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

Lisa Nicvert[1,2], Hervé Fritz[3,4], & Stéphane Dray[1]

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

Correspondence: lisa.nicvert@fondationbiodiversite.fr

Abstract

Species interactions occurring in ecological communities are often represented as networks, whose structure is thought to be linked to species' interaction niches (or Eltonian niches). Interaction niches 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, some of these methods have also been used to study the interaction niche, but they 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 suggests 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 exploratory methods to characterize species interaction profiles, provide an ecologically meaningful graphical representation of interaction niches and 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). Species' abiotic requirements are described with the Grinnellian niche (Grinnell, 1924), which centers on the physical and environmental requirements essential for a species' survival. In contrast, species' biotic requirements are described with the Eltonian niche (originally formalized in the context of food webs by Charles Elton; Elton, 1927) , which sees species' niche through the lens of their interactions with other organisms. 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.

The niche is often characterized by its optimum and breadth. 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.

One common way to describe species' niches is to approximate niche space with Gaussian curves along each niche axis (Gauch Jr. and Whittaker, 1972). Thus, niche optima are estimated as mean, and niche breadths as standard deviations, of a multivariate normal distribution along requirements gradients.

In this article, we focus in realized niches, i.e. niches that species effectively occupy, estimated from community data. Methods used to describe and analyze these niches are often different depending on whether the focus is the environmental or the interaction niche. On the one hand, interaction niches are often described with the formalism of networks (Bascompte and Jordano, 2007; Ings et al., 2009). Ecological networks are typically represented as species (nodes) linked by interactions (edges). In these networks, interaction niches are often studied through specialization or generalization of species (or niche breadth). In this context, niche breadth 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 (the 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).

On the other hand, to study environmental niches, multivariate methods are a popular tool. They allow to estimate species niche optima along environmental gradient(s). The base of these methods is weighted averaging (Whittaker, 1956): it consists in averaging the values of the environmental variable over the samples in which a species occurs, weighted by the species' abundance, to estimate its niche optimum. These multivariate methods can use implicit or explicit environmental gradients, and optionally relate these gradients with species traits: when the only data available are species abundances sampled in different sites, correspondence analysis (CA) (Hill, 1973, 1974), estimates species' niche optima on latent gradients. When multiple environmental data are recorded, canonical correspondence analysis (ter Braak, 1986) estimates species niche optima on environmental gradients. Finally, when traits data and environmental data are available, one can use fourth-corner analysis (P Legendre, Galzin, et al., 1997), RLQ (Dolédec et al., 1996) and double-constrained correspondence analysis (ter Braak et al., 2018) to link niche optima estimated on environmental gradients to species characteristics.

In this article, we focus on characterizing species interactions in ecological networks. It is generally accepted that these interactions do not arise randomly, but reflect multiple drivers such as neutral effects, trait matching, and evolutionary history: neutral effects imply that a given species is more likely to interact with abundant species than with rare ones (Vázquez et al., 2009). Trait matching posits that species with complementary traits preferentially engage in interactions. For instance, body size has been identified as a key factor influencing food web structure (Elton, 1927; Ings et al., 2009). Species evolutionary history is also thought to influence their interaction patterns, for instance by constraining traits via phylogenetic inertia or coevolution (Benadi et al., 2022; Dormann et al., 2017). Finally, these processes are not mutually exclusive, and often intersect to

shape observed patterns (Vázquez et al., 2009).

Recently, several multivariate methods originally developed to study environmental niches have been applied to study interaction networks. In particular, many methods have investigated trait matching: for instance, RLQ and fourth-corner analyses have been used to measure the correlation between two sets of partner species' traits, weighted by the interaction matrix (Albrecht et al., 2018; Bender et al., 2018; Dehling, Töpfer, et al., 2014).

Classical 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 (Godoy et al., 2018; Phillips et al., 2020) : however, trait matching clearly relies on the notion of interaction niche (Albrecht et al., 2018; Eklöf et al., 2013). Indeed, we can see trait matching as the mark of species' interaction niche, as two species will interact only if their niches match (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. 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 species' realized interaction niches in ecological networks. We use ordination scores produced by CA and interpret these scores as species' latent traits, corresponding to the optima of their interaction niches. We also use reciprocal scaling (Thioulouse and Chessel, 1992) to estimate interaction niches breadths, defined as the standard deviation of the latent traits of a species' interacting partners. Our framework also allows to measure the effect of trait matching, without using traits data, using the notion of latent traits, which are understood as a proxy of various unmeasured species characteristics. 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 4.4.2 (R Core Team, 2024) and are available at https://doi.org/10.6084/m9.figshare.28625270 (Nicvert et al., 2025).

Notations

In this part, we consider a $r \times c$ matrix $\mathbf{Y} = [y_{ij}]$ representing the bipartite interaction network 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 can encompass diverse bipartite network types (Fründ et al., 2016; Ings et al., 2009).

For the following mathematical developments, we define $\mathbf{P}$ as the table of relative frequencies of interactions: $\mathbf{P} = [y_{ij}/y_{++}]$ (where $y_{++} = \sum_{i=1}^{r}\sum_{j=1}^{c} y_{ij}$ is the grand total of $\mathbf{Y}$). We also define the weight matrices $\mathbf{D_r} = \mathrm{diag}(\mathbf{r})$ for resources (rows) and $\mathbf{D_c} = \mathrm{diag}(\mathbf{c})$ for consumers (columns), where the vectors $\mathbf{r} = \mathbf{P}\mathbf{1_r} = [p_{1+}, \ldots, p_{r+}]^\top$ and $\mathbf{c} = \mathbf{P}^\top\mathbf{1_c} = [p_{+1}, \ldots, p_{+c}]^\top$ represent respectively the rows and columns marginal sums (i.e. $p_{i+} = \sum_{j=1}^{c} p_{ij}$ and $p_{+j} = \sum_{i=1}^{r} p_{ij}$). Throughout this article, $^\top$ will denote matrix transposition.

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 $\mathrm{cor}_{\mathbf{P}}(\mathbf{x}, \mathbf{y})$, defined by fourth-corner analysis (P Legendre, Galzin, et al., 1997): this statistic measures the correlation between the traits of consumers and resources, weighted by their interaction frequencies:

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

where $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ are centered scaled versions of the traits, respectively using weights $\mathbf{D_r}$ and $\mathbf{D_c}$. This fourth-corner statistic has been used to quantify trait matching in interaction networks (Albrecht et al., 2018; Bender et al., 2018).

When the traits contributing to the matching are not known, or not available, the fourth-corner statistic cannot be computed. In that case, indirect gradient methods such as correspondence analysis, described below, can be useful to identify latent structures in ecological networks.

Correspondence analysis (CA) for interaction networks

Correspondence analysis (Hill, 1974; Hirschfeld, 1935) is a multivariate method used to analyze contingency tables. Below, we describe CA applied to contingency tables describing interaction counts.

CA is based on the generalized singular value decomposition of $\mathbf{P_0} = \mathbf{P} - \mathbf{r}\mathbf{c}^\top$ (the doubly centered matrix of interaction frequencies $\mathbf{P}$) :

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

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 respectively the left (resource) and right (consumer) generalized singular vectors. $\mathbf{\Delta}$ is the diagonal matrix containing ordered singular values. The matrix of CA eigenvalues $\mathbf{\Lambda}$ is equal to $\mathbf{\Delta}^2$.

To interpret these results, let us consider a given dimension k. The previous equation can then be rewritten as:

$$\underbrace{\delta_k}_{\mathrm{cor}_{\mathbf{P_0}}(\mathbf{u}_k, \mathbf{v}_k)} = \mathbf{u}_k^\top \mathbf{P_0} \mathbf{v}_k \tag{3}$$

As the matrix $\mathbf{P_0}$ is doubly centered (by rows and columns), its singular vectors stored in $\mathbf{U}$ and $\mathbf{V}$ are also centered. Hence, Equation (3) is strictly analogous to the trait matching equation (1) so that $\delta_k = \mathrm{cor}_{\mathbf{P_0}}(\mathbf{u}_k, \mathbf{v}_k) = \mathrm{cor}_{\mathbf{P}}(\mathbf{u}_k, \mathbf{v}_k)$. However, in Equation (3), known traits $\tilde{\mathbf{x}}$ and $\tilde{\mathbf{y}}$ are replaced with the generalized singular vectors $\mathbf{u}_k$ and $\mathbf{v}_k$. Therefore, CA amounts to finding the singular vectors $\mathbf{u}_k$ and $\mathbf{v}_k$ that are maximally correlated. Thus, scores $\mathbf{u}_k$ and $\mathbf{v}_k$ can be interpreted as latent traits, associated to resources (respectively, consumers), that maximize matching between these latent traits.

Using the relationship between $\mathbf{\Delta}$ and $\mathbf{\Lambda}$, we can show that the square-root of a CA eigenvalue $\sqrt{\lambda_k} = \delta_k$ represents the absolute value of the correlation between $\mathbf{u}_k$ and $\mathbf{v}_k$: if it is close to one, it indicates that latent traits are able to explain well the probability of interaction. The global relationship between all pairs of latent traits can also be tested statistically, using a χ^2 test on the interaction matrix (Abdi and Béra, 2017; Beh, 2004). This test can be performed analytically if its assumptions are met, but it will rarely be the case in the context of interactions networks, and a permutation approach will often be better suited.

CA has the advantage to evaluate the effect of trait matching, without using traits data, using these latent scores. However, we stress that these latent traits represent more than species unmeasured traits. Indeed, they 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.

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 B for an example of reordered interaction matrix with the bird-fruit interaction data used in this article).

Standards outputs of CA consist of two biplots where both consumers and resources are displayed in the same multivariate space. To convert consumer scores in the resouce space (and vice versa), 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.

The scores $\mathbf{U}^\star$ and $\mathbf{V}^\star$ can also be computed using scores $\mathbf{U}$ (respectively, $\mathbf{V}$) scaled by the square roots of eigenvalues (P Legendre and L Legendre, 2012):

$$\mathbf{U}^\star = \mathbf{U}\mathbf{\Lambda}^{1/2} \qquad \text{(resources)} \tag{5a}$$

$$\mathbf{V}^\star = \mathbf{V}\mathbf{\Lambda}^{1/2} \qquad \text{(consumers)} \tag{5b}$$

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 (Equation (4)). In this framework, the niche optimum of a species is defined as the mean value of its interacting partners' traits. Equivalently, these scores can be interpreted as species' latent traits (Equation (5)).

Using these scores, resources and consumers can be plotted in two spaces, corresponding to two biplots. A first biplot can be drawn by displaying the resources scores given by their latent trait ($\mathbf{U}$) and consumers scores given by weighted averaging ($\mathbf{V}^\star$). The second biplot mirrors the first, with consumers scores $\mathbf{V}$ (latent traits) and resources scores $\mathbf{U}^\star$ (weighted averaging).

These scores make CA particularly adapted to the analysis of interaction networks. However, this method summarizes species' niches with their optima, thus completely ignoring the diversity of interacting partners (niche breadth); moreover, the niche optima can only be visualized on two different biplots.

Reciprocal scaling

Reciprocal scaling, proposed by Thioulouse and Chessel (1992), is an extension of CA that solves the two issues mentioned above by providing niche breadths, in addition to optima, and defining a common multivariate space for resource and consumers. Reciprocal scaling shifts the focus of the analysis by considering the interactions, rather than the species, as statistical individuals: for that, this method uses 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 $h_k(i,j)$ for each interaction between species i and j in the k-th dimension of the multivariate space (note that it is defined only if species i and j interact). These interactions correspond to non-empty cells in the interaction matrix $\mathbf{Y}$ (they are called correspondence in Thioulouse and Chessel, 1992). Interactions scores can be computed from CA scores using the following formula:

$$h_k(i,j) = \frac{u^\star_{ik} + v^\star_{jk}}{\sqrt{2\lambda_k\mu_k}} \tag{6}$$

where λ_k is the CA eigenvalue for axis k and $\mu_k = 1 + \sqrt{\lambda_k}$. The scores $h_k(i,j)$ can be grouped into a vector $\mathbf{h}_k$ of length $n_{\bar{0}}$ for dimension k (where $n_{\bar{0}}$ is the number of unique interactions in the matrix, i.e. non-empty cells).

Using the interaction scores $h_k(i,j)$, it is possible to display the interaction niche of a species as a cloud of interactions in the multivariate space. The niche can then be summarized by estimating its optimum and its

breadth using a Gaussian approximation. If we consider species s in the multivariate plane given by axes k and l, its niche is approximated with an ellipse (see Figure 1). The center of the 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). Finally, the orientation of the ellipse represents the weighted covariance for axes k and l.

Formulas to compute niche optima, breadths and covariances are given in Appendix C (from Thioulouse and Chessel (1992) and adapted with the notation used in this article). 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, up to a scaling factor (see Equations (C.1) and (C.2) in Appendix C). Formulas to compute niche optima, breadths and covariances are given in Appendix C (from Thioulouse and Chessel (1992) and adapted with the notation used in this article).

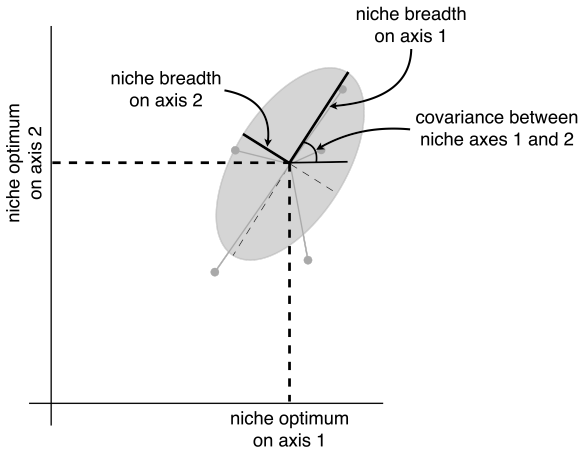


Figure 1. Visualization of the interaction niche in a 2D space. The niche of a species is depicted by an ellipse whose center represents niche optima and semi-axes represent niche breadths on each axis. The angle of the ellipse represents the covariance between niche axes. Points represent interaction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

Simulation

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

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 P 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): we compute an interaction probability due to matching using traits generated in the previous step (Figure 2b). This probability m_{ij} follows a bivariate normal distribution: its mean vector is the difference between the trait values of resource species i and consumer species j ($t^c_{jk} - t^r_{ik}$); variances in the variance-covariance matrix corresponds to the degree of generalization of consumer species (s^2_{jk}); and covariances are null:

$$\mathbf{M} = [m_{ij}] = \frac{1}{2\pi s_{j1} s_{j2}} \exp\left(-\frac{(t^c_{j1} - t^r_{i1})^2}{2s^2_{j1}} - \frac{(t^c_{j2} - t^r_{i2})^2}{2s^2_{j2}} \right) \tag{7}$$

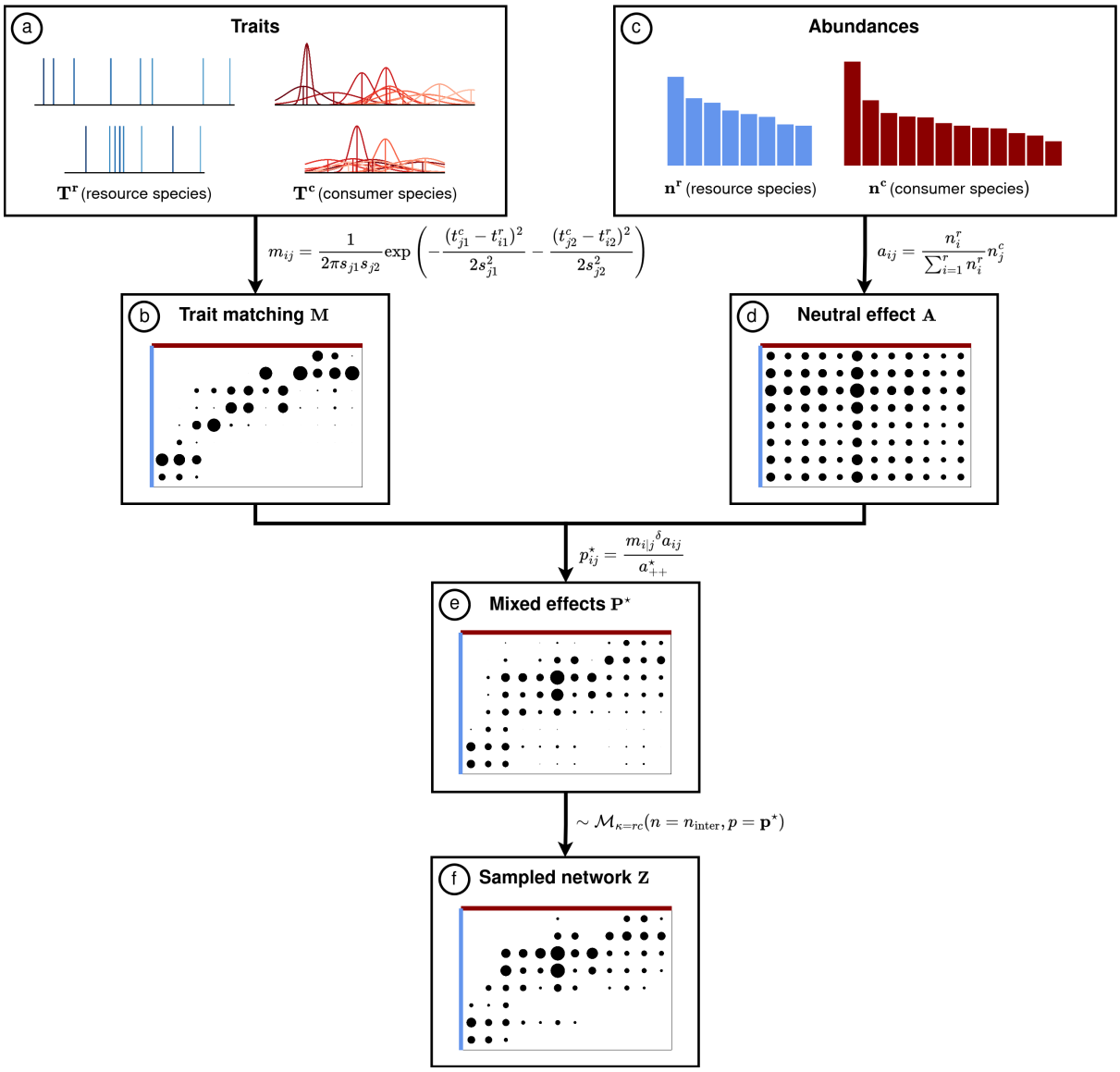


Figure 2. Model used to simulate interactions between consumer and resource species. (a) Two traits are generated for each species, corresponding to their niche optima. For consumers, standard deviations are also associated to traits to generate the niche breadths. (b) Interaction probabilities due to 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.

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 abundances (Figure 2c):

$$\mathbf{A} = [a_{ij}] = \left[\frac{n_i^r}{\sum_{i=1}^r n_i^r} n_j^c \right] \tag{8}$$

Here, we model interactions due to neutral effects as the result of consumer choices, driven 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 (7)) and neutral abundance effects ($\mathbf{A}$, equation (8)) (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 $\mathbf{A}^\star = [a_{ij}^\star]$ is computed as:

$$a_{ij}^\star = {m_{i|j}}^\delta a_{ij} \tag{9}$$

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 $\mathbf{P}^\star = [p_{ij}^\star]$ is then computed from the mixed interaction frequency matrix:

$$p_{ij}^\star = \frac{a_{ij}^\star}{a_{++}^\star} \tag{10}$$

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 $\mathbf{P}^\star$.

$$\mathbf{z} \sim \mathcal{M}_{\kappa=rc}(n = n_{\text{inter}}, p = \mathbf{p}^\star) \tag{11}$$

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 5 experiments to study the influence of 5 parameters (the detailed parameter values are presented in Table 1):

1. Sampling (n_{inter}): we vary the total number of interactions in the matrix n_{inter}, which represents the sampling intensity. This 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 or 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 communities of generalist (high standard deviation) and 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).

5. Trait matching strength (δ): we vary the value of the parameter δ, which determines the strength of trait matching compared to neutral species abundances (Equation (9)).

Each experiment described above studies the influence of one parameter by varying this parameter, and other parameters are fixed to a default value (see Table 1). In experiment (μ_{breadth}), the standard deviation is also changed to keep a constant ratio of $\sigma_{\text{breadth}} = \mu_{\text{breadth}}/2$. For all simulations, we generated interactions between $r = 50$ resources and $c = 60$ consumers.

Experiment	n_{inter}	Distribution of $\mathbf{n^r}$ and $\mathbf{n^c}$	μ_{breadth}	σ_{breadth}	δ
(n_{inter})	200 500 5000 1000 10 000	$\text{Lognormal}(\mu = \ln(3), \sigma = \ln(1.5))$	10	5	0.2
($\mathbf{n^r}$, $\mathbf{n^c}$)	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	0.2
(μ_{breadth})	5000	$\text{Lognormal}(\mu = \ln(3), \sigma = \ln(1.5))$	2 10 20 50	1 5 10 25	0.2
(σ_{breadth})	5000	$\text{Lognormal}(\mu = \ln(3), \sigma = \ln(1.5))$	10	0.1 1 5 10 50	0.2
(δ)	5000	$\text{Lognormal}(\mu = \ln(3), \sigma = \ln(1.5))$	10	5	0 0.2 0.4 0.6 0.8 1

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. Experiment (δ) changes the strength ot trait matching.

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 A). 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 was 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 interaction modules, which should be analyzed separately when using CA (van Dam et al., 2021). 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 D).

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). We test the relationship between latent traits using a χ^2 test, and evaluate the effect size using Cramér's V (Ellis, 2010).

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 latent traits (optima) obtained by reciprocal scaling (two axes for resources' and for consumers' niches). To keep simple models while also allowing a flexible in the shape, 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 5 experiments to test the influence of various parameters on the inference of

niche parameters (see Table 1). The results are presented in Figure 4 for trait matching strength (δ) and Figure 3 for others parameters.

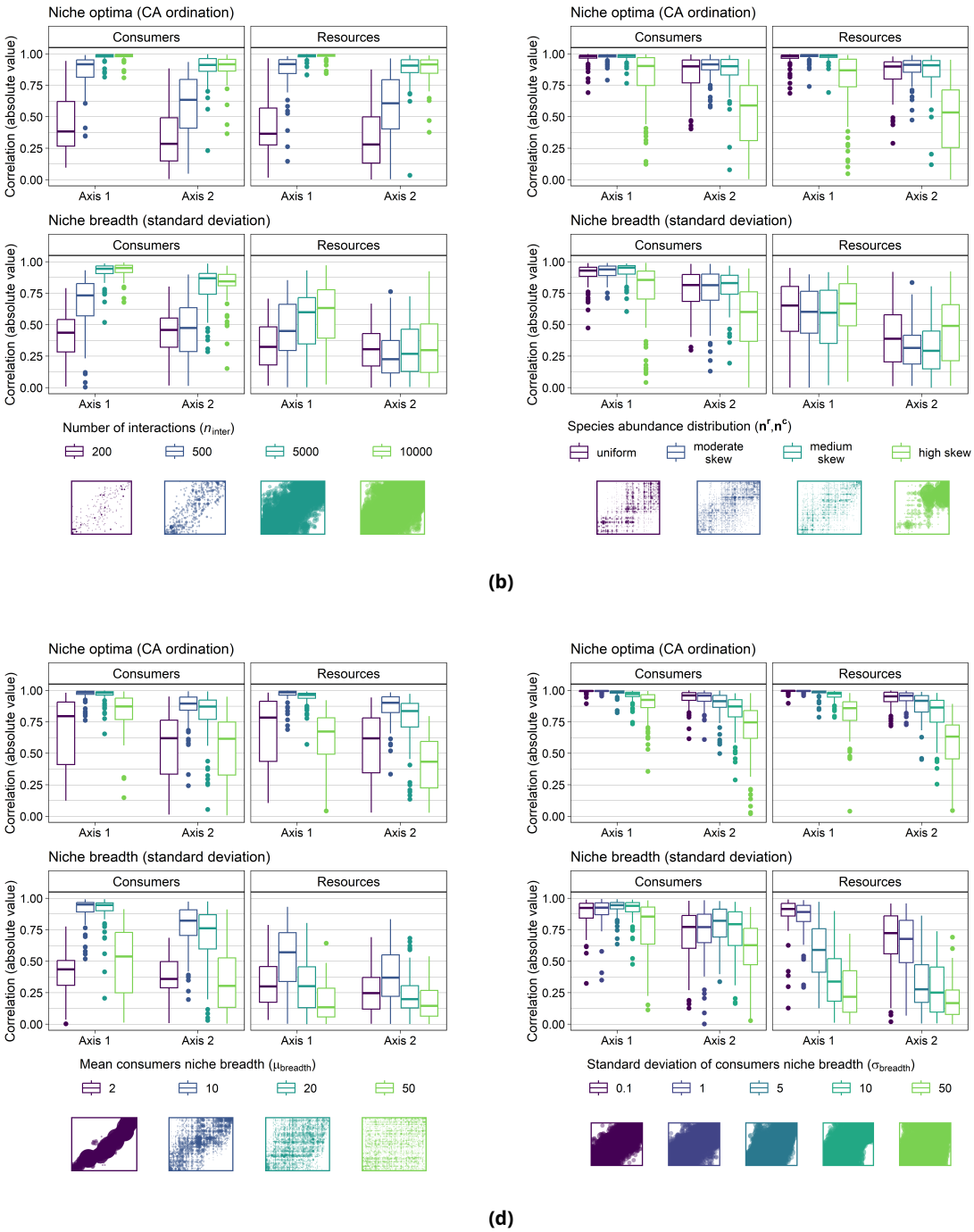


Figure 3. Results of the simulation study. Each subfigure explores the influence of one parameter on the model performance. (a) Effect of sampling intensity (number of interactions) (n_{inter}). (b) Effect of species abundance distributions ($\mathbf{n^r}$, $\mathbf{n^c}$). (c) Effect of the mean consumers niche breadths (μ_{breadth}). (d) Effect of heterogeneity (standard deviation) 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 (resources in rows, consumer in columns) generated with the corresponding parameter value. Points represent non-null interactions, with size proportional to the number of observed interactions censored at the 1st and 99th percentiles.

First, as expected, performance consistently improves with sampling completeness for all niche measures

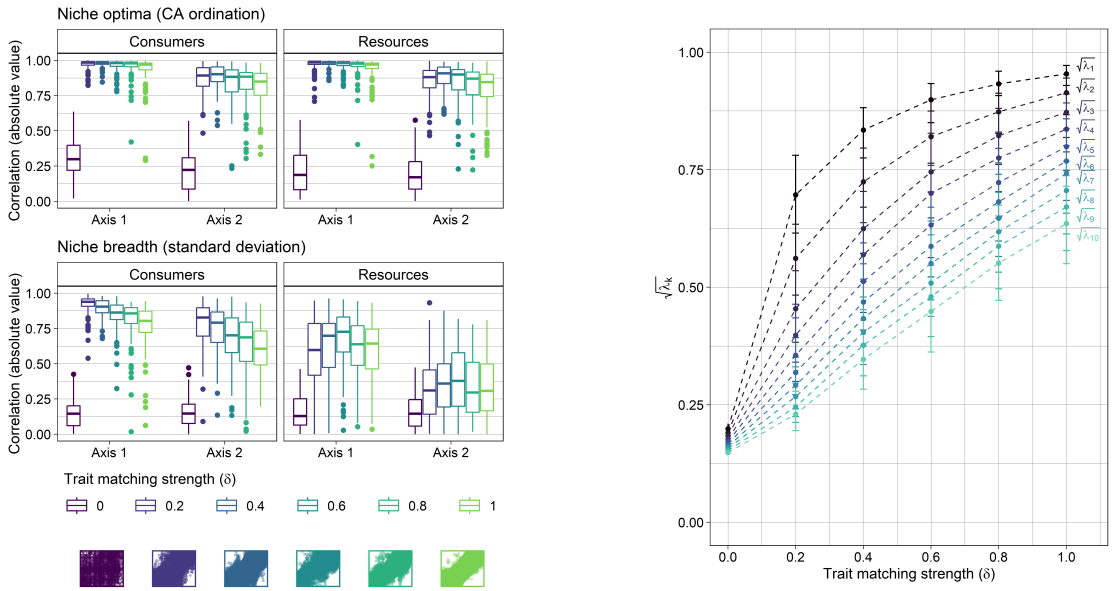
(Experiment (n_{inter}), Figure 3a). Second, when there are enough data (here, more than 500 sampled interactions), the niche optima are recovered correctly on both axes (median correlation value above 0.875), and niche breadths are recovered less precisely (Figure 3a). We can also note that for all 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.

Regarding the effect of abundance distribution (Experiment ($\mathbf{n^r}$, $\mathbf{n^c}$), 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).

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.

Surprisingly, niche parameters are optimally estimated for an intermediate matching strength δ (around 0.2-0.4) (Figure 4a). However, Figure 4b shows that the square roots of eigenvalues (which represent the absolute value of the correlation between the k-th pair of latent traits) increase with trait matching strength δ, thus indicating a higher matching signal as δ increases.



(a) Effect of matching strength (δ) on niche parameters **(b)** Eigenvalues evolution with δ

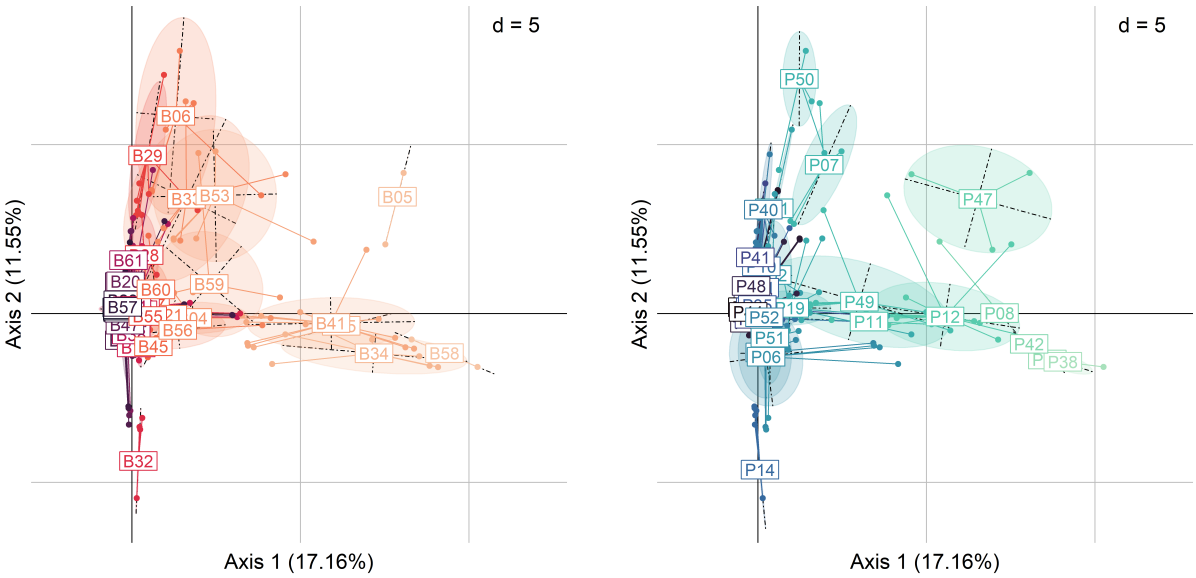
Figure 4. Influence of the strength of trait matching δ on niche parameters (a) and latent traits correlation (b). (a) Quality of the inference of niche parameters depending on values of δ. (b) Absolute value of latent traits correlation (square-root of eigenvalues $\sqrt{\lambda_k}$) depending on different values of δ. In Figure (a), 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 (resources in rows, consumer in columns) generated with the corresponding parameter value. Points represent non-null interactions, with size proportional to the number of observed interactions censored at the 1st and 99th percentiles. In Figure (b), colored points and dotted line represent the square root of an eigenvalue, and error bars represent the 2.5 and 97.5-th percentiles on the eigenvalues over 100 repetitions.

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 use them to visualize species niches. On Figure 5, interactions (points) are grouped by interacting partner (either birds or plants). Figures 5a and 5b show interactions grouped by bird and plant species, respectively (we separate the two groupings to avoid a cluttered representation, but the two graphs use 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

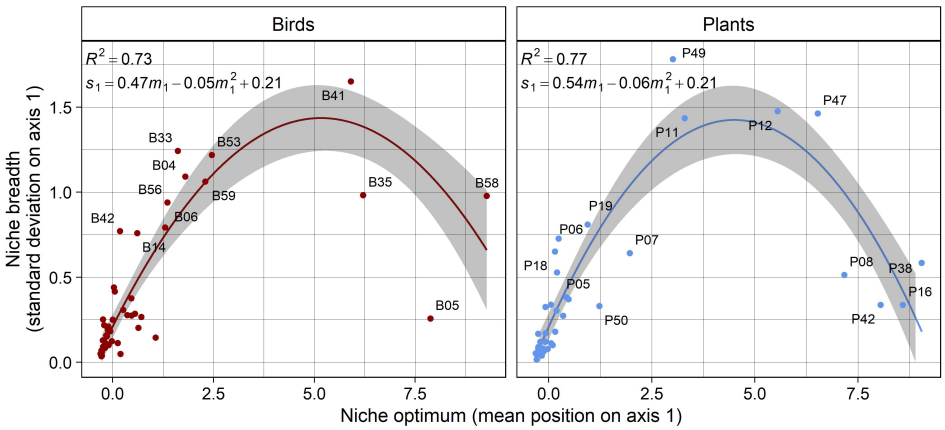
Figure 5. Reciprocal scaling of the birds-plants interaction network. Points correspond to the interaction scores $h_k(i, j)$ in the first 2 dimensions, grouped by bird (a) or plant species (b). Species label are placed at the reciprocal scaling mean, and ellipses correspond to the bivariate normal distribution 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 close niche optima. For instance, the interaction between B05 (*Cephalopterus ornatus*) and P47 (*Symplocos arechea*), which are close to each other, represents respectively 66% and 33% of their total interactions. Finally, the ellipses' semi-major and semi-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.

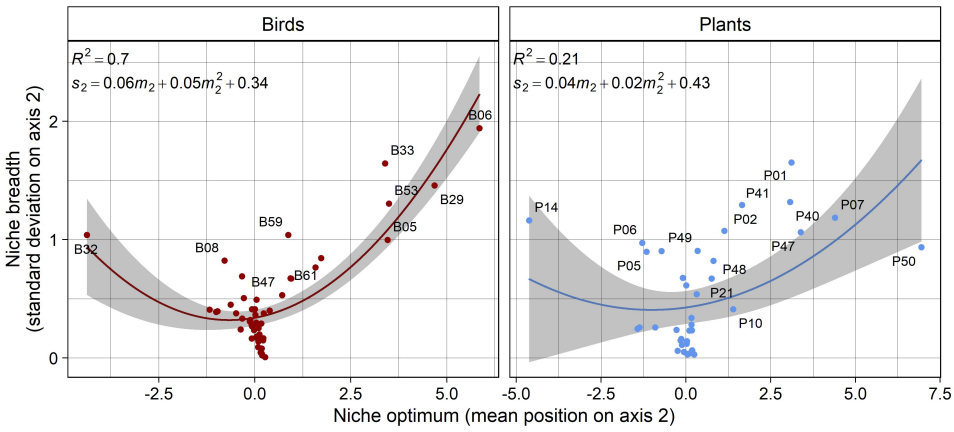
These figures also show how niches are distributed in the latent traits space. For birds (Figure 5a), many have a niche located around the origin (they interact with average plants, in the sense of their latent traits). Then, we can distinguish two clusters along the positive portion of the two main axes, which show two strategies 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 6. Niche breadth versus niche optimum on the 2 multivariate axes. Niche breadths and optima 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 niche position on the axes, which also represent the latent traits of a species, and niche breadth, we plotted the relationship between these two variables for each axis in Figure 6. We also modeled these relationships using linear models (see material and methods, and the models' equations written on the figures). On axis 1, for birds, there is a concave relationship with a positive linear component between niche optimum and niche breadth (but the variance of the residuals increases along axis 1). For plants, there is a concave relationship only. On axis 2, for birds, we have a convex relationship 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 and latent traits (measured with CA coordinates). Figure 7 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.

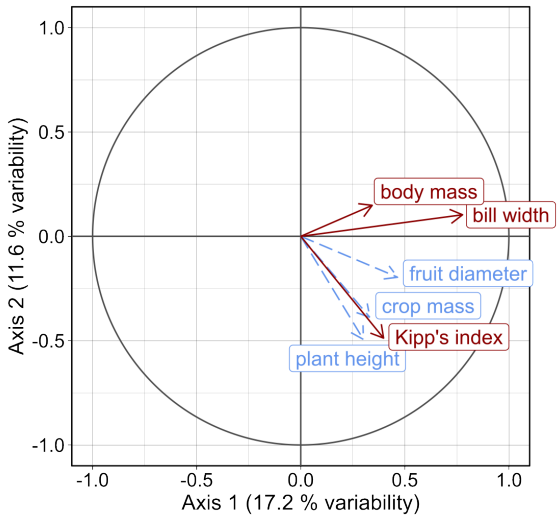


Figure 7. Correlation circle measured a posteriori between species measured and latent traits. Bird traits are shown in solid red lines and plant traits in dashed blue lines. Axis 1 shows a size effect while axis 2 is defined by the matching between plant height and Kipp's index.

Discussion

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

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. This method also works best for intermediate trait matching strength.

Unsurprisingly, performance increases with sampling intensity (Figure 3a). Moreover, when the proportion of variation explained by traits is smaller (on axis 2 of the simulation study), performance drops.

The estimation is robust to skewed data, but performance decreases with extremely skewed data (Figure 3b): 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 (Figure 3c), 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 between niches of specialist species and those of other species', which might bias the placement of niche optima and the estimation of niche breadths for specialist species and their interacting partners.

Surprisingly, a stronger matching between traits does not always yield a better recovery of niche parameters (Figure 4a, experiment (δ)). As expected, when there is no trait matching ($\delta = 0$), niche parameters are poorly recovered, because the true niche is random. However, the performance is best when the matching strength is intermediate ($\delta = 0.2$), and then decreases with a stronger matching, although it remains quite good even when the matching is maximal ($\delta = 1$).

This phenomenon can be explained by the requirements of CA to perform a latent ordination. Indeed, if the matching is too strong, the niche of each species becomes increasingly separated from the others in the ordination space. Strong trait matching essentially translates into smaller niche overlap (see discussion for Experiment (σ_{breadth})): in that case, it is more difficult for CA to find the common structuring driver without external information. The small overlap between species niches (few common resources or consumers) makes species ordination difficult as species become equidistant in the multivariate space. Here, when $\delta = 1$, all eigenvalues are also closer to 1 (Figure 4b): the structure of the network is recovered in more than two axes, whereas the true structure used to generate data is driven by two matching traits (i.e. two axes). For this reason, the latent traits on the first two axes are less close from to the two true traits used in the simulation, thus yielding less accurate estimations of the niche parameters. Therefore, if the CA of an interaction network yields eigenvalues that are all close to one, species latent traits should be interpreted with caution: we could still conclude that there is a strong trait matching, but would suspect that the dimensionality of trait matching is lower than the one that is inferred.

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.

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 (7)) 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), its validity 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 the medium effect size measured with Cramér's V.

This non-random structure is quantified as interaction counts that deviate from their expected value in

the sense of the chi-squared statistic. However, the chi-squared statistic uses the matrix margins to assess the expected number of interactions, 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. 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). Moreover, 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). These two axes represent 28.7% of the total variation in the network, so the patterns duscussed below explain onlly are only a small part of network structure.

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 both 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 correlate latent and measured traits. The first latent trait (axis) is mostly correlated with birds and plants size. For birds (Figure 6a, 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 expensesTherefore, they 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 7). 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 6a, right), the convex relationship between niche breadth and the first latent trait (axis) is even 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 their niches on Figure 5b. 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 6b (left) shows a concave effect with a large positive linear component. In this discussion, we will focus on this linear component, because the convexity 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. 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 7), whereas birds with rounded wings also supplement their diet with other food sources such as insects (Foster, 1990; Thiel et al., 2023), 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.

A limitation of the results discussed above is that latent traits are not driven by species traits only. Indeed, due to their latent nature, these traits may conflate several processes (e.g. phylogeny or spatio-temporal heterogeneity).

Moreover, the network analyzed here is an aggregation of four observed samples collected over one year. Therefore, it is a mean version of the interaction network, which is not a static object and may vary in time (Poisot et al., 2015; Singer and McBride, 2012). Here, the inferred niche is a mean of the interactions a species engages in within one year, although the niche may vary dynamically depending on various temporal processes (e.g. fruiting times).

Finally, we must keep in mind that we observe realized niches here, so these niches might not reflect the "true" fundamental preference of birds. Here, 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. In Appendix A, 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).

Although the analysis produces sensible results, the two axes (latent traits) 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 (i.e. 6 latent traits) explains more than 60% of the variation of the network.

Conclusion and perspectives

This article describes correspondence analysis and reciprocal scaling applied to ecological networks. These methods are well-suited for exploratory analysis of the structure of interaction networks. In particular, when no traits data are available, CA allow to evaluate a potential signal and the dimensionality of trait matching (Eklöf et al., 2013) , without requiring trait data (using the square roots of eigenvalues). This allows to take into account all potential species traits to evaluate trait matching, even if no traits measures are available.

Reciprocal scaling provides a quantitative measure of the interaction niche, using a Gaussian approximation: species niche optima are approximated using the centroid of their interactions, and species niche breadths using their standard deviation in the multivariate space. 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 of the niche as a hypervolume (Hutchinson, 1957).

The framework proposed by CA and reciprocal scaling also links trait matching and species specialization/generalization, through the notion of interaction niches: trait matching is viewed as the alignment be-

tween species niche optima, defined by their traits; and specialization and generalization are viewed as the diversity of traits of a species' interacting partners. This conceptual link between network metrics and species niche has also been proposed in other works (Albrecht et al., 2018; Dehling, Lai, et al., 2024; Godoy et al., 2018).

CA and reciprocal scaling are proposed as a powerful tool for analyzing species interactions, but they are also limited within a specific scope. First, latent traits conflate several processes, like species' spatial and temporal variations in abundances, phylogeny, or interaction sampling biases. Therefore, the estimated niche corresponds to species' realized niche, i.e. the interacting partners a species effectively interacts with (Devictor et al., 2010; Hutchinson, 1957), which can be influenced by a variety of factors beyond their traits.

Second, reciprocal scaling estimates a Gaussian approximation of the species niche. However, real niches might not be Gaussian. While this approximation will likely be sensible for symmetric and unimodal niches, the Gaussian estimate might be irrelevant for species with highly skewed or multimodal niches. For example, Gaussian niches would likely be a poor estimate of asymmetric niches constrained by exploitation barriers (Santamaría and Rodríguez-Gironés, 2007).

Third, CA estimates many latent traits (their number depend on the dimensionality of the original data, and is equal to $\min(r-1, c-1)$): when interpreting data, it is necessary to choose the number latent traits to investigate. In practice, ecologists should explore the first k axes/latent traits, until eigenvalues associated to these traits decrease abruptly (i.e. the marginal gain in explained variation becomes negligible). They may also use methods guiding the choice of dimensionality in multivariate analyses (similar to the method proposed by Dray, 2008, for principal component analysis).

Niche breadths have been studied using other measures in the literature, often based on the identity of interacting partners, as opposed to their traits (e.g. the index d', which also takes into account species availability, proposed by Blüthgen, Menzel, et al., 2006). A notable exception 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).

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 our data analysis, we removed species interacting with only one species, as CA is known to be sensitive to singleton observations and it has been advised to analyze modules separately with CA (Greenacre, 2013; van Dam et al., 2021). However, since specialization might hold an important biological signal, we kept other specialist species. Yet, the niches of specialist species are harder to estimate, as shown by simulation (Figure 3c) and empirical data analyses (Blüthgen and Staab, 2024).

In this article, we describe point estimates of niche breadths and optima, and do not consider the uncertainty around these measures. However, it would be interesting to evaluate the sensitivity of these measures to sampling effects, for instance by bootstrapping the data (i.e. iteratively removing species from the network and estimate niche parameters on these bootstrapped data). This would allow to estimate a confidence interval around niche breadths and optima.

Here, we use the parallel between environmental and interaction niches (Gravel et al., 2019), and apply two methods originally developed to study environmental niches (CA and reciprocal scaling) 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 applied to the study of species interactions (Dehling, Lai, et al., 2024; Godoy et al., 2018; Gravel et al., 2019).

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.

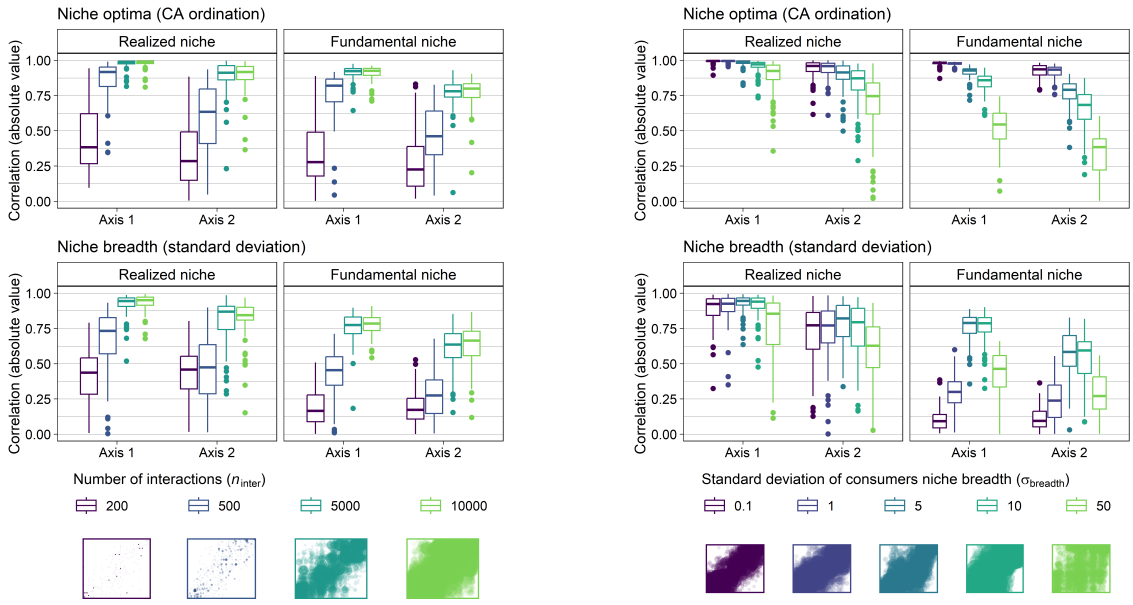
An alternative would be to use constrained versions of CA, like canonical correspondence analysis (using one trait table) (ter Braak, 1986) or double-constrained correspondence analysis (using both trait tables) (ter Braak et al., 2018). With canonical CA, we can explicitly quantify how much variation of the network is explained by known traits of resource or consumer species (Peres-Neto et al., 2006). Double-constrained CA opens the way to quantifying the part of variation explained by trait matching in ecological networks. These canonical methods would also prevent the separation of specialized species on different axes, even when a common latent trait can explain the network structure (Figures 3c and 4). Indeed, performing ordination with an explicit gradient would eliminate the need for niche overlap, which is necessary with latent gradients.

To our knowledge, constrained versions of correspondence analyses have never been applied to interaction networks. Moreover, for all methods above, there is no known way to measure niche breadth, because reciprocal scaling is restricted to CA. The part of variation explained by trait matching, that could be quantified with double-constrained CA, would also likely be biased, as this is the case for canonical CA (Peres-Neto et al., 2006) and fourth-corner analysis (Dray and P Legendre, 2008). Therefore, further methodological developments in multivariate analyses could benefit network analysis.

Appendices

A 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

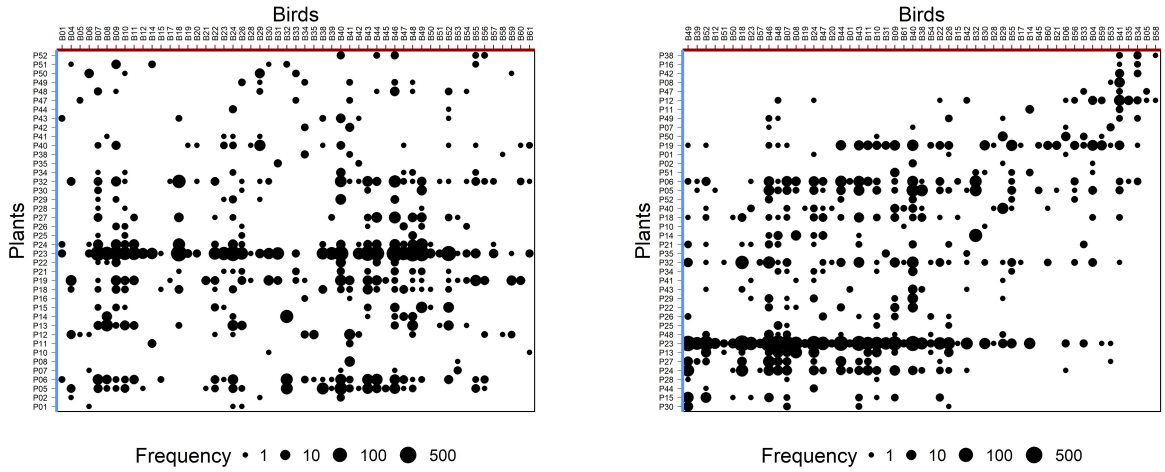
Figure A.1. Performance of the model to infer fundamental niche parameters for consumers. (a) Shows the effect of the sampling intensity and (b) shows the effect of the heterogeneity of consumers niche breadths on the model performance. The y-axis is the absolute value of the correlation between true and estimated values of niche optima and niche breadths (respectively top and bottom of each subplot). The correlation is presented separately for realized and fundamental niches (respectively left and right panels) on axes 1 and 2. Below the legend, matrices exemplify network data (resources in rows, consumer in columns) generated with the corresponding parameter value. Point sizes are explained in Figure 3.

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

B 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 B.1a, rows and columns are ordered according to the alphabetical order, while in Figure B.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 6.



(a) Raw matrix

(b) Reordered matrix

Figure B.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. Points represent non-null interactions, with size proportional to the number of observed interactions on a log scale.

C 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:

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

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

Similarly, for consumer species j, we have:

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

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:

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

Similarly, for consumer species j, we have:

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

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:

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

For consumer species j, the covariance is written as:

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

D 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 D.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 D.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 they have no conflict of interest relating to the content of this article.

Data, script, code, and supplementary information availability

Code and data are available at https://doi.org/10.6084/m9.figshare.28625270 (Nicvert et al., 2025).

References

Abdi H and M Béra (2017). Correspondence Analysis. In: *Encyclopedia of Social Network Analysis and Mining*. Ed. by Alhajj R and Rokne J. New York, NY: Springer, pp. 1–12. ISBN: 978-1-4614-7163-9. https://doi.org/10.1007/978-1-4614-7163-9_140-2.

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.

Beh EJ (2004). Simple Correspondence Analysis: A Bibliographic Review. *International Statistical Review* 72, 257–284. ISSN: 1751-5823. https://doi.org/10.1111/j.1751-5823.2004.tb00236.x.

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.

Blüthgen N and M Staab (July 26, 2024). A Critical Evaluation of Network Approaches for Studying Species Interactions. https://doi.org/10.1146/annurev-ecolsys-102722-021904.

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.

Dormann CF, J Fründ, and HM Schaefer (Nov. 2, 2017). Identifying Causes of Patterns in Ecological Networks: Opportunities and Limitations. *Annual Review of Ecology, Evolution, and Systematics* 48. ISSN: 1543-592X, 1545-2069. https://doi.org/10.1146/annurev-ecolsys-110316-022928.

Dray S (Jan. 10, 2008). On the number of principal components: A test of dimensionality based on measurements of similarity between matrices. *Computational Statistics & Data Analysis* 52, 2228–2237. ISSN: 0167-9473. https://doi.org/10.1016/j.csda.2007.07.015.

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.

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.

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.

Legendre P and L Legendre (Aug. 6, 2012). *Numerical Ecology*. Elsevier. 1008 pp. ISBN: 978-0-444-53868-0.

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.

Nicvert L, H Fritz, and S Dray (2025). *Code and data for: "Trait matching without traits: using correspondence analysis to analyze the latent structure of interaction networks"*. https://doi.org/10.6084/m9.figshare.28625270.

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.

Peres-Neto PR, P Legendre, S Dray, and D Borcard (2006). Variation Partitioning of Species Data Matrices: Estimation and Comparison of Fractions. *Ecology* 87, 2614–2625. ISSN: 1939-9170. https://doi.org/10.1890/0012-9658(2006)87[2614:VPOSDM]2.0.CO;2.

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.

Poisot T, DB Stouffer, and D Gravel (2015). Beyond species: Why ecological interaction networks vary through space and time. *Oikos* 124, 243–251. ISSN: 16000706. https://doi.org/10.1111/oik.01719.

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

Santamaría L and MA Rodríguez-Gironés (Jan. 23, 2007). Linkage Rules for Plant–Pollinator Networks: Trait Complementarity or Exploitation Barriers? *PLOS Biology* 5, e31. ISSN: 1545-7885. https://doi.org/10.1371/journal.pbio.0050031.

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.

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

Singer MC and CS McBride (2012). Geographic mosaics of species' association: a definition and an example driven by plant–insect phenological synchrony. *Ecology* 93, 2658–2673. ISSN: 1939-9170. https://doi.org/10.1890/11-2078.1.

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.

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.