RESEARCH ARTICLE

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

A multistate dynamic occupancy model to estimate local colonization–extinction rates and patterns of co-occurrence between two or more interacting species

Mason Fidino[1] 🆔 | Juniper L. Simonis[2] | Seth B. Magle[1]

[1]Urban Wildlife Institute, Lincoln Park Zoo, Chicago, Illinois

[2]Dapper Stats, Portland, Oregon

Correspondence
Mason Fidino
Email: mfidino@lpzoo.org

Abstract

1. Although ecology is rife with theory that explores how multiple species co-occur through space and time, the field lacks robust statistical models to parameterize this theory with empirical data, particularly when species are detected imperfectly and data are collected as a time-series.

2. We address this need by developing an occupancy model that estimates local colonization and extinction rates for two or more interacting species when data are collected across multiple sampling occasions. This model estimates how community composition at a site may change across sampling occasions by assuming the latent occupancy state is a categorical random variable. We used a multinomial-logit model to parameterize species specific parameters and pairwise interactions between species, both of which can be made a function of covariates. These transition probabilities between community states can then be converted to occupancy or co-occurrence probabilities to determine how community composition varies along an environmental gradient or through time.

3. As an example, we estimate patterns of co-occurrence between coyote *Canis latrans*, Virginia opossum *Didelphis virginiana*, and raccoon *Procyon lotor* in Chicago, Illinois, USA with data from a multiyear camera trapping study. Models with pairwise interactions between species greatly out performed models that assumed independence between species. Opossum and raccoon, for example, were far less likely to go extinct in habitat patches where coyotes were present.

4. Community composition at a site depends on species interactions and the local environment. Our model can separate such effects by estimating the underlying processes that define species occurrence patterns. As a result, our model can more explicitly quantify a wide range of ecological dynamics and therefore be used to empirically test ecological theory, such as estimating priority effects at a site or turnover rates between species, both of which can be made to vary as a function of covariates.

KEYWORDS
Bayesian models, camera traps, co-occurrence, occupancy models, species interactions

1 | INTRODUCTION

Many ecological investigations focus on patterns of species occupancy rather than abundance, as occupancy-based sampling designs require less effort and are well-suited to answer questions about species distributions (MacArthur & Wilson, 1967), metapopulation (Levins, 1969) or metacommunity dynamics (Holyoak, Leibold, & Holt, 2005), and species invasions (Elton, 1958). However, providing robust estimates of species occupancy patterns requires correctly partitioning observed variation between underlying ecological and observational processes. Estimated occupancy probabilities can be highly biased when observational error or sampling design are not properly addressed in a model (MacKenzie et al., 2002). Indeed, the realization of ecologists' fallibility and the impact this may have on our inferential capacity has been a motivating force in the development of statistical methods that account for observational error (MacKenzie, Royle, Pollock, Bailey, & Hines, 2006).

Among the statistical techniques that correct for species detectability, the patch occupancy modelling framework developed by MacKenzie et al. (2002) has proven to be an effective technique to understand and model species distributions (Bailey, MacKenzie, & Nichols, 2014; hereafter referred to as occupancy models). Since their introduction over 15 years ago (MacKenzie et al., 2002), occupancy models have been developed to incorporate multiple seasons (MacKenzie et al., 2006), patterns of co-occurrence for any number of species within a single season (Rota, Ferreira, et al., 2016), patterns of co-occurrence for a dominant and subordinate species across multiple seasons (MacKenzie et al., 2006), and dynamic processes for entire communities through the inclusion of continuous and discrete covariates (Dorazio, Kery, Royle, & Plattner, 2010). These models produce robust likelihood estimates, facilitating the testing of hypotheses on species occupancy or colonization-extinction dynamics, all while accounting for the vagaries of sampling.

Although Rota, Ferreira, et al. (2016) developed an occupancy model to test how species interactions and features of the physical environment influence species occupancy for ≥2 species, their model lacks the ability to estimate these factors across multiple primary sampling periods (i.e., through time). This is unfortunate as many re-introduction programmes, biodiversity surveys, and metapopulation studies monitor the occupancy status of species at sites for multiple years to quantify temporal changes in species' distributions or their colonization-extinction dynamics (Broms, Hooten, Johnson, Altwegg, & Conquest, 2016; Fidino & Magle, 2017). Furthermore, while models exist that can estimate the impact one species may have on another's colonization or extinction rate they either lack the ability to include continuous covariates (Miller, Brehme, Hines, Nichols, & Fisher, 2012), can only estimate dominant-subordinate interactions (Jones & Kroll, 2016), or can only be fit to two species (Chambert et al., 2018; Yackulic et al., 2014). As patch colonization and extinction are the processes that define a species' occupancy pattern across a landscape there is a clear need to develop a class of model that can estimate how interspecific interactions and changes in the local environment influence local colonization and extinction rates.

Here, we introduce a multivariate occupancy model for the analysis of potentially interacting species when data are collected across multiple seasons. Similar to Rota, Ferreira, et al. (2016), our model assumes that the partially observed true occupancy state is a multivariate Bernoulli random variable. We further generalize this framework so that data collected at locations across multiple sampling periods can be used to estimate species local colonization and extinction probabilities. After introducing our model, we use it to estimate co-occurrence patterns between coyote *Canis latrans*, Virginia opossum *Didelphis virginiana*, hereafter opossum, and raccoon *Procyon lotor* with data from a large-scale long-term camera trapping study of medium to large mammals in Chicago, Illinois, USA. Overall, our model is a robust and general way to estimate patterns of co-occurrence and interspecific associations between species across space and time.

2 | MATERIALS AND METHODS

2.1 | Sampling protocol and data structure

The sampling protocol for our dynamic co-occurrence model (DCOM) is similar to the multiple-season protocol developed by MacKenzie et al. (2006) where k in 1, ..., K sampling locations (i.e., sites) are randomly selected within an area of interest and each site is surveyed over t in 1, ..., T sampling occasions (hereafter seasons). Within each season j in 1, ..., J repeat surveys occur wherein detection/nondetection data of I species across sites are recorded. Like the multiple-season protocol, we assume closure within a single season (i.e., the occupancy status of species i at a site does not change across the J repeat surveys conducted during season t).

This sampling protocol allows for the estimation of local colonization and extinction events that occur between seasons which could vary via local habitat covariates or the presence of other species. We define colonization as the probability that a species occupies a site at time t that it did not at $t − 1$ and extinction as the probability that a species is absent from a site at time t that it occupied at $t − 1$. To model changes in community composition through time we deviate from how these data are represented in the classic multi-season sampling protocol to use the Categorical distribution, an alternative parameterization of the multivariate Bernoulli distribution (Robert, 2014). Detections ($y_{iktj} = 1$) and nondetections ($y_{iktj} = 0$) are typically modelled as Bernoulli trials of the underlying occupancy state, $y_{iktj}|z_{ikt} \sim \text{Bernoulli}(p_{iktj}z_{ikt})$, where p_{iktj} is the conditional probability of detecting species i at site k during season t on survey j and z_{ikt} is the true occupancy state of species i at site k during season t. Instead, we denote the 2^I possible combinations of species observed on a single survey as r in 1, ..., R community states. Two species, A and B, have $R = 4$ possible community states at a site: unoccupied (U), species A (A), species B (B), and both species (AB). These community states can be represented as the vector [1, 2, 3, 4] and the probability of transitioning from one state to the next between seasons can be estimated. For tractability, we describe this model for two species but it can be generalized to 3 or more species as seen in our example.

2.2 | The dynamic co-occurrence model

For ease of presentation, we assume that occupancy (ψ), colonization (γ), extinction (ε), and detection (p) probabilities are constant across time, surveys, and sites. Let z_{kt} denote the true community state at site k on time t, which could take one of $R = 4$ states for two species. We assume that z_{kt} depends on the community state at time $t - 1$ and a mixture of local colonization and extinction rates. Because there is no information on the community state at a site before sampling begins we model the latent community state at time $t = 1$ as the following Categorical random variable:

state from one time step to the next. If each transition is estimated separately in ϕ then $R \times R$ parameters must be estimated from the data. However, varying constraints may be used to represent different ecological hypotheses or reduce the number of estimated parameters (MacKenzie et al., 2006). In this two species example we model transitions between community states, or columns of ϕ, with eight parameters (Table 1).

Different combinations of the parameters in Table 1 generate the component rates of ϕ that describe how the local community transitions between the R states:

$$\text{From state}$$

$$\phi = \begin{array}{c} \\ \text{To state} \end{array} \begin{array}{cccc} & 1_{[U]} & 2_{[A]} & 3_{[B]} & 4_{[AB]} \\ 1_{[U]} \\ 2_{[A]} \\ 3_{[B]} \\ 4_{[AB]} \end{array} \begin{bmatrix} (1 - \gamma^A)(1 - \gamma^B) & \varepsilon^A(1 - \gamma^{B|A}) & (1 - \gamma^{A|B})\varepsilon^B & \varepsilon^{A|B}\varepsilon^{B|A} \\ \gamma^A(1 - \gamma^B) & (1 - \varepsilon^A)(1 - \gamma^{B|A}) & \gamma^{A|B}\varepsilon^B & (1 - \varepsilon^{A|B})\varepsilon^{B|A} \\ (1 - \gamma^A)\gamma^B & \varepsilon^A\gamma^{B|A} & (1 - \gamma^{A|B})(1 - \varepsilon^B) & \varepsilon^{A|B}(1 - \varepsilon^{B|A}) \\ \gamma^A\gamma^B & (1 - \varepsilon^A)\gamma^{B|A} & \gamma^{A|B}(1 - \varepsilon^B) & (1 - \varepsilon^{A|B})(1 - \varepsilon^{B|A}) \end{bmatrix} \quad (2)$$

$$z_{kt=1} \sim \text{Categorical}(\psi), \quad (1)$$

where ψ is a vector of length R that contains the probability $z_{kt=1}$ is in each of the R community states such that $\sum_{r=1}^{R} \psi_r = 1$. To do so, $R - 1$ of these state probabilities are estimated independently while one is obtained through subtraction (e.g., $\psi = [(1 - \psi_2 - \psi_3 - \psi_4), \psi_2, \psi_3, \psi_4]$).

Following the first season, the community state at a site can either stay the same or transition to a new community state in the following season. Assuming that transitions between states are conditional on the community state in the previous time step we can store these transition probabilities in square $R \times R$ community transition matrix (ϕ) whose columns sum to 1. Similar to Miller et al. (2012), the matrix ϕ contains the component rates that describe the probability a community moves between states or stays in the same

In Equation 2, we assume a species colonization and extinction rate depends on the presence or absence of other potentially interacting species in the previous time step. This parameterization can be used to test a variety of ecological hypotheses. For example, the impact that one or multiple species in a habitat patch has on the likelihood of other species colonizing that patch (i.e., priority effects) could be estimated from the posteriors of a Bayesian analysis as $P(\gamma^A > \gamma^{A|B})$ or $P(\gamma^B > \gamma^{B|A})$ for inhibitory effects and $P(\gamma^A < \gamma^{A|B})$ or $P(\gamma^B < \gamma^{B|A})$ for facilitative effects. Additionally, while Equation 2 assumes the probability one species replaces another (i.e., turnover rates) are either $\varepsilon^A\gamma^{B|A}$ or $\gamma^{A|B}\varepsilon^B$, it may be of interest to estimate these rates separately when modelling the spread of an invasive species (MacKenzie et al., 2006). Equation 2 also illustrates the primary motivation for describing community states numerically as the community state at time $t - 1$ indexes which column of ϕ is used to estimate the state at time t:

TABLE 1 Probabilities used to estimate changes between community states through time in a two species dynamic co-occurrence model

Parameter	Verbal description
γ^A	The colonization probability for species A given species B absence at time $t - 1$
γ^B	The colonization probability for species B given species A absence at time $t - 1$
$\gamma^{A|B}$	The colonization probability for species A given species B presence at time $t - 1$
$\gamma^{B|A}$	The colonization probability for species B given species A presence at time $t - 1$
ε^A	The extinction probability for species A given species B absence at time $t - 1$
ε^B	The extinction probability for species B given species A absence at time $t - 1$
$\varepsilon^{A|B}$	The extinction probability for species A given species B presence at time $t - 1$
$\varepsilon^{B|A}$	The extinction probability for species B given species A presence at time $t - 1$

$$z_{kt}|z_{kt-1},\gamma,\varepsilon \sim \text{Categorical}(\phi_{z_{kt-1}}) \text{ for } t=2,\dots,T. \tag{3}$$

Here $\phi_{z_{kt-1}}$ is a vector of R transition probabilities from column z_{kt-1} of ϕ. The partially observed process model that estimates the true community state, z_{kt}, is linked to the observed data, y_{ktj}, through a detection model:

$$y_{ktj}|z_{kt} \sim \text{Categorical}(\lambda_{z_{kt}}) \tag{4}$$

where $\lambda_{z_{kt}}$ is column z_{kt} of an $R \times R$ community detection matrix (λ) whose columns sum to 1. The elements in λ contain the probability of detecting each state given the true state:

$$
\lambda =
\begin{array}{c}
\text{True state} \\
\begin{array}{cccc}
1_{[U]} & 2_{[A]} & 3_{[B]} & 4_{[AB]}
\end{array} \\
\begin{array}{c}
1_{[U]} \\ 2_{[A]} \\ 3_{[B]} \\ 4_{[AB]}
\end{array}
\begin{bmatrix}
1 & (1-\rho^A) & (1-\rho^B) & (1-\rho^{A|B})(1-\rho^{B|A}) \\
0 & \rho^A & 0 & \rho^{A|B}(1-\rho^{B|A}) \\
0 & 0 & \rho^B & (1-\rho^{A|B})\rho^{B|A} \\
0 & 0 & 0 & \rho^{A|B}\rho^{B|A}
\end{bmatrix}
\end{array} \tag{5}
$$

(where the matrix row labels read "Observed state")

Here ρ^A and $\rho^{A|B}$ are respectively the probabilities of detecting species A given the absence (ρ^A) or presence ($\rho^{A|B}$) of species B, while ρ^B and $\rho^{B|A}$ are the probabilities of detecting species B given the absence (ρ^B) or presence ($\rho^{B|A}$) of species A. If detection probabilities are thought to be independent of the presence or absence of other species, $\rho^{A|B}$ and $\rho^{B|A}$ can simply be exchanged with ρ^A and ρ^B.

2.3 | Including covariates

The presence of one species at a site is not likely the sole factor influencing the presence of another species. Local environmental characteristics may also be important predictors of the presence of wildlife. Likewise, the influence one species may have on another could vary across environmental gradients (LaManna, Hemenway, Boccadori, & Martin, 2015) or habitat types (Atuo & O'Connell, 2017). Covariate information of this type can be included in any of the parameters of this DCOM.

Following Rota, Ferreira, et al. (2016), the R community state probabilities at site k and time $t = 1$ (Equation 1) are modelled as a function of covariates with the multinomial-logit link:

$$\psi_k^U = \psi_k^1 = \frac{1}{1+\exp(\beta_k^{\psi_2})+\exp(\beta_k^{\psi_3})+\exp(\beta_k^{\psi_2}+\beta_k^{\psi_3}+\beta_k^{\psi_4})}$$

$$\psi_k^A = \psi_k^2 = \frac{\exp(\beta_k^{\psi_2})}{1+\exp(\beta_k^{\psi_2})+\exp(\beta_k^{\psi_3})+\exp(\beta_k^{\psi_2}+\beta_k^{\psi_3}+\beta_k^{\psi_4})}$$

$$\psi_k^B = \psi_k^3 = \frac{\exp(\beta_k^{\psi_3})}{1+\exp(\beta_k^{\psi_2})+\exp(\beta_k^{\psi_3})+\exp(\beta_k^{\psi_2}+\beta_k^{\psi_3}+\beta_k^{\psi_4})} \tag{6}$$

$$\psi_k^{AB} = \psi_k^4 = \frac{\exp(\beta_k^{\psi_2}+\beta_k^{\psi_3}+\beta_k^{\psi_4})}{1+\exp(\beta_k^{\psi_2})+\exp(\beta_k^{\psi_3})+\exp(\beta_k^{\psi_2}+\beta_k^{\psi_3}+\beta_k^{\psi_4})}$$

where β_k are site-level logit-linear predictors of the R states

$$\beta_k^{\psi_1} = \exp(0) = 1$$
$$\beta_k^{\psi_2} = x_{ak}^\top \alpha_2$$
$$\beta_k^{\psi_3} = x_{ak}^\top \alpha_3 \tag{7}$$
$$\beta_k^{\psi_4} = x_{ak}^\top \alpha_4.$$

We have set the first probability, ψ_1, as a reference category (i.e., $\beta_k^{\psi_1} = 0$). Calculating the reference outcome in this way is identical to obtaining it through subtraction. For Equation 7, x_k is a vector of habitat covariates measured at site k whereas α_2, α_3, and α_4 are conformable parameter vectors, including the intercept. To accommodate the intercept, the first element of x_k must be 1. For ease of exposition x_k has the same dimensions and values within Equation 7 but this is not a requirement. Different covariates could be used for each probability.

The remaining parameters in Equation 7 can be best understood by determining the number of species they are associated with, which Dai, Ding, and Wahba (2013) define as the 'order' of a parameter. For S species and order level O there are $\binom{S}{O}$ parameters of that order. With two species, Equation 7 has two 1st order occupancy parameters, $\beta_k^{\psi_2}$ and $\beta_k^{\psi_3}$, which represent the log odds that species A and B occupy site k in the absence of one another. $\beta_k^{\psi_4}$, a 2nd order parameter, is the difference in log odds when both species are present compared to when one species is present at a site. Therefore, 2nd order parameters are associated with the probability that two species occur together via Equation 6. A three species system would have three 1st order parameters (A, B, C), three 2nd order parameters (AB, BC, AC), and one 3rd order parameter (ABC). Setting 2nd or higher order parameters equal to zero is equivalent to assuming independence between pairwise or higher order groups. For example, if $\beta_k^{\psi_4}$ is set to zero in Equation 7, the model assumes species A and B are independently distributed.

As with the first season, each probability within Equation 2 can may be made a function of covariates and the order of each parameter can be determined. To do so, we generate logit-linear predictors for each parameter:

$$\beta_k^{\gamma^A} = x_k^\top b_A \quad \beta_k^{\gamma^{A|B}} = x_k^\top g_{A|B}$$
$$\beta_k^{\gamma^B} = x_k^\top b_B \quad \beta_k^{\gamma^{B|A}} = x_k^\top g_{B|A}$$
$$\beta_k^{\varepsilon^A} = x_k^\top d_A \quad \beta_k^{\varepsilon^{A|B}} = x_k^\top h_{A|B} \tag{8}$$
$$\beta_k^{\varepsilon^B} = x_k^\top d_B \quad \beta_k^{\varepsilon^{B|A}} = x_k^\top h_{B|A}$$

where x_k denotes a vector of habitat covariates while b, d, g and h are coefficients that describe each species colonization–extinction rates in the presence (g and h) or absence (b and d) of one another. These logit-linear predictors are then combined and exponentiated

to represent the numerator of the multinomial-logit link for each transition probability:

The definitions for covariates (v) and parameter vectors (l and m) are the same as in Equation 8 except they (a) describe the logit

$$
\boldsymbol{\omega}_k = \underset{\text{To state}}{}
\begin{array}{c}
 \\
1_{[\mathrm{U}]} \\
2_{[\mathrm{A}]} \\
3_{[\mathrm{B}]} \\
4_{[\mathrm{AB}]}
\end{array}
\overset{\text{From state}}{
\begin{bmatrix}
1 & \exp(\beta_k^{\varepsilon^A}) & \exp(\beta_k^{\varepsilon^B}) & \exp\left(\begin{array}{c}\beta_k^{\varepsilon^A} + \beta_k^{\varepsilon^{A|B}} + \\ \beta_k^{\varepsilon^B} + \beta_k^{\varepsilon^{B|A}}\end{array}\right) \\[3mm]
\exp(\beta_k^{\gamma^A}) & 1 & \exp\left(\begin{array}{c}\beta_k^{\gamma^A} + \beta_k^{\gamma^{A|B}} + \\ \beta_k^{\varepsilon^B}\end{array}\right) & \exp(\beta_k^{\varepsilon^B} + \beta_k^{\varepsilon^{B|A}}) \\[3mm]
\exp(\beta_k^{\gamma^B}) & \exp\left(\begin{array}{c}\beta_k^{\varepsilon^A} + \beta_k^{\gamma^B} + \\ \beta_k^{\gamma^{B|A}}\end{array}\right) & 1 & \exp(\beta_k^{\varepsilon^A} + \beta_k^{\varepsilon^{A|B}}) \\[3mm]
\exp(\beta_k^{\gamma^A} + \beta_k^{\gamma^B}) & \exp(\beta_k^{\gamma^B} + \beta_k^{\gamma^{B|A}}) & \exp(\beta_k^{\gamma^A} + \beta_k^{\gamma^{A|B}}) & 1
\end{bmatrix}}
\tag{9}
$$

Note that transitions with 2nd order parameters include 1st order parameters. For example, the log odds species A colonizes a site that had species B in the last time step is $\beta_k^{\gamma^A} + \beta_k^{\gamma^{A|B}}$. Using the coefficients in Equation 8, this would be $\boldsymbol{x}_k^\top \boldsymbol{b}_A + \boldsymbol{x}_k^\top \boldsymbol{g}_{A|B}$. Given this parameterization, $\boldsymbol{g}_{A|B}$ denotes the logit-scale difference in species A colonization rate given the presence of species B. If $\boldsymbol{g}_{A|B} = 0$, then $\gamma_k^{A|B} = \gamma_k^A$ and the presence of species B does not influence species A colonization rate. Dividing each element of $\boldsymbol{\omega}_k$ by its respective column sum completes the multinomial-logit link function, converts each transition to a probability between 0 and 1, ensures that the columns sum to 1, and results in an $R \times R$ community transition matrix for each site, or $\boldsymbol{\phi}_k$. Note that the diagonal of the matrix in Equation 9 (i.e., the probability that a site remains in the same state) is set as reference categories. For more information on how Equation 9 relates to the community transition matrix in Equation 2 see supplemental material.

A near identical process can be used to include covariates for site and/or survey-specific detection probabilities:

probability of detecting each species given the true community state and (b) include covariates that vary by site and sampling occasion. Just as Equations 10 and 11 illustrate how survey specific covariates can be included, temporally varying covariates may be added to ϕ or λ in the same way. Doing so would result in an $R \times R$ transition matrix for each site and season of the latent state ($\boldsymbol{\phi}_{kt}$) and an $R \times R$ detection matrix for each site, season, and survey (λ_{ktj}).

2.4 | Deriving occupancy probabilities from the transition probability matrix

Each element in ϕ represents the probability a community transitions from one state to another and ϕ does not contain absorbing states (i.e., states that, once entered, cannot be left). Therefore, ϕ is an irreducible stochastic matrix and it is possible to derive a stationary probability vector δ of length R that describes the expected

$$
\boldsymbol{\zeta}_{kj} = \underset{\text{Observed state}}{}
\begin{array}{c}
1_{[\mathrm{U}]} \\
2_{[\mathrm{A}]} \\
3_{[\mathrm{B}]} \\
4_{[\mathrm{AB}]}
\end{array}
\overset{\text{True state}}{
\begin{bmatrix}
1 & 1 & 1 & 1 \\[2mm]
0 & \exp(\beta_{kj}^{\rho^A}) & 0 & \exp(\beta_{kj}^{\rho^A} + \beta_{kj}^{\rho^{A|B}}) \\[2mm]
0 & 0 & \exp(\beta_{kj}^{\rho^B}) & \exp(\beta_{kj}^{\rho^B} + \beta_{kj}^{\rho^{B|A}}) \\[2mm]
0 & 0 & 0 & \exp(\beta_{kj}^{\rho^A} + \beta_{kj}^{\rho^{A|B}} + \beta_{kj}^{\rho^B} + \beta_{kj}^{\rho^{B|A}})
\end{bmatrix}}
\tag{10}
$$

Unlike Equation 9, we have set the observed state U as the reference category as it is the only observable state that can occur across all true community states. Equation 10 results in a $R \times R$ matrix for each site and survey, or λ_{kj}, which therefore requires the following linear predictors:

$$
\begin{aligned}
\beta_{kj}^{\rho^A} &= \boldsymbol{v}_{kj}^\top \boldsymbol{l}_A \\
\beta_{kj}^{\rho^B} &= \boldsymbol{v}_{kj}^\top \boldsymbol{l}_B \\
\beta_{kj}^{\rho^{A|B}} &= \boldsymbol{v}_{kj}^\top \boldsymbol{m}_{A|B} \\
\beta_{kj}^{\rho^{B|A}} &= \boldsymbol{v}_{kj}^\top \boldsymbol{m}_{B|A}
\end{aligned}
\tag{11}
$$

occupancy probability of each community state by solving the equation $\delta = \delta\phi$ where $\Sigma\delta = 1$. Ecologically, δ is the expected probability of observing each community state given the estimated colonization-extinction rates. As with Rota, Ferreira, et al. (2016), marginal occupancy probabilities for each species can be derived from δ by summing together each state a species is part of. For example, with two species, $P(z = A) = \delta_A + \delta_{AB}$.

The stationary probability vector δ can also be made a function of covariates (e.g., δ_k) so long as they are incorporated into the

transition matrix through Equation 9. Furthermore, under a Bayesian analysis, parameter uncertainty may be propagated through δ or δ_k by calculating it during each step in an MCMC chain. Thus, not only does our model estimate local colonization-extinction rates between species, it can also be used to estimate the marginal occupancy probabilities of each species within a community and the occupancy probability of each community state across space and time.

3 | EXAMPLE: CO-OCCURRENCE OF COYOTE, OPOSSUM, AND RACCOON IN CHICAGO, IL

Here, we apply our DCOM to coyote, opossum, and raccoon detection/nondetection data from a large-scale long-term biodiversity monitoring survey throughout the greater Chicago metropolitan area. To obtain these data, Bushnell motion-triggered camera traps were placed in urban green spaces during the spring, summer, fall, and winter for 28 days per deployment at sites along three separate 50-km transects. These transects start in downtown Chicago and radiate outwards to the northwest, west, and southwest along a gradient of urbanization. At each site, one camera was attached to a tree roughly 1.5 m from the ground. Sites were located at least 1 km from one another. For additional details on this sampling protocol see Magle, Lehrer, and Fidino (2016).

Although sampling is still ongoing, we focus our analysis on a subset of seasons whose photos were identified to the species-level by experts. This includes 13 seasons from spring 2011 to spring 2013. Sites were included if they had ≥2 seasons worth of data (n = 103). Each 28 days deployment served as a primary sampling period and detection histories were generated at each site by treating each week of sampling as a repeat survey (median days active per deployment = 27). This resulted in a total of four repeat surveys within each primary sampling period per site and season. For three species (A, B, C) there are eight observable community states ($1_{[U]}$, $2_{[A]}$, $3_{[B]}$, $4_{[C]}$, $5_{[AB]}$, $6_{[AC]}$, $7_{[BC]}$, and $8_{[ABC]}$). A community state was considered detected at a site if the species attributed to that group were identified in photographs that occurred within a given sampling week (U = no species, A = coyote, B = opossum, C = raccoon). For an example of a three species community transition matrix, see supplemental material.

These three species were selected because the presence of coyote within urban remnant canyon habitat has been shown to decrease the relative abundance and occupancy of opossum and raccoon (Crooks & Soulé, 1999). Thus, we sought to determine if this trend generalized across urban environments and whether or not it varied across levels of urbanization. We fit six candidate models that represent varying hypotheses as to how coyote, opossum, and raccoon may influence each others' colonization-extinction rates. We used three factors to quantify the level of urbanization around each site: mean tree cover, mean impervious cover, and mean housing density. These variables were calculated within a 1,000 m buffer centred on a camera trap with v. 2.14 of QGIS (QGIS Development Team, 2016). We then applied principal component analysis to these data and used the first principal component which explained 78.76% of the variation. This metric (hereafter URB) was then multiplied by −1 so that positive values indicated sites with higher housing densities (more urban) while negative values were sites with higher levels of tree cover (more available natural habitat).

Because a small number of sites were sampled in the first season (n = 34) we kept parameterizations for initial occupancy simple and assumed all species were independently distributed during the first time step (i.e., $\beta_k^{\psi^{ABC}} = \beta_k^{\psi^{AB}} = \beta_k^{\psi^{AC}} = \beta_k^{\psi^{BC}} = 0$). For the detection model, we also did not include interactive effects between all species. Instead, because coyote are the largest predator in the region we only included the influence that they may have on opossum and raccoon detection.

Our first and largest model, M_1, reflects the hypothesis that coyote, opossum, and raccoon influence one another's colonization-extinction rates and that these rates vary with urbanization (Table 2). To test this, the nine 1st order occupancy, colonization, and extinction parameters (e.g., ψ^A, γ^B, ε^C, etc.) and 12 2nd order colonization and extinctions parameters (e.g., $\gamma^{A|B}$, $\varepsilon^{C|A}$, etc.) were modelled individually as functions of URB. Our next models, M_2 and M_3, also assume that species influence each other's colonization–extinction rates. However, M_2 models 2nd order colonization and extinction parameters as a function of URB while species-specific 1st order parameters are kept at their intercept. M_3 does the opposite of M_2 and assumes 2nd order colonization and extinction parameters do not vary with URB but 1st order parameters do (Table 2).

TABLE 2 The six models fit to the coyote, opossum, and raccoon detection/nondetection data[a]

Model components	Models					
	M_1	M_2	M_3	M_4	M_5	M_6
Includes species interaction parameters	Yes	Yes	Yes	Yes	—	—
URB on 1st order parameters (e.g., ε^C)	Yes	—	Yes	—	Yes	—
URB on 2nd order parameters (e.g., $\gamma^{A|B}$)	Yes	Yes	—	—	—	—
Coyote influences opossum and raccoon detection	Yes	Yes	Yes	Yes	—	—

Note. [a]URB is the first principal component of the mean canopy cover, mean impervious surface, and mean housing density within 1,000 m of a camera trap.

Our fourth model, M_4, is an intercept-only DCOM which assumes that species influence one another's colonization–extinction rates but no rates vary with URB. M_5 is further simplified and assumes species do not influence each others' colonization-extinction rates but species occupancy, colonization, and extinction rates vary with URB. This model is similar to a standard dynamic occupancy model that assumes independence between species (Dorazio et al., 2010). M_6 assumes that all species-specific occupancy, colonization, and extinction rates are kept at their intercept (i.e., a null model). We assumed Logistic(0,1) priors for all logit-scale parameters. Note that for all of these models, we assumed 3rd order interactions did not occur (i.e., 3rd order coefficients were set to zero). This meant that probabilities conditioning on two species being present in a habitat patch (e.g., $\gamma^{A|BC}$) were a function of the species-specific 1st order and pairwise 2nd order parameters.

To compare the relative fit of each model, we calculated the conditional predictive ordinate of each data point (CPO; Hooten & Hobbs, 2015). Following this, the summary statistic $-\Sigma_{kt} \log(CPO_{kt})$ for site k and time t was used to assess the overall model performance, where lower values indicate better model fit (Hooten & Hobbs, 2015). Models were fit in JAGS v. 4.2.0 (Plummer, 2003) with the runjags package (Denwood, 2016) in R v. 3.4.3 (R Core Team, 2017). After a 400-step adaptation and 100,000 step burn-in posteriors were sampled 500,000 times. MCMC chains were thinned by 10 so that a total of 50,000 samples were obtained. To check model convergence, we inspected trace plots of MCMC chains to ensure proper mixing and confirmed that Gelman–Rubin diagnostics for each parameter were <1.10 (Gelman et al., 2014). Stationary occupancy probability vectors were derived from the transition probability matrix of the best fit model.

4 | RESULTS

A total of 3958 trap weeks of data were collected between spring 2011 and spring 2013 from a maximum possible 5,356 trap weeks

TABLE 3 Model selection results of the six candidate models fit data[a]

Model	No. latent parameters	No. detection parameters	CPO	ΔCPO
M_3	30	5	5,246.34	0.00
M_4	21	5	5,289.46	43.12
M_5	18	3	5,345.42	99.08
M_2	33	5	5,476.11	229.77
M_6	9	3	5,479.06	232.72
M_1	42	5	5,792.01	545.67

Notes. [a]CPO, conditional predictive ordinate. Models were compared with the summary statistic $-\Sigma_{kt} \log(CPO_{kt})$ for site k and time t. Lower values indicate better model fit. ΔCPO is the number of CPO units each model's CPO is from the best-fit model.

(103 sites × 13 seasons × 4 weeks per deployment). Coyote, opossum, and raccoon were collectively detected on 1,063, 872, and 1,425 trap weeks. Model selection results indicated that pairwise interactions between species influenced their colonization-extinction rates (Table 3). Two models that estimated interactions between species (i.e. M_3 and M_4; Table 2) outperformed models that assumed species were independently distributed by >40 CPO units (Table 3). The top model, M_3, assumed that URB did not influence pairwise interactions but did influence species specific colonization–extinction rates.

For species-specific parameters, the best fit model indicated that (a) all species colonization rates decreased with URB and (b) raccoon and opossum extinction rates increased with URB (Table 4). There was weak evidence that coyote extinction rates decreased at higher levels of URB.

Pairwise interactions influenced colonization rates between four species pairs as evidenced by 95% credible intervals of pairwise intercept terms not overlapping zero (Figure 1). On average, opossums were more likely to colonize sites with raccoon than sites without raccoon ($g_{opo|rac} = 0.81$; 95% CI = 0.02–1.59).

TABLE 4 First order colonization and extinction estimates from the best fit model. Estimates are on the logit scale

Process	Species	Parameter	Estimate	Quantile 2.5%	97.5%
Colonization	Coyote	Intercept	−1.75	−2.44	−1.16
		URB slope	−0.48	−0.76	−0.21
	Opossum	Intercept	−2.00	−2.73	−1.32
		URB slope	−0.42	−0.71	−0.19
	Raccoon	Intercept	−1.16	−1.73	−0.61
		URB slope	−0.67	−0.95	−0.42
Extinction	Coyote	Intercept	−0.18	−0.77	0.38
		URB slope	−0.21	−0.45	0.02
	Opossum	Intercept	−0.01	−0.72	0.69
		URB slope	0.22	0.00	0.46
	Raccoon	Intercept	−0.28	−0.90	0.32
		URB slope	0.31	0.06	0.58
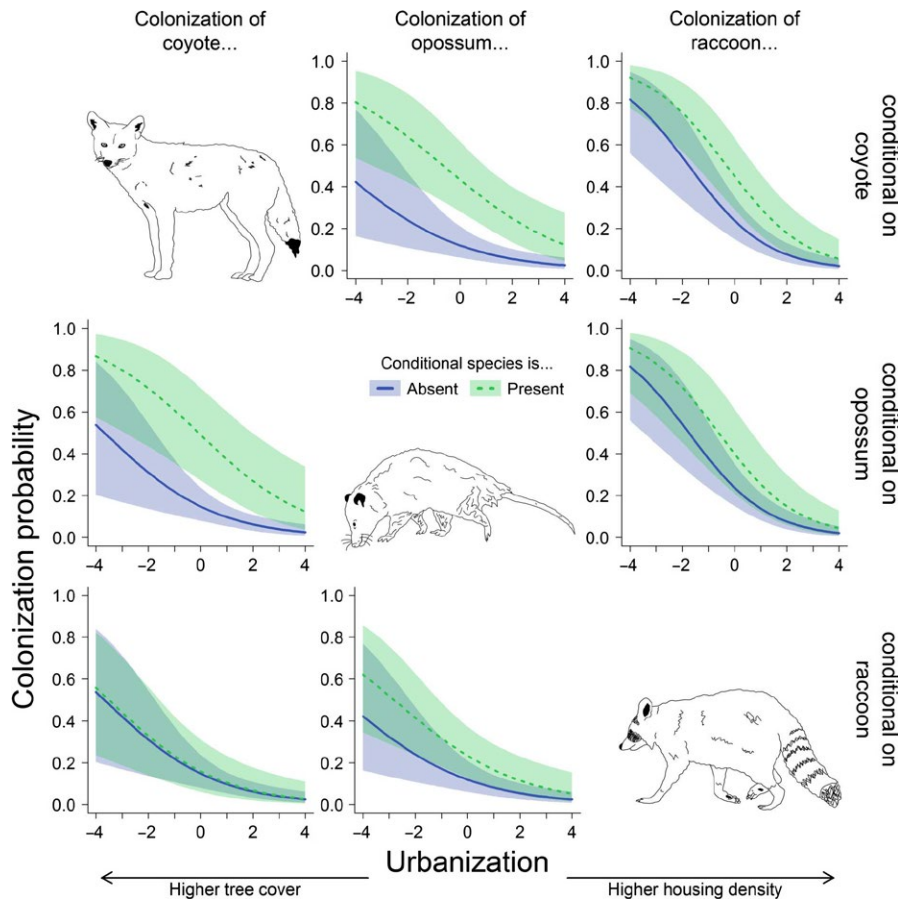
FIGURE 1 Colonization probabilities of coyote, opossum, and raccoon conditional on the presence or absence of each other estimated from thirteen seasons of camera trapping data in Chicago, Illinois, USA. Colonization probabilities in each column are conditional on the presence or absence of the species in a given row. Lines in each subplot are posterior means while the shaded ribbons represent 95% credible intervals

Opossum were also more likely to colonize sites with coyote ($g_{opo|coy}$ = 1.73; 95% CI = 0.96–2.60; Figure 1). Similar to the opossum, raccoon were more likely to colonize sites with coyote present ($g_{rac|coy}$ = 0.96; 95% CI = 0.18–1.78). Finally, coyotes were more likely to colonize sites with opossum ($g_{coy|opo}$ = 1.70; 95% CI = 0.75–2.79; Figure 1).

Three pairwise interactions influenced species extinction rates. Contrary to previous research (Crooks & Soulé, 1999), opossum were less likely to go extinct in habitat patches with coyote present ($h_{opo|coy}$ = −0.72; 95% CI = −1.29 to −0.14). Raccoon were less likely to go extinct in habitat patches that contained coyote ($h_{rac|opo}$ = −1.54; 95% CI = −2.21 to −0.91). Finally, coyote were less likely to go extinct in habitat patches with raccoon present ($h_{coy|rac}$ = −1.54; 95% CI = −2.27 to −0.84) (Figure 2).

As the transition probability matrix of the best fit model varied as a function of urbanization (i.e., ϕ_k), we derived the stationary occupancy vector δ_k for each community state across this covariate over all samples of the posterior distribution (Figure 3). At low levels of urbanization, sites were far more likely to have all three species present or opossum and raccoon present than any other community state (Figure 3). As urbanization increased, the likelihood of observing all three species at a site decreased and sites were most likely to not have these three species (Figure 3a).

With regard to detectability, coyote had the highest weekly detection probability at 39.89% (95% CI = 37.08–42.81) followed by raccoon (33.60%; 95% CI = 26.74–40.38) and opossum (19.31%; 95% CI = 13.61–26.15). Opossum and raccoon were much more detectable when coyote were present at a site. Opossum detection probability increased to 62% (95% CI = 59.31–66.37) with coyote present while the probability of detecting raccoon with coyote present was 61.79% (95% CI = 58.49–64.97).

5 | DISCUSSION

We have presented here a DCOM that can estimate species interactions between any number of species, can allow interactions to vary as a function of covariates, and corrects for imperfect detection. We also illustrate how the estimated colonization–extinction rates of a community can be converted into occupancy probabilities so that co-occurrence patterns between species may be explored across environmental gradients. Given these capabilities, our DCOM may be an especially useful tool to empirically test different metacommunity paradigms (Holyoak et al., 2005). While Dorazio et al. (2010) have already developed dynamic occupancy models to test some aspects of metacommunity theory (e.g., species sorting along a gradient), our proposed model adds the ability to test for additional relevant components such as priority effects and other types of species interactions that could not have been estimated previously. Thus, our model adds even more versatility to the occupancy modelling framework
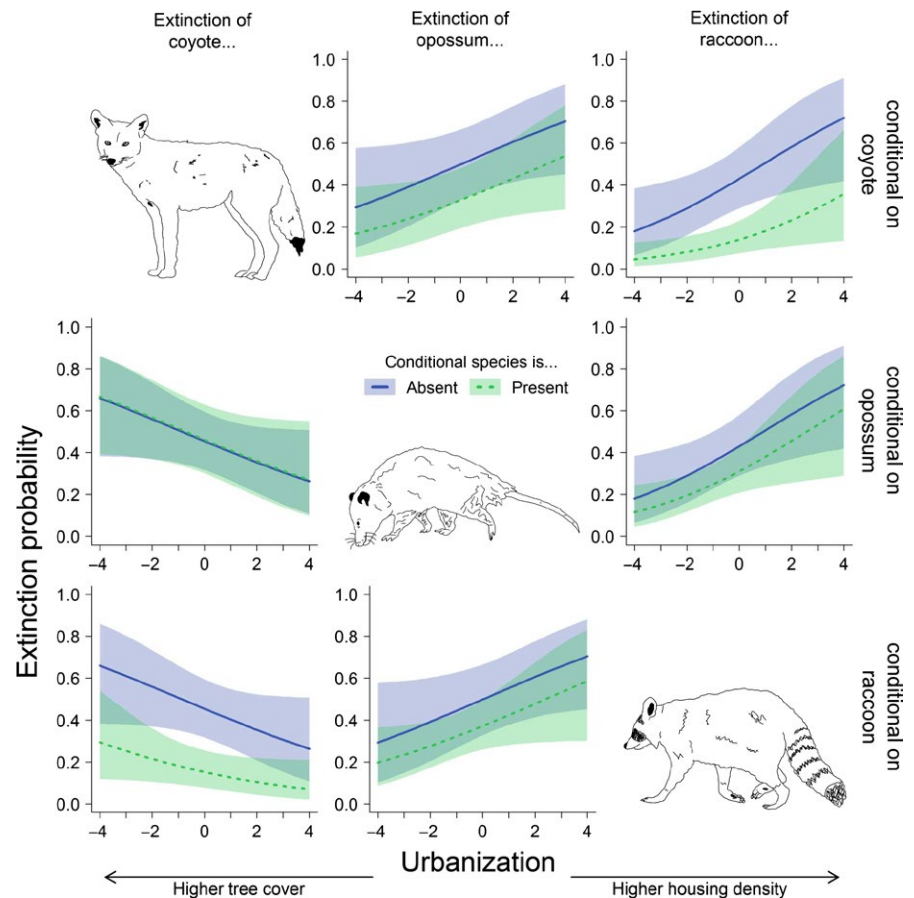
FIGURE 2 Extinction probabilities of coyote, opossum, and raccoon conditional on the presence or absence of all other species estimated from thirteen seasons of camera trapping data in Chicago, Illinois, USA. Extinction probabilities in each column are conditional on the presence or absence of the species in a given row. Lines in each sub-plot are posterior means while the shaded ribbons represent 95% credible intervals

while simultaneously allowing for more explicit tests of predictions from ecological theory.

Similar to Rota, Ferreira, et al. (2016), our DCOM can require a large number of parameters to estimate species associations. Fortunately, there are many methods available for dimension reduction that have already been implemented in BUGS or JAGS. For example, a variety of variable selection techniques could be incorporated into the model by modifying the priors of a Bayesian analysis (Hooten & Hobbs, 2015). In addition, like other joint species distribution models (JSDMs), it is critical to remember that the results do not necessarily prove the existence of causal species interactions. Other environmental variables not included within a model may be the reason why the presence of one species at a site influences another if those two covary along an unmodelled environmental gradient. In our own example, it is more likely that habitat patches suitable for opossum are also suitable for raccoon because they have overlapping diets (Feldhamer, Thompson, & Chapman, 2003). As such, pairwise interactions should be considered in context of both the species modelled and the environment they are sampled within. Regardless, our example illustrates the utility of estimating pairwise interactions between species as models that included interactive effects greatly outperformed standard models that did not (Table 3). Further, derived occupancy probabilities from our example closely track the observed data (Figure 3) and help tease apart the processes that drive future occupancy probabilities in a flexible model based framework.

In addition to being critical of the estimated biotic interactions from JSDMs, ecologists have an ever growing number of JSDMs to choose from when modelling species codistributions (for a review see Dormann et al., 2018). For example, there are many approaches that can estimate species spatio-temporal patterns for large, species rich, datasets (e.g., Ovaskainen et al., 2017; Thorson et al., 2016; Warton et al., 2015). However, such techniques do not currently correct for imperfect detection and do not explicitly model temporal changes in species co-distributions by estimating colonization and extinction rates. Since estimating local colonization and extinction rates is typically a better predictor of future occupancy patterns (Yackulic, Nichols, Reid, & Der, 2015), our approach may be more useful to predict how species distributions may change into the future under new habitat or climatic conditions. Likewise, as species interactions could vary along environmental gradients (LaManna et al., 2015), a major strength of our DCOM is the ability to model interactions as a function of covariates. JDSMs that use a covariance matrix approach to estimate co-occurrence make it more difficult to include covariates on species interactions, though such techniques make it easier to estimate associations between a greater number of species. Therefore, the choice of model to be used will likely depend on the goals of a study, the data that are available, and the number of species associations to be estimated.

Our DCOM could be used to answer a variety of applied ecological questions. For example, our model could determine how invasive
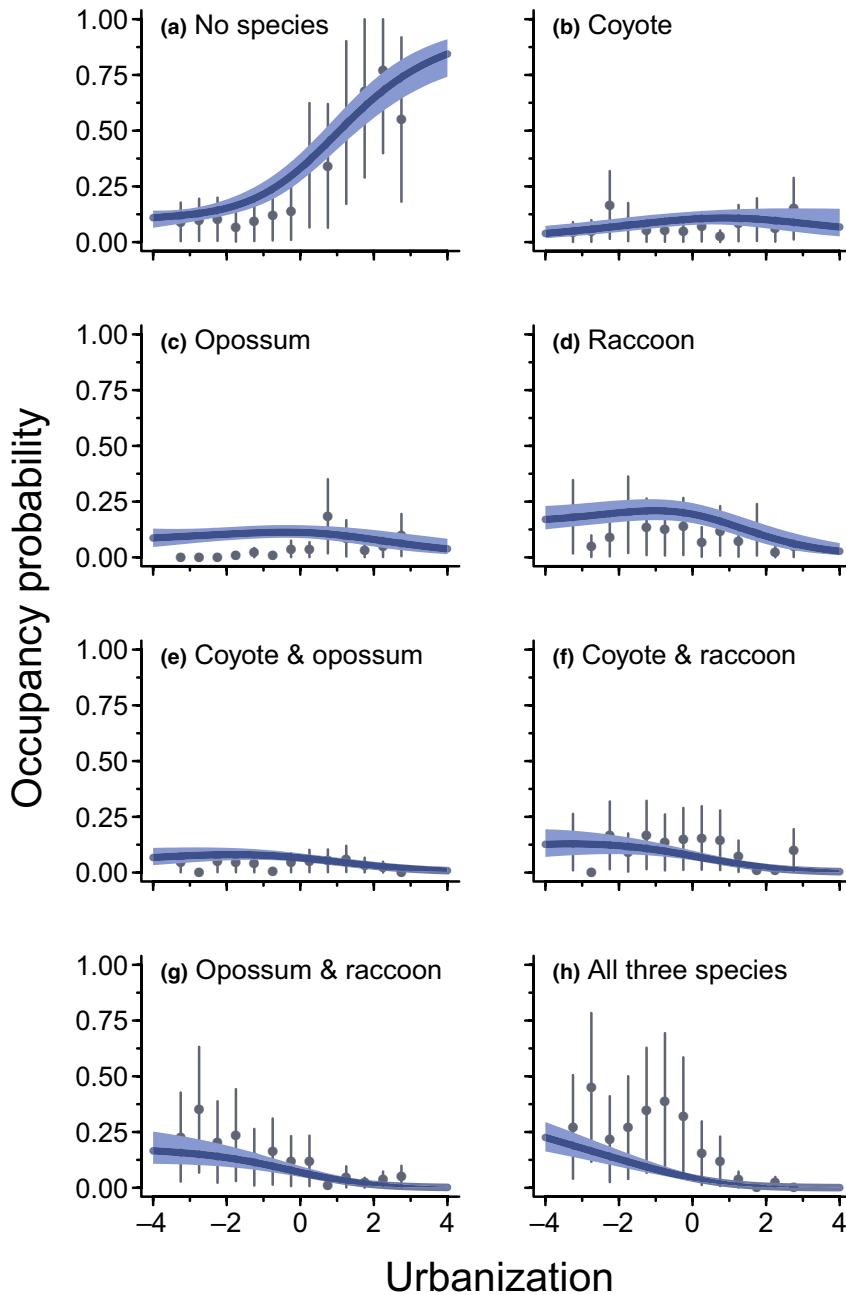
FIGURE 3 The stationary occupancy probabilities of all eight possible community states (a–h) as a function of urbanization at a site. Solid horizontal lines indicate posterior means. Shaded ribbons represent 95% credible intervals. Naïve occupancy estimates were calculated as the proportion of sites in which each community state was detected from the raw data and was discretized along the urbanization gradient in partitions of 0.5 from −4 to 4. For each partition, points represent the mean naïve occupancy of a community state relative to all other states across all seasons. Vertical lines represent the standard error associated to each naïve occupancy estimate.

species influence wildlife communities by estimating species-specific colonization-extinction rates in the presence and absence of an invasive competitor (Sanders, Gotelli, Heller, & Gordon, 2003). Because standard occupancy models assume that species are independently distributed they cannot elucidate interactive effects when they occur, whereas our DCOM can estimate interactions between any number of species. This DCOM could also determine the effectiveness of monitoring surrogate species in conservation biology (Caro & O'doherty, 1999). While understanding the habitat associations of surrogate species may suggest habitat patches to conserve, estimating the probability of observing one or more surrogate species in a habitat patch given the presence of others could provide a more robust metric for use in conservation applications. Finally, we anticipate that this approach can be used to estimate where species are more or

less likely to co-occur, facilitating follow-up smaller scale experiments to test potential mechanisms of coexistence (Chesson, 2000).

Numerous extensions to our modelling framework are possible. In our current formulation, we only assume a species local colonization or extinction probabilities depend on the modelled environmental gradients and not the spatial location of or distances between habitat patches. To make our model more spatially explicit, concepts from metapopulation ecology could be incorporated so that the probability a site is colonized is partly a function of its euclidean distance to other habitat patches that contain the target species (Hanski, 2001). Likewise, a least cost path approach could be used in favour of euclidean distance and methods exist to estimate resistance coefficients from patch occupancy data (Howell, Muths, Hossack, Sigafus, & Chandler, 2018). Our model also assumes

independence between sampling points, which may not hold true for species with relatively large home ranges such as the coyote in our example. Although Rota, Wikle, et al. (2016) relax this assumption with a two species occupancy model, the question remains whether such an approach can be generalized to >2 species or through time.

As data become more accessible and cheaper to collect over larger areas and longer time spans, ecologists are able to test theories over scales that were previously unattainable. This is facilitated through advances in computational speed and ecological statistics. One major hurdle that limits Big Data-based ecological research is the ability to understand how to correctly quantify uncertainty related to the patterns and processes underlying the data. The model we present is a versatile approach that can fit a multitude of sampling designs and hypotheses. It can also be used to provide empirical support to ecological theory that explores how potentially interacting species co-occur. Thus, our DCOM is an important improvement to modelling species distributions and interactions through space and time.

ACKNOWLEDGEMENTS

AUTHORS' CONTRIBUTIONS

M.F. and J.L.S. developed the idea, methodology and analysed the data for the manuscript. M.F. and S.B.M. collected the data. M.F. wrote the initial draft of the manuscript and all authors contributed to later drafts.

DATA ACCESSIBILITY

Data, R, and JAGS code are archived in Dryad Digital Repository https://doi.org/10.5061/dryad.k5mp137 (Fidino, Simonis, & Magle, 2018). Data and code are also available on github at www.github.com/mfidino/dcom.

ORCID

Mason Fidino http://orcid.org/0000-0002-8583-0307

REFERENCES

Atuo, F. A., & O'Connell, T. J. (2017). Spatial heterogeneity and scale-dependent habitat selection for two sympatric raptors in mixed-grass prairie. *Ecology and Evolution*, 7, 6559–6569. https://doi.org/10.1002/ece3.3182.

Bailey, L. L., MacKenzie, D. I., & Nichols, J. D. (2014). Advances and applications of occupancy models. *Methods in Ecology and Evolution*, 5, 1269–1279. https://doi.org/10.1111/2041-210X.12100

Broms, K. M., Hooten, M. B., Johnson, D. S., Altwegg, R., & Conquest, L. L. (2016). Dynamic occupancy models for explicit colonization processes. *Ecology*, 97, 194–204. https://doi.org/10.1890/15-0416.1

Caro, T. M., & O'doherty, G. (1999). On the use of surrogate species in conservation biology. *Conservation Biology*, 13, 805–814. https://doi.org/10.1046/j.1523-1739.1999.98338.x

Chambert, T., Grant, E. H. C., Miller, D. A., Nichols, J. D., Mulder, K. P., & Brand, A. B. (2018). Two-species occupancy modelling accounting for species misidentification and non-detection. *Methods in Ecology and Evolution*, 9, 1468–1477. https://doi.org/10.1111/2041-210x.12985

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, 31, 343–366. https://doi.org/10.1146/annurev.ecolsys.31.1.343

Crooks, K. R., & Soulé, M. E. (1999). Mesopredator release and avifaunal extinctions in a fragmented system. *Nature*, 400, 563. https://doi.org/10.1038/23028

Dai, B., Ding, S., & Wahba, G. (2013). Multivariate bernoulli distribution. *Bernoulli*, 19, 1465–1483. https://doi.org/10.3150/12-bejsp10

Denwood, M. J. (2016). `runjags`: An `r` package providing interface utilities, model templates, parallel computing methods and additional distributions for mcmc models in `JAGS`. *Journal of Statistical Software*, 71, 1–25. https://doi.org/10.18637/jss.v071.i09

Dorazio, R. M., Kery, M., Royle, J. A., & Plattner, M. (2010). Models for inference in dynamic meta-community systems. *Ecology*, 91, 2466–2475. https://doi.org/10.1890/09-1033.1

Dormann, C. F., Bobrowski, M., Dehling, D. M., Harris, D. J., Hartig, F., Lischke, H., … Kraan, C. (2018). Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions. *Global Ecology and Biogeography*, 27, 1004–1016. https://doi.org/10.1111/geb.12759

Elton, C. S. (1958). *The ecology of invasions by animals and plants*. London, UK: Methuen.

Feldhamer, G. A., Thompson, B. C., & Chapman, J. A. (2003). *Wild mammals of North America: Biology, management, and conservation*. Batlimore, MD: JHU Press.

Fidino, M., & Magle, S. B. (2017). Using fourier series to estimate periodic patterns in dynamic occupancy models. *Ecosphere*, 8, e01944. https://doi.org/10.1002/ecs2.1944

Fidino, M., Simonis, J. L., Magle, S. B. (2018). Data from: A multi-state dynamic occupancy model to estimate local colonization-extinction rates and patterns of co-occurrence between two or more interacting species. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.k5mp137

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (Vol. 2). Boca Raton, FL: CRC Press.

Hanski, I. (2001). Spatially realistic theory of metapopulation ecology. *Naturwissenschaften*, 88, 372–381. https://doi.org/10.1007/s001140100246

Holyoak, M., Leibold, M. A., & Holt, R. D. (2005). *Metacommunities: Spatial dynamics and ecological communities*. Chicago, IL: University of Chicago Press.

Hooten, M. B., & Hobbs, N. T. (2015). A guide to bayesian model selection for ecologists. *Ecological Monographs*, 85, 3–28. https://doi.org/10.1890/14-0661.1

Howell, P. E., Muths, E., Hossack, B. R., Sigafus, B. H., & Chandler, R. B. (2018). Increasing connectivity between metapopulation ecology and landscape ecology. *Ecology*, 99, 1119–1128. https://doi.org/10.1002/ecy.2189

Jones, J. E., & Kroll, A. J. (2016). A joint model of species interaction dynamics for multistate processes incorporating imperfect detection. *Ecosphere*, 7, e01477. https://doi.org/10.1002/ecs2.1477

LaManna, J. A., Hemenway, A. B., Boccadori, V., & Martin, T. E. (2015). Bird species turnover is related to changing predation risk along a vegetation gradient. *Ecology, 96*, 1670–1680. https://doi.org/10.1890/14-1333.1

Levins, R. (1969). Some demographic and genetic consequences of environmental heterogeneity for biological control. *American Entomologist, 15*, 237–240. https://doi.org/10.1093/besa/15.3.237

MacArthur, R. H., & Wilson, E. O. (1967). *The theory of island biogeography*. Princeton, NJ: Princeton University Press.

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. https://doi.org/10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2

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

Magle, S., Lehrer, E., & Fidino, M. (2016). Urban mesopredator distribution: Examining the relative effects of landscape and socioeconomic factors. *Animal Conservation, 19*, 163–175. https://doi.org/10.1111/acv.12231

Miller, D. A., Brehme, C. S., Hines, J. E., Nichols, J. D., & Fisher, R. N. (2012). Joint estimation of habitat dynamics and species interactions: Disturbance reduces co-occurrence of non-native predators with an endangered toad. *Journal of Animal Ecology, 81*, 1288–1297. https://doi.org/10.1111/j.1365-2656.2012.02001.x

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., … Abrego, N. (2017). How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecology Letters, 20*, 561–576. https://doi.org/10.1111/ele.12757

Plummer, M. (2003). `JAGS`: A program for analysis of bayesian graphical models using gibbs sampling. In K. Hornik, F. Leisch, & A. Zeileis (Eds.), *Proceedings of the Third International Workshop on Distributed Statistical Computing* (pp. 125–133). Vienna, Austria: R Foundation for Statistical Computing.

QGIS Development Team. (2016). QGIS geographic information system. Open source geospatial foundation project.

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Robert, C. (2014). *Machine learning, a probabilistic perspective*. Cambridge, MA: The MIT Press.

Rota, C. T., Ferreira, M. A., Kays, R. W., Forrester, T. D., Kalies, E. L., McShea, W. J., … Millspaugh, J. J. (2016). A multispecies occupancy model for two or more interacting species. *Methods in Ecology and Evolution, 7*, 1164–1173. https://doi.org/10.1111/2041-210X.12587

Rota, C. T., Wikle, C. K., Kays, R. W., Forrester, T. D., McShea, W. J., Parsons, A. W., & Millspaugh, J. J. (2016). A two-species occupancy model accommodating simultaneous spatial and interspecific dependence. *Ecology, 97*, 48–53. https://doi.org/10.1890/15-1193.1

Sanders, N. J., Gotelli, N. J., Heller, N. E., & Gordon, D. M. (2003). Community disassembly by an invasive species. *Proceedings of the National Academy of Sciences, 100*, 2474–2477. https://doi.org/10.1073/pnas.0437913100

Thorson, J. T., Ianelli, J. N., Larsen, E. A., Ries, L., Scheuerell, M. D., Szuwalski, C., & Zipkin, E. F. (2016). Joint dynamic species distribution models: A tool for community ordination and spatio-temporal monitoring. *Global Ecology and Biogeography, 25*, 1144–1158. https://doi.org/10.1111/geb.12464

Warton, D. I., Blanchet, F. G., O'Hara, R. B., Ovaskainen, O., Taskinen, S., Walker, S. C., & Hui, F. K. (2015). So many variables: Joint modeling in community ecology. *Trends in Ecology and Evolution, 30*, 766–779. https://doi.org/10.1016/j.tree.2015.09.007

Yackulic, C. B., Reid, J., Nichols, J. D., Hines, J. E., Davis, R., & Forsman, E. (2014). The roles of competition and habitat in the dynamics of populations and species distributions. *Ecology, 95*, 265–279. https://doi.org/10.1890/13-0012.1

Yackulic, C. B., Nichols, J. D., Reid, J., & Der, R. (2015). To predict the niche, model colonization and extinction. *Ecology, 96*, 16–23. https://doi.org/10.1890/14-1361.1

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.