

# Estimating coypu (*Myocastor coypus*) abundance using removal data

Olivier Gimenez<sup>1\*</sup>

2024-12-26

<sup>1</sup> CEFE, Univ Montpellier, CNRS, EPHE, IRD, Montpellier, France

\* Corresponding author: [olivier.gimenez@cefe.cnrs.fr](mailto: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 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 die with six sides. The die has a 1 in 6 chance of landing on face 1, the same for face 2, and so on. If I roll the die 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 coypu  $y_1$  captured in month 1,  $y_2$  in month 2,  $y_3$  in month 3, and let  $y_4$  represent the number of coypu never captured. Let  $p$  be the probability of capturing a coypu in a given month. The probability of capturing a coypu in the first month is  $\pi_1 = p$ . The probability of capturing a coypu 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 coypu 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  $(\pi_1, \pi_2, \pi_3, \pi_4)$ . INTRODUCE  $N \sim \text{Poisson}(\lambda)$ . AND NEGATIVE BINOMIAL. ALSO SEVERAL SITES. EXPLAIN LA PARTIE HIERARCHIQUE ICI. METTRE UN EFFET

83 ALEATOIRE QUELQUE PART?

84 FAIRE LE LIEN ENTRE LE DE ET LA CAMPAGNE DE PIEGEAGE.

85 Parameters  $N$ ,  $p$ , and  $\lambda$  are unknown and need to be estimated. In a frequentist framework,  
86 marginalization is performed by summing over all possible values of  $N$ . In a Bayesian  
87 framework, all these parameters are estimated directly, which simplifies the process.

88 Both parameters,  $N$  and  $p$ , can be modeled as functions of explanatory spatial variables, in the  
89 spirit of generalized linear models, and logistic regressions for example.

90 Citer chapitre dans bouquin de Kéry and Royle qui va bien (vol I).

## 91 **Implementation**

92 For all analyses, I used the statistical language R ([R Core Team 2024](#)). I used the `tidyverse`  
93 ([Wickham et al. 2019](#)) suite of packages for data manipulation and visualization and `sf`  
94 ([Pebesma and Bivand 2023](#)) for dealing with spatial data. I fitted models within the frequentist  
95 framework... `unmarked` ([Kellner et al. 2023](#)). Bayesian framework using the `ubms` ([Kellner et al.](#)  
96 [2021](#)) package. I specified weakly informative priors for all parameters, specifically normal  
97 distributions with mean 0 and standard deviation 1.5 for regression parameters, and uniform  
98 distributions for the partial sill, range parameter and detection probability. I ran two chains for  
99 a total of 15,000 iterations with a burn-in of 5,000 iterations. I summarized posterior  
100 distributions with posterior mean and 95% credible intervals. I assessed model convergence  
101 using R-hat values ( $< 1.1$ ), effective sample size ( $> 100$ ), and visual inspection of the trace plots.

## 102 **Simulations**

103 Voir un peu comment se comporte le modèle en fonction du nombre de sites et d'occasions.

104 I conducted a simulation study to evaluate model performance by examining parameter bias

and prediction error. I simulated a stream network over  $S = 100$  sites using an exponential tail-down structure (inspired by the case study, see previous section) with partial sill  $\sigma^2 = 2$  and range parameter  $\theta = 10$ . I considered a single covariate, normally distributed with mean 0 and standard deviation 1, with a linear effect (on the logit scale) on the occupancy probability with intercept  $\beta_0 = 0.5$  and slope  $\beta_1 = 1$ . I simulated the observation process with a detection probability  $p = 0.6$  across 5 repeated visits per site. I fitted both models, with and without spatial autocorrelation, to the simulated data, and I repeated this procedure 100 times. Eventually, I calculated the relative bias for each parameter and the root mean square prediction error (RMSPE) for each model. Unmarked only to speed up process.

## Case studies

Temperature. Citer krigR. Month effect. First case study for hierarchical and covariates. Second case study for random effect.

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

## 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. Effectifs en eux-memes un peu meaningless, boucle de gestion adaptative.



174 Two main assumptions: constant capture, closed pop.

175 Pb lack of fit. Main three are overdispersion, ok with negative binomial (give results); closure  
 176 ok w/ open models w/ survival/recruitment or simple multisession (papier census aérien);  
 177 spatial autocorrelation, ok en classique (GAM ou autre), et aussi possible w/ stream networks  
 178 pour coypus; iv) Hierarchical models are well-suited for representing dependencies in space,  
 179 time, or species interactions, which are common in ecological systems.

180 Pour ouvert, papier Eleni open (mais permanent emigration) ([Matechou et al. 2016](#)), puis Zhou  
 181 pour temporary emigration ([Zhou et al. 2019](#)). Ajouter aussi le papier de Link pour dispersal  
 182 ([Link et al. 2018](#)).

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

184 Difference entre population et species.

185 Les hypothèses du modèle. Ici ou en discussion. Population fermée: pas de naissances/morts  
 186 ou émigration/immigration pendant la période de piégeage. Les sites de piégeage sont  
 187 indépendants: les prélèvements sur un site n'affectent pas ceux faits sur un autre site.

188 Lack of fit, overdispersion and negative binomial.

189 Si random effect site pas pertinent, on peut faire du spatial autocorrelation avec RSR. Easy avec  
 190 ubms, possible aussi avec NIMBLE  
 191 (<https://wildlife.onlinelibrary.wiley.com/doi/10.1002/jwmg.22296>). Pas possible avec  
 192 unmarked.

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

194 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.

- Chen, W., and G. F. Ficetola. 2019. Conditionally autoregressive models improve occupancy analyses of autocorrelated data: An example with environmental DNA. *Molecular Ecology Resources* 19:163–175.
- Couturier, T., J. Steinmetz, P. Defos du Rau, D. Marc, E. Trichet, R. Gomes, and A. Besnard. 2023. Intensive agriculture as the main limiting factor of the otter's return in southwest france. *Biological Conservation* 279:109927.
- Diagne, C., L. Ballesteros-Mejia, R. Cuthbert, T. Bodey, J. Fantle-Lepczyk, E. Angulo, A. Bang, G. D. G, and F. Courchamp. 2023. Economic costs of invasive rodents worldwide: The tip of the iceberg. *PeerJ* 11:e14935.
- Dorazio, R. M., H. L. Jelks, and F. Jordan. 2005. Improving removal-based estimates of abundance by sampling a population of spatially distinct subpopulations. *Biometrics* 61:1093–1101.
- Kellner, K. F., N. L. Fowler, T. R. Petroelje, T. M. Kautz, D. E. Beyer, and J. L. Belant. 2021. [ubms: An R package for fitting hierarchical occupancy and n-mixture abundance models in a bayesian framework](#). *Methods in Ecology and Evolution* 13:577–584.
- Kellner, K. F., A. D. Smith, J. A. Royle, M. Kery, J. L. Belant, and R. B. Chandler. 2023. [The unmarked R package: Twelve years of advances in occurrence and abundance modelling in ecology](#). *Methods in Ecology and Evolution* 14:1408–1415.
- Kervellec, M., T. Couturier, S. Bauduin, D. Chenesseau, P. D. du Rau, N. Drouet-Hoguet, C. Duchamp, J. Steinmetz, J.-M. Vandel, and O. Gimenez. 2024. Bringing circuit theory into spatial occupancy models to assess landscape connectivity. *Methods in Ecology and Evolution*.
- 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.

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

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

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

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

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

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

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

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

260 R Core Team. 2024. [R: A language and environment for statistical computing](#). R Foundation for  
 261 Statistical Computing, Vienna, Austria.

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

264 Roy, H. E., A. Pauchard, P. J. Stoett, T. Renard Truong, L. A. Meyerson, S. Bacher, B. S. Galil, P. E.  
 265 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.

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.

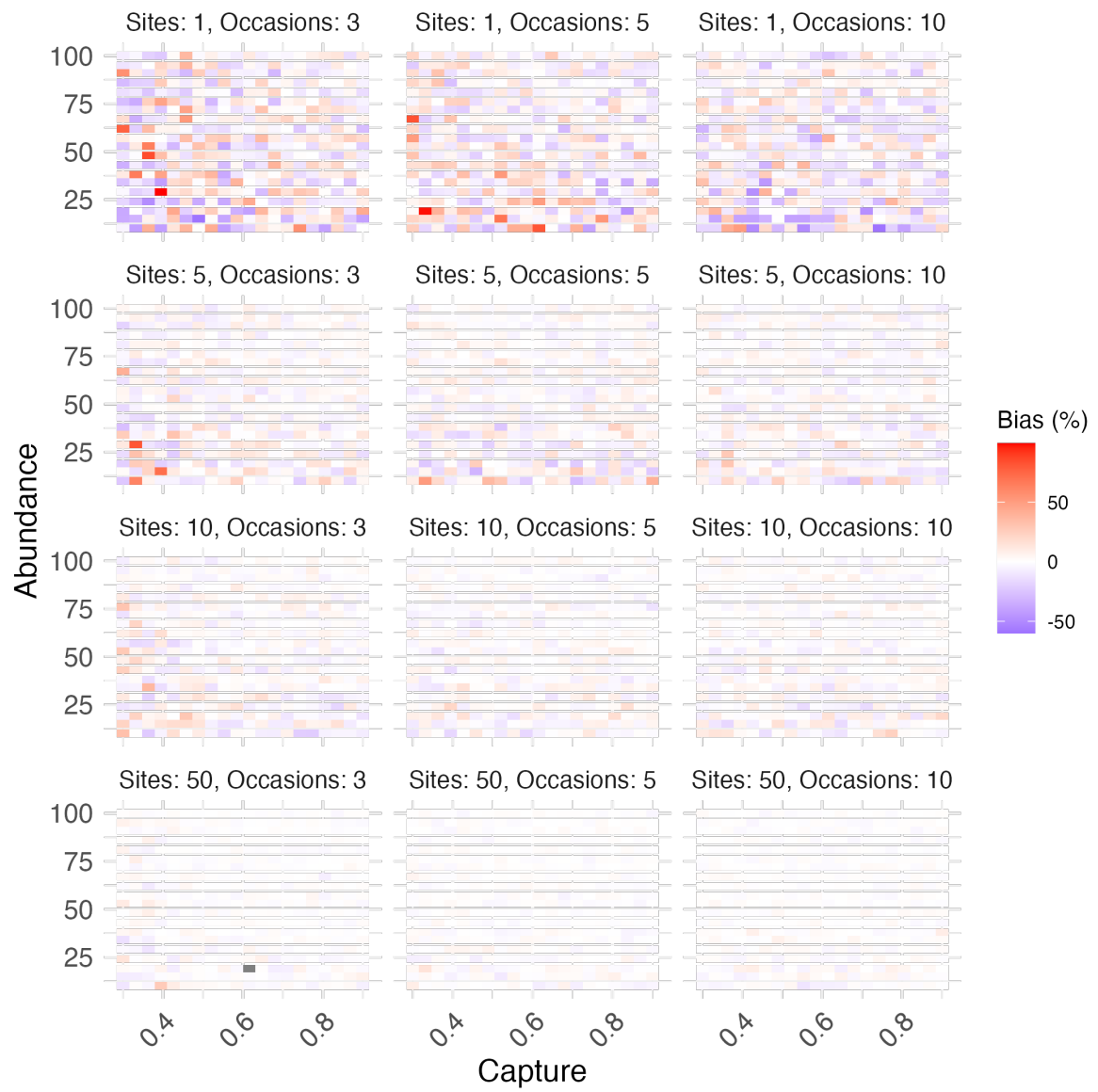


Figure 1: Relative bias.

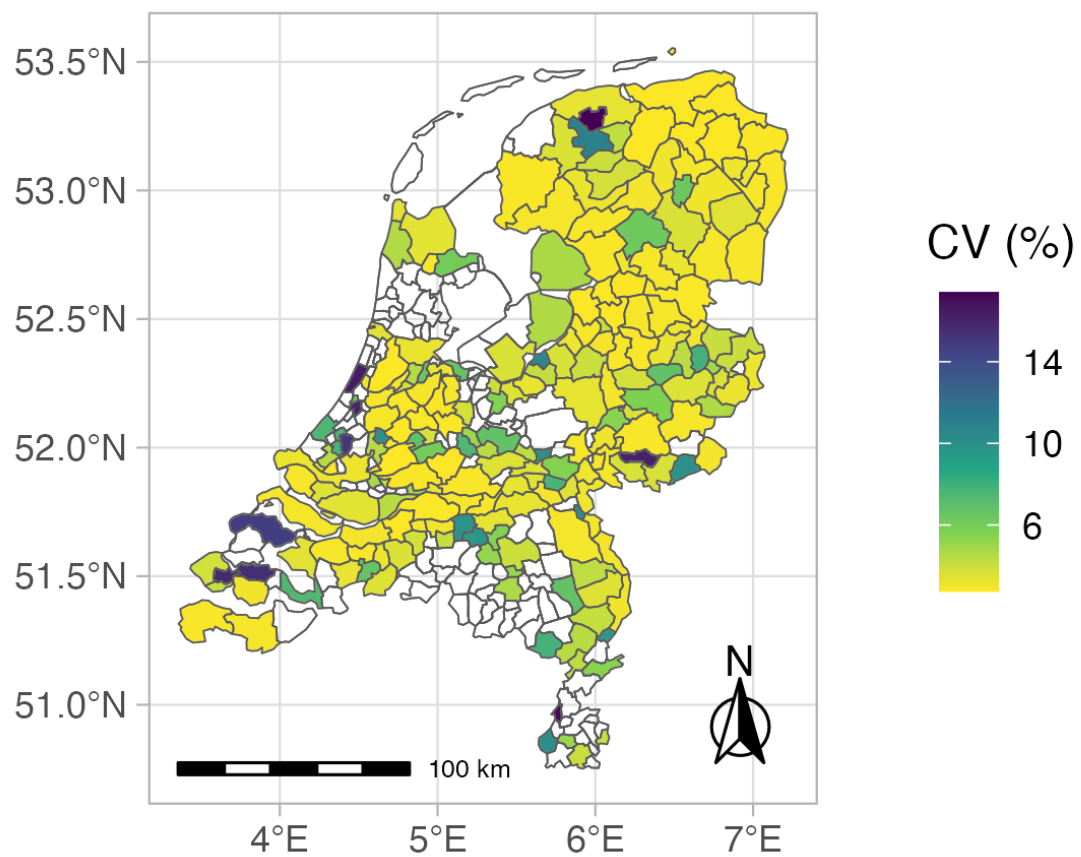
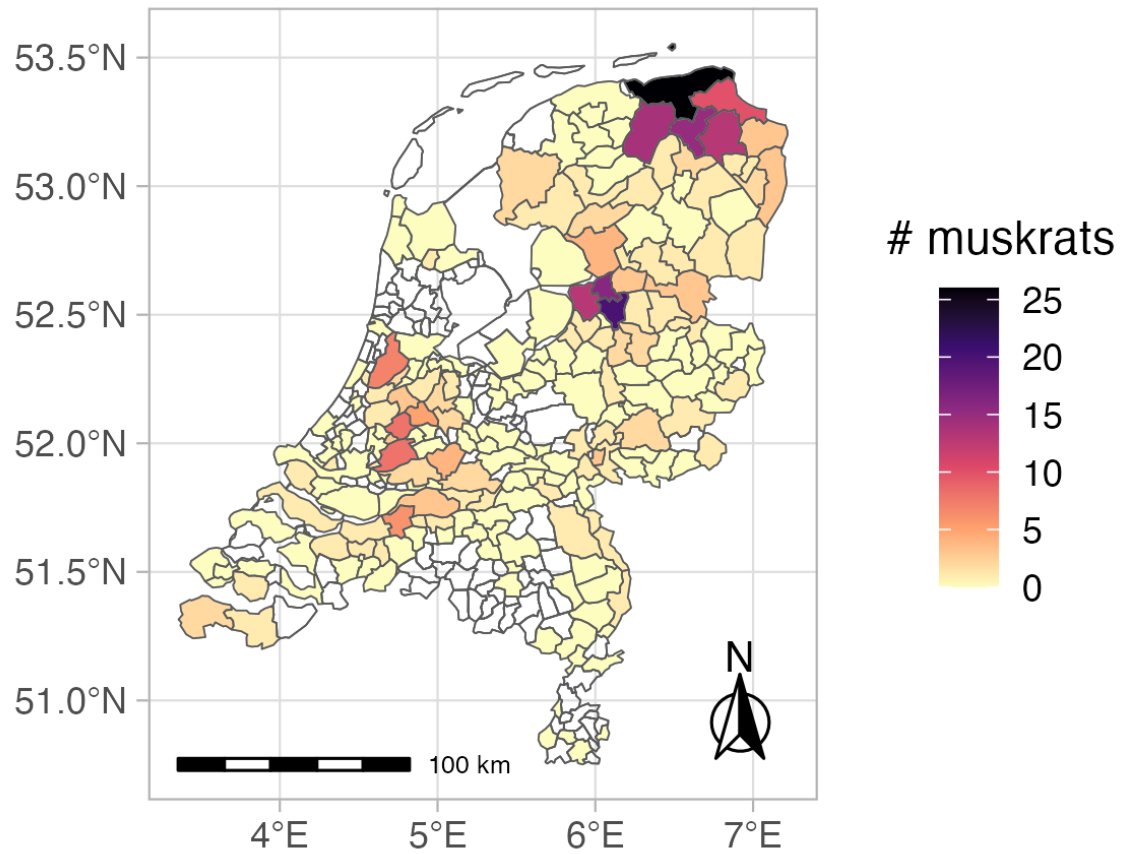


Figure 2: Muskrats.