

# ECOGRAPHY

## Forum

### Do joint species distribution models reliably detect interspecific interactions from co-occurrence data in homogenous environments?

Damaris Zurell, Laura J. Pollock and Wilfried Thuiller

*D. Zurell (http://orcid.org/0000-0002-4628-3558) (damaris.zurell@hu-berlin.de), Swiss Federal Research Inst. WSL, Dept Landscape Dynamics, Birmensdorf, Switzerland, and Humboldt-Univ. zu Berlin, Geography Dept, Berlin, Germany. – L. J. Pollock and W. Thuiller, Univ. Grenoble Alpes, Univ. Savoie-Mont Blanc, CNRS, Laboratoire d'Écologie Alpine (LECA), Grenoble, France.*

#### Ecography

41: 1812–1819, 2018

doi: 10.1111/ecog.03315

Subject Editor: Pablo A. Marquet

Editor-in-Chief: Miguel Araújo

Accepted 19 February 2018

Whether species interactions influence species response to environment and species ranges has always been a central question in ecology. Joint species distribution models (JSDMs) simultaneously model the species–environment relationships of multiple species and the residual correlation between these species. These residual correlations are assumed to depict whether species co-occur less or more often than expected by the modelled species–environment relationships, which could ultimately be attributed to species interactions, or hidden environmental information. Here, we propose to specifically test the capacity of JSDMs to detect species interactions from co-occurrence data, at different scales of data aggregation. Using a recently published point-process model, we simulated equilibrium co-occurrence patterns of species pairs by varying the strength and type of interactions (e.g. competition, predator–prey, mutualism) as well as the prevalence of the interacting species in homogeneous environments (assuming the environment does not influence the species responses and co-occurrence). Then, we fitted JSDMs without environmental predictors, and compared the estimated residual correlations against the known interaction coefficients. JSDMs detected competition and mutualism well, but failed with predator–prey interactions. For the latter, JSDMs predicted both negative and positive residual correlations for these kinds of interactions, depending on the prevalence of the interacting species. Interestingly, the estimated residual correlation was strongly influenced by species' prevalence and can thus not be translated to interaction strength. At increasingly coarser data resolution, the signals of negative and positive interactions became indiscernible by JSDMs, but – reassuringly – were rarely confounded. The underlying point-process model simulates the consequences rather than the mechanisms of interspecific interactions, and thus is better at corroborating rather than discrediting JSDMs. Nevertheless, our simple theoretical exercise pinpoints important limitations of JSDMs. In conclusion, we caution against interpreting residual correlations from JSDMs as interaction strength and against comparing these across different species and communities.

Keywords: residual correlation, scale dependence, species covariance



www.ecography.org

© 2018 The Authors. Ecography © 2018 Nordic Society Oikos

## Introduction

Determining the factors that influence species' ranges is a prerequisite for deriving reliable biodiversity scenarios (Urban et al. 2016). The importance of abiotic factors has long been recognized (Pulliam 2000, Chase and Leibold 2003, Soberón and Nakamura 2009), and correlative species distribution models (SDMs) that relate species' occurrence to prevailing environmental factors have become the most widely used tools in biogeography (Guisan and Zimmermann 2000, Guisan and Thuiller 2005, Elith and Leathwick 2009). However, it is now generally accepted that interspecific interactions likely influence species' distributions at larger spatial and temporal scales than previously thought (Blois et al. 2013, Svenning et al. 2014). Ignoring these interactions could thus hamper the ability of SDMs to predict current and future species ranges and communities (Leathwick and Austin 2001, Meier et al. 2010, Zurell et al. 2016a). Hence, over the last couple of years, several approaches have been discussed or proposed to incorporate interspecific interactions between multiple species into distribution models (Kissling et al. 2012, Pellissier et al. 2013, Thuiller et al. 2013, Wisz et al. 2013).

Aided by computational advances, methods have been introduced that combine species distribution modelling with co-occurrence analyses that originate from community ecology (Ovaskainen et al. 2017). These joint species distribution models (JSDMs) decompose species co-occurrence patterns into shared environmental responses and residual correlation (Ovaskainen et al. 2010, Clark et al. 2014, Pollock et al. 2014, Warton et al. 2015, Hui 2016, Ovaskainen et al. 2016). The latter correlation relates to the patterns of co-occurrence that are unexplained by the environmental information given to the model. The temptation is then to attribute this residual correlation to biotic interactions, such as competition and facilitation. Although this residual correlation can reflect intuitive ecological (Latimer et al. 2009, Ovaskainen et al. 2010) or evolutionary processes (Pollock et al. 2015), there could also be many non-biological explanations such as missing environmental variables or poor model fit. Theoretically, JSDMs should predict positive residual correlation between a pair of species if these co-occur more often than expected given the environment (or by chance in a homogenous environment) and predict negative residual correlation when the species co-occur less often than expected given the environment (or by chance in a homogenous environment).

To date, no study exists that explicitly tests JSDM's ability to identify the signal and the strength of species interactions and the effects of complicating factors such as species' prevalence and spatial scale. Also, the limits of JSDMs to detect asymmetric (positive-negative) interactions such as predator-prey relationships have not been tested. As JSDMs basically evaluate whether (residual) co-occurrence deviates from the null expectation (given by environmental overlap), they can per definition only capture symmetric interactions. However,

a recent study by Araújo and Rozenfeld (2014) showed that range overlaps of species pairs resulting from predator-prey relationships could resemble both co-occurrence patterns of competition (-/-) and mutualism (+/+). It is thus an important question whether JSDMs could possibly distinguish asymmetric interactions such as predator-prey relationships (+/-) from symmetric interactions (-/- and +/+).

Here, we build on the point-process model of Araújo and Rozenfeld (2014) that simulates co-occurrence of species pairs at steady state across all interspecific interaction types (-/-, +/-, +/+) and across all possible combinations of interspecific interaction strength ( $0 \leq I \leq 1$ ). We use this point-process model to generate co-occurrence data of species pairs in homogeneous environments, and then use these data to fit JSDMs. Other models could be used to simulate the distribution of interacting species, for example spatially explicit population models such as Lotka-Volterra-type models, coupled map-lattice models (Zurell et al. 2009) or individual-based models (Travis et al. 2005). However, by using data from the steady-state point-process model and from homogeneous environments, we ensure that species co-occurrence patterns are not confounded by other factors such as history, missing environmental variables or demographic stochasticity. We thus provide a simple proof-of-concept and ask whether JSDMs can detect the signal of interspecific interactions from co-occurrence data under idealised conditions, which we regard as prerequisite for applying these models in complex real-world applications. Specifically, we test whether JSDMs can detect different interspecific interactions ranging from negative to positive, how JSDMs cope with positive-negative interactions such as predator-prey relationships, and how the ability to detect interactions is influenced by the resolution of the data used for modelling (i.e. scale-dependence).

## Methods

### Simulating co-occurrence patterns

We used the point-process model developed by Araújo and Rozenfeld (2014) to simulate co-occurrence of species at steady state across all possible interaction types (+/+, +/-, -/-, +/0, -/0) and all possible combinations of interaction strengths ( $0 \leq |I_x| \leq 1$ ). If there is no interaction between species, then the expected probability of co-occurrence  $P(A \text{ and } B)$  of a species pair is simply given by the product of their prevalence  $P(A) \times P(B)$ . This corresponds to the null expectation of co-occurrence. With interspecific interaction present, the probability of co-occurrence is a function of species' prevalence and the strength of their interaction, which could be repulsive or attractive. Mutualistic interactions (+/+) will cause the species to co-occur more often than expected under the null model, whereas competitive interactions (-/-) will cause them to co-occur less often than expected. In case of predator-prey relationships (+/-), both positive

and negative interactions will cause deviation from the null model. Whether this results in higher or lower co-occurrence than expected by the null model, depends on both the relative strength of positive and negative interactions and on the prevalence of the species. Thus, the resulting co-occurrence (relative repulsion and attraction of species pairs) is deterministic while the spatial distribution of the species is stochastic. For a more detailed description of the model formulation, please see Araújo and Rozenfeld (2014). For simplicity, the environment is assumed homogenous and does not influence species' ranges and prevalence.

Here, we simulated co-occurrence patterns of two species A and B in homogeneous environments of  $100 \times 100$  cells for all combinations of potential interactions. Species' occurrence in space was random without any spatial autocorrelation. Araújo and Rozenfeld (2014) showed that the co-occurrence patterns and their scale dependence were comparable between spatially correlated and spatially uncorrelated landscapes. We varied the interaction coefficients  $I$  from  $-1$  to  $1$  in  $0.1$  increments. This resulted in 441 simulations per prevalence level. As an extension to previous results (Araújo and Rozenfeld 2014), we systematically varied prevalence levels. Specifically, in the first set of five scenarios, species A and B had equal prevalence of varying magnitude  $[0.1, 0.2, 0.3, 0.4, 0.5]$ . In a second set of two scenarios, species A and B had unequal prevalence  $[A=0.1$  and  $B=0.5$ ;  $A=0.2$  and  $B=0.4]$ .

For all scenarios, we calculated the co-occurrence index and the null expectation of co-occurrence, to evaluate whether species co-occurred more or less often than expected by chance. The co-occurrence index was defined as the number of cells where both species occur together divided by the total number of occupied cells. As the null probability of co-occurrence  $P(A$  and  $B)$  is given by the product of species' prevalence  $P(A) \times P(B)$ , the co-occurrence value

corresponding to the null expectation is given by  $(P(A) \times P(B)) / ((P(A) + P(B) - P(A) \times P(B)))$  where the dividend is the proportion of cells occupied by both species together and the divisor is the overall proportion of occupied cells. Hence, the exact co-occurrence value corresponding to the null expectation, where both species co-occur simply by chance, varies with prevalence (Fig. 1A–F).

To test for scale dependence in co-occurrence patterns and in the estimation of JSDMs, we aggregated the original landscape grid to two increasingly coarser resolutions by aggregating  $2 \times 2$  cells and  $4 \times 4$  cells, respectively. In these cases, species interactions (repulsion and attraction) are still modelled at the initial resolutions, but JSDMs are fitted at coarser resolution mimicking large scale and coarse resolution data such as atlas data (Gotelli et al. 2010, Zurell et al. 2016b), or continental and global datasets (McGill 2010, Jetz et al. 2012).

All simulations were carried out in R (R Core Team), and the code is provided in the Supplementary material Appendix 1.

### Estimating joint species distribution models

The spatial distributions of species A and B as simulated by the point-process model over the full biotic interaction space served as input for JSDMs. We thus assumed perfect detection of the species and ideal sample sizes, meaning that the entire space was sampled although we also tested for sample size effects (Supplementary material Appendix 1). We fitted JSDMs using the code provided by Pollock et al. (2014). These jointly estimate the occurrence probability of multiple species with a hierarchical multivariate probit regression, and the residual correlation between those species' presences by means of an unstructured covariance matrix. As we were mainly interested in the ability of JSDMs to detect

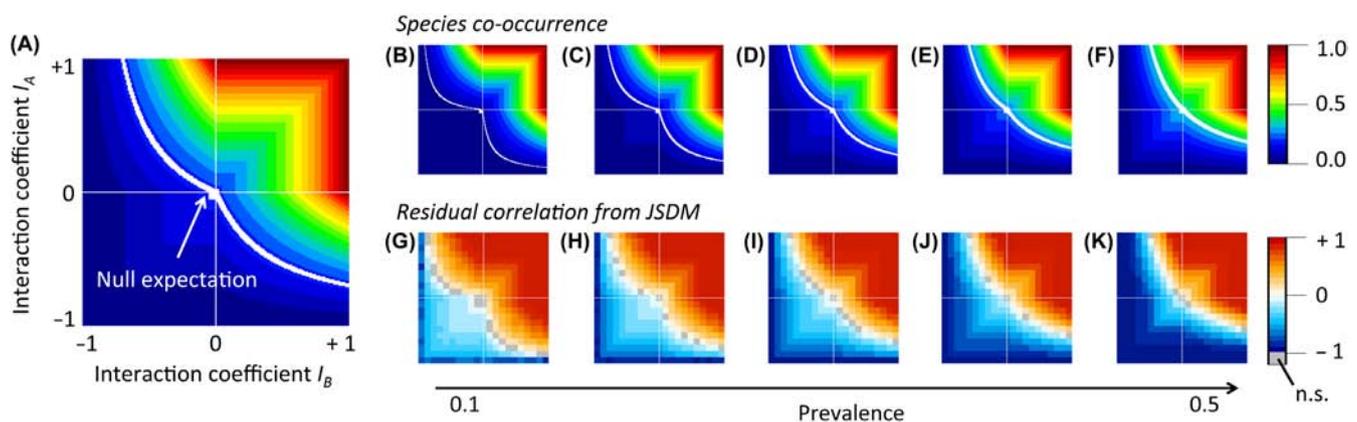


Figure 1. Expected co-occurrence (A–F) and residual correlation (estimated by JSDM; G–K) across the biotic interaction space of two species A and B and for different prevalence levels. (A) Species A and B can have negative to positive effects on each other. The co-occurrence index is defined as the number of cells where species A and B occur together divided by the total number of occupied cells. The null probability of co-occurrence  $P(A$  and  $B)$  is given by the product of species' prevalence  $P(A) \times P(B)$ , and the corresponding co-occurrence value is given by  $(P(A) \times P(B)) / ((P(A) + P(B) - P(A) \times P(B)))$  (white line). (B–F) Show the co-occurrence values obtained at different prevalence levels (0.1, 0.2, 0.3, 0.4, 0.5) with equal prevalence for both species. Again the null expectation of co-occurrence is shown in white. (G–K) show the corresponding residual correlations estimated by JSDMs, with grey areas indicating non-significant residual correlations.

underlying interactions, we assumed homogeneous environments and fitted the environmental response with an intercept only model that controls for prevalence. The JSDBMs were run with JAGS from within R (R Core Team) with 3 chains. As the simulated data were very simple and without noise (meaning no sampling error was included), for most combinations of interaction strengths ( $-1 < I < +1$ ) a comparably low number of 10 000 iterations with a burn-in of 5000 and a thinning rate of 20, were sufficient for reaching convergence with  $r_{hat}$  values below 1.1 for all parameters. For extreme interaction strengths (with an interaction coefficient  $I$  of  $-1$  or  $+1$ ), convergence was not achieved under 10 000 iterations and we subsequently tested stability of parameter values for increasing numbers of iterations. The residual correlation between species' presences was given by the mean of the posterior distribution and was judged as significant when the range between the 5 and 95% posterior quantiles did not include zero.

## Results

Different prevalence levels produced different co-occurrence patterns across the biotic interaction space (Fig. 1A–F). Co-occurrence was always lower than the null expectation for competitive interactions ( $-/-$ ) and for amensalism ( $-/0$ ) and always higher than the null expectation for mutualism ( $+/+$ ) and commensalism ( $+/0$ ). However, the exact asymmetric interaction coefficients ( $+/-$ ), at which co-occurrence was not different from the null expectation, strongly depended on prevalence (of the more common species of the species pair). When prevalence was low in both species, then co-occurrence was higher than the null expectation for most combinations of positive-negative ( $+/-$ ) interactions (Fig. 1B), while co-occurrence was mostly lower than the null expectation if at least one species in the predator–prey relationship was more common (Fig. 1F; Supplementary material Appendix 1 Fig. A1).

JSDBMs correctly assigned negative residual correlations to competitive interactions, and correctly assigned positive residual correlations to mutualistic interactions. More broadly, the difference between observed co-occurrence and the null expectation determined the residual correlation estimated by JSDBMs. Negative residual correlations were estimated when species co-occurred less often than the null expectation (meaning the co-occurrence value corresponding to the case when both species co-occur only by chance), and positive residual correlations when the species co-occurred more often than the null expectation. However, the magnitude of residual correlations varied across scenarios with different prevalence of the interacting species (Fig. 1G–K). For example, for two species A and B with an intermediate and symmetric negative interaction ( $I_A = I_B = -0.5$ ), JSDBMs estimated a residual correlation of approx.  $-0.19$  if both species had a prevalence of 0.1, and a residual correlation of approx.  $-0.71$  if both species had a prevalence of

0.5. These results were robust, as we obtained convergence with  $r_{hat}$  values below 1.1 for most combinations of interaction strengths except for the extreme interaction coefficients of  $[I] = 1.0$  (Supplementary material Appendix 1 Fig. A2). For these extreme cases, convergence was achieved for larger numbers of iterations, whereby the model parameters were stable across different numbers of iterations (Supplementary material Appendix 1 Fig. A3). Also, prevalence levels were correctly predicted in all cases. All these results correspond to ideal sampling without detection error and very large sample sizes. Reduced sample sizes will increase type II errors but the general patterns reported above remain unchanged (Supplementary material Appendix 1 Fig. A4).

JSDBMs were not able to separate predator–prey relationships from competitive or mutualistic interactions, and assigned both negative and positive residual correlations to such asymmetric ( $+/-$ ) interactions (Fig. 1). Here, residual correlations from JSDBMs exhibited the same sensitivity to prevalence as co-occurrence patterns, meaning that the exact asymmetric interaction coefficients ( $+/-$ ) that divided positive from negative residual correlations estimated by JSDBMs depended on prevalence (Fig. 1G–K; Supplementary material Appendix 1 Fig. A1).

Last, JSDBMs had difficulties at depicting the underlying co-variation between species at increasingly coarser resolution of the data (Fig. 2A–F). In line with Araújo and Rozenfeld (2014) co-occurrence patterns were highly scale dependent, which led to more insignificant residual correlation estimates in JSDBMs (meaning that the parameters credible interval between the 5 and 95% quantiles of the posterior distribution included zero; Fig. 2G–J). Thereby, chains mixed well and parameter convergence was well achieved. Generally, for competitive interactions, increasingly coarser data resolution led to lower estimates of absolute residual correlation than at the original resolution (at which the interaction outcome was being modelled). This scaling effect was much less pronounced for facilitative interactions (Fig. 2–3). Signals of interspecific interactions were lost most readily for more prevalent species (Fig. 3). The probability of confounding negative and positive residual correlations seemed to increase with decreasing prevalence and with increasingly coarser resolution of the data (Fig. 3). The insignificant residual correlations were only partially explained by the reduced sample sizes when aggregating data. When repeating the data aggregation for larger sample sizes, the type II errors were slightly smaller but the general patterns reported above remained unchanged (Supplementary material Appendix 1 Fig. A5).

## Discussion

In this paper, we propose to systematically test the capacity of JSDBMs to retrieve species interactions from co-occurrence patterns in the case of simple bi-partite interactions within a homogenous environment, and test whether this capacity was influenced by the resolution at which species were modelled.

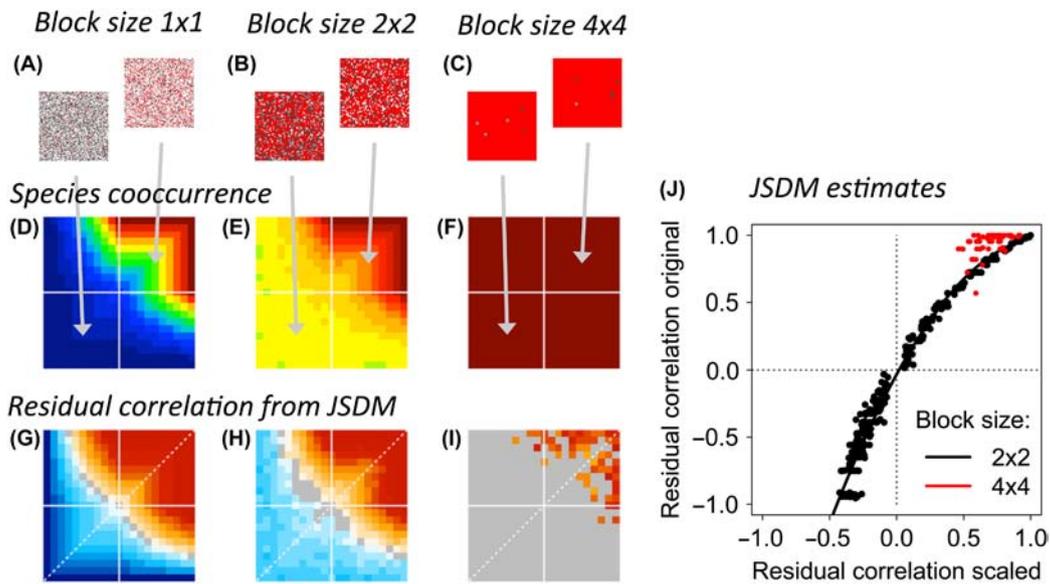


Figure 2. Scale dependence of co-occurrence patterns and corresponding residual correlations estimated by JSDMs. Both species A and B have prevalence of 0.3. (A–C) Show the spatial distribution of species at increasingly coarser resolution resulting from intermediate competition and intermediate mutualism (species A in black, species B in grey, species A and B co-occurring in red). (D–F) Show the resulting co-occurrences across the biotic interaction space of species A and B, and (G–I) the corresponding residual correlations estimated by JSDMs (please refer to Fig. 1 for legend description). (J) Shows the relationship between (significant) residual correlations estimated at the original resolution and at coarser resolution.

The key results are that 1) JSDMs reliably predict negative residual correlations for competitive interactions and positive residual correlations for mutualistic interactions. However, the magnitude of residual correlation does not directly relate to interaction strength because it is mostly driven by species prevalence. 2) Positive-negative interactions such as predator–prey relationships can be problematic for model interpretation because JSDMs cannot disentangle these from competitive and mutualistic interactions. 3) At increasingly coarser spatial resolution of the data, both the signals of negative and positive interactions become indiscernible by JSDMs, but reassuringly the signals rarely were confounded. Overall, our results have important implications for the

interpretation of JSDMs, and pinpoint important aspects that will need further model development. At the same time, we want to stress that these results were obtained under idealised conditions, and real-world applications will be complicated by even more factors. If JSDMs fail to detect interspecific interactions in very simple cases such as the one used here, there is no reason why they should work better at detecting interactions (occurring within local communities) in more complex real world applications.

First, we tested JSDMs' ability to detect negative and positive interactions from co-occurrence data. And indeed, JSDMs reliably detected competitive and mutualistic interactions in our simulated data, which is promising. However,

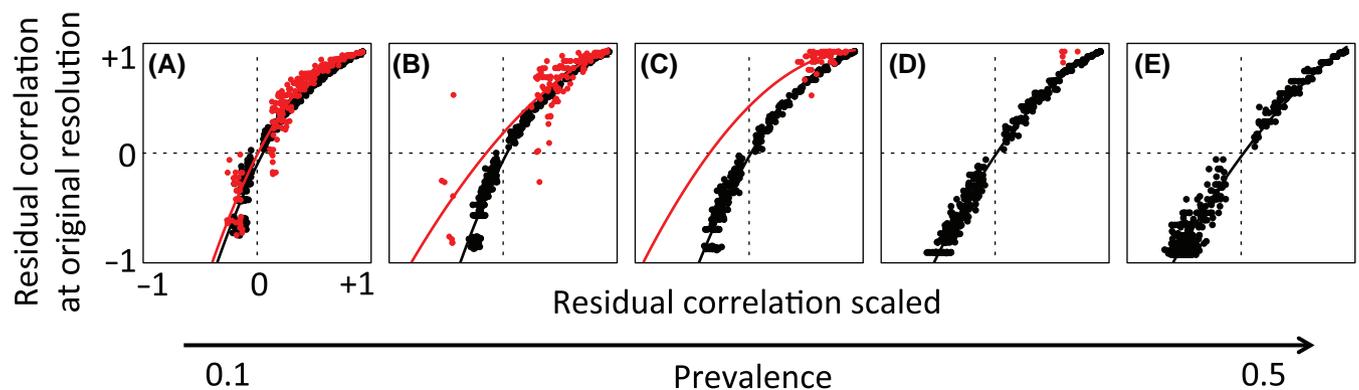


Figure 3. Scale dependence of JSDMs for different prevalence levels. Shown is the significant residual correlation estimated by JSDMs at the original scale against significant residual correlation estimated at coarser resolution (block size  $2 \times 2$  in black, block size  $4 \times 4$  in red) across biotic interaction space. Scale dependence of JSDMs increases with prevalence from left to right such that no significant residual correlations were estimated for the  $4 \times 4$  block size in (E).

interpretation of model residuals as interaction coefficients proved non-trivial. Importantly, prevalence strongly affected the magnitude of residual correlations estimated by JSDMs. This implies that residual correlations from JSDMs do not directly relate to nor can be interpreted as interaction strength, and also they cannot be compared across different species (with different prevalence levels). In real-world applications, this may be further confounded by the species' environmental response, detection bias, and by (direct and indirect) interactions between multiple species, which should be further tested.

Second, we asked how JSDMs deal with competitive and mutualistic interactions from asymmetric interactions such as predator–prey relationships. Per definition, JSDMs will model symmetric correlations and should not be able to detect asymmetric interactions. Nevertheless, it is important to understand whether and how residual correlations estimated from JSDMs differ between symmetric and asymmetric interactions especially when applying JSDMs to taxa from multiple trophic levels (such as birds and mammals). As has been discussed earlier (Araújo and Rozenfeld 2014, Cazelles et al. 2015), asymmetric positive-negative interactions can result in co-occurrence patterns that could be either higher or lower than expected by chance, meaning in both positive and negative net associations between pairs of species. Here, we showed that this also depends on interaction strength between species A and B and their prevalence. Specifically, predator–prey relationships are more likely to produce positive residual correlations when prevalence of both species is low. By contrast, when prevalence of at least one species is high, then predator–prey relationships are more likely to produce negative residual correlations in JSDMs. As a consequence, the problem that JSDMs are not able to tease apart predator–prey relationships from competitive or mutualistic interactions using co-occurrence data is exacerbated by the fact that asymmetric interactions could show up both as positive and negative residual correlations, which has been discussed previously but never been shown explicitly (Cazelles et al. 2015, Morales-Castilla et al. 2015, Warton et al. 2015). These results suggest that a useful next step would be to incorporate directional conditionality into future joint species distribution modelling, extending their capacity beyond pairwise symmetric correlations (Pollock et al. 2014, Warton et al. 2015, Hui 2016). Recent examples are beginning to consider more complex (direct and indirect) interactions for small datasets (Harris 2016, with up to 20 species). Schliep et al. (2017) recently introduced a method for incorporating temporal processes into JSDMs that allows inferring directionality from temporal co-abundance data. These advances would be extremely useful not only for distinguishing predator–prey relationships but also for detecting asymmetric interaction strengths between competing and facilitating species (Kissling et al. 2012, Harris 2016). Still, it remains to be tested in how far co-occurrence data will allow estimation of asymmetric residual correlation or under which circumstances co-abundance data

will be needed (Cazelles et al. 2015, Gallien et al. 2018). Also additional information could be used to distinguish competitive and mutualistic interactions from predator–prey relationships, for example trait information may aid simple plausibility checks (Morales-Castilla et al. 2015), and behavioural observations and manipulative experiments could help informing prior distributions (Harris 2016, Staniczenko et al. 2017).

Last, we tested for scale dependence of JSDM estimates by analysing the co-occurrence data at increasingly coarse resolution while the interactions took place at the original, fine spatial resolution. We found that with increasingly coarser resolution, both the signals of positive and negative interactions became indiscernible by JSDMs, whereby the signal of negative interactions was more sensitive to scale. This is in line with previous findings that co-occurrence patterns are highly scale-dependent (Araújo and Rozenfeld 2014) and also that the effects of local interspecific interactions vanish at coarser spatial scales (Thuiller et al. 2015). Thus, JSDMs are unlikely to give any insight on the potential interactions between species if the scale of the data does not match the process scale. At best, they might give better and more reliable models and the residual correlation might shed light on missing environmental variables and historical factors. Also, the species lists from monitoring schemes or atlas data are often at comparably coarse resolution, for example for many animals, such that the signal of interspecific interactions and sub-scale environmental heterogeneity can get easily confounded (Zurell et al. 2016b). When we apply JSDMs to vegetation plots, the analysis of the residual correlation in respect to interspecific interactions might thus be more meaningful since the interactions between species, for instance for light competition, might here be at the right scale (e.g. a  $10 \times 10$  m plot). Empirical analyses testing whether the residual correlations from JSDMs are related to known interactions between species might thus be of particular interest. Furthermore, our results showed that less prevalent species are less sensitive to mismatches between process scale and data scale. At the same time, at increasingly coarse resolution negative and positive interactions were confounded easier. Thus, we recommend interpreting JSDMs cautiously when modelling rare species.

We deliberately chose to use the steady-state point process model from Araújo and Rozenfeld (2014) to simulate spatial co-occurrence patterns. In contrast to spatially explicit population models (Kot 2001, Travis et al. 2005), this model does not explicitly simulate the spatial dynamics of interacting species, but directly generates non-random co-occurrence patterns that represent the spatial effects of interspecific interactions at equilibrium. The point-process model characterises the consequences rather than the mechanisms of interspecific interactions. Similarly, JSDMs are phenomenological and thus model the consequences of interspecific interactions on species co-occurrence (while accounting for environment). Therefore, we expect this set-up to corroborate rather than discredit JSDMs. This was mostly the case, except for

some important limitations of JSDMs when trained on co-occurrence data. Although this is a basic test, such proof-of-concept is still worthwhile because if a model fails in simple virtual examples, chances are that it will also fail in the real world (Zurell et al. 2010). In the future, JSDMs and their extensions (e.g. to incorporate directional conditionality) should be tested using more complex and spatially explicit population models.

In summary, our simulation study indicates that JSDMs might be reliable in inferring simple symmetric species interactions if the species are modelled at the resolution of the interactions. However, even in our idealised and error-free data, JSDMs could easily confound predator–prey relationships with either competitive or mutualistic interactions, they do not give an estimate of interactions strengths, they are not easily comparable across species, they are affected by prevalence and by scale. Thus, we should be very cautious when interpreting JSDM results as interactions (Clark et al. 2014, Pollock et al. 2014). Nevertheless, they have the ability of enhancing our analyses and allow us to form hypotheses about potential interspecific interactions when used cautiously (Ovaskainen et al. 2010). Our results provide a first test of potentially important or confounding factors in JSDM analyses. In the future, more efforts are needed to test JSDMs in more complicated settings, for example to test their ability to disentangle environmental response and residual correlation for multiple interacting species, for cases when species interact directly and indirectly, and for other confounding factors such as detection bias and missing environmental covariates. Thereby, further theoretical tests under controlled conditions could be useful, but we also emphasise the need for more empirical tests, for example comparing JSDM estimates against known interactions (e.g. from experiments). Hence, JSDMs constitute an important step forward for testing hypotheses of how interspecific interactions affect species distributions. Nevertheless, there are important limits to their capability of detecting interspecific interactions, for example for trophic and indirect interactions, highlighting the need for further model improvements in this respect.

*Funding* – DZ acknowledges funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. 624958, and from the Swiss National Science Foundation SNSF (grant: PZ00P3\_168136/1). WT acknowledges support from the European Research Council (ERC-2011-StG-281422-TEEMBIO). LJP acknowledges funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement no. 659422.

## References

- Araújo, M. B. and Rozenfeld, A. 2014. The geographic scaling of biotic interactions. – *Ecography* 37: 406–415.
- Blois, J. L. et al. 2013. Climate change and the past, present, and future of biotic interactions. – *Science* 341: 499–504.
- Cazelles, K. et al. 2015. A theory for species co-occurrence in interaction networks. – *Theor. Ecol.* 9: 39–48.
- Chase, J. M. and Leibold, M. A. 2003. Ecological niches: linking classical and contemporary approaches. – Univ. Chicago Press.
- Clark, J. S. et al. 2014. Competition–interaction landscapes for the joint response of forests to climate change. – *Global Change Biol.* 20: 1979–1991.
- 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.
- Gallien, L. et al. 2018. Frequency and intensity of facilitation reveal opposing patterns along a stress gradient. – *Ecol. Evol.* 8: 2171–2181.
- Gotelli, N. J. et al. 2010. Macroecological signals of species interactions in the Danish avifauna. – *Proc. Natl Acad. Sci. USA* 107: 5030–5035.
- Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in ecology. – *Ecol. Model.* 135: 147–186.
- Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. – *Ecol. Lett.* 8: 993–1009.
- Harris, D. J. 2016. Inferring species interactions from co-occurrence data with Markov networks. – *Ecology* 97: 3308–3314.
- Hui, F. K. C. 2016. Boral – Bayesian ordination and regression analysis of multivariate abundance data in R. – *Methods Ecol. Evol.* 7: 744–750.
- Jetz, W. et al. 2012. Integrating biodiversity distribution knowledge: toward a global map of life. – *Trends Ecol. Evol.* 27: 151–159.
- Kissling, W. D. et al. 2012. Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. – *J. Biogeogr.* 39: 2163–2178.
- Kot, M. 2001. Elements of mathematical ecology. – Cambridge Univ. Press.
- Latimer, A. M. et al. 2009. Hierarchical models facilitate spatial analysis of large data sets: a case study on invasive plant species in the northeastern United States. – *Ecol. Lett.* 12: 144–154.
- Leathwick, J. R. and Austin, M. P. 2001. Competitive interactions between tree species in New Zealand's old-growth indigenous forests. – *Ecology* 82: 2560–2573.
- McGill, B. 2010. Matters of scale. – *Science* 328: 575–576.
- Meier, E. S. et al. 2010. Biotic and abiotic variables show little redundancy in explaining tree species distributions. – *Ecography* 33: 1038–1048.
- Morales-Castilla, I. et al. 2015. Inferring biotic interactions from proxies. – *Trends Ecol. Evol.* 30: 347–356.
- Ovaskainen, O. et al. 2010. Modeling species co-occurrence by multivariate logistic regression generates new hypotheses on fungal interactions. – *Ecology* 91: 2514–2521.
- Ovaskainen, O. et al. 2016. Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models. – *Methods Ecol. Evol.* 7: 428–436.
- Ovaskainen, O. et al. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. – *Ecol. Lett.* 20: 561–576.
- Pellissier, L. et al. 2013. Combining food web and species distribution models for improved community projections. – *Ecol. Evol.* 3: 4572–4583.
- Pollock, L. J. et al. 2014. Understanding co-occurrence by modelling species simultaneously with a joint species distribution model (JSDM). – *Methods Ecol. Evol.* 5: 397–406.
- Pollock, L. J. et al. 2015. The roles of ecological and evolutionary processes in plant community assembly: the environment,

- hybridization, and introgression influence co-occurrence of *Eucalyptus*. – *Am. Nat.* 185: 784–796.
- Pulliam, H. R. 2000. On the relationship between niche and distribution. – *Ecol. Lett.* 3: 349–361.
- Schliep, E. M. et al. 2017. Joint species distribution modelling for spatio-temporal occurrence and ordinal abundance data. – *Global Ecol. Biogeogr.* 27: 142–155.
- Soberón, J. and Nakamura, M. 2009. Niches and distributional areas: concepts, methods, and assumptions. – *Proc. Natl Acad. Sci. USA* 106: 19644–19650.
- Staniczenko, P. P. A. et al. 2017. Linking macroecology and community ecology: refining predictions of species distributions using biotic interaction networks. – *Ecol. Lett.* 20: 693–707.
- Svenning, J.-C. et al. 2014. The influence of interspecific interactions on species range expansion rates. – *Ecography* 37: 1198–1209.
- Thuiller, W. et al. 2013. A road map for integrating eco-evolutionary processes into biodiversity models. – *Ecol. Lett.* 16: 94–105.
- Thuiller, W. et al. 2015. From species distributions to meta-communities. – *Ecol. Lett.* 18: 1321–1328.
- Travis, J. M. J. et al. 2005. The interplay of positive and negative species interactions across an environmental gradient: insights from an individual-based simulation model. – *Biol. Lett.* 1: 5–8.
- Urban, M. C. et al. 2016. Improving the forecast for biodiversity under climate change. – *Science* 353: aad8466.
- Warton, D. I. et al. 2015. So many variables: joint modeling in community ecology. – *Trends Ecol. Evol.* 30: 766–779.
- Wisz, M. S. et al. 2013. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. – *Biol. Rev.* 88: 15–30.
- Zurell, D. et al. 2009. Static species distribution models in dynamically changing systems: how good can predictions really be? – *Ecography* 32: 733–744.
- Zurell, D. et al. 2010. The virtual ecologist approach: simulating data and observers. – *Oikos* 119: 622–635.
- Zurell, D. et al. 2016a. Benchmarking novel approaches for modelling species range dynamics. – *Global Change Biol.* 22: 2651–2664.
- Zurell, D. et al. 2016b. Effects of functional traits on the prediction accuracy of species richness models. – *Divers. Distrib.* 22: 905–917.

Supplementary material (Appendix ECOG-03315 at < [www.ecography.org/appendix/ecog-03315](http://www.ecography.org/appendix/ecog-03315) >). Appendix 1.