Trait matching without traits: using correspondence analysis to analyze the latent structure of interaction networks

Lisa Nicvert[1], Hervé Fritz[2,3], & Stéphane Dray[1]

[1] Universite Claude Bernard Lyon 1, LBBE, UMR 5558, CNRS, VAS – Villeurbanne, France
[2] REHABS, International Research Laboratory, CNRS-NMU-UCBL, Nelson Mandela University – George, South Africa
[3] Sustainability Research Unit, Nelson Mandela University – George, South Africa

Correspondence: lisa.n@orange.fr

Abstract

Species interactions occurring in ecological communities are often represented as networks. The structure of these networks is thought to be linked to species' interaction niches (or Eltonian niches), which are intimately related to the notion of trait matching, which posits that a species interacts preferentially with partners whose traits are complementary to their own.
Multivariate methods are commonly used to quantify species environmental niches (or Grinnellian niches). More recently, they have also been used to study the interaction niche, but these methods consider only the niche optimum and require trait data.
In this paper, we use the correspondence analysis (CA) framework to study interaction networks and investigate trait matching without requiring traits data, using the notion of latent traits. We use reciprocal scaling, a method related to CA, to estimate niche optima and breadths, defined respectively as the mean and standard deviation of latent traits of interacting partners in CA space. We present the method, test its performance using a simulation model we designed, and analyze a real mutalistic bird-fruits interaction network.
The simulation study shows that the method is able to recover niche breadths and optima for data generated with parameters values typical of ecological networks. The bird-fruit network analysis shows strong relationships between species niche optima and niche breadths: a posteriori correlation with species traits suggest that this latent structure is related to measured traits. In this network, birds and plants of intermediate size tend to have the widest niches and birds with pointed wings (preferentially foraging in the canopy) have smaller niches than birds with rounded wings (preferentially foraging in the understory).
CA and reciprocal scaling are described as fruitful methods to characterize species interaction profiles, provide an ecologically meaningful graphical representation of interaction niches and allow to explore the effect of latent traits on network structure.

Keywords: Interaction niche; Trait matching; Correspondence analysis; Reciprocal scaling; Ecological network

Introduction

The ecological niche constitutes the pool of environmental conditions and resources required for species persistence (Vandermeer, 1972). These factors can be abiotic (e.g., temperature, precipitation, light) and summarized with the notion of Grinnellian niche (Grinnell, 1924), which centers on the physical and environmental requirements essential for a species' survival. The niche can also be seen through the lens of species' interactions with other organisms (Eltonian niche) originally formalized in the context of food webs by Charles Elton (Elton, 1927) but that can be extended to other types of interactions like competition and symbiosis. By encapsulating the ecological requirements and constraints of a species, the ecological niche is a crucial concept contributing to our understanding of species distributions, coexistence or exclusion.

Hutchinson (1957) famously formalized the niche concept as an n-dimensional hypervolume. One common way to describe these hypervolumes is to approximate the niche space with Gaussian curves along each niche axis (Gauch Jr. and Whittaker, 1972) to estimate species niche optimum and breadth along environmental gradients (as mean and standard deviation parameters of a multivariate normal distribution). Niche optimum describes the conditions where species growth is maximized, whereas niche breadth defines the range of conditions tolerated by a species: species that have a wide niche are often called generalists and species that have a narrow niche are called specialists. The factors influencing niche breadth are still being investigated (Devictor et al., 2010; Sexton et al., 2017).

In this paper, we focus on species interaction (or Eltonian) niches, i.e. we characterize species interactions with other species in the community. We also focus on the niche effectively occupied by a species, i.e. its realized niche, by opposition to its fundamental niche, which is the range of theoretical interaction partners of a species constrained by physiological or morphological characteristics (Devictor et al., 2010; Hutchinson, 1957). The realized niche has often been defined as the fundamental niche restricted to a narrower range by biotic interactions with other species in the community (Futuyma and Moreno, 1988). However, this characterization has environmental niches in mind and here, we use a broader definition of the realized niche as the range of effective interacting partners of a species, affected by a range of processes including for instance competition with other species but also species abundances.

Methods used to describe and analyze interaction niches often use the formalism of networks (Bascompte and Jordano, 2007; Ings et al., 2009). Networks are mathematical objects composed of a set of nodes (species) which are linked by edges (interactions). Ecological networks generally display non-random structures (Bascompte, Jordano, et al., 2003; Ings et al., 2009; Jordano et al., 2003) such as nestedness, where interactions of specialist species tend to be subsets of interactions of more generalist species (Bascompte, Jordano, et al., 2003). Various factors can influence the structure of ecological networks (Ings et al., 2009; Peralta et al., 2024; Vázquez et al., 2009). Neutral effects imply that a given species is more likely to interact with abundant species than with the rarest ones (Vázquez et al., 2009). Variations in abundance, influenced for instance by environmental conditions or species phenology, can thus induce structures in ecological networks. Species evolutionary history is also thought to influence their interaction patterns (Benadi et al., 2022). Trait matching, wherein species with complementary traits preferentially engage in interactions, is also a well studied mechanism: for instance, body size has been identified as a key factor influencing food web structure (Elton, 1927; Ings et al., 2009). Importantly, these processes are not mutually exclusive and often work together to shape the observed network patterns.

Regarding environmental (or Grinnellian) niches, multiple methods have also been developed to study them. In particular, weighted averaging (Whittaker, 1956) is a very simple method that estimates species optimum for a single environmental gradient by averaging the values of the environmental variable over the samples in which a species occurs (weighted by species abundance). Weighted averaging is at the basis of several popular multivariate techniques routinely used when the data at hand are species abundance (or presence) sampled across different sites. These methods include reciprocal averaging (Hill, 1973) (a.k.a. correspondence analysis (CA) Hill, 1974), to estimate latent gradients when no environmental data are avail-

able; canonical correspondence analysis (ter Braak, 1986), when multiple environmental data are recorded; or fourth-corner analysis (Legendre et al., 1997), RLQ analysis (Dolédec et al., 1996) and double-constrained correspondence analysis (ter Braak, Šmilauer, et al., 2018), to link niche optima to species characteristics when traits data are also available. Note that these methods focus on species niche optima and ignore the aspects related to niche breadth (but see Thioulouse and Chessel (1992), who notably proposed reciprocal scaling as a method to estimate niche breadth in the context of CA using weighted variances).

More recently, several multivariate methods originally developed to study environmental niches have been applied to study trait matching. Indeed, as networks can be encoded with matrices, many methods applicable to species-by-sites matrices are readily applicable to networks (van Dam et al., 2021). For instance, RLQ and fourth-corner analyses have been used to describe trait matching by measuring the correlation between species traits of two sets of partner species weighted by the interaction matrix (Albrecht et al., 2018; Bender et al., 2018; Dehling, Töpfer, et al., 2014). Even if niche theory is usually not explicitly integrated in the study of ecological networks (Godoy et al., 2018; Phillips et al., 2020), it is clear that trait matching implicitly relies on the notion of interaction niche (Albrecht et al., 2018; Eklöf et al., 2013). Indeed, trait matching posits that a species with given traits will interact only with species with trait values lying in a range defining its interaction niche. The average trait value of their partners defines a species niche optimum, while their partners' trait diversity correspond to the species niche breadth. Niche breadth, which can be seen as the degree of generalization of a species, has often been defined as the number of interacting partners, potentially considering their abundance in the network (Devictor et al., 2010). Other studies have also defined niche breadth with respect to the traits of interacting species (functional niche, Dehling, Jordano, et al., 2016), considering multiple traits separately (Maglianesi et al., 2015) or simultaneously (Dehling, Jordano, et al., 2016; Dehling, Töpfer, et al., 2014).

Methods to analyze ecological networks have brought invaluable insights on trait matching and degree of specialization, but also have shortcomings. First, trait matching in ecological networks has mainly been investigated separately from interaction niches. However, trait matching can be seen through the lens of ecological niches (Eklöf et al., 2013): the term "matching" then refers to the alignment of a species' traits with their interacting partners' traits (as considered in several simulations models, e.g. Benadi et al., 2022; Fründ et al., 2016). Second, trait matching studies only consider the average relationship between species traits across the network. Hence, they generally ignore the variability around matching, which can be related to the concept of niche breadth. Lastly, methods to study trait matching require the collection of traits data, which has been recognized as a challenge (Vázquez et al., 2009), because it is tedious to collect multiple traits and identify a priori which ones are relevant drivers of interactions.

In this paper, we propose to use correspondence analysis (CA) (Hill, 1974; Hirschfeld, 1935), a method often used in ecology to study species' environmental niches, to study trait matching and specialization/generalization in interaction networks. As previously pointed out by Lewinsohn et al. (2006) and van Dam et al. (2021), CA can be useful to analyze interaction networks. In particular, here, we show how CA can be used to infer species' interaction niches from interaction network data. Scores produced by CA allow to ordinate species along gradients determined by their interacting partners, and we show that these scores estimate the optima of species interaction niches and can be used as proxies for unmeasured traits determining the occurrences of interactions. We also use reciprocal scaling (Thioulouse and Chessel, 1992) to estimate interaction niches breadths. Our framework also allows to measure the effect of trait matching without using traits data (using the notion of latent traits). Lastly, these methods also enable ecologically meaningful graphical representation of species interaction niches.

In this paper, we first present CA and reciprocal scaling and their interpretation in the context of interaction networks. Then, we evaluate the performance of our framework to estimate species niches (optimum and breadth) using simulated data. Finally, we analyze a real bird-fruit interaction network (Dehling, Bender, et al., 2021) to show how our framework can help to measure and represent interaction niches, highlight trait matching and explore the drivers of niche breadth.

Material and methods

All analyses were performed with R 4.4.1 (R Core Team, 2024) and are stored in a GitHub repository (https://github.com/LisaNicvert/RSnetwork) and will be archived on Figshare upon publication.

Notations

In this part, we consider a $r \times c$ matrix $\mathbf{Y} = [y_{ij}]$ representing the interactions between r resource species and c consumer species (either abundances or presences). We use the terms "resource" and "consumer" in accordance with the literature, as these terms describe broad categories that encompass diverse networks such as pollination, herbivory or parasitism: but as long the interaction network is bipartite, the two sets of species could belong to any category.

Let $\mathbf{P}$ be the table of relative frequencies ($\mathbf{P} = [y_{ij}/y_{++}]$, where $y_{++} = \sum_{i=1}^{r}\sum_{j=1}^{c} y_{ij}$ is the grand total of $\mathbf{Y}$). We define the weight matrices for resources (rows) $\mathbf{D_r} = \mathrm{diag}(\mathbf{r})$ and consumers (columns) $\mathbf{D_c} = \mathrm{diag}(\mathbf{c})$, where the vectors $\mathbf{r} = \mathbf{P}\mathbf{1_r} = [p_{1+}, \dots, p_{r+}]^\top$ and $\mathbf{c} = \mathbf{P}^\top\mathbf{1_c} = [p_{+1}, \dots, p_{+c}]^\top$ represent respectively the rows and columns marginal sums and $p_{i+} = \sum_{j=1}^{c} p_{ij}$ and $p_{+j} = \sum_{i=1}^{r} p_{ij}$.

Quantify trait matching

We consider two known traits $\mathbf{x}$ and $\mathbf{y}$ measured, respectively, on the resource and consumer species. A simple measure of trait matching is given by the fourth-corner statistic, which measures the correlation between the traits of consumers and resources, weighted by the interaction matrix:

$$\mathrm{cor}_{\mathbf{P}}(\mathbf{x}, \mathbf{y}) = \tilde{\mathbf{x}}^\top \mathbf{P} \tilde{\mathbf{y}} \tag{1}$$

where $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ contains normalized traits ($\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ are centered scaled versions of the traits, respectively using weights $\mathbf{D_r}$ and $\mathbf{D_c}$). When the traits are not known, indirect gradient methods such as correspondence analysis are useful to identify latent structures in ecological networks.

Correspondence analysis (CA)

Correspondence analysis (Hill, 1974; Hirschfeld, 1935) is a multivariate method used to analyze contingency tables. CA is based on the generalized singular value decomposition of the doubly centered frequency interaction matrix $\mathbf{P}$, $\mathbf{P_0} = \mathbf{P} - \mathbf{r}\mathbf{c}^\top$:

$$\mathbf{P_0} = \mathbf{U}\mathbf{\Delta}\mathbf{V}^\top \tag{2}$$

The matrices $\mathbf{U}$ and $\mathbf{V}$ are orthonormal with respect to weights $\mathbf{D_r}$ and $\mathbf{D_c}$ ($\mathbf{U}^\top\mathbf{D_r}\mathbf{U} = \mathbf{I}$ and $\mathbf{V}^\top\mathbf{D_c}\mathbf{V} = \mathbf{I}$) and contain the left (resource) and right (consumer) generalized singular vectors (respectively). $\mathbf{\Delta}$ is the diagonal matrix containing ordered singular values. The matrix of CA eigenvalues $\mathbf{\Lambda}$ is equal to $\mathbf{\Delta}^2$.

If we consider a given dimension k, rewriting the previous equation shows that CA finds singular vectors $\mathbf{u}_k$ and $\mathbf{v}_k$ so that the singular value δ_k is maximized:

$$\delta_k = \mathbf{u}_k^\top \mathbf{P_0} \mathbf{v}_k \tag{3}$$

Equation (3) is analogous to the trait matching equation (1) (except for the use of $\mathbf{P_0}$ instead of $\mathbf{P}$, which does not affect results if traits data are centered), where known traits are replaced with the generalized singular vectors $\mathbf{u}_k$ and $\mathbf{v}_k$. CA amounts to finding the singular vectors with maximal correlation: the square-root of a CA eigenvalue λ_k represents the absolute value of the correlation between $\mathbf{u}_k$ and $\mathbf{v}_k$ ($\sqrt{\lambda_k} = |\mathrm{cor}_{\mathbf{P}}(\mathbf{u}_k, \mathbf{v}_k)|$). Thus, scores $\mathbf{u}_k$ and $\mathbf{v}_k$ can be interpreted as latent traits, associated to resources (respectively, consumers), that maximizes trait matching. The eigenvalue λ_k gives an indication on the agreement of consumer and resource species scores on dimension k: if eigenvalues are high, it indicates that latent traits are able to explain well the probability of interaction.

CA assigns similar scores to species with the same interaction profile: in other words, two consumer species interacting with the same resources will be positioned nearby in the CA multivariate space (and vice versa for resource species). Species scores can thus be used to reorder rows and columns of the interaction matrix and highlight the structure of the network (Lewinsohn et al., 2006) (see Appendix C).

CA has the advantage to evaluate the effect of trait matching, without using traits data, using these latent scores. However, latent traits can be interpreted in a broader sense as proxies for unmeasured properties for species beyond their traits (for instance, as phylogenetic signal like in Benadi et al., 2022). To help the interpretation, it is possible to link latent scores identified by CA to known properties, such as traits or phylogeny, a posteriori.

Standards outputs of CA consist of two biplots where both consumers and resources are displayed. For that, we use the following transition formulas:

$$\mathbf{U}^\star = \mathbf{D_r}^{-1}\mathbf{P}\mathbf{V} \qquad \text{(resources)} \tag{4a}$$

$$\mathbf{V}^\star = \mathbf{D_c}^{-1}\mathbf{P}^\top\mathbf{U} \qquad \text{(consumers)} \tag{4b}$$

Equation (4a) expresses resources scores as a weighted mean of the scores of consumers they interact with. Reciprocally, Equation (4b) expresses consumers scores as a weighted mean of the scores of resources they interact with. The transformation described in Equations (4) is called weighted averaging.

A first biplot can be drawn by displaying the resources scores given by the latent trait ($\mathbf{U}$) and consumers scores given by weighted averaging ($\mathbf{V}^\star$). On this plot, a consumer is located at the centroid of the resource it uses, and a resource is located according to its latent traits.

On the second biplot, scores are inverted so that consumers are represented with scores $\mathbf{V}$ (given by the latent traits) and resources with scores $\mathbf{U}^\star$ (given by weighted averaging). On this plot, a resource is located at the centroid of the consumers that use it, and a consumer is located according to its latent traits.

The coordinates of $\mathbf{U}^\star$ and $\mathbf{V}^\star$ can be interpreted as the niche optima of resource (respectively, consumer) species in the interaction network, measured with its interacting partners. This interpretation of the interaction niche is conceptually similar to the process-related niche quantified with the method of Dehling and Stouffer (2018). However, here niche measures are defined with latent traits, and in the context of CA (as discussed by (Braak and Prentice, 2004) and (ter Braak and Verdonschot, 1995) for environmental niches).

With these scores, CA seems particularly adapted to the analysis of interaction networks. However, CA completely ignores the diversity of interacting partners (niche breadth), and the niche optima can only be represented on two different biplots.

Reciprocal scaling

Reciprocal scaling (Thioulouse and Chessel, 1992) is an extension of CA that solves the two issues mentioned above by providing niche optima and breadths of consumers and resources in a common multivariate space. Reciprocal scaling shifts the focus of the analysis by considering the interactions, rather than the species, as statistical individuals, using the interpretation of CA as a special case of canonical correlation analysis (for complete mathematical development, see Thioulouse and Chessel, 1992).

Reciprocal scaling defines a score $\mathbf{h}_k$ for the interactions in the matrix in the k-th dimension of CA: these interactions, represented by non-empty cells in $\mathbf{Y}$, are called correspondence by Thioulouse and Chessel (1992). $\mathbf{h}_k$ is of length $n_{\bar{0}}$, where $n_{\bar{0}}$ is the number of unique interactions in the matrix (i.e. non-empty cells). Interactions (or correspondences) scores can be computed using the CA scores for the k-th dimension:

$$h_k(i,j) = \frac{u^\star_{ik} + v^\star_{jk}}{\sqrt{2\lambda_k\mu_k}} \tag{5}$$

where λ_k is the CA eigenvalue for axis k and $\mu_k = 1+\sqrt{\lambda_k}$. Here, a double indexing is used for the elements of $\mathbf{h}_k$, which is a one-dimensional vector: $h_k(i,j)$ corresponds to the x-th element of $\mathbf{h}_k$ ($x = 1 \dots n_{\bar{0}}$), containing

the interaction between consumer i and resource j. Note that the score $h_k(i,j)$ is defined only when species i and j interact.

Using the interaction scores $h_k(i,j)$, it is possible to display the interaction niche of a consumer (or a resource) as a cloud of interactions in the multivariate space. The niche can then be summarized by estimating its optimum and its breadth: to do so, ellipses corresponding to the bivariate normal distribution summarizing species niches can be drawn in the multivariate plane (see Figure 1). If we consider species s in the plane given by axes k and l, the center of its ellipse represents the niche optima of species s on axes k and l (weighted averages of interaction scores), the semi-minor and -major axes represent its niche breadths in axes k and l (weighted standard deviations) and the orientation of the ellipse represents the weighted covariance for axes k and l. Moreover, we can show that the niche optima computed from these scores correspond to the scores $\mathbf{U}^\star$ and $\mathbf{V}^\star$ defined above, with a scaling factor (see Equations (D.1) and (D.2) in Appendix D). Formulas to compute niche optima, breadths and covariances are given in Appendix D (from Thioulouse and Chessel (1992) and adapted with the notation used in this article).

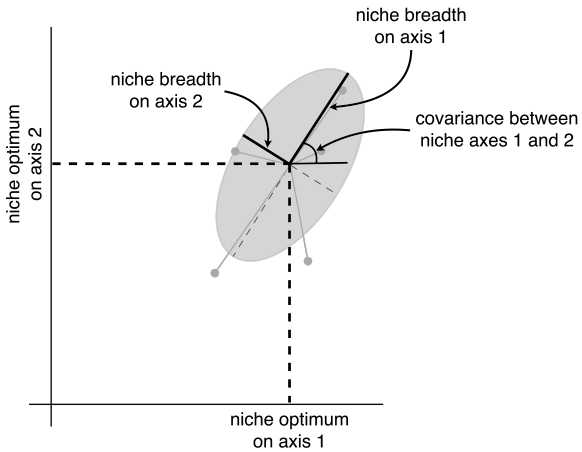


Figure 1. Visualization of the interaction niche in a 2D space. The niche of a species is depicted by an ellipse whose center represents niche optima and semi-axes represent niche breadths on each axis. The angle of the ellipse represents the covariance between niche axes.

Simulation

We simulated interaction data to evaluate the performance of CA and reciprocal scaling in estimating parameters of species interaction niches. We describe the model and the different experiments we conducted below.

Simulation model

Our simulation model takes into account species traits (via trait matching) and their abundances (via neutral effects) and combines these two processes to generate an observed interaction matrix (see Figure 2). This model is inspired from the models of Fründ et al. (2016) and Benadi et al. (2022) for the simulation of trait-based interaction networks, and of the models of Dray and Legendre (2008) and Minchin (1987) for niche modeling in the context of species/sites associations.

We consider the interactions between the i-th resource ($i = 1 \dots r$) and the j-th consumer species ($j = 1 \dots c$). The simulation procedure consists in different steps described below and illustrated in Figure 2:

Step (a): we simulate two traits for both consumer and resource species (Figure 2a). Resource species traits are stored in matrix $\mathbf{T^r} = [t^r_{ik}]$ ($r \times 2$) and consumer traits are stored in matrix $\mathbf{T^c} = [t^c_{jk}]$ ($c \times 2$). The first trait for both resources and consumers is drawn from a uniform distribution between 0 and 100, and the second trait is generated similarly but in the interval 15-85 (see length of the traits axes on Figure

2a), so that the second trait has less weight in driving matching. For consumers, we define niche breadth by associating the two columns vectors of $\mathbf{T^c}$ to those of $\mathbf{S} = [s_{jk}]$ (values of $\mathbf{S}$ are represented on Figure 2a as the standard deviation of a Gaussian curve). The elements of $\mathbf{S}$ contain the absolute values of numbers drawn from a normal distribution with mean μ_{breadth} and standard deviation σ_{breadth}.

Step (b): traits generated in the previous step are used to compute an interaction probability due to matching (Figure 2b). Species traits are assumed to define resources' and consumers' interaction niche with a bivariate normal distribution: the probability of interaction due to matching m_{ij} follows a normal distribution influenced by the difference between the trait values of resource species i and consumer species j ($t^c_{jk} - t^r_{ik}$) and the degree of generalization of consumer species (s_{jk}) as follows:

$$\mathbf{M} = [m_{ij}] = \frac{1}{2\pi s_{j1} s_{j2}} \exp\left(-\frac{(t^c_{j1} - t^r_{i1})^2}{2s^2_{j1}} - \frac{(t^c_{j2} - t^r_{i2})^2}{2s^2_{j2}} \right) \tag{6}$$

Step (c): we generate species abundances for consumers and resource species independently (Figure 2c) from a uniform or a log-normal distribution (see Table 1 for more details). These abundances are stored in vectors $\mathbf{n^c}$ ($c \times 1$) for consumers and $\mathbf{n^r}$ ($r \times 1$) for resources.

Step (d): we then compute a matrix of predicted interaction counts based solely on the abundances (Figure 2c):

$$\mathbf{A} = [a_{ij}] = \left[\frac{n^r_i}{\sum_{i=1}^r n^r_i} n^c_j \right] \tag{7}$$

Here, we model neutral effects where interactions are the result of consumer choices influenced only by the relative abundance of resources (representing the availability of resource species in the landscape).

Step (e): we compute a composite interaction probability resulting from the combined effects of trait matching ($\mathbf{M}$, equation (6)) and neutral abundance effects ($\mathbf{A}$, equation (7)) (Figure 2e). A factor δ ($0 \leq \delta \leq 1$) allows to control the relative weight of the matching and the abundance processes: if $\delta = 0$, then the strength of the matching is null, and interactions are driven only by abundances. If $\delta = 1$, abundances and matching concur to produce the observed pattern, and the strength of matching is maximal compared to other values of δ. The mixed interaction frequency matrix $\mathbf{A}^\star = [a^\star_{ij}]$ is computed as:

$$a^\star_{ij} = {m_{i|j}}^{\delta} a_{ij} \tag{8}$$

where $m_{i|j} = m_{ij}/m_{+j}$ represents the probability of interaction with resources per consumer species and $m_{+j} = \sum_{i=1}^r m_{ij}$ is the marginal probability of interaction for species j.

The interaction probability matrix $\mathbf{P}^\star = [p^\star_{ij}]$ is then computed from the abundance matrix:

$$p^\star_{ij} = \frac{a^\star_{ij}}{a^\star_{++}} \tag{9}$$

where $a^\star_{++}$ is the grand total of $\mathbf{A}^\star$.

Step (f): finally, we sample observed interactions counts from the interaction probability matrix $\mathbf{P}^*$ (Figure 2f). To do so, we sample n_{inter} interactions from a multinomial distribution with $\kappa = rc$ outcomes corresponding to the pairwise interactions. The probability vector $\mathbf{p}^\star$ (of length κ) corresponds to the flattened matrix $\mathbf{P}^\star$.

$$\mathbf{z} \sim \mathcal{M}_{\kappa=rc}(n = n_{\text{inter}}, p = \mathbf{p}^\star) \tag{10}$$

The interactions counts obtained after sampling in the vector $\mathbf{z}$ are finally rearranged in the matrix $\mathbf{Z}$ ($r \times c$) which corresponds to the sampled interaction network.

Simulation parameters

To evaluate the performance of our method in estimating parameters of species interaction niches, we simulated data under the model described above. We conducted 4 experiments to study the influence of 4 parameters (the detailed parameter values are presented in Table 1):

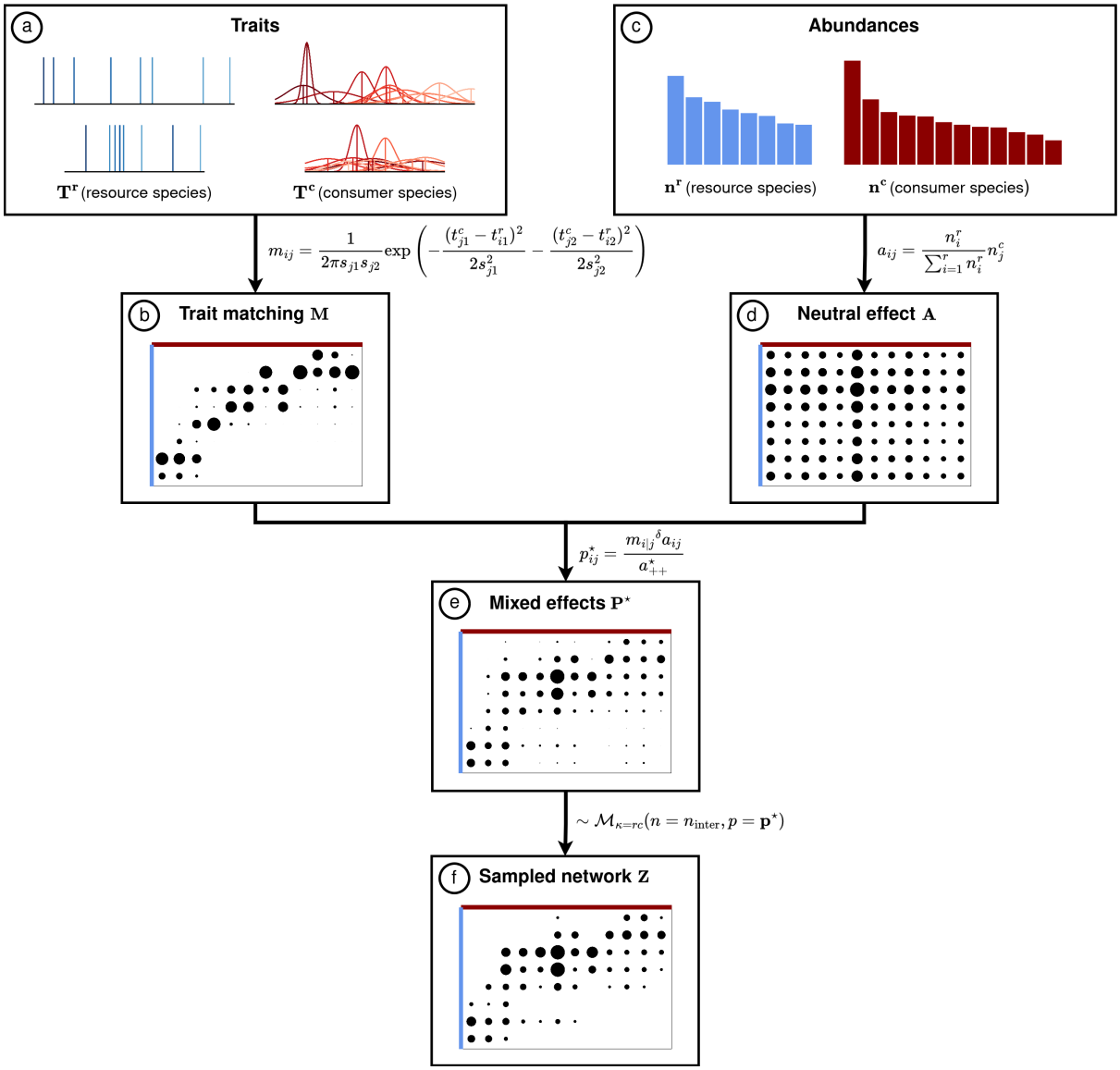


Figure 2. Model used to simulate interactions between consumer and resource species. (a) Two traits are generated for each species, corresponding to their niche optima. For consumers, standard deviations are also associated to traits to generate the niche breadths. (b) Interaction probabilities due to trait matching are computed from species traits. (c) In parallel, resource and consumers abundances are generated. (d) Interactions counts representing the neutral effect of abundances on interaction probability are computed from the abundances. (e) Matching and abundance-driven interactions are combined to get the interactions counts based on both processes. (f) Observed interactions are then sampled from the mixed interaction probabilities. Equations are explained in the main text.

1. Sampling (n_{inter}): we vary the total number of interactions in the matrix n_{inter}, which represents the sampling intensity and allows to evaluate the robustness to incomplete sampling.

2. Species abundance distributions ($\mathbf{n^r}$, $\mathbf{n^c}$): we sample $\mathbf{n^r}$ and $\mathbf{n^c}$ from a uniform distribution, and log-normal distributions either mildly, moderately- or very skewed. The aim is to evaluate the robustness of the method to skewed abundances typically encountered in ecological communities.

3. Average degree of specialization of consumers (μ_{breadth}): we vary the mean of the normal distribution in which the standard deviations of consumers traits s_{jk} ($k = 1, 2$) are sampled. This allows to contrast generalists consumers (high standard deviation) with specialized consumers (low standard deviation).

4. Heterogeneity of the degree of specialization (σ_{breadth}): we vary the standard deviation of the normal distribution in which the consumers traits' standard deviations s_{jk} ($k = 1, 2$) are sampled. This allows to contrast communities with homogeneous consumers (low standard deviation, same degree of specialization) to communities with heterogeneous consumers (high standard deviation, diverse degrees of specialization).

Each experiment described above studies the influence of one parameter by varying this parameter, and in experiment (μ_{breadth}) the standard deviation is also changed to keep a constant ratio of $\sigma_{\text{breadth}} = \mu_{\text{breadth}}/2$. Other parameters are fixed to a default value (see Table 1). For all simulations, we generated interactions between $r = 50$ resources and $c = 60$ consumers. We set the matching strength to $\delta = 0.2$, a value that was fixed empirically because it generated data where the relative strengths of abundance- and matching-driven processes seemed to both structure the observed matrix (results with other values of δ are presented in Appendix A).

Experiment	n_{inter}	Distribution of $\mathbf{n^r}$ and $\mathbf{n^c}$	μ_{breadth}	σ_{breadth}
(n_{inter})	200 500 5000 1000 10 000	$\text{Lognormal}(\mu = \ln(3), \sigma = \ln(1.5))$	10	5
($\mathbf{n^r}$, $\mathbf{n^c}$)	5000	$\mathcal{U}(a = 0, b = 100)$ $\text{Lognormal}(\mu = \ln(10), \sigma = \ln(1.5))$ $\text{Lognormal}(\mu = \ln(3), \sigma = \ln(1.5))$ $\text{Lognormal}(\mu = 0, \sigma = \ln(10))$	10	5
(μ_{breadth})	5000	$\text{Lognormal}(\mu = \ln(3), \sigma = \ln(1.5))$	2 10 20 50	1 5 10 25
(σ_{breadth})	5000	$\text{Lognormal}(\mu = \ln(3), \sigma = \ln(1.5))$	10	0.1 1 5 10 50

Table 1. Parameter values used for the simulation study. Experiment (n_{inter}) varies the total number of interactions. Experiment ($\mathbf{n^r}$,$\mathbf{n^c}$) progressively increases the skewness of the species abundances. Experiment (μ_{breadth}) changes the mean value of the consumers' niche breadth. Experiment (σ_{breadth}) changes the standard deviation of the consumers' niche breadth.

For each experiment, we generated 100 datasets. For each dataset, we computed the true niche optimum as the mean of the traits of the interacting partners of a species, weighted by their interaction frequency. Similarly, the true niche breadth is computed as the standard deviation of the traits of the interacting partners of a species weighted by the interaction frequency. These true values quantify species' realized niche (inference of the fundamental niche is discussed in Appendix B). Then, we performed reciprocal scaling on each dataset using the R package `ade4` (Thioulouse, Dray, et al., 2018) and obtained estimates of the realized niches optima and breadths with these methods. To evaluate the method's performance, we measured the correlation between the niches optima and breadths estimated by reciprocal scaling and the true values of the realized niche.

Real data analysis

Dataset

To illustrate CA and reciprocal scaling on real data, we used a bird-fruit interaction network from the ANDEAN frugivory dataset (Dehling, Bender, et al., 2021, network Peru1 from the dataset). In this dataset, consumers are bird species and resources are the fruiting plants they feed on and whose seeds they disperse. This network was sampled in the lower montane rainforest in Peru. Data were collected by direct repeated observation conduced throughout an entire year (4 times between 2009 and 2010) in 3 plots of 100 m × 30 m. A transect was used to determine focus fruiting plants inside each plot, and seed removal by birds were recorded on these plants.

Along with interactions, this dataset also includes some species traits thought to have an important role in trait matching processes, which we use a posteriori to interpret the analysis (although CA and reciprocal scaling do not require traits). Plant traits were collected in the field, and bird traits on museum specimens. We use three plant traits: fruit diameter, crop mass (mean number of fruits per plant multiplied by mean fruit mass) and plant height. For birds, we use three traits as well: bill width, body mass and Kipp's index, which is a measure of the pointedness of the wing: it is the Kipp's distance (distance from the tip of the first secondary to the wing tip) divided by wing length (Dehling, Töpfer, et al., 2014). A low Kipp's index indicates rounded wings, and a high Kipp's index indicates pointed wings.

Before data analysis, we filter out the birds or plants that interact only with one other species (8 birds and 12 plants). Indeed, since CA is based on species ordination, these species are problematic from a methodological point of view (Greenacre, 2013). These very specialized species can be seen as a part of an interaction module, as indicated in van Dam et al. (2021), who also advise analyzing modules separately. The final data has 40 plants and 53 birds (the list of these species, their traits and their corresponding codes can be found in Appendix E).

Analyses

To visualize species niches in the multivariate space and to compute niche breadths and optima, we performed CA and reciprocal scaling on this dataset, using the R package `ade4` (Thioulouse, Dray, et al., 2018).

To show how our method can answer ecological questions, we investigated the drivers of niche breadth using the latent variables given by the multivariate axes. In total, we modeled 4 relationships between niche breadths and optima obtained by reciprocal scaling (two axes for resources' and for consumers' niches). We considered linear and quadratic regression models and chose the best model with a likelihood ratio test.

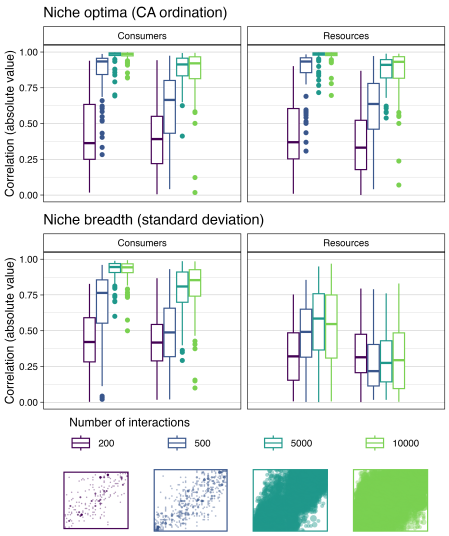
Finally, to interpret CA and reciprocal scaling more easily, we computed a posteriori correlations between the measured species traits and two first latent traits given by CA.

Results

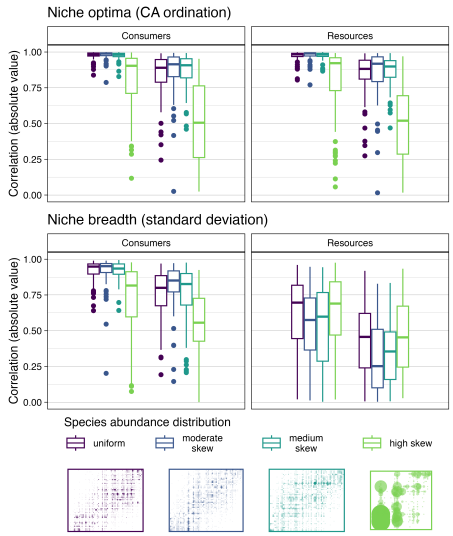
Simulation

We first carried out simulation studies to evaluate the capacity of CA and reciprocal scaling to infer species latent niches. We performed 4 experiments to test the influence of various parameters on the inference of niche parameters (see Table 1). The results are presented in Figure 3.

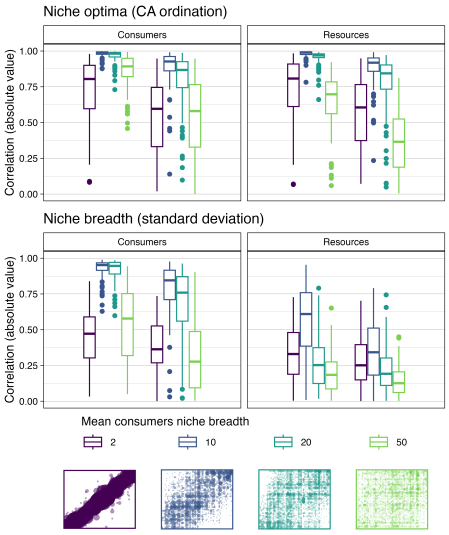
First, as expected, performance consistently improves with sampling completeness for all niche measures (Experiment (n_{inter}), Figure 3a). Second, when there are enough data, the niche optima are recovered correctly on both axes (median correlation value above 0.875), and niche breadths are recovered less precisely (Figure 3a). We can also note that for all experiments, niche parameters are always better recovered on the first axis, which is consistent with the fact that, by construction, trait matching with the first trait explains more structure. Finally, consumers niche breadths are always better recovered than resource niche breadths.



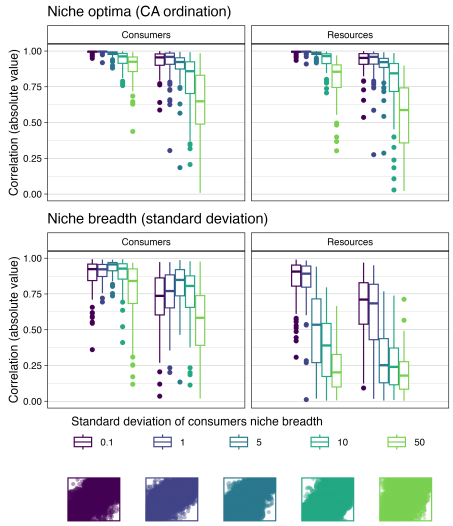
(a) Sampling intensity (n_{inter})



(b) Abundance distribution ($\mathbf{n}^{\mathbf{r}}$, $\mathbf{n}^{\mathbf{c}}$)



(c) Consumers niche breadth (μ_{breadth})



(d) Heterogeneity of consumers niche breadth (σ_{breadth})

Figure 3. Results of the simulation study. Each subfigure explores the influence of one parameter on the model performance. (a) Effect of sampling intensity (n_{inter}). (b) Effect of species abundance distributions ($\mathbf{n}^{\mathbf{r}}$, $\mathbf{n}^{\mathbf{c}}$). (c) Effect of the mean consumers niche breadths (μ_{breadth}). (d) Effect of heterogeneity of consumers niche breadths (σ_{breadth}). The y-axis is the absolute value of the correlation between true and estimated values of niche optima and niche breadths (respectively top and bottom of each subplot). Below the legend, matrices exemplify network data generated with the corresponding parameter value (with the size of points proportional to the number of observed interactions censored at the 1st and 99th percentiles).

Regarding the effect of abundance distribution (Experiment ($\mathbf{n}^{\mathbf{r}}$, $\mathbf{n}^{\mathbf{c}}$), Figure 3b), the model is robust to skewed abundance distributions, although performance drops a little with very skewed distributions.

Regarding consumer niche breadth (Experiment (μ_{breadth}), Figure 3c), the performance is better for niche breadth of intermediate size (the optimal performance is reached with a niche breadth of 10, representing 10% of the total length of the trait gradient in our setting).

Finally, regarding the heterogeneity of consumers' niche breadths (Experiment (σ_{breadth}), Figure 3d), homogeneous niche breadths improve the performance. In particular, the niche breadth of resources is correctly

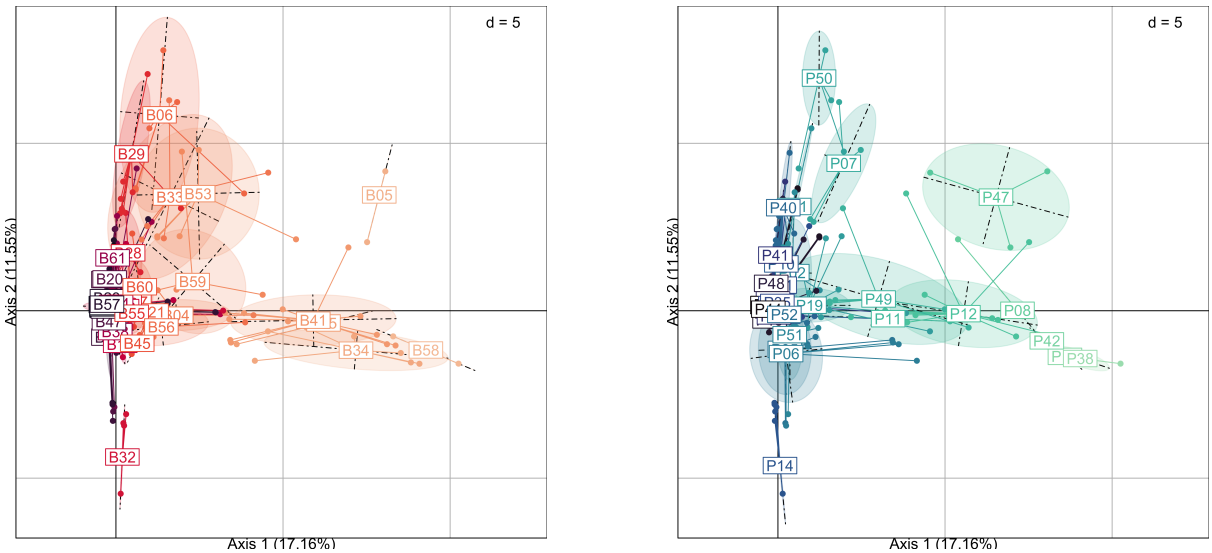
recovered only with very homogeneous consumer niches.

Data analysis

The permutation chi-squared test performed on the interaction table shows that there is a non-random structure in the network (χ^2 = 17784, 2000 permutations, p-value = 5.0×10^{-4}). The effect size computed with Cramér's V ($V = 0.29$, $IC_{95} = [0.26, 1.00]$), suggests a medium effect (Ellis, 2010).

CA eigenvalues suggest that matching between latent traits is quite strong. We recall that the absolute value of the correlation between latent traits is given by the square-root of the CA eigenvalues: here, for the first three eigenvalues, we have $\sqrt{\lambda_1} = 0.79$, $\sqrt{\lambda_2} = 0.65$ and $\sqrt{\lambda_3} = 0.63$.

With reciprocal scaling, we can position interactions in the multivariate space, and visualize species niches as the mean and variance of these interaction scores. On Figure 4, interactions (points) are grouped by interacting partner (either birds or plants). Figures 4a and 4b show interactions grouped by bird and plant species, respectively (here, we separate the two groupings to avoid a cluttered representation, but the two graphs show niches in the same multivariate space). The first 2 axes together explain 28.7% of the variance in the data (17.2% on axis 1 and 11.5% on axis 2). Eigenvalues suggest that the third axis also holds structure (10.9% variability), but here we do not show results on the third axis for concision.



(a) Bird niches **(b)** Plant niches

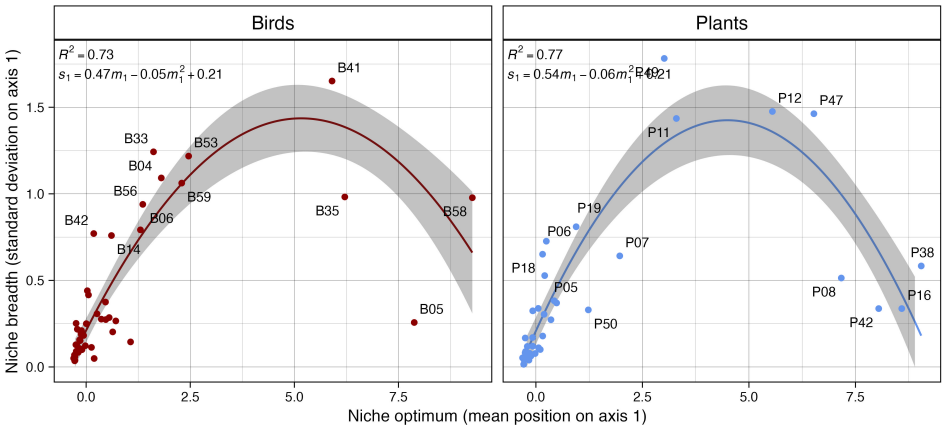
Figure 4. Reciprocal scaling of the birds-plants interaction network. Points correspond to the interaction scores $h_k(i, j)$ in the first 2 dimensions, grouped by bird (a) or plant species (b). Species label are placed at the reciprocal scaling mean, and ellipses correspond to the bivariate normal distribution with variances and covariances given by reciprocal scaling (with a scaling factor of 1.5, i.e. the ellipse axes lengths are equal to $1.5\sqrt{\lambda_k}$ on axis k, corresponding to around 67% of the points contained in the ellipse). Ellipses are colored according to the rank of their centroid on the first axis.

On these figures, two birds (or plants) whose means are located nearby have similar niche optima. Moreover, a bird and a plant that interact preferentially have their niche optima nearby. For instance, if we consider B05 (*Cephalopterus ornatus*) and P47 (*Symplocos arechea*), their interaction with each other represent respectively 66% and 33% of their interactions. Finally, the ellipses' semi-major and minor axes correspond to the niche width (standard deviation of interaction scores) along each multivariate axis (latent trait), and the inclination of the ellipse represents the covariance between the niche axes.

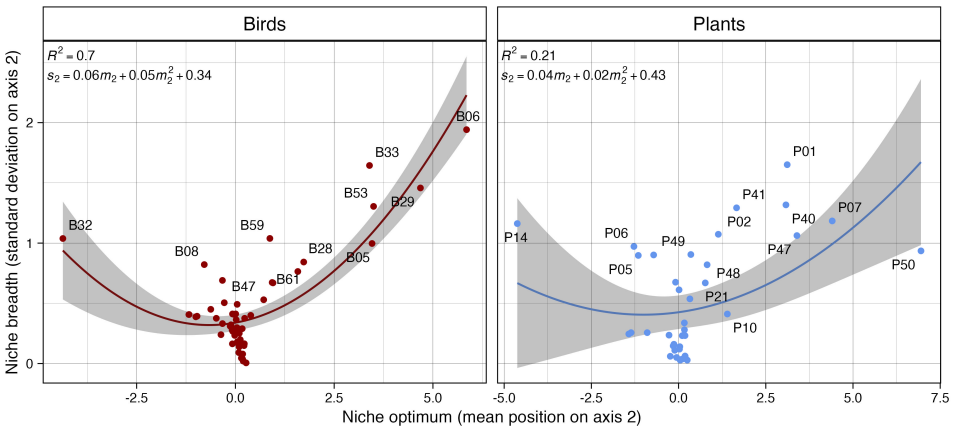
These figures also show how niches are distributed in the latent traits space. For birds (Figure 4a), many have a niche located around the origin (they interact with average plants, in the sense of their latent traits). Then, we can distinguish two clusters along the positive portion of the two main axes, which show two strate-

gies of birds along the two latent plant traits axes: either interact with plants with above-average values on axis 1, and small values on axis 2, or the reverse. This open triangle is completed by a few bird species around the diagonal (e.g. B33 *Penelope montagnii*, B53 *Thraupis bonariensis* and B59 *Turdus chiguanco*) that seem to have an intermediate niche optimum and a high niche breadth.

For plants, many also have a niche optimum close to the origin (they interact with average birds, in the sense of their latent traits). As for birds, two clusters emerge along the two main axes, which also reflects two plant strategies. However, contrary to birds, we cannot clearly see plant species around the diagonal, except P47 (*Symplocos arechea*) that stands out against the two clusters. Species further from the origin also seem to have wider niches for both birds and plants; but the cluttered niches around the origin make it hard to investigate this pattern on these plots.



(a) Axis 1



(b) Axis 2

Figure 5. Niche breadth versus niche optimum on the 2 multivariate axes. Niche breadths (standard deviations) and niche optima (means) were computed with reciprocal scaling. The top graphs (a) shows the relationship on axis 1 and the bottom graphs (b) on axis 2. The left and right panels represent bird and plant species, respectively. The solid line is the prediction of the best linear model and the gray shading represents the 95% confidence interval around the predicted mean value. The coefficient of determination (R^2) and the linear model equation are shown in the top left corner.

To better visualize the relationship between axis position and niche breadth, we plotted the relationship between these two variables for each axis in Figure 5. We also modeled these relationships as described in the material and methods (the models' equations are also written on the figures). On axis 1, there is a concave relationship with a positive linear component between niche optimum and niche breadth for birds, but the variance of the residuals increases along axis 1. For plants, there is a concave relationship only. On axis 2, we

have a convex relationship for birds, with a strong linear positive component (the convex form seems to be driven mainly by B32 (*Patagioenas plumbea*). For plants, we have a very weak relationship ($R^2 = 0.21$) with convex and linear positive components.

To interpret the latent trait axes, we correlated measured traits with plant and birds CA coordinates a posteriori. Figure 6 shows how traits are correlated with the multivariate axis and with each other. All traits correlate positively with axis 1, which strongly suggests a size effect along this axis. Axis 2 is characterized by its negative correlation with plant height, Kipp's index and crop mass. All plant traits are also strongly correlated with each other. For birds, body mass and bill width are highly correlated. Regarding cross-trophic level traits, Kipp's index is strongly positively correlated with crop mass and plant height.

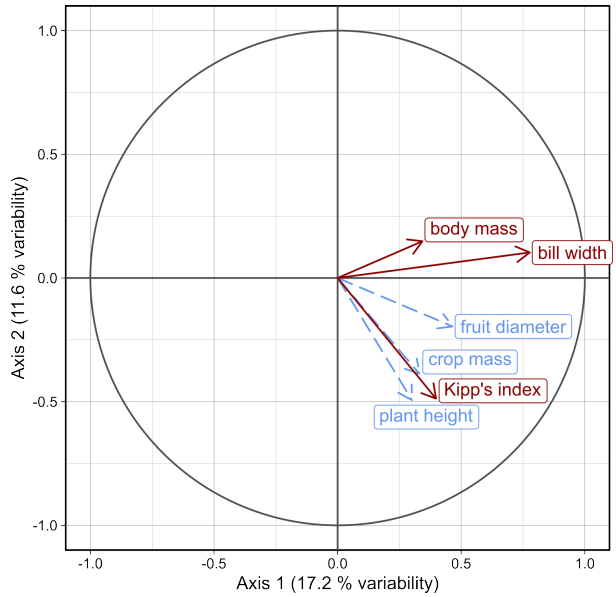


Figure 6. Correlation circle measured a posteriori between species traits and multivariate axes. Bird traits are shown in solid red lines and plant traits in dashed blue lines. Axis 1 shows a size effect while axis 2 is defined by the matching between plant height and Kipp's index.

Discussion

We evaluated CA and reciprocal scaling on both simulated and real data to quantify species interaction niches. We also showed how this approach can be used to test hypotheses on the drivers of species position on the specialist/generalist gradient with the case study.

Simulation

The simulation study shows good performance of the method to recover niche optima, but mixed performance regarding niche breadth. In particular, the niche breadths of resource species are poorly recovered overall, except in a few particularly favorable cases. Unsurprisingly, performance increases with sampling intensity (Figure 3a). When the proportion of variation explained by latent traits is smaller (axis 2), performance drops. The estimation is robust to skewed data, but performance decreases with extremely skewed data: in that case, we can hypothesize that the niches of rare species are estimated less precisely because they have very few observed interactions.

Consumers' niche breadth has an effect on the performance, which was expected because of the way CA orders individuals on a latent gradient. For instance, if we consider consumers, they are ordered so that two consumers interacting with similar resource species are positioned nearby. Hence, CA needs some overlap in interactions to position consumers on the gradient, but not too much or the signal weakens. Therefore, there

is a trade-off between narrow and wide niches: they need to overlap enough to order adjacent species, but not too much so that there is still a structure in the network (this trade-off is well reflected in Figure 3c).

The heterogeneity of consumers niche breadth decreases the performance of the estimation for both niche optima and niche breadth (Figure 3d). We can hypothesize that this is due to insufficient overlap with other species' interactions for specialist species, which might bias the placement of niche optima and the estimation of niche breadths for specialist species and their interacting partners.

Finally, niche parameters are better recovered for consumers than for resources, especially niche breadths. This might arise from the asymmetric way niche breadths are specified in our simulation model: indeed, niche breadths for resources are not controlled and driven entirely by the niche breadth of the consumers they interact with.

CA is able to recover a signal of matching rather accurately for any trait matching strength, which we can evaluate by performing chi-squared tests and examining eigenvalues (see Appendix A). Paradoxically, a stronger matching does not translate to a better recovery of the latent traits: indeed, if the matching is too strong, each species becomes increasingly separated of the others on the ordination, and CA is unable to find the common structuring driver without external information (see Appendix A).

Regarding the simulation model, it makes a number of simplifying hypotheses on interaction networks. First, in our model, only consumers have a preferred niche breadth. This makes sense in our case study with birds (as consumers) and plants (as resources), because birds' cognitive processes and movement abilities allow them to choose the plants with which they interact: but on an evolutionary timescale, plants might evolve to attract more diverse or more similar birds, thereby effectively evolving a preferred niche breadth. However, specifying a model taking into account both preferred niche breadths to determine matching (i.e. adapting Equation (6)) is difficult, because it requires to combine both niche breadths in the trait matching step. In fact, to our knowledge, no such model exists in the literature. The model also specifies trait matching with a multivariate normal distribution, which is a reasonable choice if the real-world interactions are at least symmetrical. Even though theory and observations support this hypothesis for environmental niches, especially for plants (e.g. Shelford's law of tolerance, Erofeeva, 2021; Shelford, 1931 and observations of Gauch Jr. and Whittaker, 1972), this hypothesis is debated for interaction networks (Benadi et al., 2022; Blonder et al., 2014). Despite these limitations, we think that the simulation model used here is useful to provide a first evaluation of the method's ability to recover niche parameters, before applying it to real data.

Data analysis

Regarding real data analysis, a permutation chi-squared test suggests that there is a non-random structure in the network. However, the network is also very noisy, as suggested by visual examination of the network and a medium effect size (Cramér's $V = 0.29$, $IC_{95} = [0.26, 1.00]$). This non-random structure is quantified as interaction counts that deviate from their expected value in the sense of the chi-squared statistic. However, to assess the expected number of interactions, the chi-squared statistic uses the matrix margins, thus assuming that the number of times a species is observed in the network is a surrogate of its abundance. This hypothesis might not be valid for real data due to sampling effects: in particular, the interactions range of the most abundant species might be better sampled, thus artificially increasing their generality, and reciprocally under-sampling of rare species might artificially increase their specialization. This is a well-known problem highlighted in simulation studies (Blüthgen, Fründ, et al., 2008; Fründ et al., 2016). Finally, this network represents a subsample of species, and some species in the network were probably not sampled, which may affect the network structure.

In the results, we explored the niche patterns for the first two axes (or latent traits). We find relationships between niche optimum and niche breadth for birds and plants. There is a concave relationship between niche breadth and the first latent trait for birds and plants, and a convex relationship between birds niche breadth and the second latent trait for birds.

To interpret latent traits given by the multivariate axes, we correlated species coordinates on the multi-

variate axes (giving their latent traits) with their measured traits. The first axis is mostly correlated with birds and plants size. For birds (Figure 5a, left), there is a convex relationship with a positive linear component between niche breadth and axis 1, which suggests that the larger birds are, the wider their niches are, but that medium-sized birds have the widest niches. Here, a wide niche means a high diversity of fruit sizes eaten. To interpret this result biologically, we can hypothesize that birds are limited by beak size. Small-billed birds may not be able to grasp large fruits, while bigger-billed birds can swallow even small fruits (Wheelwright, 1985). However, birds that are too large do not eat very small fruits because they need to satisfy their high absolute energy expenses and cannot spend too much time and energy foraging, so they focus on fruits that are larger or found in higher abundance on the plant (crop mass) that minimize foraging time (Albrecht et al., 2018; Schoener, 1971). This limiting effect of energy requirements is weak: the graph rather shows a diversity of strategies for large birds, with for instance B34 (*Pharomachrus antisianus*) and B41 (*Rupicola peruvianus*) eating more diverse-sized fruits and B05 (*Cephalopterus ornatus*) and B58 (*Trogon personatus*) choosing fruits of more uniform sizes. In addition, the correlation between birds' body mass and indices of plant productivity such as crop mass or fruit diameter is weak (Figure 6). This weak effect of energy traits may be due to the fact that birds foraging on fruits is partially modulated by whether they are fruit-gulpers or pulp-feeders (Palacio et al., 2017).

For plants (Figure 5a, right), the convex relationship between niche breadth and axis 1 is clearer. It suggests that plants of intermediate size attract more diverse-sized birds, while small fruits are only consumed by small birds, and large plants are reserved to big birds. To characterize plants with intermediate coordinates and wide niches on the first axis, we examine the reciprocal scaling plot 4b. We can see that P11 (*Elaeagia mariae*) and P49 (*Turpinia occidentalis*) predominantly interact with small and medium-sized birds, while P12 (*Endlicheria sp.*) and P47 (*Symplocos arechea*) interact with medium to large birds. A closer examination of their respective traits reveal that the large size of P11 and P49 on the first axis is predominantly driven by their height, so it does not impose a harsh barrier for interactions. On the contrary, P12 and P47 produce large fruits, in particular larger than the mean bird bill size, so the interaction barrier for small birds is stronger.

On axis 2, we interpret the relationship for birds only, because the relationship between the latent trait and niche breadth for plants is very weak. Figure 5b (left) shows a concave effect with a large positive linear component. In this discussion, we will focus on this linear component, because the convex form seems to be driven by a single species (*Patagioenas plumbea* B32). The second axis is negatively correlated to Kipp's index (wing pointedness) and plant height: birds with pointed wings have smaller niches on the second axis, i.e. they consume fruits from plants of homogeneous heights, while birds with rounded wings have wider niches (i.e. they consume fruits from plants of heterogeneous heights). We can interpret these results as preferences driven by foraging space. Indeed, rounded wings confer more maneuverability in dense vegetation (Thiel et al., 2023), so birds with rounded wings primarily feed on small plants from the dense understory. On the contrary, birds with pointed wings preferentially forage on tall plants from the more open canopy. Here, we find a higher specialization for birds with a high Kipp's index (pointed wings). We hypothesize that these birds might find it difficult to navigate dense understory where vegetation is denser, while birds with rounded wings might move to the less dense canopy at a lesser cost. It is also possible that top canopy bird focus on highly productive fruit patches in the canopy by choosing plants with high crop yields (as suggested by the high correlation between crop mass and Kipp's index in Figure 6), whereas birds with rounded wings also supplement their diet with other food sources such as insects (Foster, 1990; Thiel et al., 2023), which are not accounted for in this analysis. The higher specialization of canopy birds seems in contradiction with the findings of Schleuning et al. (2011): however, in this study, the authors measure specialization with an index considering the diversity of interacting partners, and not the diversity of their (latent) traits.

Finally, we must keep in mind that we observe realized niches here, so these niches might not reflect the "true" fundamental preference of birds. In Appendix B, we explore the performance of CA and reciprocal scaling to infer species' fundamental niches. We show that, in our simulation setting, the method can also recover fundamental niche parameters, although not as well as the realized niches. The realized niches of

birds and plants and might be constrained by two external factors. First, the distribution of species traits in the community might constrain species to settle for suboptimal interacting partners. For instance, here most plants species yield small fruits and have a small crop mass, so large birds might be constrained to eat smaller fruits. Second, competition may also constrain species to differentiate their interaction niche from other species in the community.

Although the analysis produces sensible results, the two axes discussed above explain roughly 30% of the variability, so these patterns cover only a small fraction of the non-random network structure. In particular, discussing patterns on the third axis would be important, as it explains almost the same variability as the second axis. Moreover, the eigenvalues' distribution shows a gap between the 3rd and 4th and the 6th and 7th eigenvalues, thus suggesting that underlying structures are found up to the 6th axis: considering the first 6 axes explains more than 60% of the variation of the network.

Conclusion and perspectives

Reciprocal scaling provides a quantitative measure of the interaction niche: species niche optima are approximated using the centroid of their interactions in the multivariate space, and species niche breadths using their standard deviation. The resulting niche is then a hyperellipsoid in n dimensions, corresponding to the isocontours of a multivariate normal distribution in the same hyperspace. This definition of the niche bears a strong parallel with Hutchinson's definition as a hypervolume (Hutchinson, 1957).

CA and reciprocal scaling could also be extended to measure niche overlap by quantifying the ellipses overlapping area (similarly to Pappas and Stoermer, 1997). This would allow to gain insights regarding the similarity, or dissimilarity, of species niches, thereby potentially allowing to measure competition strength.

In this article, we use the parallel between environmental and interaction niches (Gravel et al., 2019), and apply CA and reciprocal scaling, two methods originally developed to study environmental niches, to study interaction niches. More generally, a rich conceptual and methodological toolbox has been developed to study environmental niches and some of these tools could be or have been fruitfully be applied to the study of species interactions (Dehling, Lai, et al., 2024; Godoy et al., 2018; Gravel et al., 2019).

Other methods to quantify interaction niches have been proposed in the literature, in particular for niche breadths. These niche breadths are usually quantified using the identity of interacting partners (e.g. the index d', which also takes into account species availability, proposed by Blüthgen, Menzel, et al., 2006), as opposed to their traits. A notable exception to this trend is the process-related niche proposed by Dehling and Stouffer (2018), which measures the functional diversity of a species' interacting partners using their traits projected in a multivariate space. Reciprocal scaling bears similarities with both methods, as it quantifies niche breadth from the interaction matrix only, but also uses their latent traits. Compared with these two approaches, our measure hence provides a complementary way to define niche breadth, that takes into account both species characteristics (using latent traits) and their availability (using species' relative frequencies in the niche breadth measure).

The latent approach of CA allows the analysis to be data-driven, suited for exploratory analyses and not limited by costly trait collection. That being said, when traits are present, it would also be interesting to constrain the multivariate axes with bird, plant or both trophic levels traits. Such constrained analyses have been performed with RLQ or fourth-corner in the context of network data (Albrecht et al., 2018; Bender et al., 2018; Dehling, Töpfer, et al., 2014). However, fourth-corner considers only pairs of traits (not multivariate associations) and does not produce information on species, whereas RLQ is covariance-based and does not allow to partition the variation explained by traits in interaction networks.

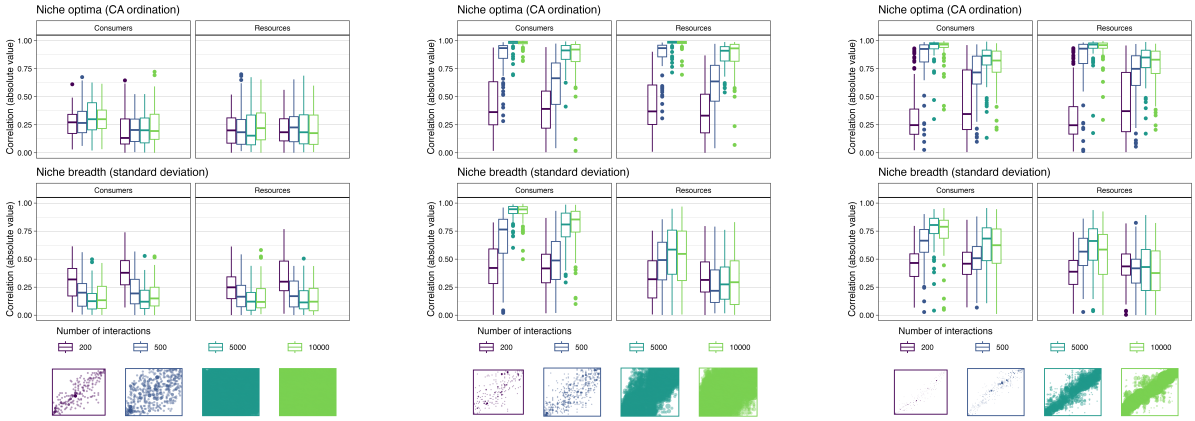
An alternative would be to use constrained versions of CA, like canonical correspondence analysis (using one trait table) (ter Braak, 1986) or double-constrained correspondence analysis (using both trait tables) (ter Braak, Šmilauer, et al., 2018). This would allow to 1) examine the relationship between measured traits of consumers and resources; 2) quantify the part of variation in networks that is explained by these traits and 3) quantify trait matching between measured traits (with double-constrained correspondence analysis). Moreover,

canonical methods would prevent the separation of specialized species on different axes when a common trait explains the network structure (results described in Appendix A and discussed above). To our knowledge, constrained versions of correspondence analyses have never been applied to interaction networks. Moreover, for all methods above, there is not a known way to measure niche breadth because reciprocal scaling is restricted to CA. Further methodological developments in multivariate analyses could therefore benefit network analysis.

Appendices

A Effect of the relative importance of matching on the recovery of niche parameters

Here, we investigate how reciprocal scaling performs when the network is more or less structured by trait matching compared to neutral processes. For that, we simulate networks with different values of the parameter δ from Equation (8): $\delta = 0$ (network structured only by neutral processes), $\delta = 0.2$ (intermediate strength of trait matching) or $\delta = 1$ (maximum strength of trait matching). For each value of δ, we generate networks with different sampling intensities (corresponding to experiment n_{inter} in the main text). Results are shown in Figure A.1 below.



(a) $\delta = 0$ **(b)** $\delta = 0.2$ **(c)** $\delta = 1$

Figure A.1. Effect of the relative importance of matching on the inference of niche parameters. Each subfigure shows the result of the inference for a different value of the trait matching strength δ. For a more complete description of the figure axes, panels and example matrices, refer to Figure 3 from the main text.

As expected, when there is no trait matching ($\delta = 0$), niche parameters are poorly recovered by CA and reciprocal scaling, because the true niche is random. Surprisingly, the performance is best when the matching strength is intermediate ($\delta = 0.2$), and decreases with a stronger matching ($\delta = 1$).

If we examine the associated eigenvalues in Figure A.2, we can see that, as expected, the matching signal is better recovered when $\delta = 1$. This plot shows the square root of eigenvalues (correspond to the absolute value of correlation between latent traits) for different values of matching strength δ and different sampling intensities. For a given sampling intensity, the eigenvalues are closer to one as δ increases. When $\delta = 0$, the (spurious) signal detected at low sampling intensity vanishes for all axes when the sampling improves. No matching is detected, so the model fails to recover niche parameters that are random and cannot be inferred from the matrix when $\delta = 0$. When $\delta = 0.2$, the two first eigenvalues become increasingly separated from other eigenvalues as sampling intensity increases. This means that two latent traits are inferred, that are well correlated to the two true traits used in the simulation, so the niche parameters are better recovered. When $\delta = 1$, we have high eigenvalues, but on all axes, thus indicating that the model recovers structuring latent traits on all axes. Therefore, the latent traits on the first two axes correspond less closely to the two true traits used in the simulation, thus yielding less accurate estimations of the niche parameters.

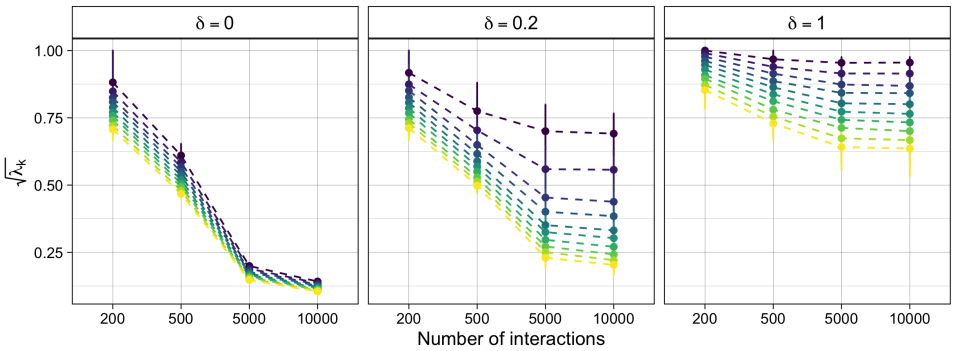
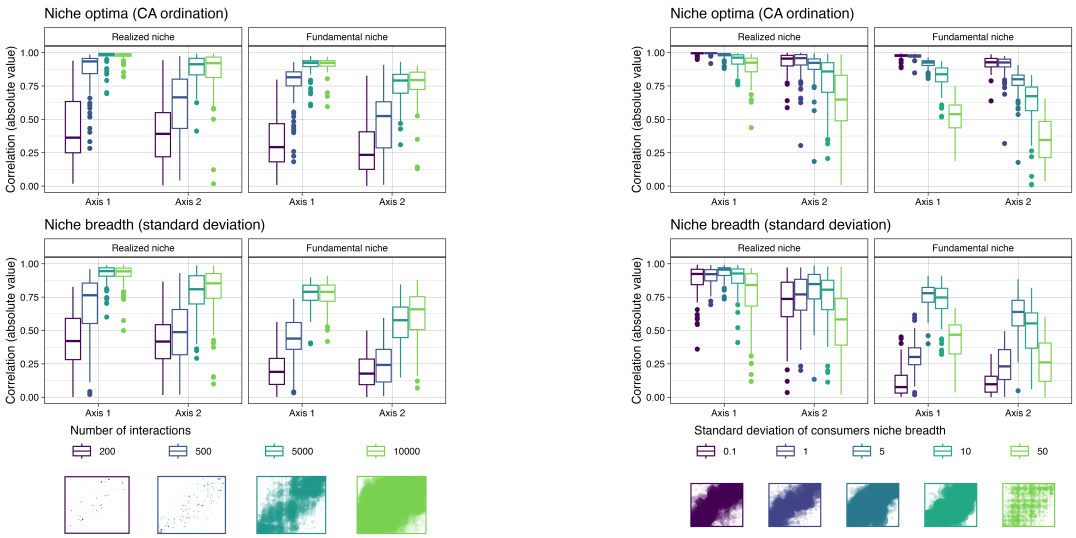


Figure A.2. Eigenvalues evolution with sampling intensity for different values of δ. This plot shows how eigenvalues evolve with sampling strength, for different values of δ (panels). Colored points and dotted line represent an eigenvalue, and error bars represent the 2.5 and 97.5-th percentiles on the eigenvalues over 100 repetitions.

B Realized and fundamental niches

In the article, we measured reciprocal scaling performance by comparing the inferred niche parameters to the realized niches. However, we could also compare the inferred niche parameters to the fundamental niches, i.e. the trait values (and niche breadths, for consumers) fixed in the model. Below, we present the performance of the model when we consider the fundamental niches as the ground truth (results are shown only for consumers). For simulation parameters, we choose the same parameters as described in 1.



(a) Sampling intensity **(b)** Heterogeneity of consumers niche breadth

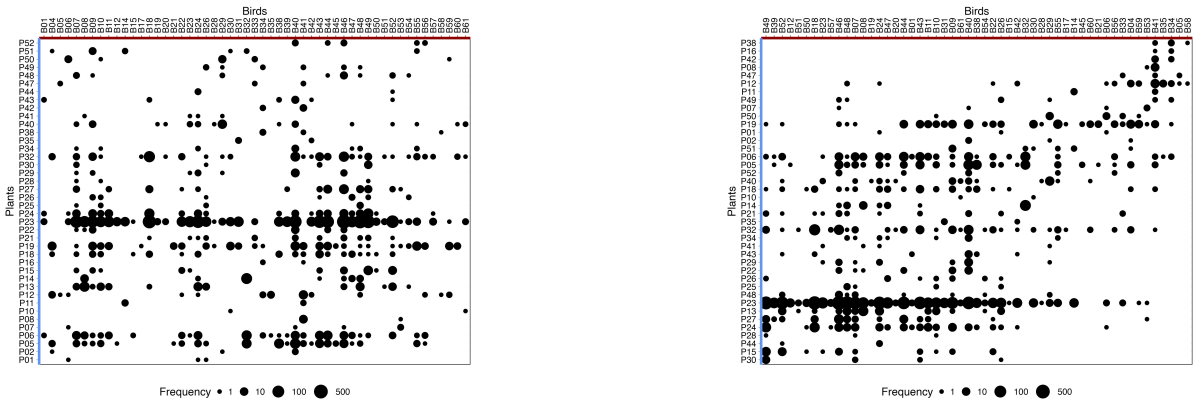
Figure B.1. Performance of the model to infer fundamental niche parameters for consumers. (a) Shows the effect of the sampling intensity and (b) shows the effect of the heterogeneity of consumers niche breadths on the model performance. y-axis and matrices have the same meaning as in Figure 3. The correlation is presented separately for realized and fundamental niches (respectively left and right panels) on axes 1 and 2.

In most cases, the performance of the model to recover fundamental niche is a reflection of the realized niche performance, but with slightly less good performance. As an example, we present the model performance for realized and fundamental niches depending on sampling intensity in Figure B.1a.

However, the performance pattern is different when the true niche breadth varies: whereas the realized niche is less well recovered as the standard deviation of niche breadth increases, the fundamental niche is better recovered for intermediate values of niche breadths (Figure B.1b). We can see it as an artifact of the performance metric used here: when niches breadths are very homogeneous, the model is able to recover the absolute niche breadths, but not their relative order because they are very homogeneous.

C Networks visualization using CA

Correspondence analysis can also be useful to reorder rows and columns of the analyzed interaction matrix. Many ways to visualize interaction networks have been proposed, with all of them highlighting a specific aspect of network structure like centrality, modularity or nestedness (Araujo et al., 2010; Miele et al., 2019). CA can also serve this purpose. The figure below shows the interaction matrix corresponding to the bipartite birds-plants interaction network analyzed in this article. In Figure C.1a, rows and columns are ordered according to the alphabetical order, while in Figure C.1b they are ordered according to their rank on the first axis of the CA. This means that species are positioned in an ecologically meaningful way, according to their interaction niche optima on the first latent trait axis. Here, we can see two weakly connected modules, corresponding to the bulk of small plant and bird species that interact together (bottom-left corner), and fewer large birds and plants that also interact together (top-right corner). A few species interact with species from both modules (like P12 (*Endlicheria sp.*) and B41 (*Rupicola peruvianus*)), and correspond to the generalist intermediate species from Figure 5.



(a) Raw matrix

(b) Reordered matrix

Figure C.1. Interaction matrix reordered using correspondence analysis. This figure shows how an interaction matrix can be reordered using CA. Rows and columns of matrix (a) are ordered alphabetically and for matrix (b) they are ordered following the coordinates of species on the first axis of CA.

D Reciprocal scaling formulas

Reciprocal scaling allows to compute niche optima and niche breadths on each axis k, as well as the covariance between two niche axes k and l. Each of these quantities can be expressed either with the scores $\mathbf{h}_k$ derived from reciprocal scaling or equivalently using the scores $\mathbf{u}_k^\star$ and $\mathbf{v}_k^\star$ of correspondence analysis, so each quantity has two equivalent expressions. The corresponding formulas are written below, adapted from Thioulouse and Chessel (1992) to match our notation.

Niche optima

Niche optima can be expressed either as a weighted mean of the correspondences scores, or from the CA scores. For the resource species i, we have:

$$m_{ik} = \begin{cases} \frac{1}{y_{i+}} \sum_{j=1}^{c} y_{ij} h_k(i,j) \\ \frac{\sqrt{\mu_k}}{\sqrt{2\lambda_k}} u_{ik}^\star \end{cases} \tag{D.1}$$

where $\mu_k = 1 + \sqrt{\lambda_k}$.

Similarly, for consumer species j, we have:

$$m_{jk} = \begin{cases} \frac{1}{y_{+j}} \sum_{i=1}^{r} y_{ij} h_k(i,j) \\ \frac{\sqrt{\mu_k}}{\sqrt{2\lambda_k}} v_{jk}^\star \end{cases} \tag{D.2}$$

Niche breadths

The niche breadth can be expressed as the weighted standard deviation of the correspondences scores or from the CA scores as well. For resource species i, we have:

$$s_{ik} = \begin{cases} \sqrt{\frac{1}{y_{i+}} \sum_{j=1}^{c} y_{ij} \left(h_k(i,j) - m_{ik}\right)^2} \\ \sqrt{\frac{1}{2\lambda_k \mu_k} \left(\frac{1}{y_{i+}} \sum_{j=1}^{c} \left(y_{ij} {v_{jk}^\star}^2 \right) - \lambda_k {u_{ik}^\star}^2 \right)} \end{cases} \tag{D.3}$$

Similarly, for consumer species j, we have:

$$s_{jk} = \begin{cases} \sqrt{\frac{1}{y_{+j}} \sum_{i=1}^{r} y_{ij} \left(h_k(i,j) - m_{jk}\right)^2} \\ \sqrt{\frac{1}{2\lambda_k \mu_k} \left(\frac{1}{y_{+j}} \sum_{i=1}^{r} \left(y_{ij} {u_{ik}^\star}^2 \right) - \lambda_k {v_{jk}^\star}^2 \right)} \end{cases} \tag{D.4}$$

Covariances

We can express the covariance between two niche axes from the reciprocal scaling of the CA scores as well. The covariance between niche axes k and l for resource species i is written:

$$c_{kl}(i) = \begin{cases} \frac{1}{y_{i+}} \sum_{j=1}^{c} y_{ij} h_k(i,j) h_l(i,j) - m_{ik} m_{il} \\ \frac{1}{2\sqrt{\lambda_k \lambda_l}\sqrt{\mu_k \mu_l}} \left(\frac{1}{y_{i+}} \sum_{j=1}^{c} y_{ij} v_{jk}^\star v_{jl}^\star - \sqrt{\lambda_k \lambda_l} u_{ik}^\star u_{il}^\star \right) \end{cases} \tag{D.5}$$

For consumer species j, the covariance is written as:

$$c_{kl}(j) = \begin{cases} \frac{1}{y_{+j}} \sum_{i=1}^{r} y_{ij} h_k(i,j) h_l(i,j) - m_{jk} m_{jl} \\ \frac{1}{2\sqrt{\lambda_k \lambda_l}\sqrt{\mu_k \mu_l}} \left(\frac{1}{y_{+j}} \sum_{i=1}^{r} y_{ij} u_{ik}^\star u_{il}^\star - \sqrt{\lambda_k \lambda_l} v_{jk}^\star v_{jl}^\star \right) \end{cases} \tag{D.6}$$

E Species names

Throughout the article, we use codes to describe species. Below are the correspondence between these codes and the species names, alongside with their traits (all data come from the ANDEAN frugivory dataset (Dehling, Bender, et al., 2021), network Peru1).

Code	Species name	Fruit diameter (mm)	Plant height (m)	Crop mass (g)
P01	*Anthurium obtusum*	2.9	5	0.8
P02	*Banara arguta*	9	5.2	241.5
P05	*Cecropia polystachya*	20.4	13	6813.4
P06	*Cecropia strigosa*	9.2	12.1	3813.1
P07	*Cestrum parqui*	5.9	3.9	470.8
P08	*Cestrum stipulatum*	6.9	4	225
P10	*Croton rusbyi*	6.5	6.8	7.2
P11	*Elaeagia mariae*	2.8	9.8	22.4
P12	*Endlicheria sp*	12.3	13.8	3986.1
P13	*Ficus americana*	3.7	15	606.7
P14	*Ficus coerulescens*	6.2	15	5400
P15	*Graffenrieda cucullata*	4.7	8.3	58.1
P16	*Guatteria duodecima*	17.4	13.3	7372.2
P18	*Guettarda crispiflora*	8.5	10.5	1448.6
P19	*Hedyosmum tepuiense*	7.7	5.3	583.5
P21	*Isertia laevis*	11.6	10.1	1100.5
P22	*Miconia affinis*	5	6.3	222.6
P23	*Miconia aprica*	3.2	7	333.7
P24	*Miconia barbeyana*	4	3.3	32.6
P25	*Miconia calvescens*	4.2	4.7	121.4
P26	*Miconia cyanocarpa*	3.6	4.1	416.3
P27	*Miconia egensis*	2.2	6.5	76.3
P28	*Miconia hygrophila*	3.3	4.2	18.4
P29	*Miconia longifolia*	4.2	12.3	293.3
P30	*Miconia myriantha*	2.7	7.3	7.9
P32	*Miconia sp3*	10.1	4	488.1
P34	*Miconia spennerostachya*	5.1	5.2	632.4
P35	*Myrcia fallax*	6.2	10.7	544.6
P38	*Ocotea sp*	8.5	12	850
P40	*Peristethium polystachyum*	3.7	3.5	139.4
P41	*Phoradendron chrysocladon*	3.2	7	16.4
P42	*Pleurothyrium trianae*	10.7	10.8	574.5
P43	*Rubus floribundus*	9.9	5.4	851
P44	*Schefflera mathewsii*	5.2	7.5	1951.3
P47	*Symplocos arechea*	12.6	9.8	3650.6
P48	*Trema micrantha*	2.2	9.1	1.5
P49	*Turpinia occidentalis*	8.3	7.2	235.9
P50	*Viburnum hallii*	6.1	5.5	249.3
P51	*Vismia gracilis*	10.4	9.5	1087
P52	*Zanthoxylum mantaro*	3.8	6.1	27.9

Table E.1. Correspondence between plant codes and species names.

Code	Species name	Kipp's index	Bill width (mm)	Body mass (g)
B01	*Anisognathus somptuosus*	0.19	10.6	42
B04	*Catharus ustulatus*	0.32	9.5	30.3
B05	*Cephalopterus ornatus*	0.17	30.2	380
B06	*Chiroxiphia boliviana*	0.16	9.2	17.2
B07	*Chlorochrysa calliparaea*	0.24	7.5	17
B08	*Chlorophonia cyanea*	0.28	6.5	14
B09	*Chlorospingus flavigularis*	0.17	10.1	25.6
B10	*Chlorospingus ophthalmicus*	0.16	7.7	15.7
B11	*Chlorospingus parvirostris*	0.16	8.5	24
B12	*Colaptes rubiginosus*	0.23	11.2	65.4
B14	*Cyanocorax yncas*	0.08	15	78.5
B15	*Dacnis cayana*	0.25	7.7	13
B17	*Diglossa cyanea*	0.17	7.2	17.1
B18	*Diglossa glauca*	0.19	5.4	12
B19	*Elaenia albiceps*	0.23	8.4	15.5
B20	*Elaenia pallatangae*	0.21	7.1	17.8
B21	*Entomodestes leucotis*	0.21	12.9	61.7
B22	*Eubucco versicolor*	0.14	12.6	32.8
B23	*Euphonia mesochrysa*	0.21	7.4	13
B24	*Euphonia xanthogaster*	0.22	7.4	13
B26	*Iridosornis analis*	0.16	8.4	26
B28	*Lophotriccus pileatus*	0.13	7.6	8.1
B29	*Mionectes striaticollis*	0.15	8.8	15
B30	*Myadestes ralloides*	0.21	12	29.1
B31	*Ortalis guttata*	0.15	13.5	550
B32	*Patagioenas plumbea*	0.35	8.1	179
B33	*Penelope montagnii*	0.15	16.5	706
B34	*Pharomachrus antisianus*	0.42	21.8	153
B35	*Pharomachrus auriceps*	0.37	22.1	180
B38	*Psarocolius angustifrons*	0.25	17.3	218.5
B39	*Psarocolius atrovirens*	0.25	15.9	152
B40	*Ramphocelus carbo*	0.15	11.4	28
B41	*Rupicola peruvianus*	0.12	20.9	244
B42	*Saltator maximus*	0.18	13	47.7
B43	*Tangara arthus*	0.22	8.5	22
B44	*Tangara chilensis*	0.23	7.4	23
B45	*Tangara chrysotis*	0.26	8.2	24
B46	*Tangara cyanicollis*	0.22	7.7	17
B47	*Tangara gyrola*	0.2	7.7	21
B48	*Tangara nigroviridis*	0.23	7.8	17
B49	*Tangara punctata*	0.21	6.8	15
B50	*Tangara ruficervix*	0.24	7.6	19
B51	*Tangara vassorii*	0.23	8	18
B52	*Tangara xanthocephala*	0.22	7.3	19
B53	*Thraupis bonariensis*	0.2	11	36
B54	*Thraupis cyanocephala*	0.16	10.7	36

Code	Species name	Kipp's index	Bill width (mm)	Body mass (g)
B55	*Thraupis episcopus*	0.23	10.2	35
B56	*Thraupis palmarum*	0.22	9.6	39
B57	*Trichothraupis melanops*	0.19	9.2	24.3
B58	*Trogon personatus*	0.37	16	63.4
B59	*Turdus chiguanco*	0.22	13.5	93.3
B60	*Xenopipo unicolor*	0.19	10	15.5
B61	*Zimmerius bolivianus*	0.14	6.5	11

Table E.2. Correspondence between bird codes and species names.

Acknowledgements

We thank Mahendra Mariadassou, Vincent Miele, Sara Puijalon, Stéphane Robin and Jan Venter for their input on this work during PhD commitee meetings. We also thank Elisa Thébault, Marie-Pierre Etienne, Simon Chamaillé-Jammes and in particular Anne-Béatrice Dufour for their input on this article initially written as a thesis chapter.

Fundings

This work was partially supported by the grant ANR-18-CE02-0010 of the French National Research Agency ANR (project EcoNet).

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Data, script, code, and supplementary information availability

Script and codes are available on GitHub (https://github.com/LisaNicvert/RSnetwork) and will be archived on Figshare upon publication.

References

Albrecht J, A Classen, MGR Vollstädt, A Mayr, NP Mollel, D Schellenberger Costa, HI Dulle, M Fischer, A Hemp, KM Howell, M Kleyer, T Nauss, MK Peters, M Tschapka, I Steffan-Dewenter, K Böhning-Gaese, and M Schleuning (2018). Plant and animal functional diversity drive mutualistic network assembly across an elevational gradient. *Nature Communications* 9, 3177. ISSN: 2041-1723. https://doi.org/10.1038/s41467-018-05610-w.

Araujo AIL, AM de Almeida, MZ Cardoso, and G Corso (2010). Abundance and nestedness in interaction networks. *Ecological Complexity* 7, 494–499. ISSN: 1476-945X. https://doi.org/10.1016/j.ecocom.2010.02.004.

Bascompte J and P Jordano (2007). Plant-Animal Mutualistic Networks: The Architecture of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics* 38, 567–593. ISSN: 1543-592X, 1545-2069. https://doi.org/10.1146/annurev.ecolsys.38.091206.095818.

Bascompte J, P Jordano, CJ Melián, and JM Olesen (2003). The nested assembly of plant–animal mutualistic networks. *Proceedings of the National Academy of Sciences* 100, 9383–9387. https://doi.org/10.1073/pnas.1633576100.

Benadi G, CF Dormann, J Fründ, R Stephan, and DP Vázquez (2022). Quantitative Prediction of Interactions in Bipartite Networks Based on Traits, Abundances, and Phylogeny. *The American Naturalist* 199, 841–854. ISSN: 0003-0147. https://doi.org/10.1086/714420.

Bender IMA, WD Kissling, PG Blendinger, K Böhning-Gaese, I Hensen, I Kühn, MC Muñoz, EL Neuschulz, L Nowak, M Quitián, F Saavedra, V Santillán, T Töpfer, T Wiegand, DM Dehling, and M Schleuning (2018). Morphological trait matching shapes plant–frugivore networks across the Andes. *Ecography* 41, 1910–1919. ISSN: 1600-0587. https://doi.org/10.1111/ecog.03396.

Blonder B, C Lamanna, C Violle, and BJ Enquist (2014). The n-dimensional hypervolume. *Global Ecology and Biogeography* 23, 595–609. ISSN: 1466-8238. https://doi.org/10.1111/geb.12146.

Blüthgen N, J Fründ, DP Vázquez, and F Menzel (2008). What do interaction network metrics tell us about specialization and biological traits. *Ecology* 89, 3387–3399. ISSN: 0012-9658. https://doi.org/10.1890/07-2121.1.

Blüthgen N, F Menzel, and N Blüthgen (2006). Measuring specialization in species interaction networks. *BMC Ecology* 6, 1–12. ISSN: 1472-6785. https://doi.org/10.1186/1472-6785-6-9.

Braak CJ ter and I Prentice (2004). A Theory of Gradient Analysis. In: *Advances in Ecological Research*. Vol. 34. Elsevier, pp. 235–282. ISBN: 978-0-12-013934-7. https://doi.org/10.1016/S0065-2504(03)34003-6.

Dehling DM, IMA Bender, PG Blendinger, MC Muñoz, M Quitián, F Saavedra, V Santillán, K Böhning-Gaese, EL Neuschulz, and M Schleuning (2021). ANDEAN frugivory: data on plant–bird interactions and functional traits of plant and bird species from montane forests along the Andes. Version 2. https://doi.org/10.5061/DRYAD.WM37PVMN5.

Dehling DM, P Jordano, HM Schaefer, K Böhning-Gaese, and M Schleuning (2016). Morphology predicts species' functional roles and their degree of specialization in plant–frugivore interactions. *Proceedings of the Royal Society B: Biological Sciences* 283, 20152444. ISSN: 0962-8452, 1471-2954. https://doi.org/10.1098/rspb.2015.2444.

Dehling DM, HR Lai, and DB Stouffer (2024). Eltonian niche modelling: applying joint hierarchical niche models to ecological networks. https://doi.org/10.1101/2024.08.14.607929.

Dehling DM and DB Stouffer (2018). Bringing the Eltonian niche into functional diversity. *Oikos* 127, 1711–1723. ISSN: 1600-0706. https://doi.org/10.1111/oik.05415.

Dehling DM, T Töpfer, HM Schaefer, P Jordano, K Böhning-Gaese, and M Schleuning (2014). Functional relationships beyond species richness patterns: trait matching in plant–bird mutualisms across scales. *Global Ecology and Biogeography* 23, 1085–1093. ISSN: 1466-8238. https://doi.org/10.1111/geb.12193.

Devictor V, J Clavel, R Julliard, S Lavergne, D Mouillot, W Thuiller, P Venail, S Villéger, and N Mouquet (2010). Defining and measuring ecological specialization. *Journal of Applied Ecology* 47, 15–25. ISSN: 1365-2664. https://doi.org/10.1111/j.1365-2664.2009.01744.x.

Dolédec S, D Chessel, CJF ter Braak, and S Champely (1996). Matching species traits to environmental variables: a new three-table ordination method. *Environmental and Ecological Statistics* 3, 143–166. ISSN: 1573-3009. https://doi.org/10.1007/BF02427859.

Dray S and P Legendre (2008). Testing the Species Traits–Environment Relationships: The Fourth-Corner Problem Revisited. *Ecology* 89, 3400–3412. ISSN: 1939-9170. https://doi.org/10.1890/08-0349.1.

Eklöf A, U Jacob, J Kopp, J Bosch, R Castro-Urgal, NP Chacoff, B Dalsgaard, C de Sassi, M Galetti, PR Guimarães, SB Lomáscolo, AM Martín González, MA Pizo, R Rader, A Rodrigo, JM Tylianakis, DP Vázquez, and S Allesina (2013). The dimensionality of ecological networks. *Ecology Letters* 16, 577–583. ISSN: 1461-0248. https://doi.org/10.1111/ele.12081.

Ellis PD (2010). *The Essential Guide to Effect Sizes: Statistical Power, Meta-Analysis, and the Interpretation of Research Results*. Cambridge University Press.

Elton C (1927). The animal community. In: *Animal Ecology*. London: Sidgwick and Jackson.

Erofeeva EA (2021). Plant hormesis and Shelford's tolerance law curve. *Journal of Forestry Research* 32, 1789–1802. ISSN: 1993-0607. https://doi.org/10.1007/s11676-021-01312-0.

Foster MS (1990). Factors Influencing Bird Foraging Preferences among Conspecific Fruit Trees. *The Condor* 92, 844–854. ISSN: 1938-5129. https://doi.org/10.2307/1368720.

Fründ J, KS McCann, and NM Williams (2016). Sampling bias is a challenge for quantifying specialization and network structure: lessons from a quantitative niche model. *Oikos* 125, 502–513. ISSN: 1600-0706. https://doi.org/10.1111/oik.02256.

Futuyma DJ and G Moreno (1988). The Evolution of Ecological Specialization. *Annual Review of Ecology and Systematics* 19, 207–233. ISSN: 0066-4162.

Gauch Jr. HG and RH Whittaker (1972). Coenocline Simulation. *Ecology* 53, 446–451. ISSN: 1939-9170. https://doi.org/10.2307/1934231.

Godoy O, I Bartomeus, RP Rohr, and S Saavedra (2018). Towards the Integration of Niche and Network Theories. *Trends in Ecology & Evolution* 33, 287–300. ISSN: 0169-5347. https://doi.org/10.1016/j.tree.2018.01.007.

Gravel D, B Baiser, JA Dunne, JP Kopelke, ND Martinez, T Nyman, T Poisot, DB Stouffer, JM Tylianakis, SA Wood, and T Roslin (2019). Bringing Elton and Grinnell together: a quantitative framework to represent the biogeography of ecological interaction networks. *Ecography* 42, 401–415. ISSN: 16000587. https://doi.org/10.1111/ecog.04006.

Greenacre M (2013). The contributions of rare objects in correspondence analysis. *Ecology* 94, 241–249. ISSN: 1939-9170. https://doi.org/10.1890/11-1730.1.

Grinnell J (1924). Geography and Evolution. *Ecology* 5, 225–229. ISSN: 0012-9658. https://doi.org/10.2307/1929447.

Hill MO (1973). Reciprocal Averaging: An Eigenvector Method of Ordination. *The Journal of Ecology* 61, 237. ISSN: 00220477. https://doi.org/10.2307/2258931.

— (1974). Correspondence Analysis: A Neglected Multivariate Method. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 23, 340–354. ISSN: 0035-9254. https://doi.org/10.2307/2347127.

Hirschfeld HO (1935). A Connection between Correlation and Contingency. *Mathematical Proceedings of the Cambridge Philosophical Society* 31, 520–524. ISSN: 1469-8064, 0305-0041. https://doi.org/10.1017/S0305004100013517.

Hutchinson GE (1957). Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22, 415–427. ISSN: 0091-7451, 1943-4456. https://doi.org/10.1101/SQB.1957.022.01.039.

Ings TC, JM Montoya, J Bascompte, N Blüthgen, L Brown, CF Dormann, F Edwards, D Figueroa, U Jacob, JI Jones, RB Lauridsen, ME Ledger, HM Lewis, JM Olesen, FF Van Veen, PH Warren, and G Woodward (2009). Ecological networks – beyond food webs. *Journal of Animal Ecology* 78, 253–269. ISSN: 1365-2656. https://doi.org/10.1111/j.1365-2656.2008.01460.x.

Jordano P, J Bascompte, and JM Olesen (2003). Invariant properties in coevolutionary networks of plant–animal interactions. *Ecology Letters* 6, 69–81. ISSN: 1461-0248. https://doi.org/10.1046/j.1461-0248.2003.00403.x.

Legendre P, R Galzin, and ML Harmelin-Vivien (1997). Relating Behavior to Habitat: Solutions to The fourth-Corner Problem. *Ecology* 78, 547–562. ISSN: 1939-9170. https://doi.org/10.1890/0012-9658(1997)078[0547:RBTHST]2.0.CO;2.

Lewinsohn TM, P Inácio Prado, P Jordano, J Bascompte, and J M. Olesen (2006). Structure in plant–animal interaction assemblages. *Oikos* 113, 174–184. ISSN: 1600-0706. https://doi.org/10.1111/j.0030-1299.2006.14583.x.

Maglianesi MA, K Böhning-Gaese, and M Schleuning (2015). Different foraging preferences of hummingbirds on artificial and natural flowers reveal mechanisms structuring plant–pollinator interactions. *Journal of Animal Ecology* 84, 655–664. ISSN: 1365-2656. https://doi.org/10.1111/1365-2656.12319.

Miele V, C Matias, S Robin, and S Dray (2019). Nine quick tips for analyzing network data. *PLOS Computational Biology* 15, e1007434. ISSN: 1553-7358. https://doi.org/10.1371/journal.pcbi.1007434.

Minchin PR (1987). Simulation of multidimensional community patterns: towards a comprehensive model. *Vegetatio* 71, 145–156. ISSN: 0042-3106. https://doi.org/10.1007/BF00039167.

Palacio F, M Valoy, F Bernacki, M Sánchez, M Núñez-Montellano, O Varela, and M Ordano (2017). Bird fruit consumption results from the interaction between fruit-handling behaviour and fruit crop size. *Ethology Ecology & Evolution* 29, 24–37. ISSN: 0394-9370. https://doi.org/10.1080/03949370.2015.1080195.

Pappas JL and EF Stoermer (1997). Multivariate measure of niche overlap using canonical correspondence analysis. *Écoscience* 4, 240–245. ISSN: 1195-6860, 2376-7626. https://doi.org/10.1080/11956860.1997.11682401.

Peralta G, PJ CaraDonna, D Rakosy, J Fründ, MP Pascual Tudanca, CF Dormann, LA Burkle, CN Kaiser-Bunbury, TM Knight, J Resasco, R Winfree, N Blüthgen, WJ Castillo, and DP Vázquez (2024). Predicting plant–pollinator interactions: concepts, methods, and challenges. *Trends in Ecology & Evolution*, S0169534723003361. ISSN: 01695347. https://doi.org/10.1016/j.tree.2023.12.005.

Phillips RD, R Peakall, T Van Der Niet, and SD Johnson (2020). Niche Perspectives on Plant–Pollinator Interactions. *Trends in Plant Science* 25, 779–793. ISSN: 13601385. https://doi.org/10.1016/j.tplants.2020.03.009.

R Core Team (2024). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Version 4.3.3. Vienna, Austria.

Schleuning M, N Blüthgen, M Flörchinger, J Braun, HM Schaefer, and K Böhning-Gaese (2011). Specialization and interaction strength in a tropical plant–frugivore network differ among forest strata. *Ecology* 92, 26–36. ISSN: 1939-9170. https://doi.org/10.1890/09-1842.1.

Schoener TW (1971). Theory of Feeding Strategies. *Annual Review of Ecology, Evolution and Systematics* 2 (Volume 2, 1971), 369–404. ISSN: 1543-592X, 1545-2069. https://doi.org/10.1146/annurev.es.02.110171.002101.

Sexton JP, J Montiel, JE Shay, MR Stephens, and RA Slatyer (2017). Evolution of Ecological Niche Breadth. *Annual Review of Ecology, Evolution, and Systematics* 48, 183–206. ISSN: 1543-592X, 1545-2069. https://doi.org/10.1146/annurev-ecolsys-110316-023003.

Shelford VE (1931). Some Concepts of Bioecology. *Ecology* 12, 455–467. ISSN: 1939-9170. https://doi.org/10.2307/1928991.

ter Braak CJF (1986). Canonical Correspondence Analysis: A New Eigenvector Technique for Multivariate Direct Gradient Analysis. *Ecology* 67, 1167–1179. ISSN: 1939-9170. https://doi.org/10.2307/1938672.

ter Braak CJF, P Šmilauer, and S Dray (2018). Algorithms and biplots for double constrained correspondence analysis. *Environmental and Ecological Statistics* 25, 171–197. ISSN: 1352-8505, 1573-3009. https://doi.org/10.1007/s10651-017-0395-x.

ter Braak CJF and PFM Verdonschot (1995). Canonical correspondence analysis and related multivariate methods in aquatic ecology. *Aquatic Sciences* 57, 255–289. ISSN: 1420-9055. https://doi.org/10.1007/BF00877430.

Thiel S, F Willems, N Farwig, F Rehling, DG Schabo, M Schleuning, N Shahuano Tello, T Töpfer, M Tschapka, EW Heymann, and K Heer (2023). Vertically stratified frugivore community composition and interaction frequency in a liana fruiting across forest strata. *Biotropica* 55, 650–664. ISSN: 1744-7429. https://doi.org/10.1111/btp.13216.

Thioulouse J and D Chessel (1992). A Method for Reciprocal Scaling of Species Tolerance and Sample Diversity. *Ecology* 73, 670–680. ISSN: 00129658. https://doi.org/10.2307/1940773.

Thioulouse J, S Dray, AB Dufour, A Siberchicot, T Jombart, and S Pavoine (2018). *Multivariate Analysis of Ecological Data with {ade4}*. Springer.

van Dam A, M Dekker, I Morales-Castilla, MÁ Rodríguez, D Wichmann, and M Baudena (2021). Correspondence analysis, spectral clustering and graph embedding: applications to ecology and economic complexity. *Scientific Reports* 11, 8926. ISSN: 2045-2322. https://doi.org/10.1038/s41598-021-87971-9.

Vandermeer JH (1972). Niche Theory. *Annual Review of Ecology and Systematics* 3, 107–132. ISSN: 0066-4162.

Vázquez DP, N Blüthgen, L Cagnolo, and NP Chacoff (2009). Uniting pattern and process in plant–animal mutualistic networks: a review. *Annals of Botany* 103, 1445–1457. ISSN: 0305-7364. https://doi.org/10.1093/aob/mcp057.

Wheelwright NT (1985). Fruit-Size, Gape Width, and the Diets of Fruit-Eating Birds. *Ecology* 66, 808–818. ISSN: 0012-9658, 1939-9170. https://doi.org/10.2307/1940542.

Whittaker RH (1956). Vegetation of the Great Smoky Mountains. *Ecological Monographs* 26, 2–80. ISSN: 0012-9615. https://doi.org/10.2307/1943577.