

# ECOGRAPHY

## Research article

### Evaluating the influence of neighborhood connectivity and habitat effects in dynamic occupancy species distribution models

Oriol Solà<sup>1</sup>, Núria Aquilué<sup>2</sup>, Sara Fraixedas<sup>1</sup> and Lluís Brotons<sup>1,2,3</sup>

<sup>1</sup>Centre for Ecological Research and Forestry Applications (CREAF), Autonomous University of Barcelona (UAB), Cerdanyola del Vallès, Spain

<sup>2</sup>Forest Science and Technology Centre of Catalonia (CTFC), Solsona, Spain

<sup>3</sup>CSIC, Cerdanyola del Vallès, Spain

Correspondence: Oriol Solà ([o.sola@creaf.uab.cat](mailto:o.sola@creaf.uab.cat))

Ecography

2024: e06985

doi: [10.1111/ecog.06985](https://doi.org/10.1111/ecog.06985)

Subject Editor: Jason Pither  
Editor-in-Chief: Miguel Araújo  
Accepted 14 April 2024



Exploring new approaches and methodologies to characterize species distribution dynamics, instead of solely relying on static spatial patterns, should be a priority for species distribution modelling research. Dynamic occupancy models (here, 'dynocc models') are a promising tool to capture temporal patterns of distribution change but their spatial accuracy has been shown to be limited. In this study, we evaluated the effectiveness of incorporating neighborhood connectivity effects into the colonization and extinction functions of dynocc models. To accomplish this, we compared dynocc models accounting either for neighborhood connectivity only, for site-level habitat covariates only, or combining both neighborhood and habitat explanations in the models for species extinction and colonization. All models were evaluated for a total of 46 bird species typical of forests and shrublands using data at 1 km<sup>2</sup> scale from two Catalan breeding bird atlases (CBBA2: 1999–2002 and CBBA3: 2015–2018). Models' predictive performance varied across species between dynocc models incorporating habitat covariates alone and those considering neighborhood connectivity alone. Among species, 68% exhibited a predominant response to habitat effects, 24% showed similar responses for habitat and connectivity effects, and 9% were mostly associated with connectivity effects. Dynocc models combining connectivity and habitat covariates achieved the best predictive performance for most species, with bigger gains for species with similar results from habitat-only and connectivity-only models. However, relative performance gains compared to dynocc models using only habitat or connectivity variables were generally modest for most species. This study shows the benefits of considering more spatially explicit formulations in dynocc models, specifically incorporating neighborhood connectivity into the extinction and colonization processes. Our work also highlights the importance of evaluating different model formulations and assessing which aspects of the model are more important depending on the study species.

Keywords: colonization, distribution change, dynamic occupancy models, extinction, hierarchical models, neighborhood connectivity, presence–absence, range change



[www.ecography.org](http://www.ecography.org)

© 2024 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Page 1 of 12

## Introduction

In the context of global change, species are shifting their spatial distributions mostly in association with changes in climate and anthropogenic land use (Parmesan and Yohe 2003). Understanding the underlying factors of distribution change and predicting how and where these changes will occur is key to effectively plan and implement biodiversity conservation measures. Distribution dynamics result from complex interactions between multiple biotic and abiotic factors, which makes it especially challenging to identify processes associated with changes in species distributions (Hodgson et al. 2009, Urban et al. 2013, Ehrlén and Morris 2015). In this regard, modelling emerges as a crucial tool to provide robust and policy-relevant information by linking environmental changes with shifts in species distributions (IPBES 2016).

Among the suite of available modelling approaches, species distribution models (SDMs) are a powerful set of methods to predict species changes in response to different biotic and abiotic factors (Guisan and Thuiller 2005). A large variety of SDM approaches are currently available, ranging from static correlative methods to purely process-based models in a gradient of increasing complexity, degree of mechanism, and data requirements. Static correlative SDMs (Guisan and Zimmermann 2000, Elith and Leathwick 2009) are relatively simple to fit and only require occurrence or abundance data. However, such models can lead to misleading predictions of range dynamics since they do not explicitly account for the processes that actually cause species distribution changes, and instead assume that species are in equilibrium at any given moment (Elith et al. 2010, Sofaer et al. 2018, Piirainen et al. 2023). At the other end of the model gradient, the so-called 'process-based models' explicitly identify and model the mechanisms that drive range dynamics at the population or individual level. Consequently, these models offer a more comprehensive understanding of the drivers of change and have the potential to yield more accurate predictions (Zurell 2017, Briscoe et al. 2019). Nonetheless, such models tend to be very complex and normally require different types of data (e.g. on dispersal distances, species interactions, thermal reaction norms). This makes them difficult to implement because in most cases the data required to calibrate or parametrize such models are not available or the species-specific behavior is not known (Urban et al. 2016). Thus, the real challenge of process-based models is to find a model formulation that incorporates the most important processes affecting species dynamics (e.g. dispersal or local extinction) without making them too complex (Fordham et al. 2018), especially with respect to the available data.

Dynamic occupancy models (MacKenzie et al. 2003) (hereafter dynocc models) represent a healthy tradeoff between model complexity and data requirements, and the incorporation of key processes describing range dynamics (Briscoe et al. 2021). These models characterize occupancy dynamics by explicitly modelling the probability of local colonization and extinction between successive time steps using a statistical approach, exactly as in a metapopulation model

(Hanski 1998). Dynocc models are calibrated with occurrence data of different 'seasons' (typically breeding seasons for birds). These data are arguably easier to obtain than abundance and demographic data required by purely process-based models (Zurell 2017). One distinctive feature distinguishing dynocc models from traditional metapopulation models is their ability to account for imperfect detection when multiple surveys per season are available (MacKenzie et al. 2003, Guillera-Arroita 2017). Additionally, these models have the flexibility to define different functions to describe the probability of colonization, extinction, and detection. This enables one to study drivers on each process.

In a recent study evaluating different SDMs (Briscoe et al. 2021), it was shown that dynocc models were the best suited to capture short-term temporal trends using presence-absence data, although their spatial accuracy was low when looking only at sites that showed some occupancy change (i.e. the models failed at identifying the specific sites where range changes are likely to occur). Moreover, in recent years, several studies have proposed modifications to the original formulation of the dynocc model in order to incorporate the influence of the occupancy state of the focal cells' neighbors (neighborhood connectivity) into both colonization and extinction processes (Bled et al. 2011, Risk et al. 2011, Chandler et al. 2015, Broms et al. 2016). In this way, dynocc models turn out to be more spatially explicit since, on one hand, for a colonization process to happen the presence of neighbors around the focal site is needed; and, on the other hand, extinction risk diminishes for sites with a higher neighborhood connectivity due to the rescue effect (i.e. higher chances of immediate recolonization after extinction) (Brown and Kodric-Brown 1977). All in all, incorporating spatial effects into the dynocc models by including connectivity-dependent colonization and extinction was found to enhance models' performance compared to dynocc models assuming independence of sites (Yackulic et al. 2012, Eaton et al. 2014, Hall et al. 2018).

Aside from the neighborhood information, incorporating habitat variables at the site level is key to obtain realistic predictions about occupancy dynamics because species distributions are strongly influenced by environmental conditions (Ehrlén and Morris 2015). Studies investigating species dynamics using dynocc models with spatial effects often incorporate habitat covariates at the site level together with neighborhood connectivity indicators as predictors for the colonization and extinction functions, for example Molinari-Jobin et al. (2018) and Zylstra et al. (2019). However, there are no studies at present assessing the relative importance of each predictor type individually. More importantly, it is yet to be assessed whether adding spatial effects into dynocc models already incorporating a good predictive set of site-level habitat covariates can significantly improve model performance, and how such patterns may differ among species.

In the present study, we defined three objectives to evaluate how the incorporation of neighborhood connectivity to the colonization and extinction terms of a dynocc model affects predictive performance for avian species: 1) we analyze

how model performance of spatial dynocc models is affected by the choice of neighborhood connectivity measures, 2) we compare the predictive power of neighborhood connectivity and site-level habitat covariates (hereafter habitat covariates) as predictors of colonization and extinction, and finally 3) we evaluate the performance gain obtained by using both connectivity and habitat covariates together as predictors as compared to using them separately. To this end, we fitted dynocc models with five different formulations for the probability of extinction or colonization: two with only spatial effects (each with a different formulation for connectivity), one with only habitat covariates, and two with both spatial and habitat covariates. We fitted all models for 46 bird species that breed in Catalonia (north-eastern Spain) using binary occurrence data from the second and third Catalan Breeding Birds Atlases (hereafter, CBBA2 and CBBA3, respectively) over a period of 16 years (Estrada et al. 2004, Franch et al. 2021). Furthermore, we conducted a simulation exercise to validate the efficacy of our study design and model formulations, ensuring their ability to accurately capture both neighborhood connectivity and habitat effects in species colonization and extinction patterns.

## Material and methods

### Data collection, study area, and study species

We used survey data from the second and third Catalan Breeding Bird Atlases that cover the periods 1999–2002 (CBBA2) and 2015–2018 (CBBA3), respectively, thus with an interval of approximately 16 years. The survey data include five 1 km<sup>2</sup> squares within each 100 km<sup>2</sup> square in Catalonia, a 32 114 km<sup>2</sup> region in north-eastern Spain. They contain a total of 1506 sites that were sampled in both atlas projects. In each sampling site, two surveys were conducted during the breeding season, each lasting for one hour. The first survey was conducted in March/April and the second in May/June. Most squares (86% in CBBA2 and 97% in CBBA3) were fully surveyed in a single year whereas the remaining squares were surveyed in consecutive years. Fieldwork was carried out between sunrise and 11:00 or between 6:00 and sunset (Franch et al. 2021).

Among the 214 species of birds detected in the atlases, we selected a subset of species based on the following criteria: 1) species that are mostly associated with Mediterranean forests and shrublands according to the CBBA2 (Estrada et al. 2004); that is, 53 out of 214 species and 2) species that were observed in at least 100 sites in the CBBA2. Based on these criteria, we retained 46 species for the analysis (see Supporting information for a list of species).

### Site-level habitat covariates

We worked with a large set of predictor variables (Supporting information) to capture the wide-ranging environmental and anthropogenic variation across Catalonia to fit the habitat

models (see ‘Model fitting’ sub-section). We selected variables mainly related to vegetation and landscape due to their dynamic nature, the direct implications for management-oriented thinking, and potential future model applications. All covariates were measured during the period of the CBBA3. Below we provide an overview of all the variables used, but for a detailed description of each variable see the Supporting information.

Land cover variables were generated using the Habitat Cartography of Catalonia for 2018 (agricultura.gencat.cat/ca/serveis/cartografia-sig/bases-cartografiques/habitats/habitats-catalunya/index.html). A total of 29 land cover categories (Supporting information) were aggregated from the original 279 classes of the map (Franch et al. 2021). For each 1-km square we calculated the proportion of each land cover category and, based on those, the Shannon diversity index as an indicator of landscape heterogeneity (Franch et al. 2021). We also included a variable indicating the proportion of each square affected by wildfires based on the land cover map of Catalonia for 2017 (mesc.creaf.cat/index\_usa.htm).

We calculated three vegetation productivity indices (Radeloff et al. 2019) in the study squares (Supporting information) based on Modis images of 500 m resolution (MOD13A1), and seven variables describing forest structure characteristics (Supporting information) from the 2016 tree biophysical variables at 20 m resolution (icgc.cat/Administracio-i-empresa/Descarregues/Capes-de-geoinformacio/Mapes-de-variables-biofisiques-de-l-arbrat-de-Catalunya) based on LiDAR2 data (aero-transported sensors).

Finally, using the 2018 population census of Catalonia (Idescat 2021) and topographic maps we calculated the distance from the study square centroids to the closest large town or city (> 10 000 inhabitants) to account for at least one anthropogenic factor (Franch et al. 2021).

### Survey covariates

For the detection component of the dynocc models we used survey date (day of the year as Julian day) because some species can have breeding seasons shorter than the atlas sampling period (from March to June) and atlas period (CBBA2 or CBBA3) to allow differences in probability of detection between atlases. We defined five different formulations for the detection probability: only the intercept, date as the only predictor (with its linear and quadratic terms), atlas period as the only predictor, and both date and atlas period as predictors with or without an interaction term between the two (Table 1).

### Dynamic occupancy modelling

We based our occupancy modelling framework on the multi-season occupancy model accounting for imperfect detection as described by MacKenzie et al. (2003) (also known as the dynocc model). The model is formulated as a first-order hidden Markov model where the probability of occupancy ( $\psi$ ) at time  $t+1$  of a focal site depends on its

Table 1. Overview of the 10 dynamic occupancy models fitted to the data of 46 bird species. SDM-CBBA2, species distribution model of the second Catalan Breeding Bird Atlases; BRM, buffer radius model; IFM, incidence function model.

Model	Occurrence $t_0$	Colonization	Extinction	Detection
Fixed coext	SDM-CBBA2	Intercept	Intercept	Intercept
Fixed coext	SDM-CBBA2	Intercept	Intercept	Date
Fixed coext	SDM-CBBA2	Intercept	Intercept	Atlas
Fixed coext	SDM-CBBA2	Intercept	Intercept	Date + Atlas
Fixed coext	SDM-CBBA2	Intercept	Intercept	Date: Atlas
Spatial coext BRM	SDM-CBBA2	BRM	BRM	Best fixed coext
Spatial coext IFM	SDM-CBBA2	IFM	IFM	Best fixed coext
Habitat coext	SDM-CBBA2	Habitat covariates	Habitat covariates	Best fixed coext
Habitat-Spatial coext BRM	SDM-CBBA2	Habitat covariates + BRM	Habitat covariates + BRM	Best fixed coext
Habitat-Spatial coext IFM	SDM-CBBA2	Habitat covariates + IFM	Habitat covariates + IFM	Best fixed coext

occupancy status at the previous time step  $t$  and the corresponding probabilities of colonization ( $\gamma$ ) and extinction ( $\epsilon$ ) (Eq. 1):

$$\psi_{t+1} = \psi_t (1 - \epsilon) + (1 - \psi_t) \gamma \quad (1)$$

Since the state 'occupied' in the Markov chain might not be observed, the model hierarchically accounts for this by inclusion of the probability of observation, which is defined as the product of the probability of occurrence and the probability of detection ( $p$ ). The probability of detection is included as a logistic submodel depending on survey covariates (see 'Survey covariates' sub-section).

Effects of covariates can be incorporated in the form of simple GLMs or more complex mechanistic models in each of the four parameter types ( $\psi_t$ ,  $\epsilon$ ,  $\gamma$ , and  $p$ ), we take advantage of this model flexibility to compare different formulations for the extinction and colonization models.

### Spatial formulations of the dynocc model

We used two different spatial formulations of the models wherein we incorporated the effects of neighborhood occupancy on extinction and colonization probability: buffer radius models (BRMs) and incidence function models (IFMs) connectivity measures. The main difference between the two spatial formulations is the number of neighbors they consider in order to calculate neighborhood connectivity. BRMs only consider cells within a fixed radius from the focal cell, which is defined a priori, whereas IFMs consider all the cells in the landscape, but with different influence based on the distance to the focal cell.

BRMs were formulated similar to Yackulic et al. (2012) and Eaton et al. (2014), where neighborhood connectivity is calculated as the mean occupancy probability at time  $t$  of all the neighbor sites within a chosen buffer distance (radius)  $d$  from the focal site ( $\hat{\psi}_{i,t}^{N_{i,d}}$ ). The neighbour connectivity variable is then incorporated as a linear term into the extinction and/or colonization probability functions (Eq. 2):

$$\text{logit}(\Theta_{i,t}) = \beta_0 + \beta_1 \hat{\psi}_{i,t}^{N_{i,d}} + \beta_{1+n} X_n \quad (2)$$

where  $\Theta_{i,t}$  is the probability of extinction or colonization for site  $i$  between seasons  $t$  and  $t + 1$ ,  $\beta$  is a parameter vector to be estimated, and  $X_n$  the optional site-specific habitat covariates.

IFMs were formulated similar to Chandler et al. (2015), where the probability of colonization or persistence ( $1 - \epsilon$ ) between seasons  $t$  and  $t + 1$  is calculated as the cumulative probability Eq. 5 taken from each pairwise probability of colonization or persistence Eq. 3 between the squares in the study area and the focal square (Eq. 3, 4, 5):

$$\delta_{i,j,t} = \frac{\delta_{B,i,t} \exp(-d_{i,j} / \alpha) \psi_{j,t}}{\Omega} \quad (3)$$

$$\text{logit}(\delta_{B,i,t}) = \beta_0 + \beta_n X_n \quad (4)$$

$$\Theta_{i,t} = 1 - \left\{ \prod_{j=1}^M 1 - \delta_{i,j,t} \right\} \quad (5)$$

Here,  $\delta_{B,i,t}$  is the 'baseline' colonization or persistence probability for site  $i$  (which might be constant or depend on habitat covariates  $X_n$  Eq. 4),  $d_{i,j}$  is the distance between sites, and  $\alpha$  is the rate of exponential decay (often interpreted as the mean dispersal distance),  $\psi_{j,t}$  is the occupancy probability of site  $j$  at time  $t$ , and  $\Omega$  is a scaling parameter  $\geq 1$  to help model fitting (see 'Model fitting' sub-section). In Eq. 5,  $M$  is the total number of neighbor sites of the focal site  $i$ .

Our model modifies the Chandler et al. (2015) model as follows: 1) we used an exponential dispersal kernel instead of a Gaussian kernel so the parameter  $\alpha$  could be more comparable to the parameter  $d$  (distance between sites) of the logistic models, 2) we used a deterministic formulation instead of a stochastic one for the autocovariates so the model could be fitted using maximum likelihood estimation (MLE) methods, and 3) we defined the probability of persistence (representing the rescue effect) independently from the probability of colonization so we could analyze the effects of the probability of extinction by itself. Additionally, for computational reasons, we used a bounding box to limit the neighbors to

consider, and only those neighbors at  $< 30$  km from the focal square were considered (i.e. local evaluation of the likelihood, Milleret et al. 2019).

## Model fitting

All models were fitted as two-season dynocc models using CBBA2 and CBBA3 data. The Atlas data were divided into five training and test datasets, respectively, with an 80–20% split in order to perform k-fold cross validation. Each model was separately trained on all training sets for each species, except for the models used to select the best detection and habitat covariates that were fitted using 100% of the data. We assumed that the occupancy status did not change within the breeding seasons of an atlas period, thus treating each atlas period as a single season. For all models, initial occupancy ( $\psi_0$ ) was estimated using the values of expert validated SDMs calculated in the CBBA2 for each species as the only covariate. These models were developed by integrating various algorithms like generalized linear models, boosted models, and maximum entropy models, with predictions weighted by model quality. Expert ornithologists rigorously assessed the final results, considering species distribution, ecology, and potential survey biases, ensuring robustness and accuracy (detailed modelling methodology can be found in Estrada et al. 2004).

We fitted four types of dynocc models for all the selected study species with different formulations for the colonization and extinction modules (Table 1): 1) without habitat covariates or spatial effects (only intercept): *fixed coext* models, 2) with spatial effects only: *spatial coext* models, 3) with habitat covariates only: *habitat coext* models, and 4) with both habitat covariates and spatial effects: *habitat-spatial coext* models. Below we describe the characteristics of each model type.

First, we fitted *fixed coext* dynocc models for each species in order to select the best detection model to be used in the other model types, as well as to have a baseline metric for model comparison. *Fixed coext* models were fitted using fixed colonization and extinction parameters (only the intercept) and selecting the best detection formulation (see 'Survey covariates' sub-section) in terms of the Akaike information criterion (AIC). We used AIC for model selection instead of directly comparing model predictions to the observed data, because the latter can be biased by imperfect detection, whereas the former is based on the model likelihood which in dynocc models explicitly accounts for it. On the other hand, we note that it could be possible that the best detection model changed when covariates are added to the colonization and extinction modules. Nevertheless, it seemed reasonable and practical to select a single detection model for each species.

Then, we fitted spatial models without habitat covariates. For each species, we fitted dynocc models with a spatial term for the colonization and extinction terms using either BRM (*spatial coext BRM* models) or IFM (*spatial coext IFM* models) formulations. To assess how the distance parameters affected model performance, for each species we used eight values of

the parameter  $d$  (1, 2, 3, 5, 7.5, 10, 15, and 20 km) for BRMs and five values of the parameter  $\alpha$  (1, 2.5, 5, 10, and 20 km) for IFMs (i.e. for each species we fitted eight *spatial coext BRM* models and five *spatial coext IFM* models in total). We used a gridded design for the sites and, for both spatial models, connectivity was calculated using the probability of occurrence values ( $\psi$  in Eq. 2–3) of the CBBA2's SDMs of all neighboring sites (i.e. the unconditional probability of occupancy).

For the *habitat coext* models, we did not have a priori hypotheses of the variables most influencing colonization or extinction, hence we performed a variable selection from the set of 42 site-level habitat covariates for each term separately. Starting from the *fixed coext* model, we used a forward selection based on AIC to select up to four variables per species and model term (extinction or colonization). Both linear and quadratic terms were evaluated for each variable, and we only kept the one resulting in a largest decrease of AIC. While we acknowledge the existence of alternative model selection approaches, such as stepwise selection, we opted for forward selection and a maximum of four variables per species to mitigate the risk of model overfitting. Finally, we fitted *habitat coext* models with the best variables selected for the extinction and colonization terms.

Finally, we fitted *habitat-spatial coext* models for each species by incorporating both spatial BRM or IFM variables and the habitat covariates selected in the *habitat coext* models to the extinction and colonization terms. For each species, we fitted *habitat-spatial coext* models for the same set of five  $\alpha$  (IFM) and eight  $d$  (BRM) values used in the *spatial coext* models. All models other than the spatial IFM models were fitted using the function *coext* in the R-package 'unmarked' (Fiske and Chandler 2011, Kellner et al. 2023) to obtain maximum likelihood estimates of the parameters. For the spatial IFM models, we modified the source code of the *coext* function so that the probability of colonization and extinction could be expressed as a function of connectivity as described in the sub-section Spatial formulations of the dynocc model. In the case of the BRM models, implementation was straightforward using the *coext* function, since the neighborhood information of each study square is calculated a priori (using the occupancy values of the CBBA2's SDMs), and it is just incorporated as another site-level covariate.

A caveat of using MLE to fit complex models with multiple variables is that the optimization algorithms can sometimes get stuck in local optima instead finding the global optima (Kéry and Royle 2021). To avoid this, we fitted all models with 10 sets of starting values for the regression parameters (uniformly distributed between  $-1.5$  and  $1.5$ ) and kept the best fit in terms of AIC. Additionally, for the spatial IFM models we used a set of different scaling parameters  $\Omega$  for the  $\delta_{i,j,t}$  values, which we found to be necessary for the models to correctly converge, especially for large values of alpha.

## Model performance

We evaluated the spatial performance of the models by comparing models' predicted probability of occurrence with the

collapsed observed data (i.e. assigning a presence if the species was observed in at least one survey, and absence otherwise). For each trained model we calculated the AUC (area under the ROC curve) using the model predictions at the second time step and the respective test data observed in the CBBA3. AUC provides a measure of model discrimination (whether occupied sites are consistently ranked higher than unoccupied sites). Moreover, for each trained model, we calculated the AUC difference relative to the *fixed coext* models (null colonization and extinction models). By calculating the AUC increase, we could directly evaluate the relative influence of habitat and neighborhood connectivity in the predictions of occurrence change. Finally, for each species and model type, we averaged the absolute AUC and AUC increase results for the five test sets.

Furthermore, for each species, we visually evaluated the predicted occurrence change from each dynocc model by plotting the predicted probability of occurrence at  $t_1$  minus the probability of occurrence at  $t_0$  for all the squares in the study area.

### Simulation analysis

We aimed to compare dynocc models incorporating habitat and neighborhood connectivity effects on the colonization and extinction probability. However, we were unsure if – with our modelling approach applied to this specific case study, where only 5% of cells in the landscape were surveyed – we could accurately capture the effects of each variable type on species colonization and extinction. To address this uncertainty, we conducted a simulation study to assess our ability to evaluate the neighborhood connectivity and habitat effects given the limited data available in the study area.

We simulated species occupancy dynamics for four distinct scenarios: one scenario with colonization and extinction rates depending on habitat covariates only, one depending on neighborhood connectivity only, and two scenarios with both habitat and connectivity effects with strong and weak connectivity effects, respectively (Fig. 1). These simulations were conducted under the same landscape size and survey design as in our case study, and we evaluated the goodness of fit of the model in terms of AUC.

To conduct these simulations, we created a  $180 \times 180$  grid and followed these steps: first, we used a Gaussian random field as a covariate to determine the initial probability of occupancy. Then, we simulated occupancy changes over one time step. To achieve this, we simulated two more Gaussian field variables representing the emergence of new habitat and the loss of existing habitat, which we used as single covariates for the colonization and extinction, respectively, for all the scenarios including habitat effects. For the scenarios including connectivity effects, we used the BRM model (see ‘Spatial formulations of the dynocc model’ sub-section) with an arbitrary dispersal parameter  $d$  of 10 km for the probabilities of colonization and extinction. We ran 10 replicates of occupancy data for each model type and scenario.

For each replicate, we randomly selected 5% of the landscape cells and simulated two visits with a fixed detection probability of 50% (mean average detection probability of the species in our case study). Subsequently, we fitted the resulting detection data of each scenario to all the model types used in the study (Table 1): *fixed coext*, *habitat coext*, *spatial BRM coext*, *spatial IFM coext*, *habitat-spatial BRM coext*, and *habitat-spatial IFM coext*.

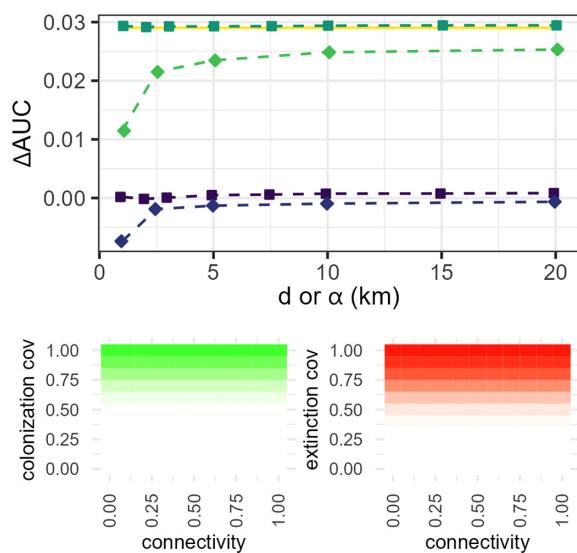
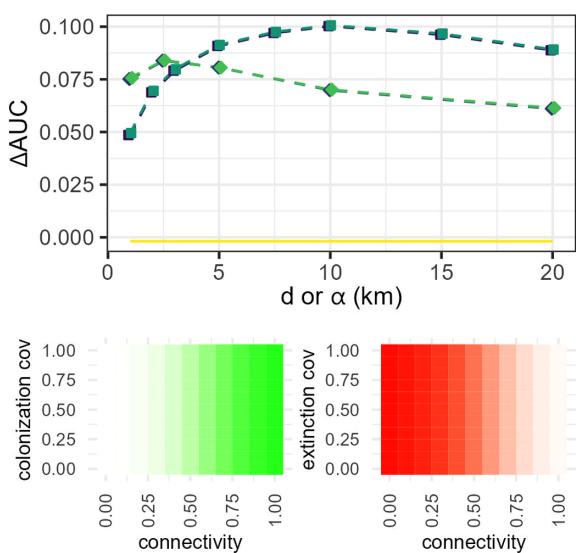
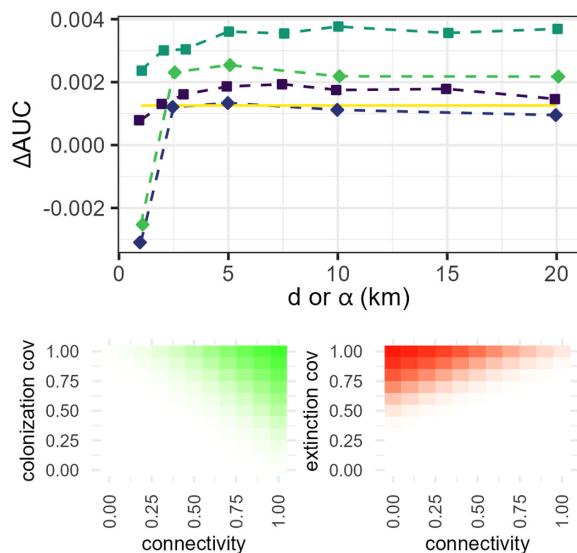
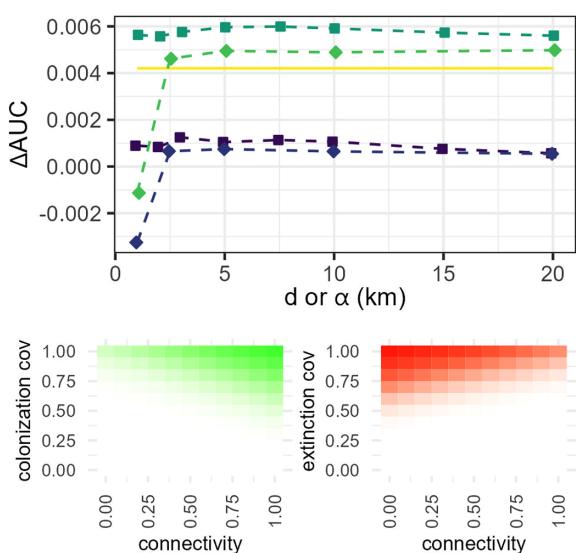
To assess the spatial predictive performance of our models, we followed the same methodology as for the Catalan bird species models (see ‘Model performance’ sub-section). We calculated both the AUC and the AUC increase respective to the *fixed coext* models using the model predictions of occupancy at the second time period and the observed simulated data.

## Results

### Simulation analysis

Model performance increase relative to the *fixed coext* model was well correlated between *coext* model type and scenarios of species colonization and extinction. However, the increase in AUC of the best model fit was highly variable between scenarios, being much smaller for the two scenarios incorporating both connectivity and habitat effects (0.029 only habitat scenario, 0.101 only connectivity scenario, 0.004 habitat and strong connectivity scenario, and 0.006 habitat and weak connectivity scenario; Fig. 1). This variability can be attributed to differences in the absolute AUC values of the *fixed coext* models for each scenario (0.78 only habitat scenario, 0.70 only connectivity scenario, 0.86 habitat and strong connectivity scenario, and 0.85 habitat and weak connectivity scenario), resulting in varying margins of improvement between scenarios.

As expected, *spatial coext* models showed a negligible performance increase for the habitat only scenario and *habitat coext* models for the connectivity only scenario, respectively (Fig. 1). On the other hand, *habitat-spatial BRM coext* models performed almost identically to the *habitat coext* models for the habitat only scenario while *habitat-spatial IFM* were slightly inferior (Fig. 1). For the connectivity only scenario *habitat-spatial BRM* and *IFM* models performed identically to the *spatial BRM* and *IFM* models, respectively, with *IFMs* performing better for smaller dispersal distance parameters and *BRMs* for bigger dispersal distance parameters (Fig. 1). For the scenarios incorporating both connectivity and habitat effects, *habitat-spatial BRM coext* models performed better than the *habitat* and *spatial BRM coext* models alone ( $\times 3$  and  $\times 2$ , respectively for the strong connectivity scenario and  $\times 1.4$  and  $\times 4.8$ , respectively for the weak connectivity scenario in terms of AUC increase; see Fig. 1). As expected, *habitat-spatial IFM* models performed worse than *spatial-habitat BRM*, but nonetheless they also performed better than the *habitat* and *spatial BRM* models alone (Fig. 1). Finally, the effects of the dispersal distance parameter were less noticeable for the scenarios incorporating both habitat and connectivity effects compared to the scenario incorporating only connectivity effects.

**1: Only Habitat effects****2: Only Connectivity effects****3: Habitat and strong Connectivity effects****4: Habitat and weak Connectivity effects****model**

- spatial BRM coext
- habitat-spatial BRM coext
- ◆ spatial IFM coext
- ◆ habitat-spatial IFM coext
- habitat coext

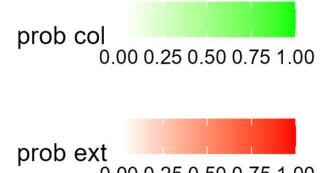


Figure 1. Model performance results for the four simulated colonization and extinction scenarios. Model performance is shown as the mean increment of AUC relative to the *fixed coext* models of all the replicates. Y-axes are different for the distinct scenarios in order to compare the relative performance increase of each model type. For each scenario, the true effects of connectivity and habitat for the probabilities of colonization and extinction used for the simulations are also shown below. See 'Model fitting' sub-section for detailed model specifications.

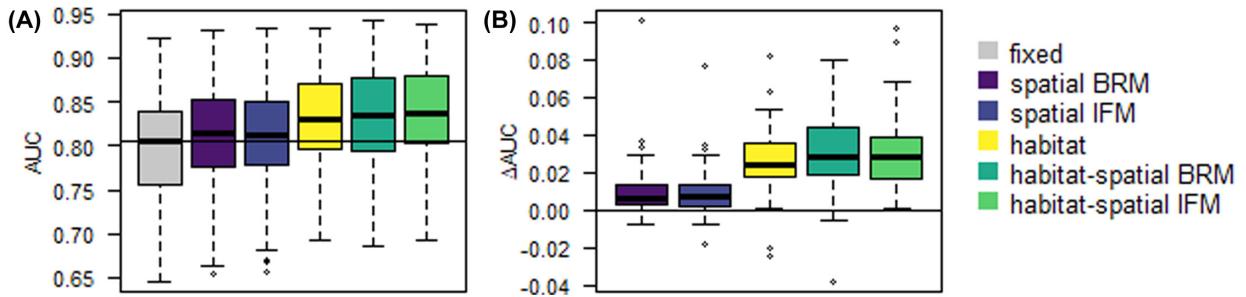


Figure 2. Boxplots of mean absolute AUC (A) and mean relative AUC increase (B). Each model type was fitted to the 46 study bird species. Models were fit to five different training sets and evaluated as the mean absolute AUC for the five test sets (A) or as the mean AUC increase ( $\Delta\text{AUC}$ ) compared to the *fixed coext* models for the five test sets (B). Only the best-performing spatial models were considered (i.e. with the best dispersal parameters in terms of AUC). The horizontal line in (A) indicates the median AUC of the *fixed coext* models, and the horizontal line in (B) indicates 0 AUC increase.

In conclusion, findings from the simulation study show that our study design, coupled with the chosen model formulations, is adequate at capturing both neighborhood connectivity and habitat effects in species colonization and extinction patterns. Moreover, it enables a robust assessment of the relative importance of each variable type.

### Model performance overview

All model types (including the *fixed coext* models) had relatively high spatial performance for most species across the study landscape, as measured by AUC calculated at the second time period using the test datasets, with mean AUC for all methods  $> 0.8$  (Fig. 2A) which indicates strong discrimination between presence and absence predictions.

When spatial effects were included in the colonization and extinction terms (*spatial coext* models), model performance for the best model selected for each species (in terms of distance parameter) was increased for most species (39/46 for BRMs and 37/46 for IFMs) compared to *fixed coext* models (i.e.  $\text{AUC}_{\text{spatial coext}} > \text{AUC}_{\text{fixed coext}}$ ). Both spatial model types had similar performance results. For BRMs, the mean  $\Delta\text{AUC}$  was 0.010 with a range of  $-0.007$  to  $0.101$ , while IFMs exhibited a mean  $\Delta\text{AUC}$  of 0.009, ranging from  $-0.018$  to  $0.076$  (Fig. 2B). Moreover, mean absolute AUC difference between IFMs and BRMs *spatial coext* models was relatively small in general (mean 0.004, SD 0.004) with BRMs performing better for 25/46 species. As expected, across all species, the probability of colonization increased for higher connectivity values, whereas the probability of extinction decreased for lower connectivity values (Supporting information).

Including site-level habitat covariates to the colonization and extinction terms (i.e. *habitat coext* models) increased model performance for most species compared to the *fixed coext* models (44/46), with mean  $\Delta\text{AUC}$  0.025 ranging from  $-0.025$  to  $0.082$  (Fig. 2B). The best variables selected were mostly different for the colonization and extinction terms: for 15 species there were no variables in common, for 22 species only 1/4 variables were in common, for 7 species 2/4 variables were in common, and for 2 species 3/4 were in common

(see Supporting information for all the variables selected for each species and model term).

### Habitat effects versus connectivity effects

Comparing the performance of *habitat coext* models with that of the best *spatial coext* models, *habitat coext* models outperformed the latter for most species (37/46); see Supporting information for the results of each species individually. Additionally, predicted occurrence change showed different spatial patterns between the *habitat coext* and the *spatial coext* models at the landscape level for most species (Supporting information). Furthermore, *habitat coext* models generally predicted higher occurrence probability increases and decreases at specific areas where occurrence change was high (Supporting information).

Species were classified based on their relative association with habitat and neighborhood connectivity effects alone, revealing distinct patterns. Approximately 33% of the species showed a stronger association with habitat effects ( $\Delta\text{AUChabitat} > 5\Delta\text{AUCspatial}$ ), 35% displayed associations with both variable types, although more prominently with habitat effects ( $2\Delta\text{AUCspatial} < \Delta\text{AUChabitat} < 5\Delta\text{AUCspatial}$ ), and 24% showed no clear dominance of either habitat or connectivity effects ( $2\Delta\text{AUCspatial} > \Delta\text{AUChabitat}$  and  $2\Delta\text{AUChabitat} > \Delta\text{AUCspatial}$ ). A smaller proportion (2%) exhibited associations with both but leaned towards connectivity ( $2\Delta\text{AUChabitat} < \Delta\text{AUCspatial} < 5\Delta\text{AUChabitat}$ ), and 7% were primarily associated with connectivity effects ( $\Delta\text{AUCspatial} > 5\Delta\text{AUChabitat}$ ). See Supporting information for the classification of each species.

### Joint habitat and connectivity effects

When the colonization and extinction terms incorporated both neighbor and habitat covariates (i.e. *habitat-spatial coext* models), the best model selected for each species (in terms of distance parameter and spatial model type) generally performed better compared to either *spatial coext* models or *habitat coext* models (40/46 species). However, performance

gains relative to the best model between the *spatial coext* model and the *habitat coext* model of each species were relatively low for most species (mean AUC gain 0.007 ranging from -0.004 to 0.020). Furthermore, for species for which the performance increase was the lowest, the *habitat-spatial coext* models yielded very similar predictions (i.e. similar spatial patterns of occurrence change) compared to the *habitat coext* models (Fig. 3, Supporting information).

Of particular interest was the relative performance increase in  $\Delta\text{AUC}$  compared to the best of the *spatial coext* model and *habitat coext* model for each species ( $\Delta\text{AUC}_{\text{habitat-spatial}} / \max(\Delta\text{AUC}_{\text{habitat}}, \Delta\text{AUC}_{\text{spatial}})$ ); see Supporting information. For species that responded mainly to habitat effects, we still observed an increase of performance for *habitat-spatial coext* models (mean 1.25X). On the other hand, relative increase in  $\Delta\text{AUC}$  was generally bigger for species that showed no clear dominance between habitat and connectivity effects compared to species that responded more to habitat effect (mean 1.51X for the former and mean 1.21X for the latter). Finally, for the few species that were primarily driven by connectivity effects, there was a large variability in the performance increase of the *habitat-spatial coext* models (range 0.89, 3.4X).

## Discussion

In the present study, we evaluated whether incorporating neighborhood connectivity as a predictor for the

colonization and extinction terms in a dynamic occupancy model (MacKenzie et al. 2003) could improve model predictions of changes in species occurrence between two time periods. For that, we fitted a variety of models to data from 46 breeding bird species typical of forest and shrublands from two atlas projects carried out in Catalonia and separated by 16 years. We compared dynocc models that only accounted for neighborhood connectivity with dynocc models that only incorporated site-level habitat covariates to assess the predictive power of each variable type individually. Additionally, we evaluated the joint effects of neighborhood connectivity and site-level habitat covariates in predicting extinction and colonization. All models were tested for the 46 bird species showing different temporal trends. We found that, for most species, habitat models performed better than models accounting for connectivity alone. In addition, adding site connectivity information to habitat models generally improved model performance, especially for species with similar performance gains for habitat and connectivity models alone. These results highlight the importance of evaluating different model formulations and assessing which aspects of the model are more important depending on the study species and the study design.

Incorporating neighborhood connectivity information to the colonization and extinction probabilities improved dynocc model performance for most species compared to models where colonization and extinction rates were constant. These results go in line with metapopulation theory: colonization is positively correlated with neighborhood connectivity and

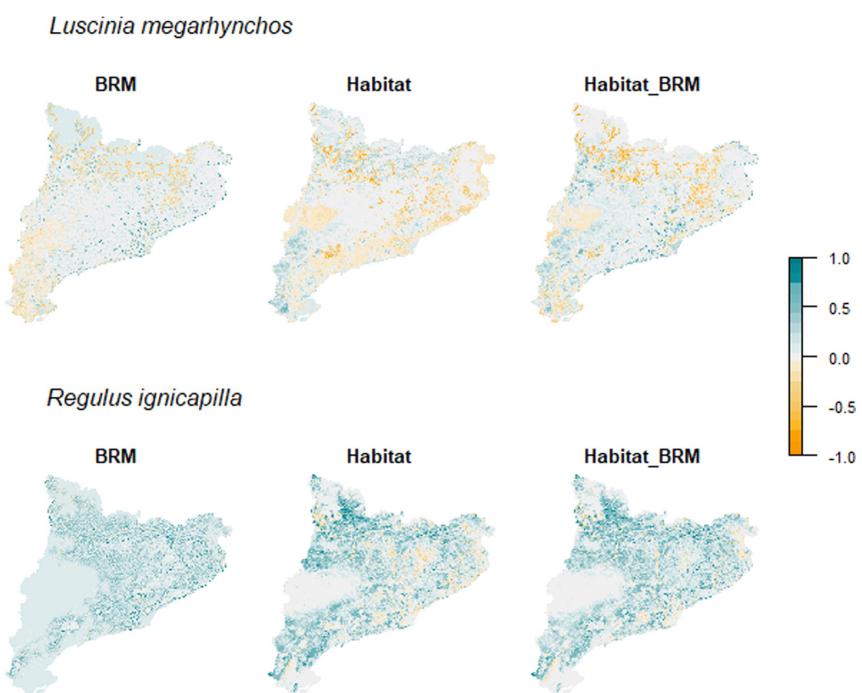


Figure 3. Predicted change in probability of occurrence of the different models. Examples of the predictions of occurrence change between the two atlas periods (occurrence at  $t_1$  – occurrence at  $t_0$ ) of the different dynocc models analysed. On the one hand, for *Luscinia megarhynchos*, as an example of a species where *habitat-spatial coext* models achieved a substantial increase in performance compared to *habitat coext* models, the predictions of distribution change showed different patterns between the two models. On the other hand, for *Regulus ignicapilla*, as an example of a species where *habitat-spatial coext* models only achieved a small increase in performance compared to *habitat coext* models, we can see that the model predictions are very similar between the two models.

extinction is negatively correlated with connectivity due to the rescue effect (Hanski 1999). We analyzed two of the most commonly used dispersal kernel formulations in spatial dynocc models with connectivity-dependent colonization/extinction: BRMs and IFMs. These two model formulations result in contrasting connectivity measures, which have been found previously to perform differently (Moilanen and Nieminen 2002, Nathan et al. 2012, Hall et al. 2018). However, in our study we found that for most species the differences in model predictive performance in terms of AUC between the two spatial formulations of colonization and extinction were minimal. For a minority of species for which there was a marked difference, BRMs generally performed better than IFMs. Aside from the model formulation, the dispersal parameter appears as an important factor affecting model performance. In BRMs it is the buffer radius ( $d$ ) which defines the entire neighborhood that can affect dispersal, whereas in IFMs the parameter  $\alpha$  represents the mean dispersal distance of the species in the exponential kernel (which is often unknown and must be estimated). For most of the study species, differences in model performance were minimal for the different values of dispersal parameters used (ranging from 1 to 20 km) in the connectivity measures, especially when habitat covariates were included. This might be because connectivity values were highly correlated for the different dispersal distances analyzed.

Dynocc models with only site-specific habitat covariates for the colonization and extinction terms (*habitat coext* models) performed better than models only accounting for neighborhood connectivity (*spatial coext* models) for most species (80%). Nevertheless, the relative effects of neighborhood connectivity compared to habitat covariates were still relevant for 67% of the species ( $\Delta$ AUC of *spatial coext* models was at least 20% of the  $\Delta$ AUC of the *habitat coext* models). These results suggest that, generally, in dynocc models, considering site-level habitat covariates alone leads to superior outcomes compared to accounting for site connectivity alone when predicting the occupancy dynamics of bird species prevalent in a region (i.e. present for an extended period). Nevertheless, including the latter is still important, because colonization and extinction processes can be influenced by dispersal constraints. Our results could be explained by the fact that common bird species are in quasi-equilibrium with the environment and therefore if the environment does not suffer major changes, birds will not experience large changes in their distributions no matter how good or poor their neighborhood connectivity is. Conversely, invasive or reintroduced species experiencing an expansion process or species linked to active disturbance regimes such as fires (Zozaya et al. 2012) will typically not be in equilibrium with the environment. Therefore, their dynamics will probably be more influenced by their dispersal capabilities, and neighborhood connectivity might be a better predictor of colonization and extinction than habitat covariates themselves. Finally, we tested if the joint information of connectivity and habitat would improve models' performance as compared to accounting for them individually.

We observed performance gains for almost all species, with higher gains observed for species that showed similar performance improvements for habitat and connectivity models alone. However, the performance increase in terms of absolute AUC was relatively low for most species. This suggests that while the models can identify habitat and connectivity effects on species colonization and extinction processes, there remains considerable stochasticity in the occupancy dynamics patterns of the studied species. This is further exemplified in our simulation study, where for species with known habitat and connectivity effects, models that incorporated both variables showed a limited increase in AUC. This observation arises from the landscape's inherent complexity, where gradients of habitat suitability and neighborhood connectivity values vary widely. Consequently, the probability of colonization and extinction at each site becomes highly stochastic, ranging between 0 and 1. This unpredictability underscores the challenges in accurately capturing and modelling species occupancy dynamics in heterogeneous landscapes. Finally, since the models we used are entirely data-driven and require a large amount of data to obtain reliable estimates, for rare species with limited presence records, models might fail to converge. In such cases, the application of alternative sampling designs and modelling methodologies should be considered to obtain reliable predictions (Zhang et al. 2020, Jeliazkov et al. 2022).

There is an often-heard claim within the scientific community to make models of species distribution change more process-based in order to improve their predictions (Urban et al. 2016, Zurell 2017). Explicitly incorporating ecological processes into models goes at the expense of requiring extra data and computational resources, as well as additional expertise in order to implement them (given their complexity) compared to classic correlative methods (Hefley et al. 2017). Therefore, studies evaluating semi-process-based models (where ecological processes are simplified and added to the model following a statistical approach), like the present one, are valuable to assess which processes are worth adding into a model and for which study designs or study species. Our results showed that spatial dynocc models incorporating neighborhood connectivity as a representation of the dispersal process can improve model performance for most species. However, for some species, the performance gains relative to non-spatial dynocc models were only marginal and probably incorporating other key processes affecting distribution dynamics into the model (i.e. species interactions) would be necessary to improve model predictions. Further studies are needed to discern the species properties (e.g. specialists versus generalists or species migratory behavior) that makes their models more prone to show a positive response to the introduction of spatial effects. Furthermore, studies comparing model performance for different study designs (i.e. gridded cells versus genuine metapopulation patches) and survey efforts (i.e. the coverage of surveyed cells within the study landscape) should be encouraged since these are major factors affecting model performance aside from the study species. Overall, our study shows that spatial dynamic distribution models are a

promising tool to improve model predictions for most species compared to non-spatial dynocc models. Furthermore, we establish that site-level habitat covariates are strong predictors of species colonization and extinction for the majority of species, suggesting their inclusion in dynocc models whenever feasible. Exploring new approaches and methodologies to better characterize species dynamics, rather than focusing solely on static spatial patterns, should also be a priority for ecologists working on species distribution models.

**Acknowledgements** – We thank Marc Kéry for his insightful comments and suggestions.

**Funding** – This study was partially funded by MCIN/AEI through the projects GREENRISK (PID2020-119933RB-C22) and CEX-2018-000828-S ‘Centro de Excelencia Severo Ochoa’. OS was supported by a pre-doctoral grant no. PRE2020-092082 by the Spanish Ministry of Science, Innovation and Universities. NA was supported by a Juan de la Cierva fellowship of the Spanish Ministry of Science and Innovation (FCJ2020-046387-I). SF was funded by the Beatriu de Pinós postdoctoral program of the Secretariat of Universities and Research, Ministry of Enterprise and Knowledge of the Government of Catalonia (2021 BP 00134) funded by the EU R+I Horizon 2020 program.

## Author contributions

**Oriol Solà**: Conceptualization (equal); Data curation (lead); Formal analysis (lead); Methodology (lead); Writing – original draft (lead); Writing – review and editing (equal).

**Núria Aquilué**: Conceptualization (equal); Formal analysis (supporting); Supervision (equal); Writing – original draft (supporting); Writing – review and editing (equal).

**Sara Fraixedas**: Writing – original draft (supporting); Writing – review and editing (equal).

**Lluís Brotons**: Conceptualization (equal); Formal analysis (supporting); Supervision (equal); Writing – review and editing (equal).

## Transparent peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ecog.06985>.

## Data availability statement

Code and model evaluation data are available from the Zenodo Digital Repository: <https://doi.org/10.5281/zenodo.11059379> (Solà et al. 2024). Our deposition includes a worked example to fit the *Spatial coext IFM* models used in the study. Species occurrence and environmental data used for fitting models are available upon reasonable request from [o.sola@creaf.uab.cat](mailto:o.sola@creaf.uab.cat).

## Supporting information

The Supporting information associated with this article is available with the online version.

## References

Bled, F., Royle, J. A. and Cam, E. 2011. Hierarchical modeling of an invasive spread: the Eurasian collared-dove *Streptopelia decaocto* in the United States. – *Ecol. Appl.* 21: 290–302.

Briscoe, N. J., Elith, J., Salguero-Gómez, R., Lahoz-Monfort, J. J., Camac, J. S., Giljohann, K. M., Holden, M. H., Hradsky, B. A., Kearney, M. R., McMahon, S. M., Phillips, B. L., Regan, T. J., Rhodes, J. R., Vesk, P. A., Wintle, B. A., Yen, J. D. L. and Guillera-Arroita, G. 2019. Forecasting species range dynamics with process-explicit models: matching methods to applications. – *Ecol. Lett.* 22: 1940–1956.

Briscoe, N. J., Zurell, D., Elith, J., König, C., Fandos, G., Malchow, A. K., Kéry, M., Schmid, H. and Guillera-Arroita, G. 2021. Can dynamic occupancy models improve predictions of species’ range dynamics? A test using Swiss birds. – *Global Change Biol.* 27: 4269–4282.

Broms, K. M., Hooten, M. B., Johnson, D. S., Altweig, R. and Conquest, L. L. 2016. Dynamic occupancy models for explicit colonization processes. – *Ecology* 97: 194–204.

Brown, J. H. and Kodric-Brown, A. 1977. Turnover rates in insular biogeography: effect of immigration on extinction. – *Ecology* 58: 445–449.

Chandler, R. B., Muths, E., Sigafus, B. H., Schwalbe, C. R., Jarchow, C. J. and Hossack, B. R. 2015. Spatial occupancy models for predicting metapopulation dynamics and viability following reintroduction. – *J. Appl. Ecol.* 52: 1325–1333.

Eaton, M. J., Hughes, P. T., Hines, J. E. and Nichols, J. D. 2014. Testing metapopulation concepts: effects of patch characteristics and neighborhood occupancy on the dynamics of an endangered lagomorph. – *Oikos* 123: 662–676.

Ehrlén, J. and Morris, W. F. 2015. Predicting changes in the distribution and abundance of species under environmental change. – *Ecol. Lett.* 18: 303–314.

Elith, J. and Leathwick, J. R. 2009. Species distribution models: ecological explanation and prediction across space and time. – *Annu. Rev. Ecol. Evol. Syst.* 40: 677–697.

Elith, J., Kearney, M. and Phillips, S. 2010. The art of modelling range-shifting species. – *Methods Ecol. Evol.* 1: 330–342.

Estrada, J., Pedrocchi, V., Brotons, L. and Herrando, S. 2004. Catalan breeding bird atlas (1999–2002). – *Lynx*.

Fiske, I. and Chandler, R. 2011. unmarked: an R package for fitting hierarchical models of wildlife occurrence and abundance. – *J. Stat. Softw.* 43: 1–23.

Fordham, D. A., Bertelsmeier, C., Brook, B. W., Early, R., Neto, D., Brown, S. C., Ollier, S. and Araújo, M. B. 2018. How complex should models be? Comparing correlative and mechanistic range dynamics models. – *Global Change Biol.* 24: 1357–1370.

Franch, M., Herrando, S., Anton, M., Villero, D. and Brotons, L. 2021. Catalan breeding bird atlas: distribution and abundance 2015–2018 and change since 1980. – *Cossetània*.

Guillera-Arroita, G. 2017. Modelling of species distributions, range dynamics and communities under imperfect detection: advances, challenges and opportunities. – *Ecography* 40: 281–295.

Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – *Ecol. Lett.* 8: 993–1009.

Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – *Ecol. Modell.* 135: 147–186.

Hall, L. A., Van Schmidt, N. D. and Beissinger, S. R. 2018. Validating dispersal distances inferred from autoregressive occupancy

models with genetic parentage assignments. – *J. Anim. Ecol.* 87: 691–702.

Hanski, I. 1998. Metapopulation dynamics. – *Nature* 396: article 6706.

Hanski, I. 1999. Metapopulation ecology. – OUP.

Hefley, T. J., Hooten, M. B., Russell, R. E., Walsh, D. P. and Powell, J. A. 2017. When mechanism matters: Bayesian forecasting using models of ecological diffusion. – *Ecol. Lett.* 20: 640–650.

Hodgson, J. A., Moilanen, A. and Thomas, C. D. 2009. Metapopulation responses to patch connectivity and quality are masked by successional habitat dynamics. – *Ecology* 90: 1608–1619.

IPBES. 2016. The methodological assessment report on scenarios and models of biodiversity and ecosystem services. – Zenodo, <https://doi.org/10.5281/zenodo.3235429>.

Jeliazkov, A., Gavish, Y., Marsh, C. J., Geschke, J., Brummitt, N., Rocchini, D., Haase, P., Kunin, W. E. and Henle, K. 2022. Sampling and modelling rare species: conceptual guidelines for the neglected majority. – *Global Change Biol.* 28: 3754–3777.

Kellner, K. F., Smith, A. D., Royle, J. A., Kéry, M., Belant, J. L. and Chandler, R. B. 2023. The unmarked R package: twelve years of advances in occurrence and abundance modelling in ecology. – *Methods Ecol. Evol.* 14: 1–8.

Kéry, M. and Royle, J. 2021. Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in r and bugs: volume 2: dynamic and advanced models, 1st edn. – Elsevier.

MacKenzie, D. I., Nichols, J. D., Hines, J. E., Knutson, M. G. and Franklin, A. B. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. – *Ecology* 84: 2200–2207.

Milleret, C., Dupont, P., Bonenfant, C., Brøseth, H., Flagstad, Ø., Sutherland, C. and Bischof, R. 2019. A local evaluation of the individual state-space to scale up Bayesian spatial capture–recapture. – *Ecol. Evol.* 9: 352–363.

Moilanen, A. and Nieminen, M. 2002. Simple connectivity measures in spatial ecology. – *Ecology* 83: 1131–1145.

Molinari-Jobin, A., Kéry, M., Marboutin, E., Marucco, F., Zimermann, F., Molinari, P., Frick, H., Fuxjäger, C., Wölfl, S., Bled, F., Breitenmoser-Würsten, C., Kos, I., Wölfl, M., Černe, R., Müller, O. and Breitenmoser, U. 2018. Mapping range dynamics from opportunistic data: spatiotemporal modelling of the lynx distribution in the Alps over 21 years. – *Anim. Conserv.* 21: 168–180.

Nathan, R., Klein, E., Robledo-Arnuncio, J. J. and Revilla, E. 2012. Dispersal kernels: review. – In: Clobert, J., Baguette, M., Benton, T. G. and Bullock, J. M. (eds), *Dispersal ecology and evolution*. Oxford Academic, pp. 185–210.

Parmesan, C. and Yohe, G. 2003. A globally coherent fingerprint of climate change impacts across natural systems. – *Nature* 421: 37–42.

Piirainen, S., Lehikoinen, A., Husby, M., Kålås, J. A., Lindström, Å. and Ovaskainen, O. 2023. Species distributions models may predict accurately future distributions but poorly how distributions change: a critical perspective on model validation. – *Divers. Distrib.* 29: 654–665.

Radeloff, V. C., Dubinin, M., Coops, N. C., Allen, A. M., Brooks, T. M., Clayton, M. K., Costa, G. C., Graham, C. H., Helmers, D. P., Ives, A. R., Kolesov, D., Pidgeon, A. M., Rapacciulo, G., Razenkova, E., Suttidate, N., Young, B. E., Zhu, L. and Hobi, M. L. 2019. The Dynamic Habitat Indices (DHIs) from MODIS and global biodiversity. – *Remote Sens. Environ.* 222: 204–214.

Risk, B. B., de Valpine, P. and Beissinger, S. R. 2011. A robust design formulation of the incidence function model of metapopulation dynamics applied to two species of rails. – *Ecology* 92: 462–474.

Sofaer, H. R., Jarnevich, C. S. and Flather, C. H. 2018. Misleading prioritizations from modelling range shifts under climate change. – *Global Ecol. Biogeogr.* 27: 658–666.

Solà, O., Aquilué, N., Fraixedas, S. and Brotons, L. 2024. Data from: Evaluating the influence of neighborhood connectivity and habitat effects in dynamic occupancy species distribution models. – Zenodo Digital Repository, <https://doi.org/10.5281/zenodo.11059379>.

Urban, M. C., Zarnetske, P. L. and Skelly, D. K. 2013. Moving forward: dispersal and species interactions determine biotic responses to climate change. – *Ann. N. Y. Acad. Sci.* 1297: 44–60.

Urban, M. C. et al. 2016. Improving the forecast for biodiversity under climate change. – *Science* 353: aad8466.

Yackulic, C. B., Reid, J., Davis, R., Hines, J. E., Nichols, J. D. and Forsman, E. 2012. Neighborhood and habitat effects on vital rates: expansion of the barred owl in the Oregon Coast ranges. – *Ecology* 93: 1953–1966.

Zhang, C., Chen, Y., Xu, B., Xue, Y. and Ren, Y. 2020. Improving prediction of rare species' distribution from community data. – *Sci. Rep.* 10: 12230.

Zozaya, E. L., Brotons, L., Saura, S., Pons, P. and Herrando, S. 2012. Connectivity determines post-fire colonisation by open-habitat bird species: the case of the ortolan bunting *Emberiza hortulana*. – *Ardeola* 59: 57–74.

Zurell, D. 2017. Integrating demography, dispersal and interspecific interactions into bird distribution models. – *J. Avian Biol.* 48: 1505–1516.

Zylstra, E. R., Swann, D. E., Hossack, B. R., Muths, E. and Steidl, R. J. 2019. Drought-mediated extinction of an arid-land amphibian: insights from a spatially explicit dynamic occupancy model. – *Ecol. Appl.* 29: e01859.