WILEY **Global Ecology and Biogeography** | A Journal of Macroecology

Biotic interactions in species distribution modelling: 10 questions to guide interpretation and avoid false conclusions

Carsten F. Dormann[1] | Maria Bobrowski[2] | D. Matthias Dehling[3] |
David J. Harris[4] | Florian Hartig[1,5] | Heike Lischke[6] | Marco D. Moretti[7] |
Jörn Pagel[8] | Stefan Pinkert[9] | Matthias Schleuning[10] |
Susanne I. Schmidt[11] | Christine S. Sheppard[8] | Manuel J. Steinbauer[12,13] |
Dirk Zeuss[14] | Casper Kraan[15,16]

[1]Biometry and Environmental System Analysis, Faculty of Environment and Natural Resources, University of Freiburg, Freiburg, Germany

[2]Institute of Geography/CEN Center for Earth System Research and Sustainability, University of Hamburg, Hamburg, Germany

[3]Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

[4]Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, Florida

[5]Theoretical Ecology, University of Regensburg, Regensburg, Germany

[6]Landscape Dynamics, Dynamic Macroecology, Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

[7]Biodiversity and Conservation Biology, Swiss Federal Research Institute WSL, Birmensdorf, Switzerland

[8]Institute of Landscape and Plant Ecology, University of Hohenheim, Stuttgart, Germany

[9]Faculty of Biology, Ecology – Animal Ecology, Philipps-University Marburg, Marburg, Germany

[10]Senckenberg Biodiversity and Climate Research Centre (BiK-F), Frankfurt (Main), Germany

[11]Institute of Environmental Sciences, University of Koblenz-Landau, Landau, Germany

[12]Department of Geography and Geosciences, GeoZentrum Nordbayern, Friedrich-Alexander University Erlangen-Nürnberg (FAU), Erlangen, Germany

[13]Section for Ecoinformatics and Biodiversity, Department of Bioscience, Aarhus University, Aarhus, Denmark

[14]Department of Zoology, Stockholm University, Stockholm, Sweden

[15]Department of Functional Ecology, Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Bremerhaven, Germany

[16]Helmholtz Institute for Functional Marine Biodiversity at the University Oldenburg, Oldenburg, Germany

Correspondence
Carsten F. Dormann, Biometry and Environmental System Analysis, Faculty of Environment and Natural Resources, University of Freiburg, Tennenbacher Straße 4, 79106 Freiburg, Germany.
Email: carsten.dormann@biom.uni-freiburg.de

Funding information
German Research Foundation, Grant/Award Number: DO786/10-1, SH 924/1-1 and DE 2754/1-1; Marie Skłodowska-Curie action, Grant/Award Number: 700796 and 707491; Danish Carlsbergfondet, Grant/Award Number: CF14-0148; Ministerium für Wissenschaft, Weiterbildung und Kultur Rheinland-Pfalz

Abstract

Aim: Recent studies increasingly use statistical methods to infer biotic interactions from co-occurrence information at a large spatial scale. However, disentangling biotic interactions from other factors that can affect co-occurrence patterns at the macroscale is a major challenge.

Approach: We present a set of questions that analysts and reviewers should ask to avoid erroneously attributing species association patterns to biotic interactions. Our questions relate to the appropriateness of data and models, the causality behind a correlative signal, and the problems associated with static data from dynamic systems. We summarize caveats reported by macroecological studies of biotic interactions and examine whether conclusions on the presence of biotic interactions are supported by the modelling approaches used.

Findings: Irrespective of the method used, studies that set out to test for biotic interactions find statistical associations in species' co-occurrences. Yet, when compared with our list of questions, few purported interpretations of such associations as biotic interactions hold up to scrutiny. This does not dismiss the presence or importance of biotic interactions, but it highlights the risk of too lenient interpretation of the data. Combining model results with information from experiments and

functional traits that are relevant for the biotic interaction of interest might strengthen conclusions.

Main conclusions: Moving from species- to community-level models, including biotic interactions among species, is of great importance for process-based understanding and forecasting ecological responses. We hope that our questions will help to improve these models and facilitate the interpretation of their results. In essence, we conclude that ecologists have to recognize that a species association pattern in joint species distribution models will be driven not only by real biotic interactions, but also by shared habitat preferences, common migration history, phylogenetic history and shared response to missing environmental drivers, which specifically need to be discussed and, if possible, integrated into models.

1 | INTRODUCTION

Ecological theory aims to explain the observed distribution of an organism from its environmental requirements, its physiological tolerances, its ecological (including dispersal, biogeographical and evolutionary) history and its interactions with other organisms (Begon, Townsend, & Harper, 2006). Sometimes, an organism will be little affected by competitors, mutualists, pathogens or predators; a mature tree may appear to grow unfettered by its neighbours for decades. In other cases, an individual may struggle to survive the pressure of its community during every hour of its life; for example, a fish preying and escaping predators in a small river. Biotic interactions such as these can have major effects on local processes and population dynamics, and some, but not necessarily all, scale up to co-determining the range of a species (reviewed, e.g., by Wisz et al., 2013).

Recently, considerable interest has arisen in inferring biotic interactions from large-scale data through the statistical analysis of species (co)distributions. The methods for this task are in substantial flux (González-Salazar, Stephens, & Marquet, 2013; Louthan, Doak, & Angert, 2006; Morales-Castilla, Matias, Gravel, & Araújo, 2015; Nieto-Lugilde, Maguire, Blois, Williams, & Fitzpatrick, 2018; Ovaskainen et al., 2017; Staniczenko, Sivasubramaniam, Suttle, & Pearson, 2017; Warton et al., 2015; Wisz et al., 2013; Zhang, Kissling, & He, 2012). The general idea of these methods is to look for an excess in co-occurrences (possibly indicating facilitation) and deficits in co-occurrence (possibly indicating competition). Setting aside the question of whether correlation in species occurrence patterns is sufficient to infer that a particular causal mechanism is operating (Roughgarden, 1983), a basic requirement for calculating excesses and deficits is to establish some baseline level of co-occurrence.

The most commonly used baseline, particularly in null models (Connor & Simberloff, 1979; Gotelli, 2000; Veech, 2012), is based on each species' landscape-wide occurrence rate. For example, if two species each occur at 50% of the sites, then these methods would expect them to co-occur at 25% of sites (.50 × .50), and any statistically significant deviations from this baseline would be interpreted as evidence for species interactions. This inference would clearly be problematic if, as is often the case, other factors (such as environmental filtering) also influence species occurrence patterns (e.g., Giannini, Chapman, Saraiva, Alves-dos-Santos, & Biesmeijer, 2013) or if data are spatially autocorrelated (Morueta-Holme et al., 2016).

Researchers increasingly use joint species distribution models (jSDMs) to set each species' local baseline according to the abiotic environment instead of making it constant. In this framework, evidence of species interactions can be found in the residuals of the model, after abiotic factors have been controlled for (Pollock et al., 2014). In principle, controlling for abiotic factors should make inferences more robust, because researchers would be able to rule out some alternative explanations for species' co-occurrence patterns based on environmental constraints. As we discuss below, however, it may not always be possible to control for these factors (questions 4 and 5).

One reason is that, even when controlling for the abiotic environment, inferences from co-occurrence data can be misleading as a result of missing biotic or abiotic variables. Harris (2016) demonstrated this point with an example based on two simulated species of shrub and one species of tree (omitted from analysis, surrogating an unknown interactor). The two shrub species impeded one another's growth, yet co-occurred frequently (which would normally imply facilitation or mutualism). In this example, correctly inferring the sign of their interaction required controlling for resource limitation imposed by the tree species, which forced the two shrubs into the same subset of habitats where the tree did not grow. The key point is that indirect biotic interactions may lead to patterns that can be statistically, but ecologically incorrectly, explained by the environment, thereby complicating interpretation (discussed as apparent competition by Connell, 1990, or apparent facilitation by Levine, 1999).

Thus, the question that arises is not so much about the statistical methods as such, but about their interpretation. We have a

proliferation of increasingly complex methods that can detect co-occurrence patterns that deviate from some baseline expectations, which could be a signal of biotic interactions. The dispute, however, is whether such statistical fluctuations in co-occurrence are helpful for inferring biotic interactions from observational data (Faust & Raes, 2012; Hastings, 1987; Meier et al., 2010; Morales-Castilla et al., 2015; Morueta-Holme et al., 2016; Ulrich, Jabot, & Gotelli, 2017). We have few empirical tests of the quality of such inference, as direct empirical evidence for biotic interactions from observational studies is rare and often limited to simple systems (e.g., Ives, Dennis, Cottingham, & Carpenter, 2003; Pfister, 1995; Rees, Grubb, & Kelly, 1996; Schooler, Salau, Julien, & Ives, 2011). Thus, if even such highly detailed local time-series studies struggle to infer local interactions between species, why should the same be so easy for macroecological biotic interaction species distribution models (BI-SDMs)? In other words, how justified are claims of model-based attributions of biotic interactions? Bearing in mind their conceptual advances at the frontiers of this field, we want to raise awareness about the limitations of BI-SDMs. The real litmus test for the correctness of any BI-SDM would be an experimental confirmation of inferred biotic interactions, which has, to our knowledge, rarely been addressed (e.g., Stephens et al., 2016).

In the present work, we review ecological aspects that need to be taken into account when judging the level of evidence for true biotic interactions generated from macroecological patterns, and here we mainly focus on direct biotic interactions and on caveats reported in other BI-SDM studies (Supporting Information Appendices S1 and S2). Most studies have so far ignored the more complex indirect biotic interactions, as illustrated above by Harris (2016). We present a set of questions (Box 1) that scientists and reviewers should ask, and answer, to avoid premature attribution of species association patterns to biotic interactions. Our review

---

> ### BOX 1 Questions aiding careful interpretation and paper review of potential biotic interactions
>
> **Data and models**
> 1. Can a causal chain be drawn, a priori, from a purported interaction to the pattern of interest?
> 2. Are the data suitable for detecting interactions?
> 3. Is the modelling approach suitable for representing the expected structure of the biotic interaction?
>
> **Confounding biotic interactions and environment**
> 4. Is the purported biotic interaction a surrogate for a missing predictor?
> 5. Are biotic interactions dependent on the environment?
> 6. Did the inclusion of joint abiotic and biotic interactions improve the model?
> 7. Can phylogeny confound biotic interactions?
>
> **Dynamics, dispersal, and food web context**
> 8. Are inferred biotic interactions confounded with dispersal limitation?
> 9. Are purported biotic interactions attributable to temporal dynamics?
> 10. Which food web components should be considered in a biotic interactions species distribution model?

extends that of Wisz et al. (2013), where most of these key problems were mentioned, but were not yet transformed into concrete suggestions for interpretation and reporting. With this review, we want to call analysts back to the drawing board, re-aligning their methods with the ecological and evolutionary processes they seek to embrace. We present the questions in a sequence from data/methodological considerations, over confounding biotic interactions with environmental covariates, to challenges posed by the dynamics of ecological systems.

2 | REVIEW OF PUBLISHED STUDIES

To obtain an overview of how species interactions were interpreted, and whether and how the listed questions were addressed, we scanned publications retrieved from Google Scholar (and cross-referencing) dealing with biotic interactions in species distribution modelling (SDM; Supporting Information Appendices S1 and S2).

In line with the cautionary remark with which we started in the Introduction, many studies mention that spatial associations quantified by BI-SDMs should not be interpreted as biotic interactions, but rather as the patterns that may also result from similar environmental niches (e.g., Clark, Gelfand, Woodall, & Zhu, 2014; Giannini et al., 2013; Ovaskainen, Hottola, & Siitonen, 2010; Ovaskainen, Roy, Fox, & Anderson, 2016; Sebastián-González, Dalsgaard, Sandel, & Guimarães, 2015; Warton et al., 2015). They recommend BI-SDMs as hypothesis-generating tools (i.e., for identifying an association that could be worth further ecological investigation for understanding or prediction, but not as interaction detectors). Few studies used population dynamic models to represent biotic effects. One of those studies found biotic interactions to be little supported by the data; the Bayes factor suggested that the evidence was 'barely worth mentioning' (Mutshinda, O'Hara, & Woiwod, 2010).

Moreover, studies commonly offered missing predictors as an alternative explanation to purported biotic interactions, acknowledging that missing predictions would lead to inflated type I errors for detecting biotic interactions. Meier et al. (2010) suggest that the missing predictors might particularly represent environmental conditions changing at small scales, such as soil, microclimate or land use. For instance, they found less effect of purported biotic interactions on shade-intolerant species, although these should exhibit a larger biotic interaction signal (see Supporting Information Appendix S2 for the full list). Morueta-Holme et al. (2016) offer some estimates on type II errors, which increased with decreasing data set size.

Some issues additionally emerged from this review, which are partly covered by our questions. For example, Williams, van der Meer, Dekker, Beukema, and Holmes (2004) notice that many species are too rare to be useful predictors, and Giannini et al. (2013) suggest that an interactor may correlate with sampling effort.

3 | QUESTIONS FOR CAREFUL INTERPRETATION OF BIOTIC INTERACTIONS

Given our review of the methodological and applied literature, we suggest that the following 10 questions should be asked by each analyst
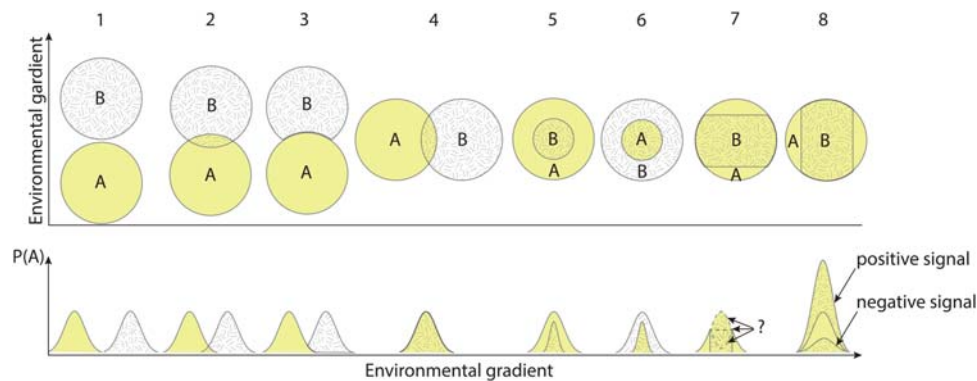
DORMANN ET AL.

Global Ecology
and Biogeography

A Journal of
Macroecology

WILEY | 1007

FIGURE 1 Schematic illustration of eight alternative occurrence scenarios of the target species A and the interacting species B along a simplified environmental gradient (e.g., temperature). Scenarios 1–3 represent species with different niches. In scenario 3, species A excludes species B from the intersection. In scenarios 4–8, species have the same niche and overlap to different degrees. In scenario 5, the effect of B on A can be evaluated, as there is an area of A without B. In contrast, in scenario 6 the effect of B on A cannot be evaluated, as A does not occur outside the niche of B. This need not indicate an interaction, only a smaller niche. In scenario 7, we can assess the effect of B on A only at the margin, but not in the core niche area, as again there are no sites with A but without B. In contrast, scenario 8 allows for an evaluation of the effect of B throughout A's niche, and this effect can be positive or negative. Note the subtle differences between scenarios 5, 7 and 8, which severely affect our ability to interpret an association signal [Colour figure can be viewed at wileyonlinelibrary.com]

to ensure an appropriate interpretation of state-of-the-art methods for analysing biotic interactions or species covariances.

3.1 | Data and models

3.1.1 | Question 1: Can a causal chain be drawn, a priori, from a purported interaction to the pattern of interest?

We realize that the inference of biotic interaction is currently often of an exploratory nature, meaning that there are no clear hypotheses formulated for the data analysis. Although this is generally permissible, one should realize and communicate the ample opportunities for false positives in such an approach. In an exploratory analysis with a large number of tests for pairwise interactions among multiple species pairs, relying on p-values alone is a misguided way to draw inferences (e.g., Peres-Neto, 1999). Correction for multi-testing can solve this problem, but costs power. We suggest that, before data analysis, analysts should first try to reason which interactions between species pairs are ecologically plausible and how they link to the pattern at hand.

3.1.2 | Question 2: Are the data suitable for detecting interactions?

The majority of large-scale data are collected or compiled with pertinent biases (Guillera-Arroita et al., 2015). Records from opportunistic observations [e.g., GBIF (www.gbif.org), eBird (www.ebird.org) or iNaturalist (www.iNaturalist.org)], data gridded to arbitrary resolutions, and expert-drawn range maps (International Union for Conservation of Nature) may provide reasonable information on the distribution of species but may obscure real co-occurrences at ecologically relevant spatial and temporal scales. On what spatial or temporal scale and resolution can we expect a signal of competition between trees, symbiotic effects of mycorrhiza, or herbivory by leaf miners using such data? As more surveys devoted to specific interactions (e.g., pollination or seed dispersal, host–parasitoid networks, herbivory) become available,

such signals may or may not emerge (e.g., Tylianakis et al., 2008), suggesting that indeed local interactions may be detectable in specific large-scale data. At present, however, the burden of proof for the biotic interactions as cause of this pattern lies with the claimant.

A necessary, but not sufficient, condition for detecting an effect of species B on species A from presence or presence–absence data is that within the range of A there exist regions without B (see Figure 1, scenarios 5 and 8; for a full picture on how one species affects the other, all possible combinations of A and B are needed). As an illustration, assume that B occurs whenever, say, the probability (P) of A to occur is < 0.3. We have no statistical tool for telling us whether $P(A) < 0.3$ because B is present, or whether these sites are, say, simply too cold. The situation is similar for data on abundances, except that we require variation in the abundance of B across different expected abundances of A (Figure 1, scenario 6). This condition was used to quantify Lotka–Volterra-style interactions among microbiota in grasslands (Shang et al., 2017). Computing and visualizing the overlap of A and B in (environmental) space can address this issue.

3.1.3 | Question 3: Is the modelling approach suitable for representing the expected structure of the biotic interaction?

Before putting our faith in analytical results, we must verify whether the modelling approach is suitable for detection and representation of expected biotic interactions. For example, as Box 2 and Table 1 detail, most joint species distribution models (jSDMs) use a parameterization of a symmetrical 'association matrix' (representing the correlation in the residuals among the different species). This is in contradiction to ecological experiments at the local scale, which show that the effect of A on B typically is not the same as that of B on A (think of predator–prey relationships, or competition for light; Lawton, & Hassell, 1981; Weiner, 1990). Likewise, coefficients for the effect of a species on the

---

BOX 2 Methods for estimating species associations in species distribution models

There are five main approaches for modelling species associations in macroecological species distribution models (see Table 1 for more details; Nieto-Lugilde et al., 2018 and Ovaskainen et al., 2017 for a recent review). All essentially look for excess or deficits in co-occurrence, relative to a baseline occurrence rate set by the environment or a random null model.

Correlation methods are the simplest case. They evaluate correlations between species' occurrences or abundances, which confounds environment and interactions. Instead of raw data correlations, residuals of regression models should be used to avoid this problem. Partial correlations additionally control for indirect correlations among species (Harris, 2016; Morueta-Holme et al., 2016; Schäfer & Strimmer, 2005).

Interactor-as-predictor approaches use potentially interacting species as covariates in the model of the target species. In the case of occurrence data, this is:

$$\text{logit}(P(A)) = X\beta + B\alpha + \varepsilon,$$

Where α is the estimate of the effect of B on the probability of occurrence of A, P(A). B can be either occurrence or abundance. β = slope of the regression coefficients; ε = error term. Environmental and biotic effects are estimated simultaneously. Originally conceived for one target species (Leathwick & Austin, 2001), this can be extended to many species (Harris, 2016; Sierra & Stephens, 2012).

Ordination approaches, probably the most common approach to representing species associations, consider all species simultaneously and allow for species-specific environmental effects (Legendre & Legendre, 2013).

Spatial factor analysis (Thorson et al., 2015) describes species associations by regressing occurrence (or abundance) of multiple species against common but unknown (latent) covariates, which are modelled as random spatial fields.

Joint hierarchical multispecies models finally estimate species-specific environmental responses of all interacting species and simultaneously describe residual covariance in species occurrence (here for two species, A and B):

$$\text{logit}(P(A)) = X\beta_A + \varepsilon_A$$

$$\text{logit}(P(B)) = X\beta_B + \varepsilon_B$$

$$(\varepsilon_A, \varepsilon_B) \sim \text{MVN}((0, 0), \textstyle\sum)$$

$\sum$ represents the positive and negative covariances in species occurrence after accounting for environment. The approaches differ in how the variance–covariance matrix $\sum$ is computed. As the number of species (hence regressions) increases, $\sum$ grows quadratically in size, and applications for many species become computationally expensive. To overcome this problem, one can make the entries of $\sum$ a function of a few random auxiliary (latent) covariates (Warton et al., 2015). Spatial structural constraints can be put on these covariates (Ovaskainen, Roy, et al., 2016), similar to those in spatial factor analysis.

---

occurrence of another in a species-as-predictor approach (Box 2) have little to do with abundance-affecting interaction coefficients in a Lotka–Volterra model (see question 9). This is not necessarily a problem, as one would in any case expect that the local interaction structure does not translate one to one to the interactions at the macroscale, but such considerations are crucial when interpreting macroscale patterns as local interactions.

3.2 | Confounding biotic interactions and environment

3.2.1 | Question 4: Is the purported biotic interaction a surrogate for a missing predictor?

The presence or absence of a species, through habitat requirements, conveys information about the local environment (leading to the concept of 'indicator species'; Dufrêne & Legendre, 1997). Hence, using species as covariates in BI-SDMs potentially confounds their role as environmental indicators with that of biotic interactors. The results of Giannini et al. (2013) showed that species known not to interact are nevertheless good predictors for each other, suggesting that it was not the species, but the proxy for missing environmental variables that improved their models. Or, in the words of Clark et al. (2011; p. 1283): 'in the absence of detailed environmental information, species provide most information about one another'. The implication of this insight is

that the failure to include an important environmental predictor (in the correct functional form) for a target species will result in an apparent biotic interaction if other species can 'indicate' the missing environmental information (Bar-Massada, 2015; Blois et al., 2014). Species with similar habitat requirements will appear to interact positively, whereas species that have contrasting requirements will appear to interact negatively. To avoid this problem, both the distributional and the environmental data on which the BI-SDMs are based must be of high quality. Scepticism is warranted: (a) if the number of predictors is unreasonably low; (b) if the predictive quality of the single-species SDMs is much lower than SDMs for the same or related species in other studies; and (c) if substitute variables for environmental predictors, such as a spatial field or spatial eigenvector (Borcard & Legendre, 2002), can remove the estimate of the biotic interaction.

3.2.2 | Question 5: Are biotic interactions dependent on the environment?

Most BI-SDMs assume that biotic interactions are constant in space and time. It is more likely, however, that interaction strengths are modulated by the biotic and abiotic environment. Experimental studies showed that rates of biological interactions depend on environmental conditions: rates of herbivory are temperature dependent (Burnside, Erhardt, Hammond, & Brown, 2014); mycorrhizal symbioses change

TABLE 1 Overview of statistical modelling methods that have been used to explore biotic interactions

Method	Type	Comments	Spatial	Computing pace	Asymmetrical effects
Correlations	Correlation of co-occurrences (Ulrich, Almeida-Neto, & Gotelli, 2009)	No environment	No	Gallop	No
	Partial correlations (of residuals from separate GLMs controlling for environmental effects; Harris, 2016)	Conservative; removes all biotic effects collinear with environmental effects	Yes	Gallop	No
Interactor as predictor	Single target species (Leathwick & Austin, 2001)	The 'original' approach to including biotic interactions	Yes	Gallop	Yes
	Many independent GLMs (all other species as predictor; Harris, 2016)	Convergence issues owing to high collinearity	Yes	Gallop	Yes
Ordination	Canonical Correspondence Analysis (CCA), distance-based Redundancy Analysis (db-RDA) (e.g., Legendre & Legendre, 2013), multivariate regression trees (De'ath, 2002)	Same environmental response for all species	Yes	Gallop	No
	Community-level models (Maguire et al., 2016; Nieto-Lugilde et al., 2018)	Extension of GLM; no explicit modelling of interactions	No	Gallop	?
Spatial factor analysis	Spatial factor analysis (Thorson et al., 2015)	Includes spatial dimension	Yes	Trot	No
	Spatial dynamic factor analysis (Thorson et al., 2016)	Uses multiple time slices and allows for temporal variation in latent factors	Yes	Walk	No
Joint hierarchical multispecies model	Artificial neural network (Özesmi & Özesmi, 1999); mistnet (Harris, 2015)	Non-hierarchical, but same idea; uses back-propagation instead	No	Canter	Yes
	jSDM (Pollock et al., 2014)	Hierarchical modelling of environmental effects across species	No	Walk	No
	BayesComm (Golding & Purse, 2016); gjam (Clark, Nemergut, Seyednasrollah, Turner, & Zhang, 2017)	Independent environmental effects across species	No	Canter	No
	Spatial jSDM (Finley, Banerjee, & Carlin, 2007; Latimer, Banerjee, Sang, Mosher, & Silander, 2009)	Spatial prediction allows for large sample size (but for few species)	Yes	Walk	No
	Spatio-phylogenetic jSDM (Kaldhusdal et al., 2015)	Spatio-phylogenetic models	Yes	Walk	No
	Latent variable jSDMs (Warton et al., 2015)	Simplified fitting of covariance matrix through latent covariates	No	Trot	Yes
	Spatial latent variable jSDM (Ovaskainen, Roy, et al., 2016; Tikhonov et al., 2017)	Spatial structure in latent covariates. Interactions are changing along environmental gradients (Tikhonov et al., 2017, only)	Yes	Walk	Yes

GLM = generalized linear model; jSDM; joint species distribution model. *Note.* In contrast to Nieto-Lugilde et al. (2018), we focus on methods that allow, in some way, a quantification of species associations. 'Spatial' refers to whether space could be incorporated in some form. 'Computing speed' will depend on the size of the data set, so we indicate only whether it will be fast (gallop; within seconds), intermediate (canter to trot) or slow (walk; days to weeks). This assessment is based on our own experience with these methods and self-reported computation times. The last column refers to the ability of the method to allow for asymmetrical effects (i.e., the effect of A on B being different from that of B on A).

with resource availability (Cox, Barsoum, Lilleskov, & Bidartondo, 2010; Johnson, Wilson, Bowker, Wilson, & Miller, 2010); and plant–plant interactions can shift from competitive to facilitative along a stress gradient (Bertness & Callaway, 1994). Additionally, environmental change can modulate biotic interactions via phenological mismatches, such as for plant–pollinator interactions (Berg et al., 2010; Elzinga et al., 2007). In these conditions, the model must also allow for the interplay of environmental and biotic interactions (Clark et al., 2011), which has only recently started to be explored in BI-SDMs (Tikhonov, Abrego, Dunson, & Ovaskainen, 2017).

3.2.3 | Question 6: Did the inclusion of a biotic interactor improve the model?

A question that should also be analysed more frequently is, whether and by how much are predictions improved by including a biotic predictor? Recent approaches to partition the variation build on comparison of pure biotic, pure environmental, and joint environmental and biotic models (e.g., de Araújo, Marcondes-Machado, & Costa, 2014; Lobo, Castro, & Moreno, 2001; Meier et al., 2010). Biotic and abiotic predictors are separated, and the joint contribution of biotic and abiotic predictors to the explained deviance can be evaluated. This allows quantification of the magnitude of overlap (i.e., the joint contribution), providing an indication of how much we cannot uniquely attribute to environmental or biotic interaction. In itself, this is not evidence for a biotic interaction, but it helps us to put its potential importance into perspective.

3.2.4 | Question 7: Can phylogeny confound biotic interactions?

As a result of a phylogenetic signal in the ecological requirements of species, the distributions of closely related species should also show greater overlap than those of distantly related species. This phylogenetic signal should result in positive associations one could falsely interpret as evidence of biotic interactions. A growing number of studies in community ecology have documented that at broader spatial scales the phylogenetic relatedness of ecological communities is often structured along environmental gradients (but see Cavender-Bares, Kozak, Fine, & Kembel, 2009; Davies, Meiri, Barraclough, & Gittleman, 2007). For instance, the phylogenetic diversity of dragonfly assemblages decreases with decreasing temperature, so that closely related species tend to have more similar distributions than expected by chance, coinciding with a phylogenetic signal in thermal tolerances (Pinkert, Brandl, & Zeuss, 2017). Similar patterns are known for many different taxa (cf. Hortal et al., 2011; Svenning & Skov, 2004). However, phylogeny is only problematic if we ignore this potential source of bias in the associational signal. We suggest following the lead of Kaldhusdal, Brandl, Müller, Möst, and Hothorn (2015) on the integration of a method to detect a phylogenetic signal underlying co-occurrence patterns into SDMs, and Pollock, Morris, and Vesk (2012) for integration of measures of phylogenetic diversity into SDMs (but see de Bello et al., 2017).

3.3 | Dynamic populations: Dispersal and food web context

3.3.1 | Question 8: Are inferred biotic interactions confounded with dispersal limitation?

Migration lags following environmental change may lead to species associations not developing as expected from biotic interactions and species' niches alone (Gilman, Urban, Tewksbury, Gilchrist, & Holt, 2010). This is particularly true if the direction of migration is correlated with environmental gradients, making a differentiation of dispersal limitation and environment statistically impossible (Figure 2). For instance, ranges of neotropical marmosets (primates) are separated by the arms of river systems, although the estimated potential distributions are much broader (Ayres & Clutton-Brock, 1992). Such incomplete range filling might be interpreted as the outcome of competition among species, even if the reason is primarily a dispersal barrier (see also Morueta-Holme et al., 2016). Likewise, post-glacial recolonization, during which many species lag environmental change, left detectable historical signals underlying contemporary distribution patterns of taxa with very different biology, such as trees, amphibians, dung beetles and dragonflies (Araújo et al., 2008; Hortal et al., 2011; Pinkert, Dijkstra, et al., 2017; Svenning & Skov, 2004).

Potential dispersal constraints on species ranges could be explored by systematically comparing association coefficients (in Box 2) between species of different range sizes (small ranges may indicate dispersal constraints), between species with differing dispersal syndromes (Normand et al., 2011) or between species of different distance to glacial refugia; if one of these drivers matters, interactions may well be inferred incorrectly. An indirect implementation in a BI-SDM with only one time-step could be achieved by masking unreachable habitat, down-weighting probabilities with the distance to expert-knowledge maps or by including dispersal barriers as additional predictors for each species (Gilman et al., 2010; Merow, Wilson, & Jetz, 2017).

3.3.2 | Question 9: Are purported biotic interactions attributable to temporal dynamics?

Local environmental dynamics can lead to false estimation of both interaction strengths and environmental dependencies (Figure 3; Yackulic, Nichols, Reid, & Der, 2015). At one location, environmental
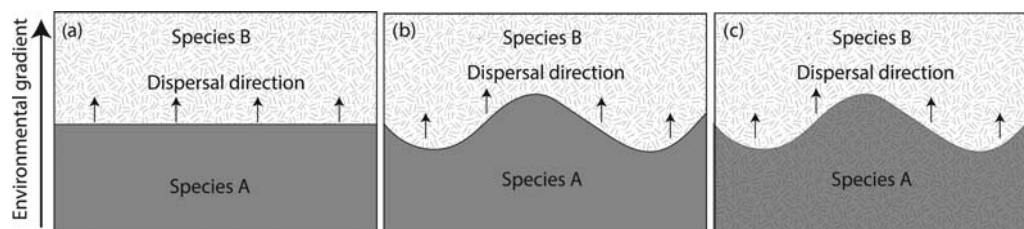

FIGURE 2 Negative biotic interactions may cause dispersal constraints that are difficult to differentiate from environmental filtering if dispersal follows an environmental gradient (e.g., postglacial recolonization). (a) Species A (grey shading) would be interpreted as occupying a different niche from B, although species A might eventually outcompete B (stippled) but is delayed in its dispersal by the presence of species B (priority effects). (b) This scenario could be detected in case (b), where areas with similar environmental conditions but differing species composition may indicate biotic interactions or dispersal constraints. (c) Species A may also be limited in its dispersal only by species B without negative effects on the occurrence of species B.

dynamics (e.g., disturbances) can drive transient shifts in community compositions (e.g., succession), with different sets of species co-occurring at different times (e.g., pioneer tree species at the beginning of forest succession). Large-scale BI-SDMs may easily attribute these co-occurrence patterns to direct biotic interactions.

To a certain extent, problems related to dynamic systems can be overcome by using time-lagged drivers or time since a triggering event (e.g., disturbance, start of colonization) as additional covariates. Species abundance data from several time slices, or ideally time series, can give more information about the dynamics and allow more reliable estimation of such interactions (Thorson et al., 2016). Another solution is to fit explicitly dynamic models or their static surrogates (see Box 3), with the challenge that the interaction and environmental niche parameters

---

BOX 3   Dynamic approaches to simulate distributions of interacting species

Many dynamic models explicitly consider species interactions, dispersal, or both (Cabral, Valente, & Hartig, 2017; Gravel, 2013; Snell et al., 2014; Svenning et al., 2014). Species interactions can be described at different levels of abstraction and complexity, ranging from individuals encountering and interacting directly or exploiting a shared resource (e.g., trees competing for light; Lischke, Zimmermann, Bolliger, Rickebusch, & Löffler, 2006) to highly aggregated interaction coefficients (Allesina & Tang, 2012). These interactions influence species and community dynamics (e.g., Cabral & Kreft, 2012; Lischke et al., 2006; Morin, Fahse, de Mazancourt, Scherer-Lorenzen, & Bugmann, 2014) and, combined with dispersal, species ranges (e.g., Epstein, Kaplan, Lischke, & Yu, 2007; Lischke, 2005). Yet, many of these models are complex and time consuming, and thus applying them for many species and fitting them to macroecological data remains challenging (but see Evans, Merow, Record, McMahon, & Enquist, 2016; Hartig et al., 2012). Hence, the current situation is that mechanistic models of biotic interactions could be ecologically more informative, but parameterization and validation can be difficult and may ultimately lead to problems similar to those of correlational approaches (Dormann et al., 2012).

A possible avenue to solve that dilemma is to create meta- or hybrid models that aggregate the known mechanisms of species interactions into a simpler structure that is easier to connect to large-scale data. An example is the possibility of directly determining the equilibria of dynamic Lotka–Volterra models (Lischke & Löffler, 2017). It incorporates the dynamic effects of species interactions and yields a very simple static model required for parameter estimation, complementing correlational approaches with dynamic add-ons (leading to hybrid models, such as KISSMIG; Nobis & Normand, 2014). Another approach is to start from statistical models but try to include essential dynamic mechanisms. An example is the hierarchical statistical state–space models that simultaneously estimate species distribution model and dynamic parameters (e.g., Pagel & Schurr, 2012), but only if they can be extended to many species (Evans et al., 2016).

---

of even these simple models are more difficult to estimate than those of a BI-SDM.

3.3.3 | Question 10: Which food web components should be considered in a BI-SDM?

Most species are positioned in food webs consisting of several trophic levels, which should create additional indirect effects. Sometimes, environmental change cascades through the food web to all trophic levels (Schmitz & Price, 2011). At other times, changes remain within a food web module (see Gilman et al., 2010; Tylianakis et al., 2008; Wardle, Barker, Yeates, Bonner, & Ghani, 2001).

A limited number of studies have attempted to integrate higher-order trophic interactions in macroecological SDMs (e.g., Pellissier et al., 2013; Rohr, Scherer, Kehrli, Mazza, & Bersier, 2010). This approach is referred to as a community-level model (CLM; Maguire, Nieto-Lugilde, Fitzpatrick, Williams, & Blois, 2015) or 'assemble and predict together' (Nieto-Lugilde et al., 2018). In communities with a large number of species, this leads to an inflation of covariates, which can be counteracted by including as predictors only the most abundant species (Le Roux, Pellissier, Wisz, & Luoto, 2014), or only those species that are part of the immediate food web of the focal species (Ovaskainen, Abrego, Halme, & Dunson, 2016; Pellissier et al., 2013). However, implementing these specifications requires considerable prior knowledge.

For food webs, a promising approach is the trophic interaction distribution model (TIDM), based on resource-selection theory (Trainor, Schmitz, Ivan, & Shenk, 2014). In any case, reflecting on the food web context forces the analyst to present a causal chain before interpreting a statistical pattern as biotic interaction. Here, the developmental stages of species (e.g., larvae versus adults), ontogeny and sex can also affect rates of biological interactions through niche and diet displacement (Berg et al., 2010).

4 | ONE AVENUE TOWARDS A SOLUTION: GENERALIZATIONS USING TRAITS

How should we move forwards? Technical advances may ameliorate some of the issues we have listed (also see Supporting Information Appendices S1 and S2), but we believe at a fundamental level that the problem of inferring biotic interactions will rise and fall with the availability of high-quality data and information from dedicated experiments. More specifically, we believe that the data currently used for BI-SDMs must be combined with other, ideally independent, data sources. One such data source that could presumably guide improvements in BI-SDMs is species traits (including surrogates, such as phylogenetic or trophic position). A growing number of studies show that species with similar functional traits interact with similar sets of species (e.g., Dehling, Jordano, Schaefer, Böhning-Gaese, & Schleuning, 2016; Garibaldi et al., 2015). Such trait matching of interacting species pairs is detectable across scales, from species interactions to co-variation of functional trait combinations at the regional scale (Dehling et al., 2014). Species traits can also be used to suggest competitive exclusion of species that are ecologically too similar (Davies et al., 2007). The
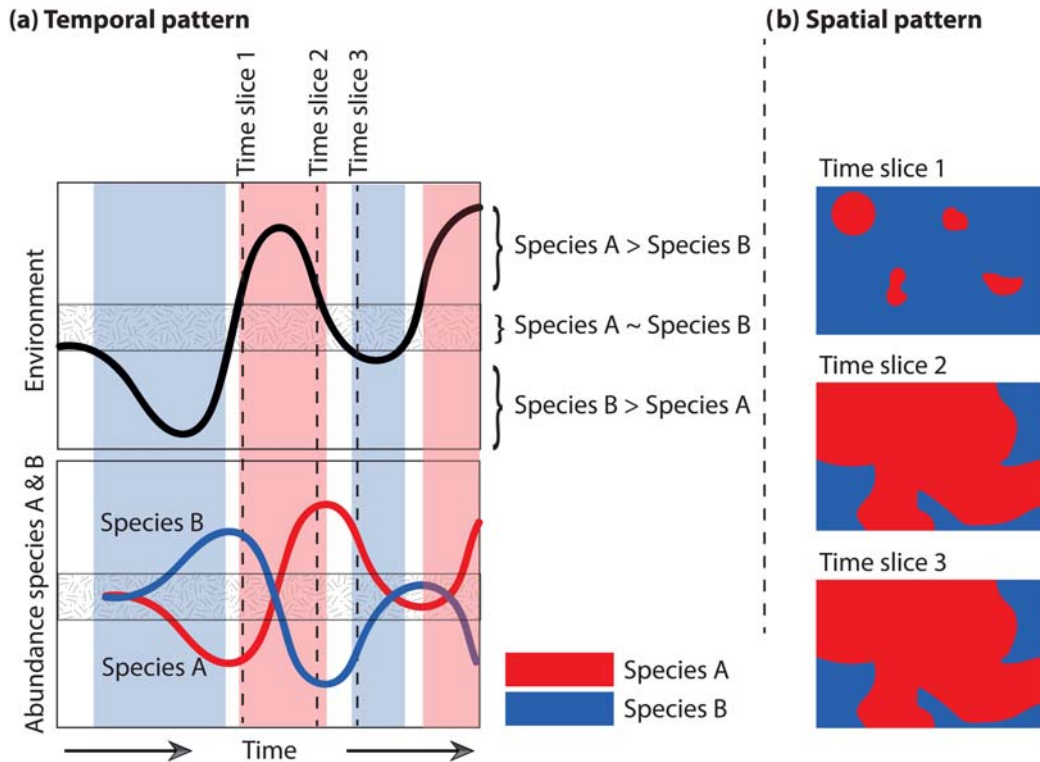
FIGURE 3 Temporal variation in the environment can, in combination with a time-delayed reaction of species abundance changes, enable long-term species coexistence despite strong negative biotic interactions (Chesson, 2000). Light red indicates time periods with environmental conditions where species A would outcompete species B, and light blue indicates time periods when species B would outcompete species A. Time delay in these negative species interactions (e.g., by priority effects) may delay species exclusion and thus allow coexistence in fluctuating environments [Colour figure can be viewed at wileyonlinelibrary.com]

incorporation of species traits into BI-SDMs (e.g., by using hierarchical modelling frameworks; Ovaskainen et al., 2017) would be most informative if the functional relevance of traits and trait surrogates is carefully tested before they are used to infer and quantify species interactions in ecological communities (Brousseau, Gravel, & Handa, 2017; Dehling et al., 2014; Ibanez, Lavorel, Puijalon, & Moretti, 2013).

Another, more challenging, obstacle is the prediction of interactions in communities (Gravel, Poisot, Albouy, Velez, & Mouillot, 2013; Pearse, Harris, Karban, & Sih, 2013), because we lack in-depth information on species interactions for most communities. By relating interactions to real and latent traits, it is conceivable to predict unobserved interactions (Naisbit, Rohr, Rossberg, Kehrli, & Bersier, 2012; Rohr, Naisbit, Mazza, & Bersier, 2016; Rohr et al., 2010), including those in novel communities in the future. Despite a growing interest in such predictive models (Garibaldi et al., 2015; Gravel, 2013; Kissling et al., 2012), there is not yet a fully established method to predict interactions between species pairs from traits (for attempts, see Dehling et al., 2016; Ibanez et al., 2013; Moretti et al., 2013; Morueta-Holme et al., 2016). The further refinement of this methodology for a large number of species and across spatial scales (Morales-Castilla et al., 2015; Morueta-Holme et al., 2016) offers a great opportunity for improving BI-SDMs. Nevertheless, this requires the same questions as outlined above to be asked and answered, but it might offer an avenue for a better mechanistic understanding of why macroecological signals of

species co-occurrences should be attributed to biotic interactions, or not.

5 | CONCLUSIONS

Macroecology is at a point of deciding between barging ahead, accepting by-and-large positive and negative associations in BI-SDMs as biotic interactions, or remaining sceptical until multiple lines of evidence confirm such patterns as genuinely representing biotic interactions. Our review and caveats reported in previous studies give many reasons why such scepticism seems warranted. We recommend careful and reflective interpretation of BI-SDMs, ideally complemented by and compared with experimental evidence, multiple independent data sets, models representing alternative ecological hypotheses, and consideration of the questions we have raised in this review. It seems easy to misinterpret the findings of a superficial analysis, and it is much harder to make a solid ecological case for it. Macroecologists should demand rigorous evidence for any community process inferred at large spatial scales and for methodological soundness of new approaches (particularly type I error rates). At present, BI-SDMs are perceived as promising, particularly for generating hypotheses about which interactions occur among species; yet asking, and answering, the questions we outlined here may lead to an improved ecological integration of biotic interactions in SDMs in the future.

ACKNOWLEDGMENTS

ORCID

Carsten F. Dormann http://orcid.org/0000-0002-9835-1794

Marco D. Moretti https://orcid.org/0000-0002-5845-3198

Stefan Pinkert http://orcid.org/0000-0002-8348-2337

Susanne I. Schmidt https://orcid.org/0000-0003-0051-6480

Christine S. Sheppard http://orcid.org/0000-0002-7847-2310

Manuel J. Steinbauer https://orcid.org/0000-0002-7142-9272

Dirk Zeuss http://orcid.org/0000-0001-6457-2866

Casper Kraan https://orcid.org/0000-0003-2062-6222

REFERENCES

Allesina, S., & Tang, S. (2012). Stability criteria for complex ecosystems. *Nature*, *483*, 205–208.

Araújo, M. B., Nogués-Bravo, D., Diniz-Filho, J. A. F., Haywood, A. M., Valdes, P. J., & Rahbek, C. (2008). Quaternary climate changes explain diversity among reptiles and amphibians. *Ecography*, *31*, 8–15.

Ayres, J. M., & Clutton-Brock, T. H. (1992). River boundaries and species range size in Amazonian primates. *The American Naturalist*, *140*, 531–537.

Bar-Massada, A. (2015). Complex relationships between species niches and environmental heterogeneity affect species co-occurrence patterns in modelled and real communities *Proceedings of the Royal Society B: Biological Sciences*, *282*, 20150927.

Bar-Massada, A., & Belmaker, J. (2017). Non-stationarity in the co-occurrence patterns of species across environmental gradients. *Journal of Ecology*, *105*, 391–399.

Begon, M., Townsend, C. R., & Harper, J. L. (2006). *Ecology – From individuals to ecosystems*. (4th ed.). Oxford, UK: Blackwell.

Berg, M. P., Kiers, E. T., Driessen, G., van der Heijden, M., Kooi, B. W., Kuenen, F., ... Ellers, J. (2010). Adapt or disperse: Understanding species persistence in a changing world. *Global Change Biology*, *16*, 587–598.

Bertness, M. D., & Callaway, R. M. (1994). Positive interactions in communities. *Trends in Ecology and Evolution*, *9*, 191–193.

Blois, J. L., Gotelli, N. J., Behrensmeyer, A. K., Faith, J. T., Lyons, S. K., Williams, J. W., ... Wing, S. (2014). A framework for evaluating the influence of climate, dispersal limitation, and biotic interactions using fossil pollen associations across the late Quaternary. *Ecography*, *37*, 1095–1108.

Borcard, D., & Legendre, P. (2002). All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling*, *153*, 51–68.

Brousseau, P. M., Gravel, D., & Handa, T. I. (2017). Trait-matching and phylogeny as predictors of predator–prey interactions involving ground beetles. *Functional Ecology*, *32*, 192–202.

Burnside, W. R., Erhardt, E. B., Hammond, S. T., & Brown, J. H. (2014). Rates of biotic interactions scale predictably with temperature despite variation. *Oikos*, *123*, 1449–1456.

Cabral, J. S., & Kreft, H. (2012). Linking ecological niche, community ecology and biogeography: Insights from a mechanistic niche model. *Journal of Biogeography*, *39*, 2212–2224.

Cabral, J. S., Valente, L., & Hartig, F. (2017). Mechanistic simulation models in macroecology and biogeography: State-of-art and prospects. *Ecography*, *40*, 267–280.

Cavender-Bares, J., Kozak, K. H., Fine, P. V. A., & Kembel, S. W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters*, *12*, 693–715.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology and Systematics*, *31*, 343–366.

Clark, J. S., Bell, D. M., Hersh, M. H., Kwit, M. C., Moran, E., Salk, C., ... Zhu, K. (2011). Individual-scale variation, species-scale differences: Inference needed to understand diversity. *Ecology Letters*, *14*, 1273–1287.

Clark, J. S., Gelfand, A. E., Woodall, C. W., & Zhu, K. (2014). More than the sum of the parts : Forest climate response from joint species distribution models. *Ecological Applications*, *24*, 990–999.

Clark, J. S., Nemergut, D., Seyednasrollah, B., Turner, P. J., & Zhang, S. (2017). Generalized joint attribute modeling for biodiversity analysis: Median-zero, multivariate, multifarious data. *Ecological Monographs*, *87*, 34–56.

Connell, J. H. (1990). Apparent versus "real" competition in plants. In J. B. Grace & D. Tilman (Eds.), *Perspectives in plant competition* (pp. 9–27). San Diego, CA: Academic Press.

Connor, E. F., & Simberloff, D. (1979). The assembly of species communities: Chance or competition? *The American Naturalist*, *60*, 1132–1140.

Cox, F., Barsoum, N., Lilleskov, E. A., & Bidartondo, M. I. (2010). Nitrogen availability is a primary determinant of conifer mycorrhizas across complex environmental gradients. *Ecology Letters*, *13*, 1103–1113.

Davies, T. J., Meiri, S., Barraclough, T. G., & Gittleman, J. L. (2007). Species co-existence and character divergence across carnivores. *Ecology Letters*, *10*, 146–152.

De'ath, G. (2002). Multivariate regression trees: A new technique for modeling species-environment relationships. *Ecology*, *83*, 1105–1117.

de Araújo, C. B., Marcondes-Machado, L. O., & Costa, G. C. (2014). The importance of biotic interactions in species distribution models: A test of the Eltonian noise hypothesis using parrots. *Journal of Biogeography*, *41*, 513–523.

de Bello, F., Šmilauer, P., Diniz-Filho, J. A. F., Carmona, C. P., Lososová, Z., Herben, T., & Götzenberger, L. (2017). Decoupling phylogenetic and functional diversity to reveal hidden signals in community assembly. *Methods in Ecology and Evolution*, *8*, 1200–1211.

Dehling, D. M., Jordano, P., Schaefer, H. M., Böhning-Gaese, K., & Schleuning, M. (2016). Morphology predicts species' functional roles and their degree of specialization in plant–frugivore interactions. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20152444.

Dehling, D. M., Töpfer, T., Schaefer, H. M., Jordano, P., Böhning-Gaese, K., & Schleuning, M. (2014). Functional relationships beyond species richness patterns: Trait matching in plant–bird mutualisms across scales. *Global Ecology and Biogeography*, *23*, 1085–1093.

Dormann, C. F., Schymanski, S. J., Cabral, J., Chuine, I., Graham, C. H., Hartig, F., ... Singer, A. (2012). Correlation and process in species

distribution models: Bridging a dichotomy. *Journal of Biogeography*, 39, 2119–2131.

Dufrêne, M., & Legendre, P. (1997). Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs*, 67, 345–366.

Elzinga, J. A., Atlan, A., Biere, A., Gigord, L., Weis, A. E., & Bernasconi, G. (2007). Time after time: Flowering phenology and biotic interactions. *Trends in Ecology and Evolution*, 22, 432–439.

Epstein, H. E., Kaplan, J. O., Lischke, H., & Yu, Q. (2007). Simulating future changes in arctic and subarctic vegetation. *Computing in Science and Engineering*, 9, 12–23.

Evans, M., Merow, C., Record, S., McMahon, S. M., & Enquist, B. J. (2016). Towards process-based range modeling of many species. *Trends in Ecology and Evolution*, 31, 860–871.

Faust, K., & Raes, J. (2012). Microbial interactions: From networks to models. *Nature Reviews Microbiology*, 10, 538–550.

Finley, A. O., Banerjee, S., & Carlin, B. P. (2007). spBayes: An R package for univariate and multivariate hierarchical point-referenced spatial models. *Journal of Statistical Software*, 19, 1–24.

Garibaldi, L. A., Bartomeus, I., Bommarco, R., Klein, A. M., Cunningham, S. A., Aizen, M. A., ... Woyciechowski, M. (2015). Trait matching of flower visitors and crops predicts fruit set better than trait diversity. *Journal of Applied Ecology*, 52, 1436–1444.

Giannini, T. C., Chapman, D. S., Saraiva, A. M., Alves-dos-Santos, I., & Biesmeijer, J. C. (2013). Improving species distribution models using biotic interactions: A case study of parasites, pollinators and plants. *Ecography*, 36, 649–656.

Gilman, S. E., Urban, M. C., Tewksbury, J. J., Gilchrist, G. W., & Holt, R. D. (2010). A framework for community interactions under climate change. *Trends in Ecology and Evolution*, 25, 325–331.

Golding, N., & Purse, B. V. (2016). Fast and flexible Bayesian species distribution modelling using Gaussian processes. *Methods in Ecology and Evolution*, 7, 598–608.

González-Salazar, C., Stephens, C. R., & Marquet, P. A. (2013). Comparing the relative contributions of biotic and abiotic factors as mediators of species' distributions. *Ecological Modelling*, 248, 57–70.

Gotelli, N. J. (2000). Null model analysis of species co-occurrence patterns. *Ecology*, 81, 2606–2621.

Gravel, D. (2013). *Assembly models*. Oxford, UK: Oxford University Press.

Gravel, D., Poisot, T., Albouy, C., Velez, L., & Mouillot, D. (2013). Inferring food web structure from predator–prey body size relationships. *Methods in Ecology and Evolution*, 4, 1083–1090.

Guillera-Arroita, G., Lahoz-Monfort, J., Elith, J., Gordon, A., Kujala, H., Lentini, P., ... Wintle, B. (2015). Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography*, 24, 276–292.

Harris, D. J. (2015). Generating realistic assemblages with a joint species distribution model. *Methods in Ecology and Evolution*, 6, 465–473.

Harris, D. J. (2016). Inferring species interactions from co-occurrence data with Markov networks. *Ecology*, 97, 3308–3314.

Hartig, F., Dyke, J. G., Hickler, T., Higgins, S., O'Hara, R. B., Scheiter, S., & Huth, A. (2012). Connecting dynamic vegetation models to data – An inverse perspective. *Journal of Biogeography*, 39, 2240–2252.

Hastings, A. (1987). Can competition be detected using species co-occurrence data? *Ecology*, 68, 117–123.

Heikkinen, R. K., Luoto, M., Virkkala, R., Pearson, R. G., & Körber, J.-H. (2007). Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, 16, 754–763.

Hortal, J., Diniz-Filho, J. A. F., Bini, L. M., Rodríguez, M. Á., Baselga, A., Nogués-Bravo, D., ... Lobo, J. M. (2011). Ice age climate, evolutionary constraints and diversity patterns of European dung beetles. *Ecology Letters*, 14, 741–748.

Ibanez, S., Lavorel, S., Puijalon, S., & Moretti, M. (2013). Herbivory mediated by coupling between biomechanical traits of plants and grasshoppers. *Functional Ecology*, 27, 479–489.

Ives, A. R., Dennis, B., Cottingham, K. L., & Carpenter, S. R. (2003). Estimating community stability and ecological interactions from time-series data. *Ecological Monographs*, 73, 301–330.

Johnson, N. C., Wilson, G. W. T., Bowker, M. A., Wilson, J. A., & Miller, R. M. (2010). Resource limitation is a driver of local adaptation in mycorrhizal symbioses. *Proceedings of the National Academy of Sciences USA*, 107, 2093–2098.

Kaldhusdal, A., Brandl, R., Müller, J., Möst, L., & Hothorn, T. (2015). Spatio-phylogenetic multispecies distribution models. *Methods in Ecology and Evolution*, 6, 187–197.

Kissling, W. D., Dormann, C. F., Groeneveld, J., Hickler, T., Kühn, I., McInerny, G. J., ... O'Hara, R. B. (2012). Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *Journal of Biogeography*, 39, 2163–2178.

Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., & Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, 29, 592–599.

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

Lawton, J. H., & Hassell, M. P. (1981). Asymmetrical competition in insects. *Nature*, 289, 793–795.

Le Roux, P. C., Pellissier, L., Wisz, M. S., & Luoto, M. (2014). Incorporating dominant species as proxies for biotic interactions strengthens plant community models. *Journal of Ecology*, 102, 767–775.

Leathwick, J. R., & Austin, M. P. (2001). Competitive interactions between tree species in New Zealand's old-growth indigenous forests. *Ecology*, 82, 2560–2573.

Legendre, P., & Legendre, L. (2013). *Numerical ecology* (3rd ed.). Amsterdam, The Netherlands: Elsevier Science.

Levine, J. M. (1999). Indirect facilitation: Evidence and predictions from a riparian community. *Ecology*, 80, 1762–1769.

Lischke, H. (2005). Modeling tree species migration in the Alps during the Holocene: What creates complexity? *Ecological Complexity*, 2, 159–174.

Lischke, H., & Löffler, T. (2017). Finding all multiple stable fixpoints of n-species Lotka–Volterra competition models. *Theoretical Population Biology*, 115, 24–34.

Lischke, H., Zimmermann, N. E., Bolliger, J., Rickebusch, S., & Löffler, T. J. (2006). TreeMig: A forest-landscape model for simulating spatio-temporal patterns from stand to landscape scale. *Ecological Modelling*, 199, 409–420.

Lobo, J. M., Castro, I., & Moreno, J. C. (2001). Spatial and environmental determinants of vascular plant species richness distribution in the Iberian Peninsula and Balearic Islands. *Biological Journal of the Linnean Society*, 73, 233–253.

Louthan, A. M., Doak, D. F., & Angert, A. L. (2006). Where and when do species interactions set range limits? *Trends in Ecology and Evolution*, 30, 780–792.

Maguire, K. C., Nieto-Lugilde, D., Blois, J. L., Fitzpatrick, M. C., Williams, J. W., Ferrier, S., & Lorenz, D. J. (2016). Controlled comparison of species- and community-level models across novel climates and

communities. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20152817.

Maguire, K. C., Nieto-Lugilde, D., Fitzpatrick, M. C., Williams, J. W., & Blois, J. L. (2015). Modeling species and community responses to past, present, and future episodes of climatic and ecological change. *Annual Review of Ecology, Evolution, and Systematics*, *46*, 343–368.

Meier, E. S., Kienast, F., Pearman, P. B., Svenning, J.-C., Thuiller, W., Araújo, M. B., ... Zimmermann, N. E. (2010). Biotic and abiotic variables show little redundancy in explaining tree species distributions. *Ecography*, *33*, 1038–1048.

Merow, C., Wilson, A. M., & Jetz, W. (2017). Integrating occurrence data and expert maps for improved species range predictions. *Global Ecology and Biogeography*, *26*, 243–258.

Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic interactions from proxies. *Trends in Ecology and Evolution*, *30*, 347–356.

Moretti, M., De Bello, F., Ibanez, S., Fontana, S., Pezzatti, G. B., Dziock, F., ... Lavorel, S. (2013). Linking traits between plants and invertebrate herbivores to track functional effects of land-use changes. *Journal of Vegetation Science*, *24*, 949–962.

Morin, X., Fahse, L., de Mazancourt, C., Scherer-Lorenzen, M., & Bugmann, H. (2014). Temporal stability in forest productivity increases with tree diversity due to asynchrony in species dynamics. *Ecology Letters*, *17*, 1526–1535.

Morueta-Holme, N., Blonder, B., Sandel, B., McGill, B. J., Peet, R. K., Ott, J. E., ... Svenning, J. C. (2016). A network approach for inferring species associations from co-occurrence data. *Ecography*, *39*, 1139–1150.

Mutshinda, C. M., O'Hara, R. B., & Woiwod, I. P. (2010). A multispecies perspective on ecological impacts of climatic forcing. *Journal of Animal Ecology*, *80*, 101–107.

Naisbit, R. E., Rohr, R. P., Rossberg, A. G., Kehrli, P., & Bersier, L.-F. (2012). Phylogeny versus body size as determinants of food web structure. *Proceedings of the Royal Society B: Biological Sciences*, *279*, 3291–3297.

Nieto-Lugilde, D., Maguire, K. C., Blois, J. L., Williams, J. W., & Fitzpatrick, M. C. (2018). Multiresponse algorithms for community-level modelling: Review of theory, applications, and comparison to species distribution models. *Methods in Ecology and Evolution*, *9*, 834–848.

Nobis, M. P., & Normand, S. (2014). KISSMig – A simple model for R to account for limited migration in analyses of species distributions. *Ecography*, *37*, 1282–1287.

Normand, S., Ricklefs, R. E., Skov, F., Bladt, J., Tackenberg, O., & Svenning, J.-C. (2011). Postglacial migration supplements climate in determining plant species ranges in Europe. *Proceedings of the Royal Society B: Biological Sciences*, *278*, 3644–3653.

Ovaskainen, O., Abrego, N., Halme, P., & Dunson, D. (2016). Using latent variable models to identify large networks of species-to-species associations at different spatial scales. *Methods in Ecology and Evolution*, *7*, 549–555.

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

Ovaskainen, O., Roy, D. B., Fox, R., & Anderson, B. J. (2016). Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models. *Methods in Ecology and Evolution*, *7*, 428–436.

Ovaskainen, O., Tikhonov, G., Norberg, A., Blanchet, F. G., 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.

Özesmi, S. L., & Özesmi, U. (1999). An artificial neural network approach to spatial habitat modelling with interspecific interaction. *Ecological Modelling*, *116*, 15–31.

Pagel, J., & Schurr, F. M. (2012). Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography*, *21*, 293–304.

Pearse, I. S., Harris, D. J., Karban, R., & Sih, A. (2013). Predicting novel herbivore–plant interactions. *Oikos*, *122*, 1554–1564.

Pellissier, L., Rohr, R. P., Ndiribe, C., Pradervand, J.-N., Salamin, N., Guisan, A., & Wisz, M. (2013). Combining food web and species distribution models for improved community projections. *Ecology and Evolution*, *3*, 4572–4583.

Peres-Neto, P. R. (1999). Conducting multiple ecological inferences revisited. *Marine Ecology Progress Series*, *176*, 303–306.

Pfister, C. A. (1995). Estimating competition coefficients from census data: A test with field manipulations of tidepool fishes. *The American Naturalist*, *146*, 271–291.

Pinkert, S., Brandl, R., & Zeuss, D. (2017). Colour lightness of dragonfly assemblages across North America and Europe, *Ecography*, *40*, 1110–1117.

Pinkert, S., Dijkstra, K.-D., Zeuss, D., Reudenbach, C., Brandl, R., & Hof, C. (2017). Evolutionary processes, dispersal limitation and climatic history shape current diversity patterns of European dragonflies, *Ecography*, *41*, 795–804.

Pollock, L. J., Morris, W. K., & Vesk, P. A. (2012). The role of functional traits in species distributions revealed through a hierarchical model. *Ecography*, *35*, 716–725.

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

Rees, M., Grubb, P. J., & Kelly, D. (1996). Quantifying the impact of competition and spatial heterogeneity on the structure and dynamics of a four-species guild of winter annuals. *The American Naturalist*, *147*, 1–32.

Rohr, R. P., Naisbit, R. E., Mazza, C., & Bersier, L.-F. (2016). *Matching–centrality* decomposition and the forecasting of new links in networks. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 20152702.

Rohr, R. P., Scherer, H., Kehrli, P., Mazza, C., & Bersier, L.-F. (2010). Modeling food webs: Exploring unexplained structure using latent traits. *The American Naturalist*, *176*, 170–177.

Roughgarden, J. (1983). Competition and theory in community ecology. *The American Naturalist*, *122*, 583–601.

Schäfer, J., & Strimmer, K. (2005). A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics. *Statistical Applications in Genetics and Molecular Biology*, *4*, 32.

Schmitz, O. J., & Price, J. R. (2011). Convergence of trophic interaction strengths in grassland food webs through metabolic scaling of herbivore biomass. *Journal of Animal Ecology*, *80*, 1330–1336.

Schooler, S. S., Salau, B., Julien, M. H., & Ives, A. R. (2011). Alternative stable states explain unpredictable biological control of *Salvinia molesta* in Kakadu. *Nature*, *470*, 86–89.

Sebastián-González, E., Dalsgaard, B., Sandel, B., & Guimarães, P. R. (2015). Macroecological trends in nestedness and modularity of seed-dispersal networks: Human impact matters. *Global Ecology and Biogeography*, *24*, 293–303.

Shang, Y., Sikorski, J., Bonkowski, M., Kandeler, E., Marhan, S., Boedding-haus, R. S., … Overmann, J. (2017). Inferring interactions in complex microbial communities from nucleotide sequence data and environmental parameters. *PLoS One*, *12*, e0173765.

Sierra, R., & Stephens, C. R. (2012). Exploratory analysis of the interrelations between co-located Boolean spatial features using network graphs. *International Journal of Geographical Information Science*, *26*, 441–468.

Snell, R. S., Huth, A., Nabel, J. E. M. S., Bocedi, G., Travis, J. M. J., Gravel, D., … Lischke, H. (2014). Using dynamic vegetation models to simulate plant range shifts. *Ecography*, *37*, 1184–1197.

Staniczenko, P. P. A., Sivasubramaniam, P., Suttle, K. B., & Pearson, R. G. (2017). Linking macroecology and community ecology: Refining predictions of species distributions using biotic interaction networks. *Ecology Letters*, *20*, 693–707.

Stephens, C. R., González-Salazar, C., Sánchez-Cordero, V., Becker, I., Rebollar-Tellez, E., Á., Rodríguez-Moreno, A., … Ramírez Martínez, M. M. (2016). Can you judge a disease host by the company it keeps? Predicting disease hosts and their relative importance: A case study for *Leishmaniasis*. *PLoS Neglected Tropical Diseases*, *10*, e0005004.

Svenning, J.-C., Gravel, D., Holt, R. D., Schurr, F. M., Thuiller, W., Münkemüller, T., … Normand, S. (2014). The influence of interspecific interactions on species range expansion rates. *Ecography*, *37*, 1198–1209.

Svenning, J.-C., & Skov, F. (2004). Limited filling of the potential range in European tree species. *Ecology Letters*, *7*, 565–573.

Thorson, J. T., Ianelli, J. N., Larsen, E. A., Ries, L., Scheuerell, M. D., Szu-walski, 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.

Thorson, J. T., Scheuerell, M. D., Shelton, A. O., See, K. E., Skaug, H. J., & Kristensen, K. (2015). Spatial factor analysis: A new tool for estimating joint species distributions and correlations in species range. *Methods in Ecology and Evolution*, *6*, 627–637.

Tikhonov, G., Abrego, N., Dunson, D., & Ovaskainen, O. (2017). Using joint species distribution models for evaluating how species-to-species associations depend on the environmental context. *Methods in Ecology and Evolution*, *8*, 443–452.

Trainor, A., Schmitz, O., Ivan, J., & Shenk, T. (2014). Enhancing species distribution modeling by characterizing predator–prey interactions. *Ecological Applications*, *24*, 204–216.

Tylianakis, J. M., Rand, T. A., Kahmen, A., Klein, A.-M., Buchmann, N., Perner, J., & Tscharntke, T. (2008). Resource heterogeneity moderates the biodiversity-function relationship in real world ecosystems. *PLoS Biology*, *6*, e122.

Ulrich, W., Almeida-Neto, M., & Gotelli, N. J. (2009). A consumer's guide to nestedness analysis. *Oikos*, *118*, 3–17.

Ulrich, W., Jabot, F., & Gotelli, N. J. (2017). Competitive interactions change the pattern of species co-occurrences under neutral dispersal. *Oikos*, *126*, 91–100.

Veech, J. A. (2012). A probabilistic model for analysing species co-occurrence. *Global Ecography and Briography*, *22*, 252–260.

Wardle, D. A., Barker, G. M., Yeates, G. W., Bonner, K. I., & Ghani, A. (2001). Introduced browsing mammals in New Zealand natural forests: Aboveground and belowground consequences. *Ecological Monographs*, *71*, 587–614.

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

Weiner, J. (1990). Asymmetric competition in plant populations. *Trends in Ecology and Evolution*, *5*, 360–364.

Williams, I. D., van der Meer, J., Dekker, R., Beukema, J. J., & Holmes, S. P. (2004). Exploring interactions among intertidal macrozoobenthos of the Dutch Wadden Sea using population growth models. *Journal of Sea Research*, *52*, 307–319.

Wisz, M. S., Pottier, J., Kissling, W. D., Pellissier, L., Lenoir, J., Damgaard, C. F., … Svenning, J.-C. (2013). The role of biotic interactions in shaping distributions and realised assemblages of species: Implications for species distribution modelling. *Biological Reviews*, *88*, 15–30.

Yackulic, C. B., Nichols, J. D., Reid, J., & Der, R. (2015). To predict the niche, model colonization and extinction. *Ecology*, *96*, 16–23.

Zhang, J., Kissling, W. D., & He, F. (2012). Local forest structure, climate and human disturbance determine regional distribution of boreal bird species richness in Alberta, Canada. *Journal of Biogeography*, *40*, 1131–1142.

BIOSKETCH

Most authors were participants of the workshop, 'Putting biotic interactions into correlative species distribution analyses: is it possible, and if so, how?', held at the Annual Meeting of the GfÖ (The Ecological Society of Germany, Austria and Switzerland) 2016 in Marburg with the aim of discussing the potential of SDMs to capture biotic interactions. The focus was not so much on the technical challenges, but rather on the ecological question, whether we really believe that what is estimated are indeed biotic interactions. Authors contributed complementary experience with statistical and process-based analyses of species distributions, from very different systems and types of interactions, enabling the wide review presented here.

SUPPORTING INFORMATION

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