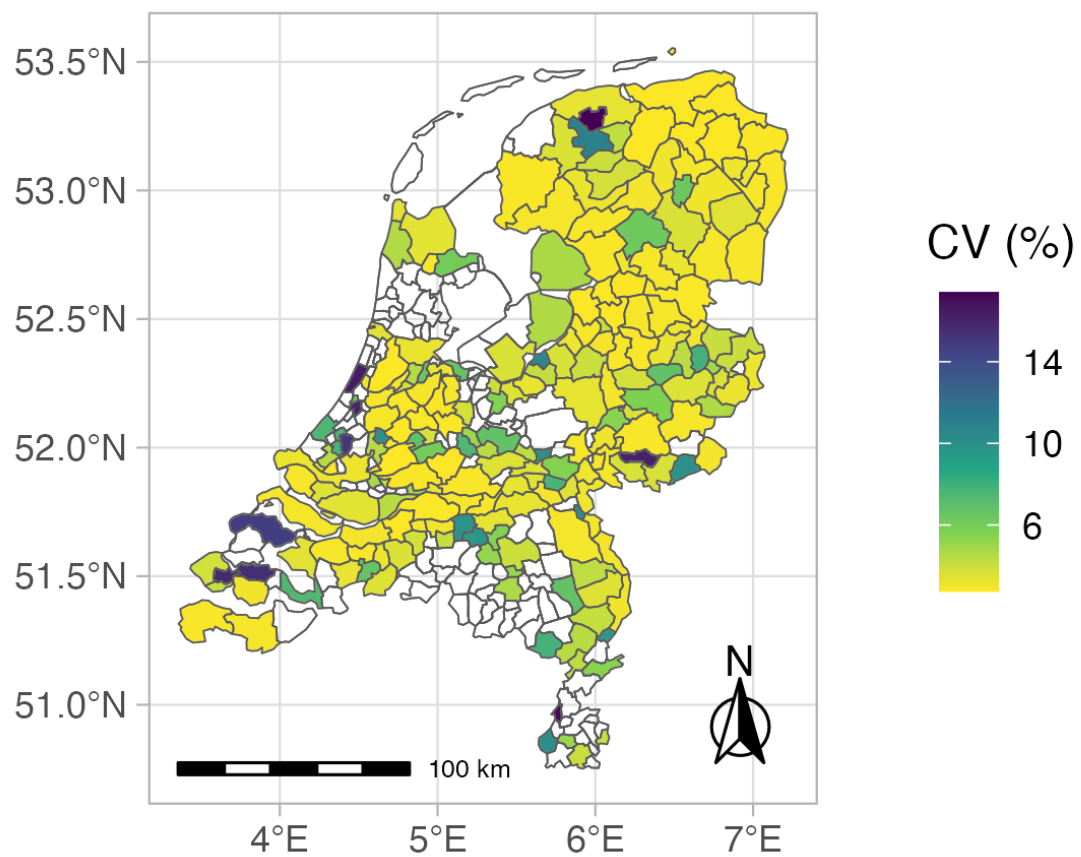
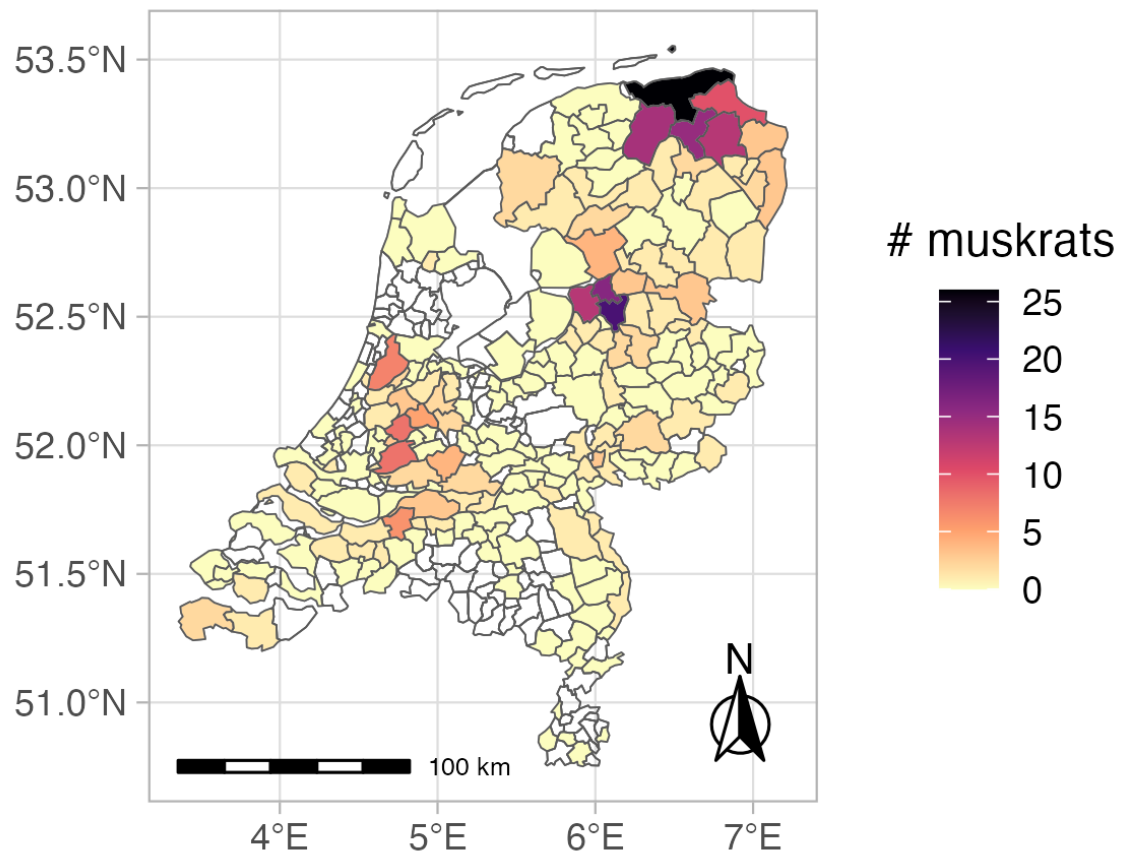


Figure 2: Muskrats.