

A Gibbs sampler for multi-species occupancy models

Allan Ernest Clark (✉ allan.clark@uct.ac.za)

University of Cape Town

Res Altwegg

University of Cape Town

Research Article

Keywords: Bayesian multi-species occupancy model, occupancy model, species richness, imperfect detection, reversible-jump Markov chain MonteCarlo

Posted Date: June 13th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1737410/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

A Gibbs sampler for multi-species occupancy models

Allan Ernest Clark^{1,2*} and Res Altwegg^{1,2}

¹*Department of Statistical Sciences, University of Cape Town, Private Bag X3, Rondebosch, 7701, Cape Town, South Africa.

²Centre for Statistics in Ecology, Environment and Conservation, University of Cape Town, Cape Town, South Africa.

*Corresponding author(s). E-mail(s): allan.clark@uct.ac.za;
Contributing authors: Res.Altweg@gmail.com;

Abstract

Multi-species occupancy (MSO) models use detection-nondetection data from several species observed at different locations to estimate the probability that a particular species occupies a particular geographical region. The models are particularly useful for estimating the occupancy probabilities associated with rare species since they are seldom observed when undertaking field surveys. In this paper, we develop Gibbs sampling algorithms that can be used to fit various Bayesian MSO models to detection-nondetection data. Bayesian analysis of these models can be undertaken using statistical packages such as *JAGS*, *Stan* and *NIMBLE*, however, since these packages were not developed specifically to fit occupancy models, one often experiences long run-times when undertaking analysis.

In a single season (single species) nonspatial and spatial occupancy modelling context, [Clark and Altwegg \(2019\)](#), show that special purpose Gibbs samplers can produce posterior chains that mix faster and have larger expected sampling rates ([Holmes & Held, 2006](#)) than those obtained using *JAGS* and *Stan*. These results suggest that such algorithms could potentially lead to significant reductions in the run-times of MSO models.

This paper illustrates how to fit MSO models when the detection and occupancy processes are modelled using logistic link functions and apply these methods to a camera-trapping study undertaken by

[Drouilly, Clark, and O’Riain \(2018\)](#). Variable selection is undertaken using a reversible-jump Markov chain Monte Carlo ([Barker & Link, 2013](#)) algorithm. We found that the Gibbs sampling algorithm developed produces posterior samples that are identical to those obtained when using *Stan*, resulting in faster run-times and has a larger expected sampling rate than *Stan* when analysing the above-referenced data set.

Keywords: Bayesian multi-species occupancy model, occupancy model, species richness, imperfect detection, reversible-jump Markov chain Monte Carlo.

1 Introduction

Species richness (defined as the number of distinct species in a region) is a measure of biodiversity (Dorazio, Royle, Söderström, & Glimskär, 2006) and is important from a conservation and management point of view since ecologists and nature conservationists use these estimates to prioritize conservation action (Gojman, Conroy, Bernardos, & Zaccagnini, 2015; Zipkin, DeWan, & Andrew Royle, 2009). In most cases, species richness in an area is unknown and has to be estimated using observed data. Such data are collected by sampling a region a finite number of times. Some species might be easily observed while others might be more cryptic, thus making it more difficult to draw inferences about the community structure as a whole. Multi-species occupancy (MSO) models account for variable detection probabilities amongst species and are used to investigate the community structure of a group of species (Broms, Hooten, & Fitzpatrick, 2016; Dorazio, Gotelli, & Ellison, 2011; Dorazio & Royle, 2005; Dorazio et al., 2006).

MSO models use detection-nondetection data from several species observed at different locations to estimate the probability that a species occupies a location. The models are particularly useful for estimating the occupancy probabilities associated with rare species (Devarajan, Morelli, & Tenan, 2020). The observed detections of these species are usually low which results in imprecise occupancy probabilities if single species occupancy models are fit to such data. MSO models utilize random effects and allow the parameters associated with the regression effects of the model to be shared. This allows one to estimate detection and occupancy regression effects for species that would otherwise not be estimable (Broms et al., 2016; Tobler, Zúñiga Hartley, Carrillo-Percecastegui, & Powell, 2015).

MSO models are mainly used to obtain estimates of species richness in a certain geographical region but can also be used to compare the biodiversity of different regions (Sauer, Blank, Zipkin, Fallon, & Fallon, 2013; Zipkin, Royle, Dawson, & Bates, 2010). In certain geographical regions, species richness is known or well-understood (e.g. in well-studied areas) in which case these models are also used to uncover the biological drivers of biodiversity in a region (Drouilly et al., 2018; Ruiz-Gutiérrez, Zipkin, & Dhondt, 2010; Zipkin et al., 2009).

Two formulations of the model have been developed in the literature. The first version assumes that **we know the species richness** in a certain geographical region (Broms et al., 2016) and it is assumed that all species of interest have been observed at least once. The second version of the model assumes that **species richness is an unknown random variable** where the aim is to estimate the number of unobserved species (Dorazio & Royle, 2005). Both model formulations are generally undertaken using Bayesian methods. They have extensively been used and computer code to fit these models are freely available online (Broms et al., 2016; Dorazio et al., 2006).

In the literature to date, the modelling process is mostly undertaken using statistical packages such as *WinBUGS* (Lunn, Thomas, Best, & Spiegelhalter,

2000), *OpenBUGS* (Thomas, O'Hara, Ligges, & Sturtz, 2006), *JAGS* (Plummer, 2003), *NIMBLE* (de Valpine et al., 2017) and *Stan* (Carpenter et al., 2017). These packages allow the user to apply Bayesian methods to many different application areas and have not been developed specifically to undertake MSO models. *WinBUGS*, *OpenBUGS* and *JAGS* use a combination of Gibbs sampling (Geman & Geman, 1984), Metropolis Hastings (denoted as 'MH') steps (Hastings, 1970; Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953) as well as other custom Markov Chain Monte Carlo (MCMC) samplers (Brooks, Gelman, Jones, & Meng, 2011) to obtain posterior samples from the posterior distribution of the parameters of a Bayesian model. *NIMBLE* is flexible and allows the user to specify the sampling algorithms that should be used when undertaking posterior sampling. The above software packages with the exception of *Stan*, allow the user to sample from both discrete and continuous random variables. *Stan* in contrast only allows posterior sampling from continuous random variables which are undertaken using the *No-U-turn* Hamiltonian Monte Carlo sampler (Hoffman & Gelman, 2014).

Below we develop Gibbs sampling algorithms that can be used to fit MSO models. The Gibbs sampler is a MCMC method that is used to obtain samples from the posterior distribution of the parameters of a statistical model. The method can be used when all of the conditional posterior distributions of the parameters of a model are of a known form and sampling from the conditional posterior distributions are straightforward. The Gibbs sampler is known to exhibit slow mixing in certain cases although this shortcoming is often overcome by sampling groups of parameters in a block which improves mixing (Roberts & Sahu, 1997). In a single season (single species) nonspatial and spatial occupancy modelling context, Clark and Altwegg (2019), show that special purpose Gibbs samplers can produce posterior chains that mix faster and have larger expected sampling rates (Holmes & Held, 2006) than those obtained using *JAGS* and *Stan*. These results suggest that such algorithms could potentially lead to significant reductions in the run-times of MSO models. Doser, Finley, Kéry, and Zipkin (2021) extended Clark and Altwegg (2019) and developed algorithms to fit various Bayesian occupancy models. As one of their developments, they assume that the spatial effects in the occupancy process of a single-season, single-species occupancy model can be modelled through a Gaussian process prior distribution with a Matérn covariance function. They found that for data sets with more than 1 000 sites, their algorithm was prohibitively slow in which case they used a nearest-neighbourhood Gaussian process prior distribution (Datta, Banerjee, Finley, & Gelfand, 2016) to significantly reduce run-times when analysing large detection-nondetection data sets.

Recently Dorazio and Rodriguez (2012) and Johnson, Conn, Hooten, Ray, and Pond (2013) developed Gibbs algorithms to obtain posterior samples for the parameters of a nonspatial and spatial single species occupancy model respectively. Both approaches assume that detection and occupancy processes

are modelled using probit link functions, which enables the use of data augmentation (Tanner & Wong, 1987) to obtain closed form expressions of the conditional posterior distributions of the parameters of an occupancy model. Below we develop Gibbs sampling algorithms to undertake MSO models when species richness is known and thereafter we describe how the algorithm would change when species richness is unknown. Initially, we assume that the regression effects of the detection and occupancy processes are modelled using probit link functions and thereafter we extend the analysis to the use of logistic link functions. We do so since we observe that in certain model formulations, the use of probit link functions does not lead to a simple Gibbs sampling algorithm when sampling from the posterior distribution of the parameters of an MSO model. In these cases, MH steps, which require the monitoring of various additional tuning parameters, could be used. Another reason for using the logistic link function is that it allows one to more easily interpret the regression effects of the model compared to the probit model. Take note that below we do not discuss joint species distribution models (Pollock et al., 2014; Tikhonov et al., 2020) and restrict the focus to occupancy modelling. A detailed review of MSO and its use can be found in Devarajan et al. (2020).

The paper commences by discussing the Bayesian formulation of an MSO model when species richness is assumed to be known in a geographical region while in Section 3 we develop a Gibbs sampler for the MSO model when species richness is unknown. In both cases, it is assumed that the detection and occupancy processes are modelled using probit link functions. In Section 4 we extend the previous work and consider the case when the detection and occupancy processes are modelled using logistic functions. We describe how the Gibbs algorithms are constructed and briefly comment how MSO models can be applied to different taxa or species groups. Before concluding, we re-analyse the data collected from a camera-trapping study undertaken by Drouilly et al. (2018) in order to highlight the use of the methods developed as well.

2 Known species richness - probit link function

The following formulation of the MSO model is similar to Broms et al. (2016)¹. It is assumed that each site (J in total) can be surveyed numerous times potentially with unequal survey effort and that n_s species have been observed in total. The observed data is stored in a three-dimensional ragged-array \mathbf{y} where $y_{i,j,k}$ represent the detection-nondetection of species i , at site j during survey k ($i = 1, \dots, n_s$; $j = 1, \dots, J$; $k = 1, \dots, K_j$). The random variable $y_{i,j,k} = 1$ if species i is observed at site j during visit k and 0 otherwise. The partially observed occupancy states are stored in a matrix \mathbf{z} where $z_{i,j} = 1$ if species i occupies site j while $z_{i,j} = 0$ if species i does not occupy site j . We assume that the occupancy process can be modelled using site specific covariates (\mathbf{X})

¹ n_s = the number of observed species; J = the number of sites; K_j = the number of surveys undertaken at site j ; $\sum_j K_j$ = the total surveys undertaken; n_d = the number of detection covariates (including an intercept); n_o = the number of occupancy covariates (including an intercept); \mathbf{X} is a $J \times n_o$ matrix; \mathbf{z} is a $n_s \times J$ matrix; \mathbf{V} is a $(\sum_{j=1}^J K_j \times J) \times n_d$ matrix.

4 *A Gibbs sampler for multi-species occupancy models*

while the detection process can be modelled using survey-specific covariates (\mathbf{V}) where it is assumed that these covariates do not vary by species.

The model can now be specified using the following Bayesian hierarchical model,

$$\begin{aligned} z_{i,j}|\psi_{i,j} &\sim \text{Bernoulli}(\psi_{i,j}), \quad i = 1, \dots, n_s; \quad j = 1, \dots, J, \\ y_{i,j,k}|z_{i,j}, p_{i,j,k} &\sim \text{Bernoulli}(z_{i,j}p_{i,j,k}), \quad k = 1, \dots, K_j \end{aligned}$$

where occupancy and detection probabilities are modelled using probit link functions, $\psi_{i,j} = \Phi(\mathbf{x}_j^T \boldsymbol{\beta}_i)$ and $p_{i,j,k} = \Phi(\mathbf{v}_{j,k}^T \boldsymbol{\alpha}_i) \quad \forall (i, j, k)$. The Bayesian formulation is completed by specifying the prior distributions used. The regression effects of the detection and occupancy processes are modelled using multivariate Gaussian distributions and we specifically assume that

$$\begin{aligned} \boldsymbol{\alpha}_i|\boldsymbol{\mu}_\alpha, \sigma_\alpha^2 &\sim \mathcal{N}(\boldsymbol{\mu}_\alpha, \sigma_\alpha^2 \mathbf{I}_{n_d}) \text{ with } \boldsymbol{\mu}_\alpha \sim \mathcal{N}(\mathbf{0}, a^2 \mathbf{I}_{n_d}) \text{ and} \\ \boldsymbol{\beta}_i|\boldsymbol{\mu}_\beta, \sigma_\beta^2 &\sim \mathcal{N}(\boldsymbol{\mu}_\beta, \sigma_\beta^2 \mathbf{I}_{n_o}) \text{ with } \boldsymbol{\mu}_\beta \sim \mathcal{N}(\mathbf{0}, b^2 \mathbf{I}_{n_o}). \end{aligned}$$

The coefficient vectors $\boldsymbol{\mu}_\alpha$ and $\boldsymbol{\mu}_\beta$ represent the community-level effects associated with the detection and occupancy processes respectively and are each assigned a common Gaussian hyper-prior distribution. By doing so, one allows the regression effects to *share information* since often the detection and occupancy probabilities are not estimable for rare animal species due to lack of detections (Tobler et al., 2015).

Finally, σ_α and σ_β are modelled using a half-Cauchy distribution (Gelman, 2006) and can be specified hierarchically as

$$\sigma_\alpha^2|\nabla_\alpha \sim \text{IG}\left(0.5, \frac{1}{\nabla_\alpha}\right), \quad \nabla_\alpha \sim \text{IG}\left(0.5, \frac{1}{A^2}\right) \text{ and} \quad (1)$$

$$\sigma_\beta^2|\nabla_\beta \sim \text{IG}\left(0.5, \frac{1}{\nabla_\beta}\right), \quad \nabla_\beta \sim \text{IG}\left(0.5, \frac{1}{B^2}\right) \quad (2)$$

where A and B are known hyper-parameters.

Posterior samples can be obtained from the above MSO model by using a MH algorithm although we instead derive a Gibbs algorithm that uses the data augmentation scheme developed by Albert and Chib (1993) to represent a probit link function by means of latent Gaussian random variables. The full conditional distributions needed to undertake a Gibbs sampler are all of a familiar form and are shown in the Supplementary information (refer to Appendix 1.A.1²). In Appendix 1.A.2 we also show how similar conditional posterior distributions can be obtained if site-specific covariates are available but no survey-specific covariates are available for the detection process. In this case, site-specific covariates are used as proxies for the survey-specific covariates.

²All references to Appendices are appendices in the Supplementary information.

3 Unknown species richness - probit link function.

Dorazio and Royle (2005) developed a model that could be used to estimate the species richness (denoted as N) in a geographical region. They assume that there exists a super-population of species (denoted as S) that consists of the observed species, n_s , as well as additional unseen species, $S - n_s$. Recall that in Section 2 survey occasion was explicitly included in the model formulation. Here all information relating to survey occasions are grouped together such that the observed data \mathbf{y}_{n_s} consists of a $n_s \times J$ matrix of the total number of detections for each species at each location (from K_j surveys, $j = 1, \dots, J$). Since N is unknown, an $(S - n_s) \times J$ matrix of zeros are introduced which represents the data associated with the unobserved species.

A latent indicator variable w_i is introduced which takes on the value 1 if species i in the super-population occurs in the geographical region under investigation and 0 otherwise. From the above discussion, $w_i = 1$ for $i = 1, \dots, n_s$ and the species richness in a geographical region is $N = \sum_{i=1}^S w_i$. The latent variable $z_{i,j}$ takes on the value 1 if species i occupies location j and 0 otherwise. The w_i indicator is modelled using a Bernoulli distribution with success parameter Ω while $z_{i,j}|w_i$ is modelled as a Bernoulli random variable with success probability, $w_i\psi_{i,j}$. The detection process (conditional on $w_i = 1, z_{i,j} = 1$) is modelled using a binomial distribution such that

$$p(y_{i,j}|z_{i,j} = 1, p_{i,j}, w_i = 1) = \binom{K_j}{y_{i,j}} p_{i,j}^{y_{i,j}} (1 - p_{i,j})^{K_j - y_{i,j}}. \quad (3)$$

Using the above description it is clear that the model can be formulated hierarchically as

$$\begin{aligned} w_i &\sim \text{Bernoulli}(\Omega), \quad \forall i = 1, \dots, S \\ z_{i,j}|w_i, \psi_{i,j} &\sim \text{Bernoulli}(w_i\psi_{i,j}), \quad \forall i = 1, \dots, S; \quad \forall j = 1, \dots, J, \\ y_{i,j}|z_{i,j}, w_i, p_{i,j} &\sim \text{Binomial}(K_j, w_i z_{i,j} p_{i,j}), \\ p_{i,j} &= \Phi(\mathbf{v}_j^T \boldsymbol{\alpha}_i) \text{ with } \boldsymbol{\alpha}_i \sim \mathcal{N}(\boldsymbol{\mu}_\alpha, \sigma_\alpha^2 \mathbf{I}_{n_d}), \\ \psi_{i,j} &= \Phi(\mathbf{x}_j^T \boldsymbol{\beta}_i) \text{ with } \boldsymbol{\beta}_i \sim \mathcal{N}(\boldsymbol{\mu}_\beta, \sigma_\beta^2 \mathbf{I}_{n_o}), \end{aligned}$$

where \mathbf{v}_j and \mathbf{x}_j variables are site-specific covariates used to model the detection and occupancy processes respectively.

Since the detection process is a binomial distribution we are unable to use the data augmentation method developed previously to sample from the detection regression effects although the remaining parameters do have simple conditional posterior distributions if we use data augmentation to obtain the posterior distribution of the occupancy regression coefficients. A derivation of the resulting conditional posterior distributions can be found in Appendix 1.B.

4 The use of logistic link functions in multi-species occupancy models.

In Appendix 1.C and 1.D we show that if logistic link functions are used to model the regression effects of the detection and occupancy processes, a Gibbs sampling algorithm, with closed form conditional posterior distributions for **all parameters** of the MSO models in Sections 2 and 3, can be derived. A Gibbs sampler is preferable to the use of a MH algorithm since the former does not require any ‘tuning’ when undertaking the MCMC sampling.

As an example, consider the MSO model developed in Section 3 but assume that the regression effects for detection and occupancy are modelled as

$$\begin{aligned}\text{logit}(p_{i,j}) &= \mathbf{v}_j^T \boldsymbol{\alpha}_i \text{ where } \boldsymbol{\alpha}_i \sim \mathcal{N}(\boldsymbol{\mu}_\alpha, \sigma_\alpha^2 \mathbf{I}_{n_d}), \\ \text{logit}(\psi_{i,j}) &= \mathbf{x}_j^T \boldsymbol{\beta}_i \text{ where } \boldsymbol{\beta}_i \sim \mathcal{N}(\boldsymbol{\mu}_\beta, \sigma_\beta^2 \mathbf{I}_{n_o})\end{aligned}$$

and both $\boldsymbol{\mu}_\alpha$ and $\boldsymbol{\mu}_\beta$ are modelled hierarchically as done previously.

The conditional posterior distributions of most of the parameters remain the same as those defined in Section 3 although the conditional posterior distributions of $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_i$ are slightly amended as detailed below.

When $w_i = 0$, $\boldsymbol{\beta}_i | \cdot \sim \mathcal{N}(\boldsymbol{\mu}_\beta, \sigma_\beta^2 \mathbf{I}_{n_o})$ while when $w_i = 1$, posterior samples of $\boldsymbol{\beta}_i$ can be obtained by sampling from a Pólya-Gamma distribution and Gaussian distribution in turn. Specifically,

$$\begin{aligned}\omega_{i,j}^{(\beta)} | \boldsymbol{\beta}, w_i = 1, \cdot &\sim \text{PG}(1, \mathbf{x}_j^T \boldsymbol{\beta}_i), \forall j = 1, \dots, J \text{ and} \\ \boldsymbol{\beta}_i | \mathbf{y}, \omega_i^{(\beta)}, w_i = 1, \cdot &\sim \mathcal{N}(\boldsymbol{\mu}_{\beta_i}, \boldsymbol{\Sigma}_{\beta_i}),\end{aligned}$$

where the ‘PG’ notation denotes a Pólya-Gamma random variable, $\omega_i^{(\beta)} = [\omega_{i,1}^{(\beta)}, \dots, \omega_{i,J}^{(\beta)}]^T$, $\tau_\beta = \frac{1}{\sigma_\beta^2}$ and

$$\begin{aligned}\boldsymbol{\Sigma}_{\beta_i} &= (\tau_\beta \mathbf{I}_{n_o} + \mathbf{X}^T \mathbf{D}_\beta \mathbf{X})^{-1}, \\ \boldsymbol{\mu}_{\beta_i} &= \boldsymbol{\Sigma}_{\beta_i} (\mathbf{X}^T \boldsymbol{\Theta}_\beta + \tau_\beta \boldsymbol{\mu}_\beta).\end{aligned}$$

The matrices \mathbf{D}_β and $\boldsymbol{\Theta}_\beta$ are defined as, $\mathbf{D}_\beta = \text{diagonal} [\omega_{i,1}^{(\beta)}, \dots, \omega_{i,J}^{(\beta)}]$ and $\boldsymbol{\Theta}_\beta = [z_{i,1} - 1/2, \dots, z_{i,J} - 1/2]^T$.

The conditional posterior distribution of $\boldsymbol{\alpha}_i$ can be obtained similarly such that

$$\begin{aligned}\omega_{i,j}^{(\alpha)} | \boldsymbol{\alpha}, w_i = 1, \cdot &\sim \text{PG}(K_j, \mathbf{x}_j^T \boldsymbol{\alpha}_i), \forall j = 1, \dots, j^*, \\ \boldsymbol{\alpha}_i | \mathbf{y}, \omega_i^{(\alpha)}, w_i = 1, \cdot &\sim \mathcal{N}(\boldsymbol{\mu}_{\alpha_i}, \boldsymbol{\Sigma}_{\alpha_i})\end{aligned}$$

where $\omega_i^{(\alpha)} = [\omega_{i,1}^{(\alpha)}, \dots, \omega_{i,j^*}^{(\alpha)}]^T$, $\tau_\alpha = \frac{1}{\sigma_\alpha^2}$ and

$$\boldsymbol{\Sigma}_{\alpha_i} = \left(\tau_\alpha \mathbf{I}_{n_d} + \tilde{\mathbf{V}}^T \mathbf{D}_\alpha \tilde{\mathbf{V}} \right)^{-1},$$

$$\mu_{\alpha_i} = \Sigma_{\alpha_i} \left(\tilde{\mathbf{V}}^T \Theta_{\alpha} + \tau_{\alpha} \mu_{\alpha} \right),$$

where $\tilde{\mathbf{V}}$ is defined as the elements of \mathbf{V} associated with $z_{i,j} = 1$ while the matrices \mathbf{D}_{α} and Θ_{α} are defined as, $\mathbf{D}_{\alpha} = \text{diagonal} \left[\omega_{i,1}^{(\alpha)}, \dots, \omega_{i,j^*}^{(\alpha)} \right]$ and $\Theta_{\alpha} = [y_{i,1} - K_1/2, \dots, y_{i,j^*} - K_{j^*}/2]^T$ for an appropriate value of j^* .

Complete derivations of the conditional posterior distributions of the MSO model can be found in Appendix 1.C while Appendix 1.D contains the required conditional distributions when species richness is assumed known. In the above formulations of the MSO model, it is assumed that both the detection and occupancy regression effects are modelled as realisations from a common Gaussian distribution. These models can easily be extended by adding one additional level of complexity and assume that the regression effects of different groups (e.g. bird guilds, taxa or species groups) are modelled as realisations from different Gaussian distributions. The required algorithm is included in Appendix 1.D.2. The above model formulations can easily be extended to account for a Bernoulli detection process or the inclusion of spatial random effects in either the detection or occupancy processes.

5 Application

As an application of the methods developed, we re-analyse the data collected from a camera-trapping study undertaken by Drouilly et al. (2018). The Anysberg Nature Reserve (33° 31' S, 20° 37' E) in South Africa (referred to as *Anysberg*) is the study area. The camera trapping design consisted of 156 camera trap sites (deployed from the end of September 2013-May 2014) each placed at a location within a 2 km grid design (Kinnaird & O'Brien, 2012). Data were collected using Bushnell Trophy CAM HG (Bushnell Outdoor Products, Overland Park, Kansas) camera traps which consisted of the processing of camera images obtained when cameras were triggered in the field. The study specifically focused on terrestrial vertebrates ≥ 0.5 kilograms in mass. A detailed explanation of camera trap survey design can be found in Section 2.2 of Drouilly et al. (2018). A list of the species observed at least once at Anysberg during the course of the study can be found in Appendix 1.5.

Since the number of detections obtained daily were low, it was decided to define a sampling occasion as a 6 day period. From the captured images an observation matrix was constructed which consisted of the observed number of times species i (for $i = 1, \dots, n_s = 35$) was detected at site j ($j = 1, \dots, J = 156$) during K_j sampling occasions. The observed data is stored as $[y_{i,j}]$.

Below we use the observed species at Anysberg and relate the occupancy and detection processes to explanatory variables. We assume that **species richness is known** and fit the MSO model described in Section 2 (assuming logistic link functions) to the data. The occupancy process is modelled as a function of *elevation* (measured as metres above sea-level), the *modified soil-adjusted index* (*MSAVI2*) (Qi, Chehbouni, Huerte, Kerr, & Sorooshian,

1994) and a *PreyIndex*³ MSAVI2 is a vegetation index similar to the normalised difference vegetation index. *MSAVI2* was converted into a binary indicator variable (named *MSAVI2ind*) such that *MSAVI2ind* equals 1 if the standardised variable *MSAVI2* is positive. The detection process is modelled as a function of a binary indicator variable, *trail* and a habitat factor variable. The *trail* variable equals 1 when a camera is placed on a trail and 0 otherwise. The habitat factor variable has three levels namely, *plain*, *river* and *mountain* where the *mountain* level is used as the baseline.

We fit all model combinations (32 in total, see Appendix 1.6) and assume that all models include an intercept in the detection and occupancy process. We further assume that we are in a \mathcal{M} -complete setting (Bernardo & Smith, 1994) and acknowledge that other variables might well be the true drivers of both the occupancy and detection process but choose to use the ones listed above. All MCMC sampling was undertaken using the authors' code. For all models, three chains of 100 000 draws were used with a burn-in proportion of one-third. The following hyper-parameters were used: $a^2 = 2.25^2 = b^2$ and $A^2 = 2.25^2 = B^2$. MSO models produce significant amounts of outputs and are computationally expensive. Because of this, we thinned the resulting MCMC chains using a factor of ten (Link & Eaton, 2012). Various convergence tests were undertaken to assess that the MCMC chains had converged (Smith, 2007).

Exploratory (with-in sample) model selection was initially undertaken by first considering the WAIC (Watanabe, 2010) and the approximate leave-one-out cross-validation scores (LOO) using Pareto smoothed importance sampling (Vehtari, Gelman, & Gabry, 2017) for each of the candidate models. In order to assess the variability of the model selection statistics ten independent MCMC runs were undertaken and two goodness-of-fit statistics namely the Bayesian p-value (Meng et al., 1994) and the sampled Bayesian p-value (Zhang, 2014) were also calculated. Gosselin (2011) show that asymptotically the distribution of the sampled Bayesian p-value is uniformly distributed over the interval [0, 1] which suggests that misspecified models will not have uniformly distributed sample Bayesian p-values. The Bayesian posterior probability of each of the models were calculated using the version of the Reversible jump MCMC (RJMCMC) algorithm developed by Barker and Link (2013) which were then used to undertake Bayesian model averaging (Hoeting, Madigan, Raftery, & Volinsky, 1999). We assume a-priori that each model is equally likely. Refer to Appendix 1.7 for more details regarding the above mentioned model selection methods.

5.1 Application Results

Both model selection measures require the calculation of the variance of the logarithm of the integrated likelihood evaluated using the posterior samples at each MCMC iteration. Based on ten independent MCMC runs, we found that

³The *PreyIndex* variable was constructed as the number of independent pictures for certain types of prey species considered in the study divided by the number of camera trap nights across the whole survey multiplied by 100. Pictures were assumed to be independent if they were taken more than 30 minutes apart.

both model selection measures were not accurately estimated since the previous mentioned variance calculation was often greater than 0.4 thus rendering both model selection methods unreliable (Vehtari et al., 2017).

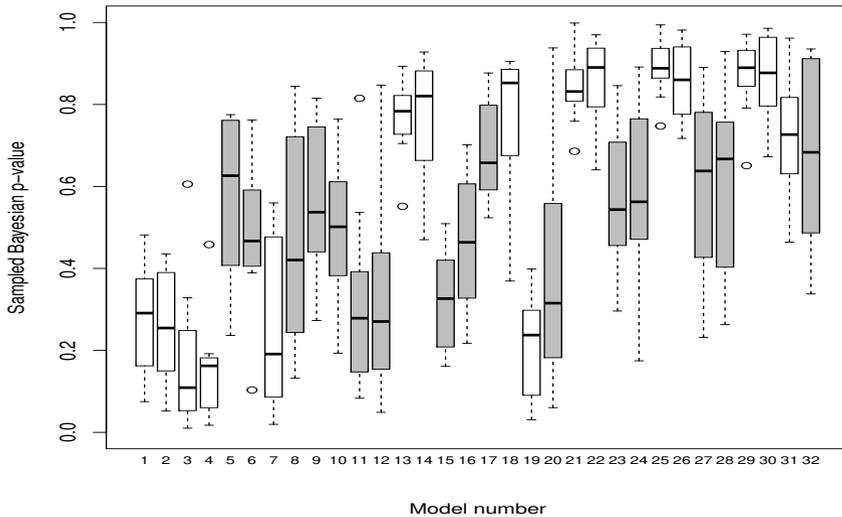


Fig. 1 The sampled Bayesian p-values of each model based on 10 independent replications. The highlighted boxplots indicate that the assumption that the distribution of the sampled Bayesian p-values are uniformly distributed over the interval $[0, 1]$ cannot be rejected for these models when testing at a 30% significance level. The model associated with the ‘Model number’ can be found in Appendix 1.6.

The variability of the Bayesian p-values were of an order of magnitude smaller than those observed for the sampled Bayesian p-values (see Appendix 1.8). Anderson Darling tests (Anderson & Darling, 1954) using the sampled Bayesian p-values obtained from ten independent MCMC runs, indicate that the assumption that the distribution of the sampled Bayesian p-values are uniformly distributed over the interval $[0, 1]$ cannot be rejected for fifteen of the thirty two models (assuming a 30% significance level⁴). These models have been highlighted in grey in Figure 1, which displays the boxplots of the sampled Bayesian p-values based on ten independent MCMC runs. The models and the p-values of the Anderson-Darling tests are tabulated in Table 1. Given the small sample size (of ten), these tests tentatively suggests that models that only include an intercept term in the occupancy process are misspecified. These results also suggest that none of the variables considered always lead to misspecified models and thus all variables do appear to have some predictive ability of occupancy for the species considered in this data set.

⁴A liberal significance level was selected since the sample size is very small.

Table 1 The list of models associated with non-significant Anderson-Darling tests when testing the null hypothesis that the distribution of the sampled Bayesian p-values are uniformly distributed over the range [0, 1]; the respective p-value of the test as well as the median sampled Bayesian p-value of the associated models.

Model	Occupancy covariates	Detection covariates	p-value	Median sampled Bayesian p-value
5	elevation	1	0.67	0.63
6	elevation	trail	0.78	0.47
8	elevation	trail + plain + river	0.76	0.42
9	MSAVI2ind	1	0.74	0.54
10	MSAVI2ind	trail	0.82	0.5
11	MSAVI2ind	plain + river	0.30	0.28
12	MSAVI2ind	trail + plain + river	0.31	0.27
15	elevation + MSAVI2ind	plain + river	0.42	0.33
16	elevation + MSAVI2ind	trail + plain + river	0.80	0.46
17	PreyIndex	1	0.43	0.66
20	PreyIndex	trail + plain + river	0.50	0.32
23	elevation + PreyIndex	plain + river	0.76	0.54
24	elevation + PreyIndex	trail + plain + river	0.64	0.56
27	MSAVI2ind + PreyIndex	plain + river	0.62	0.61
28	MSAVI2ind + PreyIndex	trail + plain + river	0.58	0.67
32	elevation + MSAVI2ind + PreyIndex	trail + plain + river	0.31	0.68

Conn, Johnson, Williams, Melin, and Hooten (2018) suggests that the Bayesian p-values should not be used to undertake model selection, but should rather be used to identify models that display clear signs of model misspecification. In our context, models 3, 4, 21, 25, 26, 29, and 30 appear to be misspecified since they either have a Bayesian p-value above 0.8 or are smaller than 0.2. These models either only include an intercept term in the detection or occupancy process; do not include the *elevation* variable in the occupancy process or only includes the *trail* variable in the detection process.

Table 2 The posterior model probabilities obtained using RJMCMC

Model	Occupancy covariates	Detection covariates	Posterior probability	Posterior odds
6	elevation	trail	0.8543	1
8	elevation	trail + plain + river	0.1442	5.92
22	elevation + PreyIndex	trail	0.0011	776.64
24	elevation + PreyIndex	trail + plain + river	0.0002	4 271.5
5	elevation	1	0.0001	8 543

The resulting posterior model probabilities from the RJMCMC algorithm are displayed in Table 2. The models listed contain a combination of the *elevation* and *PreyIndex* variable in the occupancy process as well as a combination of the indicator variables *trail*, *plain* and *river* in the detection process. From these results, we observed that Model 6 has the largest posterior probability

with a value of 0.85, Model 8 has the second largest posterior probability with a value of 0.14 while all other models have very low posterior mass associated with them.

From an examination of the population regression effects in Table 4 in Appendix 1.9 it can be seen that the mean population regression effects for plain and river in the detection process (of Models 8 and 24) both have 95% credibility intervals that contain zero. Similarly, the PreyIndex variable in Models 22 and 24 also have 95% credibility intervals that contain zero. Table 3 contains the Bayesian model averaged population regression effects⁵ as well as the posterior effect probabilities (the probability that a population regression effect is non-zero). The results clearly indicate that trail and elevation have non-zero regression effects in the detection and occupancy processes respectively while the remaining regression effects have a high probability of being zero. These results indicate that on average one is more likely to detect one of the animal species investigated off a trail than on it and that on average no significant differences in detectability of animal species are observed between the different habitat types. We found that the mean population regression effect for elevation is negative indicating that on average higher occupancy probabilities are observed at lower elevation levels than at higher ones.

Table 3 The Bayesian model averaged community-level regression effects. (Posterior mean, Standard deviation as well as posterior effect probabilities.) Regression effects cannot be calculated if $\Pr(\beta \neq 0 | \text{Data})$ and is captured as ‘-’.

Coefficient	Mean	Std	$\Pr(\beta \neq 0 \text{Data})$
α_{trail}	-0.54	0.21	1
α_{plain}	0.02	0.10	0.14
α_{river}	0.02	0.11	0.14
$\alpha_{\text{MSAVI2ind}}$	-	-	0
$\beta_{\text{elevation}}$	-0.59	0.22	1
$\beta_{\text{PreyIndex}}$	0.0003	0.01	0.001
$\beta_{\text{MSAVI2ind}}$	-	-	0

Figure 2 displays caterpillar plots highlighting the mean community regression effects as well as the 95% credibility intervals for the species specific regression effects of *trail* (a) and *elevation* (b) for the detection and occupancy processes respectively. Of the 35 species photographed, the occupancy probability was strongly related to *elevation* for 19 species. The results indicate that Smith’s rock hare (*Pronolagus rupestris*), klipspringers (*Oreotragus oreotragus*) and leopard’s (*Panthera pardus*) were more likely to occupy high lying regions of the reserve compared to many of the antelope species that preferred low-lying regions of the reserve. In general, the detection of most

⁵From Raftery (1993), $\hat{\Delta}_k = \mathbb{E}(\Delta_k | \mathcal{M}_k, \text{Data})$ such that $\mathbb{E}(\Delta | \text{Data}) = \sum_k \mathbb{E}(\Delta_k | \mathcal{M}_k) \Pr(\mathcal{M}_k | \text{Data})$ and $\text{var}(\Delta | \text{Data}) = \left(\sum_k \left(\text{var}(\Delta | \mathcal{M}_k, \text{Data}) - \hat{\Delta}_k^2 \right) \Pr(\mathcal{M}_k | \text{Data}) \right) - \mathbb{E}(\Delta_k | \mathcal{M}_k, \text{Data})^2$. The posterior effect probabilities are obtained by summing the posterior model probabilities across models for each variable (Hoeting et al., 1999).

carnivore species were negatively related to *trail* while the detection of species such as the hyrax (*Procavia capensis*) and the duiker (*Sylvicapra grimmia*) were positively related to *trail*. Plots displaying the estimated mean occupancy probability of the different species can be found in Appendix 1.10.

We found that the above results were not sensitive to the specification of the prior distribution. To this end, in Appendix 1.11 we display statistics pertaining to the community-level regression effects for Model 6 when run using three chains of 100 000 in length. A burn-in sample of one-third is specified thereafter the resulting chains were thinned by retaining every 10th observation. The different prior distributions were specified as follows: Prior 1: $a^2 = 2.25^2 = b^2$, $A^2 = 2.25^2 = B^2$; Prior 2: $a^2 = 2.25^2 = b^2$, $A^2 = 2.25 = B^2$; Prior 3: $a^2 = 5^2 = b^2$, $A^2 = 5^2 = B^2$; Prior 4: $a^2 = 5^2 = b^2$, $A^2 = 5 = B^2$.

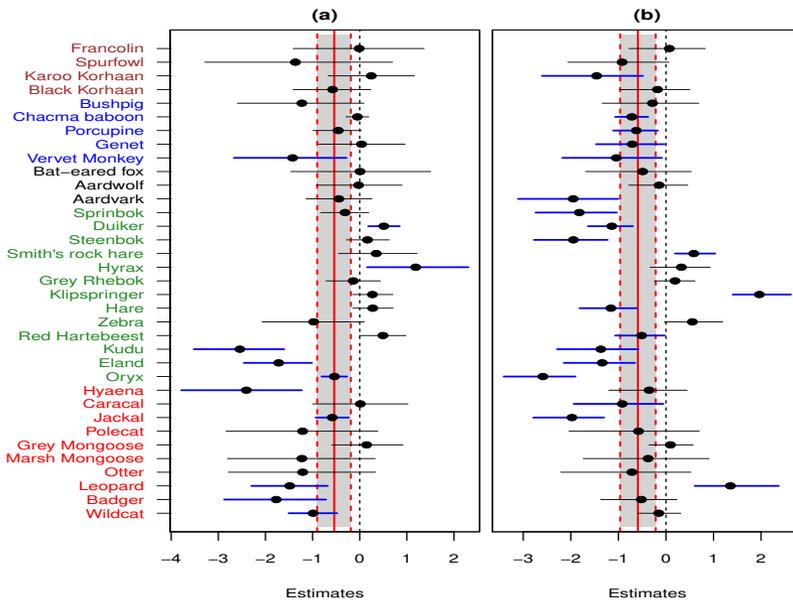


Fig. 2 Caterpillar plots highlighting the mean community regression effects as well as the 95% credibility intervals for the species specific regression effects of *trail* (a) and *elevation* (b). Credibility intervals in bold do not overlap 0. The thick dashed lines indicate the 95% equal tail credibility interval for the mean community response to each variable. All quantities were calculated using the mixture distributions obtained when undertaking Bayesian model averaging. The results were grouped into the five groups namely carnivores (red), herbivores (green), insectivores (black), omnivores (blue) and birds (brown).

6 Discussion and Conclusions

Multi-species occupancy models are important tools for statistical ecologists since they allow one to undertake community-level as well as species-specific inference by using detection-nondetection data of many species at different

locations (Broms et al., 2016; Dorazio & Royle, 2005; Dorazio et al., 2006). The information across species are pooled together which allows one to obtain occupancy probabilities of rare species which would otherwise not be possible due to the low detectability of these species.

Two types of MSO models exist in the literature. The first model-type is used to estimate species richness (Dorazio & Royle, 2005) in an area while the second model-type is used to obtain species specific regression effects when the species richness is known (Broms et al., 2016). Several studies have undertaken studies that use both model-types. They however either use probit link functions to obtain the regression effects of their models; use general Bayesian analysis software such as *JAGS*, *WinBUGS*, *NIMBLE* or *Stan* to undertake their analysis or use the MH algorithm to sample from the parameters of their models. Here, we develop Gibbs sampling algorithms when probit or logistic link functions are used to model the detection and occupancy processes respectively and show how this can be done for a number of different versions of the MSO models. We also specifically show that the algorithms for MSO models (that use a logistic link function) are similar to those developed for single season occupancy models as well as single season spatial occupancy models where all algorithms use the work of Polson, Scott, and Windle (2013) and Clark and Altwegg (2019) to develop efficient Gibbs sampling algorithms.

In our application we found that model selection methods such as WAIC (Watanabe, 2010) and the approximate leave-one-out cross-validation scores using Pareto smoothed importance sampling (Vehtari et al., 2017) were unreliable for the data considered. A possible reason for this being the sparsity of data at a number of sites. Model selection was undertaken using Reversible jump MCMC (Barker & Link, 2013) while the occupancy surfaces for all species considered were obtained using Bayesian model averaging (see Appendix 1.10).

A single MSO model took approximately 12 – 13 minutes to fit compared to approximately 4 hours when using *Stan*. All calculations were performed using R 4.0.4 (R Core Team, 2014) on a Windows 10 desktop with a AMD Ryzen 7 3700X 8-Core Processor, a clock speed of 3.59 GHz and 64 GB of RAM. On the same system, the RJMCMC algorithm took approximately one hour to complete. These run-times were **not found to be prohibitively long** and thus we would encourage others to consider using the technique when undertaking model selection. The authors used their own code to undertake the RJMCMC algorithm although the R package *rjmc* (Gelling, Schofield, & Barker, 2019) could also be used.

Supplementary information

All appendices referred to in the manuscript were stored as supplementary information.

Declarations

We have no conflict of interest to declare. Data and can be shared on request.

References

- Albert, J.H., & Chib, S. (1993). Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, *88*(422), 669-679.
- Anderson, T.W., & Darling, D.A. (1954). A test of goodness of fit. *Journal of the American Statistical Association*, *49*(268), 765-769.
- Barker, R.J., & Link, W.A. (2013). Bayesian multimodel inference by RJMCMC: A Gibbs sampling approach. *The American Statistician*, *67*(3), 150-156.
- Bernardo, J.M., & Smith, A.F. (1994). *Bayesian theory* (Vol. 405). John Wiley and Sons.
- Broms, K.M., Hooten, M.B., Fitzpatrick, R.M. (2016). Model selection and assessment for multi-species occupancy models. *Ecology*, *97*(7), 1759-1770.
- Brooks, S., Gelman, A., Jones, G., Meng, X.-L. (2011). *Handbook of Markov Chain Monte Carlo*. Chapman and Hall/CRC.
- Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., . . . Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1-32.
- Clark, A.E., & Altwegg, R. (2019). Efficient Bayesian analysis of occupancy models with logit link functions. *Ecology and evolution*, *9*(2), 756-768.
- Conn, P.B., Johnson, D.S., Williams, P.J., Melin, S.R., Hooten, M.B. (2018). A guide to Bayesian model checking for ecologists. *Ecological Monographs*, *88*(4), 526-542.
- Datta, A., Banerjee, S., Finley, A.O., Gelfand, A.E. (2016). Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. *Journal of the American Statistical Association*, *111*(514), 800-812.

- de Valpine, P., Turek, D., Paciorek, C., Anderson-Bergman, C., Temple Lang, D., Bodik, R. (2017). Programming with models: writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics*, 26(2), 403-413.
- Devarajan, K., Morelli, T.L., Tenan, S. (2020). Multi-species occupancy models: review, roadmap, and recommendations. *Ecography*, 43(11), 1612-1624.
- Dorazio, R.M., Gotelli, N.J., Ellison, A.M. (2011). Modern methods of estimating biodiversity from presence-absence surveys. *Biodiversity loss in a changing planet*. (p. 277-302). InTech Rijeka, Croatia.
- Dorazio, R.M., & Rodriguez, D.T. (2012). A Gibbs sampler for Bayesian analysis of site-occupancy data. *Methods in Ecology and Evolution*, 3(6), 1093-1098.
- Dorazio, R.M., & Royle, J.A. (2005). Estimating size and composition of biological communities by modeling the occurrence of species. *Journal of the American Statistical Association*, 100(470), 389-398.
- Dorazio, R.M., Royle, J.A., Söderström, B., Glimskär, A. (2006). Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology*, 87(4), 842-854.
- Doser, J.W., Finley, A.O., Kéry, M., Zipkin, E.F. (2021). spOccupancy: An R package for single species, multispecies, and integrated spatial occupancy models. *arXiv preprint arXiv:2111.12163*.
- Drouilly, M., Clark, A., O’Riain, M.J. (2018). Multi-species occupancy modelling of mammal and ground bird communities in rangeland in the karoo: A case for dryland systems globally. *Biological Conservation*, 224, 16-25.
- Gelling, N., Schofield, M.R., Barker, R.J. (2019). R package rjmc: reversible jump mcmc using post-processing. *Australian and New Zealand Journal of Statistics*, 61(2), 189-212.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian analysis*, 1(3), 515-534.
- Geman, S., & Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6), 721-741.

- Goijman, A.P., Conroy, M.J., Bernardos, J.N., Zaccagnini, M.E. (2015). Multi-season regional analysis of multi-species occupancy: implications for bird conservation in agricultural lands in east-central argentina. *PLoS one*, *10*(6), e0130874.
- Gosselin, F. (2011). A New Calibrated Bayesian Internal Goodness-of-Fit Method: Sampled Posterior p-Values as Simple and General p-Values That Allow Double Use of the Data. *PLoS one*, *6*(3), e14770.
- Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, *57*(1), 97–109.
- Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T. (1999). Bayesian model averaging: a tutorial (with comments by M. Clyde, David Draper and E. I. George, and a rejoinder by the authors). *Statistical Science*, *14*(4), 382-417.
- 10.1214/ss/1009212519
- Hoffman, M.D., & Gelman, A. (2014). The No-U-turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*(1), 1593-1623.
- Holmes, C.C., & Held, L. (2006). Bayesian auxiliary variable models for binary and multinomial regression. *Bayesian Analysis*, *1*(1), 145-168.
- Johnson, D.S., Conn, P.B., Hooten, M.B., Ray, J.C., Pond, B.A. (2013). Spatial occupancy models for large data sets. *Ecology*, *94*(4), 801-808.
- Kinnaird, M.F., & O'Brien, T.G. (2012). Effects of private-land use, livestock management, and human tolerance on diversity, distribution, and abundance of large African mammals. *Conservation Biology*, *26*(6), 1026-1039.
- Link, W.A., & Eaton, M.J. (2012). On thinning of chains in MCMC. *Methods in Ecology and Evolution*, *3*(1), 112-115.
- Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D. (2000). WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and computing*, *10*(4), 325-337.
- Meng, X.-L., et al. (1994). Posterior predictive p-values. *The Annals of Statistics*, *22*(3), 1142-1160.
- Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E. (1953). Equation of state calculations by fast computing machines. *The*

- Journal of Chemical Physics*, 21(6), 1087-1092.
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Hornik, K., Leisch, F. and Zeileis, A. (Ed.), *Proceedings of the 3rd international workshop on distributed statistical computing*. Retrieved from <http://www.ci.tuwien.ac.at/Conferences/DSC-2003/Proceedings/>
- Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., ... McCarthy, M.A. (2014). Understanding co-occurrence by modelling species simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and Evolution*, 5(5), 397-406.
- Polson, N.G., Scott, J.G., Windle, J. (2013). Bayesian inference for logistic models using Pólya-Gamma latent variables. *Journal of the American Statistical Association*, 108(504), 1339-1349.
- Qi, J., Chehbouni, A., Huerte, A., Kerr, Y., Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, 48(2), 119-126.
- R Core Team (2014). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from <http://www.R-project.org/>
- Raftery, A.E. (1993). Bayesian model selection in structural equation models. *Sage Focus Editions*, 154, 163-163.
- Roberts, G.O., & Sahu, S.K. (1997). Updating schemes, correlation structure, blocking and parameterization for the Gibbs sampler. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 59(2), 291-317.
- Ruiz-Gutiérrez, V., Zipkin, E.F., Dhondt, A.A. (2010). Occupancy dynamics in a tropical bird community: unexpectedly high forest use by birds classified as non-forest species. *Journal of Applied Ecology*, 47(3), 621-630.
- Sauer, J.R., Blank, P.J., Zipkin, E.F., Fallon, J.E., Fallon, F.W. (2013). Using multi-species occupancy models in structured decision making on managed lands. *The Journal of Wildlife Management*, 77(1), 117-127.
- Smith, B.J. (2007). boa: an R package for MCMC output convergence assessment and posterior inference. *Journal of Statistical Software*, 21(11), 1-37.

- Tanner, M.A., & Wong, W.H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, *82*(398), 528-540.
- Thomas, A., O'Hara, B., Ligges, U., Sturtz, S. (2006). Making BUGS open. *R news*, *6*(1), 12-17.
- Tikhonov, G., Opedal, Ø.H., Abrego, N., Lehtikoinen, A., de Jonge, M.M., Oksanen, J., Ovaskainen, O. (2020). Joint species distribution modelling with the R-package Hmsc. *Methods in ecology and evolution*, *11*(3), 442-447.
- Tobler, M.W., Zúñiga Hartley, A., Carrillo-Percestequi, S.E., Powell, G.V. (2015). Spatiotemporal hierarchical modelling of species and occupancy using camera trap data. *Journal of Applied Ecology*, *52*(2), 413-421.
- Vehtari, A., Gelman, A., Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413-1432.
- Watanabe, S. (2010). Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, *11*, 3571-3594.
- Zhang, J.L. (2014). Comparative investigation of three Bayesian p values. *Computational Statistics and Data Analysis*, *79*, 277-291.
- Zipkin, E.F., DeWan, A., Andrew Royle, J. (2009). Impacts of forest fragmentation on species richness: a hierarchical approach to community modelling. *Journal of Applied Ecology*, *46*(4), 815-822.
- Zipkin, E.F., Royle, J.A., Dawson, D.K., Bates, S. (2010). Multi-species occurrence models to evaluate the effects of conservation and management actions. *Biological Conservation*, *143*(2), 479-484.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [snjnlSupplInfo.cls](#)
- [EnvMSOSupplInfo.pdf](#)
- [EnvMSOSupplInfo.bbl](#)
- [MultispeciesSupplIntroPage.tex](#)
- [SupportingInformationPaper4.tex](#)