A multispecies occupancy model for two or more interacting species

Christopher T. Rota[1,2], Marco A. R. Ferreira[3], Roland W. Kays[4,5], Tavis D. Forrester[6], Elizabeth L. Kalies[7], William J. McShea[6], Arielle W. Parsons[4] and Joshua J. Millspaugh[2]

[1]*School of Natural Resources, West Virginia University, Morgantown, WV 26506, USA;* [2]*Department of Fisheries and Wildlidfe Sciences, University of Missouri, Columbia, MO 65211, USA;* [3]*Department of Statistics, Virginia Tech, Blacksburg, VA 24061, USA;* [4]*North Carolina Museum of Natural Sciences, Raleigh, NC 27601, USA;* [5]*Department of Forestry and Environmental Resources, North Carolina State University, Raleigh, NC 27695, USA;* [6]*Smithsonian Conservation Biology Institute, Front Royal, VA 22630, USA;* and [7]*The Nature Conservancy, Durham, NC 27701, USA*

Summary

1. Species occurrence is influenced by environmental conditions and the presence of other species. Current approaches for multispecies occupancy modelling are practically limited to two interacting species and often require the assumption of asymmetric interactions. We propose a multispecies occupancy model that can accommodate two or more interacting species.

2. We generalize the single-species occupancy model to two or more interacting species by assuming the latent occupancy state is a multivariate Bernoulli random variable. We propose modelling the probability of each potential latent occupancy state with both a multinomial logit and a multinomial probit model and present details of a Gibbs sampler for the latter.

3. As an example, we model co-occurrence probabilities of bobcat (*Lynx rufus*), coyote (*Canis latrans*), grey fox (*Urocyon cinereoargenteus*) and red fox (*Vulpes vulpes*) as a function of human disturbance variables throughout 6 Mid-Atlantic states in the eastern United States. We found evidence for pairwise interactions among most species, and the probability of some pairs of species occupying the same site varied along environmental gradients; for example, occupancy probabilities of coyote and grey fox were independent at sites with little human disturbance, but these two species were more likely to occur together at sites with high human disturbance.

4. Ecological communities are composed of multiple interacting species. Our proposed method improves our ability to draw inference from such communities by permitting modelling of detection/non-detection data from an arbitrary number of species, without assuming asymmetric interactions. Additionally, our proposed method permits modelling the probability two or more species occur together as a function of environmental variables. These advancements represent an important improvement in our ability to draw community-level inference from multiple interacting species that are subject to imperfect detection.

Key-words: community, competition, eMammal, interspecific interactions, multinomial logit, multinomial probit, multivariate Bernoulli, occupancy modelling, predation

Introduction

Understanding factors that influence the distribution of species is of fundamental interest in wildlife conservation and management. A common sampling design when studying factors that influence where species occur is to repeatedly survey a selection of sample units and record detection/non-detection of that species. Failure to detect a species may occur because the species is absent from the sample unit or because the species was present within the sample unit but remained undetected (MacKenzie *et al.* 2002). A class of zero-inflated binomial models (hereafter occupancy models, MacKenzie *et al.* 2002, 2006) have proven useful for modelling factors influencing species occupancy probabilities in the face of imperfect detection.

Species occupancy probabilities are often influenced by environmental factors and the presence or absence of interacting species; for example, habitat selection in prey species can be influenced by the presence of predators (Willems & Hill 2009; Coleman & Hill 2014), competitive exclusion may prevent inferior competitors from occupying habitat (Sherry & Holmes 1988; Wang *et al.* 2015a), or habitat use may be facilitated by the presence of another species (Veblen 2012). Failure to account for interspecific interactions may lead to biased inference when modelling habitat associations (McLoughlin *et al.* 2010) and has been suggested as a primary cause for poor performance of some species distribution models at large geographic scales (Byholm *et al.* 2012). Several methods that account for dependence between two or more interacting species have been proposed (e.g. Latimer *et al.* 2009; Ovaskainen, Hottola & Siitonen 2010; Pollock *et al.* 2014) though none of

*Correspondence author. E-mail: christopher.rota@mail.wvu.edu

these methods account for imperfect detection. Additionally, several multispecies occupancy models that account for imperfect detection have been proposed (e.g. MacKenzie, Bailey & Nichols 2004; Richmond, Hines & Beissinger 2010; Waddle *et al.* 2010; Rota *et al.* 2016b), but these methods have been limited to two interacting species.

The first approach for modelling occupancy probabilities of two interacting species while accounting for imperfect detection was proposed by MacKenzie, Bailey & Nichols (2004). While this model has proven useful for modelling co-occurrence probabilities of a variety of species (e.g. Bailey *et al.* 2009; Steen *et al.* 2014), it nonetheless has several limitations. In particular, their parameterization included a 'species interaction factor' (SIF) term that, while attempting to quantify the degree of dependence in occurrence probability between two species, can also lead to numerical difficulties (MacKenzie, Bailey & Nichols 2004). In fact, MacKenzie, Bailey & Nichols (2004) were unable to achieve convergence when fitting a covariate model of co-occurrence between two salamander species. In response, both Waddle *et al.* (2010) and Richmond, Hines & Beissinger (2010) proposed alternative two-species occupancy models that are numerically stable, but which assume asymmetric interactions between two species (e.g. one species is dominant over the other). Interspecific interactions occur between numerous species within a community, and as the size and scope of many ecological monitoring programmes increase, there is a clear need to simultaneously model occupancy dynamics for numerous species; for example, recent studies by Steen *et al.* (2014), Lesmeister *et al.* (2015), and Wang, Allen & Wilmers (2015b) all investigate occupancy dynamics of more than two interacting species. Although these investigators used state-of-the art techniques, they were nonetheless constrained to fit multiple two-species models, when a single model would have been able to simultaneously accommodate interactions among all species.

We propose a generalization of single-season occupancy models that can accommodate two or more interacting species. Our model assumes the latent occupancy state is a multivariate Bernoulli random variable (Dai, Ding & Wahba 2013), which allows construction of numerically stable covariate models of species co-occurrence that do not require *a priori* assumptions of asymmetric interactions. Furthermore, assuming a latent multivariate Bernoulli occupancy state provides explicit conditions for interspecific independence without the need to include additional parameters (i.e. MacKenzie, Bailey & Nichols [2004]'s SIF) and pairwise covariance/correlation can be obtained using results from basic probability theory. Finally, our model permits ecologists to model the probability that two or more species occupy the same site as a function of covariates, a novel development among occupancy models and joint species distribution models. As an example, we model co-occurrence probability of bobcat (*Lynx rufus*), coyote (*Canis latrans*), grey fox (*Urocyon cinereoargenteus*) and red fox (*Vulpes vulpes*) using camera trap data from 6 Mid-Atlantic states in the eastern United States.

Materials and methods

SAMPLING PROTOCOL

The basic sampling protocol and assumptions for a multispecies occupancy model are identical to the single-species case (MacKenzie *et al.* 2006). Briefly, a set of n sites is randomly selected from a population of interest, and each site i is surveyed J_i times. During each survey, detection/non-detection of S focal species is recorded. Detection/non-detection data are partial observations of the underlying occupancy state, which we assume can be modelled as a multivariate Bernoulli random variable.

MULTISPECIES OCCUPANCY MODEL

The multispecies occupancy model we propose generalizes MacKenzie et al.'s (2002) single-species model to 2 or more species. As with the single-species case, we link a detection model with a partially observed, latent process model. We model detections ($y_{sit} = 1$) and non-detections ($y_{sit} = 0$) of species s at site i during survey t, conditional on the presence of species s ($z_s = 1$), as a Bernoulli random variable:

$$y_{sit}|z_{si} \sim \text{Bernoulli}(z_{si}p_{sit}).$$

The conditional detection probability p_{sit} can be modelled as a function of covariates with a variety of link functions, for example, $p_{sit} = \text{logit}^{-1}(\mathbf{w}'_{sit}\boldsymbol{\alpha}_s)$, where $\mathbf{w}_{sit}$ is a D-dimensional vector of covariates thought to influence detection probability and $\boldsymbol{\alpha}_s$ is D-dimensional vector of slope parameters. We model the latent occupancy state of species s at site i as a multivariate Bernoulli random variable:

$$\mathbf{Z}_i \sim \text{MVB}(\boldsymbol{\Psi}_i),$$

where $\mathbf{Z}_i = \{z_{1i}, z_{2i}, ..., z_{Si}\}$ is an S-dimensional vector of 1's and 0's denoting the latent occupancy state of all S species and $\boldsymbol{\Psi}_i$ is a 2^S-dimensional vector denoting the probability of all possible sequences of 1's and 0's $\mathbf{Z}_i$ can attain, such that $\sum_{a=1}^{2^S} \psi_{ai} = 1$. Note that conditional detection probabilities can be allowed to vary both as a function of the presence/absence of other species, as in MacKenzie, Bailey & Nichols (2004) and Waddle *et al.* (2010) or as a function of detection/non-detection of other species as in Richmond, Hines & Beissinger (2010).

THE MULTIVARIATE BERNOULLI DISTRIBUTION

The multivariate Bernoulli distribution is a generalization of the well-known Bernoulli distribution to >1 dimension (Dai, Ding & Wahba 2013). We begin by describing the univariate Bernoulli distribution. Then, for ease of exposition, we present results from the bivariate Bernoulli distribution, though this model readily generalizes to 3 or more dimensions. When $S = 1$ (i.e. single-species occupancy models),

$$z \sim \text{Bernoulli}(\psi),$$

with corresponding probability mass function

$$
\begin{aligned}
f(z|\psi) &= \psi^z(1-\psi)^{1-z} \\
&= \exp\left(z \log\left(\tfrac{\psi}{1-\psi}\right) + \log(1-\psi)\right).
\end{aligned}
$$

The quantity $f = \log\left(\tfrac{\psi}{1-\psi}\right)$ is the log odds a species occupies a site and is often referred to as the 'natural parameter' (Schabenberger & Pierce 2002, p. 305). In applications assuming a logit link, the natural parameter is modelled as a function of covariates:

$$f = \log\left(\frac{\psi}{1-\psi}\right) = \mathbf{x}'\boldsymbol{\beta},$$

where $\mathbf{x}$ is a vector of covariates and $\boldsymbol{\beta}$ is a conformable vector of parameters. When modelling the natural parameter as a linear function, slope parameters are interpreted as the log odds ratio of occupancy probability resulting from a 1-unit change in their associated variables. When $S = 2$,

$$\mathbf{Z} \sim \text{MVB}(\psi_{11}, \psi_{10}, \psi_{01}, \psi_{00}),$$

with corresponding probability mass function

$$f(\mathbf{Z}|\psi_{11},\psi_{10},\psi_{01},\psi_{00}) = \psi_{11}^{z_1 z_2}\psi_{10}^{z_1(1-z_2)}\psi_{01}^{(1-z_1)z_2}\psi_{00}^{(1-z_1)(1-z_2)}$$
$$= \exp\left(\log(\psi_{00}) + z_1\log\left(\frac{\psi_{10}}{\psi_{00}}\right) + z_2\log\left(\frac{\psi_{01}}{\psi_{00}}\right)\right.$$
$$\left. + z_1 z_2\log\left(\frac{\psi_{11}\psi_{00}}{\psi_{01}\psi_{10}}\right)\right).$$

We now define the natural parameters f_1, f_2, f_{12} as

$$f_1 = \log\left(\frac{\psi_{10}}{\psi_{00}}\right)$$
$$f_2 = \log\left(\frac{\psi_{01}}{\psi_{00}}\right)$$
$$f_{12} = \log\left(\frac{\psi_{11}\psi_{00}}{\psi_{01}\psi_{10}}\right).$$

Covariate information can be included by modelling the natural parameters as linear functions:

$$\begin{aligned} f_1 &= \mathbf{x}'_\alpha\boldsymbol{\alpha} \\ f_2 &= \mathbf{x}'_\beta\boldsymbol{\beta} \qquad\qquad \text{eqn 1} \\ f_{12} &= \mathbf{x}'_\gamma\boldsymbol{\gamma}, \end{aligned}$$

where $\mathbf{x}_\alpha, \mathbf{x}_\beta$, and $\mathbf{x}_\gamma$ are vectors of covariates (of possibly different dimensions) and $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$ are conformable vectors of slope parameters. Note that pairwise independence between z_1 and z_2 occurs if $f_{12} = 0$. Finally, the natural parameters can be used to obtain the probability of each combination of 1's and 0's via the multinomial logit link:

$$\begin{aligned} \psi_{11} &= \frac{\exp(f_1+f_2+f_{12})}{1+\exp(f_1)+\exp(f_2)+\exp(f_1+f_2+f_{12})}, \\ \psi_{10} &= \frac{\exp(f_1)}{1+\exp(f_1)+\exp(f_2)+\exp(f_1+f_2+f_{12})}, \\ \psi_{01} &= \frac{\exp(f_2)}{1+\exp(f_1)+\exp(f_2)+\exp(f_1+f_2+f_{12})}, \qquad \text{eqn 2} \\ \psi_{00} &= \frac{1}{1+\exp(f_1)+\exp(f_2)+\exp(f_1+f_2+f_{12})}. \end{aligned}$$

Quantities of interest, such as the marginal probability of occurrence for species s; pairwise covariance; or the probability of occurrence of one species, conditional on the presence or absence of another species, can be obtained using results from basic probability theory; for example, when $S = 2$, the marginal probability of occurrence for species 1 is:

$$P(z_1 = 1) = \psi_{11} + \psi_{10}, \qquad\qquad \text{eqn 3}$$

the covariance between species 1 and 2 is:

$$\text{cov}(z_1, z_2) = \psi_{11}\psi_{00} - \psi_{10}\psi_{01},$$

and the probability of occurrence for species 1, conditional on the presence or absence of species 2, is:

$$P(z_1 = 1|z_2) = \frac{\psi_{1z_2}}{\psi_{1z_2} + \psi_{0z_2}}. \qquad\qquad \text{eqn 4}$$

Slope parameters of linear models from eqn 1 have direct interpretations in terms of conditional occupancy probabilities. To see this, assume the natural parameters are linear functions of 1 covariate each:

$$\begin{aligned} f_1 &= \alpha_0 + \alpha_1 x, \\ f_2 &= \beta_0 + \beta_1 x, \\ f_{12} &= \gamma_0 + \gamma_1 x. \end{aligned}$$

From eqn 4, and subsequent algebraic manipulation of probabilities presented in eqn 2, we obtain all possible conditional probabilities of occurrence:

$$\begin{aligned} P(z_1 = 1|z_2 = 0) &= \frac{\psi_{10}}{\psi_{10}+\psi_{00}} = \text{logit}^{-1}(\alpha_0 + \alpha_1 x), \\ P(z_1 = 1|z_2 = 1) &= \frac{\psi_{11}}{\psi_{11}+\psi_{01}} = \text{logit}^{-1}((\alpha_0 + \gamma_0) + (\alpha_1 + \gamma_1)x), \\ P(z_2 = 1|z_1 = 0) &= \frac{\psi_{01}}{\psi_{01}+\psi_{00}} = \text{logit}^{-1}(\beta_0 + \beta_1 x), \\ P(z_2 = 1|z_1 = 1) &= \frac{\psi_{11}}{\psi_{11}+\psi_{10}} = \text{logit}^{-1}((\beta_0 + \gamma_0) + (\beta_1 + \gamma_1)x). \end{aligned}$$

Writing the conditional occupancy probabilities in this manner demonstrates that slope parameters associated with f_1 and f_2 (i.e. α_1 and β_1 in our example) can be interpreted as log odds ratios of occupancy probabilities of each species, conditional on the absence of the other species, resulting from a 1-unit change in x. Additionally, it can be seen that slope parameters associated with f_{12} (i.e. γ_1 in our example) modify the relationship between x and the occupancy probability of one species in the presence of the other species. More formally, γ_1 can be interpreted as the difference in log odds ratios of occupancy probability of one species resulting from a 1-unit change in x when the other species is present and absent (Appendix S1).

Different hypotheses regarding the influence of environmental variables and species interactions can be explored by assuming different functional forms for the natural parameters; for example, there may be interest in exploring evidence for species interactions while accounting for the effects of environmental variables. This may be accomplished by comparing a model assuming species occur independently (e.g. assuming $f_1 = \alpha_0 + \alpha_1 x$; $f_2 = \beta_0 + \beta_1 x$; and $f_{12} = 0$) with a model that assumes conditional probabilities of occurrence of one species are different in the presence or absence of the other species (e.g. assuming $f_1 = \alpha_0 + \alpha_1 x$; $f_2 = \beta_0 + \beta_1 x$; and $f_{12} = \gamma_0$). An advantage of the multivariate Bernoulli model relative to other joint species distribution models is that it allows the influence of environmental variables on one species to differ in the presence and absence of another species; for example, there may be interest in evaluating evidence that the probability two species occur together varies along an environmental gradient. This may be accomplished by comparing a model assuming the probability of one species, conditional on the presence of the other, is also function of x (e.g. assuming $f_1 = \alpha_0 + \alpha_1 x$; $f_2 = \beta_0 + \beta_1 x$; and $f_{12} = \gamma_0 + \gamma_1 x$).

There may also be interest in quantifying the relative importance of environmental variables vs. interspecific interactions in influencing the probability of occurrence for individual species. This can be accomplished by calculating odds ratios of predicted probability of use of species 1 conditional on the presence and absence of species 2, which gives an interpretation identical to slope parameters of linear predictors in occupancy and detection models. If q denotes the probability an event will happen, the odds of that event is defined as $q/(1-q)$. The odds ratio is then defined as the odds of one event occurring divided by the odds of another event occurring; for example, if covariate h takes baseline value x, we can calculate the odds ratio associated with the presence or absence of species 2:

$$\text{OR}_{\text{sp}} = \frac{\text{odds}(z_1 = 1|z_2 = 1, h = x)}{\text{odds}(z_1 = 1|z_2 = 0, h = x)},$$

and the odds ratio associated with a change in covariate h of Δ_x units:

$$\text{OR}_h = \frac{\text{odds}(z_1 = 1|z_2 = 0, h = x + \Delta_x)}{\text{odds}(z_1 = 1|z_2 = 0, h = x)}.$$

If desired, logarithms of odds ratios can be obtained, which produces a quantity with an interpretation similar to that of slope parameters of linear predictors. Insight into the relative strength of interspecific interactions vs. environmental variables can then be gained by comparing the absolute magnitude of $\log(OR_{sp})$ relative to $\log(OR_h)$, for example by calculating the probability $|\log(OR_h)| > |\log(OR_{sp})|$ from posterior samples in a Bayesian analysis. Note that care must be taken in selection of a meaningful Δ_x when h is a continuous variable.

The multispecies occupancy model developed above assumes detection and occurrence probabilities are modelled as a function of covariates via a multinomial logit link function. Alternatively, this model can be specified as multinomial probit regression, which allows exploitation of data-augmentation techniques and Gibbs sampling (Albert & Chib 1993; Dorazio & Rodríguez 2012; Johnson *et al.* 2013; Rota *et al.* 2016b). See Appendix S2 for a description of a multivariate probit regression model and details for constructing a Gibbs sampler.

EXAMPLE: CO-OCCURRENCE OF MESOCARNIVORES IN THE MID-ATLANTIC, USA

As an example, we fit the multispecies occupancy model to detection/non-detection data of bobcat, coyotes, grey fox and red fox obtained from camera traps deployed as part of the eMammal citizen science project (McShea *et al.* 2016; Rota *et al.* 2016a). These species are common throughout the south-eastern United States, are of similar body size and compete directly for small prey. Camera traps with motion detectors and infrared flash (Reconyx models RC55, PC800 and PC900; Reconyx, Inc. Holmen, WI, USA; and Bushnell Trophy Cam HD, Bushnell Outdoor Products, Overland Park, KS, USA) were deployed in 32 parks and protected areas across 6 Mid-Atlantic States and the District of Columbia. Between 27 and 80 camera traps were deployed within each park or protected area, generally in groups of 3. Within each group of 3, one camera was placed on a trail, another camera was placed 50 m from a trail, and the third was placed 100–200 m from a trail. Camera clusters in urban parks were spaced at least 100 m apart, while camera clusters in larger non-urban parks and protected areas were spaced at least 200 m apart. The assumption of independence between sites is unlikely to hold exactly for this data set (Rota *et al.* 2016b), though accounting for spatial dependence when modelling >2 species is still an open research problem. Cameras were attached to trees 40 cm above the ground and were deployed for 1–71 (mean = 22) consecutive days between August 2012 and December 2013. Data from cameras that were inadvertently placed too high or too low were discarded. At each camera site, the maximum distance the camera would trigger on a person (detection distance) was recorded. Citizen science volunteers set all cameras and identified animals in pictures, which were later confirmed by expert reviewers. A species was then considered detected at a camera trap during a given date if at least 1 photograph was obtained on that date, and was considered undetected otherwise. See Kays et al. (2015) and McShea *et al.* (2016) for additional details.

We fit a small set of candidate models meant to reflect hypotheses regarding the effects of interspecific interactions on occupancy and detection processes. For this example, we were primarily interested in the influence of human disturbance in occupancy probability as each of these species differ in their ability to infiltrate human-dominated landscapes (Lesmeister *et al.* 2015). We included three measures of human disturbance as covariates: the average daily number of hikers photographed at each camera trapping site; housing density in the surrounding 5 km (Stewart & Radeloff 2012); and the proportion of area disturbed or modified within a 5-km radius between 2001 and 2006

(Fry *et al.* 2011). We controlled for geographic variation in occupancy probability by including latitude and longitude of camera trap sites, and their interaction, as covariates. Our first set of models (models M_1, M_2, and M_3) all assume species-specific detection probabilities are a function of whether a camera was on or off a trail and the total detection distance of the camera. Model M_1 reflects the hypothesis that all four species occur independently and that marginal occupancy probabilities for each species are a function of geographic coordinates and a single human disturbance covariate (Table 1). We modelled coyote and grey fox marginal occupancy probabilities as a function of housing density in the surrounding 5 km because Lesmeister *et al.* (2015) found marginal occupancy probabilities were related to density of urban patches and distance to the nearest human structure, respectively. We modelled bobcat and red fox marginal occupancy probabilities as a function of hiker density because Wang, Allen & Wilmers (2015b) found bobcat activity was related to intensity of human trail use and Rota *et al.* (2016b) found red fox marginal occupancy probability was related to the number of hikers. Model M_2 still assumes occupancy probabilities are a function of geographic coordinates and human disturbance, but additionally reflects the hypothesis that species exhibited constant pairwise dependence. Model M_3 reflects the hypothesis that the relationship between human disturbance and occupancy probabilities for each species varies in the presence and absence of each of the other species. We modelled the probability that bobcats and coyote, coyote and grey fox, and grey fox and red fox occur together as a function of the proportion of recently disturbed habitat within 5 km; we modelled the probability that bobcat and grey fox, bobcat and red fox, and coyote and red fox occur together as a function of housing density within 5 km. There is little to guide decisions of what variables to use when modelling the probability that two or more of these species occur together, as this is one of the unique features of our model. Although we were interested in evaluating how the probability two species occupy the same site varies as a function of human disturbance, we were also interested in building parsimonious models for the purpose of this example. A more complete analysis could include more variables thought to influence the probability two or more species occur together. Our second set of models (models M_4, M_5 and M_6) reflect the same set of hypotheses regarding the occupancy process outlined for models M_1, M_2 and M_3, respectively, but now assume that detection probabilities of bobcat, grey fox and red fox are also a function of the latent presence/absence of coyotes. Coyotes are larger and potentially competitively superior to the other species considered and may alter detectability of interacting species when present. Note that we did not include higher-order interactions in any of our models (i.e. we assumed $f_{123} = f_{124} = f_{134} = f_{234} = f_{1234} = 0$), meaning we assumed the conditional probability 3 or more species occurred together was purely a function of species-specific (f_1, f_2, f_3, f_4) and pairwise interaction (f_{12}, f_{13}, f_{14}, f_{23}, f_{24}, f_{34}) parameters. We assumed logistic (0, 1) prior distributions for all parameters. We selected this prior distribution for slope parameters because it effectively induces a uniform distribution on the probability scale when assuming a logit link; that is, if $\theta \sim$ logistic(0,1), $\text{logit}^{-1}(\theta) \sim$ uniform(0, 1).

We compared our candidate models with Watanabe–Akaike Information Criterion (WAIC), a fully Bayesian information criterion analogous to Akaike Information Criterion (AIC) (Gelman *et al.* 2014; Hooten & Hobbs 2015). We fit models in Stan v. 2.8.0 (Stan Development Team 2015b) via the rstan v. 2.8.0 interface (Stan Development Team 2015a) in R v. 3.2.2. (R Core Team 2015). We ran 2 chains for each model, using trace plots to determine an adequate burn-in phase and subsequently running chains until the Brooks–Gelman–Rubin convergence diagnostic indicated adequate convergence ($\hat{R} \leq 1{\cdot}1$,

Table 1. Specification of natural parameters and conditional probabilities for the 6 candidate models considered in this study

Species	Natural parameter / Conditional probability	Models[1]		
		M_1 & M_4	M_2 & M_5	M_3 & M_6
Bobcat	$f_1 =$	$\alpha_1 + \beta_1\text{hike} + \mathbf{x}'\boldsymbol{\gamma}_1$	$\alpha_1 + \beta_1\text{hike} + \mathbf{x}'\boldsymbol{\gamma}_1$	$\alpha_1 + \beta_1\text{hike} + \mathbf{x}'\boldsymbol{\gamma}_1$
	$P(z_1 = 1\|\mathbf{Z}_{-1} = 0) =$	$\text{logit}^{-1}(f_1)$	$\text{logit}^{-1}(f_1)$	$\text{logit}^{-1}(f_1)$
Coyote	$f_2 =$	$\alpha_2 + \beta_2\text{hden} + \mathbf{x}'\boldsymbol{\gamma}_2$	$\alpha_2 + \beta_2\text{hden} + \mathbf{x}'\boldsymbol{\gamma}_2$	$\alpha_2 + \beta_2\text{hden} + \mathbf{x}'\boldsymbol{\gamma}_2$
	$P(z_2 = 1\|\mathbf{Z}_{-2} = 0) =$	$\text{logit}^{-1}(f_2)$	$\text{logit}^{-1}(f_2)$	$\text{logit}^{-1}(f_2)$
Grey fox	$f_3 =$	$\alpha_3 + \beta_3\text{hden} + \mathbf{x}'\boldsymbol{\gamma}_3$	$\alpha_3 + \beta_3\text{hden} + \mathbf{x}'\boldsymbol{\gamma}_3$	$\alpha_3 + \beta_3\text{hden} + \mathbf{x}'\boldsymbol{\gamma}_3$
	$P(z_3 = 1\|\mathbf{Z}_{-3} = 0) =$	$\text{logit}^{-1}(f_3)$	$\text{logit}^{-1}(f_3)$	$\text{logit}^{-1}(f_3)$
Red fox	$f_4 =$	$\alpha_4 + \beta_4\text{hike} + \mathbf{x}'\boldsymbol{\gamma}_4$	$\alpha_4 + \beta_4\text{hike} + \mathbf{x}'\boldsymbol{\gamma}_4$	$\alpha_4 + \beta_4\text{hike} + \mathbf{x}'\boldsymbol{\gamma}_4$
	$P(z_4 = 1\|\mathbf{Z}_{-4} = 0) =$	$\text{logit}^{-1}(f_4)$	$\text{logit}^{-1}(f_4)$	$\text{logit}^{-1}(f_4)$
Bobcat & Coyote	$f_{12} =$	0	δ_{12}	$\delta_{12} + \eta_{12}\text{dist}$
	$P(\mathbf{Z}_{1,2} = 1\|\mathbf{Z}_{3,4} = 0) =$	$\text{logit}^{-1}(f_1 + f_2)$	$\text{logit}^{-1}(f_1 + f_2 + f_{12})$	$\text{logit}^{-1}(f_1 + f_2 + f_{12})$
Bobcat & Grey fox	$f_{13} =$	0	δ_{13}	$\delta_{13} + \eta_{13}\text{hden}$
	$P(\mathbf{Z}_{1,3} = 1\|\mathbf{Z}_{2,4} = 0) =$	$\text{logit}^{-1}(f_1 + f_3)$	$\text{logit}^{-1}(f_1 + f_3 + f_{13})$	$\text{logit}^{-1}(f_1 + f_3 + f_{13})$
Bobcat & Red fox	$f_{14} =$	0	δ_{14}	$\delta_{14} + \eta_{14}\text{hden}$
	$P(\mathbf{Z}_{1,4} = 1\|\mathbf{Z}_{2,3} = 0) =$	$\text{logit}^{-1}(f_1 + f_4)$	$\text{logit}^{-1}(f_1 + f_4 + f_{14})$	$\text{logit}^{-1}(f_1 + f_4 + f_{14})$
Coyote & Grey fox	$f_{23} =$	0	δ_{23}	$\delta_{23} + \eta_{23}\text{dist}$
	$P(\mathbf{Z}_{2,3} = 1\|\mathbf{Z}_{1,4} = 0) =$	$\text{logit}^{-1}(f_2 + f_3)$	$\text{logit}^{-1}(f_2 + f_3 + f_{23})$	$\text{logit}^{-1}(f_2 + f_3 + f_{23})$
Coyote & Red fox	$f_{24} =$	0	δ_{24}	$\delta_{24} + \eta_{24}\text{hden}$
	$P(\mathbf{Z}_{2,4} = 1\|\mathbf{Z}_{1,3} = 0) =$	$\text{logit}^{-1}(f_2 + f_4)$	$\text{logit}^{-1}(f_2 + f_4 + f_{24})$	$\text{logit}^{-1}(f_2 + f_4 + f_{24})$
Grey fox & Red fox	$f_{34} =$	0	δ_{34}	$\delta_{34} + \eta_{34}\text{dist}$
	$P(\mathbf{Z}_{3,4} = 1\|\mathbf{Z}_{1,2} = 0) =$	$\text{logit}^{-1}(f_3 + f_4)$	$\text{logit}^{-1}(f_3 + f_4 + f_{34})$	$\text{logit}^{-1}(f_3 + f_4 + f_{34})$

[1]$\mathbf{x}$ is a vector of length 3 corresponding to latitude, longitude and their interaction, and $\boldsymbol{\gamma}$ is a conformable vector of slope parameters.
hike represents the average daily number of hikers photographed at a camera trap site.
hden represents housing density in the surrounding 5 km.
dist represents the proportion of recently disturbed habitat in the surrounding 5 km.

Gelman *et al.* 2014). All models achieved adequate convergence by running for 1000 iterations following a burn-in phase of 1000 iterations.

Results

The entire data set consisted of 1906 camera trap sites and 42 556 trap days. We obtained 429, 1237, 176 and 590 detection of bobcats, coyotes, grey fox and red fox, respectively, at 264, 540, 102 and 218 sites, respectively. Model selection provided clear evidence for interspecific dependence among the four species considered. The two top models (M_3 and M_2), ranked approximately 1 WAIC unit apart (Table 2), both assumed pairwise dependence between species. The only differences between these models were that M_3 assumed the probabilities of two species occupying the same site varied as a function of human disturbance, while M_2 assumed constant probabilities of two species of occupying the same site. The remaining models were ranked ≥ 10 WAIC units from models

Table 2. WAIC model ranks

Model	No. occupancy parameters	No. detection parameters	WAIC
M_3	32	12	20 065
M_2	26	12	20 066
M_4	20	15	20 076
M_5	26	15	20 093
M_1	20	12	20 106
M_6	32	15	20 119

M_3 and M_2, indicating there was no strong evidence that occupancy probabilities were independent between species (model M_1) or detection probabilities of bobcat, grey fox and red fox varied as a function of latent presence/absence of coyotes (models M_4, M_5 and M_6). Hereafter, we report results from our top ranked model, M_3, while acknowledging the presence of model selection uncertainty between models M_3 and M_2.

Daily detection probability of all species was higher on-trail than off-trail and varied as a function of camera detection distance for coyote and grey fox (Fig. 1). Marginal occupancy probabilities of all species in parks and protected areas varied across the study region. Bobcat and coyote were the two most widespread species: occupancy probability of bobcat was greatest in parks and protected areas in the north-western portion of the study area (Virgina and Tennessee), and occupancy probability of coyote was greatest in parks and protected areas in the western portion of the study area (Tennessee; Fig. 2). While grey fox and red fox were both detected throughout the study area, occupancy probability of grey fox in parks and protected areas was greatest in the southern portion of the study area (North Carolina and South Carolina), occupancy probability of red fox in parks and protected areas was greatest in the north-eastern portion of the study area (Maryland and Virginia), and these two species exhibited limited geographic overlap.

Bobcat, grey fox and red fox exhibited consistent, though modest, relationships between marginal occupancy probability and anthropogenic disturbance variables. Bobcat mean marginal occupancy probability consistently declined with increasing levels of anthropogenic disturbance, with the strongest
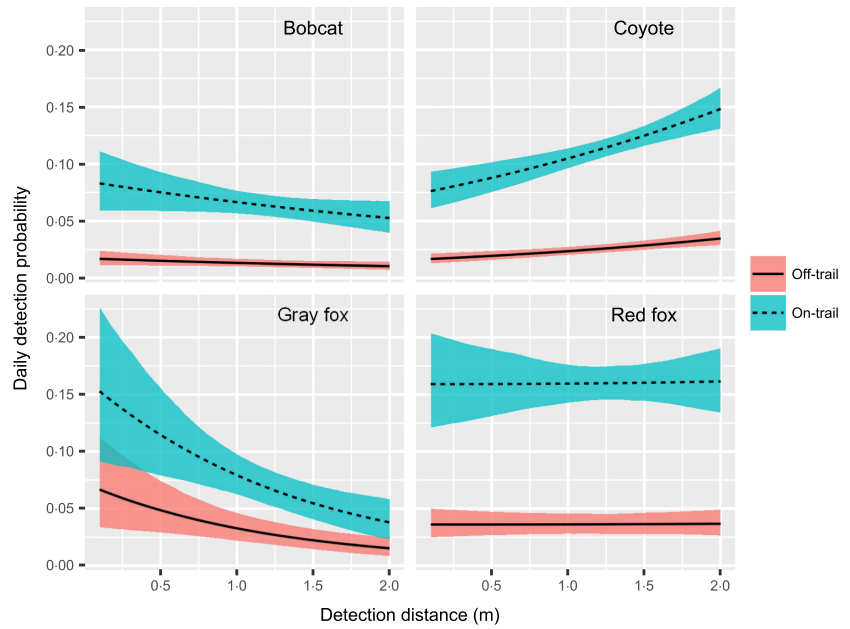
Fig. 1. Daily detection probability of bobcat (*Lynx rufus*), coyote (*Canis latrans*), grey fox (*Urocyon cinereoargenteus*) and red fox (*Vulpes vulpes*) across six states and the District of Columbia in the Mid-Atlantic, USA. Lines represents the mean posterior distributions and ribbons envelop 95% credible intervals.
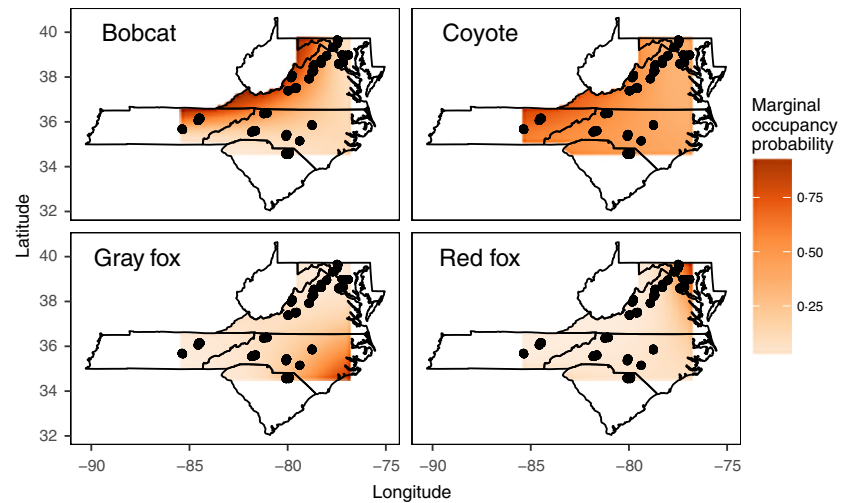

Fig. 2. Posterior mean marginal occupancy probability of bobcat (*Lynx rufus*), coyote (*Canis latrans*), grey fox (*Urocyon cinereoargenteus*) and red fox (*Vulpes vulpes*) in parks and protected areas across six states and the District of Columbia in the Mid-Atlantic, USA. Black dots represent camera trap sites, and polygons outline the six states within which camera traps were located (Maryland, North Carolina, South Carolina, Tennessee, Virginia and West Virginia). All other variables are assumed fixed at their observed mean.

relationship between marginal occupancy probability and the proportion of area recently disturbed within 5 km (Fig. 3). Grey fox and red fox mean marginal occupancy probability consistently increased with increasing levels of anthropogenic disturbance. Grey fox exhibited the strongest relationship between marginal occupancy probability and proportion of area recently disturbed, and red fox exhibited the strongest relationship between marginal occupancy probability and hiker density. Coyotes exhibited no strong relationships between marginal occupancy probability and any of the anthropogenic disturbance variables we examined.

We found evidence for interspecific interactions among most species pairs. We found evidence that occupancy probability of 1 species varied only in the presence and absence of another species for three species pairs: bobcat and red fox, coyote and red fox, and grey fox and red fox (i.e. credible intervals of intercept parameters in the linear models for f_{14}, f_{24} and f_{34} did not overlap 0). These constant pairwise interspecific interactions

are evident from the approximately parallel slopes presented in Fig. 4; for example, bobcats were less likely to occur at sites where red foxes were also present, regardless of the amount of disturbance within 5 km. We also found evidence that the probability two species occurred together varied as a function of the proportion of area recently disturbed within 5 km for two species pairs: coyote and bobcat, and coyote and grey fox (Fig. 4). The relationship between the probability of coyote occurrence and the proportion of area recently disturbed within 5 km varied markedly depending on whether bobcats and grey fox were present. At low levels of recent disturbance, coyotes were more likely to occupy sites if bobcats were present, and occurred largely independently of grey fox. At high levels of recent disturbance within 5 km, however, coyotes were more likely to occur at sites where grey fox were also present, and occurred largely independently of bobcats. The probability that any of the two species occurred together did not vary appreciably as a function of housing density (i.e. credible
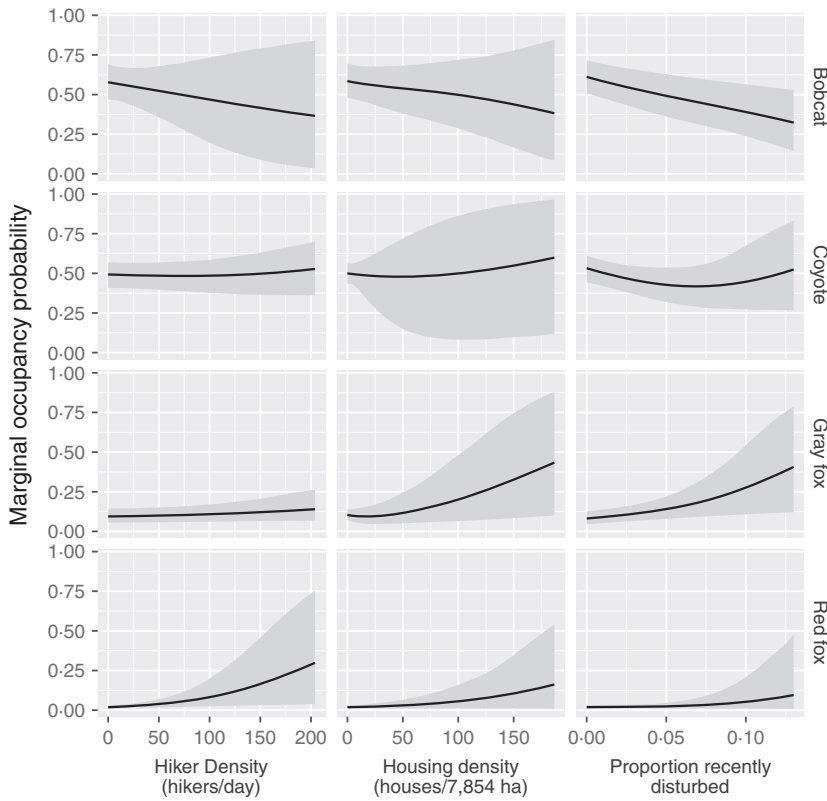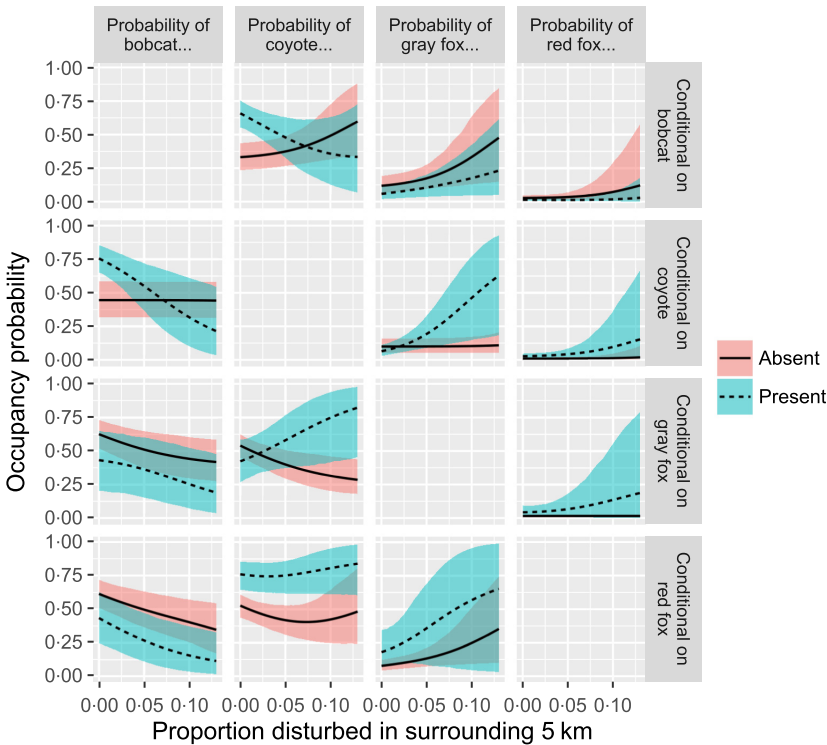
Fig. 3. Marginal occupancy probability of bobcat (*Lynx rufus*, row 1), coyote (*Canis latrans*, row 2), grey fox (*Urocyon cinereoargenteus*, row 3) and red fox (*Vulpes vulpes*, row 4) in the Mid-Atlantic USA as a function of the average number of hikers photographed per day (column 1), housing density within a 5 km radius (column 2) and the proportion of area recently disturbed within 5 km (column 3). Solid lines represents the mean posterior distributions, and grey ribbons envelop 95% credible intervals. All variables not included in a plot are assumed fixed at their observed mean.

Fig. 4. Occupancy probability of bobcat (*Lynx rufus*), coyote (*Canis latrans*), grey fox (*Urocyon cinereoargenteus*) and red fox (*Vulpes vulpes*) conditional on the presence and absence of each of the other species. The occupancy probability of the species in each column is conditional on the presence and absence of the species in each row; for example, the plot in column 1, row 2 represents bobcat occupancy probability, conditional on the presence and absence of coyote. Lines represent posterior means, and ribbons envelop 95% credible intervals. All variables not included in a plot are assumed fixed at their observed mean. Additionally, conditional plots are marginalized over the 2 species that do not occur in a plot; for example, the plot in column 1, row 2 sums over all combinations of grey fox and red fox presence and absence.

intervals of slope parameters for the housing density variables in linear models for f_{13}, f_{14} and f_{24} overlapped 0; see Table 1) or hiker density.

Our ability to model the probability two species occupy the same site as a function of covariates provided insight

into factors driving marginal occupancy probabilities that might not have been evident otherwise; for example, we assumed bobcat occupancy probability only varied as a function of the proportion of area recently disturbed within 5 km when coyotes were also present (i.e. we assumed the

natural parameter f_{12} was a function of the proportion of area recently disturbed within 5 km, while f_1 was not a function of this variable). The observed relationship between marginal occupancy probability of bobcats and disturbance within 5 km (Fig. 3) is therefore strongly linked to interspecific interactions with coyotes. Additionally, the observed relationship between coyote marginal occupancy probability and recent disturbance within 5 km (Fig. 3) masks the underlying interactions with bobcat and grey fox along this gradient (Fig. 4). This ability to model interspecific interactions as a function of covariates is a novel advancement of our proposed model and allows deeper insights into factors driving marginal occupancy probabilities that are not possible with other joint species distribution models.

Our results demonstrate that occupancy probabilities of individual species were related both to environmental variables and the presence or absence of other interacting species. The relative importance of environmental variables vs. interspecific interactions in shaping occupancy probabilities varied strongly, however, depending on the nature of interspecific interactions; for example, at low proportions (0·015) of recent disturbance within 5 km coyote and grey fox occurred largely independently of each other. Therefore, at this level of recent disturbance, moving west from a reference location of $\{37\cdot2°N, -79\cdot7°W\}$ (mean geographic coordinates of study sites) to a new location at $\{37\cdot2°N, -82\cdot0°W\}$ (approximately 200 km, equivalent to 1 standard deviation of longitude variables) had a stronger influence on coyote occupancy probability than the presence or absence of grey fox $(P[|\log(OR_h)| > |\log(OR_{sp})|] = 0\cdot97)$. At high proportions (0·12) of recent disturbance within 5 km, however, coyote are much more likely to occur at sites where grey fox are also present. At this level of recent disturbance, moving west as before from the same reference location has a weaker influence on coyote occupancy probability than the presence or absence of grey fox $(P[|\log(OR_h)| > |\log(OR_{sp})|] = 0\cdot01)$. Though there are many more relationships we could report, this example demonstrates the additional insights that can be gained by quantifying the relative strength of environmental variables vs. interspecific interactions in shaping occupancy probabilities when both processes are acting simultaneously.

Discussion

The distribution of species in space is simultaneously influenced by environmental variables and interactions with other species. Recent statistical advances have allowed ecologists to model detection/non-detection data for two species while accounting for imperfect detection (MacKenzie, Bailey & Nichols 2004; Richmond, Hines & Beissinger 2010; Waddle *et al.* 2010; Rota *et al.* 2016b) or for >2 species without accounting for imperfect detection (Pollock *et al.* 2014). Our proposed method overcomes the limitations of these recent advances and permits simultaneous modelling of detection/non-detection data for an arbitrary number of species while accommodating imperfect detection. The utility of our proposed model is highlighted in our example, where we found strong evidence for interactions among multiple species pairs that varied along environmental gradients. The ability of our proposed method to model the probability two or more species occur together as a function of covariates is an additional advancement relative to current approaches, which have either proved numerically unstable when modelling co-occurrence probabilities as a function of covariates (MacKenzie, Bailey & Nichols 2004) or can only model constant pairwise covariance between species pairs (Pollock *et al.* 2014). Although we cannot directly infer mechanisms that lead to observed patterns of co-occurrence, the ability to model the probability that two or more species occur together does allow greater insight into processes that shape these patterns and represents an important improvement in our ability to model species interactions. Finally, even though we develop this model in the context of interspecific interactions, it could potentially be used to evaluate intraspecific interactions, such as between individuals of different sexes or age classes.

The dimension of the probability vector Ψ, which describes the probability of all possible combinations species of presence/absence, grows exponentially as the number of species increases. While the total number of possible combinations of species presence/absence can be large when modelling many species simultaneously, the number of parameters required to fit such a model need not be large; for example, modelling detection/non-detection data for 10 species would result in a $2^{10} = 1024$-dimensional probability vector Ψ. Nonetheless, a model for the latent occupancy state could be fit with as few as 10 parameters by assuming independence between all species and constant occupancy probabilities. A more complex model could continue to assume independence between species, but also assume marginal occupancy probabilities of each species are a function of covariates. Such a model would be equivalent to fitting 10 independent occupancy models. Adding yet more complexity, a model assuming constant pairwise covariance between all species pairs adds $\binom{10}{2} = 45$ additional parameters and produces a model of the latent occupancy state that would require as many covariance parameters as Pollock *et al.* (2014). Our proposed model has the flexibility to add yet more complexity, as appropriate, by assuming that the probability two or more species occupy the same site is a function of covariates. This level of complexity can also quickly increase the number of parameters because of the high number of pairwise (and higher-order) interactions. Finding a parsimonious model in such situations can prove challenging, but can be aided with model-based variable selection techniques such as indicator variable selection and reversible jump MCMC (reviewed in O'Hara & Sillanpää 2009 and Hooten & Hobbs 2015).

We developed our proposed method as a multispecies 'occupancy model' because we believe that detection/non-detection data are subject to imperfect detection in general. While accounting for imperfect detection is appropriate in

many settings, Welsh, Lindenmayer & Donnelly (2013 see also Guillera-Arroita *et al.* 2014) describe situations where the mean squared error of estimated occupancy probability is lower when assuming perfect detection. If there is reason to assume perfect detection, the multivariate Bernoulli model can still be used to model detection/non-detection data of multiple species simultaneously. Indeed, the 'MVB' package (Dai 2013) within R (R Core Team 2015) can be used to fit loglinear multivariate Bernoulli models that implicitly assume perfect detection.

Ecologists have long recognized the role of interspecific interactions in structuring ecological communities (e.g. MacArthur 1972). Interspecific interactions may often be as important as biotic and abiotic conditions in determining where species occur in space and ecologists are increasingly recognizing the need to accommodate such interactions when modelling habitat selection (McLoughlin *et al.* 2010) and species distributions (Godsoe & Harmon 2012). Accounting for interspecific interactions will be important when predicting future distributions, particularly in response to global climate change, as rates of range expansion and contraction are likely to be influenced by co-occurring species within a community (Svenning *et al.* 2014). By modelling occupancy probability of multiple co-occurring species, all of which are imperfectly detected, this model represents a step towards a better understanding of how interspecific interactions shape habitat selection and species distributions.

Acknowledgments

Data accessibility

Data, R and Stan code are deposited in the Dryad repository http://dx.doi.org/10.5061/dryad.pq624.

References

Albert, J.H. & Chib, S. (1993) Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, **88**, 669–679.

Bailey, L.L., Reid, J.A., Forsman, E.D. & Nichols, J.D. (2009) Modeling cooccurrence of northern spotted and barred owls: Accounting for detection probability differences. *Biological Conservation*, **142**, 2983–2989.

Byholm, P., Burgas, D., Virtanen, T. & Valkama, J. (2012) Competitive exclusion within the predator community influences the distribution of a threatened prey species. *Ecology*, **93**, 1802–1808.

Coleman, B.T. & Hill, R.A. (2014) Living in a landscape of fear: the impact of predation, resource availability and habitat structure on primate range use. *Animal Behaviour*, **88**, 165–173.

Dai, B. (2013) MVB: Mutivariate Bernoulli log-linear model. R package version 1.1.

Dai, B., Ding, S. & Wahba, G. (2013) Multivariate bernoulli distribution. *Bernoulli*, **19**, 1465–1483.

Dorazio, R.M. & Rodríguez, D.T. (2012) A gibbs sampler for bayesian analysis of site-occupancy data. *Methods in Ecology and Evolution*, **3**, 1093–1098.

Fry, J., Xian, G., Jin, S., Dewitz, J., Homer, C., Yang, L., Barnes, C., Herold, N. & Wickham, J. (2011) Completion of the 2006 national land cover database for the conterminous United States. *Photogrammetric Engineering & Remote Sensing*, **77**, 858–864.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A. & Rubin, D.B. (2014) *Bayesian data analysis*, 3rd edn. CRC Press, Boca Raton.

Godsoe, W. & Harmon, L.J. (2012) How do species interactions affect species distribution models? *Ecography*, **35**, 811–820.

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**, e99571.

Hooten, M.B. & Hobbs, N.T. (2015) A guide to Bayesian model selection for ecologists. *Ecological Monographs*, **85**, 3–28.

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**, 801–808.

Kays, R., Costello, R., Forrester, T., Baker, M.C., Parsons, A.W., Kalies, E.L., Hess, G., Millspaugh, J.J. & McShea, W. (2015) Cats are rare where coyotes roam. *Journal of Mammalogy*, **96**, 981–987.

Latimer, A.M., Banerjee, S., Sang, H., Mosher, E.S. & Silander Jr, J.A. (2009) Hierarchical models facilitate spatial analysis of large data sets: A case study on invasive plant species in the northeastern united states. *Ecology Letters*, **12**, 144–154.

Lesmeister, D.B., Nielsen, C.K., Schauber, E.M. & Hellgren, E.C. (2015) Spatial and temporal structure of a mesocarnivore guild in midwestern North America. *Wildlife Monographs*, **191**, 1–61.

MacArthur, R. (1972) *Geographical Ecology: Patterns in the Distribution of Species*. Princeton University Press, Princeton.

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**, 2248–2255.

MacKenzie, D.I., Bailey, L.L. & Nichols, J.D. (2004) Investigating species cooccurrence patterns when species are detected imperfectly. *Journal of Animal Ecology*, **73**, 546–555.

MacKenzie, D.I., Nichols, J.D., Royle, J.A., Pollock, K.H., Bailey, L.L. & Hines, J.E. (2006) *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*. Elsevier, Burlington.

McLoughlin, P.D., Morris, D.W., Fortin, D., Vander Wal, E. & Contasti, A.L. (2010) Considering ecological dynamics in resource selection functions. *Journal of Animal Ecology*, **79**, 4–12.

McShea, W.J., Forrester, T., Costello, R., He, Z. & Kays, R. (2016) Volunteer run cameras as distributed sensors for macrosystem mammal research. *Landscape Ecology*, **31**, 55–66.

O'Hara, R.B. & Sillanpää, M.J. (2009) A review of bayesian variable selection methods: What, how and which. *Bayesian Analysis*, **4**, 85–118.

Ovaskainen, O., Hottola, J. & Siitonen, J. (2010) Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. *Ecology*, **91**, 2514–2521.

Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk, P.A. & McCarthy, M.A. (2014) Understanding co-occurrence by modelling species simultaneously with a joint species distribution model (jsdm). *Methods in Ecology and Evolution*, **5**, 397–406.

R Core Team (2015) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.

Richmond, O.M.W., Hines, J.E. & Beissinger, S.R. (2010) Two-species occupancy models: a new parameterization applied to co-occurrence of secretive rails. *Ecological Applications*, **20**, 2036–2046.

Rota, C.T., Ferreira, M.A.R., Kays, R.W., Forrester, T.D., Kalies, E.L., McShea, W.J., Parsons, A.W. & Millspaugh, J.J. (2016a) Data from: A multispecies occupancy model for two or more interacting species. *Dryad Digital Repository*, http://dx.doi.org/10.5061/dryad.pq624.

Rota, C.T., Wikle, C.K., Kays, R.W., Forrester, T.D., McShea, W.J., Parsons, A.W. & Millspaugh, J.J. (2016b) A two-species occupancy model accommodating simultaneous spatial and interspecific dependence. *Ecology*, **97**, 48–53.

Schabenberger, O. & Pierce, F.J. (2002) *Contemporary statistical models for the plant and social sciences*. CRC Press, Boca Raton.

Sherry, T.W. & Holmes, R.T. (1988) Habitat selection by breeding american redstarts in response to the dominant competitor, the least flycatcher. *Auk*, **105**, 350–364.

Stan Development Team (2015a) Rstan: the R interface to Stan, version 2.8.0.

Stan Development Team (2015b) Stan: a C++ library for probability and sampling, version 2.8.0.

Steen, D.A., McClure, C.J.W., Brock, J.C., Craig Rudolph, D., Pierce, J.B., Lee, J.R. *et al.* (2014) Snake co-occurrence patterns are best explained by habitat and hypothesized effects of interspecific interactions. *Journal of Animal Ecology*, **83**, 286–295.

Stewart, S. & Radeloff, V.C. (2012) 2010 U.S. housing density block group level. http://silvis.forest.wisc.edu/maps/blk_pla/2010/download. Accessed 21 January, 2016.

Svenning, J.C., Gravel, D., Holt, R.D., Schurr, F.M., Thuiller, W., Mun¨kemül-ler, T. *et al.* (2014) The influence of interspecific interactions on species range expansion rates. *Ecography*, **37**, 1198–1209.

Veblen, K.E. (2012) Savanna glade hotspots: Plant community development and synergy with large herbivores. *Journal of Arid Environments*, **78**, 119–127.

Waddle, J.H., Dorazio, R.M., Walls, S.C., Rice, K.G., Beauchamp, J., Schuman, M.J. & Mazzotti, F.J. (2010) A new parameterization for estimating co-occurrence of interacting species. *Ecological Applications*, **20**, 1467–1475.

Wang, F., McShea, W.J., Wang, D. & Li, S. (2015a) Shared resources between giant panda and sympatric wild and domestic mammals. *Biological Conservation*, **186**, 319–325.

Wang, Y., Allen, M.L. & Wilmers, C.C. (2015b) Mesopredator spatial and temporal responses to large predators and human development in the Santa Cruz mountains of California. *Biological Conservation*, **190**, 23–33.

Welsh, A.H., Lindenmayer, D.B. & Donnelly, C.F. (2013) Fitting and interpreting occupancy models. *PLoS One*, **8**, e52015.

Willems, E.P. & Hill, R.A. (2009) Predator-specific landscapes of fear and resource distribution: effects on spatial range use. *Ecology*, **90**, 546–555.

Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Appendix S1. Interpretation of slope parameters from a multivariate Bernoulli logit-linear model.

Appendix S2. Description of a multinomial probit multi-species occupancy model.