Appendix S1. Formulation of a Dynamic Occupancy Model

Here we more thoroughly describe the structure of a dynamic occupancy model for a single species at k in 1,2,…,K sites and t in 1,2,…T sampling seasons using matrix notation. All notation is described in Table S1 at the end of this appendix. For simplicity, we assume that we would like to quantify the effect of Q covariates collected at all K sites on the probability of initial occupancy (ψ), colonization (γ), and persistence (ϕ) and the effect of R covariates collected at all K sites on the conditional probability of detection (p). However, the model is general enough such that different covariates could be used for each process, including those that temporally vary.

The number of surveys a species is detected at a site across multiple seasons is:

$$y_{k,t}|z_{k,t}, p_{k,t} \sim Binomial(j_{k,t}, z_{k,t}p_{k,t}). \tag{S1}$$

The probability of occupancy at the first time step is

$$Z_{k,t=1}|\psi_k \sim Bernoulli(\psi_k) \tag{S2}$$

Which can be made a function of Q covariates through the logit link

$$logit(\psi_k) = \boldsymbol{x}_k\boldsymbol{b} \tag{S3}$$

Following the first time step we assume a first-order autoregressive relationship in that the future occupancy state in the next season (t+1) depends on the current occupancy status (time t) via a mixture of local colonization (γ) and persistence (ϕ) probabilities.

$$z_{k,t}|z_{k,t-1}, \phi_{k,t}, \gamma_{k,t} \sim \tag{S4}$$
$$Bernoulli\left(\phi_{k,t}z_{k,t-1} + \gamma_{k,t}(1 - z_{k,t-1})\right) \; for \; t = 2,3, \dots, T$$

where

$$logit(\phi_{k,t}) = \boldsymbol{x}_k\boldsymbol{d} \tag{S5}$$

and

$$logit(\gamma_{k,t}) = \boldsymbol{x}_k \boldsymbol{m} \tag{S6}$$

are the logit-scale probabilities for persistence (i.e., 1 – probability of extinction, Eq. S5) and colonization (Eq. S6), which are both functions of Q covariates. Finally, the conditional probability of detecting a species given its presence can also be made a function of R covariates such that

$$logit(p_{k,t}) = \boldsymbol{s}_k \boldsymbol{f}. \tag{S7}$$

The above model allows for the estimation of how spatial features of a landscape influence ψ, ϕ, γ, and p but could easily be generalized to include time-varying parameters and covariates.

Table S1. Notation used in a dynamic occupancy model for a single species.

K	The total number of sites.
T	The total number of seasons of data.
$y_{k,t}$	The number of surveys a species was detected at site k at time t.
$z_{k,t}$	The occupancy state of a species at site k at time t. If present than $Z_{k,t} = 1$.
$p_{k,t}$	The conditional probability of detecting a species given its presence at site k at time t.
$j_{k,t}$	The total number of surveys conducted at site k at time t.
ψ_k	The probability of initial occupancy at site k at time $t = 1$.
Q	The total number of covariates thought to influence Ψ, γ, and Φ.
$\boldsymbol{x}_k$	A row-vector of length $Q + 1$. The first value in this vector is 1, which is multiplied by the constant term in the regression, while the rest the elements in $\boldsymbol{x}_k$ are the Q covariate values collected at site k.
$\boldsymbol{b}$	A column vector of $Q + 1$ regression coefficients which includes the constant term and the effect of Q covariates on Ψ.
$\phi_{k,t}$	The probability of persistence at site k at time t.
$\boldsymbol{d}$	A column vector of $Q + 1$ regression coefficients that includes the constant term and the effect of Q covariates on Φ.
$\gamma_{k,t}$	The probability of colonization at site k at time t.
$\boldsymbol{m}$	A column vector of $Q + 1$ regression coefficients which includes the constant term and the effect of Q covariates on γ.
R	The number of covariates thought to influence p_k
$\boldsymbol{S}$	A $K \times (R + 1)$ matrix of covariates (the first column of $\boldsymbol{G}$ being a column of 1's for the constant term in the regression). If covariates temporally vary $\boldsymbol{S}$ becomes a three-dimensional array.
$\boldsymbol{s}_k$	A row-vector of length $R + 1$. The first value in this vector is 1, which is multiplied by the constant term in the regression, while the rest of the elements in $\boldsymbol{s}_k$ are the R covariate values collected at site k.
$\boldsymbol{f}$	A column vector of $R + 1$ regression coefficients which includes the constant term and the effect of R covariates on p.