8-1-2018

Disentangling biotic interactions, environmental filters, and dispersal limitation as drivers of species co-occurrence

Manuela D'Amen
Université de Lausanne (UNIL)

Heidi K. Mod
Université de Lausanne (UNIL)

Nicholas J. Gotelli
University of Vermont

Antoine Guisan
Université de Lausanne (UNIL)

Recommended Citation

ECOGRAPHY

Research

Disentangling biotic interactions, environmental filters, and dispersal limitation as drivers of species co-occurrence

EDITOR'S CHOICE

Manuela D'Amen, Heidi K. Mod, Nicholas J. Gotelli* and Antoine Guisan*

M. D'Amen (http://orcid.org/0000-0001-5910-4267) (manuela.damen@msn.com), H. K. Mod and A. Guisan, Dept of Ecology and Evolution, Univ. of Lausanne, Lausanne, Switzerland. AG also at: Inst. of Earth Surface Dynamics, Geopolis, Univ. of Lausanne, Switzerland. – N. J. Gotelli, Dept of Biology, Univ. of Vermont, Burlington, VT, USA.

A key focus in ecology is to search for community assembly rules. Here we compare two community modelling frameworks that integrate a combination of environmental and spatial data to identify positive and negative species associations from presence–absence matrices, and incorporate an additional comparison using joint species distribution models (JSDM).

The frameworks use a dichotomous logic tree that distinguishes dispersal limitation, environmental requirements, and interspecific interactions as causes of segregated or aggregated species pairs. The first framework is based on a classical null model analysis complemented by tests of spatial arrangement and environmental characteristics of the sites occupied by the members of each species pair (Classic framework). The second framework, (SDM framework) implemented here for the first time, builds on the application of environmentally-constrained null models (or JSDMs) to partial out the influence of the environment, and includes an analysis of the geographical configuration of species ranges to account for dispersal effects.

We applied these approaches to examine plot-level species co-occurrence in plant communities sampled along a wide elevation gradient in the Swiss Alps. According to the frameworks, the majority of species pairs were randomly associated, and most of the non-random positive and negative species associations could be attributed to environmental filtering and/or dispersal limitation. These patterns were partly detected also with JSDM. Biotic interactions were detected more frequently in the SDM framework, and by JSDM, than in the Classic framework. All approaches detected species aggregation more often than segregation, perhaps reflecting the important role of facilitation in stressful high-elevation environments.

Differences between the frameworks may reflect the explicit incorporation of elevational segregation in the SDM framework and the sensitivity of JSDM to the environmental data. Nevertheless, all methods have the potential to reveal general patterns of species co-occurrence for different taxa, spatial scales, and environmental conditions.

NORDIC SOCIETY OIKOS

www.ecography.org

Introduction

Understanding the causes of species co-occurrence patterns is a major research focus in community ecology. Diamond (1975) proposed that species coexistence is regulated by 'community assembly rules' based primarily on species interactions, and inferred these rules from the pattern of species co-occurrence in replicated island assemblages. However, Connor and Simberloff (1979) used null model analysis to argue that co-occurrence structure may be no different than expected by chance. Since then, the search for community assembly rules has dominated community ecology (Gotelli and Graves 1996, Ovaskainen et al. 2010, Boulangeat et al. 2012, HilleRisLambers et al. 2012, Blois et al. 2014), and null model analysis has become a standard tool to search for patterns that may reflect processes of community assembly (Gotelli and Ulrich 2012). Early null model analyses were often based on a single summary metric for an entire assemblage, such as the number of species pairs that form perfect checkerboards (Graves and Gotelli 1993), or the number of missing species combinations in an archipelago (Simberloff and Connor 1981). More recent approaches view individual species pairs as a more informative unit of co-occurrence, allowing for a classification of all of the unique species pairs in an assemblage as random, aggregated, or segregated (Boulangeat et al. 2012, Gotelli and Ulrich 2012, Veech 2014).

Patterns of species aggregation may reflect positive biotic interactions such as mutualism, commensalism, and, under some circumstances, predation (Sih 1984). However, positive species associations could also reflect shared environmental requirements or historical factors such as sympatric speciation or common dispersal barriers. Similarly, negative species associations may reflect negative biotic interactions (as Diamond (1975) assumed), but could also reflect distinctive environmental niches or historical factors, such as allopatric speciation or isolation on opposite sides of dispersal barriers. However most null model analyses are not spatially explicit and assume that environmental requirements of species are similar. As a consequence, simple null models can identify non-random patterns of species association, but they cannot, by themselves, distinguish the separate effects of species interactions, dispersal limitation, and environmental filtering as causes of this non-randomness (Gotelli and Ulrich 2012, Blois et al. 2014).

In this study, we use primarily null model analysis to describe the pattern of pairwise species co-occurrence for a well-resolved data set on the occurrence of Swiss alpine plants using high resolution species and environmental data. We then use additional data on the spatial organization of species occurrences and the environmental structure of occupied and unoccupied sites to classify non-random species pairs that reflect the processes of biotic interactions, environmental associations, or dispersal limitations.

Recently, Blois et al. (2014) analysed the spatial arrangement and environmental characteristics of the sites occupied by the members of a species pair to infer the mechanisms responsible for pairwise aggregation or segregation. Their framework leads to a logic tree with nine tips or outcomes based on the co-occurrence pattern of each species pair (random, aggregated, or segregated) (Fig. 1a, b left box), the pattern of spatial overlap in pairwise occurrences (Fig. 1c left box), and the environmental structure of occupied and unoccupied sites (Fig. 1d left box), (Blois et al. 2014).

Figure 1. Comparison of workflows from Classic (left) and SDM (right) frameworks. Acronyms: D: either disjunct ranges or disjunct in elevation; POV: partially overlapping ranges; FOV: fully overlapping ranges.

Here, we propose a complementary set of analyses to understand the causes of species co-occurrence patterns (Fig. 1, right box). This framework integrates three different modules. 1) A null model analysis of pairwise species co-occurrence that estimate species occurrence probabilities from species distribution models (SDMs) to place species randomly in sites (environmentally-constrained null model, as proposed by Peres-Neto et al. 2001; Fig. 1a right box). Because the SDMs represent the application of an environmental filter on the community, this step is expected to maximize the chance of distinguishing between the influence of environmental preferences and biotic interactions. 2) A comparative null model analysis with the classic null model that does not incorporate environmental variables (Fig. 1b, c right box). 3) An analysis of the spatial and elevational geographic configuration of the species-pair distributions to distinguish species pairs that may be limited by dispersal (disjoint distributions) from those that may be limited by species interactions or niche differentiation (contiguous distributions; Fig. 1d right box and Fig. 2).

The main similarity of the two frameworks is that they both use null models. The main difference is in the use of SDMs to create an environmentally constrained null model versus the use of a classic null model. Thus, hereafter we refer to these as the 'SDM framework' and the 'Classic framework', respectively.

Because another emerging SDM tool – joint species distribution modelling (JSDM; Pollock et al. 2014, Warton et al. 2015) – is increasingly presented to infer community assembly rules, and could thus be integrated into our test, we present an analytical alternative in which we use JSDMs in modules 1) and 2) of the SDM framework. This test is made possible because JSDMs also use data on species occurrences and environment. The main difference with the SDM framework is that JSDM is not based on null models. Rather,

it fits a multivariate hierarchical model that uses generalized linear model, ordination techniques and latent variables to describe species association patterns. The significantly associated species-pairs can be classified as cases of environmental preferences or biotic interactions. The detailed description of the implemented JSDM is provided in Supplementary material Appendix 1.

Both null model frameworks and the JSDM approach are applied to an extensive inventory of alpine vegetation in the Swiss Alps.

Methods

Community data and environmental variables

The study area is located in the western Swiss Alps, Canton de Vaud (46°10′–46°30′N, 6°50′–7°10′E). It covers approximately 700 km², with a strong elevation gradient from 375 to 3210 m a.s.l. Plant species were exhaustively inventoried in 912 plots of 4 m² between 2002 and 2009. Site selection followed a balanced stratified random sampling design based on elevation, slope, and aspect and was restricted to open vegetation grasslands. Plots were visited at increasing elevations across the sampling season to follow as much as possible peak flowering (occurring later at higher elevations where snow persists longer) to optimize species identification. The specimens were identified following Aeschimann and Burdet (1994) but not collected (for more information on the sampling, see Dubuis et al. 2011). Geographic range and occurrence data were made available through the National Database for the Swiss Flora (InfoFlora < www.infoflora.ch >). We analysed the 175 most frequent species that each had 22 or more occurrences throughout all surveyed plots (out of several hundred
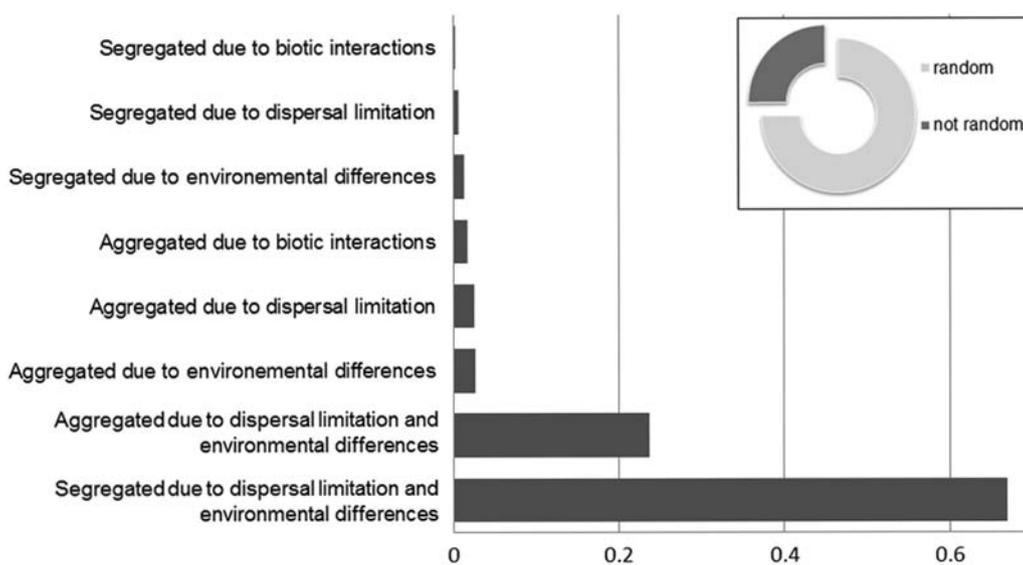
Figure 2. Visual summary of the results from the Classic framework. The pie chart shows the proportion of pairs that were random and not random. The bars show the relative frequencies of each category of the not random fraction.

species for the whole area, e.g. 260 used in Dubuis et al. 2011). This minimum sample size was used to ensure that SDMs were fit reliably. The data were organized as a binary presence–absence matrix with 175 columns (= species) and 912 rows (= plots), with each entry indicating the presence (1) or absence (0) of a particular species in a plot.

For each plot, we recorded the exact geographic coordinates and the elevation. Site environmental characteristics were extracted from temperature and precipitation data recorded by the Swiss network of meteorological stations and from a digital elevation model with a spatial resolution of 25 m (Dubuis et al. 2011). Specifically, we analysed: 1) growing degree-days (above 0°C), 2) a moisture index calculated as the difference between precipitation and evapotranspiration (values summed over the growing season), which expresses the amount of soil water potentially available, 3) solar radiation summed over the year, 4) plot slope (in degrees), and 5) topography. These five continuous variables have been previously shown to be important predictors of the distribution of the same set of plant species in the study area (Dubuis et al. 2011). For each pair of species, we additionally used these five variables to calculate niche overlap (see Supplementary material Appendix 2 for a detailed explanation of this procedure).

Null model analysis of species co-occurrence

We applied a pairwise null model analysis (Gotelli and Ulrich 2010) to the presence–absence plant species matrix to determine which species associations are significant across the study area (hereafter 'classic null model'). We quantified the strength of association between each pair of species a and b by the C-score index (Stone and Roberts 1990) rescaled between 0 and 1:

$$Cscore_{ab} = \frac{(R_a - S) \times (R_b - S)}{(R_a \times R_b)}$$

where R_a is the number of occurrences of species a, R_b is the number of occurrences of species b, and S is the number of sites that contain both species a and b. An index of 0.0 means that the species pair is maximally aggregated (occurrences are perfectly nested), and an index of 1.0 means that the species pair is maximally segregated (no co-occurrences, forming a perfect 'checkerboard pair' in Diamond's (1975) terminology).

C-scores for individual species pairs were compared to the statistical expectation for a set of 10 000 null communities generated with a 'fixed-equiprobable' null model algorithm. This algorithm preserves species occurrence frequencies (column totals), but allows species richness per plot (row totals) to vary randomly and equiprobably, which is appropriate for data sampled from plots of constant area (Gotelli 2000). The null model imposes constraints on the randomization to try and remove some obvious sampling factors, but it is not a mechanistic model with parameters that specify particular processes. Simple null models analyzed without any environmental or spatial data assume that differences in environmental conditions among sites and differences in dispersal potential among species are weak or not important. These null models generate patterns expected in the absence of all of these forces (Gotelli and Ulrich 2012).

We apply the C-score because of its large use in the co-occurrence literature. Alternatively, Schluter's (1984) V-ratio could have been used with the fixed-equiprobable model, but its performance is no better than the C-score (Gotelli 2000). We always compare the observed C-score with a null model that preserves the frequencies of species occurrences, which controls for shared absences. This means that for a given set of species occurrence frequencies, the more shared absences there are, the more likely the pattern will be identified as aggregated. Conversely, the fewer shared presences there are, the more likely the pattern will be identified as segregated.

With 175 species, there are 15 225 possible species pairs (175 × 174/2) and associated significance tests. We applied the empirical Bayes approach, which assumes independence of probabilities (Efron 2005), to control for the potentially large number of false positives that can emerge with the analysis of many species pairs (Gotelli and Ulrich 2010). The empirical Bayes approach is based on the comparison of the scores calculated for each species pair with those obtained for the same species pair in randomized matrices. By doing this, the test seeks to impose a realistic cutoff to identify 'interesting' cases, while still controlling for false discovery (Gotelli and Ulrich 2010). Operationally, to apply the empirical Bayes approach, each pairwise C-score was rescaled on a 0–1 interval and grouped into 22 classes of evenly spaced bins. Within each bin, the scores were ranked from smallest to largest. We then randomized the original matrix 1000 times, and we calculated the average number of species pairs in each bin. This produced a null distribution of the frequencies of species pairs with different scores. In each bin, we retained as significant pairs only the ones that had observed C-scores greater than the mean number of simulated species pairs. Further, within this selection, we retained only the species pairs for which the simple null hypothesis was also rejected by the standard randomization procedure. These steps should reduce the number of false positives and the probability of a type I error (Bayes M criterion; see Gotelli and Ulrich 2010 for further explanations of the method). This empirical Bayes method turns out to give results that are similar to false discovery rate calculations (Efron 2005).

To distinguish between patterns of segregation and aggregation among the selected non-random pairs, we compared the observed score with the mean of the simulated scores for a particular species pair. To allow comparisons across species pairs, we re-scaled the p-values for each species pair in units of standard deviations (Gurevitch et al. 1992). The relative standardized effect size (Z-score) for the co-occurrence indices is calculated as $(I\ obs - I\ sim)/(\sigma\ sim)$, where I_{obs} is the observed index, I_{sim} and σ_{sim} are respectively the mean and the standard deviation of the 10 000 indices from the simulated communities (Gotelli and McCabe 2002). Large positive Z-scores indicate segregation (relatively large C-score

values) and large negative Z-scores indicate aggregation (relatively small C-score values).

Results from this pairwise null model analysis are used by both the Classic framework and the SDM framework, described below.

Classic framework

The empirical Bayes approach allowed us to classify each species pair as aggregated, segregated, or random (Fig. 1a, b left box). We next inferred the roles of species interactions, dispersal limitation, and environmental association as potential causes of non-randomness. For each significantly segregated species pair, we compared spatial locations (spatial test) and environmental characteristics (environmental test) of the set of sites that contained only species a (1,0) versus only species b (0,1). Similarly, for the significantly aggregated species pairs, we performed the spatial and the environmental tests to compare the set of sites that contained both species (1,1) with the set of sites that contained neither species (0,0) (Fig. 1c, d left box).

1) Spatial test (Fig. 1c left box). We used a MANOVA to tests for the spatial overlap of allotopic sites ((1,0) and (0,1)) in the case of segregated pairs and for the spatial overlap of syntopic (1,1) and empty (0,0) sites in the case of aggregated pairs. We treated the geographic coordinates as a two-element response vector, and then we used a one-way MANOVA to compare the distance in coordinate space between the group centroids of the different site types. Because our study area is characterized by a steep elevational gradient, we also tested with a one-way analysis of variance (ANOVA) the elevational overlap between the same classes of sites. The null hypothesis here is that the spatial overlap of the different classes is random. If this null hypothesis is rejected, dispersal limitation is implicated as a primary cause of the aggregated (or segregated) occurrence observed pattern for a particular species pair. We used this spatial test to represent the original method proposed by Blois et al. (2014), but other methods to account spatial patterns that are more complex than the ones solely captured by the geographical coordinates could also be applied to refine these results (Wagner and Dray 2015).

2) Environmental test (Fig. 1d left box). We compared environmental characteristics between the allotopic sites of the segregated pairs (occupancy classes (1,0) vs (0,1)) and between the syntopic sites and empty sites of the aggregated pairs (occupancy classes (1,1) vs (0,0)). We first performed a principal components analysis (PCA) with the five continuous environmental variables measured at each site. We then treated the first two principal component scores as a two-element response vector in a one-way multiple analysis of variance (MANOVA) to test the null hypothesis that the frequencies of different environmental types did not differ among the site classes. Rejecting this null hypothesis implies that environmental differences among plots are at least partly responsible for patterns of species co-occurrence.

Results from the environmental and spatial tests allowed us to assign each non-random aggregated or segregated species pair to one of the following categories: 1) signal of dispersal limitation; 2) signal of environmental differences; 3) signals of both processes; 4) signals of neither processes (Fig. 1E left box, Fig. 2 and Supplementary material Appendix 4 Table A2). The last category represents cases of non-random pairwise co-occurrence that can be attributed exclusively to biotic interactions because there was no evidence for spatial or environmental effects.

SDM framework

The SDM framework is composed of three modules

1) Environmentally-constrained null model (Fig. 1a right box). This module is based on the 'environmentally-constrained' null model approach proposed by Peres-Neto et al. (2001). We ran a co-occurrence analysis with the C-score index, using an environmentally-constrained null model. We generated the environmentally-constrained null communities following the algorithm Ct-RA1, in which the species presences were reassigned to sites in the matrix according to species-specific relative probability values calculated for each site. Like the fixed-equiprobable model, this algorithm preserves the species occurrence frequencies in the original matrix (see Peres-Neto et al. 2001 for further explanations).

The probability values for the sites were obtained by fitting species distribution models (SDMs; Guisan et al. 2017) to the presence–absence data for each plant species using the five environmental predictors described above. A high predicted probability indicates suitable environmental conditions for species occurrence, and vice versa. We applied three modelling techniques in R (2.14.1) with the BIOMOD package (Thuiller et al. 2009): generalized linear models (GLMs), generalized additive models (GAMs) and generalized boosted models (GBMs). The resulting three projections were averaged to implement a single ensemble estimate of the probability of presence for each species in each plot (see Supplementary material Appendix 4 Table A1 for summary evaluation scores of the SDMs). Finally, we applied the Bayes correction to adjust for the large number of tests (Bayes CL criterion; Gotelli and Ulrich 2010), and we scaled the results in units of standard deviations (Gurevitch et al. 1992), as described for the classic null models. The inclusion of the likelihood of species occupancy of a site during the generation of 'null communities' facilitates the distinction between environmental filtering and biotic interactions as causes of species associations.

2) Classic null models (Fig. 1b right box). The description of this analysis is provided above in section Null model analysis of species co-occurrence. Following Peres-Neto et al. (2001), we contrasted results from classic and environmentally-constrained null models. This comparison provides the necessary information to derive a simple biological interpretation (dispersal limitation, environmental niche differences, biotic interactions, and/or random co-occurrence) of the

observed pattern for each of the 15 225 unique species pairs (Supplementary material Appendix 4 Table A3).

3) Analysis of the geographic configuration of species ranges (Fig. 1c right box and Fig. 3). To classify the species pairs on the basis of their range configuration in the study area, we created a logic tree with three forks and four possible outcomes. The first level of branching separates species pairs that do or do not overlap in their spatial convex hulls, calculated as the minimum convex polygon (MCP) encompassing the occupied plots for each species (IUCN 1994), to represent an approximation of the species distributional range. We label the non-overlapping species pairs as 'Disjunct in the ranges' (or allotopic) (DR). The species pairs that do overlap in their convex hulls are further divided based on whether or not they show overlap in their elevation range. We label the non-overlapping pairs as 'Disjunct in elevation' (DE). The final branching for each pair overlapping both in the MCP and elevation is based on a MANOVA analysis to test whether the spatial centroids of the two species occurrences differ significantly. The null hypothesis is that the spatial centroids do not differ, and the species are syntopic (labelled as 'Fully overlapping' – FOV). The alternative hypothesis is that the spatial centroids differ, and the species are parapatric in the study area ('Partly overlapping' – POV) (Fig. 3). Results from this last module allowed us to assign each non-random aggregated or segregated pair to one (or more) of the following categories: Habitat filtering, Dispersal limitation, Positive or Negative biotic interactions (Fig. 1*E* right box, Fig. 4 and Supplementary material Appendix 4 Table A3).

The analyses to run both the SDM framework and the Classic framework were coded in R (ver. 3.1.2, R Core Team)
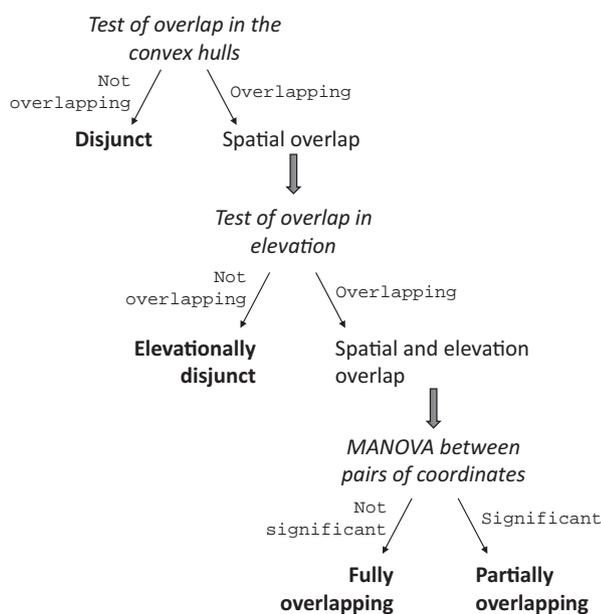
using the libraries 'ade4' (Dray et al. 2007), 'adehabitatHR' (Calenge 2006), 'ecospat' (Di Cola et al. 2017), 'maptools' (Bivand and Lewin-Koh 2014), 'raster' (Hijmans 2015), 'rgdal' (Bivand et al. 2017), 'rgeos' (Bivand and Rundel 2015) and 'vegan' (Oksanen et al. 2014). For both null model analyses, we adapted the procedures in the FORTRAN program PAIRS (Ulrich 2010) for analysis in R with the fixed-equiprobable null model. These functions are available in the 'ecospat' library, and here we use a modified version to rescale the C-score index between 0 and 1. R scripts to apply the empirical Bayes approach to null model results and to run the Classic framework and the SDM framework are available in the on line supplementary materials (Supplementary material Appendix 3).

JSDM approach

The JSDMs were applied within the SDM framework as an alternative in the modules 1) and 2). This methodology was applied as implemented in the 'boral' R-package by Hui (2016) and suggested also by Warton et al. (2015). It uses a model-based approach to ordination, in which latent variables are included to estimate species interactions (namely significant residual correlations). The non-random species associations due to the environment are inferred from significant correlations in environmental responses after fitting a generalized linear model that uses species occurrence matrix as response variable and environmental covariates as explanatory predictors (here the five environmental variables as linear and quadratic coefficients). We used the function 'boral' to first fit a pure latent variable model (mimicking a classic null model) and then a model including both environmental covariates and latent variables (mimicking an environmentally-constrained null model). Thus, based on the significant residual and environmental correlations, we were able to use JSDM to classify species pairs as random or associated, with further discrimination of segregation or aggregation resulting from habitat filtering and/or biotic interactions (as in modules i and ii in the SDM framework). See Supplementary material Appendix 1 for a detailed explanation of the methods.

Elevation effect

To examine the potential effect of the elevation gradient on species co-occurrence patterns, we classified each species according to its elevational range extent into four exclusive categories: ele1 – range spanning the alpine and subalpine to alpine belts; ele2 – range spanning the montane to subalpine and colline to subalpine belts; ele3 – range spanning montane and colline to montane belts; ele 4 – species with broad elevation range spanning the montane to alpine and colline to alpine belts (Supplementary material Appendix 4 Table A4). We also calculated how often each species formed pairs pertaining to the co-occurrence classes identified by either the SDM framework or the Classic framework. We tested whether plants with different elevation ranges have different co-occurrence class frequencies.

Figure 3. Decision tree for the SDM framework showing the steps to classify the geographical arrangement of species pairs ranges in the study area.
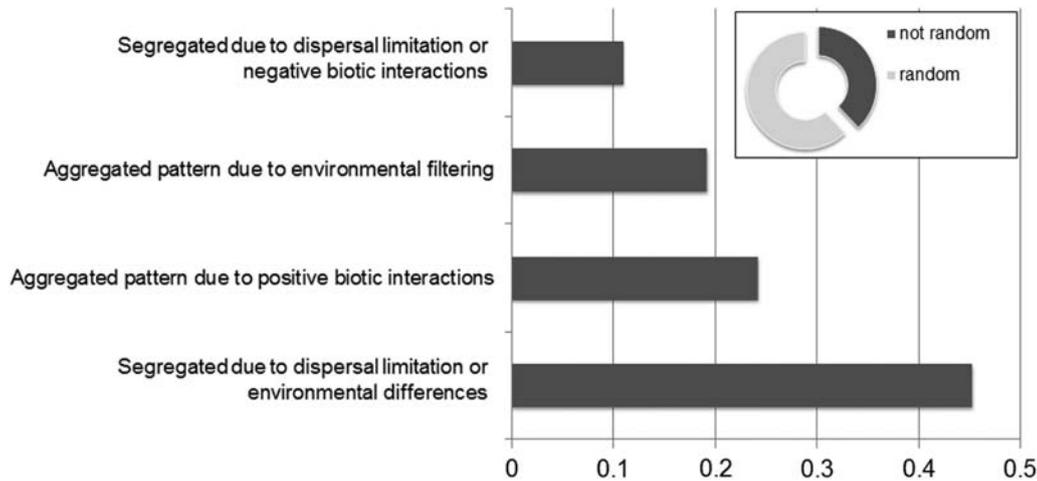
Figure 4. Visual summary of the results from the SDM framework. The pie chart shows the proportion of pairs that were random and not random. The bars show the relative frequencies of each category of the not random fraction.

Data deposition

Results

Classic null model

The pairwise classic null model identified 3825 significantly associated pairs (around 25% of all possible pairs) after the Bayesian correction. Of these non-random pairs, 1179 were aggregated and 2646 were segregated, each with an approximately normal distribution of association strengths.

Classic framework

The spatial and environmental tests performed within the Classic framework were significant for the majority of non-random species pairs. This means that most of the significantly aggregated or segregated species pairs identified with the classic null model differed in geographic distance, elevational range, or environmental characteristics of occupied and unoccupied sites, and usually differed in both spatial and environmental aspects (Fig. 1E left box, Fig. 2 and Supplementary material Appendix 4 Table A2). In the Classic framework, species pairs that did not differ in these site characteristics imply that biotic interactions produce non-random patterns of aggregation or segregation (Blois et al. 2014). Although these cases were relatively few, the Classic framework detected a larger number of aggregated species pairs (68) versus segregated species pairs (8) that imply biotic interactions.

SDM framework

The pairwise environmentally-constrained null model identified 2048 significantly associated pairs (around 13% of all possible pairs) after the Bayesian correction. Of these non-random pairs, 1448 were aggregated and 640 were segregated (Fig. 1a right box). From the analyses of spatial configuration of ranges, we tallied the number of species pairs found for each of the four branch tips of the SDM framework (Fig. 1d right box, Fig. 2 and Fig. 4). 723 species pairs formed disjunct ranges in space that did not overlap in their convex hulls. An additional 127 species pairs did not overlap in elevation. These species pairs with elevational disjunctions had significantly less niche overlap than species pairs that overlapped partly or completely in elevation (see Supplementary material Appendix 4 for a detailed description of niche overlap results). For these latter two groups, there were 8668 species pairs that were syntopic (no significant segregation in their spatial occurrences and high niche overlap values), and 5705 species pairs that overlapped partially (Fig. 2).

Partitioning the SDM results from the classic and environmentally-constrained null models and spatial configuration analyses, we categorized the 15 225 species pairs into different classes on the basis of which factor predominated in shaping the pairwise co-occurrence pattern (Fig. 1E right box). The largest fraction of species pairs (9429 pairs – 62% of the total pairs) showed no significant association (both null models detected a random pattern). Allotopic species pairs (no overlap in the convex hull either in space or elevation) were either randomly distributed (370 pairs) or were identified by the classic null model as segregated (405 pairs disjunct in space and 74 pairs disjunct in elevation). 3268 pairs were classified either as segregated or aggregated by the classic null model, but were classified as randomly associated by the environmentally-constrained null model; these cases imply the role of the environment in determining the observed patterns.

The syntopic pairs identified as significantly aggregated or segregated by the environmentally-constrained null models are those for which biotic interactions can be invoked to explain patterns of species association (12%; 1279 pairs). Only 5% of these pairs were also classified as significantly segregated or aggregated by the classic null model. Almost

half of the species pairs with partially overlapping ranges (parapatric in the study area) were not significantly associated. Around 30% of the pairs with partially overlapping ranges were identified by the classic null model as spatially segregated, this pattern being likely due to dispersal limitation. The environmentally-constrained null models classified 12% of the parapatric pairs as significantly aggregated, which could be attributed to positive biotic interaction at the overlapping range borders. Only a small percentage of partially overlapping species pairs were identified as segregated by the environmentally-constrained null model, which suggests either species interactions or dispersal limitation are responsible for most cases of parapatry in the study area.

In comparison, JSDMs only classified 2476 species pairs as random (Supplementary material Appendix 1). The associated species pairs were more often aggregated than segregated. Most of the associations among species pairs were at least partly due to habitat filtering, yet biotic interactions were detected as the only cause for segregation in 903 cases and for aggregation in 36 cases.

Frameworks comparison

Both null model-based frameworks detected a random spatial pattern for 62% of species pairs, which is considerably less than the frequency of randomness detected in other plant and animal assemblages using a more restrictive fixed-fixed null model (over 90% – Gotelli and Ulrich 2010). JSDM classified even fewer species pairs as randomly associated. All segregated species pairs with disjunct ranges, which were attributed to dispersal limitation by the SDM framework, were significant for both the spatial and environmental tests in the Classic framework. In contrast, almost all species pairs with co-occurrence patterns attributed exclusively to dispersal limitation (124 pairs) by the Classic framework had fully overlapping ranges (103 pairs), thus were attributed by the SDM framework to environmental affinities/differences. Almost all segregated pairs (97%) explained by environmental differences in the SDM framework also yielded a significant environmental test (and 84% also yielded a significant spatial test) in the Classic framework. Similarly, 86% of aggregated pairs explained by environmental affinities in the SDM framework yielded a significant environmental test (and 77% also yielded a significant spatial test) in the Classic framework. Most of the significant species pairs exhibiting partially overlapping (POV) ranges (1841 pairs) were segregated according to the classic null model. Among these, only 4 were segregated also according to the environmentally-constrained null model. These POV species pairs were classified into different categories by both frameworks. As with the null model frameworks, the JSDM highlighted the role of environment in explaining species association patterns. However, environmental preferences in the JSDM were identified more often as the cause of aggregated rather than segregated patterns.

Overall, the SDM framework, and even more the separate JSDM analyses, revealed a greater importance of biotic interactions than the Classic framework. Both frameworks also differed in the mechanisms attributed to significant aggregated or segregated species pairs. Only two non-random species pairs were identified by both the Classic framework and the SDM framework as being caused by biotic interactions (both segregation patterns), but neither of them were classified as such by JSDM. All the other pairs attributed to biotic interactions in the Classic framework (74 pairs) were instead attributed to environmental factors by the SDM framework. Most of the significantly associated pairs that imply the influence of biotic interaction according to the SDM framework were not significant in the classic null model (97% of segregated and 95% of aggregated pairs). On this last fraction of species pairs, we performed environmental and spatial tests from the Classic framework: for both segregated and aggregated pairs, around half of the tests were not significant.

Elevation effect

We did not detect significant differences in plants elevation range for any of the co-occurrence classes identified by the Classic framework. In contrast, for some of the co-occurrence classes identified by the SDM framework, there were significant trends along elevational gradients. In particular, species with higher elevation ranges formed more positive biotic interactions than species pairs of species from lower elevations (ANOVA test $p < 0.01$). In contrast, there was no significant trend with elevation for the species forming negative biotic interactions. Species pairs segregated due to dispersal limitation were significantly more common at lower elevation (ANOVA test $p < 0.01$). Species with broad elevation ranges were more frequently classified as aggregated due to common environmental conditions (ANOVA tests, $p < 0.01$), whereas no significant trend was detected for species segregated due to environmental differences.

Discussion

Commonalities or differences in species dispersal or environmental requirements may lead to spatial patterns of species similar to those produced by biotic interactions; thus, untangling the effect of these processes is particularly challenging. To date, a common approach to detect biotic interactions has been to explicitly or implicitly limit the analyses to a set of partially adjacent sites with similar environments, so that dispersal limitation and environmental filters could be considered unimportant (Phillips et al. 2003, Zhang et al. 2011). When considering larger spatial extents, such simplification cannot be applied. Several recent studies developed analytical approaches to disentangle the roles of environmental filtering and interspecific interactions in co-occurrence matrices. However, these frameworks did not explicitly consider the effect of dispersal limitation (Ovaskainen et al. 2010, Pollock et al. 2014, Bar-Massada 2015, Harris 2016). Developing a deductive framework that

considers together these three processes within field observation data is expected to improve our understanding of the mechanisms structuring ecological communities.

Here we compared the performance of two such frameworks (and one alternative approach using the recently promoted JSDMs, provided in the Supplementary material Appendix 1). The first 'Classic framework' has been recently proposed and applied to examine taxon associations through time for late Quaternary fossil pollen assemblages at relatively large spatial extent and small spatial grain (Blois et al. 2014). The second 'SDM framework' is here proposed for the first time and builds on the Peres-Neto et al. (2001) algorithm, but complemented with an analysis of spatial configuration of species ranges. The main affinity between the two frameworks is that they are based on null models and a logic decision tree structure: the species pairs that can be attributed to interspecific interactions are identified by a deductive process of elimination that excludes spatial and/or environmental effects. Both frameworks and the JSDM apply to presence/absence data: this makes their use more generally applicable, as quality abundance data are not often available. A core difference between them is the inclusion of the environmentally-constrained null model in the SDM framework, which is expected to maximize the chance of distinguishing between the influence of environmental preferences and biotic interactions (Peres-Neto et al. 2001). As a separate and not fully comparable alternative, the JSDM method has also been suggested to differentiate the causes of species association patterns (Pollock et al. 2014, Warton et al. 2015), but does not include a null model component.

The outcomes of the SDM framework and JSDM are both highly dependent on the performance of SDMs. Many factors involved in the construction of an SDM affect its prediction performance, such as sample size, spatial arrangement of occupied and unoccupied sites, choice of modelling technique, and sampling bias. In this study, we used a robust data set, which includes both presences and absences of plants, sampled with a random stratified design (Hirzel and Guisan 2002). The models are run with environmental correlates that previous studies reported to be relevant predictors of the distribution of alpine plant species. We also verified that there is no spatial autocorrelation for the very large majority of species (Schmid 2014, Di Cola et al. pers. comm.). In addition, a study conducted in this specific area, using virtual species based on real plant species observations, revealed that spatial autocorrelation has only negligible effect compared to other factors possibly affecting spatial predictions (Thibaud et al. 2014).

The use of SDMs to apply an environmental filter on the species assemblages is not new in community ecology (D'Amen et al. 2015b). For instance, SDMs have been integrated in a recently proposed modelling framework aimed at reconstructing the community assembly process (SESAM framework, Guisan and Rahbek 2011). This community-level approach includes different modelling steps that account for dispersal limitation, habitat filtering (using SDMs) and species interactions, similarly to our Classic and SDM frameworks (D'Amen et al. 2015a). However, the goal of SESAM is different than the one intended here: whereas SESAM aims to predict spatio-temporal patterns of species assemblages from individual species, here we depart from the assemblage level and aim to estimate the drivers of pattern for individual species pairs.

We implemented the environmentally-constrained null model by using as weights the probabilities of occurrence generated by SDMs, and, in addition to the original implementation, we included also an analysis of spatial range configuration (Blois et al. 2014). We classified ranges of species pairs using the biogeographical distinction among allopatric, parapatric and sympatric distributions, applied at the spatial scale of our study area (which does not include full species ranges). The condition of ranges disjunction (allopatry) implies the role of dispersal limitation as a primary cause of the observed pairwise pattern. Conversely, a pattern of fully overlapping ranges (sympatry) excludes the mechanism of dispersal limitation. For the pattern of partially overlapping ranges (parapatry), two main biological explanations have been proposed. It has been shown in different taxonomic groups that the location of a species border is often determined by the border of a related species, particularly towards the milder environmental edge of species distributions (the more stressful end being expected to be closer to an abiotic limit; Normand et al. 2009), leading to parapatric pattern (Darwin 1869, Miller 1967, Jaeger 1970). In our study, this translates into geographic exclusion along an elevation gradient in which the parapatric overlap zone will occur at the physiological limit of the competitively dominant species. The competitively inferior species, often with wider physiological tolerance, may nevertheless be restricted to a smaller area in which the dominant species cannot persist (Bull 1991, Gaston 2003, Bridle and Vines 2007). On the other hand, many studies have shown the important role of environmental factors alone in preventing overlap: the strong change in abiotic conditions along the environmental gradient could exceed the physiological capacities of the species involved and thereby limit ranges and co-occurrence with only a weak role for species interactions (Barton and Hewitt 1985, Hewitt 1988). In these cases, literature on how some of these species behave in botanical gardens worldwide, in common garden experiments or in reciprocal transplant experiments at different elevations would be helpful to further clarify the process (Anderson et al. 1996, Vetaas 2002, Hautier et al. 2009), but such data is currently not available for nearly all of the species considered here.

In our study, the classic null model produced a relatively high fraction of species pairs with non-random associations. Compared to other similar analyses, this proportion is quite large: 25% of all possible pairs in our analysis compared with an average of 1–5% in other studies (Blois et al. 2014, Lopes et al. 2015, Lyons et al. 2016). This difference could be due to our use of the fixed-equiprobable algorithm, which may be less conservative than the more standard

fixed-fixed algorithm used in previous pairwise tests (Gotelli and Ulrich 2010). Another reason for this result could be that we restricted the analysis to common species (i.e. species with more than 30 occurrences). When both species in a pair are rare, it may be difficult to detect non-randomness. As in other studies of modern assemblages (Gotelli and Ulrich 2010, Lyons et al. 2016), segregated pairs were more common than aggregated pairs. However, with environmental information incorporated into the null model, most non-random species pairs were aggregated rather than segregated.

The strong elevational gradient in the study area creates very different environmental conditions between low and high elevations. In such a situation, one can expect environmental effects to be the primary driver of species distributions because it will filter species with the morphological and physiological traits necessary to survive at particular elevations (de Bello et al. 2012). In addition, the mountainous terrain can represent a serious barrier to dispersal. In a previous study of spatial patterns and species traits for the same plant communities (Dubuis et al. 2013), the topo-climatic factors explained most of the variation in species occurrence. With environmental filtering and dispersal limitation being so strong, we expect relatively few cases in which biotic interactions are responsible for non-random co-occurrence patterns, unless they are so strong that they drive species into extreme subsets of their environmental niche. Both the Classic framework and the SDM framework confirmed the primary importance of environment and dispersal limitation as the primer drivers of species co-occurrence. This does not mean that biotic interactions are absent, but that they are less likely to be the dominant driver of associations in high-elevation alpine plants. Nonetheless, when species interactions are ignored, stacked SDMs based only on climate variables typically overestimate species co-occurrence and total species richness. This bias (Dubuis et al. 2011, Pottier et al. 2013, D'Amen et al. 2015a, b, but see Calabrese et al. 2014) might propagate in community projections of species co-occurrence (Di Febbraro et al. pers. comm.).

In mountainous landscapes, biotic interactions – when present – are expected to vary with elevation (Pottier et al. 2013): the frequency of positive and negative interactions is predicted to change across the stress gradient, with facilitation being more common in places with high abiotic stress (Bertness and Callaway 1994, Michalet 2006, He et al. 2013, Chamberlain et al. 2014). Moreover, facilitation among plants has often been demonstrated in severe environments, such as high altitudes (Carlsson and Callaghan 1991, Choler et al. 2001, Callaway et al. 2002). The results from the SDM framework do indeed provide evidence for an increasing importance of facilitation at higher elevation, where positive biotic interactions were more frequently detected. In contrast, the frequency of negative associations that could be attributed to species interactions did not vary along the elevational gradient. For plant assemblages, evidence for competitive effects may be more clearly detected by analysing plant life forms. For example,

small-stature alpine plant species are frequently excluded by large-stature species at lower elevations (Gotzenberger et al. 2012, Wisz et al. 2013).

Between the results of the null models and JSDM analyses, there were two main differences. First, the JSDMs identified substantially more species pairs as non-random. This is likely due to the more conservative nature of the C-scores metric compared to the ordination technique used by JSDMs. Second, JSDMs classified more non-random associations as due to biotic interactions. This result presumably comes from JSDM's sensitivity to omitted predictor variables (Pollock et al. 2014, Hui 2016): because the identification of biotic interactions in JSDMs is based on species residual correlations, unexplained deviance due to a missing predictor may also translate as the effect of biotic interactions.

As a consequence, the JSDM detection of species associations is sensitive to the choice of environmental covariates. We consider this as the main drawback compared to the other two frameworks, which rely on a comparison of expected co-occurrence frequencies from constrained randomizations. On the other hand, JSDM's advantage is its efficiency: within one (iterative) model fit, it can distinguish the roles of environment and other species in defining communities. Recent implementations of JSDMs also account for the effects of species traits and phylogeny, and the spatial and/or temporal structure of the data, allowing for interpretations at multiple spatial, temporal, and phylogenetic scales (Clark et al. 2017, Ovaskainen et al. 2017). Here only a part of the JSDM analysis was exploited to allow for a simple comparison with the null model based methods. Results might differ with a more comprehensive implementation of JSDMs.

The approaches such as the Classic framework, SDM framework and JSDMs improve our understanding of how species interactions, habitat filtering, and dispersal limitation affect species associations and range limits (Araújo and Rozenfeld 2014, Morueta-Holme et al. 2016). These methods can be readily applied to other assemblages and at different spatial scales to improve forecasting accuracy and reveal the relative importance of different community assembly processes. Further, they narrow down the sets of species for which biotic interactions are most likely to be important.

However, no matter how sophisticated the analysis, it is still a challenge to infer cause-and-effect from statistical patterns of species co-occurrence and environmental associations. Without experimental verification, caution is needed. Even for cases in which segregated species pairs are not explained by dispersal limitation or habitat filtering, we cannot exclude all explanations other than negative biotic interactions. For example, missing predictors in the SDMs could have prevented us from recognizing a signal of habitat filtering in the pattern of co-occurrence. Additional effects of successional or non-equilibrium dynamics, as well as idiosyncratic biogeographic can all contribute to co-occurrence. The inclusion of information coming from trait data or phylogeny can represent an interesting perspective to further disentangle the co-occurrence patterns (Ovaskainen et al.

2017). Whatever the ultimate causes of species co-occurrence, the use of explicit hypothesis-testing frameworks as we have employed here sharpens the process of inference.

References

Aeschimann, D. and Burdet, H. M. 1994. Flore de la Suisse et des territoires limitrophes. Le nouveau Binz, 2nd ed. – Le Griffon, Neuchâtel.

Anderson, J. E. et al. 1996. Correlations between carbon isotope discrimination and climate of native habitats for diverse eucalypt taxa growing in a common garden. – Aust. J. Plant Physiol. 23: 311–320.

Araújo, M. B. and Rozenfeld, A. 2014. The geographic scaling of biotic interactions. – Ecography 37: 406–415.

Bar-Massada, A. 2015. Complex relationships between species niches and environmental heterogeneity affect species co-occurrence patterns in modelled and real communities. – Proc. R. Soc. B 282: 20150927.

Barton, N. H. and Hewitt, G. M. 1985. Analysis of hybrid zones. – Annu. Rev. Ecol. Syst. 16: 113–148.

Bertness, M. D. and Callaway, R. M. 1994. Positive interactions in communities. – Trends Ecol. Evol. 9: 191–193.

Bivand, R. and Lewin-Koh, N. 2014. maptools: tools for reading and handling spatial objects. – R package ver. 0.8-30.

Bivand, R. and Rundel, C. 2015. rgeos: interface to geometry engine – open source (GEOS). – R package ver. 0.3-15.

Bivand, R. et al. 2017. rgdal: bindings for the geospatial data abstraction. – R package ver. 1.2-7.

Blois, J. L. et al. 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.

Boulangeat, I. et al. 2012. Accounting for dispersal and biotic interactions to disentangle the drivers of species distributions and their abundances. – Ecol. Lett. 15: 584–593.

Bridle, J. R. and Vines, T. H. 2007. Limits to evolution at range margins: when and why does adaptation fail? – Trends Ecol. Evol. 22: 140–147.

Bull, C. 1991. Ecology of parapatric distributions. – Annu. Rev. Ecol. Syst. 22: 19–36.

Calabrese, J. M. et al. 2014. Stacking species distribution models and adjusting bias by linking them to macroecological models. – Global Ecol. Biogeogr. 23: 99–112.

Calenge, C. 2006. The package adehabitat for the R software: a tool for the analysis of space and habitat use by animals. – Ecol. Model. 197: 516–519.

Callaway, R. M. et al. 2002. Positive interactions among alpine plants increase with stress. – Nature 417: 844–848.

Carlsson, B. Å and Callaghan, T. V. 1991. Positive plant interactions in tundra vegetation and the importance of shelter. – J. Ecol. 79: 973–983.

Chamberlain, S. A. et al. 2014. How context dependent are species interactions? – Ecol. Lett. 17: 881–890.

Choler, P. et al. 2001. Facilitation and competition on gradients in alpine plant communities. – Ecology 82: 3295–3308.

Clark, J. S. et al. 2017. Generalized joint attribute modeling for biodiversity analysis: median-zero, multivariate, multifarious data. – Ecol. Monogr. 87: 34–56.

Connor, E. F. and Simberloff, S. 1979. The assembly of species communities: chance or competition? – Ecology 60: 1132–1140.

D'Amen, M. et al. 2015a. Using species richness and functional traits predictions to constrain assemblage predictions from stacked species distribution models. – J. Biogeogr. 42: 1255–1266.

D'Amen, M. et al. 2015b. Spatial prediction of community structure: state of the art and future perspectives. – Biol. Rev. doi:10.1111/brv.12222

D'Amen, M. et al. 2017. Data from: Disentangling biotic interactions, environmental filters, and dispersal limitation as drivers of species co-occurrence. – Dryad Digital Repository, < http://dx.doi.org/10.5061/dryad.8mv11 >.

Darwin, C. 1869. On the origin of species by means of natural selection. – Murray, London.

de Bello, F. et al. 2012. Functional species pool framework to test for biotic effects on community assembly. – Ecology 93: 2263–2273.

Di Cola, V. et al. 2017. ecospat: an R package to support spatial analyses and modeling of species niches and distributions. – Ecography 40: 774–787.

Diamond, J. M. 1975. Assembly of species communities. – In: Cody, M. L. and Diamond, J. M. (eds), Ecology and evolution of communities. Harvard Univ. Press, pp. 342–344.

Dray, S. et al. 2007. The ade4 package-II: two-table and K-table methods. – R News 7: 47–52.

Dubuis, A. et al. 2011. Predicting spatial patterns of plant species richness: a comparison of direct macroecological and species stacking modelling approaches. – Divers. Distrib. 17: 1122–1131.

Dubuis, A. et al. 2013. Predicting current and future spatial community patterns of plant functional traits. – Ecography 36: 1158–1168.

Efron, B. 2005. Bayesians, frequentists, and scientists. – J. Am. Stat. Assoc. 100: 1–5.

Gaston, K. J. 2003. The structure and dynamics of geographic ranges. – Oxford Univ. Press.

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

Gotelli, N. J. and Graves, G. R. 1996. Null models in ecology. – Smithsonian Inst., Washington.

Gotelli, N. J. and McCabe, D. J. 2002. Species co-occurrence: a meta-analysis of Diamond's assembly rules model. – Ecology 83: 2091–2096.

Gotelli, N. J. and Ulrich, W. 2010. The empirical Bayes approach as a tool to identify non-random species associations. – Oecologia 162: 463–477.

Gotelli, N. J. and Ulrich, W. 2012. Statistical challenges in null model analysis. – Oikos 121: 171–180.

Gotzenberger, L. et al. 2012. Ecological assembly rules in plant communities – approaches, patterns and prospects. – Biol. Rev. 87: 121–127.

Graves, G. R. and Gotelli, N. J. 1993. Assembly of avian mixed-species flocks in Amazonia. – Proc. Natl Acad. Sci. USA 90: 1388–1391.

Guisan, A. and Rahbek, C. 2011. SESAM – a new framework integrating macroecological and species distribution models for predicting spatio-temporal patterns of species assemblages. – J. Biogeogr. 38: 1433–1444.

Guisan, A. et al. 2017. Habitat suitability and distribution models. – Cambridge Univ. Press.

Gurevitch, J. et al. 1992. A meta-analysis of field experiments on competition. – Am. Nat. 140: 539–572.

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

Hautier, Y. et al. 2009. Changes in reproductive investment with altitude in an alpine plant. – J. Plant Ecol. 2: 125–134.

He, Q. et al. 2013. Global shifts towards positive species interactions with increasing environmental stress. – Ecol. Lett. 16: 695–706.

Hewitt, G. M. 1988. Hybrid zones? Natural laboratories for evolutionary studies. – Trends Ecol. Evol. 3: 58–67.

Hijmans, R. J. 2015. raster: geographic data analysis and modeling. – R package ver. 2.3-40.

HilleRisLambers, J. et al. 2012. Rethinking community assembly through the lens of coexistence theory. – Annu. Rev. Ecol. Evol. Syst. 43: 227–248.

Hirzel, A. and Guisan, A. 2002. Which is the optimal sampling strategy for habitat suitabilitymodelling. – Ecol. Model. 157: 331–341.

Hui, F. K. C. 2016. Boral – Bayesian ordination and regression analysis of multivariate abundance data in r. – Methods Ecol. Evol. 7: 744–750.

IUCN 1994. IUCN Red List categories. – International Union for the Conservation of Nature, Species Survival Commission, Gland, Switzerland.

Jaeger, R. G. 1970. Potential extinction through competition between two species of terrestrial salamanders. – Evolution 24: 632–642.

Lopes, G. N. et al. 2015. Temporal overlap and co-occurrence in a guild of sub-tropical tephritid fruit flies. – PLoS One 10: e0132124.

Lyons, S. K. et al. 2016. Holocene shifts in the assembly of plant and animal communities implicate human impacts. – Nature 529: 80–83.

Michalet, R. 2006. Is facilitation in arid environments the result of direct or complex interactions? – New Phytol. 169: 3–6.

Miller, R. S. 1967. Pattern and process in competition. – Adv. Ecol. Res. 4: 1–74.

Morueta-Holme, N. et al. 2016. A network approach for inferring species associations from co-occurrence data. – Ecography 39: 1–12.

Normand, S. U. et al. 2009. Importance of abiotic stress as a range-limit determinant for European plants: insights from species responses to climatic gradients. – Global Ecol. Biogeogr. 18: 437–449.

Oksanen, J. et al. 2014. vegan: community ecology package. – R package ver. 2.2-0.

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. 2017. How to make more out of community data? A conceptual framework and its implementation as models and software. – Ecol. Lett. 20: 561–576.

Peres-Neto, P. R. et al. 2001. Environmentally constrained null models: site suitability as occupancy criterion. – Oikos 93: 110–120.

Phillips, O. L. et al. 2003. Habitat association among Amazonian tree species: a landscape-scale approach. – J. Ecol. 91: 757–775.

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.

Pottier, J. et al. 2013. Accuracy of plant assemblage predictions from species distribution models varies along environmental gradients. – Global Ecol. Biogeogr. 22: 52–63.

Schluter, D. 1984. A variance test for detecting species associations, with some example applications. – Ecology 65: 998–1005.

Schmid, S. 2014. Assessing spatialautocorrelation in biodiversity data in the western Swiss Alps. – First step project of Master in Science in Behaviour, Evolution and Conservation. Univ. of Lausanne, Switzerland.

Sih, A. 1984. The behavioral response race between predator and prey. – Am. Nat. 123: 143–150.

Simberloff, D. and Connor, E. F. 1981. Missing species combinations. – Am. Nat. 118: 215–239.

Stone, L. and Roberts, A. 1990. The checkerboard score and species distributions. – Oecologia 85: 74–79

Thibaud, E. et al. 2014. Measuring the relative effect of factors affecting species distribution model predictions. – Methods Ecol. Evol. 5: 947–955.

Thuiller, W. et al. 2009 BIOMOD – a platform for ensemble forecasting of species distributions. – Ecography 32: 369–373.

Ulrich, W. 2010. Pairs – a FORTRAN program for studying pair-wise species associations in ecological matrices. – <www.keib.umk.pl/pairs/?lang=en>.

Veech, J. A. 2014. The pairwise approach to analysing species co-occurrence. – J. Biogeogr. 41: 1029–1035.

Vetaas, O. R. 2002. Realized and potential climate niches: a comparison of four *Rhododendron* tree species. – J. Biogeogr. 29: 545–554.

Wagner, H. H. and Dray, S. 2015. Generating spatially-constrained null models for irregularly spaced data using Moran spectral randomization methods. – Methods Ecol. Evol. 6: 1169–1178.

Warton, D. I. et al. 2015. So many variables: joint modeling in community ecology. – Trends Ecol. Evol. 30: 766–779.

Wisz, M. et al. 2013. The role of biotic interactions in shaping distributions and realized assemblages of species: implications for species distribution modelling. – Biol. Rev. 88: 15–30.

Zhang, L. et al. 2011. Strong plant–soil associations in a heterogeneous subtropical broad-leaved forest. – Plant Soil 347: 211–220.

Supplementary material (Appendix ECOG-03148 at <www.ecography.org/appendix/ecog-03148>). Appendix 1–4.