National-Scale Asymmetric Biotic Associations Reveal Habitat Structure and Species Roles

Authors: Matthias Rohr[1], Sara Si-Moussi[1], Tamara Münkemüller[1] and Wilfried Thuiller[1]

- 1: Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LECA, F-38000 Grenoble, France

Abstract:

Habitat classifications such as EUNIS have proven powerful tools for mapping and monitoring biodiversity, but they remain fundamentally compositional: they describe which species are present, not how those species interact. Whether habitats differ in their underlying biotic functioning — the network of associations structuring communities from within, remains largely unknown. Here, we infer a large directed and asymmetric network of plant–plant associations from 59,390 open habitats plots across France (DivGrass), using a framework that explicitly separates how each species affects others from how it is itself affected by them. We aggregate these associations into a national metaweb of 672,411 significant links and derive local association networks for each site to characterize community biotic structure across EUNIS habitat classes.

The national network is overwhelmingly dominated by negative associations (98%), with asymmetric competitive interactions accounting for the majority of links — a pattern that would have been invisible to symmetric inference frameworks. Local network structure varies along a primary gradient opposing small, densely connected and relatively symmetric networks to larger, more asymmetric and centralized ones. This gradient partly reflects EUNIS habitat classification: wetland and coastal habitats consistently occupy one end of the spectrum, while heathlands show meaningful internal differentiation. However, the correspondence is imperfect, grasslands, the dominant habitat type, show little internal differentiation in network structure, and biotic organization does not map cleanly onto the EUNIS hierarchy. At the species level, nearly all species exert significant negative effects on their communities, while responses are

weaker and more variable. Species effects are best explained by habitat specialization, with specialists exerting stronger suppressive effects, and by leaf dry matter content (LDMC), likely through litter-mediated inhibition rather than direct resource competition. Together, these results demonstrate that the biotic interaction structure of plant communities captures ecologically meaningful variation that is only partially reflected in composition-based habitat classifications, supporting the integration of network-based approaches into habitat characterization and functioning.

1 Introduction

Habitats are spatially and environmentally coherent units within which species co-occur, interact, and collectively shape ecosystem processes. There are the fundamental units of biodiversity conservation and ecological research (Davies et al., 2004). Their identity emerges from two inseparable dimensions: the abiotic context (climate, soil, hydrology) that filters which species can persist, and the network of interactions among those species that determines how the community actually functions (Gaüzère et al., 2022; Tilman, 1982). These interactions, including competition, facilitation, and the full spectrum between them, underpin the ecological processes that generate the economic, cultural, and societal services habitats provide (Díaz et al., 2018; Pollock et al., 2012). Yet while the abiotic dimension of habitat identity is relatively well characterized, the biotic dimension, the network of interactions structuring communities within habitats, remains largely invisible in current classification frameworks.

The dominant approach to formalizing habitats, exemplified by the EUNIS classification, is rooted in phytosociology: habitat types are defined by the presence or dominance of diagnostic species under characteristic abiotic and geographic conditions (Chytrỳ et al., 2016, 2020; De Cáceres et al., 2015). This framework has proven powerful for mapping and monitoring biodiversity across Europe, and recent advances have substantially improved our ability to predict habitat types from environmental conditions (Si-Moussi et al., 2025b) or community composition alone (Leblanc et al., 2025). However, it remains fundamentally a compositional approach since it merely captures who is there, not how species interact. Two habitats sharing similar species richness and environmental conditions could differ profoundly in their competitive dynamics, facilitation networks, or the degree of asymmetry in dominant–subordinate relationships. Whether habitats differ in their underlying biotic functioning, and whether such variation maps onto existing classifications, are questions that compositional approaches cannot address. Doing so requires moving beyond species inventories to explicitly characterize the network of interactions among co-occurring species.

This gap is not merely conceptual. The structure of biotic interactions within a community has direct consequences for its stability, its response to perturbation, and its capacity to sustain ecosystem processes (Tilman, 1982; Desallais et al., 2025). Do productive grasslands exhibit stronger competitive interactions than coastal or wetland systems? Are alpine habitats more strongly structured by facilitative interactions, as suggested by the stress-gradient hypothesis (Callaway et al., 2002)? Addressing these questions requires a shift in perspective, from cataloguing species to characterizing the ecological roles they play relative to one another. This ambition is consistent with a broader movement in community ecology toward the biogeography of ecological networks (Thuiller et al., 2024), which asks how network structure varies across space and whether these changes are associated with classical shifts in habitat type (Pellissier et al., 2018; Rohr et al., 2026).

Inferring biotic interactions from observational community data, however, remains one of the central methodological challenges in ecology (Morales-Castilla et al., 2015; Blanchet et al., 2020), and is particularly difficult in plant communities where interactions are diffuse, context-dependent, and rarely directly observable (Rohr et al., 2026). Joint Species Distribution Models (JSDMs) have been proposed as a means of capturing signals of biotic interactions through residual covariance matrices (Ovaskainen et al., 2016), but their practical ability to reliably recover true interactions has proven limited (Poggiato et al., 2021). A deeper conceptual limitation is that residual correlations are by construction symmetric, and therefore can only represent a restricted subset of the full spectrum of biotic interactions, missing, in particular, the directional, asymmetric relationships that characterize competition and facilitation (Morales-Castilla et al., 2015).

Here, we build on a recently developed method (Si-Moussi et al., 2025a) that, moves beyond symmetric covariance by explicitly modelling, for each species, both how it affects other species and how it is in turn affected by them. This dual perspective draws on the response–effect framework long central to functional ecology (Lavorel and Garnier, 2002). Distinguish a species'

response to its community from its effect on that community offers a principled way to extract directional, asymmetric associations from co-occurrence data. Pairwise associations estimated in this way can be aggregated into a regional "metaweb" of plant–plant associations (Saravia et al., 2022), encoding all potential interactions across the study region. Local community compositions can then be sampled from this metaweb to characterize the structure of association networks within and across habitats (Losapio et al., 2019).

Representing species as nodes and asymmetric associations as directed links, network theory provides a rich framework for quantifying emergent community properties (e.g., degree, clustering, centralization) and for relating them to community functioning (Delmas et al., 2019), independently of species identity. For instance, strong asymmetry between species effects and responses combined with high centrality may indicate the dominance of a few strong competitors exerting disproportionate influence on the community (González et al., 2010; Daniel et al., 2024). Conversely, a highly symmetric network may reflect balanced fitness differences among species, more conducive to stable coexistence (Chesson, 2000). Comparing network structures across EUNIS habitat classes therefore offers a way to test whether habitats that differ compositionally also differ in their biotic organization and potentially to define a typology of habitats grounded in their functional structure rather than their species lists.

Beyond habitat level, directed networks also allows the characterization of individual species' roles within their communities (Delmas et al., 2019). The separation of effects from responses provides a means to identify species that exert strong competitive pressure on their community, those primarily involved in facilitative interactions, or those occupying highly central positions (Chillo et al., 2021; González et al., 2010) — offering a more mechanistic understanding of community organization than species identity or abundance alone.

These roles can then be related to species' functional traits and phylogenetic affiliations, which serve as proxies for environmental preferences and resource-use strategies (Violle et al., 2007; Mayfield and Levine, 2010; Rohr et al., 2025). In particular, this framework may shed

light on how ecological specialization shapes a species' role: specialist species may be expected to engage in stronger or more directional interactions within their communities, potentially acting as stronger competitors, while generalists may occupy more diffuse and symmetric positions in the network (Denelle et al., 2020).

We apply this framework to a large open-habitats dataset spanning the whole France (DIV-GRASS; Violle et al., 2015), inferring a directed, asymmetric network of plant–plant associations and using it to address three questions:

1. What is the overall structure of the regional association network? In particular, what is the balance between positive and negative associations, and how prevalent are asymmetric interactions?

2. Do habitats share common patterns of biotic structure, and can a typology of EUNIS habitat classes be defined based on their association networks structure?

3. Are species' roles within communities, in terms of effects and responses, associated with their functional traits, phylogeny, or ecological strategies such as specialization versus generalism?

2 Methods

2.1 Data

2.1.1 Vegetation data

The vegetation data originate from the DivGrass project (Violle et al., 2015), which compiled 96,441 botanical plots of permanent open-habitats across France, sampled at various points over the period 1913–2012. To ensure consistency with environmental data, we excluded plots with imprecise geolocation (uncertainty > 1,500 m) and those sampled before 1981, resulting in the removal of 31,296 plots.

Because our study focuses on herbaceous vegetation, we further excluded tree species based on growth form information from the GIFT database (Weigelt et al., 2020) and removed plots containing more than 10% tree cover (5,754 plots). The final dataset comprised 59,390 plots and 3,945 species (see Appendix S.1 for the spatial distribution of plots).

Most plots (43,582) included species relative cover data derived from phytosociological relevés. These consist of visually estimated cover values for all species within homogeneous plots, typically ranging from 25 to 100 m^2, using the six-level Braun–Blanquet abundance scale (Braun-Blanquet, 1932): 0%–1%, 1%–5%, 5%–25%, 25%–50%, 50%–75% and 75%–100%. The median of each class provides a quantitative scale. We converted these ordinal classes into quantitative values using their median of each interval. The remaining plots contained presence–absence data (Violle et al., 2015).

To ensure robust inference of biotic associations, we retained only species with quantitative cover data and at least 20 occurrences, resulting in a final set of 2,160 species.

2.1.2 Environmental data

To prevent species co-occurrence patterns driven by shared environmental preferences from being misattributed to biotic interactions, we included 32 environmental variables in the model (see

Appendix S.2 for the complete list of variables).

Climatic variables consisted of monthly means of air temperatures (°C) and monthly precipitation sums (mm) for the period 1981-2010, drawn from the 1 km resolution MeteoFrance Aurelhy gridded dataset (Canellas et al., 2014). Following Borgy et al. (2017), we derived a set of bioclimatic variables from a one-bucket water-balance model, including available water (AW) mean potential evapotranspiration, and two estimates of growing season length: the number of days with mean daily temperature above 5°C and the number of days satisfying both mean daily temperature above 5°C and available water exceeding 20% of total soil water holding capacity (GSL_{tw}).

Soil covariates were extracted from the 1km resolution SoilGrid database (Hengl et al., 2017).

2.1.3 Traits and taxonomic data

We selected five functional traits representing both resource-use strategies and competitive ability: leaf area (LA), leaf nitrogen content (LNC), specific leaf area (SLA), leaf dry matter content (LDMC), and vegetative plant height (Kunstler et al., 2012; Díaz et al., 2016). Trait data were obtained from the TRY 6.0 database (Kattge et al., 2020). Complete trait information was available for 1,665 species (77%), with trait-specific coverage ranging from 81% for LNC to 84% for plant height.

Taxonomic classification at the family level was retrieved from the World Flora Online (WFO) backbone (`10.5281/zenodo.18007552`).

To characterize the functional composition of communities, we computed community-weighted means (CWMs) for each trait, as well as functional richness, functional divergence, and functional evenness (Mason et al., 2005) for each site.

2.2 Habitat classification

To explore how plant associations networks vary across habitats, we used the EUNIS habitat classification (Chytrỳ et al., 2020). This three-level nested classification defines habitat types based on local dominance of diagnostic species, making it directly relevant to our aim of characterizing the roles these species play within their communities. Since the expert-based EUNIS assignment were not available for the DivGrass plot, we predicted habitat types used the Pl@ntBERT classifier (Leblanc et al., 2025), which accurately infers EUNIS classes from species composition alone. Habitats predictions supported by fewer than five plots occurrences were discarded, as these likely reflect classification errors..

2.3 Species strategies

To investigate how position along the generalist–specialist continuum relates to species' roles within communities, we quantified species specialization following (Fridley et al., 2007). This approach infers specialization from co-occurrence patterns: generalist species are expected to co-occur with a broad range of species, whereas specialists tend to co-occur with a narrower set of species, independently of their overall occurrence frequency.

For each of the 2,160 species, we thus computed a generality index θ using the R implementation provided by Fridley et al. (2007).

2.4 Species associations

2.4.1 Model overview

This subsection briefly introduces the key components of the modelling framework. For a complete description, see Si-Moussi et al. (2025a).

The framework jointly models species–environment responses and species–species associations, capturing both how species respond to environmental abiotic conditions and how they influence one another's abundance.

To infer bidirectional associations, the framework explicitly separates two complementary quantities for each species: its textiteffect on others, and its *response* to them. The effect embedding of a species characterizes the types of species it promotes or excludes when present. The response embedding describes the types of biotic contexts the species can tolerate. For example, a tall competitive grass may suppress light-demanding forbs (effect) while itself being relatively insensitive to the surrounding community (response).

The inferred association matrix A emerges from the combination of these effect and response embeddings, where each coefficient α_{ij} represents the influence of species j on species i (Figure 1.1). These pairwise associations are aggregated into a national-scale directed network in which species are represented as nodes and links correspond to positive or negative associations. Links may be either unidirectional (asymmetric associations) or bidirectional (symmetric associations).

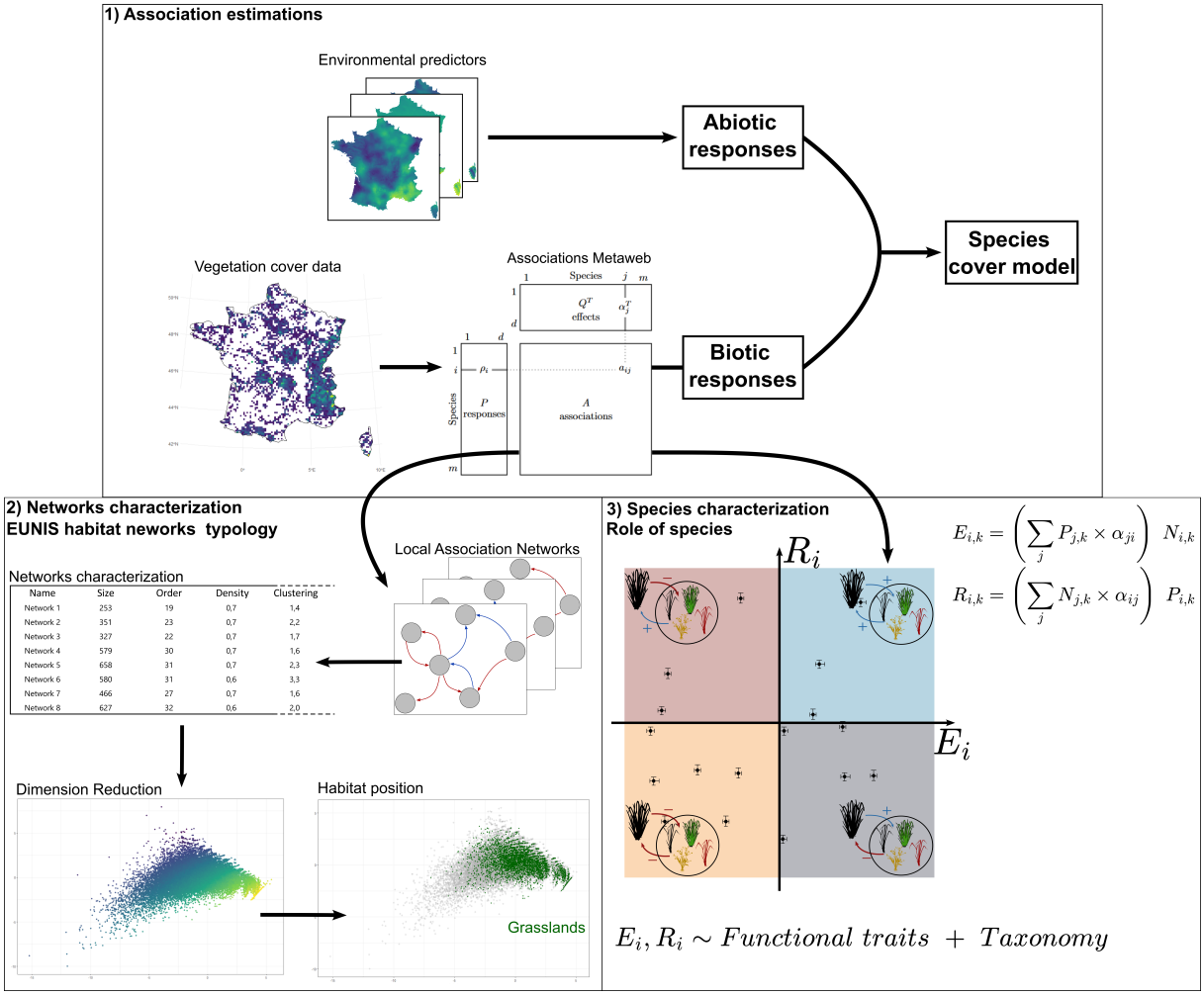


Figure 1: Overview of the analytical framework:
1) Estimation of species associations: The model predict the species vegetation cover by the combination of the species abiotic response and the biotic associations, which are estimate as a parameter of of the model.
2) EUNIS habitat network typology: The national-scale association network inferred in step (1) is treated as a metaweb, from which local association networks are derived for each sampling site. For each local network, a set of network metrics is computed to summarize its structure, and these metrics are projected into a reduced multidimensional space. EUNIS habitat classes are then mapped into this space to identify similarities and differences in their biotic structure.
3) Species roles: Using the directional nature of associations, we quantify for each species i its effect on community k ($E_{i,k}$) as the sum of its effects on co-occurring species (α_{ji}), weighted by their presence ($P_{j,k}$) and its relative abundance in the community ($N_{i,k}$). Conversely, the response of species i ($R_{i,k}$) is computed as the sum of the effects of co-occurring species on i (α_{ij}), weighted by their abundances ($N_{j,k}$). By summarizing the distribution of E_i and R_i across communities, species are positioned in an effect–response space that characterizes their overall interaction profile. These roles are then related to species' functional traits and taxonomic characteristics.

2.4.2 Generative model

We adopted a *hierarchical filtering hypothesis* in which the presence or absence of a species is governed solely by environmental suitability (Keddy, 1992; Kraft et al., 2015). Conditional on the species being present, its cover is shaped by biotic associations with co-occurring species. This two-stage hypothesis translates into a **Hurdle Logit-Normal** generative model for species cover.

Let $\mathbf{E} \in \mathbb{R}^{n \times p}$ denote the matrix of raw environmental covariates for n sites and p predictors, and let $\mathbf{Y} \in [0,1]^{n \times s}$ denote the matrix of observed species cover for s species. Additionally, let $\mathbf{X} \in \mathbb{R}^{n \times d}$ denote the matrix of *environmental embeddings*, i.e. low-dimensional representations of abiotic conditions obtained by transforming the raw environmental data through a neural network f_{enc} optimised for predicting species presence–absence:

$$\mathbf{X} = f_{\text{enc}}(\mathbf{E};\, W, B)\,, \tag{1}$$

where $\mathbf{W}$ and $\mathbf{B}$ collect the encoder weights and biases.

For a given species s, its cover in site i denoted as y_{is} follows a mixture distribution governed by a latent binary habitat suitability indicator h_{is} as follows:

Case $h = 1$ (presence). Conditional on presence, species cover follows a logit-Normal distribution on $(0, 1)$:

$$\operatorname{logit}(y_{i,k}) \sim \mathcal{N}\Big(\eta_{\text{biotic}},\, \phi_i^2\Big)\,, \tag{2}$$

where $\phi_i^2 > 0$ is the species-specific variance parameter and the biotic linear predictor is

$$\eta_{\text{biotic}} = o_i + \rho_i \sum \alpha_j y_{j,k} \tag{3}$$

Case $h = 0$ (absence). Conditional on absence, species cover is identically zero:

$$y_{i,k} \sim \delta_0, \tag{4}$$

where δ_0 denotes the Dirac point mass at zero.

The habitat suitability indicator $h_{is} \in \{0, 1\}$ follows a Bernoulli distribution whose success probability is determined by the abiotic linear predictor η_{abiotic}:

$$h \sim \text{Bernoulli}(\sigma(\eta_{\text{abiotic}}))\,, \tag{5}$$

$$\eta_{\text{abiotic}} = \beta\,\mathbf{X} + \beta_0, \tag{6}$$

where σ is the sigmoid (logistic) function, β is the vector of species-specific abiotic response parameters. The full parameter set Θ consists of:

- **Encoder parameters**: $\{\mathbf{W}, \mathbf{b}\}$ — weights and biases of the neural network f_{enc}.
- **Abiotic coefficients**: β, β_0 — linear predictor of species occurrence probability.
- **Biotic coefficients**: species response ρ and effect α embeddings.
- **Variance**: $\phi^2 > 0$ vector of species-specific logit-scale cover variance.

Parameters are estimated by maximising the likelihood using gradient-based optimisation (e.g. Adam), with regularisation constraints equivalent to prior distribution of parameters, yielding a maximum a posteriori (MAP) estimate.

2.4.3 Model fitting

Model fitting proceeded in two stages (Figure 2). In the first stage, we pre-trained the abiotic component using presence-absence data from all 59,390 sites, conventing quantitative cover values to binary where necessary.

We implemented a residual neural network architecture (Dollinger et al., 2024) and partitioned the data into training (70%), validation (15%) and test (15%) sets using 10 km spatial blocks to limit spatial autocorrelation. Hyperparameters — network width (number of neurons per hidden layer), depth (number of residual blocks), and regularisation (neuron dropout rate) — were tuned on the validation set (see Appendix S.3 for the final architecture). The model was optimized using AdamW with an asymmetric loss function to account for class imbalance between presences and absences.

In the second stage, we jointly fitted the abiotic and biotic components, initialising the environmental encoder with the pre-trained weights from stage one (transfer learning). In this step we aggregate the abiotic response $\hat{y}^A$ and the biotic response $\hat{y}^B$ following equation 7. The prediction of the cover of the species i at site k, $Y_{i,k}^{BB}$ is assumed to follow a LogitNormal distribution with a dispersion parameter ϕ_i. The dimensions of the biotic effect and response embeddings were tuned during this stage.

$$
\begin{aligned}
P_{i,k} &\sim \text{Bernoulli}(\text{sigm}(\hat{y}_{i,k}^A)) \\
Y_{i,k}^{BB} &\sim \begin{cases} \text{LogitNormal}(\hat{y}_{i,k}^B, \phi_i), & \text{if } P_{i,k} = 1 \\ \delta_0, & \text{otherwise} \end{cases}
\end{aligned}
\tag{7}
$$

The model was trained in batches of 128 over a maximum of 200 epochs, with early stopping triggered when validation loss failed to decrease for more than 20 iterations.

2.4.4 bootstrap

To assess the significative of inferred associations, we performed a bootstrap analysis to obtain confidence intervals around fitted association values. Sites were clustered into 1 km spatial blocks to preserve the local environmental structure essential for estimating biotic effects. At each of 100 bootstrap iterations, spatial blocks were resampled with replacement, retaining all observations within selected blocks; out-of-bag samples served as the validation set. The tuned

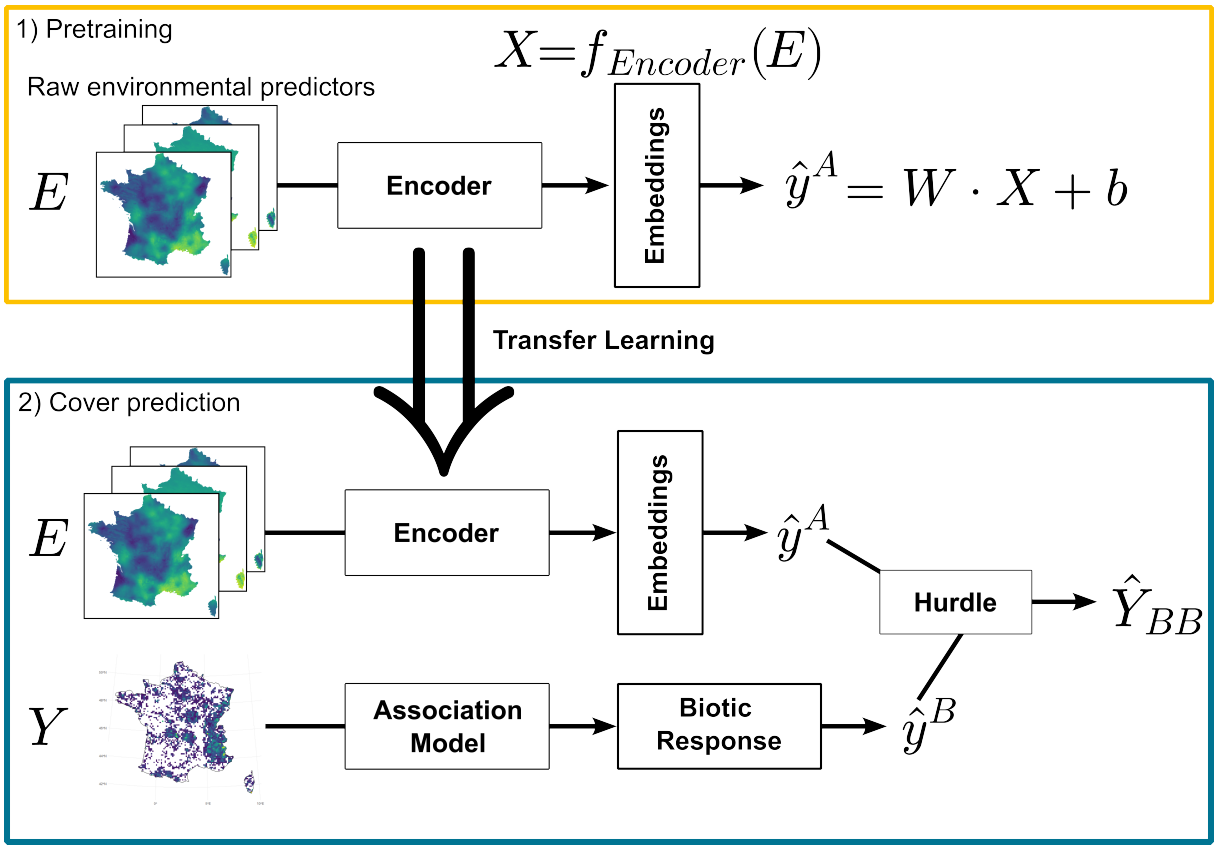


Figure 2: General model architecture. 1) In a first step we trained the model using the environmental predictors only in order to firstly estimate the species responses to the abiotic context. 2) The species cover ($\hat{Y}_{BB}$) are predicted using the joint biotic-abiotic model, with the environmental encoder weights have been initialized with the pre-trained model weight (transfer learning). The biotic ($\hat{y}^B$) and abiotic ($\hat{y}^A$) responses are integrated together through the Hurdle Logit-Normal generative model.

model architecture from Section 2.4.3 was used throughout. Associations whose 95% bootstrap confidence intervals overlapped zero were treated as non-significant and set to zero.

2.5 National network overview

All significant associations were aggregated into a single national-scale directed network. We computed a series of structual metrics for this network (Table 1) to characterize its overall structure. To assess the significance of clustering and centralization metrics, observed values were compared to null expectations derived from 99 Erdős–Rényi $G(n, m)$ random networks with identical numbers of nodes and links (Erdos et al., 1960).

Table 1: List of the network metrics and their ecological meanings used for characterised the local networks structure.

Metrics	Ecological meaning
Order (S)	Number of nodes (i.e species richness)
Size (L)	Number of links
Link density (d)	Mean number of association per node
Clustering (c)	Local cohesiveness of a group of species (the probability that two (or more) species associated with the same species are also associated with each others)
Centralisation (Ce)	Degree of concentration of link around few nodes (sum of the differences between each node's centrality, and the highest degree value). Declined in centralization "in" and "out".
Symmetric Negative associations	Proportion of symmetrical negative associations (-/-).
Asymmetric Negative association	Proportion of asymmetrical negative associations (-/0).
Symmetric Positive associations	Proportion of mutualist associations (+/+).
Asymmetric Positive associations	Proportion of asymmetric mutualist associations (+/0).
Dominance	Proportion of dominance associations (-/+).

2.6 EUNIS habitats networks typology

We used the national-scale association network as a "metaweb", a network encoding all potential associations within the regional species pool (Saravia et al., 2022). From this metaweb, local networks were derived for each site by retaining only the species present locally. To compare network structure across sites and EUNIS habitats, we applied a graph embedding approach that projects networks into a shared multidimensional space based on their structural properties (Figure 1.2; Botella et al., 2022).

Several approaches exist to represent networks in a common space. Here, we used an adapted version of the *Metrics2Vec* framework proposed by Botella et al. (2022). This method consists of computing m network metrics for each of the n sites, resulting in an $n \times m$ embedding matrix that can be analysed using multivariate methods.

Specifically, we used an adapted version of the *Metrics2Vec* framework (Botella et al., 2022), computing the set of metrics described in Table 1 for each site to produce an $n \times m$ embedding matrix. We adapted the original metric selection for non-trophic networks: we excluded the proportion of symmetric negative associations due to its redundancy with the asymmetric coun-

terpart, retained only overall centralization (rather than separate in- and out-centralization), and removed mutualism and dominance metrics due to their highly skewed distributions across sites.

Because clustering and centralization are sensitive to network size and order, we standardized these metrics using null models. For each site, we generated 99 Erdős–Rényi $G(n, m)$ random networks with identical numbers of nodes and links, and computed standardized effect sizes ($SES = (Obs - \text{mean}(random))/\text{sd}(random)$), where positive (negative) values indicate that observed metrics exceeds (falls below) random expectations

We then projected the resulting metric matrix onto its principal components (PCA) to visualize variation in network structure across sites, and mapped EUNIS habitat classes onto this space to identify similarities and differences in biotic organization. To identify which community-level properties explained variation along the primary axis of network structure space, we fitted a Random Forest regression model (Wright and Ziegler, 2017) with EUNIS habitat class and community-level functional metrics (see Section 2.1.3) as predictors, using the *Ranger 0.18.0* R package.

2.7 Species roles

For each species i present at site k, we computed two complementary quantities summarizing its position in the local association network. The *effect* of species i on community k, $E_{i,k}$ (Equation 8), is defined as the sum of its associations with co-occurring species, weighted by their presence and by species i's own relative abundance:

$$E_{i,k} = \left(\sum_{j} P_{j,k} \times \alpha_{ji} \right) \; N_{i,k} \tag{8}$$

The *response* of species i to community k, $R_{i,k}$, is defined as the sum of the effects of co-occurring species on species i, weighted by their relative abundances, conditional on the presence of species i:

$$R_{i,k} = \left(\sum_{j} N_{j,k} \times \alpha_{ij} \right) P_{i,k} \tag{9}$$

The distributions of $R_{i,k}$ and $E_{i,k}$ across communities characterize each species' general interaction profile (Figure 1.3). For example, a dominant competitor may exert an overall negative effect on co-occuring species while itself showing a neutral or positive response to the community.

To investigate the relationship between species' mean effects and responses and their functional traits, taxonomy, and position along the generalist–specialist continuum, we fitted two Random Forest regression models, one for effects and one for responses, removing highly correlated predictors prior to fitting using the *findCorrelation* function from the *caret* R package (Kuhn and Max, 2008). The model for species effects showed meaningful explanatory power while the response model did not; we therefore focused subsequent analyses on effects, fitting a generalized linear model (GLM) with a Gamma distribution applied to $(-Effect)$ (the opposite of mean effect, which was strictly negative for all species).

3 Results

Following our three research questions, we first describe the structure of the national network of plant–plant associations, then examine whether local network structure varies systematically across EUNIS habitat classes, and finally relate species' roles within communities to their characteristics to their functional and ecological characteristics..

3.1 National association network

We identified 672,411 significant biotic associations involving nearly all species, resulting in a directed national network with relatively low link density. The vast majority of these associations were negative (98%), with asymmetric competitive associations (−/0) accounting for more than half of all links (57.3%). Positive associations, whether symmetric (+/+) or asymmetric (+/0) were nearly absent (<1% combined), as were dominance interactions (−/+, 1%). This underscores the added value of allowing for asymmetric associations. Symmetric inference framework would have missed the dominant interaction type in the network.

The network exhibited significantly higher clustering and centralization than expected under random graph models ($p < 0.001$ for all metrics), indicating that species associations are organized in a strongly non-random manner, with links concentrated around a subset of highly connected nodes.

Table 2: Summary of network-level metrics at the national scale.

Metric	Size	Order	Link density	Clust.	Cent. in	Cent. out	Sym. Neg. Assoc.	Asym. Neg. Assoc.	Sym. Pos. Assoc.	Asym. Pos. Assoc.	Dom.
Value	672,411	2,158	0.14	0.44	0.72	0.38	41%	57.3%	0.01%	0.8%	1%

3.2 EUNIS habitats networks typology

The PCA performed on local network metrics (see Appendix S.4) revealed a main gradient opposing large, highly asymmetric and centralized networks to smaller, more densely connected networks characterized characterized by a higher proportion of symmetric negative associations.

The first two axes explained 63% and 16% of the total variance, respectively.

A Random Forest analysis (R^2 = 0.66, MSE(OOB) = 1.5) identified the variables most strongly associated with site positions along the first PCA axis: EUNIS habitat classes at levels 2 and 3, and functional richness of leaf area.

Projecting the median coordinates of EUNIS level 2 habitat classes (with first and third quartiles) into the PCA space (Figure 3; see Appendix S.5 for the full site projection) reveals that network structure is partly, but not fully, associated with habitat type. Coastal (N) and wetland (Q) habitats tend to exhibit smaller, more densely connected networks, occupying coherent and distinct regions of the PCA space. Grasslands, by contrast, cluster near the centre of the PCA space, reflecting less differentiated structural characteristics substantial intra-class variability.

Heathland habitats show a clearer internal structure along the gradient defined by link density and asymmetric negative associations. Garrigue systems (S6) are positioned towards more asymmetric interactions, while temperate heathlands (S4) form more densely connected networks, and maquis systems are associated with larger networks overall.

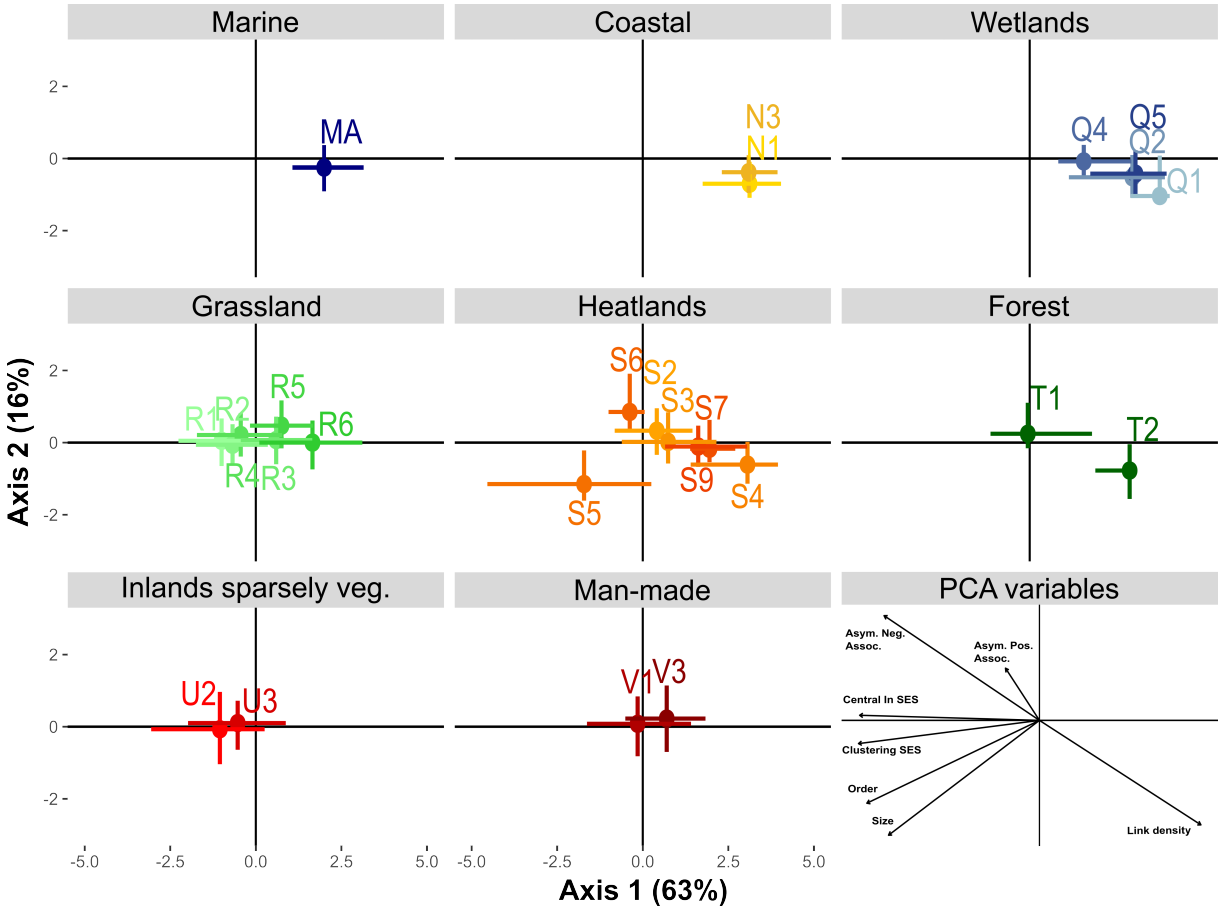


Figure 3: Projection of EUNIS habitat classes in the PCA space based on local network metrics. Each panel represents a level 1 EUNIS habitat category, with points corresponding to level 2 habitat subclasses. Points indicate median positions, with bars representing the first and third quartiles. The first two axes describe a gradient opposing large, highly asymmetric and centralized networks (upper-left) to smaller, more densely connected networks characterized by higher link density and more symmetric associations (lower-right). Arrows in the bottom-right panel indicate the direction and contribution of network metrics to the first two principal components.

3.3 Role of species

We quantified, for each species, its mean effect and response to their communities (Figure 4). Nearly all species (99% of the species, 2,151 of 2,160) exhibited significant non-zero mean effects on their communities, whereas a smaller proportion (61%, 1,327 species) showed significant responses. All species with significant effects exhibited negative effects, and the large majority of significant responses were also negative (99%). A small subset of species (23) showed a combination of negative effects and positive responses, suggesting net suppression of competitors alongside a facilitative return signal from the community.

A third of species (33%) exhibited effects that were stronger in magnitude than their responses, that is, they exerted more negative effects on their communities than the community exerted on them. These species tended to be significantly less generalist than others ($t = -9.3,\ p < 2 \times 10^{-16}$).

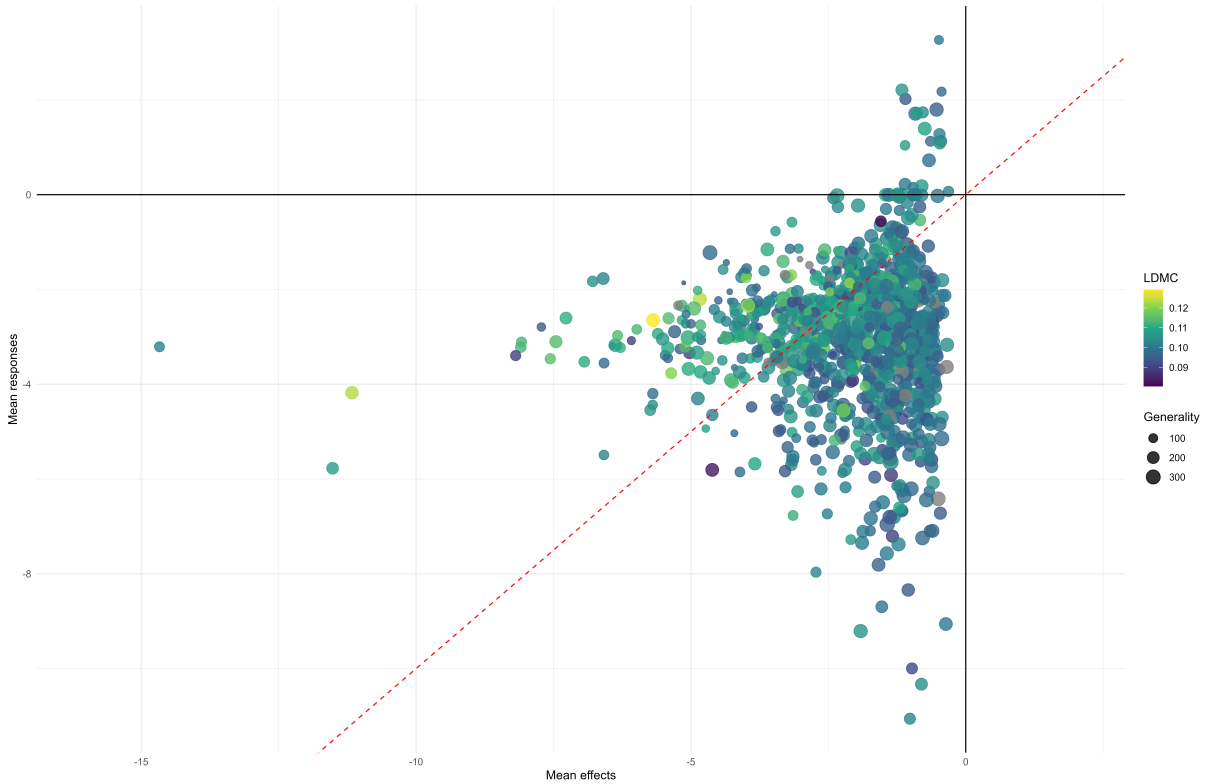


Figure 4: Species mean effects and responses to their communities. Each point represents a species; point size is proportional to its generality index, and color indicates leaf dry matter content (LDMC). The red dashed line represents the 1:1 isocline. Only species with significant (non-zero) mean effects and responses are shown.

We investigated the relationship between species' mean effects and responses and their ecological characteristics, including functional traits, taxonomy (i.e., family affiliation), and demographic strategies (i.e., specialist vs. generalist). We fitted two models, one for species effects and one for responses. Only the effect model showed satisfactory explanatory power (adjusted $R^2 = 0.34$, compared to *adj.* $R^2 = 0.06$ for the response model).

The GLM fitted to species mean effects identified two significant predictors. First, more generalist species tended to exert weaker (less negative) effects on their communities, as indicated by a positive coefficient for the generality index. Second, higher LDMC was associated with stronger negative effects on co-occurring species (Figure 5).

Taxonomic identity also contributed significantly: certain families, notably *Ericaceae* and *Poaceae*, exhibited systematically more negative effects than the overall mean .

Variable	Coefficient	Sdt Error	P
Generality	$1 \cdot 10^{-3}$	$1 \cdot 10^{-4}$	$< 2 \cdot 10^{-16}$
LDMC	-4.1	1.6	0.01

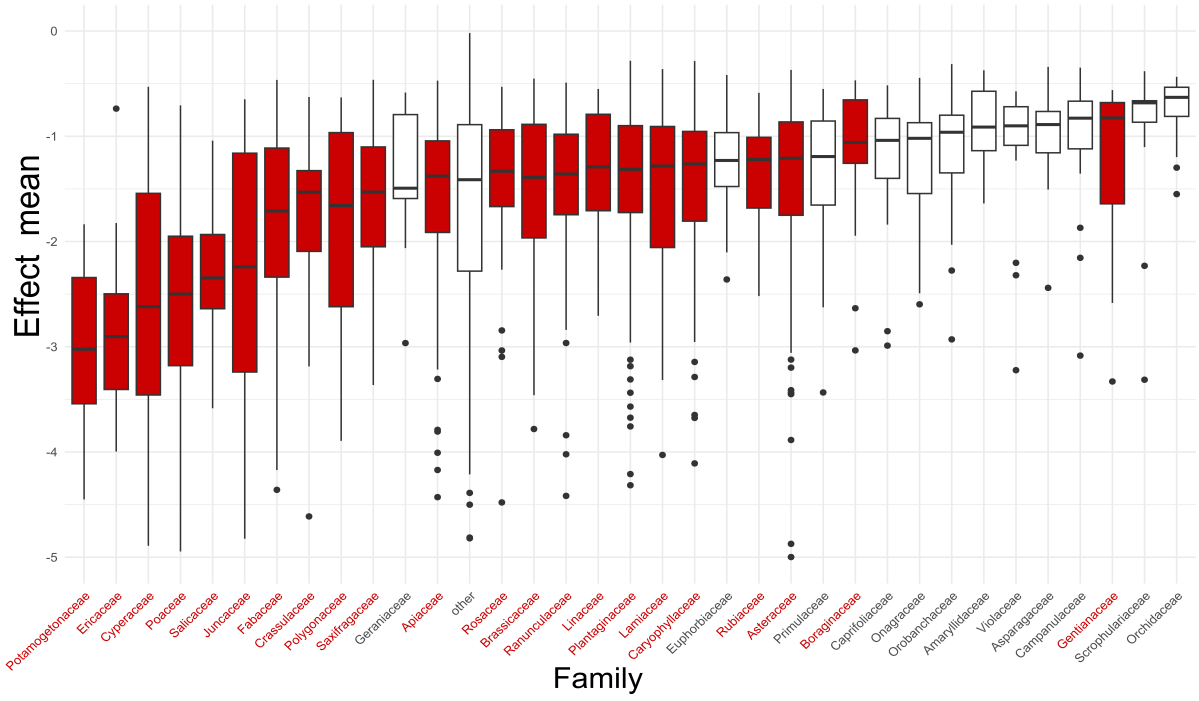


Figure 5: Results of the generalized linear model relating species mean effects to ecological characteristics. The upper panel summarizes the estimated coefficients for the main continuous predictors (generality index and LDMC). The lower panel shows the distribution of species mean effects across plant families. Red boxplots indicate families with significantly more negative effects.

4 Discussion

In this study, we inferred a large directed network of plant–plant associations from grassland communities across France, and used it to characterize the biotic structure of communities at both the habitat and species levels. Three main findings emerge. First, the national association network is overwhelmingly dominated by negative and asymmetric interactions, with positive associations nearly absent. Second, local network structure is partly associated with EUNIS habitat classification, but similarities and differences among communities do not map cleanly onto the hierarchy of habitat types. Third, nearly all species exert significant negative effects on their communities, while responses are weaker and harder to explain; species effects are best predicted by the generality index and leaf dry matter content (LDMC), as well as by family-level taxonomy.

A network dominated by negative and asymmetric associations

The near-total dominance of negative associations in the inferred national network warrants careful interpretation. While competitive exclusion is widely recognized as a fundamental structuring force in plant communities (Mayfield and Levine, 2010), the near absence of positive associations is somewhat unexpected given the substantial literature on facilitation in plant communities (Olofsson, 2004; Anthelme et al., 2014), particularly in mountain (Choler et al., 2001; Callaway et al., 2002) and drylands systems (Kéfi et al., 2008). This discrepancy may partly reflect a genuine ecological signal: grassland communities, which dominate our dataset, are generally more productive and competitive than the stress-dominated systems where facilitation is most prevalent (Bertness and Callaway, 1994). It may also, however, reflect a methodological asymmetry: our hurdle model conditions species cover on prior presence, which is governed by abiotic suitability. Facilitative interactions that primarily affect species occurrence rather than local abundance may therefore be absorbed into the abiotic component of the model and rendered invisible to the biotic association inference. This is an important limitation

to bear in mind when interpreting the prevalence of negative associations.

More ecologically informative is the predominance of asymmetric over symmetric negative associations, indicating that the effects of one species on another are frequently unidirectional. This asymmetry, combined with the high centralization of the network, suggests that a subset of dominant species exerts disproportionate suppressive effects on co-occurring species — consistent with a hierarchical competition model in which a few strong competitors structure community composition (Mayfield and Levine, 2010; Arnillas and Cadotte, 2019). This pattern stands in contrast to what a symmetric inference framework would have recovered, and underscores the importance of allowing for directional associations when characterizing biotic structure.

Biotic structure partly reflects habitat classification

Following recent calls to develop a biogeography of ecological interactions (Thuiller et al., 2024), we asked whether EUNIS habitat classes could be distinguished based on their local network structure. Our results reveal a primary axis of variation in network structure opposing small, densely connected and relatively symmetric networks to larger, more asymmetric and centralized ones. This gradient may reflect a transition from species-poor communities with relatively even competitive dynamics to more diverse systems dominated by a limited number of strong competitors (Arnillas and Cadotte, 2019).

Some habitat classes occupy coherent and interpretable regions of this network-structure space. Wetland and coastal habitats (Q and N) consistently form small, densely connected networks — a pattern potentially consistent with the importance of abiotic stress and facilitative dynamics in these systems (Bertness and Callaway, 1994), which may reduce the prevalence of strongly asymmetric competitive interactions. Heathlands show a meaningful internal gradient: garrigue systems (S6) tend towards more asymmetric structures, possibly reflecting organization around a small number of dominant shrub species (Díaz Barradas et al., 1999), while temperate heathlands (S4) form more densely connected networks, potentially indicating more even

competitive performance among co-occurring species and greater scope for stable coexistence (Chesson, 2000).

However, the overall correspondence between network structure and habitat classification is imperfect. Grasslands, the most abundant habitat class in our dataset, show the least differentiated network structure, with most subclasses clustering near the centre of the PCA space and no clear internal organization. More strikingly, the hierarchical structure of the EUNIS classification is not well reflected in the biotic structure space: ecologically contrasting habitats sometimes show unexpectedly similar network properties, while subclasses within the same level-1 category can be widely separated. This suggests that the biotic organization of communities captures an axis of habitat variation that is partly orthogonal to the compositional and abiotic axes on which the EUNIS classification is built, a finding that supports the broader argument for integrating interaction-based descriptors into habitat characterization (Thuiller et al., 2024).

Species effects are shaped by specialization and leaf economics

The finding that nearly all species exert significant negative effects on their communities, while only 61% show significant responses, mirrors the overall structure of the national network and is consistent with a community organization in which competitive suppression is widespread but asymmetric. The small subset of species combining negative effects with positive responses is particularly interesting: these species suppress competitors while themselves benefiting from the community context, a profile consistent with strong competitive dominants that exploit the biotic environment they help create.

Only species effects — not responses — could be explained by species characteristics, a result that is itself ecologically informative. Species responses to their communities are inherently more context-dependent: the same species may respond very differently to community composition depending on which co-occurring species happen to be present, making average responses across communities a noisy and difficult-to-predict quantity (Lavorel and Garnier,

2002). Species effects, by contrast, may be more consistent properties of species identity, rooted in stable functional attributes.

Among the predictors of species effects, the generality index emerged as the strongest: specialist species tend to exert stronger negative effects on their communities than generalists. This is consistent with the hypothesis that habitat specialists possess more competitive, locally optimized strategies that allow them to suppress co-occurring species more effectively within their focal habitat (Denelle et al., 2020). Generalists, by contrast, may be characterized by broader but shallower niches, resulting in weaker per-interaction effects.

The positive association between LDMC and the magnitude of negative effects deserves careful interpretation. LDMC is typically associated with a conservative, stress-tolerant resource-use strategy rather than with competitive dominance (Pierce et al., 2013), and a link between LDMC and competitive suppression may therefore seem paradoxical. However, in our framework, a negative biotic association reflects a reduction in the cover of co-occurring species rather than direct competitive exclusion in the classical sense. Species with high LDMC frequently produce slow-decomposing, recalcitrant litter that can acidify soils and inhibit the establishment or persistence of neighbouring species (Körner, 2003) — a suppressive mechanism that operates through the modification of local abiotic conditions rather than direct resource competition. This interpretation is further supported by the strong negative effects observed for *Ericaceae*, a family emblematic of high-LDMC, acidic-litter-producing species. The similarly strong effects of *Poaceae* may reflect a more conventional competitive mechanism, given that grasses are among the most effective resource competitors in grassland systems (Tilman, 1982).

Limitations and future directions

Our results depend critically on the reliability of the inferred biotic associations, and several sources of uncertainty deserve attention. Despite the model's explicit conditioning of biotic associations on abiotic responses, a non-negligible fraction of inferred links may still reflect un-

measured environmental variation rather than true species interactions (Blanchet et al., 2020; Poggiato et al., 2021). The relatively coarse spatial resolution of both the vegetation plots (approximately 1 km geolocation accuracy) and the environmental predictors means that ecologically distinct communities can share very similar environmental descriptors, increasing the risk of attributing shared compositional responses to abiotic gradients to spurious biotic associations. This is a plausible contributor to the dominance of negative associations in the inferred network, since unmodelled environmental heterogeneity tends to generate patterns of mutual exclusion in co-occurrence data.

Two complementary approaches could help address this limitation. First, incorporating habitat class at a coarser level directly as an environmental predictor could reduce the risk of inferring spurious associations between species that co-occur only because they share broad habitat affinities rather than direct biotic relationships. Second, integrating latent factors representing unmeasured ecological processes into the model framework — an approach increasingly adopted in joint species distribution modelling (Korhonen et al., 2025; Si-Moussi et al., 2026) — could absorb residual environmental variation and improve the specificity of inferred biotic signals.

Beyond inference, future work should explore how the network-based habitat typology developed here relates to other axes of habitat variation, including species compositional turnover and abiotic gradients, to better understand what additional information biotic structure provides over and above classical approaches. Extending the framework to other taxonomic groups and biomes would also allow testing of whether the dominance of negative, asymmetric associations is a general feature of plant communities or a characteristic of the productive grassland systems that dominate our dataset.

Take home message

Plant–plant association networks in French open-habitats communities are overwhelmingly dominated by negative and asymmetric interactions, suggesting that hierarchical competitive suppression is a widespread feature of community organization in these systems. Local network structure is partly, but imperfectly, reflected in EUNIS habitat classification: some habitat types — particularly wetlands and certain heathlands — occupy coherent regions of the network-structure space, while others, notably grasslands, show little internal differentiation. At the species level, competitive effects are strongest in habitat specialists and in species with high LDMC, pointing to both niche-based and litter-mediated mechanisms of community influence. Taken together, these findings support the emerging view that characterizing the biotic interaction structure of communities — rather than their composition alone — can provide new and complementary insights into habitat identity and functioning.

References

Anthelme, F., Cavieres, L. A., and Dangles, O. (2014). Facilitation among plants in alpine environments in the face of climate change. *Frontiers in plant science*, 5:387.

Arnillas, C. A. and Cadotte, M. W. (2019). Experimental dominant plant removal results in contrasting assembly for dominant and non-dominant plants. *Ecology Letters*, 22(8):1233–1242.

Bertness, M. D. and Callaway, R. (1994). Positive interactions in communities. *Trends in ecology & evolution*, 9(5):191–193.

Blanchet, F. G., Cazelles, K., and Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, 23(7):1050–1063.

Borgy, B., Violle, C., Choler, P., Denelle, P., Munoz, F., Kattge, J., Lavorel, S., Loranger, J., Amiaud, B., Bahn, M., et al. (2017). Plant community structure and nitrogen inputs modulate the climate signal on leaf traits. *Global Ecology and Biogeography*, 26(10):1138–1152.

Botella, C., Dray, S., Matias, C., Miele, V., and Thuiller, W. (2022). An appraisal of graph embeddings for comparing trophic network architectures. *Methods in Ecology and Evolution*, 13(1):203–216.

Braun-Blanquet, J. (1932). Plant sociology.

Callaway, R. M., Brooker, R. W., Choler, P., Kikvidze, Z., Lortie, C. J., Michalet, R., Paolini, L., Pugnaire, F. I., Newingham, B., Aschehoug, E. T., Armas, C., Kikodze, D., and Cook, B. J. (2002). Positive interactions among alpine plants increase with stress. *Nature*, 417(6891):844–848.

Canellas, C., Gibelin, A.-L., Lassègues, P., Kerdoncuff, M., Dandin, P., and Simon, P. (2014). Les normales climatiques spatialisées aurelhy 1981-2010: températures et précipitations. *La météorologie*, 2014(85):47–55.

Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual review of Ecology and Systematics*, 31(1):343–366.

Chillo, V., Vázquez, D. P., Tavella, J., and Cagnolo, L. (2021). Plant–plant co-occurrences under a complex land-use gradient in a temperate forest. *Oecologia*, 196(3):815–824.

Choler, P., Michalet, R., and Callaway, R. M. (2001). Facilitation and competition on gradients in alpine plant communities. *Ecology*, 82(12):3295–3308.

Chytrỳ, M., Hennekens, S. M., Jiménez-Alfaro, B., Knollová, I., Dengler, J., Jansen, F., Landucci, F., Schaminée, J. H., Aćić, S., Agrillo, E., et al. (2016). European vegetation archive (eva): an integrated database of european vegetation plots. *Applied vegetation science*, 19(1):173–180.

Chytrỳ, M., Tichỳ, L., Hennekens, S. M., Knollová, I., Janssen, J. A., Rodwell, J. S., Peterka, T., Marcenò, C., Landucci, F., Danihelka, J., et al. (2020). Eunis habitat classification: Expert system, characteristic species combinations and distribution maps of european habitats. *Applied Vegetation Science*, 23(4):648–675.

Daniel, C., Allan, E., Saiz, H., and Godoy, O. (2024). Fast–slow traits predict competition network structure and its response to resources and enemies. *Ecology Letters*, 27(4):e14425.

Davies, C. E., Moss, D., and Hill, M. O. (2004). Eunis habitat classification revised 2004. *Report to: European environment agency-European topic centre on nature protection and biodiversity*, pages 127–143.

De Cáceres, M., Chytrỳ, M., Agrillo, E., Attorre, F., Botta-Dukát, Z., Capelo, J., Czúcz, B., Dengler, J., Ewald, J., Faber-Langendoen, D., et al. (2015). A comparative framework for broad-scale plot-based vegetation classification. *Applied Vegetation Science*, 18(4):543–560.

Delmas, E., Besson, M., Brice, M.-H., Burkle, L. A., Dalla Riva, G. V., Fortin, M.-J., Gravel, D., Guimarães, P. R., Hembry, D. H., Newman, E. A., Olesen, J. M., Pires, M. M., Yeakel, J. D., and Poisot, T. (2019). Analysing ecological networks of species interactions: Analyzing ecological networks. *Biological Reviews*, 94(1):16–36.

Denelle, P., Violle, C., Consortium, D., and Munoz, F. (2020). Generalist plants are more competitive and more functionally similar to each other than specialist plants: insights from network analyses. *Journal of Biogeography*, 47(9):1922–1933.

Desallais, M., Arnoldi, J.-F., and Loreau, M. (2025). Partitioning net biodiversity effects on ecosystem resistance and resilience. *Ecology Letters*, 28(11):e70249.

Díaz, S., Kattge, J., Cornelissen, J. H., Wright, I. J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Colin Prentice, I., et al. (2016). The global spectrum of plant form and function. *Nature*, 529(7585):167–171.

Díaz, S., Pascual, U., Stenseke, M., Martín-López, B., Watson, R. T., Molnár, Z., Hill, R., Chan, K. M., Baste, I. A., Brauman, K. A., et al. (2018). Assessing nature's contributions to people. *Science*, 359(6373):270–272.

Díaz Barradas, M., Zunzunegui, M., Tirado, R., Ain-Lhout, F., and García Novo, F. (1999). Plant functional types and ecosystem function in mediterranean shrubland. *Journal of Vegetation Science*, 10(5):709–716.

Dollinger, J., Brun, P., Sainte Fare Garnot, V., and Wegner, J. D. (2024). Sat-sinr: High-resolution species distribution models through satellite imagery. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 10:41–48.

Erdos, P., Rényi, A., et al. (1960). On the evolution of random graphs. *Publ. math. inst. hung. acad. sci*, 5(1):17–60.

Fridley, J. D., Vandermast, D. B., Kuppinger, D. M., Manthey, M., and Peet, R. K. (2007). Co-occurrence based assessment of habitat generalists and specialists: A new approach for the measurement of niche width. *Journal of ecology*, 95(4):707–722.

Gaüzère, P., O'Connor, L., Botella, C., Poggiato, G., Münkemüller, T., Pollock, L. J., Brose, U., Maiorano, L., Harfoot, M., and Thuiller, W. (2022). The diversity of biotic interactions complements functional and phylogenetic facets of biodiversity. *Current Biology*, 32(9):2093–2100.e3.

González, A. M. M., Dalsgaard, B., and Olesen, J. M. (2010). Centrality measures and the importance of generalist species in pollination networks. *Ecological complexity*, 7(1):36–43.

Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M. N., Geng, X., Bauer-Marschallinger, B., et al. (2017). Soilgrids250m: Global gridded soil information based on machine learning. *PLoS one*, 12(2):e0169748.

Kattge, J., Bönisch, G., Díaz, S., Lavorel, S., Prentice, I. C., Leadley, P., Tautenhahn, S., Werner, G. D., Aakala, T., Abedi, M., et al. (2020). Try plant trait database–enhanced coverage and open access. *Global change biology*, 26(1):119–188.

Keddy, P. A. (1992). Assembly and response rules: two goals for predictive community ecology. *Journal of Vegetation Science*, 3(2):157–164.

Kéfi, S., Baalen, M. v., Rietkerk, M., and Loreau, M. (2008). Evolution of local facilitation in arid ecosystems. *The American Naturalist*, 172(1):E1–E17.

Korhonen, P., Hui, F. K., Niku, J., Taskinen, S., and van der Veen, B. (2025). gllvm 2.0: fast fitting of advanced ordination methods and joint species distribution models. *PeerJ*, 13:e20338.

Kraft, N. J. B., Adler, P. B., Godoy, O., James, E. C., Fuller, S., and Levine, J. M. (2015). Community assembly, coexistence and the environmental filtering metaphor. *Functional Ecology*, 29(5):592–599.

Kuhn and Max (2008). Building predictive models in r using the caret package. *Journal of Statistical Software*, 28(5):1–26.

Kunstler, G., Lavergne, S., Courbaud, B., Thuiller, W., Vieilledent, G., Zimmermann, N. E., Kattge, J., and Coomes, D. A. (2012). Competitive interactions between forest trees are driven by species' trait hierarchy, not phylogenetic or functional similarity: implications for forest community assembly. *Ecology Letters*, 15(8):831–840.

Körner, C. (2003). *Alpine plant life: functional plant ecology of high mountain ecosystems*. Springer, 2nd edition.

Lavorel, S. and Garnier, E. (2002). Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the holy grail. *Functional ecology*, 16(5):545–556.

Leblanc, C., Bonnet, P., Servajean, M., Thuiller, W., Chytrỳ, M., Aćić, S., Argagnon, O., Biurrun, I., Bonari, G., Bruelheide, H., et al. (2025). Learning the syntax of plant assemblages. *Nature Plants*, pages 1–15.

Losapio, G., Montesinos-Navarro, A., and Saiz, H. (2019). Perspectives for ecological networks in plant ecology. *Plant Ecology & Diversity*, 12(2):87–102.

Mason, N. W., Mouillot, D., Lee, W. G., and Wilson, J. B. (2005). Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos*, 111(1):112–118.

Mayfield, M. M. and Levine, J. M. (2010). Opposing effects of competitive exclusion on the phylogenetic structure of communities. *Ecology Letters*, 13(9):1085–1093.

Morales-Castilla, I., Matias, M. G., Gravel, D., and Araújo, M. B. (2015). Inferring biotic interactions from proxies. *Trends in Ecology & Evolution*, 30(6):347–356.

Olofsson, J. (2004). Positive and negative plant-plant interactions in two contrasting arctic-alpine plant communities. *Arctic, Antarctic, and Alpine Research*, 36(4):464–467.

Ovaskainen, O., Abrego, N., Halme, P., and Dunson, D. (2016). Using latent variable models to identify large networks of species-to-species associations at different spatial scales. *Methods in Ecology and Evolution*, 7(5):549–555.

Pellissier, L., Albouy, C., Bascompte, J., Farwig, N., Graham, C., Loreau, M., Maglianesi, M. A., Melián, C. J., Pitteloud, C., Roslin, T., Rohr, R., Saavedra, S., Thuiller, W., Woodward, G., Zimmermann, N. E., and Gravel, D. (2018). Comparing species interaction networks along environmental gradients: Networks along environmental gradients. *Biological Reviews*, 93(2):785–800.

Pierce, S., Brusa, G., Vagge, I., and Cerabolini, B. E. (2013). Allocating csr plant functional types: the use of leaf economics and size traits to classify woody and herbaceous vascular plants. *Functional Ecology*, 27(4):1002–1010.

Poggiato, G., Münkemüller, T., Bystrova, D., Arbel, J., Clark, J. S., and Thuiller, W. (2021). On the Interpretations of Joint Modeling in Community Ecology. *Trends in Ecology & Evolution*, page S0169534721000045.

Pollock, L. J., Morris, W. K., and Vesk, P. A. (2012). The role of functional traits in species distributions revealed through a hierarchical model. *Ecography*, 35(8):716–725.

Rohr, M., Thuiller, W., Chalmandrier, L., and Münkemüller, T. (2025). The role of observation scales, trait correlations, and competitive regimes in community assembly patterns. *Ecology and Evolution*, 15(4):e71272.

Rohr, M., Wendling, A., Münkemüller, T., Gravel, D., Galiez, C., Renaud, J., Consortium, O., and Thuiller, W. (2026). Beyond co-occurrence: Multi-scale evidence for segregation-dominated plant networks in the french alps. *Ecology Letters*, 29(5):e70393. e70393 ELE-01258-2025.R4.

Saravia, L. A., Marina, T. I., Kristensen, N. P., De Troch, M., and Momo, F. R. (2022). Ecological network assembly: How the regional metaweb influences local food webs. *Journal of Animal Ecology*, 91(3):630–642.

Si-Moussi, S., Galbrun, E., Hedde, M., Poggiato, G., Rohr, M., and Thuiller, W. (2025a). Uncovering symmetric and asymmetric species associations from community and environmental data. *arXiv preprint arXiv:2507.09317*.

Si-Moussi, S., Hennekens, S., Mücher, S., De Keersmaecker, W., Chytrỳ, M., Agrillo, E., Attorre, F., Biurrun, I., Bonari, G., Čarni, A., et al. (2025b). Eunis habitat maps: enhancing thematic and spatial resolution for europe through machine learning. *Scientific Data*, 12(1):1976.

Si-Moussi, S., Thuiller, W., Galbrun, E., Decaëns, T., Gérard, S., Marchán, D., Marsden, C., Capowiez, Y., and Hedde, M. (2026). Digging deeper: deep joint species distribution modeling reveals environmental drivers of earthworm communities. *Soil Biology and Biochemistry*, 212:110021.

Thuiller, W., Calderón-Sanou, I., Chalmandrier, L., Gaüzère, P., O'Connor, L. M., Ohlmann, M., Poggiato, G., and Münkemüller, T. (2024). Navigating the integration of biotic interactions in biogeography. *Journal of Biogeography*, 51(4):550–559.

Tilman, D. (1982). *Resource competition and community structure*. Number 17. Princeton university press.

Violle, C., Choler, P., Borgy, B., Garnier, E., Amiaud, B., Debarros, G., Diquelou, S., Gachet, S., Jolivet, C., Kattge, J., et al. (2015). Vegetation ecology meets ecosystem science: Permanent grasslands as a functional biogeography case study. *Science of the Total Environment*, 534:43–51.

Violle, C., Navas, M.-L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I., and Garnier, E. (2007). Let the concept of trait be functional! *Oikos*, 116(5):882–892.

Weigelt, P., König, C., and Kreft, H. (2020). Gift–a global inventory of floras and traits for macroecology and biogeography. *Journal of Biogeography*, 47(1):16–43.

Wright, M. N. and Ziegler, A. (2017). ranger: A fast implementation of random forests for high dimensional data in c++ and r. *Journal of statistical software*, 77:1–17.