


RESEARCH ARTICLE

Improving inferences and predictions of species environmental responses with occupancy data

Teresa Morán-López¹  | Sofía Ruiz-Suarez¹ | Joaquín Aldabe²  |
 Juan Manuel Morales¹ 

¹Grupo de Ecología Cuantitativa, INIBIOMA-CONICET, Universidad Nacional del Comahue, San Carlos De Bariloche, Rio Negro, Argentina

²Centro Universitario de la Región Este - CURE, Universidad de la República de Montevideo, Uruguay

Correspondence

Teresa Morán-López
 Email: tmoranlopez@gmail.com

Funding information

Ministry of Science of Argentina, Grant/Award Number: PICT-2015-0815 and PICT-2018-01566; BirdLife International; Aves Uruguay; International Programs, US Forest Service, Southern Cone Grassland Alliance Program; International Programs, US Forest Service; Argentinian Government (CONICET)

Handling Editor: Nick Isaac

Abstract

- Occupancy models represent a useful tool to estimate species distribution throughout the landscape. Among them, MacKenzie et al.'s model (2002, MC), is frequently used to infer species environmental responses. However, the assumption that detection probability is homogeneous or fully explained by covariates may limit its performance. Species should be more easily observed at sites with a higher number of individuals. We simulated data following occupancy model (RN) that accounts for abundance-driven heterogeneous detection and two variants with overdispersion in the detection probability and local abundances. Then, we compared the performance of the MC model against that of RN.
- In addition to model misspecifications, insufficient information in data (i.e. infrequent detections) can limit our ability to detect existing effects with affordable sampling designs. To deal with this source of error, we extended RN approach to a community-level joint species model (RN-JSM), where species responses and detectability depended on their traits and phylogeny. Then, we tested RN-JSM performance in simulated and out-of-sample field data.
- High abundance-driven heterogeneity in detection (i.e. common and secretive species) limited the ability of the MC model to quantify covariate effects; especially, when the number of visits was low. Both models (MC and RN), often failed to detect existing effects when data were overdispersed. Moreover, the RN model consistently lacked sufficient power when analysing data from uncommon species (even when simulations and model specifications perfectly matched). This problem was solved by our RN-JSM, which yielded more precise and accurate estimates of species environmental responses. Increased accuracy in rare species held when the RN-JSM was tested with real and out-of-sample datasets.
- In the light of our results, we propose: (i) for common and secretive species analyse occupancy data with the RN model and prioritize revisiting sites; (ii) for species that may have overdispersed detectability or local abundances (e.g. with correlated behaviours or occurring in clusters), apply RN extensions that account

for this extra-variation (e.g. Poisson-beta or zero-inflated models). Finally, (iii) for uncommon species (mean abundances <1), whenever possible, gather data at the community level and apply joint species modelling techniques.

KEYWORDS

environmental responses, heterogeneous detection, joint species modelling, occupancy models

1 | INTRODUCTION

Estimating species distribution across the landscape is relevant from an ecological and conservation perspective (Guillera-Arroita et al., 2015). By explicitly considering how species respond to the environment as a separate process from our imperfect observations, occupancy models allow us to infer species distribution with affordable sampling designs (e.g. detection–non detection records vs. marking individuals; Mackenzie & Royle, 2005). Moreover, acknowledging imperfect detection improves model estimates of species occupancies and metapopulation dynamics (Guillera-Arroita et al., 2014; Lahoz-Monfort et al., 2014; MacKenzie et al., 2002). To date, one of the most widely used models to infer species distribution is MacKenzie's occupancy-detection model (MacKenzie et al., 2002; MC, hereafter). Survey designs that revisit sites within a timeframe, where changes in site-specific occupancies are unlikely, are used to estimate the probability of detecting a species at a site given that is present. A potential limitation of MC is that detection probability is assumed to be homogeneous or fully explained by environmental covariates. However, we expect that detecting the presence of a species would be easier at sites with many individuals (i.e. intrinsic heterogeneity in detection). In this sense, Royle and Nichols (2003), proposed an occupancy model (RN, hereafter), where local abundance is related in the simplest way to detectability, assuming constant and independent detection among individuals. This approach can improve the accuracy of occupancy estimates (Dorazio, 2007; Royle & Nichols, 2003); although may not be adequate when detectability varies across individuals (Veech et al., 2016) or they are not detected independently (e.g. movement of individuals in groups, Martin et al., 2011). As in any model, violations of occupancy model assumptions can yield biased or inaccurate estimates, affecting the reliability of our inferences (Bailey et al., 2014).

The limitations of the MC model to account for abundance-driven heterogeneous detection have been largely recognized, but most studies have focused on potential errors in estimates of landscape occupancies or demographic parameters (e.g. Dorazio, 2007; Rossman et al., 2016; Royle & Nichols, 2003); while species responses to the environment have been largely overlooked. However, the MC model is frequently used to infer covariate effects on species distribution (Guillera-Arroita & Lahoz-Monfort, 2012); which can help to assess their vulnerability to disturbances or to evaluate management actions (MacKenzie, 2005). Thus, inaccurate estimates can lead to misguided management decisions (Kery & Schmidt, 2008).

In general, optimal sites for a species hold a greater number of individuals, and hence, we would expect habitat suitability and site-specific detectability to covary. It is well-established that covariation between environmental variables and detectability can strongly bias our inferences if such relationship is not explicitly modelled (Gu & Swihart, 2004; Lahoz-Monfort et al., 2014). Therefore, not accounting for abundance-driven heterogeneous detection may affect our inferences about how species respond to the environment. In principle, MC can be extended to include site-specific detection as a function of covariates affecting abundances (Mackenzie, 2006), but this approach could be noisier than explicitly linking abundance and detectability. Hence, MC model performance may be compromised even when the relationship between detection and habitat suitability is acknowledged.

Another common source of error in occupancy models is insufficient information in data (Barry & Elith, 2006). If the number of detections across sites is too scarce, estimates may be imprecise and unable to detect existing effects of covariates (Guillera-Arroita & Lahoz-Monfort, 2012). A way to overcome this problem is designing surveys that optimize the number of sites versus revisits according to species commonness and conspicuity (Mackenzie & Royle, 2005). Nonetheless, sampling efforts for rare or cryptic species may be impractical (e.g. >100 sites and 6–8 visits, Guillera-Arroita, 2017). In this context, multi-species occupancy models can help to overcome such logistic constraints. Within communities, species environmental responses are simultaneously fitted, thus, reducing uncertainty around model estimates (Ovaskainen & Soininen, 2011). In particular, joint species model (JSM) extensions, in which species responses depend on their traits and phylogeny, represent a promising tool since species coefficients borrow information from those with similar traits or close phylogenetic relatives (e.g. HMSC models, Ovaskainen et al., 2017). To evaluate the potential of this approach, we developed and tested a Royle and Nichols (2003) joint species model (RN-JSM) in which species responses and detectability depend on their traits and phylogeny. This, should provide better estimates in infrequently observed species with feasible sampling efforts.

In this paper, we evaluate the sensitivity of occupancy models to two common sources of error: model misspecifications due to heterogeneous detection and insufficient information in data. First, we assess the reliability of the MC model in scenarios of abundance-driven heterogeneous detection. Then, we evaluate the sensitivity of MC and RN models to overdispersed local abundances and detectability. Finally, we present a new RN-JSM and test its performance

with a simulation study and a dataset on grasslands bird communities from Uruguay.

2 | MATERIALS AND METHODS

2.1 | Sampling designs, occupancies and detection rates reported in literature

To design simulation experiments with sampling efforts similar to those encountered by practitioners, we searched for studies inferring species environmental responses by means of occupancy models. In the ISI Web of Science database, we searched for 150 papers targeted on birds, mammals or reptiles ($N = 50$ per taxon), using combinations of key words (occupancy AND cover AND bird*; occupancy AND habitat AND mammal*/reptile*). For each taxon, we ordered studies according to their relevance with respect to key words and number of citations. Then, we selected the first ones that used single-species occupancy models to infer species environmental responses. For each study, we recorded the sampling design (number of sites and visits), species detectability and occupancy (minimum and maximum when multiple species were monitored; mean values otherwise). Finally, we calculated median values of these parameters (as well as 0.05 and 0.95 quantiles) across studies (per taxon and in total).

2.2 | Performance of classical occupancy models

To evaluate the performance of the classical MC occupancy model in quantifying species environmental responses (MacKenzie et al., 2002), we simulated data following the occupancy model developed by Royle and Nichols (2003) and two variants that accounted for overdispersion in local abundances and variability in detection. In all cases, we compared the MC model performance against that of the RN model. We chose the RN model as a benchmark because in its formulation local abundances depend on a Poisson process, which represents a standard null model for species distributions (Royle et al., 2005). In addition, it implies that most common species also show higher variability in site-specific abundances, a pattern frequently found in nature (He & Gaston, 2003). Finally, RN links local abundance and detectability in a simple but intuitive way. Species detection increases with the number of individuals at a rate that depends on how conspicuous they are (Royle & Nichols, 2003).

2.2.1 | Simulating survey data

In scenario 1, we simulated local abundances from a Poisson distribution whose mean depended on a centred environmental covariate, conducted virtual surveys and recorded whether the species was detected or not during visits:

$$\log(\lambda_i) = \beta_0 + \beta_1 x_i \quad (1)$$

$$N_i \sim \text{Poisson}(\lambda_i) \quad (2)$$

$$p_i = 1 - (1-r)^{N_i} \quad (3)$$

$$y_i \sim \text{Binomial}(p_i, M_i) \quad (4)$$

The expected abundance of the focal species at the i -th site (λ_i) depended on its mean abundances (according to β_0) and the effects of environmental conditions ($\beta_1 x_i$). The number of individuals at a given site (N_i), was drawn from a Poisson distribution with λ_i . The site-specific probability of detection of the species (p_i) depended on the detection probability of each individual (r) and the number of individuals at each site (N_i). The number of times the species was detected at the i -th site (y_i) was a function of its probability of detection (p_i) and the number of visits to the site (M_i). According to this model, the site-specific probability of occurrence of the species at the i -th site is defined as $\psi_i = 1 - \exp(-\lambda_i)$.

To evaluate the performance of RN and MC when detectability and local abundances are overdispersed, we simulated two additional scenarios. In scenario 2, the detectability of individuals in the i -th site and m -th survey was drawn from a Beta distribution with expected values R and an overdispersion parameter τ ($r_{im} \sim \text{Beta}(R\tau, (1-R)\tau)$, Dorazio et al., 2013). For each expected value of detectability R (see below), τ parameter was adjusted so that the Beta distribution had a coefficient of variation of 0.5. This variability reflects values found in nature (e.g. manatees, Dorazio et al., 2013). In scenario 3 we accounted for overdispersed local abundances. The number of individuals at the i -th site (N_i) was sampled from a negative-binomial distribution with mean λ_i and a dispersion parameter Φ . For each lambda, we adjusted Φ so that the variance equaled two times the mean. This represents a degree of dispersion found in nature that can compromise RN model performance (Duarte et al., 2018). In addition, it provides reasonable values of local abundances.

In all three scenarios (Poisson, Poisson-beta and negative-binomial), we simulated datasets with species differing in their mean abundances (according to e^{β_0} , Equation 1) and detectability (r). Assumed levels of mean abundances were 0.05, 0.1, 0.2, 1, 2, 4 and 6. Simulated levels of landscape occupancy ranged from 0.08 to 0.94 (Poisson distribution for abundances) and from 0.05 to 0.91 (negative-binomial distribution). We assumed individual detectability of 0.15, 0.45 and 0.75. By combining both factors, we obtained a wide range of species occupancy and detection rates. Species responses to the environment (β_1) were drawn from a mixture of two normal distributions (equal proportion each) with means 1 and -1 and a standard deviation of 1. This allowed us to assess the ability of occupancy models to infer covariate effects, avoiding biases due to sign. x_i was sampled from normal distribution (mean = 0, $SD = 1$). For every abundance-detectability combination, we performed simulations using four types of sampling designs: (i) 150 sites and 6 visits, (ii) 150 sites and 3 visits, (iii) 75 sites and 6 visits and (iv) 75 sites and 3 visits. The first survey design generates an extensive

dataset, but that may be logistically challenging to obtain. The last one, a much less labour-intensive option that may hold too little information. Designs ii and iii, correspond to intermediate sampling efforts with the same number of surveys (450), but in which surveying a larger number sites or performing more visits is prioritized.

2.3 | Model fit and estimation of errors

For each scenario (Poisson, Poisson-beta and negative-binomial) and combinations of parameters and survey designs ($N = 252$) we simulated 100 replicated datasets. For each dataset we fitted the RN model, and two variants of the MC model. One with fixed species detection (p) across sites; and another, in which detectability depended on the same environmental covariates affecting abundances ($\text{logit}(p_i) = \alpha_0 + \alpha_1 x_i$). For each simulated dataset, we tracked site-specific probability of occurrence (ψ_i) and the true value of β_1 . First, we evaluated the accuracy of model estimates on site-specific probability of occurrence. To this end, we calculated the root-mean-square error between fitted ($\hat{\psi}_i$) and true (ψ_i) values of occupancy; $RMSE = \sqrt{\frac{1}{I} \sum_1^I (\hat{\psi}_i - \psi_i)^2}$, being i each site. Then, we evaluated the ability of models to detect and quantify species environmental responses. For each parameter set we calculated power as $1 - \text{Type II error rate}$ across 100 repetitions. Type II errors were found when models could not detect an existing effect (the 95% credible interval for β_1 included zero). We considered a power of 0.8 or more as adequate (Guillera-Aroita & Lahoz-Monfort, 2012). In addition, we regressed estimated site-specific probability of occurrence against true values ($\hat{\psi}_i$ vs. ψ_i). The intercept of this regression allows detecting systematic biases in estimations and the slope provides measures of under- or overestimation of effect sizes. We did not directly compare true versus estimated β_1 due to the different link functions used in RN and MC models (see section 1 Appendix C).

The Poisson scenario represents optimal conditions for a RN occupancy analysis as its specifications reflect the exact processes that generated the data (i.e. all statistical assumptions are met). In this scenario, RN performance serves as a baseline to detect problems due to insufficient information in data. MC versus RN comparisons allow us to detect errors due to model misspecification on abundance-driven heterogeneous detection (only fits from MC model are inaccurate or imprecise). In both models, results from Poisson-beta and negative-binomial scenarios allowed us to assess sensitivity to overdispersed data.

3 | DEVELOPMENT AND EVALUATION OF RN JOINT SPECIES MODEL

To improve model accuracy for rare or secretive species we designed a RN-JSM model. As in Royle and Nichols (2003), the site-specific number of individuals is sampled from a Poisson distribution whose mean depends on environmental covariates. Local abundance, in turn, modifies detectability (Equations 1–4). According to the

Hierarchical Modelling of Species in Communities (HMSC) framework (Ovaskainen et al., 2017), species responses to the environment (β) are sampled from a multivariate normal distribution with mean μ and variance-covariance matrix Σ :

$$\beta \sim N(\mu, \Sigma) \quad (5)$$

$$\mu_{jk} = \gamma_{0k} + \sum_{t=1}^T \gamma_{tk} l_{tj} \quad (6)$$

$$\Sigma = \mathbf{V} \otimes [\rho \mathbf{C} + (1 - \rho) \mathbf{I}] \quad (7)$$

The expected response of the j -th species to the k -th covariate (μ_{jk}) depends on an intercept (γ_{0k}) and on the effects of T different species traits on environmental responses ($\sum_{t=1}^T \gamma_{tk} l_{tj}$). To obtain the variance-covariance matrix Σ , the variance-covariance of the K environmental responses (\mathbf{V}) is expanded via the Kronecker product with $[\rho \mathbf{C} + (1 - \rho) \mathbf{I}]$; where \mathbf{C} is the matrix of species phylogenetic correlations and \mathbf{I} is an identity matrix of $J \times J$ size. Parameter ρ can take a value between 0 and 1 and represents the strength of the phylogenetic signal on species environmental responses. In our RN-JSM model, species detectability (r_j) was also modelled as a function of species traits and phylogeny (following a joint species approach).

3.1 | Performance of RN-JSM model with simulated data

To evaluate the performance of our RN-JSM we generated communities composed of 20 species. Local abundance of species was a function of an environmental covariate (X_{1j}), sampled from a normal distribution (mean = 0, $SD = 1$). Expected environmental responses of species (μ , Equation 5) depended on one species trait which was drawn from a normal distribution (mean = 0, $SD = 1$). Trait effects (γ_{tk}) were sampled from a mixed normal distribution with mean 1 or -1 and a standard deviation of 1 (equal proportion each). When generating Σ , the species phylogenetic matrix was constructed with a random phylogenetic tree built with the ape R package (Paradis & Schliep, 2019). The strength of the phylogenetic signal (ρ) was sampled from a uniform distribution with a range between 0.2 and 0.8 (low and high signal, avoiding extreme values). The variance-covariance matrix of environmental effects (\mathbf{V}) was constructed as the quadratic form of their variance vector and a correlation matrix. Variance was sampled from a uniform distribution (min = 0.1, max = 0.5) and correlation from a normal distribution with mean 0 and sd 0.25 (discarding values beyond |1|). Simulated species detectability (r_j), depended on a single species trait (sampled from a normal distribution, mean = 0, $SD = 1$); and trait effects (drawn from a mixed normal distribution, mean = $1/-1$, $SD = 1$). For each simulated community, we conducted virtual surveys following i-iv sampling designs (see above).

In each simulated community, we randomly selected a rare and a common species (occupancy <0.1 and ≥ 0.5 respectively) and estimated their environmental responses using RN and RN-JSM. Detectability in selected species ranged between 0.1 and 0.75. After model parameterization, we plotted true against estimated environmental responses (mean and credible intervals). The position of mean values (with respect to 1:1 line) determines the accuracy of model estimates. The breadth of the posterior distribution informs about certainty in model predictions (precision).

3.2 | Case of study: Grassland bird community from Uruguay

To illustrate the performance of the RN-JSM, we applied single and multi-species models to data from a community of 50 bird species from managed pastures in Uruguay (Aldabe et al., in prep). A total of 46 cattle farms were sampled. Farms differed in the region where they were located (low vs. highlands) and in their pasture height (cm) and type (natural vs. artificial). Within each farm, four to nine transects (mean = 6) of 300m length were established with an average spacing between them of 501m ([71, 1734] 0.05 and 0.95 quantiles). Bird communities were sampled on yearly basis from 2015 to 2018 during the austral summer (December–March). Each year, transects were visited between one and four times (with a time-lag of 2 to 4 days between visits) and observers noted all bird species detected. Fieldwork was performed in private cattle farms with permission from the landowners.

In RN-JSM, the abundance of the j -th species in the i -th transect from the f -th farm was modified by the region in which the farm was located and the local environmental conditions of the transect. In particular, the mean abundance of the j -th species at the f -th farm (according to β_{0jf}) was modelled with a normal distribution whose mean depended on whether the farm was located on a valley or a highland region ($\beta_{0jf} \sim N(\nu_{0j} + \nu_{1j}HI_f, \sigma_0^2)$, $HI_f = 1$ highland, 0 otherwise) and a σ_0^2 variance. Local abundance was then modified by pasture height (PH_i) and type (L_i , natural vs. artificial) within the i -th transect ($\log(\lambda_{ji}) = \beta_{0jf} + \beta_{1j}PH_i + \beta_{2j}L_i$). Species responses to the environment ($\nu_0, \nu_1, \beta_1, \beta_2$) were modelled with a multivariate normal distribution whose mean values depended on species body size (log-transformed) and their degree of insectivory (measured as percentage of consumption of insects in their diets). Detectability, in turn, depended on species body size and vocalization type (low or high frequency). Quantitative environmental covariates and traits were scaled (mean = 0, SD = 1) previous to the analyses. We obtained species functional traits from the Elton trait database (Wilman et al., 2014), except for vocalization rate which was classified according to Billerman et al. (2020) and JA personal experience. The phylogenetic correlations matrices were constructed from phylogenetic trees obtained from Birdtree (Jetz et al., 2014) using the R package APE (Paradis & Schliep, 2019). When fitting the RN model (Equations 1–4), we followed the same scheme for covariate effects but detectability and environmental responses were estimated separately for each species and independently of traits and phylogeny.

To evaluate the ability of both models (RN and JSM-RN) to predict occupancy at a new survey, we tested their accuracy in an out-of-sample dataset. We used field data from 2017 and 2018, which contained a higher number of revisits, for model training; and data from 2015 and 2016 to measure prediction error. For each species present in the community, we sampled 1,000 values from the joint posterior of species environmental responses and detectability (from models fitted with 2017–2018 data) and simulated data for 2015 and 2016 (i.e. species-specific occupancy and detection during surveys). We quantified root-mean-square errors of observed occupancies predicted by our model and box plotted them according to species rarity (proportion of sites detected: rare ≤ 0.1 , intermediate [0.1, 0.5] and common ≥ 0.5). In addition, in four focal species with contrasting abundance and detectability, we evaluated the precision of model estimates on covariate effects (pasture height and land use), and the accuracy of predictions. Focal species were: (i) *Vanellus chilensis* (common and conspicuous); (ii) *Sicalis luteola* (common and cryptic); (iii) *Xolmis dominicanus* (uncommon and conspicuous); and (iv) *Donacospiza albifrons* (uncommon and cryptic).

In all cases (simulated datasets and study case) we used a Bayesian approach with weak priors. We ran between $1e+05$ and $5e+05$ iterations in three chains and checked for convergence ($Rhat < 1.1$) and effective sample size (≥ 100). See Appendix B, for further information on parameterization of models and priors. Single-species models were fitted using JAGS (Plummer, 2003) and RN-JSM using Stan (Stan development team, 2020) programs. We fitted RN-JSM in Stan because it allowed us to use the Kronecker product to obtain the variance–covariance matrix of species environmental responses (Equation 7).

4 | RESULTS

4.1 | Sampling designs in observational studies

Across taxa, our sample of field surveys had median values of 100 sites and 3 revisits, although there was a high variability across studies ([37, 317] and [2, 9], [$q_{0.05}, q_{0.95}$] for sites and visits respectively). Within taxa, the number of sites surveyed went from 59 to 100 (median values for reptiles and birds respectively); and the number of visits from 3 to 7 (median values for birds and mammals respectively). Regarding landscape occupancies, maximum values went from 0.66 and 0.70. In the case of detection probabilities, minimum values ranged between 0.19 and 0.23, whereas maximum values went from 0.51 to 0.73 (Appendix A, Figure A1, Table A1).

4.2 | Assessment of MC and RN model performance

In both models, estimates of site-specific probability of occupancy (ψ_i) were less accurate for common species (mean abundance >1 and landscape occupancy >0.65). In such conditions, errors of the MC model were higher than RN (RMSE values >0.1). Especially, in

simulated surveys with three visits to sites, where RMSE of occupancy estimates were on average 1.9 times higher in MC than in RN (Figure 1, lower vs. upper panels).

Regarding environmental-driven occupancy gradients, both models (RN and MC), could correctly infer their direction (slope of $\hat{\psi}_{ij}$ vs. ψ_{ij} regression >0), but tended to overestimate their size in abundant but secretive species (mean abundance ≥ 4 , $r = 0.15$) (Figure 2). This pattern was stronger in the MC model; especially, when only three visits were conducted in simulated surveys (Figure 2 upper vs. lower panels). The MC model extension in which site-specific detections depended on covariates affecting local abundances, did not perform better for secretive species. It systematically underestimated covariate effects (Appendix D, Figure D1) and could not attain sufficient power (Power <0.8 , Figure D2).

In both models, estimates of covariate effects ($\hat{\beta}_1$) became less precise in uncommon species (mean abundances <1 ; Appendix D, Figure D4). As a result, their posteriors overlapped zero more often and power was reduced. In uncommon species, MC outperformed RN (Figure 3, upper vs. lower panels), although frequently it did not attain sufficient power (Table D2). When detectability among individuals varied (Poisson-beta scenario), in common and conspicuous species neither MC or RN reached sufficient power (Power <0.8 ;

Figure 3). Nor, in rare species with extra-variation in their local abundances (Figure 3, negative-binomial scenario).

4.3 | Evaluation of RN joint species model (RN-JSM)

In simulated communities, for rare species the precision of RN estimates decreased when the number of sites was 75 or were only surveyed three times (Figure 4b–d), while it was maintained for RN-JSM (Figure 4f–g). Results from the grassland bird communities from Uruguay showed similar patterns. When species were rare, RN-JSM resulted in far more precise predictions (Figure 5, left panels), that were also more accurate (Figure 5, right panels). In addition, for rare species, species environmental responses had low precision in the RN model, while RN-JSM strongly reduced uncertainty in model estimations (Figure D5).

5 | DISCUSSION

In general, our results show that the MC model can infer site-specific occupancies when abundance-induced heterogeneity in

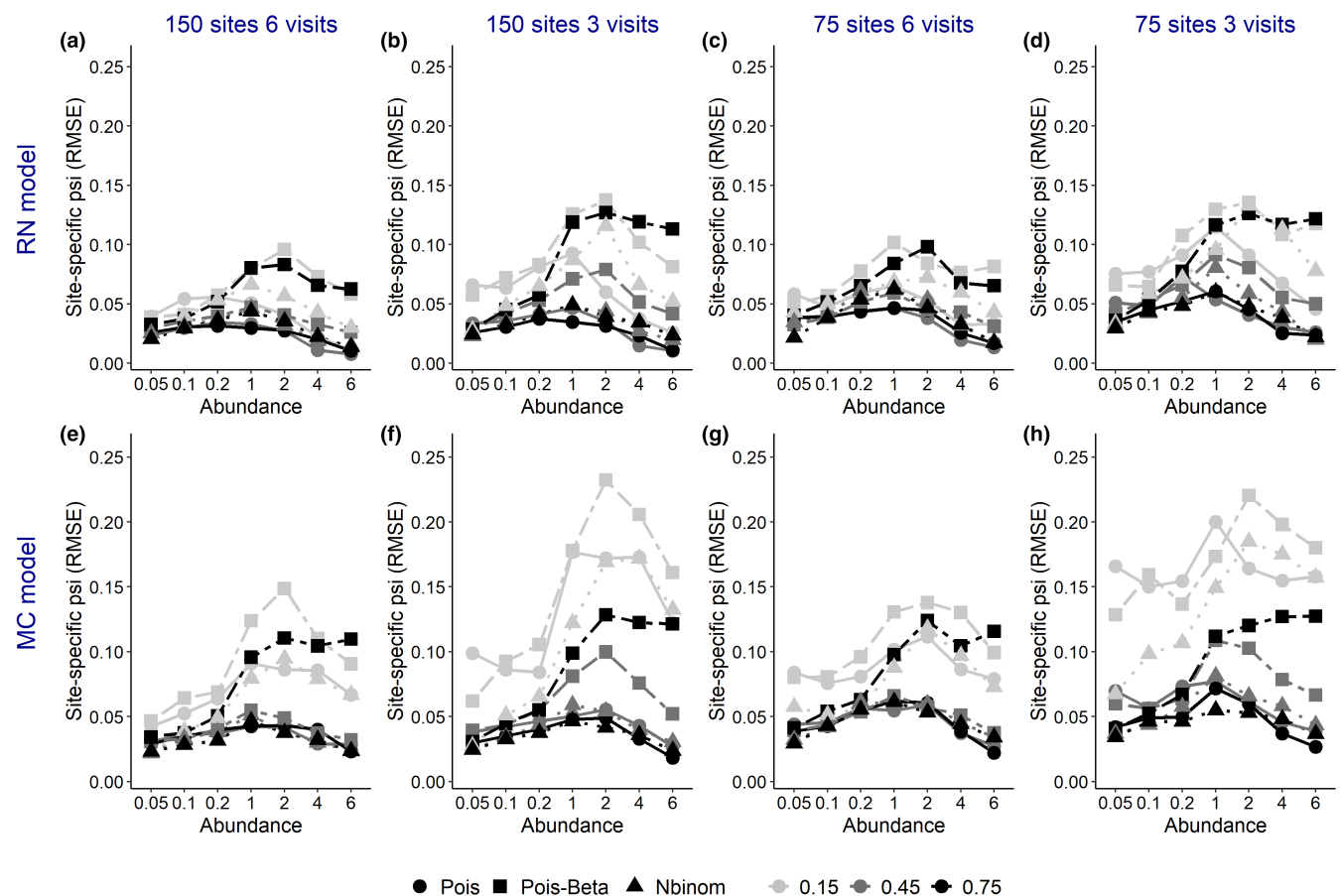


FIGURE 1 Root-mean-square error (RMSE) between estimated and true site-specific probability of occupancy (ψ_{ij}). Results from RN (a–d) and MC (e–h) models in the four types of sampling designs. Simulated species differed in their abundance and detectability. For each abundance-detectability combination RMSE was calculated median values across 100 simulated datasets. Colours depict the probability of detection of individuals (r , in Equation 3)

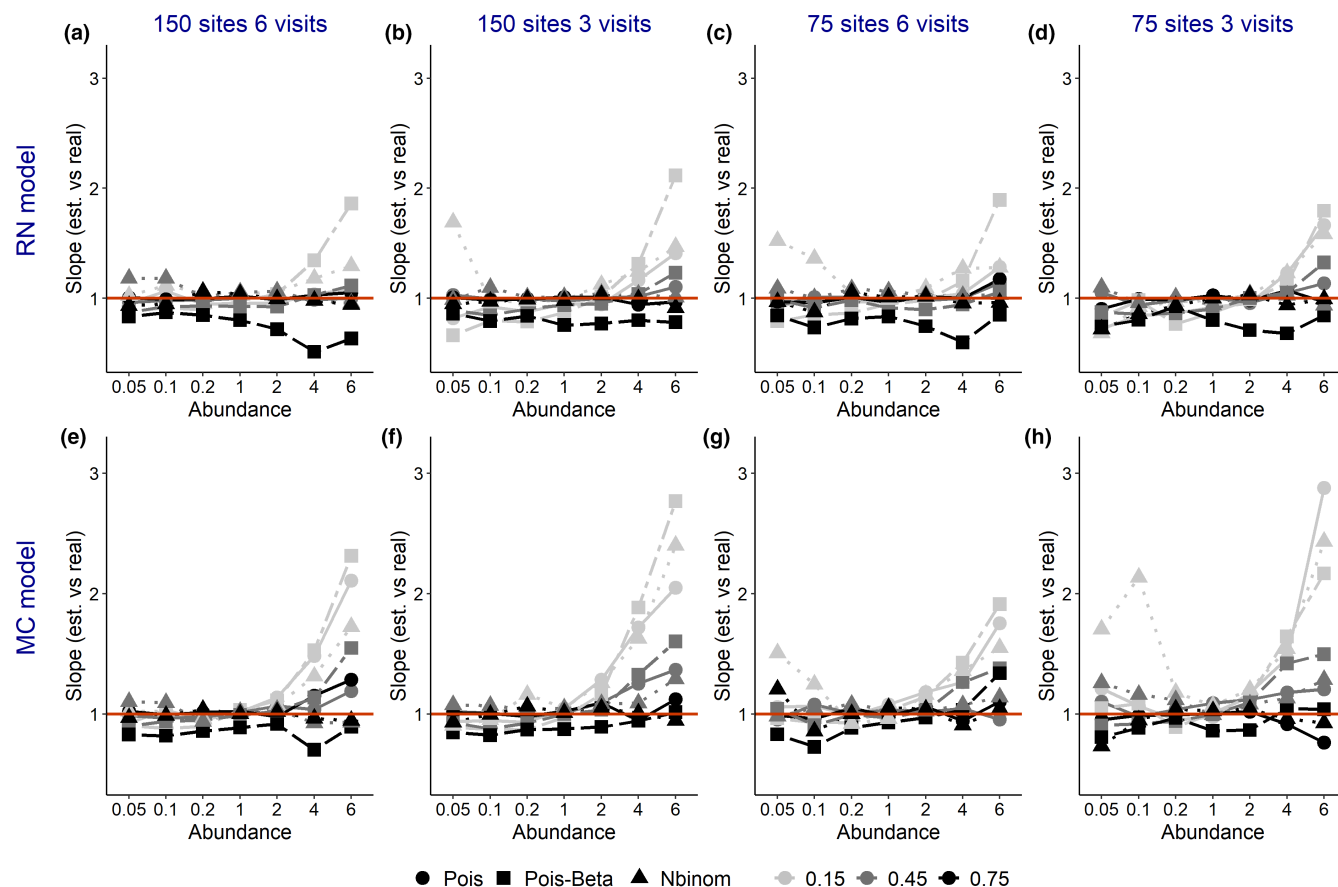


FIGURE 2 Regression slope between estimated and simulated site-specific probabilities of occupancy. Results from RN (a–d) and MC (e–h) models in the four types of sampling designs. Simulated species differed in their abundance and detectability. For each abundance–detectability combination, reported slopes were calculated as the median across 100 simulated datasets. Red line depicts a perfect match between estimated and simulated values (slope = 1). Colours depict the probability of detection of individuals (r , in Equation 3)

detection is low (i.e. rare or conspicuous species). Otherwise, its accuracy decreases (i.e. the effects of environmental covariates are overestimated (Figures 1 and 2). In species that are both common and secretive, the spatial variability of local abundances is high (He & Gaston, 2003) and detectability remains imperfect (Royle & Nichols, 2003). As a result, abundance-driven heterogeneity in detection increases (Dorazio, 2007), compromising the performance of MC. We expected that modelling species detectability as a function of covariates affecting abundances would improve its accuracy. However, this was not the case (Appendix D). In secretive species, MC extension underestimated species environmental responses (Figure D1) and had a limited ability to detect existing covariate effects (Figure D2). Modelling differences in detection by means of covariates seem to be noisier than directly accounting for abundance effects on detection. Besides, it may not reflect the shape of abundance–detection relationships (logistic vs. saturating response). Therefore, our results call for caution when assuming that including environmental effects in the observation process is sufficient to overcome problems arising from abundance-driven heterogeneous detectability.

In line with previous work, we found that the performance of MC can be improved through field sampling designs (Mackenzie &

Royle, 2005). For instance, MC provided more accurate estimates when local abundances were low (<1 individual per site) and differences across sites diminished. Since local abundances are scale dependent, accommodating the size of plots to unify them can be a strategy that homogenizes detectability across sites, improving the MC model performance (e.g. plot sizes similar to species home-ranges, Efford & Dawson, 2012). Nevertheless, in many cases there is no a priori information about local densities or mobility and these properties frequently change in response to the environmental covariates being tested. Another option is to optimize the number sampling units or revisits during surveys (Guillera-Arroita, 2017). For the same sampling effort, MC accuracy increased with the number of visits (Figures 1 and 2, f vs. g). In abundant and secretive species, detectability is a parameter with high uncertainty and new visits provide a lot of information about it. Therefore, in line with previous works (Guillera-Arroita & Lahoz-Monfort, 2012; Mackenzie & Royle, 2005), for common but secretive species we recommend investing more in revisiting sites than in increasing their number.

The fact that RN outperformed MC was not surprising since it better reflected the structure of simulations. However, our results provided valuable information about its sensitivity to overdispersed data. In some scenarios, extra-variation in detectability and local

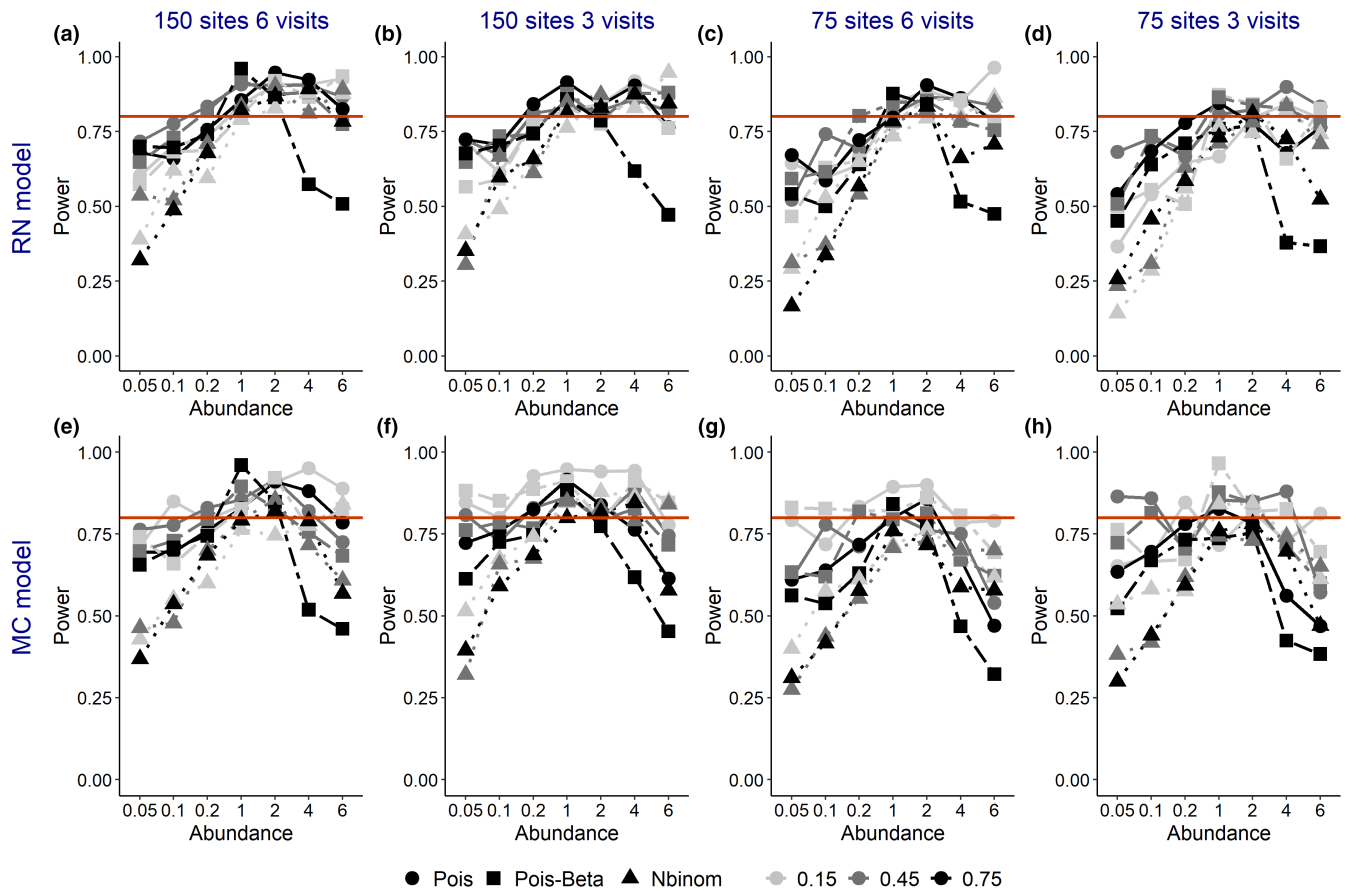


FIGURE 3 Power analyses of models for detecting existing environmental responses. For each abundance-detectability combination, power was calculated as $1 - \text{Type II errors}$ across 100 simulated datasets. Red line depicts a power of 0.8 (considered adequate). Results from RN (a–d) and MC (e–h) models in the four types of sampling designs. Colours depict the probability of detection of individuals (r , in Equation 3)

abundances, strongly reduced the power of RN (Figure 3). On one hand, variability in detection blurred the effects of local abundances on site-specific detections. Consequently, it often failed to detect covariate effects (Figure 3, Poisson-beta scenario); especially, in common and conspicuous species. On the other hand, when local abundances were overdispersed, RN was unable to accommodate for such extra-variation, leading to increased uncertainty in model estimates (Figure D4); thus, reducing its power (Figure 3, negative-binomial). This was particularly the case for uncommon species, in which information per se was scarce. Overall, these results agree with previous work showing that assuming fixed and independent detectability of individuals (as in classical N-mixture model for occupancy data, Royle & Nichols, 2003 or counts, Royle, 2004) can result in poor model performance when individuals are spatially clustered (Joseph et al., 2009), detected in groups (Martin et al., 2011), or not equally conspicuous (Veech et al., 2016). In such cases, drawing detectability from a beta distribution (Martin et al., 2011) and local abundances from zero-inflated or hurdle models (Dorazio et al., 2013; Joseph et al., 2009) may deal with these overdispersion problems. It is important to note, however, that accommodating overdispersed detection probabilities may lead to identifiability issues (i.e. different models fitting data equally well but providing

contrasting estimates of species occupancies). In such cases, it is advisable to design field surveys so that detectability is maximized and its heterogeneity reduced or explained by covariates (Royle, 2006).

Noteworthy, when model specifications matched those of simulations, RN did not achieve sufficient power (0.8) when species had, on average, less than one individual per site (Figure 3, Poisson scenario). Moreover, extensive sampling designs (150 sites, 6 visits) did not avoid this problem. In those cases, MC outperformed RN since local abundances were similar across sites and its formulation was simpler. Still, MC often failed to detect covariate effects (Appendix D, Table D2). Poor performance of single-species models for uncommon species results particularly relevant in the light of our literature search. Average local abundances of 1 correspond to landscape occupancies of 0.63 (according to $\psi = 1 - e^{-\lambda}$). These values are representative of maximum occupancies reported in field studies, which usually have more limited sampling designs than our simulated surveys (Appendix A, Figure A1). Overall, these results suggest that field studies may have had insufficient power to detect covariate effects of a similar size (or lower) to that of our simulations (mean values of $\beta_1 = |1|$). In this context, the closed formulation of power analysis for occupancy models developed by Guillera-Aroita and Lahoz-Monfort (2012) can be useful. It allows, without simulating

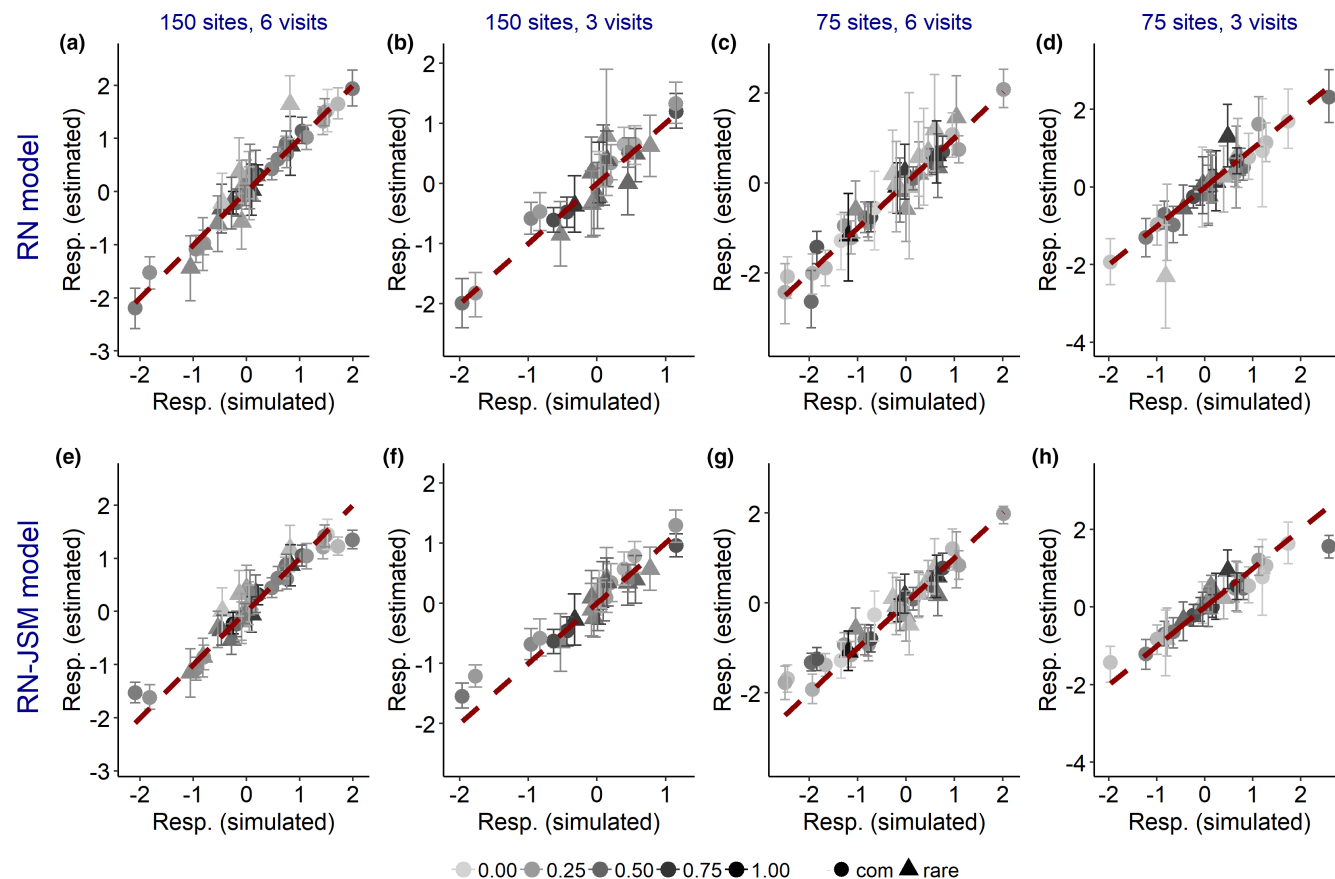


FIGURE 4 Simulated versus estimated environmental responses (β_1) in simulated communities ($N = 20$). In each community we selected a rare and a common species (landscape occupancy <0.1 and ≥ 0.5 respectively). Detectability in selected species ranged between 0.1 and 0.75. Points represent posterior means and bars depict credible intervals. Red dashed line shows the 1:1 relationship. Focal species were classified according to their abundance: rare (<0.1 of occupancy, triangles) and common (≥ 0.5 of occupancy, circles). Species detectability ranged between 0.1 and 0.85 (grey palette)

data, to calculate the number of sites and revisits needed to detect an expected environmental effect. This may be particularly relevant when conclusions drawn from occupancy models guide management decisions (MacKenzie, 2005).

The RN-JSM model presented here significantly increased the accuracy and precision of our estimates when information was scarce (Figure 4). Our model builds on previous multi-species N-mixture models, where species environmental responses (Tobler et al., 2019; Yamaura et al., 2011), detectability (Gomez et al., 2018; Tobler et al., 2015) or both (Yamaura, Connor, et al., 2016; Yamaura, Kery, & Royle, 2016) are modelled in a hierarchical way, from common distributions. Our main contribution with respect to previous approaches is that, following the HMSC framework (Ovaskainen et al., 2017), expected species responses and detectability are not treated as random effects, but rather depend on the species' functional traits and phylogeny. Species fundamental niches depend on their functional traits (McGill et al., 2006), which often reflect a shared evolutionary history (Revell et al., 2008). In addition, detectability of species arises from their behaviour (e.g. vocalization rate) and distinctiveness (e.g. size, colour; Iknayan et al., 2014). Therefore, RN-JSM model should provide reliable estimates with

more limited sampling efforts and reduce the risk of pulling rare species responses (for which information is scarce) towards community means. Accordingly, when tested in real data, RN-JSM showed an enhanced performance with respect to single-species models. For uncommon bird species, it strongly reduced the uncertainty about species responses (Appendix D, Figure D5) and predictions of landscape occupancy were by far more accurate (Figure 5). Enhanced performance of our model in the study case is important as in field studies we usually do not have a priori information about the processes generating data. Moreover, the enhanced forecasting accuracy of RN-JSM model in out-of-sample data provided solid evidence of the suitability of our approach.

As in any occupancy model (Bailey et al., 2014), we expect that RN-JSM will be sensitive to violations of its assumptions. For instance, overdispersed abundance or detectability may reduce its accuracy, as it occurred with the single-species approach. In this sense, RN-JSM could be extended to account for extra-variation (e.g. by modelling detectability with a Beta distribution and abundances from a hurdle model, Dorazio et al., 2013). In addition, choosing an adequate set of species traits will be critical to avoid pulling rare species model estimates towards community means. To ensure reliable

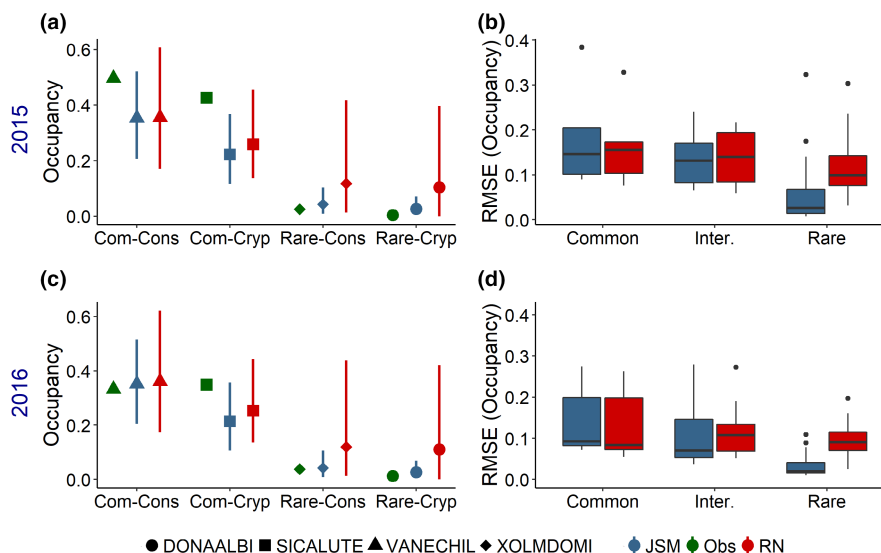


FIGURE 5 Evaluation of forecasting ability of RN and RN-JSM models using out-of-sample data from 2015 to 2016 (upper and lower panels respectively). Left panels: observed occupancy (green) versus predictions of observed occupancy of RN and RN-JSM models (red and blue respectively) in target species. Points show mean values of predictions and bars 0.05 and 0.95 quantiles. Right panels: root-mean-squared error of predicted observed occupancy for common, intermediate and rare species in the studied community (>0.5 , $[0.1, 0.5]$ and <0.1 of occupancy respectively). Acronyms: DONALBI: *Donacospiza albifrons*, SICALUTE: *Sicalis luteola*, VANECHIL: *Vanelus chilensis*, XOLMDOMI: *Xolmis dominicanus*. Cryp: cryptic, Cons: conspicuous, Com: common. Results from RN (a–d) and MC (e–h) models in the four types of sampling designs

inferences, we can take advantage of the ability of RN-JSM to obtain good estimates with reduced sampling efforts. This should facilitate split data and quantify model performance in out-of-sample sets. Within its limitations, we believe that RN-JSM provides a promising statistical tool to improve inferences with affordable and feasible sampling designs. Future work evaluating potential biases in RN-JSM predictions; and if necessary, extending its formulation will help to improve the modelling framework presented here.

6 | CONCLUSIONS

In light of our results, we provide the following recommendations when analysing covariate effects on species occupancy.

1. **For common and secretive species** analyse occupancy data with Royle and Nichols model (2003) and prioritize increasing the number of visits over sampling more sites.
2. **For species with overdispersed local abundances or detectability.** Use model extensions that accommodate extra-zeros or variable detectability among individuals (e.g. Dorazio et al., 2013) since RN and MC model may not attain sufficient power.
3. **For uncommon species** (landscape occupancy <0.65) preferentially use MC, but be aware that information may be too scarce to detect existing environmental effects on species occupancy. Whenever possible, gather data at the community level and apply joint species modelling techniques. Ideally, model species responses and detectability as a function of their traits and phylogeny (e.g. HMSC framework, Ovaskainen et al., 2017).

ACKNOWLEDGEMENTS

We are grateful for the comments raised by referees and the handling editor, which helped to improve this work. The authors are especially grateful to Federico Pérez, Sasha Hackembruck, Agustina Medina, Belén Calvete and Graciela Amorín, who helped at data gathering. Sofía Ruiz-Suarez was beneficiary of a PhD fellowship from the Argentinian Government (CONICET) and Joaquín Aldabe International Programs, US Forest Service travel award. This project was financed by International Programs, US Forest Service, Southern Cone Grassland Alliance Program, Aves Uruguay and BirdLife International and by the Ministry of Science of Argentina (PICT-2018-01566 and PICT-2015-0815).

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/2041-210X.13875>.

DATA AVAILABILITY STATEMENT

Data and code used for the analyses of this manuscript are available at the Dryad Digital Repository <https://doi.org/10.5061/dryad.vq83bk3vp> (Morán-López et al., 2022).

AUTHORS' CONTRIBUTIONS

T.M.-L. conceived the aims of this work, designed simulation experiments and analyses and drafted the first version of the manuscript; S.R.-S. helped at designing simulation studies and deciding error

quantification metrics; J.A. gathered field data, listed main difficulties when evaluating occupancy for birds; J.M.M. lead the project, tutored TML in defining the aims of the work and in designing and improving the concepts of the model. All authors revised the work and approved the final version of the article.

ORCID

Teresa Morán-López  <https://orcid.org/0000-0002-5804-5258>

Joaquín Aldabe  <https://orcid.org/0000-0002-9175-9404>

Juan Manuel Morales  <https://orcid.org/0000-0001-7269-7490>

REFERENCES

- Bailey, L. L., MacKenzie, D. I., & Nichols, J. D. (2014). Advances and applications of occupancy models. *Methods in Ecology and Evolution*, 5(12), 1269–1279.
- Barry, S., & Elith, J. (2006). Error and uncertainty in habitat models. *Journal of Applied Ecology*, 43(3), 413–423.
- Billerman, S. M., Keeney, B. K., Rodewald, P. G., & Schulenberg, T. S. (2020). *Birds of the World*. Cornell Laboratory of Ornithology. <https://birdsoftheworld.org/bow/home>
- Dorazio, R. M. (2007). On the choice of statistical models for estimating occurrence and extinction from animal surveys. *Ecology*, 88(11), 2773–2782.
- Dorazio, R. M., Martin, J., & Edwards, H. H. (2013). Estimating abundance while accounting for rarity, correlated behavior, and other sources of variation in counts. *Ecology*, 94(7), 1472–1478. <https://doi.org/10.1890/0012-9658-94.7.1472>
- 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.
- Efford, M. G., & Dawson, D. K. (2012). Occupancy in continuous habitat. *Ecosphere*, 3(4), 1–15.
- Gomez, J. P., Robinson, S. K., Blackburn, J. K., & Ponciano, J. M. (2018). An efficient extension of N-mixture models for multi-species abundance estimation. *Methods in Ecology and Evolution*, 9(2), 340–353.
- Gu, W. D., & Swihart, R. K. (2004). Absent or undetected? Effects of non-detection of species occurrence on wildlife-habitat models. *Biological Conservation*, 116(2), 195–203. [https://doi.org/10.1016/S0006-3207\(03\)00190-3](https://doi.org/10.1016/S0006-3207(03)00190-3)
- Guillera-Arroita, G. (2017). Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. *Ecography*, 40(2), 281–295.
- Guillera-Arroita, G., & Lahoz-Monfort, J. J. (2012). Designing studies to detect differences in species occupancy: power analysis under imperfect detection. *Methods in Ecology and Evolution*, 3(5), 860–869. <https://doi.org/10.1111/j.2041-210X.2012.00225.x>
- Guillera-Arroita, G., Lahoz-Monfort, J. J., Elith, J., Gordon, A., Kujala, H., Lentini, P. E., McCarthy, M. A., Tingley, R., & Wintle, B. A. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, 24(3), 276–292.
- Guillera-Arroita, G., Lahoz-Monfort, J. J., MacKenzie, D. I., Wintle, B. A., & McCarthy, M. A. (2014). Ignoring imperfect detection in biological surveys is dangerous: A response to ‘fitting and interpreting occupancy models’. *PLoS One*, 9(7), e99571.
- He, F., & Gaston, K. J. (2003). Occupancy, spatial variance, and the abundance of species. *The American Naturalist*, 162(3), 366–375.
- Iknayan, K. J., Tingley, M. W., Furnas, B. J., & Beissinger, S. R. (2014). Detecting diversity: emerging methods to estimate species diversity. *Trends in Ecology & Evolution*, 29(2), 97–106.
- Jetz, W., Thomas, G. H., Joy, J. B., Redding, D. W., Hartmann, K., & Mooers, A. O. (2014). Global distribution and conservation of evolutionary distinctness in birds. *Current Biology*, 24(9), 919–930.
- 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.
- Kery, M., & Schmidt, B. R. (2008). Imperfect detection and its consequences for monitoring for conservation. *Community Ecology*, 9(2), 207–216. <https://doi.org/10.1556/ComEc.9.2008.2.10>
- Lahoz-Monfort, J. J., Guillera-Arroita, G., & Wintle, B. A. (2014). Imperfect detection impacts the performance of species distribution models. *Global Ecology and Biogeography*, 23(4), 504–515.
- MacKenzie, D. I. (2005). What are the issues with presence-absence data for wildlife managers? *Journal of Wildlife Management*, 69(3), 849–860. [https://doi.org/10.2193/0022-541x\(2005\)069\[0849:watiw\]2.0.co;2](https://doi.org/10.2193/0022-541x(2005)069[0849:watiw]2.0.co;2)
- Mackenzie, D. I. (2006). Modeling the probability of resource use: the effect of, and dealing with, detecting a species imperfectly. *The Journal of Wildlife Management*, 70(2), 367–374.
- 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.
- Mackenzie, D. I., & Royle, J. A. (2005). Designing occupancy studies: General advice and allocating survey effort. *Journal of Applied Ecology*, 42(6), 1105–1114. <https://doi.org/10.1111/j.1365-2664.2005.01098.x>
- Martin, J., Royle, J. A., Mackenzie, D. I., Edwards, H. H., Kery, M., & Gardner, B. (2011). Accounting for non-independent detection when estimating abundance of organisms with a Bayesian approach. *Methods in Ecology and Evolution*, 2(6), 595–601.
- McGill, B. J., Enquist, B. J., Weiher, E., & Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology & Evolution*, 21(4), 178–185.
- Morán-López, T., Ruiz-Suarez, S., Aldabe, J., & Morales, J. M. (2022). Data from: Improving inferences and predictions of species environmental responses with occupancy data. *Dryad Data Repository*, <https://doi.org/10.5061/dryad.vq83bk3vp>
- Ovaskainen, O., & Soininen, J. (2011). Making more out of sparse data: Hierarchical modeling of species communities. *Ecology*, 92(2), 289–295.
- Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T., & Abrego, N. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters*, 20(5), 561–576.
- Paradis, E., & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35(3), 526–528.
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing*, Vol. 124, pp. 1–10.
- Revell, L. J., Harmon, L. J., & Collar, D. C. (2008). Phylogenetic signal, evolutionary process, and rate. *Systematic Biology*, 57(4), 591–601.
- Rossman, S., Yackulic, C. B., Saunders, S. P., Reid, J., Davis, R., & Zipkin, E. F. (2016). Dynamic N-occupancy models: Estimating demographic rates and local abundance from detection-nondetection data. *Ecology*, 97(12), 3300–3307.
- Royle, J. A. (2004). N-mixture models for estimating population size from spatially replicated counts. *Biometrics*, 60(1), 108–115.
- Royle, J. A. (2006). Site occupancy models with heterogeneous detection probabilities. *Biometrics*, 62(1), 97–102.
- Royle, J. A., & Nichols, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology*, 84(3), 777–790.
- Royle, J. A., Nichols, J. D., & Kéry, M. (2005). Modelling occurrence and abundance of species when detection is imperfect. *Oikos*, 110(2), 353–359.
- Tobler, M. W., Kery, M., Hui, F. K. C., Guillera-Arroita, G., Knaus, P., & Sattler, T. (2019). Joint species distribution models with species correlations and imperfect detection. *Ecology*, 100(8), 14. <https://doi.org/10.1002/ecy.2754>

- Tobler, M. W., Zúñiga Hartley, A., Carrillo-Percegué, S. E., & Powell, G. V. N. (2015). Spatiotemporal hierarchical modelling of species richness and occupancy using camera trap data. *Journal of Applied Ecology*, 52(2), 413–421.
- Veech, J. A., Ott, J. R., & Troy, J. R. (2016). Intrinsic heterogeneity in detection probability and its effect on N-mixture models. *Methods in Ecology and Evolution*, 7(9), 1019–1028.
- Wilman, H., Belmaker, J., Simpson, J., de la Rosa, C., Rivadeneira, M. M., & Jetz, W. (2014). EltonTraits 1.0: Species-level foraging attributes of the world's birds and mammals: Ecological Archives E095-178. *Ecology*, 95(7), 2027.
- Yamaura, Y., Connor, E. F., Royle, J. A., Itoh, K., Sato, K., Taki, H., & Mishima, Y. (2016). Estimating species-area relationships by modeling abundance and frequency subject to incomplete sampling. *Ecology and Evolution*, 6(14), 4836–4848. <https://doi.org/10.1002/ece3.2244>
- Yamaura, Y., Kery, M., & Royle, J. A. (2016). Study of biological communities subject to imperfect detection: Bias and precision of community N-mixture abundance models in small-sample situations. *Ecological Research*, 31(3), 289–305.
- Yamaura, Y., Royle, J. A., Kuboi, K., Tada, T., Ikeno, S., & Makino, S. (2011). Modelling community dynamics based on species-level abundance models from detection/nondetection data. *Journal of Applied Ecology*, 48(1), 67–75.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Morán-López, T., Ruiz-Suarez, S., Aldabe, J. & Morales, J. M. (2022). Improving inferences and predictions of species environmental responses with occupancy data. *Methods in Ecology and Evolution*, 13, 1540–1551. <https://doi.org/10.1111/2041-210X.13875>