1   **Article type**: Forum paper

2

3   **Title:** Do joint species distribution models reliably detect interspecific interactions from

4   co-occurrence data in homogenous environments?

5

6   **Authors:** Damaris Zurell[1]*, Laura J. Pollock[2], Wilfried Thuiller[2]

7   **Affiliations:**

8   [1] Swiss Federal Research Institute WSL, Dept. Landscape Dynamics, Birmensdorf,

9   Switzerland

10   [2] Univ. Grenoble Alpes, CNRS, Laboratoire d'Écologie Alpine (LECA), F-38000 Grenoble,

11   France

12

13   * Corresponding author: damaris.zurell@wsl.ch, ORCID: 0000-0002-4628-3558

14

15

16

17 **Abstract:**

18 Whether species interactions influence species response to environment and species

19 ranges has always been a central question in Ecology. Joint species distribution models

20 (JSDMs) simultaneously model the species-environment relationships of multiple

21 species and the residual correlation between these species. These residual correlations

22 are assumed to depict whether species co-occur less or more often than expected by the

23 modelled species-environment relationships, which could ultimately be attributed to

24 species interactions, or hidden environmental information. Here, we propose to

25 specifically test the capacity of JSDMs to detect species interactions, at different

26 aggregation scales of data. Using a recently published point-process model, we

27 simulated equilibrium co-occurrence patterns of species pairs by varying the strength

28 and type of interactions (e.g. competition, predator-prey, mutualism) as well as the

29 prevalence of the interacting species in homogeneous environments (assuming the

30 environment does not influence the species responses and co-occurrence). Then, we

31 fitted JSDMs without environmental predictors, and compared the estimated residual

32 correlations against the known interaction coefficients. JSDMs detected competition and

33 mutualism well, but failed with predator-prey interactions. For the latter, JSDMs

34 predicted both negative and positive residual correlations for these kinds of

35 interactions, depending on the prevalence of the interacting species. Interestingly, the

36 degree of residual correlation estimated by JSDMs was strongly influenced by species'

37 prevalence and can thus not be translated to interaction strength. At increasingly

38 coarser data resolution, the signals of negative and positive interactions became

39 indiscernible by JSDMs, but – reassuringly –were rarely confounded. Our simple

40 theoretical exercise demonstrates that when species are modelled at the scale at which

41   species interactions occur, JSDMs are able to depict symmetric interactions (+/+, -/-) but

42   fail for asymmetric interactions between trophic levels. However, when JSDMs are

43   applied at coarse resolutions, they are unlikely to give insights on the underlying biotic

44   interaction mechanisms.

45

46

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 (Chase and Leibold 2003, Pulliam 2000, 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 (Elith and Leathwick 2009, Guisan and Thuiller 2005, Guisan and

Zimmermann 2000). 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 (Clark, et al. 2014, Hui 2016, Ovaskainen, et al. 2010, Ovaskainen, et

al. 2016, Pollock, et al. 2014, Warton, et al. 2015). The latter correlation relates to the

5

69   patterns of co-occurrence that are unexplained by the environmental information given

70   to the model. The temptation is then to attribute this residual correlation to biotic

71   interactions, such as competition and facilitation. Although this residual correlation can

72   reflect intuitive ecological (Latimer, et al. 2009, Ovaskainen, et al. 2010) or evolutionary

73   processes (Pollock, et al. 2015), there could also be many non-biological explanations

74   such as missing environmental variables or poor model fit. Theoretically, JSDMs should

75   predict positive residual correlation between a pair of species if these co-occur more

76   often than expected given the environment (or by chance in a homogenous

77   environment) and predict negative residual correlation when the species co-occur less

78   often than expected given the environment (or by chance in a homogenous

79   environment). However, to date, no study exists that explicitly tests JSDM's ability to

80   identify the signal and the strength of species interactions and the effects of

81   complicating factors such as species' prevalence and spatial scale. Also, the limits of

82   JSDMs to detect asymmetric (positive-negative) interactions such as predator-prey

83   relationships have not been tested. For example, a recent study by Araújo and Rozenfeld

84   (2014) showed that range overlaps of species pairs resulting from predator-prey

85   relationships could resemble both co-occurrence patterns of competition (-/-) and

86   mutualism (+/+). The authors also showed that at increasingly coarser resolution, the

87   signal of negative interactions gets lost earlier than the signal of positive interactions.

88   Here, we build on the point-process model of Araújo and Rozenfeld (2014), and estimate

89   JSDMs from the co-occurrence patterns resulting from different types of interspecific

90   interactions (-/-, +/-, +/+) in homogeneous environments and from different species'

91   prevalence levels. That way no other reason like history or missing environmental

92   variables could explain the residual patterns. With that, we aim to answer the questions

93   of whether JSDMs can detect different interspecific interactions ranging from negative to

94   positive, how JSDMs cope with positive-negative interactions such as predator-prey

95   relationships, and how the ability to detect interactions is influenced by the resolution of

96   the data used for modelling (i.e. scale-dependence).

97   **Methods**

98   **Simulating co-occurrence patterns**

99   We used the point-process model developed by Araújo and Rozenfeld (2014) to simulate

100   co-occurrence of species at steady state across all possible interaction types (+/+, +/-, -

101   /-, +/0, -/0) and all possible combinations of interaction strengths ($0 \leq |I_x| \leq 1$). If there

102   is no interaction between species, then the expected probability of co-occurrence P(A

103   and B) of a species pair is simply given by the product of their prevalence P(A) x P(B).

104   This corresponds to the null expectation of co-occurrence. With interspecific interaction

105   present, the probability of co-occurrence is a function of species' prevalence and the

106   strength of their interaction, which could be repulsive or attractive. Mutualistic

107   interactions (+/+) will cause the species to co-occur more often than expected under the

108   null model, whereas competitive interactions (-/-) will cause them to co-occur less often

109   than expected. In case of predator-prey relationships (+/-), both positive and negative

110   interactions will cause deviation from the null model. Whether this results in higher or

111   lower co-occurrence than expected by the null model, depends on both the relative

112   strength of positive and negative interactions and on the prevalence of the species. For a

113   more detailed description of the model formulation, please see Araújo and Rozenfeld

114   (2014). For simplicity, the environment is assumed homogenous and does not influence

115   species' ranges and prevalence.

116   Here, we simulated co-occurrence patterns of two species A and B in homogeneous

117   environments of 100x100 cells for all combinations of potential interactions. Species'

118   occurrence in space was random without any spatial autocorrelation. Araújo and

119   Rozenfeld (2014) showed that the co-occurrence patterns and their scale dependence

120   were comparable between spatially correlated and spatially uncorrelated landscapes.

121   We varied the interaction coefficients I from -1 to 1 in 0.1 increments. This resulted in

122   441 simulations per prevalence level. As an extension to previous results (Araújo and

123   Rozenfeld 2014), we systematically varied prevalence levels. Specifically, in the first set

124   of five scenarios, species A and B had equal prevalence of varying magnitude [0.1, 0.2,

125   0.3, 0.4, 0.5]. In a second set of two scenarios, species A and B had unequal prevalence

126   [A=0.1 and B=0.5; A=0.2 and B=0.4].

127   To test for scale dependence in co-occurrence patterns and in the estimation of JSDMs,

128   we aggregated the original landscape grid to two increasingly coarser resolutions by

129   aggregating 2x2 cells and 4x4 cells, respectively. In these cases, species interactions still

130   happen at the initial resolutions, but JSDMs are fitted at coarser resolution mimicking

131   large scale and coarse resolution data such as atlas data (Gotelli, et al. 2010, Zurell, et al.

132   2016b), or continental and global datasets (Jetz, et al. 2012, McGill 2010).

133   All simulations were carried out in R (R Core Team 2015), and the code is provided in

134   the Supplementary Material Appendix 1.

135   **Estimating joint species distribution models**

136   The spatial distributions of species A and B as simulated by the point-process model

137   over the full biotic interaction space served as input for JSDMs. We fitted JSDMs using

138   the code provided by Pollock, et al. (2014). These jointly estimate the occurrence

139 probability of multiple species with a hierarchical multivariate probit regression, and

140 the residual correlation between those species' presences by means of an unstructured

141 covariance matrix. As we were mainly interested in the ability of JSDMs to detect

142 underlying interactions, we assumed homogeneous environments and fitted the

143 environmental response with an intercept only model that controls for prevalence. The

144 JSDMs were run with JAGS from within R (R Core Team 2015) with 3 chains. As the

145 simulated data were very simple and without noise, for most combinations of

146 interaction strengths ($-1 < I < +1$) a comparably low number of 10000 iterations with a

147 burn-in of 5000 and a thinning rate of 20, were sufficient for reaching convergence with

148 rhat values below 1.1 for all parameters. For extreme interactions strengths (with an

149 interaction coefficient I of -1 or +1), convergence was not achieved under 10000

150 iterations and we subsequently tested stability of parameters values for increasing

151 numbers of iterations. The residual correlation between species' presences was given by

152 the mean of the posterior distribution and was judged as significant when the range

153 between the 5% and 95% posterior quantiles did not include zero.

Results

155 Different prevalence levels produced different co-occurrence patterns across the biotic

156 interaction space (Fig. 1a-f). The co-occurrence index was defined as the number of cells

157 where both species occur together divided by the total number of occupied cells. The

158 null probability of co-occurrence P(A and B) is given by the product of species'

159 prevalence P(A) x P(B), and thus the co-occurrence value corresponding to the null

160 expectation is given by (P(A) x P(B)) / ((P(A) + P(B) – P(A) x P(B)) where the dividend

161 is the proportion of cells occupied by both species together and the divisor is the overall

162 proportion of occupied cells. Hence, the exact co-occurrence value corresponding to the

163   null expectation, where both species co-occur simply by chance, varied with prevalence

164   (Fig. 1a-f). Co-occurrence was always lower than the null expectation for competitive

165   interactions (-/-) and for amensalism (-/0) and always higher than the null expectation

166   for mutualism (+/+) and commensalism (+/0). However, the exact asymmetric

167   interaction coefficients (+/-), at which co-occurrence was not different from the null

168   expectation, strongly depended on prevalence (of the more common species of the

169   species pair). When prevalence was low in both species, then co-occurrence was higher

170   than the null expectation for most combinations of positive-negative (+/-) interactions

171   (Fig. 1b), while co-occurrence was mostly lower than the null expectation if at least one

172   species in the predator-prey relationship was more common (Fig. 1f; Supplementary

173   material Appendix 1, Fig. A1).

174   JSDMs correctly assigned negative residual correlations to competitive interactions, and

175   correctly assigned positive residual correlations to mutualistic interactions. Specifically,

176   the null expectation of co-occurrence determined the residual correlation estimated by

177   JSDMs. Negative residual correlations were estimated when species co-occurred less

178   often than the null expectation (meaning the co-occurrence value corresponding to the

179   case when both species co-occur only by chance), and positive residual correlations

180   when the species co-occurred more often than the null expectation. However, the

181   magnitude of residual correlations varied across scenarios with different prevalence of

182   the interacting species (Fig. 1g-k). For example, for two species A and B with an

183   intermediate and symmetric negative interaction ($I_A = I_B = -0.5$), JSDMs estimated a

184   residual correlation of approx. -0.19 if both species had a prevalence of 0.1, and a

185   residual correlation of approx. -0.71 if both species had a prevalence of 0.5. These

186   results were robust, as we obtained convergence with rhat values below 1.1 for most

10

187   combinations of interaction strengths except for the extreme interaction coefficients of

188 |I|=1.0 (Fig. A2). For these extreme cases, convergence was achieved for larger numbers

189   of iterations, whereby the model parameters were stable across different numbers of

190   iterations (Fig. A3). Also, prevalence levels were correctly predicted in all cases.

191   JSDMs were not able to separate predator-prey relationships from competitive or

192   mutualistic interactions, and assigned both negative and positive residual correlations

193   to such asymmetric (+/-) interactions (Fig. 1). Here, residual correlations from JSDMs

194   exhibited the same sensitivity to prevalence as co-occurrence patterns, meaning that the

195   exact asymmetric interactions coefficients (+/-) that divided positive from negative

196   residual correlations estimated by JSDMs depended on prevalence (Fig. 1g-k; Fig. A1).

197   Last, JSDMs had difficulties at depicting the underlying co-variation between species at

198   increasingly coarser resolution of the data (Fig. 2a-f). As shown by Araújo and Rozenfeld

199   (2014), co-occurrence patterns were scale-dependent and especially negative

200   interactions became indiscernible at increasingly coarser resolution of the data. This

201   scale dependency of co-occurrence patterns led to more insignificant residual

202   correlation estimates in JSDMs (meaning that the parameters credible interval between

203   the 5% and 95% quantiles of the posterior distribution included zero; Fig. 2 g-j).

204   Generally, chains mixed well and parameter convergence was well achieved. Thereby,

205   more prevalent species were more sensitive to scaling (Fig. 3, Fig. A4-A10).

206   Furthermore, for competitive interactions, increasingly coarser resolution of the data

207   led to lower estimates of absolute residual correlation than at the original resolution

208   (where the interaction takes place) while this effect of smaller estimates was much less

209   pronounced for facilitative interactions (Figs. 2-3). For intermediate prevalence levels,

210   negative and positive correlations between species only rarely got confounded (Fig. 2j).

211   However, for rare species, the probability of confounding negative and positive residual

212   correlations seemed to increase at increasingly coarser resolution of the data (Fig. 3a-b).

Discussion

214   In this paper, we propose to systematically test the capacity of JSDMs to retrieve species

215   interactions from co-occurrence patterns in the case of simple bi-partite interactions

216   within a homogenous environment, and test whether this capacity was influenced by the

217   resolution at which species were modelled. The key results are that (i) JSDMs reliably

218   predict negative residual correlations for competitive interactions and positive residual

219   correlations for mutualistic interactions. However, the magnitude of residual correlation

220   does not directly relate to interaction strength because it is mostly driven by species

221   prevalence. (ii) Positive-negative interactions such as predator-prey relationships can

222   be problematic for model interpretation because JSDMs cannot disentangle these from

223   competitive and mutualistic interactions. (iii) At increasingly coarser spatial resolution

224   of the data, both the signals of negative and positive interactions become indiscernible

225   by JSDMs, but reassuringly the signals rarely were confounded. Overall, our results have

226   important implications for the interpretation of JSDMs, and pinpoint important aspects

227   that will need further model development. At the same time, we want to stress that

228   these results were obtained under idealized conditions, and real-world applications will

229   be complicated by even more factors. If JSDMs fail to detect interspecific interactions in

230   very simple cases such as the one used here, they is no reason why they should work

231   better at detecting interactions (occurring within local communities) in more complex

232   real world applications.

233 First, we tested JSDMs' ability to detect negative and positive interactions from co-

234 occurrence data. And indeed, JSDMs reliably detected competitive and mutualistic

235 interactions in our simulated data, which is promising. However, interpretation of model

236 residuals as interaction coefficients proved non-trivial. Importantly, prevalence strongly

237 affected the magnitude of residual correlations estimated by JSDMs. This implies that

238 residual correlations from JSDMs do not directly relate to nor can be interpreted as

239 interaction strength, and also they cannot be compared across different species (with

240 different prevalence levels). In real-world applications, this may be further confounded

241 by the species' environmental response, detection bias, and by (direct and indirect)

242 interactions between multiple species, which should be further tested.

243 Second, we asked whether JSDMs could distinguish competitive and mutualistic

244 interactions from asymmetric interactions such as predator-prey relationships, which is

245 important when dealing with taxa from multiple trophic levels. As shown previously by

246 Araújo and Rozenfeld (2014), asymmetric positive-negative interactions can result in

247 co-occurrence patterns that could be either higher or lower than expected by chance.

248 Here, we showed that this also depends on interaction strength between species A and B

249 and their prevalence. Specifically, predator-prey relationships are more likely to

250 produce positive residual correlations when prevalence of both species is low. By

251 contrast, when prevalence of at least one species is high, then predator-prey

252 relationships are more like to produce negative residual correlations in JSDMs. As a

253 consequence, JSDMs are not able to tease apart predator-prey relationships from

254 competitive or mutualistic interactions, which has been discussed previously but never

255 been shown explicitly (Morales-Castilla, et al. 2015, Warton, et al. 2015). These results

256 suggest that a useful next step would be to incorporate directional conditionality into

257 future joint species distribution modelling, extending their capacity beyond pairwise

258 symmetric correlations (Hui 2016, Pollock, et al. 2014, Warton, et al. 2015). Recent

259 examples are beginning to consider asymmetric and more complex interactions for

260 small datasets (e.g. Harris 2016, with up to 20 species). These advances would be

261 extremely useful not only for distinguishing predator-prey relationships but also for

262 detecting asymmetric interaction strengths between competing or facilitating species

263 (Harris 2016, Kissling, et al. 2012). Still, it remains to be tested in how far co-occurrence

264 data will allow estimation of asymmetric residual correlation or under which

265 circumstances co-abundance data will be needed (Cazelles, et al. 2015). Incorporating

266 trait information could also help to distinguish competitive and mutualistic interactions

267 from predator-prey relationships, for example by simple plausibility checks (Morales-

268 Castilla, et al. 2015).

269 Last, we tested for scale dependence of JSDM estimates by analysing the co-occurrence

270 data at increasingly coarse resolution while the interactions took place at the original,

271 fine spatial resolution. We found that with increasingly coarser resolution, both the

272 signals of positive and negative interactions became indiscernible by JSDMs, whereby

273 the signal of negative interactions was more sensitive to scale. This is in line with

274 previous findings that co-occurrence patterns are highly scale-dependent (Araújo and

275 Rozenfeld 2014) and also that the effects of local interspecific interactions vanish at

276 coarser spatial scales (Thuiller, et al. 2015). Thus, JSDMs are unlikely to give any insight

277 on the potential interactions between species if the scale of the data does not match the

278 process scale. At best, they might give better and more reliable models and the residual

279 correlation might shed light on missing environmental variables and historical factors.

280 Also, the species lists from monitoring schemes or atlas data are often at comparably

14

281   coarse resolution, for example for many animals, such that the signal of interspecific

282   interactions and sub-scale environmental heterogeneity can get easily confounded

283   (Zurell, et al. 2016b). When we apply JSDMs to vegetation plots, the analysis of the

284   residual correlation in respect to interspecific interactions might thus be more

285   meaningful since the interactions between species, for instance for light competition,

286   might here be at the right scale (e.g. a 10x10m plot). Empirical analyses testing whether

287   the residual correlations from JSDMs are related to known interactions between species

288   might thus be of particular interest. Furthermore, our results showed that less prevalent

289   species are less sensitive to mismatches between process scale and data scale. At the

290   same time, at increasingly coarse resolution negative and positive interactions were

291   confounded easier. Thus, we recommend interpreting JSDMs cautiously when modelling

292   rare species.

293   In summary, our simulation study indicates that JSDMs might be reliable in inferring

294   simple symmetric species interactions if the species are modelled at the resolution of

295   the interactions. However, even in our idealized and error-free data, JSDMs could easily

296   confound predator-prey relationships with either competitive or mutualistic

297   interactions, they do not give an estimate of interactions strengths, they are not easily

298   comparable across species, they are affected by prevalence and by scale. Thus, we

299   should be very cautious when interpreting JSDM results as interactions (Clark, et al.

300   2014, Pollock, et al. 2014). Nevertheless, they have the ability of enhancing our analyses

301   and allow us to form hypotheses about potential interspecific interactions when used

302   cautiously (Ovaskainen, et al. 2010). Our results provide a first test of potentially

303   important or confounding factors in JSDM analyses. In the future, more efforts are

304   needed to test JSDMs in more complicated settings, for example to test their ability to

305 disentangle environmental response and residual correlation for multiple interacting

306 species, for cases when species interact directly and indirectly, and for other

307 confounding factors such as detection bias and missing environmental covariates.

308 Thereby, further theoretical tests under controlled conditions could be useful, but we

309 also emphasise the need for more empirical tests, for example comparing JSDM

310 estimates against known interactions (e.g. from experiments). Hence, JSDMs constitute

311 an important step forward for testing hypotheses of how interspecific interactions affect

312 species distributions. Nevertheless, there are important limits to their capability of

313 detecting interspecific interactions, for example for trophic and indirect interactions,

314 highlighting the need for further model improvements in this respect.

315 ## Acknowledgements

323

References

324 **References**

325 Araújo, M. B. and Rozenfeld, A. 2014. The geographic scaling of biotic interactions. -

326   Ecography 37: 406-415.

327 Blois, J. L., et al. 2013. Climate Change and the Past, Present, and Future of Biotic

328   Interactions. - Science 341: 499-504.

329 Cazelles, K., et al. 2015. A theory for species co-occurrence in interaction networks. -

330   Theoretical Ecology 9: 39-48.

331 Chase, J. M. and Leibold, M. A. 2003. Ecological niches: linking classical and

332   contemporary approaches. - The University of Chicago Press.

333 Clark, J. S., et al. 2014. Competition-interaction landscapes for the joint response of

334   forests to climate change. - Global Change Biology 20: 1979-1991.

335 Elith, J. and Leathwick, J. R. 2009. Species Distribution Models: Ecological Explanation

336   and Prediction Across Space and Time. - Annual Review of Ecology, Evolution, and

337   Systematics 40: 677-697.

338 Gotelli, N. J., et al. 2010. Macroecological signals of species interactions in the Danish

339   avifauna. - PNAS 107: 5030-5035.

340 Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than

341   simple habitat models. - Ecology Letters 8: 993-1009.

342 Guisan, A. and Zimmermann, N. E. 2000. Predictive habitat distribution models in

343   ecology. - Ecological Modelling 135: 147-186.

344 Harris, D. J. 2016. Inferring species interactions from co-occurrence data with Markov

345   networks. - Ecology 97: 3308-3314.

346 Hui, F. K. C. 2016. boral – Bayesian Ordination and Regression Analysis of Multivariate

347   Abundance Data in r. - Methods in Ecology and Evolution 7: 744-750.

348   Jetz, W., et al. 2012. Integrating biodiversity distribution knowledge: toward a global

349   map of life. - Trends in Ecology & Evolution 27: 151-159.

350   Kissling, W. D., et al. 2012. Towards novel approaches to modelling biotic interactions in

351   multispecies assemblages at large spatial extents. - Journal of Biogeography 39:

352   2163-2178.

353   Latimer, A. M., et al. 2009. Hierarchical models facilitate spatial analysis of large data

354   sets: a case study on invasive plant species in the northeastern United States. -

355   Ecolgy Letters 12: 144-154.

356   Leathwick, J. R. and Austin, M. P. 2001. Competitive interactions between tree species in

357   New Zealand's old-growth indigenous forests. - Ecology 82: 2560-2573.

358   McGill, B. 2010. Matters of scale. - Science 328: 575-576.

359   Meier, E. S., et al. 2010. Biotic and abiotic variables show little redundancy in explaining

360   tree species distributions. - Ecography 33: 1038-1048.

361   Morales-Castilla, I., et al. 2015. Inferring biotic interactions from proxies. - Trends in

362   Ecology & Evolution 30: 347-356.

363   Ovaskainen, O., et al. 2010. Modeling species co-occurrence by multivariate logistic

364   regression generates new hypotheses on fungal interactions. - Ecology 91: 2514-

365   2521.

366   Ovaskainen, O., et al. 2016. Uncovering hidden spatial structure in species communities

367   with spatially explicit joint species distribution models. - Methods in Ecology and

368   Evolution 7: 428-436.

369   Ovaskainen, O., et al. 2017. How to make more out of community data? A conceptual

370   framework and its implementation as models and software. - Ecology Letters: doi:

371   10.1111/ele.12757.

372   Pellissier, L., et al. 2013. Combining food web and species distribution models for

373   improved community projections. - Ecology and Evolution 3: 4572-4583.

374   Pollock, L. J., et al. 2015. The Roles of Ecological and Evolutionary Processes in Plant

375   Community Assembly: The Environment, Hybridization, and Introgression

376   Influence Co-occurrence ofEucalyptus. - The American Naturalist 185: 784-7

377   Pollock, L. J., et al. 2014. Understanding co-occurrence by modelling species

378   simultaneously with a Joint Species Distribution Model (JSDM). - Methods in

379   Ecology and Evolution 5: 397-406.

380   Pulliam, H. R. 2000. On the relationship between niche and distribution. - Ecology

381   Letters 3: 349-361.

382   R Core Team. 2015. R: A Language and Environment for Statistical Computing. - R

383   Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-

384   project.org/.

385   Soberón, J. and Nakamura, M. 2009. Niches and distributional areas: Concepts, methods,

386   and assumptions. - Proceedings of the National Academy of Sciences 106: 19644-

387   19650.

388   Svenning, J.-C., et al. 2014. The influence of interspecific interactions on species range

389   expansion rates. - Ecography 37: 1198-1209.

390   Thuiller, W., et al. 2013. A road map for integrating eco-evolutionary processes into

391   biodiversity models. - Ecology Letters 16: 94-105.

392   Thuiller, W., et al. 2015. From species distributions to meta-communities. - Ecology

393   Letters 18: 1321-1328.

394   Urban, M. C., et al. 2016. Improving the forecast for biodiversity under climate change. -

395   Science 353.

396   Warton, D. I., et al. 2015. So Many Variables: Joint Modeling in Community Ecology. -

397   Trends in Ecology & Evolution 30: 766-779.

398   Wisz, M. S., et al. 2013. The role of biotic interactions in shaping distributions and

399   realised assemblages of species: implications for species distribution modelling. -

400   Biological Reviews 88: 15-30.

401   Zurell, D., et al. 2016a. Benchmarking novel approaches for modelling species range

402   dynamics. - Global Change Biology 22: 2651-2664.

403   Zurell, D., et al. 2016b. Effects of functional traits on the prediction accuracy of species

404   richness models. - Diversity and Distributions 22: 905-917.

405

406

407   Supplementary material (Appendix EXXXXX at <www.oikosoffice.lu.se/appendix>).

408   Appendix 1.

409

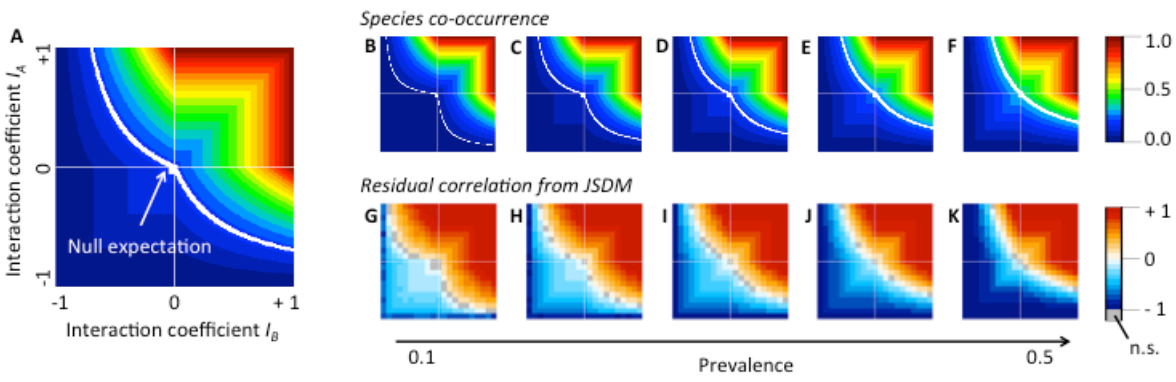410   **Figure captions**

411   Figure 1. Expected co-occurrence (A-F) and residual correlation (estimated by JSDM; G-

412   K) across the biotic interaction space of two species A and B and for different prevalence

413   levels. (A) Species A and B can have negative to positive effects on each other. The co-

414   occurrence index is defined as the number of cells where species A and B occur together

415   divided by the total number of occupied cells. The null probability of co-occurrence P(A

416   and B) is given by the product of species' prevalence P(A) x P(B), and the corresponding

417   co-occurrence value is given by (P(A) x P(B)) / ((P(A) + P(B) – P(A) x P(B)) (white

418   line). (B-F) show the co-occurrence values obtained at different prevalence levels (0.1,

419   0.2, 0.3, 0.4, 0.5) with equal prevalence for both species. Again the null expectation of co-

420   occurrence is shown in white. (G-K) show the corresponding residual correlations

421   estimated by JSDMs, with grey areas indicating non-significant residual correlations.

422

423   Figure 2. Scale dependence of co-occurrence patterns and corresponding residual

424   correlations estimated by JSDMs. Both species A and B have prevalence of 0.3. (A-C)

425   show the spatial distribution of species at increasingly coarser resolution resulting from

426   intermediate competition and intermediate mutualism (species A in black, species B in

427   grey, species A and B co-occurring in red). (D-F) show the resulting co-occurrences

428   across the biotic interaction space of species A and B, and (G-I) the corresponding

429   residual correlations estimated by JSDMs (please refer to Fig. 1 for legend description).

430   (J) shows the relationship between (significant) residual correlations estimated at the

431   original resolution and at coarser resolution.

432

433   Figure 3. Scale dependence of JSDMs for different prevalence levels. Shown is the

434   significant residual correlation estimated by JSDMs at the original scale against

435   significant residual correlation estimated at coarser resolution (block size 2x2 in black,

436   block size 4x4 in red) across biotic interaction space. Scale dependence of JSDMs

437   increases with prevalence from left to right such that no significant residual correlations

438   were estimated for the 4x4 block size in (E).

439

Figures

441

442   Figure 1
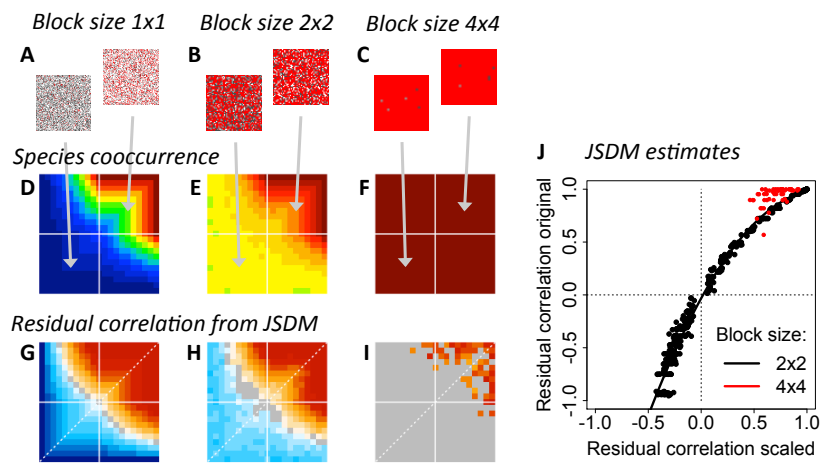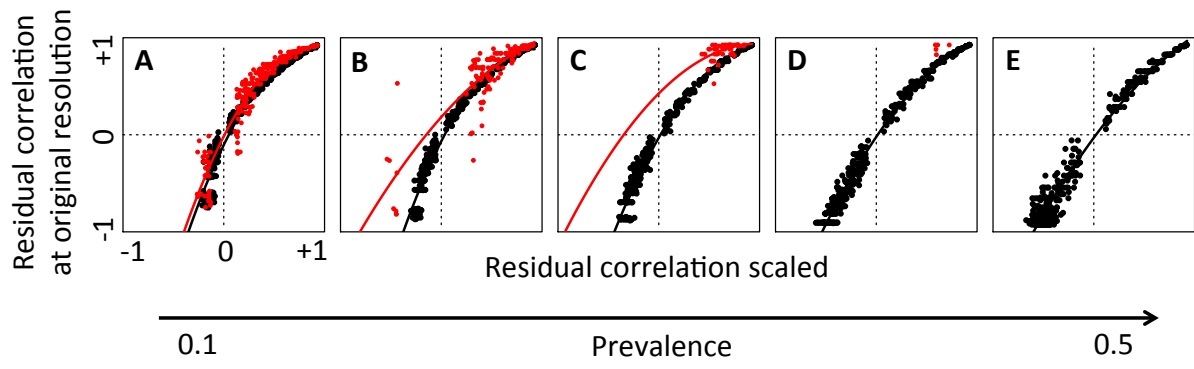
443

Block size 1x1   Block size 2x2   Block size 4x4

A **B** **C**

Species cooccurrence

D **E** **F**

Residual correlation from JSDM

G **H** **I**

J *JSDM estimates*

444

445   Figure 2

446

447

448   Figure 3