Predicting interaction frequency in plant-pollinator networks

Carsten F. Dormann *[1], William J. Castillo[1], María P. Pascual Tudanca [2], Paul CaraDonna [3,4,20], Laura A. Burkle [6], Demetra Rakosy [7,8], Rachel Winfree [9], Leana Zoller [10], Gita Benadi [11], Nico Blüthgen [12], Christopher N. Kaiser-Bunbury [13], Jochen Fründ [14], Florian Hartig [15], Tiffany M. Knight [6,16,17], Konstantinos Minachilis [18], Maximilian Pichler [19], Julian Resasco [10], Theodora Petanidou [18], Boris A. Tinoco [19], Guadalupe Peralta† [20], and Diego P. Vázquez† [2,21]

[1] *Biometry & Environmental System Analysis, University of Freiburg, Germany*
[2] *Argentine Institute for Dryland Research, National Council for Scientific and Technical Research (CONICET)–National University of Cuyo, Mendoza, Argentina*
[3] *Chicago Botanic Garden, Negaunee Institute for Plant Conservation Science and Action, Glencoe, IL, USA*
[4] *Plant Biology and Conservation, Northwestern University, Evanston, IL, USA*
[5] *Rocky Mountain Biological Laboratory, Crested Butte, CO, USA*
[6] *Department of Ecology, Montana State University, Bozeman, MT, USA*
[7] *German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany*
[8] *Institute of Biodiversity, Thünen Institute, Braunschweig, Germany*
[9] *Department of Ecology, Evolution, and Natural Resources, Rutgers University, New Brunswick, NJ, USA.*
[10] *Department of Ecology & Evolutionary Biology, University of Colorado, Boulder, USA*
[11] *Institute of Medical Biometry and Statistics, Medical Faculty and Medical Center, University of Freiburg*
[12] *Department of Biology, Technical University of Darmstadt, Germany*
[13] *Centre for Ecology and Conservation, Faculty of Environment, Science and Economy, University of Exeter, Penryn, UK*
[14] *Animal Network Ecology, Department of Biology, University of Hamburg, Germany*
[15] *Theoretical Ecology, Faculty of Biology and Pre-Clinical Medicine, University of Regensburg, Regensburg, Germany*
[16] *National Tropic Botanic Garden, Hawai'i, USA*
[17] *Institute of Biology, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany*
[18] *Laboratory of Biogeography & Ecology, University of the Aegean, Department of Geography, Mytilene, Greece*
[19] *School of Biology, Ecology and Management, University of Azuay, Cuenca, Ecuador*
[20] *Multidisciplinary Institute of Plant Biology, National Council for Scientific and Technical Research (CONICET)– National University of Córdoba, Córdoba, Argentina*
[21] *Faculty of Exact and Natural Sciences, National University of Cuyo, Mendoza, Argentina*

March 16, 2025

---

*corresponding author; email: carsten.dormann@biom.uni-freiburg.de

†: Joint senior authorship.

Abstract

Flowers and their pollinators represent a bipartite interaction system, whose links are hypothesised to be related to species traits. To explore whether we can predict the weight of this link, i.e. the frequency of interactions, in an validation network, we analysed 14 studies of pollinator-flower visitation network from around the world.

We used information on species abundances, their traits and their phylogenetic (for plants) or taxonomic (for animals) position as predictors of interaction frequency, and fitted different statistical modelling approaches. We expected to see prediction quality on validation data to decay with spatial and temporal distance to the training networks. Similarly, we expect that changes in pollinator or plant composition will negatively affect predictive performance.

Using the best-predicting modelling approach (randomForest), we indeed see a slight decay in predictive quality with plant and pollinator compositional distance. Temporal distance played little role, although predictions for one year ahead (or back) were better than across the season or across multiple years.

The overall predictive power of our models was low (Spearman's $\rho \approx 0.4$), suggesting a very noisy system. Also, the most important predictor was abundance, as revealed by a parameter-free benchmark model that only used the cross-product of abundances to predict interaction frequency. Trait and phylogenetic information did not substantially improve predictive performance beyond abundance-based predictions. Across all studies, we failed to confirm a substantial contribution of ecological characteristics to pollinator-flower interaction frequency.

One reason why predictions were relatively poor is that sampling effort is not standardised, and thus networks differed substantially in the observed number of interactions, network size, and interaction density. Also the pooling of networks across space or across time may have diluted preferences in the data, reducing their explanatory value. Finally, the majority of species in each network are rare, and the interaction information they provide may be much less relevant that that of common species.

At present, we conclude that the frequencies of interactions are very difficult to predict, and using traits we cannot really do better than simply using abundance information.

Keywords:
flower visitation, machine learning, pollination, prediction, trait-matching

1 Introduction

Many ecological phenomena exhibit strong variability, resulting in low proportions of explained variance, and poor predictive performance of statistical as well as process models (Doak *et al.* 2008; Mouquet *et al.* 2015). Aggregate phenomena, such as biomass or species richness, average out the variability observed at the level of species, populations and individuals, making them more reliable targets for ecological research. But researchers in evolution, conservation and ecology are particularly interested in individual- and population-level processes. In pollination ecology, for some plant species, and many of the pollinators, understanding which species they interact with is relevant: rare plant species (Robson 2013; Rodger *et al.* 2021) or species of commercial interest (Klein *et al.* 2007; Ollerton *et al.* 2011; Winfree *et al.* 2018) may crucially depend on these specific interactions to overcome pollen limitation. We want to understand *why* certain pollinators only visit particular plants, regardless even of how important the interaction is for the species' fitness.

The tools to achieve understanding include causal explanation, generalisation, and testing (Pickett *et al.* 2007). Thus, by attempting to predict interactions we test our causal explanations of observed interactions, generalising these explanations by testing them on a number of different study systems. Being able to predict interactions allows us to test our understanding of the processes governing interactions, and such understanding paves the way for any application including management and conservation of biodiversity in natural and human-dominated ecosystems (Peralta *et al.* 2024).

Species interact with other species in different frequencies that vary by orders of magnitude, and such differences may strongly determine (although are not the same as) interaction strengths (the ecological impact of one species on another: Berlow *et al.* 2004). Thus, considering quantitative interactions gives us more information about the ecological processes we want to understand, and in that sense prediction of links alone represents a limited gain in understanding. Either way, it is unclear how well we may be able to predict the frequency of interactions for such networks. We are aware of only two previous studies that attempted to predict interaction frequency, Pichler *et al.* (2019) and Benadi *et al.* (2022), and both are technical studies that test approaches in simulations and on a single exemplary data set. These results suggest that species abundances play an important role for observed interaction frequency: if species interact according to their relative abundances (p_i for flower i and p_j for pollinator j, each computed as the observed divided by the total number of interactions n), the expected frequency np_ip_j for any given link was close to the observed data in these studies. Note that this implies that rare species will be observed much less, infusing the data with false missing links.

When attempting to predict interactions, we should start by attempting to predict links: the presence–absence of each pairwise species interaction—who interacts with whom. For plant-pollinator networks, links *within a network* can be predicted moderately well, with AUC values of around 0.7 for the best models on simulated data (Terry & Lewis 2020). Yet, predicting the presence or absence of links between species differs from predicting the actual frequency of those interactions: how many times each pair of species interacts. Predictive interpolation of links (i.e. *within networks*) is also substantially easier than extrapolating links of new networks, i.e. when both plants and pollinator composition changes. This problem becomes even larger for interaction frequency, when not merely the existence but the intensity of an interaction is to be estimated for new sites or times.

Network ecologists assume causal links between species traits and species interactions. That is, interactions are not entirely idiosyncratic and species-specific, but rather interactions are more likely when traits match, or indeed absent for “forbidden” trait combinations (reviewed in González-Varo & Traveset 2016). In consequence, a model that is able to represent the causal drivers of pollinator interactions through the use of traits and abundances should be able to predict well to other sites even with different plant and pollinator species. The key challenge is to identify which traits are relevant. Pollinators use vision and scent to judge a flower’s attractiveness, yet ecologists primarily measure easily accessible characteristics, such as body size and flower morphology. To date, we have no comprehensive way to assess biochemical trait-matching, let alone a data base with estimates for more than a handful of pollinator species. If scent- and vision-related traits are phylogenetically conserved to some degree, one can use phylogenetic relatedness as surrogate of the missing traits (Smith 2010, but see van der Niet & Johnson 2012). That is the idea behind augmenting the measured traits with phylogenetic eigenvectors, which in their combination represent “latent” (= unobserved) traits within a branch of the phylogeny.

A strong test of prediction requires fitting (training) the models and then testing them with an independent data set. Such independence increases with spatial or temporal distance to the training data. When fitting a model to a set of networks, it is arguably more challenging to predict a network at the same site a few years into the future, or at a very different site, than to networks within the training space or time. Indeed, a model’s predictive ability should decrease with spatial or temporal distance between training and test data. This distance decay of predictive ability comes from the heterogeneous distribution of species in time and space, which causes training and test communities to become increasingly different with distance. We may hypothesise that across space, plant turnover may drive predictability: plants depend strongly on soil, climate and management, and hence may change substantially across sites, but much less over the years (but well over the course of a season). On the other hand, pollinator variability may drive predictability across time, as their abundance fluctuates from year to year, while they can fly and may thus seek out whatever their preferred floral resources are even across larger spatial scales. Thus, testing our ability to predict interactions depends on the availability of datasets that can be split into training data and (independent) test data.

In this study we explore 14 previously published studies of plant-pollinator networks to explore the connection between compositional distance and predictability. We theorise that plants and pollinators interact as a function of their respective abundances and the match of their interaction traits. A statistical model can learn which trait or trait combinations are important for interactions, and hence predict unobserved interactions based on these traits. If true, we can thus expect that (i) abundances determine the basic interaction frequency, such that the cross-product of relative abundances should give a good indication of observed frequencies. We hypothesise (ii) that this basic frequency is increased when flowers and pollinators exhibit matching traits, and reduced if they do not, leading to traits improving frequency predictions. As a result, we expect (iii) to see predictions to networks comprising similar species being more successful than to those with a very different set of plants and/or pollinators. Temporal and spatial distance are thus merely proxies for species compositional distances.

In the following, we outline the data included in our database, the methods used to analyse and predict plant–pollinator interactions, and the within-study analysis of predictive quality. Although dozens of publications report plant-pollinator networks (for example, Terry & Lewis 2020, used 48 of them), our hypotheses required strict inclusion criteria: quantitative networks replicated in space and/or time, with sufficient information to allow construction of phylogenetic or at least taxonomic trees (to construct latent traits, see Methods). Also, our focus is the prediction of interaction frequency, not merely occurrence of a pollination event as in most previous studies on predicting pollination interactions (e.g. Terry & Lewis 2020; Sydenham *et al.* 2022, 2024). We analysed each case study separately but present the trends across studies.

2 Methods

We first describe which kind of data we assembled for each case study (interaction networks, traits, abundances, phylogeny and taxonomy), before outlining the methods used for modelling them. Then we detail the training and validation strategy (using spatial and temporal cross-validation on data pooled over the respective other dimension). We close on a short description of each contributing study and the specific cross-validation setting.

2.1 Predictor variables

Abundance. The relative abundances of species are widely hypothesized to influence the probability of interaction. Most simply, more abundant species are hypothesized to interact with one another more frequenlty than rare species. Abundances may or may not be sampled independently of the network itself ("external abundances"), and they are rarely reported. Such external abundances are more common for plants (e.g. vegetation transects) than pollinators (e.g. sweep-net transects). We used external abundances as predictor whenever available, and network-based abundances (i.e.

sum of observed interactions per species) otherwise. Species recorded in networks but missing in external abundances were set to an external abundance of 1 (i.e. the lowest possible value).

Interaction traits. Specific interaction relevant traits are hypothesized to influence the formation of interactions between species and their probabilities. For example, morphological interaction traits that match, such as species morphology (e.g. length of proboscis or corolla) or olfactory or visual cues (floral scent and colour, petal markings) are hypothesized to promote interaction formation. Nevertheless, it is challenging to compute a match between traits, e.g. what shape a bee must have to be able to access the nectar of a flower. As a consequence, traits reported are typically those that are easy to measure (and match morphology), not necessarily those most important for an interaction (floral bouquet, scent sensitivities, colour perception).

Latent traits. In addition to the traits provided by the authors of the respective studies, we tried to construct latent traits based on the phylogeny or taxonomy of the species recorded. The idea is that the position of a species in the phylogenetic tree may imply the value of a trait relevant for the interaction. Thus, for each study we constructed the phylogenetic relationship of the plants based on U.PhyloMaker (Jin 2023) and the instructions on its github pages.[1] For pollinators, no such reliable source is available across the invertebrate orders (as the data also include arachnids and vertebrates), and hence we here construct a tree based on taxonomic relationships. We resolved synonyms, and referenced phylogeny/taxonomy, using taxize (Chamberlain *et al.* 2020). We plotted the resulting trees (using phytools: Revell 2024) and checked for correctness. Then we decomposed them into phylogenetic eigenvectors (using MPSEM: Guenard & Legendre 2022), of which there are one fewer than species in the tree.

The data to analyse thus consist of the observed number of interactions between plant A and pollinator B, and the abundances, traits and phylogenetic eigenvectors of plants and pollinators, respectively. Convenience functions in tapnet (`make_tapnet` and `tapnet2df`) ensures that all traits, abundances and network names are matched correctly before converting everything into a long-format data set for analysis with machine- and deep-learning tools.

2.2 Method to predict interaction frequency

We investigated three types of predictive approaches: baseline predictions based only on network size and abundances (base 1 and 2, respectively); predictions based on an explicitly defined trait-matching approach (Benadi *et al.* 2022, "tapnet"); and predictions based on three different machine/deep-learning approaches previously used for similar tasks (Pichler *et al.* 2019; Terry & Lewis 2020).

The simplest baseline (base 1) gives every interaction the same probability, i.e. $N \cdot P(\hat{O}_{ij}) = N/(nm)$,[2] where $\widehat{\mathbf{O}}$ (size $n \times m$) is the predicted network, with N interactions observed between n plant and m pollinator species. Baseline 2 predicts interactions based on abundances, i.e. as cross-product of the abundance vector a_n for plants and a_m for pollinators: $N^2 \cdot P(\hat{O}_{ij}) = a_n^\intercal a_m$.[3] Note that base 2 contains information about the network whenever no external abundances are available, because it uses marginal totals (see also *Discussion*).

The "tapnet" approach (Benadi *et al.* 2022) uses abundances in the same way as baseline 2, but adds a probability based on trait matching of pre-specified trait pairs (e.g. length of proboscis and corolla), as well as a trait match of latent traits constructed from a linear combination of (a few) phylogenetic eigenvectors. The optimisation of this latent trait-pair increases exponentially with

---

[1] `https://github.com/jinyizju/U.PhyloMaker`

[2] This means, the N interactions are spread evenly across all nm cells of the matrix. Since P is a probability, it has to be multiplied with N to give us the expected interactions.

[3] Since a_n and a_m are divided by N to give a fraction, we need to multiply P with N^2 to give us expectations.

the number of eigenvectors considered and thus had to be confined to the first 5. Also, tapnet is limited in the number of networks used for training and thus cannot use all training networks in some cases (see case study descriptions). The model is available as R-package tapnet, version 0.6 (Dormann & Benadi 2023).

The machine- and deep-learning approaches comprise random forest (see also Pichler *et al.* 2019; Ornai & Keasar 2020; Sydenham *et al.* 2022; Benadi *et al.* 2022), extreme gradient boosting, and a neural network, all implemented as R packages. Specifically, we used: (1) ranger, version 0.16.0 (Wright & Ziegler 2017), with default settings. (2) Extreme gradient boosting, using xgboost package, version 1.7.7.1 (Chen *et al.* 2024), with hyper-parameter training of nrounds (200, 1200, 2000), max_depth (1, 2, 3, 5), eta (0.01, 0.001, 0.0001) and gamma (0, 1, 10), using a Poisson-likelihood and defaults otherwise. We tuned xgboost on the first training subset of each study and used the resulting settings for all other training runs. (3) A neural network (a fully connected multi-layer perceptron: MLP) assuming a Poisson likelihood was fitted using cito, version 1.1 (Amesoeder *et al.* 2023), which internally calls the torch framework; hyper-parameter tuning included the number of nodes in the first of two hidden layers (20-150 nodes in steps of 10), a regularisation with lambda betwen 0.00001 and 0.01, an elastic net tuning alpha between 0 and 1, a batch size between 500 and 3000; and a learning rate on a log scale betwen 10^{-6} and 10^{-3}, with 200 epochs and a selu activation function. Within these limits, 150 tuning combinations were randomly selected and the resulting best hyper-parameter was used for fitting the MLP. As for extreme gradient boosting, the MLP was tuned on the first training subset per study. We do not present the MLP in the results, because it failed to provide reasonable predictions for many case studies, largely due to an insufficient sample size.

Other approaches were excluded mainly because they predicted links but not interactions (e.g. Eklöf *et al.* 2013; Elmasri *et al.* 2020; Klomberg *et al.* 2022; Young *et al.* 2021); these approaches are reviewed in Terry & Lewis (2020). Brousseau *et al.* (2018) used a GAM, which makes it necessary to pre-select phylogenetic eigenvectors to avoid unidentifiability of the model. Also, it is less flexible than the machine-learning approaches used and performed worse in an earlier study (Benadi *et al.* 2022). The approach of Crea *et al.* (2015) was not included because it predicts interactions only from the perspective of one group, not both.

2.3 Training and validation

Each study required different training and validation decisions for the cross-validation. We explain here the general strategy behind these decisions, and in the case study description below we give the details for each case study.

We fitted the models and made predictions separately for each study. We split the available data into training and test networks. When both temporal and spatially replicated networks were available, we pooled them across the dimension not under consideration using bipartite::frame2webs (Dormann *et al.* 2008). That is, we combined networks across time for spatial prediction, and we combined networks across sites for temporal prediction. We acknowledge that this dilutes possible signals at the network level, but otherwise we would be left with too small and incomplete networks for the training phase (see Table 1 for resulting average network dimensions).

Depending on the number of networks available, we used a subset for training, from which we made predictions for all other networks. For example, for 8 networks in total, we would use 4 for training, predicting the other 4. Then, we swapped the training/validation data, repeated the procedure to finally obtained 8 validation points. When the number of networks was much larger, instead of this two-fold cross-validation we used a three-fold validation, fitting one third of the networks and predicting the remaining two thirds (see Table 1).

We evaluated the quality of the prediction by calculating the correlation between predicted and observed interaction frequency (using Spearman's ρ) and the normalised Nash-Sutcliffe efficiency (NNSE), because other measures such as root mean squared error (RMSE), median absolute difference (MAD: Gauss 1816) and (negative binomial) log-likelihood depend on the number of data points and their absolute values, making results less comparable among different case studies. The normalised Nash-Sutcliffe efficiency compares the squared error of a prediction to that of the squared error of a reference, in our case the mean of observed frequencies. An NNSE > 0.5 indicates that the model has skill, i.e. it predicts better than the mean of the data.

2.4 Temporal and compositional distance

For each case study we computed the distances in time between the mean of training networks and each validation network. Thus, if a study comprised 4 years and the first two were used for training, then the distance of networks of years 3 and 4 would be 1.5 and 2.5 years, respectively. We would need to calculate spatial distance differently, as case studies included two types of designs: altitudinal gradients, where elevational distance is more relevant than geographical distance, and multi-site studies, where geographic distance could be appropriately computed. Each type of spatial study contributed too few studies to allow for a meta-analysis of either spatial or elevational distance.

Furthermore, we computed the compositional distance between all networks as Bray-Curtis distance of their abundances, separately for plants and pollinators. We averaged abundances for the training data and computed the Bray-Curtis distance between each network and mean training abundance. We used these distances to assess whether plants and/or pollinators determine prediction quality.

2.5 Statistical analysis of study-level results

Since we had the original cross-validation results from all case studies, we were able to run an "individual-participant level meta-analysis" (as it is referred to in medicine: Riley *et al.* 2010; and known in psychology more vaguely as "integrative data analysis": Curran & Hussong 2009). Thus, we did not aggregate the data per case study to an effect size, as in a conventional meta-analysis, but kept the individual validation results and analysed them as raw data. For this analysis we used a mixed-effects model with study ID as random effect, and distance in time or composition as univariate predictor (representing our respective hypothesis), with a third-order polynomial effect to account for non-linear effects.

2.6 Analysis of case studies combined

If observed or phylogenetic latent traits are relevant for plant–pollinator interactions (our hypothesis ii), we should be able to train a model also across all studies on the species-level information. To do so, we combined all data from all studies into a single data set for a large general analysis (akin to Pichler *et al.* 2019). Since several studies did not provide external abundances, and no trait was reported for either plants or pollinators for more than 20% of the species, we could only use phylogenetic eigenvectors as predictors for this general analysis.

Specifically, we combined all plant taxa (798) into a single large phylogeny in the same way as described above for the case studies. The pollinator taxa (2469) were combined into a large taxonomy, which was converted into a pseudo-phylogeny, as described above. Each pseudo-phylogeny was then used to compute $N - 1$ orthogonal phylogenetic eigenvectors, which were combined with the 89,968 network links representing 194,125 observed pairwise interactions in 513 networks.

Since every study used a different sampling intensity, we standardised the frequency of observations per study to yield the same interaction density of 10^{-5} interactions/link (which is the median interaction density observed in all studies; we tried the maximum density of 0.1, which yielded slightly poorer cross-validation errors). To do so, for each study, we divided the observed interactions by the total number of interactions (making them sum to 1), then multiplied it by the number of observed links (to give each the average study weight) and by 10^{-5}. In this way, each data set had the same expected interaction density, i.e. the same weight per observed link.

For the analysis, we did a 14-fold block cross-validation, omitting each study in turn from the training data and using them as test data. We computed Spearman's ρ and normalised Nash-Sutcliffe efficiency for comparability with the per-study analyses. As modelling approach we again used random forest, but with a larger "minimal bucket size" of 10 (default: 1) and a maximal tree depth of 20 (default: unlimited) to reduce overfitting and computation time.

2.7 Case studies

We analysed 14 pollination network case studies (detailed below and summarised in Table 1), which differed widely in the number of interactions, number of networks and number of species. For example, networks analysed had between 20 and 4070 observed interactions, between 26 and 178 plants, and between 14 and 929 pollinators. While most studies contributed fewer than 10 networks, some featured well over 50 up to even 117 networks. As a result, the data basis on which we fitted the models varied hugely among case studies.

For prediction, typically externally observed abundances were available only for plants (only one study also estimated pollinator abundances independently of the networks; several studies did not provide independent plant abundances). For prediction, we substituted those missing external abundances by marginal sums of the observed networks, inevitably leaking some information from the test network to the prediction model. Our results must therefore be regarded as somewhat optimistically biased.

Some studies were spatio-temporal, so that they contributed to both types of analyses.

2.7.1 Case study Bartomeus et al. (2008)

Bartomeus *et al.* (2008) studied the change in the structure of plant-pollinator networks under the influence of two invasive plant species in the Mediterranean coastal scrublands of the Cap de Creus Natural Park (Catalonia, Spain). A total of 1227 interaction between 32 plant and 119 pollinator species were recorded in six networks. Two 50 m × 50 m plots (with and without invasive plant species) were established at six sites in the early stages of invasion so that these plots had similar plant species diversity. External abundance for plant species was calculated using the point intercept method in four parallel transects of 50 m at each site. There is, however, no external abundance for pollinators, so we use marginal totals of the interaction networks. Plant species traits were collected by the authors from the literature and included colour, morphology, shape and inflorescence type.

We used two types of cross-validation in space. First, we combined the paired plots of each site, producing six networks with the following average properties: $n = 11.5$, $m = 42.17$, and $N = 204.5$. We then selected the three northwest plots for training and each of the remaining

Table 1: Total number (and mean per network) of interactions, pollinators and plants, number of networks used in analysis, and number of cross-validations they resulted in. Last column indicates which external abundances were reported and hence used in analysis. Note that mean network properties differ between spatial and temporal aggregation of the underlying data.

Case study	# Interactions	# Plants	# Pollinators	webs	CVs	ext. abund.
temporal:						
Benadi *et al.* (2014)	10144 (2100.8)	119 (46.8)	409 (153.4)	5	5	plants
CaraDonna *et al.* (2017)	28473 (677.9)	45 (9.2)	89 (20.0)	42	42	plants
Chacoff *et al.* (2018)	7501 (1173.8)	59 (36.2)	196 (82.7)	6	12	plants
Kaiser-Bunbury *et al.* (2017)	12235 (1529.4)	38 (18.8)	144 (56.6)	8	8	plants
Minachilis *et al.* (2023)	2779 (926.3)	151 (98.0)	335 (186.3)	3	6	plants
Olito & Fox (2015)	914 (101.6)	43 (13.9)	118 (26.8)	9	9	plants
Resasco *et al.* (2021)	3386 (41.8)	39 (7.0)	245 (14.1)	81	324	plants
Roswell *et al.* (2019)	20344 (4068.8)	111 (56)	161 (91.4)	5	6	–
Winfree *et al.* (2007)	474 (118.5)	45 (15.75)	75 (30.75)	4	4	plants
spatial:						
Bartomeus *et al.* (2008)	1227 (204.5)	32 (11.5)	119 (42.2)	6	6	plants
Benadi *et al.* (2014)	10144 (1690.7)	119 (45.3)	408 (148.3)	6	6	plants
Chacoff *et al.* (2018)	7501 (182.8)	59 (14.5)	196 (33.8)	4	4	plants
Kaiser-Bunbury *et al.* (2017)	12235 (1529.4)	38 (16.0)	144 (55.4)	8	8	plants
Minachilis *et al.* (2023)	2779 (277.9)	151 (33.7)	335 (74.5)	10	10	plants
Rakosy *et al.* (2022)	1177 (235.4)	33 (11.8)	62 (27)	5	6	plants
Roswell *et al.* (2019)	20344 (2543)	111 (32.6)	161 (69.6)	8	15	–
Tinoco *et al.* (2017)	7672 (2557.3)	32 (19.3)	14 (9.7)	3	6	–
Weiner *et al.* (2014)	97318 (832)	178 (10.6)	929 (48.5)	117	234	plants
Winfree *et al.* (2007)	474 (148.3)	45 (23)	75 (39.7)	3	6	plants
Winfree *et al.* (2014)	484 (37.2)	26 (4.8)	57 (14.3)	13	65	plants, pollinators

three for testing. Then we repeat this procedure, swapping the training and testing plots. We then end up with six validations of different spatial distances to the mean of the training data.

2.7.2 Case study Benadi et al. (2014)

Benadi *et al.* (2014) report on the phenology and interactions of plant and pollinator communities along an altitudinal gradient in the National Park Berchtesgaden, located in the German Alps. Six sites on grasslands at altitudes between 950 m and 2020 m a.s.l. were sampled from May to September 2010, whenever possible once per week. On each sampling date, flower abundances (flower counts) were sampled on each site. In total, 10144 interactions (1716 binary links) were recorded between 119 plant and 408 pollinator species, aggregated into six and five networks for spatial and temporal analysis, respectively. Diameter of flowers for all plant species was measured. Thus, the training information included external abundance of plants but marginal totals of animals as abundances, one morphological trait for plants (flower size), and all the phylogenetic/taxonomic eigenvectors for both plants and animals as predictors. Predictions made to pollinator abundances are based on the observed validation networks. We acknowledge that this may yield optimistic prediction estimates.

To analyse the correlation between networks over time, we divided the interactions for each month. This resulted in a total of five matrices. We used the first two of the chronologically ordered networks to train the models, and then tested the predictions with the last three networks. For the spatial analysis, we aggregated the interactions into six matrices representing the altitudinal gradient. The training set consisted of the networks from the three lowest altitudes, while the models' predictions were tested with the networks from the three highest altitudes. For both analyses, the networks in the training and test sets were swapped, and the process was repeated to ensure a comprehensive evaluation of model performance. As a result, we performed five cross-validations for the temporal analysis and six for the spatial analysis.

2.7.3 Case study CaraDonna et al. (2017)

CaraDonna *et al.* (2017) investigated the within-season temporal turnover of plant-pollinator interactions in a subalpine ecosystem in the Western Colorado Rocky Mountains (USA) near the Rocky Mountain Biological Laboratory. The study area can be described as a mosaic of wet and dry meadows intermixed with aspen and conifer forest; it is snow covered for much of the year, with a short summer growing season of 3–5 months (May–September). Interactions between plants and pollinators were recorded in dry meadows at weekly intervals for the majority of three summers. In total, 28959 interactions were recorded between 45 plant species and 89 pollinator taxa, resulting in 42 weekly networks (n = 12 in 2013, 15 in 2014, 15 in 2015). External abundance data was recorded for flowers at weekly intervals. Plant and pollinator trait data (functional nectar depth and pollinator proboscis length) were measured (or estimated) for most species.

For temporal cross-validation, we fitted the models to the 21 first networks and predicted to the 21 other, and than switched training and testing, yielding 42 cross-validations.

2.7.4 Case study Chacoff et al. (2018)

Chacoff *et al.* (2018) report plant-pollinator networks from four xeric shrubland sites, ca. 50 km north of Mendoza city (Argentina), two sites sampled weekly during the flowering seasons of six consecutive years and the other two sites sampled only the first year. In total, 7501 interactions were recorded, for 59 plants and 196 pollinators in total, across 14 networks. External abundance was recorded for flowers (Vázquez *et al.* 2022), and a range of flower and animal traits are available for most species (nine plant traits describing dimensions of the flower, nine pollinator traits related to head and body dimensions: Lomascolo *et al.* 2022). Missing traits were imputed using traits and phylogenetic eigenvectors (but not abundances or species names) based on 10 randomForest imputations (using R's mice-package: van Buuren & Groothuis-Oudshoorn 2011). Thus, the training used information of external abundances of plants but marginal totals of animals as abundances, all traits and the phylogenetic/taxonomic eigenvectors as predictors. Accordingly, predictions were made to pollinator abundances as based on the observed validation networks. We acknowledge that this procedure may yield optimistic prediction estimates.

We used two-fold cross-validation in space for the first year of the study, by fitting the models to two adjacent sites for training and another two for training. This procedure yielded four validations of different spatial distance to the mean of the training data.

For three-fold temporal cross-validation (only sites 1 and 4) we used two adjacent years (i.e. {1,2}, {3,4} or {5,6}) for training, and the other four years for validation. This procedure yielded twelve validations of different temporal distance to the mean training data.

2.7.5 Case study Kaiser-Bunbury et al. (2017)

Kaiser-Bunbury *et al.* (2017) report on plant-pollinator networks from eight sites on Mahé, Seychelles, assessed eight times at roughly monthly intervals, covering an entire flowering season. The data from the exotic plant removal treatment was included in this analysis but ignored. In total, 12235 interactions were recorded, for 38 plants and 144 pollinators, across 64 networks. External flower abundance data were available, but no traits. Thus, training used external abundances for flowers and marginal totals of the interaction network for pollinators as abundances, and phylogenetic/taxonomic eigenvectors as predictors. Accordingly, predictions were made to abundances as based on the observed validation networks and their external flower abundances.

For spatial cross-validation, we first summed all networks within a site across the 8 sampling times, yielding a total of 8 networks. Of those, we used the north-western four site for training and each of the south-eastern for testing; then we swapped training and testing. This yielded eight validations of different spatial distance to the mean of the training data.

The same strategy was used for temporal cross-validation: all networks were summed across space, yielding one for each of the eight time periods. The first four were used to training, the other for validation, then vice versa. Again, this yielded eight validations of different temporal distance to the mean training data.

2.7.6 Case study Minachilis et al. (2023)

Minachilis *et al.* (2023) report on pollination networks sampled in Mt. Olympus, Greece. Ten sites were sampled covering the major vegetation zones of Mt Olympus (scrub, forests and alpine meadows). Species interactions were recorded by hand netting insects visiting flowering plants in 2013, 2014, 2016. Some sites were sampled less often, when harsh weather conditions made them inaccessible. In total, 2779 interactions (1281 binary links) were recorded between 151 plant and 335 pollinator species, across 30 networks. Plant abundance was estimated as the average number of flower or inflorescences per m^2 per site (average of flower counts in twenty five quadrats of 1 m^2 per site). The training information included external abundance of plants but marginal totals of animals as abundances and all the phylogenetic/taxonomic eigenvectors for both plants and animals as predictors. No morphological trait data was available.

For temporal cross-validation, data from all sites were pooled to yield three networks, each representing a year. We fitted the models to each year separately, predicting to the other two in turn, yielding six predictions.

For spatial cross-validation, we pooled the three years' of data for each of the 10 sites and trained the model on 5 sites, predicting to each of the other. This yielded 10 predictions for validation.

2.7.7 Case study Olito & Fox (2015)

Olito & Fox (2015) studied how species abundance, phenology, and morphology predict both network structural metrics and specific pairwise interactions in plant-pollinator networks. The study assembled plant–pollinator networks in a contiguous low-alpine meadow on the east face of Mt Murray in the Canadian Rockies in Kananaskis Country, Alberta, during summer 2010. The interactions, when an insect visitor was observed contacting floral reproductive structures, were recorded in a square 1-ha plot located at 2350–2410 m elevation on every day that weather conditions were suitable for pollinator flight, from the day of first flowering, until killing frosts occurred and pollinators were no longer observed (24 June 2010–26 August 2010, a total of 32 sampling days). The authors documented 914 interactions between 42 flowering plant species and 118 insect species, across 9 networks. External plant abundance was estimated using floral density measured per square meter in five 1 m^2-quadrats randomly placed in transects, while pollinator species abundances were not independently estimated. Functional trait information for plants and pollinators was not included in the study, but phylogenetic information for plants and taxonomic data for pollinators were integrated as eigenvectors in the analysis.

For the temporal cross-validation, with nine networks representing one per week, we fitted the models using the first four networks and predicted to each of the others five networks, yielding 9 predictions. The networks in the training and test sets were then swapped, and the process was repeated to ensure comprehensive evaluation of model performance.

2.7.8 Case study Rakosy et al. (2022)

Rakosy *et al.* (2022) studied how anthropogenic changes impact the diversity, composition, and structure of plant-pollinator networks in semi-natural grasslands in the Opawskie Mountains, located along the Poland-Czech Republic border. Their study was conducted at five grassland sites within a matrix landscape characterized by mesic, acidic soils and similar altitudes. They established 10 transects per site (with one pasture limited to 6 transects), each measuring 30 × 2 m. Transects were placed with a minimum distance of 30 m between them, and towards the nearest field margins. Over the course of the study, 33 plant species were recorded interacting 1,177 times with 62 pollinator species, across 5 networks. In each transect, they visually estimated the percent cover of flowers/inflorescences of each plant species. Then, external plant abundance was provided, but no information about pollinator abundance and functional traits of both trophic levels was included.

With five networks representing each site, the spatial analysis used two networks to train the models and evaluated the predictive performance using the remaining three networks. By swapping the networks in the training and test sets and repeating the process, the analysis conducted a total of six cross-validations for each model.

2.7.9 Case study Resasco et al. (2021)

Resasco *et al.* (2021) recorded plant-pollinator interactions over five years (2015–2019) in a subalpine meadow of the Colorado Rocky Mountains (40°01’48”N, 105°32’26”W). Observations spanned approximately 16 to 18 weeks each year, beginning after snowmelt in late May to early June and continuing until late September. The study site comprised six plots at a similar elevation (2962–2978 m), five with a size of 30×2 m^2 and one with 20×2 m^2. A total of 3386 interactions were recorded among 39 plant species and 245 pollinators across 6 spatial and 81 temporal networks, respectively.

Plant-centred sampling was carried out in the mornings between 08:00 and 12:00 in good weather, with a 15-minute survey of all flowers for visitors. In addition, the number of flower units of each plant species (i.e. external abundance) was measured weekly in the six plots, while the external abundance of pollinators was not measured. Two weekly networks were removed while preparing the data (week 17 in 2016 and week 17 in 2019), as they were too small (only one plant or pollinator in a network). Plant and pollinator phenological traits (mean start (day of the year) of flowering/flying activity, mean end (day of the year) of flowering/flying activity and mean duration (in days) of flowering/flying activity were included. Lastly, phylogenetic data for plants and taxonomic classifications for pollinators were included as eigenvectors in the analysis.

For spatial analysis, we pooled the data into six interaction networks, one per site. These networks were then grouped into three pairs. For each pair, we trained the model on the two networks and made predictions on the remaining four, resulting in a total of 9 cross-validations. For the temporal analysis we applied five-fold temporal cross-validation, fitting the models to all 16-18 networks of a given year and predicting to all networks in the other years, yielding a total of 324 cross-validations.

2.7.10 Case study Roswell et al. (2019)

Roswell *et al.* (2019) collected the data for a study that sought to distinguish between the floral use and preference of male and female bees of the same species. This study took place in eight semi-natural meadows (sites) in New Jersey, where most of the flower species are native of the eastern United States. A total of 20344 interactions were recorded between 111 plant and 161 pollinator species across 33 networks.

Each site was sampled on 5 separate equally spaced sampling rounds between June 6 and Aug 20, 2016. Each sampling round consisted of three consecutive days of sampling when the weather was sunny and without precipitation. On each day, bees were netted during a minimum of six (but often more) 30-minute periods with short breaks in between, beginning in the morning and continuing until pollinator activity slowed sometime in the afternoon, but honey bees were not collected. No external observations of plant and bee abundances were made, but data on intertegular distance of pollinators was provided. In addition, phylogenetic and taxonomic trees are added to the models as eigenvectors.

For the temporal analysis, we aggregated the interactions in five interaction networks, each corresponding to a distinct sampling round. The first three rounds were used to train the models, while the last two rounds were used to test predictive performance. For the spatial analysis, the training set consisted of five networks from the southern sites, while the test set included three networks from the northern sites. In both procedures, the process was repeated by swapping the training and testing sets, resulting in six cross-validation runs for the temporal analysis and 15 cross-validation runs for the spatial analysis.

2.7.11 Case study Tinoco et al. (2017)

In their 2017 study, Tinoco *et al.* (2017) investigated the impact of resource availability and hummingbird morphology on hummingbird specialization. To this end, the researchers conducted measurements of resource availability, species traits, and hummingbird specialization over two years in three habitat types (forest, shrub vegetation, and cattle ranch) in the western Andes of Azuay (Ecuador). They conducted five surveys, with the first four occurring in February and August and the fifth in December. The available data are pooled over these five surveys, yielding three networks, recording 7672 interactions among 32 flower species and 14 hummingbirds. In each habitat type, a 2.2-km transect was established, with sampling points every 200 m. Hummingbirds were captured to measure the body mass and total bill length of males. Furthermore, three plots of 200 m × 5 m were established in each habitat type to measure nectar production and nectar sugar content as plant traits. The visits of the hummingbirds to the plants were recorded in these same plots, and a network was assembled by habitat type.

For spatial cross-validation, we fitted the models to the network of each of the three sites separately, predicting to the other two in turn, yielding a total of 6 predictions.

2.7.12 Case study Weiner et al. (2014)

Weiner *et al.* (2014) report on 119 plant-pollinator networks from three regions in Germany. We excluded two networks from one of the regions (AEG 8 and 13) because they were heavily infested by rape pollen beetle *Brassicogethes aeneus* from an adjacent oil seed rape field, with over 80,000 individuals in a single network. In total, N =61,902 (plus 416,558 from the two excluded networks) interactions were recorded, for $n = 178$ plants and $m = 929$ pollinators, across 117 networks. External abundance were recorded for flowers only, and no traits were available. Thus, the training used information of external abundances of plants but marginal totals of animals as abundances, all traits and the phylogenetic/taxonomic eigenvectors as predictors. Accordingly, predictions were made to pollinator abundances as based on the observed validation networks. We acknowledge that this may yield optimistic prediction estimates.

We used three-fold cross-validation in space, by employing approximately 40 networks of each region for training and each network of the other two regions for training. Each region was once used for training, and hence each region was also used twice for validation, yielding 234 validation points of different spatial distance to the mean of the training data. Tapnet could only be trained on a random subset of four networks per site, due to optimisation constraints. Its predictive ability will thus likely be underestimated.

2.7.13 Case study of Winfree et al. (2007)

Winfree *et al.* (2007) investigated the effects of human land use on pollinators (bees; Hymenoptera: Apiformes) at both landscape and local scales in the Pinelands Biosphere Reserve, southern New Jersey (USA). Plant-pollinator interactions (excluding honeybees) were recorded across 40 study sites, though ten were excluded from our analysis due to insufficient data. In total, 474 interactions were documented, involving 45 plant species and 75 pollinator species. These interactions were structured into three spatial and four temporal networks for analysis. All sites were within the same ecoregion, sharing similar elevation (0–63 m), geologic history, and soil type.

Each site was surveyed using a 110 m × 10 m transect, sampled four times during peak bee activity (April 14–August 14, 2003). Pollinators were hand-netted during two 30-minute sessions per sampling event. In addition, external floral abundance was assessed at each site, but external pollinator abundance was not recorded. On the other hand, pollinator intertegular distance is provided, but no information on functional traits is available in this study. Finally, phylogenetic information for plants and taxonomic data for pollinators were included as eigenvectors in the analysis.

For the temporal analysis, interactions were pooled into four matrices, each representing a different sampling round. The algorithm was trained using the matrices from the first two rounds and tested on the remaining two, and vice versa, resulting in a total of four cross-validations. For the spatial analysis, interactions from the 30 sites were grouped into three networks based on their spatial distribution. Each network was used to train the model while predictions were made on the remaining two, leading to a total of six cross-validations.

2.7.14 Case study of Winfree et al. (2014)

Winfree *et al.* (2014) studied the relationship between species linkage in pollinator networks and species persistence across a gradient of human disturbance in 13 sites surrounding native deciduous forests in New Jersey, USA. These sites span a steep gradient of land-use intensity, and data were collected on 424 interactions involving 26 plant species and 57 pollinator species, divided in 13 interaction networks. Pollinators were collected on the flower using hand-netting methods, with bees sampled for 60 minutes in each plot. At each site, data were collected in a 0.5 ha plot on four separate occasions, with sampling rounds spaced 9–13 days apart between April 10 and June 1, 2006. The sites, separated by at least 3 km, were sampled only under conditions suitable for pollinator activity: sunny or partly cloudy days, temperatures above 14°C, and wind speeds no greater than 2 m/s.

External plant abundance was assessed by counting flowers in 100 quadrats per plot, while plants with clustered small flowers were counted by clusters. External pollinator abundance was measured using pan-trapping, deployed for 2 hours during the first two sampling rounds and 4 hours in the final two rounds This dataset includes intertegular disctances as functional trait information for pollinator species but lacks such data for plants species. Additionally, eigenvectors are generated using phylogenetic information for plants and taxonomic information for pollinators

For the spatial analysis, interactions were pooled into 13 matrices, one per site. These matrices were then grouped into six sets based on spatial distribution—one group containing three matrices and the remaining five groups containing two matrices each. Each network group was used to train the model while predictions were made on the remaining networks, leading to a total of 65 cross-validations, allowing us to compare the predictions of each model with the empirical data across different spatial distances.

3 Results

3.1 Prediction methods

Across the methods compared for prediction, randomForest performed on average best (Fig. 1). For temporal validation, the best method, randomForest, achieved a correlation between prediction and observation of $\rho = 0.42$, and for spatial validation $\rho = 0.27$.[4] In space, all methods were very similar in predictive power to the abundance-only base 2, suggesting that little information

---

[4] We do not show results for the deep neural network, as its performance was extremely variable and on average poor in the first few case studies we analysed, probably due to the still relatively small sample sizes. Investing much more time to fine tune this method may yield much better fits, so we do not want to give the impression that it is an unsuitable method.

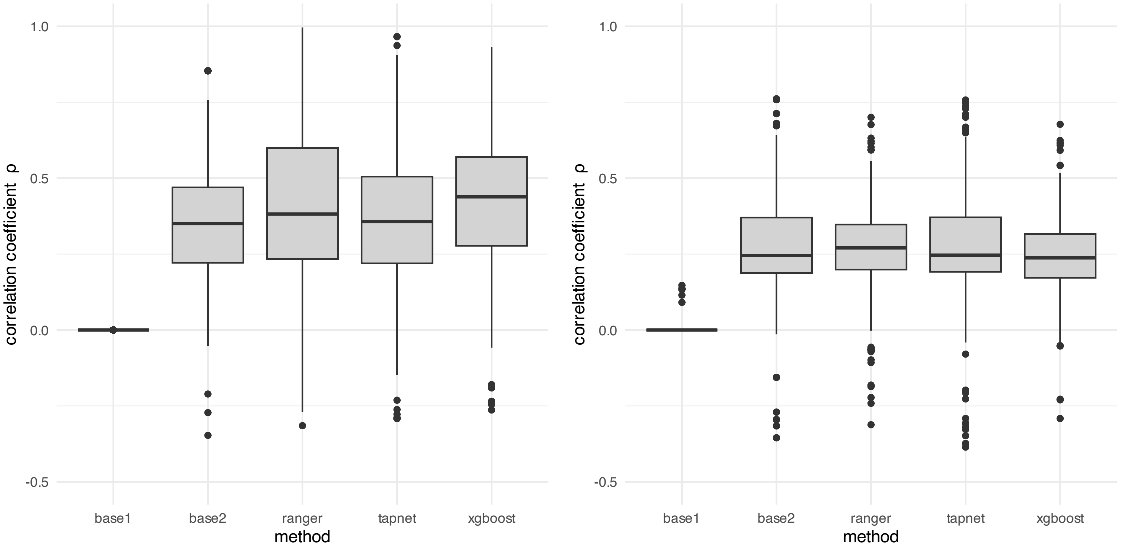

Figure 1: Predictive performance of three methods (ranger as implementation of randomForest, tapnet and XGboost) and two baseline models in temporal (*left*) and spatial (*right*) cross-validation. Correlation between prediction and observation measured as Spearman's ρ. (Note that the pure abundance-based baseline 2 has considerable skill compared to predicting a constant number of interactions (base 1).)

beyond what was coded for in the species abundances was successfully harnessed for prediction. In time, however, randomForest and XGboost slightly surpassed base 2 (and tapnet), indicating information beyond abundance enhances model predictive ability. In consequence, below we present results only for randomForest.

3.2 Predictions validated in time

Prediction quality of interaction frequency in time was weakly and non-linearly related to temporal distance between training and test data (quadratic time effect= -0.87 ± 0.31, $P = 0.0071$; Fig. 2 left). The initial increase from poor predictability of $\rho \approx 0.25$ within a season to a moderate $\rho \approx 0.4$ from one year to the other represents the "return of skill" (as it is called in atmospheric forecasting: Guo *et al.* 2012) for temporal predictions into the same season in the next or previous year. This effect of seasonal phenology decreased quickly with time, disappearing by year two.

Datasets also varied in their predictability, but much less so than the residual variance (variance estimates: 0.0079 vs 0.0565, respectively). The dataset contributing most validation points, Resasco *et al.* (2021), is also the one with the smallest networks (Table 1), explaining to some extent its large scatter.

In the temporal prediction setting, the majority of case studies exhibit a decrease of skill with increasing differences in community composition. However, distance in plant community composition had no detectable effect on skill (comparison with intercept-only model: $p = 0.1089$), being masked by the large residual variation (variance between data sets: 0.0081; residual variance: 0.066).

The effect of distance in pollinator composition was broadly similar to that of plant community composition. Here a significant trend could be detected ($P < 0.001$), showing a loss of skill only for substantially different communities (Bray-Curtis distance > 0.8; Fig. 3 right).

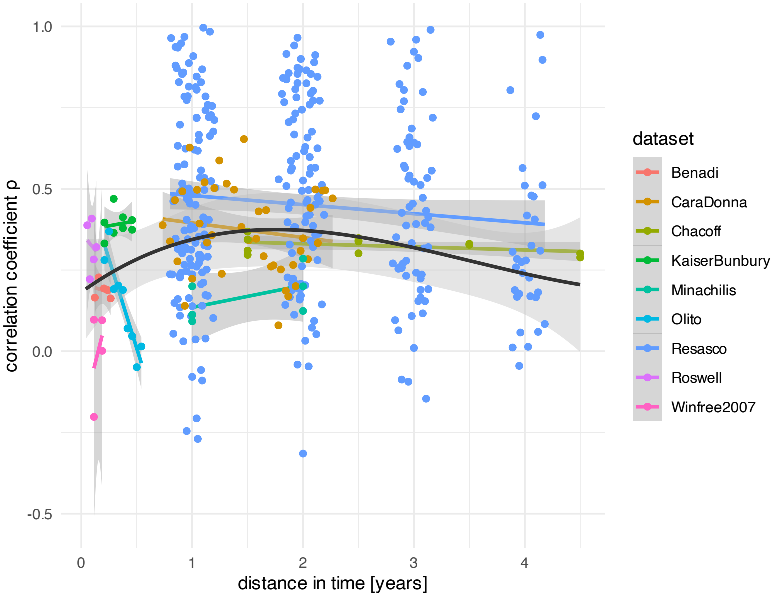

Figure 2: Effect of temporal distance between training and test data on prediction skill of randomForest models. Colours represent different data sets, black line is prediction from non-linear random-intercept model across all data sets. Ribbons around regression are 95%-confidence intervals. Note that a distance of, say, 1.5 years may arise from training on two years and predicting to a third year: the average distance between the training and test data is then 1.5 years.

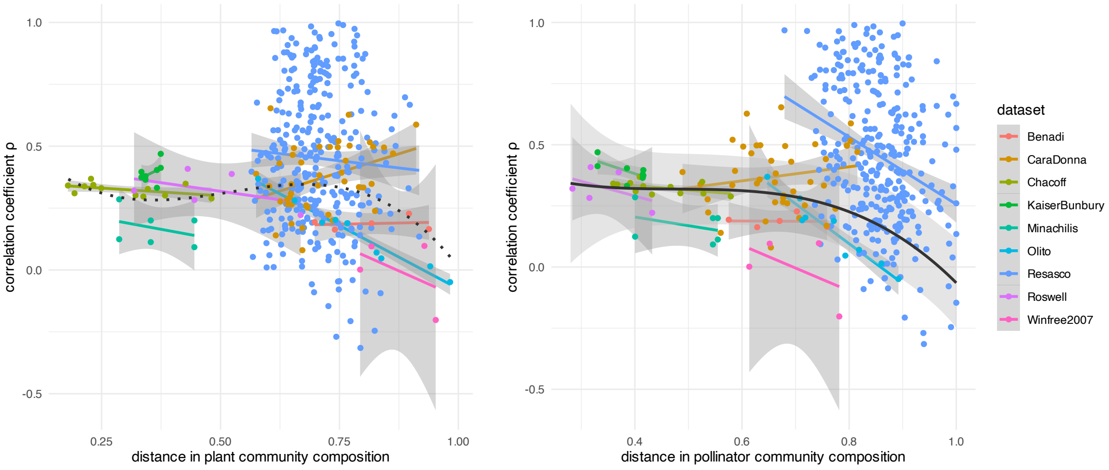

Figure 3: Effect of distance of community composition (*left*: plants; *right*: pollinators) between training and test data on temporal prediction skill of randomForest models. Colours represent different data sets, black line in pollinator plot is fit of non-linear random-intercept model across all data sets. Ribbons around regression are 95%-confidence intervals (for significant relationship only).

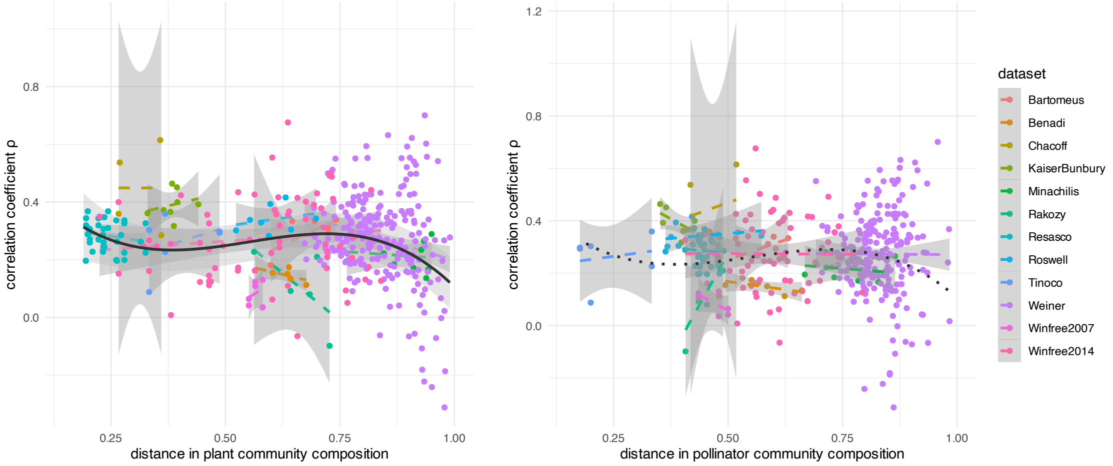

Figure 4: Effect of distance of community composition (*left*: plants; *right*: pollinators) between training and test data on spatial prediction skill of randomForest models. Colours represent different data sets, black solid/dotted line is fit of non-linear random-intercept model across all data sets. Ribbons around regression are 95%-confidence intervals (for significant trends only).

3.3 Predictions validated in space

As for temporal predictions, randomForest provided the best predictions (by a very slight margin: Fig. 1 right), and so we use those results only for testing effects of compositional distances on predictability.

We could not meaningfully evaluate the effect of spatial distance on predictability, as half of the studies were conducted along on elevational gradients, where horizontal distance is not very informative of differences in environmental conditions. The other half of studies provide too few data to reliably make any statement and hence we refrained from an analysis of the spatial distance effect.

For predictability in space we find the opposite pattern as for predictability in time with respect to the effect of plant and pollinator community composition Fig. 4). Prediction quality decreased weakly but significantly ($P < 0.001$) with plant community composition, once the Bray-Curtis distance is above approximately 0.75 (Fig. 4 left). In contrast, no significant effect could be discovered for the effect of pollinator community turnover (Fig. 4 right).

3.4 Analysis of all case studies combined

In 14-fold or leave-one-study-out cross-validation, interaction density of a link was very poorly predicted ($\bar{\rho} = 0.10 \pm 0.096$, not significantly different from 0). This means that the random forest trained on all but one study had virtually no skill. We thus refrained from analysing and interpreting variable importances.

4 Discussion

Predictions of interaction frequency in pollination networks remain a challenge. Across the 14 studies we used in spatial or temporal cross-validation, predictability was low to moderate. Consistently, it was the knowledge of the abundances in the target networks, collected either independently or computed as sum of observed interactions in the network itself, that did practically all of the prediction (base 2 in Fig. 1). In other words, the traits and phylogenetic eigenvectors used as additional predictors were of very limited use (in contrast for example to Vizentin-Bugoni *et al.* 2014). While pollinators are typically slightly less specialised than for example herbivores (Fontaine *et al.* 2009), it is still disquieting that we failed to detect clear and strong signals from traits and trait matching.

4.1 Prediction methods

This study is, to date, the largest attempt in predicting frequency of network interactions in ecology, yet we failed to achieve reliable predictions. Using cutting edge machine- and deep-learning approaches did not improve on a simple abundance-based prediction (base 2), and also the dedicated "tapnet" approach, which tries to match plant and pollinator traits explicitly, did not perform any better (or worse) than random forest or extreme gradient boosting. While we analysed hundreds of thousands of flower visitation events across the 14 studies, for any given study, sample size was typically "only" in the hundreds to thousands per network, for dozens to hundreds of plant and pollinator species (Table 1). This may be the reason for the unsatisfactory performance of deep neural networks, which were dropped from reporting here due to their erratic and, on average, poor performance (see supplementary material).

The consistently moderate performance of all modelling approaches, and their lack of improving on abundance-only predictions, suggests that the information provided by easy-to-measure traits and trait-substituting phylogenetic eigenvectors is not specific enough to describe the actual mechanisms of flower selection by pollinators. Scent and vision in particular are virtually inaccessible for these thousands of species, yet of demonstrated importance for flower-pollinator interactions (Junker *et al.* 2010; Renoult *et al.* 2015; Kantsa *et al.* 2018). We suggest that such trait information, rather than larger data sets (Lanuza *et al.* 2025) and fancier methods, may lead to improved predictions in the future.

4.2 Trait-moderated interactions?

It is a long-standing discussion, which role plant and pollinator traits play in pollination networks (Stang *et al.* 2007; Vázquez *et al.* 2009; Burkle *et al.* 2013; Maruyama *et al.* 2014; Vizentin-Bugoni *et al.* 2014; Olito & Fox 2015; CaraDonna *et al.* 2017; Weinstein & Graham 2017). The clear and obvious adaptations of some flowers to animal pollination (from corolla shape, style and pollen placement, to floral patterns and nectar chemistry: Willmer 2011) is blurred by observations of supposedly extremely specialised pollinators acting much more generalistically (lepidoptera with extremely long proboscis visiting open flowers; oil-collecting bees also visiting other flowers for pollen and nectar: Stefanescu & Traveset 2009; Schäffler & Dötterl 2011), and by non-specialised pollinators being able to access specialised flowers (by force: bumble bees; by small size: Meliponini; by hovering: chiffchaff on *Anagyris foetida*: Ortega-Olivencia *et al.* 2005).

Flower visitation does not imply pollination (e.g. Wilson & Thomson 1991; Mayfield *et al.* 2001). It may well be that, from a plant's perspective, a substantial amount of nectar and pollen production is wasted on ineffective pollinator, but the few effective pollinator visits justify this investment. The trait-matching pollination ecologists discuss may be highly adaptive in both plants and pollinators, yet not be readily detectable from patterns of flower visitation alone. In this case, flower visitation would primarily depend on availability, matching our findings that abundance (of flowers and pollinators) is of large importance (compare base 1 and 2), while trait and phylogenetic information does not substantially improve on abundance-only predictions (compare base 2 with all other model types).

4.3 Frequency vs missing links

In a seminal analysis of methods to identify missing links in ecological networks, Terry & Lewis (2020) report moderate to good success in predicting *links* within networks (AUC-values of around 0.7). That is a rather different task, as it takes a given network and tries to work out which links *within that network* are missing, based on network structure and sampling coverage. Our ambition was to predict interaction frequencies *in a new network*, of which external abundances of the participating species as well as their (observed and phylogenetic latent) traits are known. Arguably, this is a more difficult task, comparable to an extrapolation. It is thus not surprising to find our models' predictive performance to be inferior to theirs.

Since Terry & Lewis (2020) used quantitative networks (as did we), the abundances of all species can be roughly inferred from the data of each network. Their version of an abundance-only prediction (our base 2, their "sample size" model) performed similar to two of their approaches for pollination networks, but worse than the two best (connectance and matching-centrality). Neither of these two is applicable to our setting, regrettably, as they require information from the test network itself, which is what we want to predict. Note that this is a difference in their aim, no deficit of their study.

Another recent development is the attempt to pool many network observations across studies, and thereby generate a "meta-network" (Devoto *et al.* 2014). This approach overcomes the potential small-sample size limitations of each individual study. When we implemented this approach, it showed very little predictive skill for random forest or deep learning. A more fine-grained method may be able to discover more signal in the phylogenies, but it is unlikely to yield convincing and substantial predictive power. The obvious ecological problem is that such meta-networks lack information of phenological and indeed geographic co-occurrence. While trait-matching may transfer from one continent to another, the geographic separation of species must be included as forbidden links as well.

4.4 Prediction based on network-derived abundances

For our predictions we used externally recorded abundances whenever available (see Table 1). These were recorded along transects near to where network interactions were sampled. However, only one study (Winfree *et al.* 2014) also provided information on external abundances of pollinators, and several did not report flower abundances either. In these cases, we used the sum of observed interactions in the network as stand-in for abundances. However, if abundances are a consequence of network interactions, then this approach becomes circular, as we predict interactions based on abundances that are the consequence of interactions (Fort *et al.* 2016). We acknowledge this flaw in our logic, yet if this effect were strong, it should improve our ability to predict interactions, which clearly is poor even with network-based abundances.

The studies of Roswell *et al.* (2019) and Weiner *et al.* (2014) did not provide information on flower abundance, yet they do not exhibit improved predictions in Figs. 2 to 4. That is to say: for these studies, the potential circularity in the use of abundances does not seem to be a problem worth going into.

4.5 Community compositional effects on prediction

It may be tempting to interpret the difference between the non-significant effect of plant community composition on prediction skill and the significant effect of pollinator community composition (Fig. 3). However, we regard this as spurious, as the effect is exactly the other way around in the spatial setting (Fig. 4, and it is only non-significant for normalised Nash-Sutcliffe efficiency (see supplementary material, Fig. A3 and A4).

The dominant role of abundance as predictor of interactions is in line with simulations of Yahaya *et al.* (2024), who find specialisation to increase with floral resources. And it is only with high specialisation that we can hope to achieve predictive power beyond the effect of abundance. However, we regard the evidence from our synthesis as not particularly supportive of such a strong role of traits.

4.6 Interaction density, sampling intensity, and pooling of networks

Pollinator-flower interactions are recorded following different traditions and protocols. The resulting networks will thus also differ, as function of both methodology and sampling intensity (Novella-Fernandez *et al.* 2019; Brimacombe *et al.* 2023). Analyses across all data set may thus suffer from false-negative errors. That would explain the poor performance of our analysis across all data sets pooled, but it does not explain the low skill *within* each study (Figs 1-4).

Our methodological decision to pool networks either across sites or times to yield sufficiently dense networks may well have washed out differences in specialisation, phenological signals or spatial differences (as argued and shown in CaraDonna *et al.* 2021; Schwarz *et al.* 2020). For small networks, noise levels are very high and possibly unsuitable for the kind of analysis performed here. Future analyses will have to show whether data aggregation reduced prediction skill and thereby explain our failure to predict network interaction frequencies.

4.7 Outlook

There are several hypotheses our study did not investigate, which could form the basis of further investigations. For example, the majority of plants and pollinators are rare. So maybe it would be optimistic to expect high predictability for those species. In contrast, common pollinators may be competing more for floral resources so that for them the presence of a competitor may be a relevant predictor, not only the traits and abundances.

For now, however, we suggest to assume that network interaction frequencies are primarily driven by the abundance of its participants, with little generalisable signal of species traits, and instead a very high stochasticity in the recorded interactions.v Also, we hope to encourage more studies in the future to provide a similar data quality to allow extending and quantitatively summarising the emerging results

Acknowledgement

CFD, WJC and NB were funded by DFG FG 5207 Reassembly.

References

Amesoeder, C., Hartig, F. & Pichler, M. (2023) cito: An R package for training neural networks using torch. *arXiv preprint arXiv:230309599*. https://dx.doi.org/10.48550/arXiv.2303.09599.

Bartomeus, I., Vilà, M. & Santamaría, L. (2008) Contrasting effects of invasive plants in plant-pollinator networks. *Oecologia*, **155**, 761–770. https://dx.doi.org/10.1007/S00442-007-0946-1/TABLES/2.

Benadi, G., Dormann, C.F., Fründ, J., Stephan, R. & Vázquez, D.P. (2022) Quantitative prediction of interactions in bipartite networks based on traits, abundances, and phylogeny. *The American Naturalist*, **199**, 841–854. https://dx.doi.org/10.1086/714420.

Benadi, G., Hovestadt, T., Poethke, H.J. & Blüthgen, N. (2014) Specialization and phenological synchrony of plant-pollinator interactions along an altitudinal gradient. *Journal of Animal Ecology*, **83**, 639–650. https://dx.doi.org/10.1111/1365-2656.12158.

Berlow, E.L., Neutel, A.M., Cohen, J.E., de Ruiter, P.C., Ebenman, B., Emmerson, M., Fox, J.W., Jansen, V.A.A., Jones, J.I., Kokkoris, G.D., Logofet, D.O., McKane, A.J., Montoya, J.M. & Petchey, O. (2004) Interaction strengths in food webs: issues and opportunities. *Journal of Animal Ecology*, **73**, 585–598. https://dx.doi.org/10.1111/j.0021-8790.2004.00833.x.

Brimacombe, C., Bodner, K., Michalska-Smith, M., Poisot, T. & Fortin, M.J. (2023) Shortcomings of reusing species interaction networks created by different sets of researchers. *PLoS Biology*, **21**, e3002068. https://dx.doi.org/10.1371/journal.pbio.3002068.

Brousseau, P.M., Gravel, D. & Handa, I.T. (2018) Trait matching and phylogeny as predictors of predator-prey interactions involving ground beetles. *Functional Ecology*, **32**, 192–202. https://dx.doi.org/10.1111/1365-2435.12943.

Burkle, L.A., Marlin, J.C. & Knight, T.M. (2013) Plant-pollinator interactions over 120 years: loss of species, co-occurrence, and function. *Science*, **339**, 1611–1615.

CaraDonna, P.J., Burkle, L.A., Schwarz, B., Resasco, J., Knight, T.M., Benadi, G., Blüthgen, N., Dormann, C.F., Fang, Q., Fründ, J., Gauzens, B., Kaiser-Bunbury, C.N., Winfree, R. & Vázquez, D.P. (2021) Seeing through the static: the temporal dimension of plant–animal mutualistic interactions. *Ecology Letters*, **24**, 149–161. https://dx.doi.org/https://doi.org/10.1111/ele.13623.

CaraDonna, P.J., Petry, W.K., Brennan, R.M., Cunningham, J.L., Bronstein, J.L., Waser, N.M. & Sanders, N.J. (2017) Interaction rewiring and the rapid turnover of plant-pollinator networks. *Ecology Letters*, **20**, 385–394. https://dx.doi.org/10.1111/ele.12740.

Chacoff, N.P., Resasco, J. & Vázquez, D.P. (2018) Interaction frequency, network position, and the temporal persistence of interactions in a plant–pollinator network. *Ecology*, **99**, 21–28. https://dx.doi.org/10.1002/ecy.2063.

Chamberlain, S., Szöcs, E., Foster, Z., Arendsee, Z., Boettiger, C., Ram, K., Bartomeus, I., Baumgartner, J., O'Donnell, J., Oksanen, J., Tzovaras, B.G., Marchand, P., Tran, V., Salmon, M., Li, G. & Grenié, M. (2020) *taxize: Taxonomic information from around the web.* R package version 0.9.98.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., Li, Y. & Yuan, J. (2024) *xgboost: Extreme Gradient Boosting.* R package version 1.7.7.1.

Crea, C., Ali, R.A. & Rader, R. (2015) A new model for ecological networks using species level traits. *Methods in Ecology and Evolution*, **7**, 232–241. https://dx.doi.org/10.1111/2041-210X.12471.

Curran, P.J. & Hussong, A.M. (2009) Integrative data analysis: The simultaneous analysis of multiple data sets. *Psychological Methods*, **14**, 81–100. https://dx.doi.org/10.1037/a0015914.

Devoto, M., Bailey, S. & Memmott, J. (2014) Ecological meta-networks integrate spatial and temporal dynamics of plant-bumble bee interactions. *Oikos*, **123**, 714–720. https://dx.doi.org/10.1111/j.1600-0706.2013.01251.x.

Doak, D.F., Estes, J.A., Halpern, B.S., Jacob, U., Lindberg, D.R., Lovvorn, J., Monson, D.H., Tinker, M.T., Williams, T.M., Wootton, J.T., Carroll, I., Emmerson, M., Micheli, F. & Novak, M. (2008) Understanding and predicting ecological dynamics: are major surprises inevitable? *Ecology*, **89**, 952–961. https://dx.doi.org/10.1890/07-0965.1.

Dormann, C.F. & Benadi, G. (2023) *tapnet: Trait matching and abundance for predicting bipartite networks.* R package version 0.6.

Dormann, C.F., Fründ, J. & Gruber, B. (2008) Introducing the bipartite package: Analysing ecological networks. *R News*, **8**, 8–11.

Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N.P., Dalsgaard, B., de Sassi, C., Galetti, M., Guimarães, P.R., Lomáscolo, S.B., Martín González, A.M., Pizo, M.A., Rader, R., Rodrigo, A., Tylianakis, J.M., Vázquez, D.P. & Allesina, S. (2013) The dimensionality of ecological networks. *Ecology Letters*, **16**, 577–583. https://dx.doi.org/10.1111/ele.12081.

Elmasri, M., Farrell, M.J., Davies, T.J. & Stephens, D.A. (2020) A hierarchical Bayesian model for predicting ecological interactions using scaled evolutionary relationships. *The Annals of Applied Statistics*, **14**, 221–240. https://dx.doi.org/10.1214/19-AOAS1296.

Fontaine, C., Thébault, E. & Dajoz, I. (2009) Are insect pollinators more generalist than insect herbivores? *Proceedings of the Royal Society B: Biological Sciences*, **276**, 3027–3033. https://dx.doi.org/10.1098/rspb.2009.0635.

Fort, H., Vázquez, D.P. & Lan, B.L. (2016) Abundance and generalization in mutualistic networks: solving the chicken-and-egg dilemma. *Ecology Letters*, **19**, 4–11. https://dx.doi.org/10.1111/ele.12535.

Gauss, C.F. (1816) Bestimmung der Genauigkeit der Beobachtungen. *Zeitschrift für Astronomie und verwandte Wissenschaften*, **1**, 187–197.

González-Varo, J.P. & Traveset, A. (2016) The labile limits of forbidden interactions. *Trends in Ecology & Evolution*, **31**, 700–710. https://dx.doi.org/10.1016/j.tree.2016.06.009.

Guenard, G. & Legendre, P. (2022) *MPSEM: Modeling Phylogenetic Signals using Eigenvector Maps*. R package version 0.4-1.

Guo, Z., Dirmeyer, P.A., DelSole, T. & Koster, R.D. (2012) Rebound in atmospheric predictability and the role of the land surface. *Journal of Climate*, **25**, 4744–4749. https://dx.doi.org/10.1175/JCLI-D-11-00651.1.

Jin, Y. (2023) *U.PhyloMaker: Make phylogenetic hypotheses*. R package version 0.1.0.

Junker, R.R., Höcherl, N. & Blüthgen, N. (2010) Responses to olfactory signals reflect network structure of flower-visitor interactions. *Journal of Animal Ecology*, **79**, 818–23. https://dx.doi.org/10.1111/j.1365-2656.2010.01698.x.

Kaiser-Bunbury, C.N., Mougal, J., Whittington, A.E., Valentin, T., Gabriel, R., Olesen, J.M. & Blüthgen, N. (2017) Ecosystem restoration strengthens pollination network resilience and function. *Nature*, **542**, 223–227. https://dx.doi.org/10.1038/nature21071.

Kantsa, A., Raguso, R.A., Dyer, A.G., Olesen, J.M., Tscheulin, T. & Petanidou, T. (2018) Disentangling the role of floral sensory stimuli in pollination networks. *Nature Communications*, **9**, 1041. https://dx.doi.org/10.1038/s41467-018-03448-w.

Klein, A.M., Vaissière, B.E., Cane, J.H., Steffan-Dewenter, I., Cunningham, S.a., Kremen, C. & Tscharntke, T. (2007) Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society of London B*, **274**, 303–13. https://dx.doi.org/10.1098/rspb.2006.3721.

Klomberg, Y., Tropek, R., Mertens, J.E.J., Kobe, I.N., Hodeček, J., Raška, J., Fominka, N.T., Souto-Vilarós, D., Janečková, P. & Janeček, Š. (2022) Spatiotemporal variation in the role of floral traits in shaping tropical plant-pollinator interactions. *Ecology Letters*, **25**, 839–850. https://dx.doi.org/10.1111/ele.13958.

Lanuza, J.B., Knight, T.M., Montes-Perez, N., Glenny, W., Acuña, P., Albrecht, M., Artamendi, M., Badenhausser, I., Bennett, J.M., Biella, P., Bommarco, R., Cappellari, A., Castro, S., Clough, Y., Colom, P., Costa, J., Cyrille, N., de Manincor, N., Dominguez-Lapido, P., Dominik, C., Dupont, Y.L., Feldmann, R., Felten, E., Ferrero, V., Fiordaliso, W., Fisogni, A., Fitzpatrick, Ú., Galloni, M., Gaspar, H., Gazzea, E., Goia, I., Gómez-Martínez, C., González-Estévez, M.A., González-Varo, J.P., Grass, I., Hadrava, J., Hautekèete, N., Hederström, V., Heleno, R., Hervias-Parejo, S., Heuschele, J.M., Hoiss, B., Holzschuh, A., Hopfenmüller, S., Iriondo, J.M., Jauker, B., Jauker, F., Jersáková, J., Kallnik, K., Karise, R., Kleijn, D., Klotz, S., Krausl, T., Kühn, E., Lara-Romero, C., Larkin, M., Laurent, E., Lázaro, A., Librán-Embid, F., Liu, Y., Lopes, S., López-Núñez, F., Loureiro, J., Magrach, A., Mänd, M., Marini, L., Mas, R.B., Massol, F., Maurer, C., Michez, D., Molina, F.P., Morente-López, J., Mullen, S., Nakas, G., Neuenkamp, L., Nowak, A., O'Connor, C.J., O'Rourke, A., Öckinger, E., Olesen, J.M., Opedal, Ø.H., Petanidou, T., Piquot, Y., Potts, S.G., Power, E.F., Proesmans, W., Rakosy, D., Reverté, S., Roberts, S.P.M., Rundlöf, M., Russo, L., Schatz, B., Scheper, J., Schweiger, O., Serra, P.E., Siopa, C., Smith, H.G., Stanley, D., Ştefan, V., Steffan-Dewenter, I., Stout, J.C., Sutter, L., Švara, E.M., Świerszcz, S., Thompson, A., Traveset, A., Trefflich, A., Tropek, R., Tscharntke, T., Vanbergen, A.J., Vilà, M., Vujić, A., White, C., Wickens, J.B., Wickens, V.B., Winsa, M., Zoller, L. & Bartomeus, I. (2025) EuPPollNet: A European database of plant-pollinator networks. *Global Ecology and Biogeography*, **34**, e70000. https://dx.doi.org/10.1111/geb.70000.

Lomascolo, S., Chacoff, N. & Vázquez, D.P. (2022) Plant and flower visitor traits for plant-pollinator visitation network in Monte desert sites of Villavicencio Nature Reserve, Mendoza, Argentina in 2006-2011. *Figshare*. https://dx.doi.org/10.6084/m9.figshare.19424678.v4.

Maruyama, P.K., Vizentin-Bugoni, J., Oliveira, G.M., Oliveira, P.E. & Dalsgaard, B. (2014) Morphological and spatio-temporal mismatches shape a neotropical savanna plant-hummingbird network. *Biotropica*, **46**, 740–747. https://dx.doi.org/10.1111/btp.12170.

Mayfield, M.M., Waser, N.M. & Price, M.V. (2001) Exploring the 'most effective pollinator principle' with complex flowers: bumblebees and *Ipomopsis aggregata*. *Annals of Botany*, **88**, 591–596. https://dx.doi.org/10.1006/anbo.2001.1500.

Minachilis, K., Kantsa, A., Devalez, J., Vujic, A., Pauly, A. & Petanidou, T. (2023) High species turnover and unique plant–pollinator interactions make a hyperdiverse mountain. *Journal of Animal Ecology*, **92**, 1001–1015. https://dx.doi.org/10.1111/1365-2656.13898.

Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., Faure, D., Garnier, E., Gimenez, O., Huneman, P., Jabot, F., Jarne, P., Joly, D., Julliard, R., Kéfi, S., Kergoat, G.J., Lavorel, S., Le Gall, L., Meslin, L., Morand, S., Morin, X., Morlon, H., Pinay, G., Pradel, R., Schurr, F.M., Thuiller, W. & Loreau, M. (2015) Predictive ecology in a changing world. *Journal of Applied Ecology*, **52**, 1293–1310. https://dx.doi.org/10.1111/1365-2664.12482.

Novella-Fernandez, R., Rodrigo, A., Arnan, X. & Bosch, J. (2019) Interaction strength in plant-pollinator networks: Are we using the right measure? *PLoS ONE*, **14**, e0225930. https://dx.doi.org/10.1371/journal.pone.0225930.

Olito, C. & Fox, J.W. (2015) Species traits and abundances predict metrics of plant-pollinator network structure, but not pairwise interactions. *Oikos*, **124**, 428–436. https://dx.doi.org/10.1111/oik.01439.

Ollerton, J., Winfree, R. & Tarrant, S. (2011) How many flowering plants are pollinated by animals? *Oikos*, **120**, 321–326. https://dx.doi.org/10.1111/j.1600-0706.2010.18644.x.

Ornai, A. & Keasar, T. (2020) Floral complexity traits as predictors of plant-bee interactions in a Mediterranean pollination web. *Plants*, **9**, 1432. https://dx.doi.org/10.3390/plants9111432.

Ortega-Olivencia, A., Rodríguez-Riaño, T., Valtueña, F.J., López, J. & Devesa, J.A. (2005) First confirmation of a native bird-pollinated plant in europe. *Oikos*, **110**, 578–590. https://dx.doi.org/10.1111/j.0030-1299.2005.13877.x.

Peralta, G., CaraDonna, P.J., Rakosy, D., Fründ, J., Pascual Tudanca, M.P., Dormann, C.F., Burkle, L.A., Kaiser-Bunbury, C.N., Knight, T.M., Resasco, J., Winfree, R., Blüthgen, N., Castillo, W.J. & Vázquez, D.P. (2024) Predicting plant–pollinator interactions: concepts, methods, and challenges. *Trends in Ecology & Evolution*, **39**, 494–505. https://dx.doi.org/10.1016/j.tree.2023.12.005.

Pichler, M., Boreux, V., Klein, A., Schleuning, M. & Hartig, F. (2019) Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. *Methods in Ecology and Evolution*, **11**, 281–293. https://dx.doi.org/10.1111/2041-210X.13329.

Pickett, S.T.A., Kolasa, J. & Jones, C.G. (2007) *Ecological Understanding: The Nature of Theory and the Theory of Nature*. Academic Press, Amsterdam Boston, 2nd edition.

Rakosy, D., Motivans, E., Ştefan, V., Nowak, A., Świerszcz, S., Feldmann, R., Kühn, E., Geppert, C., Venkataraman, N., Sobieraj-Betlińska, A., Grossmann, A., Rojek, W., Pochrzast, K., Cielniak, M., Gathof, A.K., Baumann, K. & Knight, T.M. (2022) Intensive grazing alters the diversity, composition and structure of plant-pollinator interaction networks in Central European grasslands. *PLoS ONE*, **17**, e0263576. https://dx.doi.org/10.1371/journal.pone.0263576.

Renoult, J.P., Blüthgen, N., Binkenstein, J., Weiner, C.N., Werner, M. & Schaefer, H.M. (2015) The relative importance of color signaling for plant generalization in pollination networks. *Oikos*, **124**, 347–354. https://dx.doi.org/10.1111/oik.01361.

Resasco, J., Chacoff, N.P. & Vázquez, D.P. (2021) Plant–pollinator interactions between generalists persist over time and space. *Ecology*, **102**, e03359. https://dx.doi.org/10.1002/ecy.3359.

Revell, L.J. (2024) phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ*, **12**, e16505. https://dx.doi.org/10.7717/peerj.16505.

Riley, R.D., Lambert, P.C. & Abo-Zaid, G. (2010) Meta-analysis of individual participant data: rationale, conduct, and reporting. *British Medical Journal*, **340**, c221. https://dx.doi.org/10.1136/bmj.c221.

Robson, D.B. (2013) An assessment of the potential for pollination facilitation of a rare plant by common plants: *Symphyotrichum sericeum* (asteraceae) as a case study. *Botany*, **91**, 34–42. https://dx.doi.org/10.1139/cjb-2012-0133.

Rodger, J.G., Bennett, J.M., Razanajatovo, M., Knight, T.M., van Kleunen, M., Ashman, T.L., Steets, J.A., Hui, C., Arceo-Gómez, G., Burd, M., Burkle, L.A., Burns, J.H., Durka, W., Freitas, L., Kemp, J.E., Li, J., Pauw, A., Vamosi, J.C., Wolowski, M., Xia, J. & Ellis, A.G. (2021) Widespread vulnerability of flowering plant seed production to pollinator declines. *Science Advances*, **7**, eabd3524. https://dx.doi.org/10.1126/sciadv.abd3524.

Roswell, M., Dushoff, J. & Winfree, R. (2019) Male and female bees show large differences in floral preference. *PLoS ONE*, **14**, e0214909. https://dx.doi.org/10.1371/journal.pone.0214909.

Schwarz, B., Vázquez, D.P., CaraDonna, P.J., Knight, T.M., Benadi, G., Dormann, C.F., Gauzens, B., Motivans, E., Resasco, J., Blüthgen, N., Burkle, L.A., Fang, Q., Kaiser-Bunbury, C.N., Alarcón, R., Bain, J.A., Chacoff, N.P., Huang, S.Q., LeBuhn, G., MacLeod, M., Petanidou, T., Rasmussen, C., Simanonok, M.P., Thompson, A.H. & Fründ, J. (2020) Temporal scale-dependence of plant–pollinator networks. *Oikos*, **129**, 1289–1302. https://dx.doi.org/10.1111/oik.07303.

Schäffler, I. & Dötterl, S. (2011) A day in the life of an oil bee: phenology, nesting, and foraging behavior. *Apidologie*, **42**, 409–424. https://dx.doi.org/10.1007/s13592-011-0010-3.

Smith, S.D. (2010) Using phylogenetics to detect pollinator-mediated floral evolution. *New Phytologist*, **188**, 354–363. https://dx.doi.org/10.1111/j.1469-8137.2010.03292.x.

Stang, M., Klinkhamer, P.G.L. & van der Meijden, E. (2007) Asymmetric specialization and extinction risk in plant-flower visitor webs: a matter of morphology or abundance? *Oecologia*, **151**, 442–453.

Stefanescu, C. & Traveset, A. (2009) Factors influencing the degree of generalization in flower use by mediterranean butterflies. *Oikos*, **118**, 1109–1117. https://dx.doi.org/10.1111/j.1600-0706.2009.17274.x.

Sydenham, M.A.K., Dupont, Y.L., Nielsen, A., Olesen, J.M., Madsen, H.B., Skrindo, A.B., Rasmussen, C., Nowell, M.S., Venter, Z.S., Hegland, S.J., Helle, A.G., Skoog, D.I.J., Torvanger, M.S., Hanevik, K.A., Hinderaker, S.E., Paulsen, T., Eldegard, K., Reitan, T. & Rusch, G.M. (2024) Climatic conditions and landscape diversity predict plant–bee interactions and pollen deposition in bee-pollinated plants. *Ecography*, **2024**, e07138. https://dx.doi.org/10.1111/ecog.07138.

Sydenham, M.A.K., Venter, Z.S., Reitan, T., Rasmussen, C., Skrindo, A.B., Skoog, D.I.J., Hanevik, K., Hegland, S.J., Dupont, Y.L., Nielsen, A., Chipperfield, J. & Rusch, G.M. (2022) MetaComNet: A random forest-based framework for making spatial predictions of plant–pollinator interactions. *Methods in Ecology and Evolution*, **13**, 500–513. https://dx.doi.org/10.1111/2041-210X.13762.

Terry, J.C.D. & Lewis, O.T. (2020) Finding missing links in interaction networks. *Ecology*, **101**, e03047. https://dx.doi.org/https://doi.org/10.1002/ecy.3047.

Tinoco, B.A., Graham, C.H., Aguilar, J.M. & Schleuning, M. (2017) Effects of hummingbird morphology on specialization in pollination networks vary with resource availability. *Oikos*, **126**, 52–60. https://dx.doi.org/10.1111/oik.02998.

van Buuren, S. & Groothuis-Oudshoorn, K. (2011) mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, **45**, 1–67. https://dx.doi.org/10.18637/jss.v045.i03.

van der Niet, T. & Johnson, S.D. (2012) Phylogenetic evidence for pollinator-driven diversification of angiosperms. *Trends in Ecology & Evolution*, **27**, 353–361. https://dx.doi.org/10.1016/j.tree.2012.02.002.

Vizentin-Bugoni, J., Maruyama, P.K. & Sazima, M. (2014) Processes entangling interactions in communities: forbidden links are more important than abundance in a hummingbird – plant network. *Proceedings of the Royal Society London B: Biological Sciences*, **281**, 20132397–20132397. https://dx.doi.org/10.1098/rspb.2013.2397.

Vázquez, D.P., Chacoff, N. & Lomascolo, S. (2022) Flower abundance data for Monte desert sites in Villavicencio Nature Reserve for Austral springs in years 2006-2011. *Figshare*. https://dx.doi.org/10.6084/m9.figshare.19424129.v3.

Vázquez, D.P., Chacoff, N. & Cagnolo, L. (2009) Evaluating multiple determinants of the structure of plant-animal mutualistic networks. *Ecology*, **90**, 2039–2046.

Weiner, C.N., Werner, M., Linsenmair, K.E. & Blüthgen, N. (2014) Land use impacts on mutualistic networks: disproportional declines in specialized pollinators via changes in flower composition. *Ecology*, **95**, 466–474.

Weinstein, B.G. & Graham, C.H. (2017) On comparing traits and abundance for predicting species interactions with imperfect detection. *Food Webs*, **11**, 17–25. https://dx.doi.org/10.1016/j.fooweb.2017.05.002.

Willmer, P. (2011) *Pollination and Floral Ecology*. Princeton University Press.

Wilson, P. & Thomson, J.D. (1991) Heterogeneity among floral visitors leads to discordance between removal and deposition of pollen. *Ecology*, **72**, 1503–1507. https://dx.doi.org/10.2307/1941124.

Winfree, R., Griswold, T. & Kremen, C. (2007) Effect of human disturbance on bee communities in a forested ecosystem. *Conservation Biology*, **21**, 213–223. https://dx.doi.org/10.1111/j.1523-1739.2006.00574.x.

Winfree, R., Reilly, J.R., Bartomeus, I., Cariveau, D.P., Williams, N.M. & Gibbs, J. (2018) Species turnover promotes the importance of bee diversity for crop pollination at regional scales. *Science*, **359**, 791–793.

Winfree, R., Williams, N.M., Dushoff, J. & Kremen, C. (2014) Species abundance, not diet breadth, drives the persistence of the most linked pollinators as plant-pollinator networks disassemble. *American Naturalist*, **183**, 600–611. https://dx.doi.org/10.1086/675716.

Wright, M.N. & 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. https://dx.doi.org/10.18637/jss.v077.i01.

Yahaya, M.M., Rodger, J.G., Landi, P. & Hui, C. (2024) Emergence of structure in plant–pollinator networks: low floral resource constrains network specialisation. *Oikos*, **2024**, e10533. https://dx.doi.org/10.1111/oik.10533.

Young, J.G., Valdovinos, F.S. & Newman, M.E.J. (2021) Reconstruction of plant–pollinator networks from observational data. *Nature Communications*, **12**, 3911. https://dx.doi.org/10.1038/s41467-021-24149-x.

A Results for normalised Nash-Sutcliffe efficiency (NNSE)

In some fields of environmental science, particularly hydrology, predictive quality of a model is assessed relative to the skill of assuming a constant prediction would be. This is called, after their inventors, the Nash-Sutcliffe model efficiency NSE and is computed as:

$$\mathrm{NSE} = 1 - \frac{\sum(\hat{y}_i - y_i)^2}{\sum(\hat{y}_i - \bar{\hat{y}})^2} \tag{1}$$

In words, NSE is one minus the mean squared error of a prediction divided by the mean squared error of fitting just an intercept (or grand mean). NSE ranges between $-\infty$ (entirely unsuitable) and 1 (for a perfect model), but is frequently normalised to $[0, 1]$:

$$\mathrm{NNSE} = \frac{1}{2 - NSE}, \tag{2}$$

with 0.5 indicating same predictive skill as an intercept-only model, larger values indicating better than random skill, and values below 0.5 lack of predictive skill.

While not particularly common in ecology, NNSE has the same advantage as a correlation coefficient, namely that it can be compared across the different studies, unlike the (y-value-dependent) RMSE or (sample-size dependent) log-likelihood.

We repeat here the same analyses and figures as presented for Spearman's ρ in the main text.

Overall, results are qualitatively similar, but the lack of predictive skill is even more visible for NNSE.

A.1 Prediction method comparison

For temporal prediction, all models performed poorer than the grand average (which is identical to base 1). For spatial prediction, all model types were similar to the grand average reference, and base 2 and tapnet ever so slightly improved on it; no model was better than base 2 here (Fig A1).

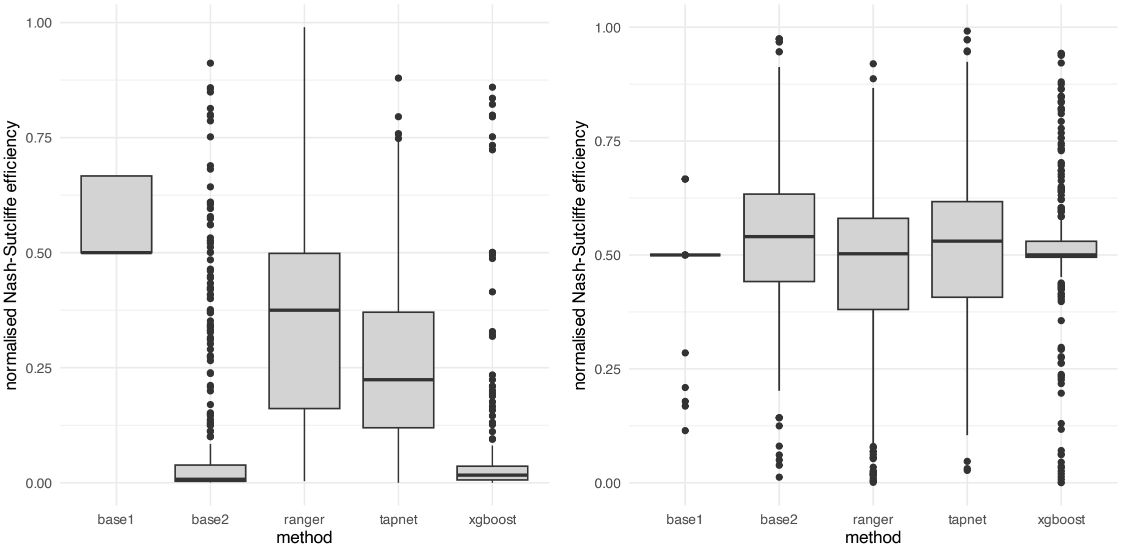

Figure A1: Predictive performance of three methods (ranger as implementation of randomForest, tapnet and XGboost) and two baseline models in temporal (*left*) and spatial (*right*) cross-validation. Values are normalised Nash-Sutcliffe efficiency, where a value < 0.5 indicate no skill in prediction.

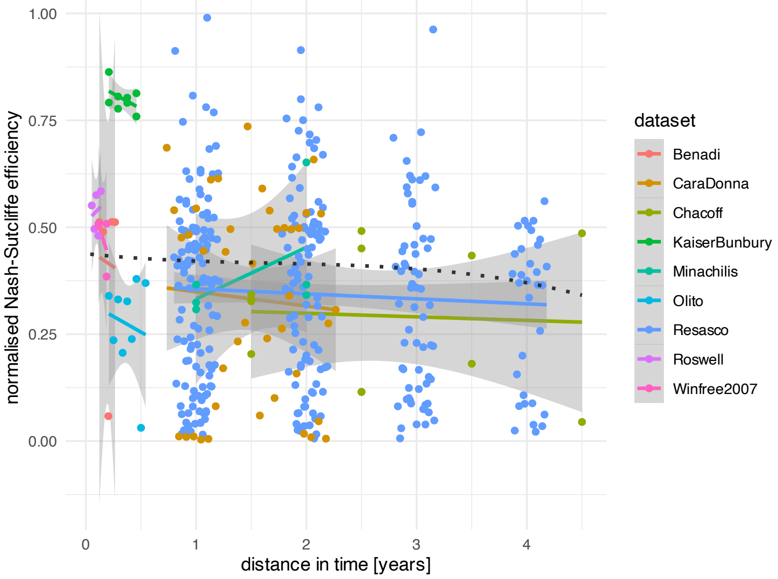

Figure A2: Effect of temporal distance between training and test data on prediction skill of random forest models, measured as normalised Nash-Sutcliffe efficiency. Colours represent different data sets, lines around regression are 95%-confidence intervals. Dotted black line is prediction from non-linear random-intercept model across all data sets (being not significantly related to temporal distance).

A.2 Random forest prediction in time

There was no detectable effect of temporal distance on predictive ability (Fig. A2), in contrast to the hump-shaped pattern we found for Spearman's ρ (Fig. 2).

Both plant and pollinator community compositional distance had a significant negative effect on the normalised Nash-Sutcliffe efficiency (Fig. A3), yet few studies had values above 0.5 (Kaiser-Bunburry, Roswell), and the qualitative picture is similar to the pattern described in the main text for Spearman's ρ (Fig. 3).

A.3 Random forest prediction in space

As for time, also in space compositional distance affected predictive skill negatively (Fig. A4). That effect was significant for plant community composition, but not so for pollinator composition. While the shape of the best-fitting polynomial was different, the trend downwards was similar to the results presented in the main text (Fig. 4).

A.4 Analysis of all data sets combined

Combining all studies and using random forest for predicting to the held-out study, yielded a normalised Nash-Sutcliffe efficiency of 0.498 ± 0.021, which is not significantly different from the no-skill reference of 0.5.

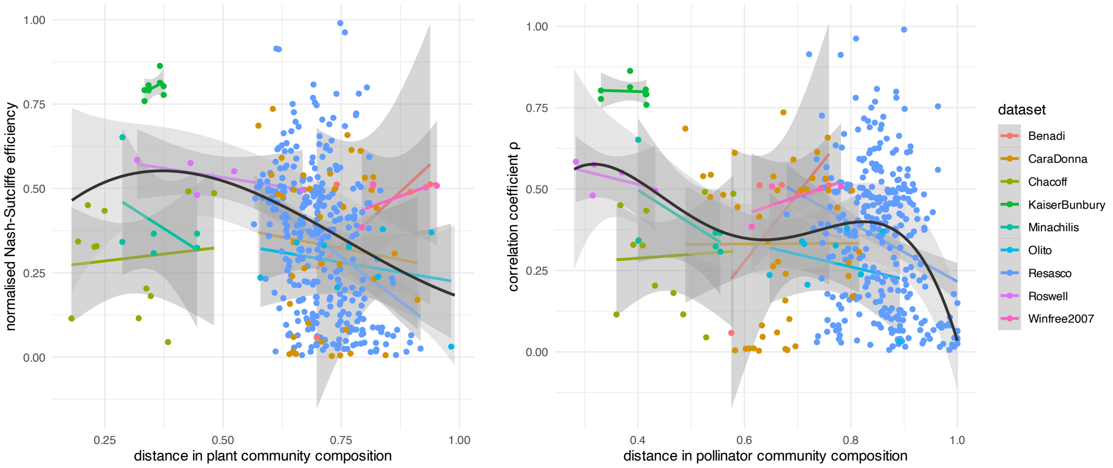

Figure A3: Effect of distance of community composition (*left*: plants; *right*: pollinators) between training and test data on temporal prediction skill (measured as normalised Nash-Sutcliffe efficiency) of randomForest models. Colours represent different data sets, black line is fit of non-linear random-intercept model across all data sets. Ribbons around regression are 95%-confidence intervals.

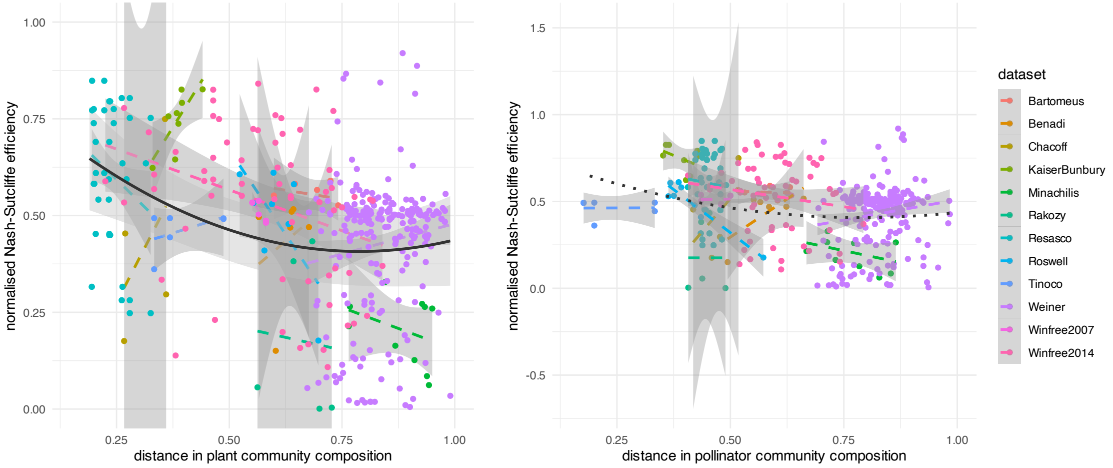

Figure A4: Effect of distance of community composition (*left*: plants; *right*: pollinators) between training and test data on spatial prediction skill (measured as normalised Nash-Sutcliffe efficiency) of randomForest models. Colours represent different data sets, black line is fit of non-linear random-intercept model across all data sets; dotted line indicates no significant relationship. Ribbons around regression are 95%-confidence intervals

B Explorative analysis of Kaiser-Bunbury et al. (2017)

This section presents the preliminary analyses of the 8 replicated networks at each of 8 sites of Kaiser-Bunbury *et al.* (2017). The point was to trial different methods and analyses before applying them consistently to all other studies.

B.1 XGBoost hyper-parameter search

The hyper-parameter search, based on the first four networks in the pooled-sites analysis, yielded the following settings: max.depth = 3, eta = 0.01, nrounds=2000 and gamma=1. These were used for all spatial and temporal validation settings of this data set. The resulting setting was substantially better than a previous best guess.[5]

B.2 MLP hyper-parameter search

The hyper-parameter search, based on the first four networks in the pooled-sites analysis, yielded the following settings: loss="poisson", batchsize=2000, epochs=200, lambda=0.01, alpha=0.47, lr = 0.0001, activation = "selu", hidden=c(20, 2). These were used for all spatial and temporal validation settings of this data set. Note that the number of nodes is rather small for this analysis. The resulting setting was substantially better than a previous best guess.

B.3 Validation score analysis

Although we pre-selected log-likelihood (ℓ) and median absolute difference (MAD) as indicators of model performance, we additionally computed others (see Methods). The correlation among these performance measures differed between temporal and spatial predictions (Table A1). The log-likelihood ℓ was least correlated with the others, while MAD and dr were well correlated with the other measures. Using both ℓ and MAD thus seems to cover most ways in which prediction errors can be assessed.

Table A1: Correlation (Pearson's r) between different measures of prediction quality. Upper triangle is prediction in time, lower triangle in space. Measures are log-likelihood (ℓ), median absolute difference (MAD), root mean squared error (RMSE), Spearman's correlation coefficient ρ, Nash-Sutcliffe efficiency (r_{NE}) and Willmott's discrepancy d_r. All correlations are significant ($p < 0.05$), strong ones are highlighted.

	ell	MAD	RMSE	ρ	r_{NE}	d_r
ell	1.000	-0.523	-0.465	0.542	0.481	0.529
MAD	−0.529	1	0.668	**−0.870**	**−0.744**	**−0.930**
RMSE	−0.592	0.550	1	−0.499	**−0.755**	**−0.681**
ρ	0.530	**−0.838**	−0.480	1	0.692	**0.932**
r_{NE}	0.282	−0.587	−0.339	0.671	1	**0.870**
d_r	0.441	**−0.814**	−0.468	**0.925**	**0.756**	1

B.4 Compositional analysis

We compute the compositional distance between all networks (for time and space separately). For time, the range of Bray-Curtis distances between the plants of any two networks was 0.20-0.44, and for pollinators 0.25-0.46. Thus, composition did not vary that much, despite the exotic plant removal treatment, sharing more than 50% of species of plants and pollinators.

---

[5] Thanks go to cito package maintainer Maximilian Pichler for helping us with this hyper-parameter optimisation!

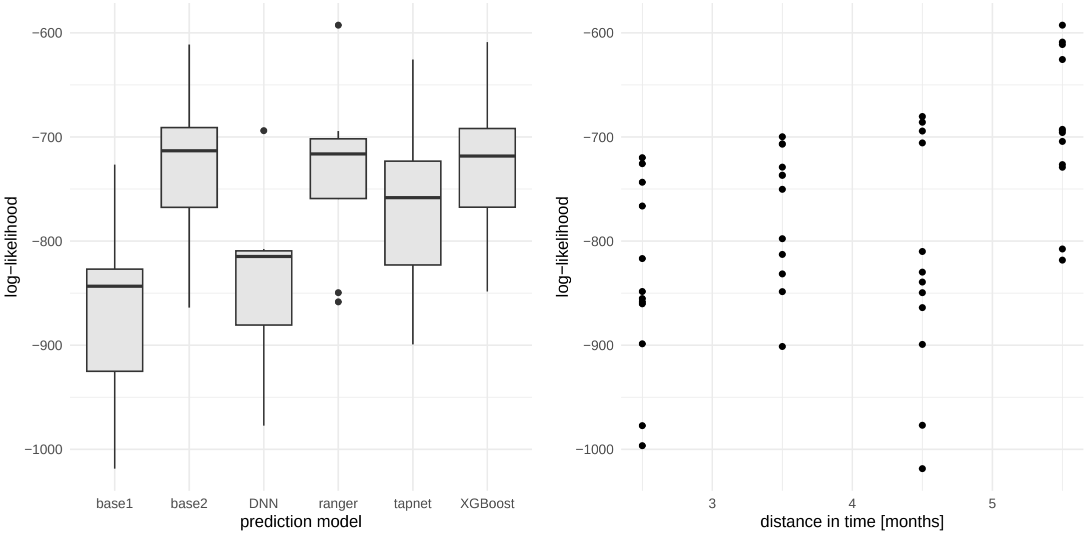

Figure A5: Quality of temporal predictions, as measured by log-likelihood (more is better). *Left*: The effect of prediction method. Base 1 refers to predicting a constant interaction intensity across the network, base 2 is predicted by abundances only. DNN is the neural network. Each box is computed from eight cross-validation values. *Right*: Predictions become *better* as more months pass between the training and the test data.

Distances showed a similar pattern for plants and pollinators, i.e. when two networks were similar in plant composition, they were also similar in pollinator composition ($r = 0.79$). That means, if we interpret a compositional effect, it should not differ dramatically between plants and pollinators.

B.4.1 Temporal analysis

The analysis show a clear effect of the method on prediction quality, as well as an effect of distance in time (Table A2, Fig. A5).

In terms of methods, the neural network (“DNN”) performed poorly, and indeed worse than the parameter-free abundance-only base 2. Random forest (“ranger”), XGboost and tapnet were very similar, with slight advantages for the machine-learning approaches.

It is not obvious, how to explain the better predictions to data sets further apart in time (Fig. A5, right). The only interpretation to offer is that the statistical models overfit the data, i.e. interpret some phylogenetic signal where in fact there is only noise. Since networks further apart in time have fewer overlapping species, this overfitting has less effect than predicting to similarly composed plant or pollinator communities in the nearer future or past. However, there is no indication in the analysis that this is a correct interpretation, since neither distance in plant nor

Table A2: ANOVA-result for cross-validation in time, using log-likelihood ℓ as measure of prediction quality. Significant effects are indicated by bold-printed p-values. The model’s $R^2 = 0.46$. See Fig. A5 for effects.

ℓ	Df	Sum Sq	Mean Sq	F value	Pr(>F)
method	5	153447	30689	5.468	**0.0007**
Dtime	1	91473	91473	16.299	**0.0002**
Dlower	1	7473	7473	1.332	0.2556
Dhigher	1	21307	21307	3.797	0.0586
Residuals	39	218873	5612		

Table A3: ANOVA-result for cross-validation in time, using MAD as measure of prediction quality. Significant effects are indicated by bold-printed p-values. The model's $R^2 = 0.94$. See Fig. A6 for effects.

MAD	Df	Sum Sq	Mean Sq	F value	Pr(>F)
method	5	15.39	3.08	153.439	**0.0000**
Dtime	1	0.02	0.02	0.916	0.3443
Dlower	1	0.02	0.02	0.760	0.3887
Dhigher	1	0.09	0.09	4.563	**0.0390**
Residuals	39	0.78	0.02		

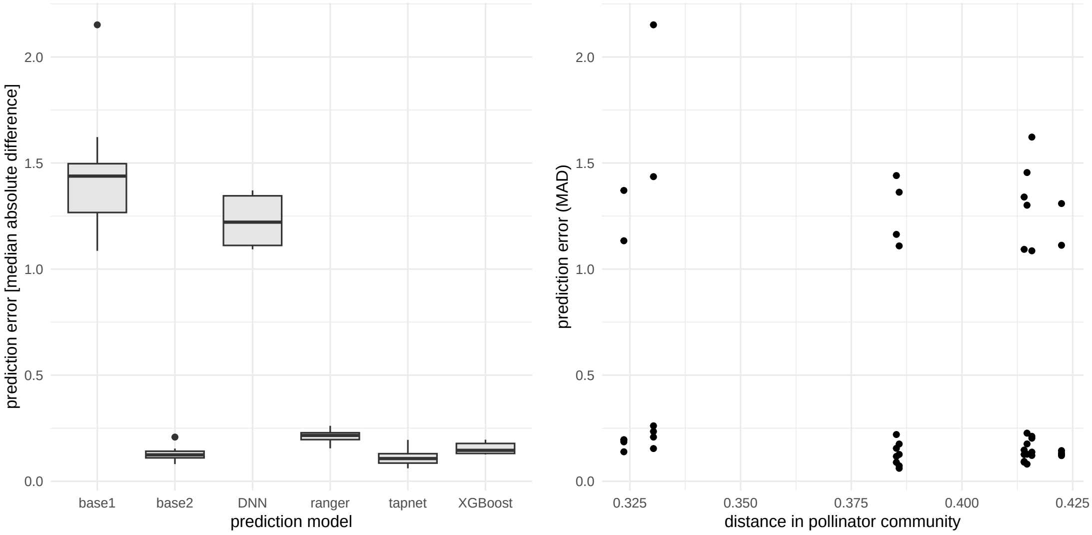

Figure A6: Quality of temporal predictions, as measured by MAD (less is better). *Left*: The effect of prediction method. *Right*: Predictions tend to have *less* error as pollinator communities differ more between training and test data. Note that higher MAD values are due to base 1 and DNN, which also drive this insignificant trend.

pollinator community yields a significant effect (Table A2).

For the alternative measure of prediction quality, the median absolute difference (MAD), the results point towards an effect of pollinator composition, although this is not significant (Table A3). Again, DNN performed poorly, while tapnet had the lowest prediction error, slightly better than XGboost.

B.4.2 Spatial analysis

The results for predictions in space are by and large similar to those in time (Tables A4, A5, Figs. A7, A8). Again the DNN had substantial difficulty in predicting well. The effect of space was practically absent. Plant similarity modified this pattern a bit, but there are too few data points to make this a consistent effect (Fig. A7, right).

For MAD, only the method could be detected as significant effect (Table A5, Fig. A8). As for time, DNNs were not able to predict well.

B.4.3 Compositional distances in time and space

Distances in time were not at all related to distances in composition of either plant or pollinator communities ($r = 0$, Fig. A9, left).

For spatial cross-validation, the picture was slightly different (Fig. A9, right) . Here, there was

Table A4: ANOVA-result for cross-validation in space, using log-likelihood as measure of prediction quality. Significant effects are indicated by bold-printed p-values. The model's $R^2 = 0.93$. See Fig. A7 for effects.

ℓ	Df	Sum Sq	Mean Sq	F value	Pr(>F)
method	5	125214	25043	33.62	**0.0000**
Dspace	1	8	8	0.01	0.9186
Dlower	1	490	490	0.66	0.4265
Dhigher	1	7293	7293	9.79	**0.0051**
method:Dspace	5	1793	359	0.48	0.7861
method:Dlower	5	1075	215	0.29	0.9140
method:Dhigher	5	1010	202	0.27	0.9238
Dspace:Dlower	1	243235	243235	326.57	**0.0000**
Dspace:Dhigher	1	89150	89150	119.70	**0.0000**
Dlower:Dhigher	1	664	664	0.89	0.3560
Residuals	21	15641	745		

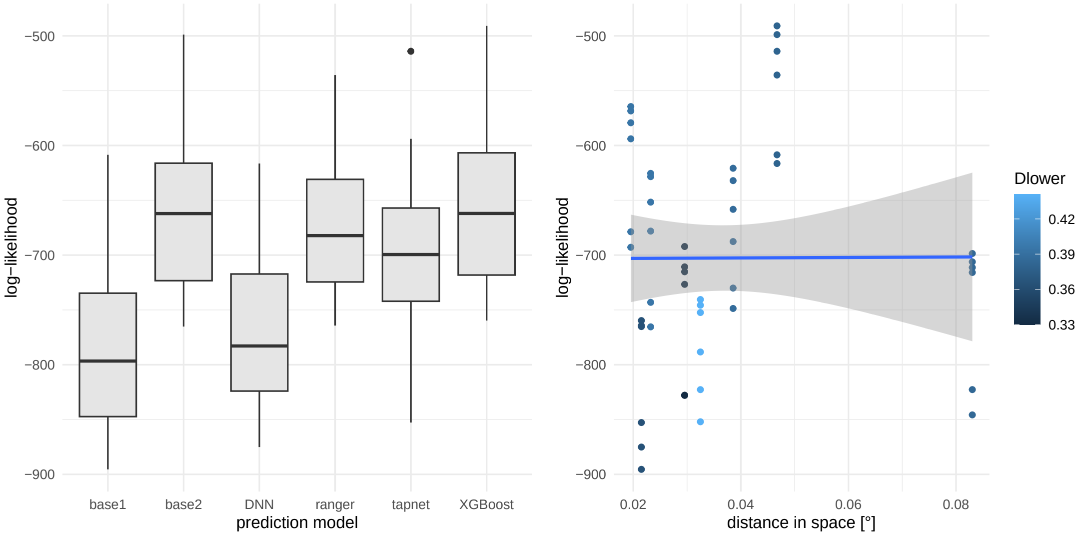

Figure A7: Quality of spatial predictions, as measured by log-likelihood (more is better). *Left*: The effect of prediction method. *Right*: The significant interaction between spatial distance and compositional distance of plants.

still no correlation of spatial distance with the dissimilarity of the plant community ($r = 0.012$), but pollinator communities were surprisingly *less* dissimilar the farther they were apart. Or, to get rid of the double negative: Pollinator communities were more similar, the farther the sites were apart ($r = -0.48$, $p < 0.001$).

B.5 Analysis of common species only

Restricting the predictions to only the most common species substantially improved temporal prediction quality. This is particularly noticeable for the NNSE, which is always and often much higher than the no-skill threshold of 0.5 (Fig. A10). However, the best performing model was base2 (NNSE > 0.9, all others < 0.8), indicating that abundance alone was responsible, not trait information or alike.

B.6 Conclusions for further analyses

Judging from these analyses, base 1 and 2 make for a good set of reference analyses.

Table A5: ANOVA-result for cross-validation in space, using MAD as measure of prediction quality. Significant effects are indicated by bold-printed p-values. The model's $R^2 = 0.85$. See Fig. A7 for effects.

MAD	Df	Sum Sq	Mean Sq	F value	Pr(>F)
method	5	23.54	4.71	51.84	**0.0000**
Dspace	1	0.05	0.05	0.50	0.4840
Dlower	1	0.54	0.54	5.96	**0.0193**
Dhigher	1	0.03	0.03	0.32	0.5743
Residuals	39	3.54	0.09		

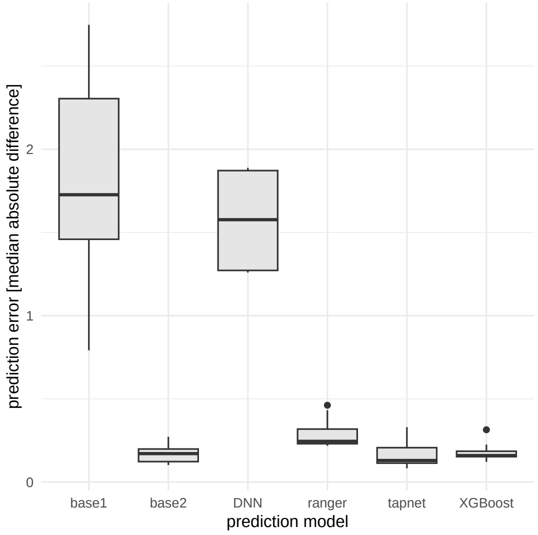

Figure A8: Quality of spatial predictions by different methods, as measured by MAD (less is better).

The neural network does not improve prediction beyond base 2. Unless a way to improve on this performance can be found, the machine-learning approaches are preferable. (It is a common phenomenon, that neural networks excel on visual data, but not on tabular data such as these.) So, DNN could be dropped from the analysis without loss of maximal prediction quality.

Tapnet did well, being slightly worse in the predictions in terms of log-likehood, but slightly better than ML in MAD. Note, however, that even there the MAD-score of base 2 was at least as good (Fig. 1 left, 3 left, 4). Given the considerable longer runtimes, tapnet could be dropped without loss of maximal prediction quality.

This leaves us with baselines 1 and 2, random forest and extreme gradient boosting as the four methods to run further predictions with.

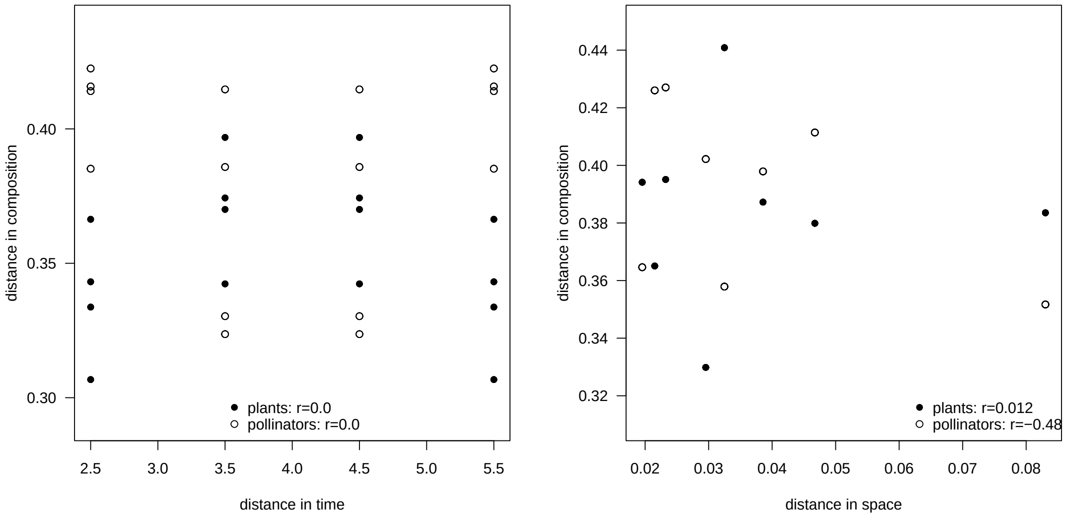

Figure A9: Distances in time (left, in months) and space (right, in geographic degrees) between the training and test networks, compared to the compositional Bray-Curtis distances in the plant (black dots) and pollinator (white dots) communities.

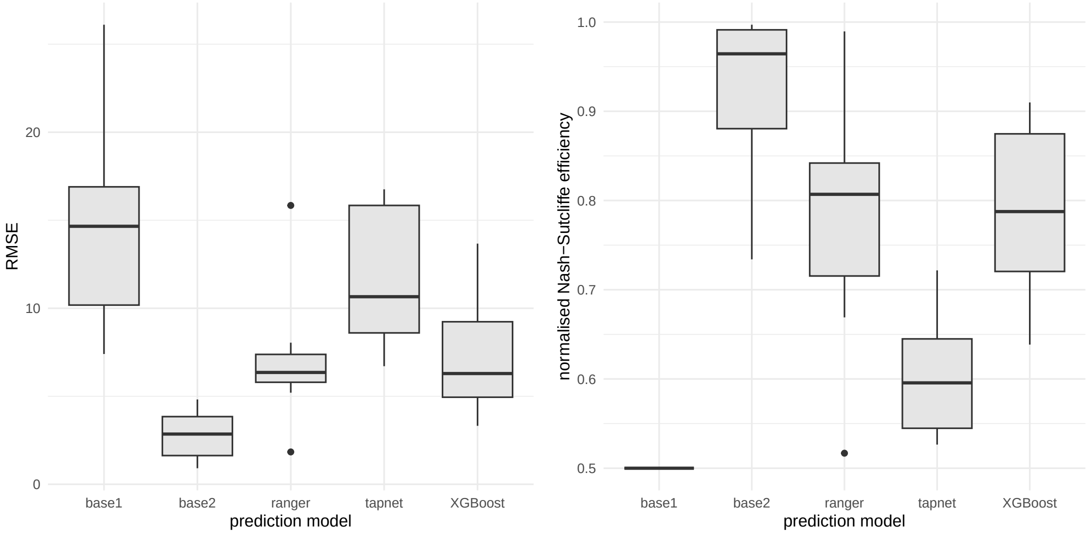

Figure A10: Effect of prediction method on RMSE (*left*) and normalised Nash-Sutcliffe efficiency (*right*) based only on the most common plant and pollinator species.

C Preparing data for analysis

Data preparation for the analysis within each study has to follow several steps.

Homogenisation of species names All elements of the analyses are linked through the names of the species. Thus, these have to be perfectly identical in all data: the trait data, the external abundances and the network (row and column names). Suggestion: Use the Latin binomial with an underscore ("_") as unique identifier. If a species is only known at genus, family or order level, use that with a unique number, e.g. "Diptera_3" or so.

Species taxonomy and trait table Provide two tables, one for plants, one for pollinators, of the species name, genus, family and order along with any trait data (each of these being a column). This is the input to working out the phylogeny and taxonomy using R's taxize. Thus, you may have to go back to this table to change names, if GBIF uses a different synonym. Maybe add another column, "original name" whenever that happens, so that we keep both the original data and the renaming.

Network names All networks will eventually form a list of networks. This list can be named, but in the analysis it will (typically) be referred to by position in the list (i.e. network 1:8). If there is a logic to the networks, e.g. years 1:8 or sites from north to south, then please organise the networks in that sequence (which makes it easier for splitting into training and testing data).

Abundance vectors per network If you provide external abundance data, these will eventually be transformed into a named vector, i.e. one such named vector per plant and pollinator per network. Make sure the data are prepared in tables such that such a vector can be computed easily.

Using `make_tapnet` to guarantee correct format In the R package "tapnet", the function `make_tapnet` pulls together all the above information and makes a single "tapnet" object of it. Doing so, it checks that names are identical and so forth. Thus, in the end of the data preparation, please use `make_tapnet` to ensure everything is organised well. This is a moment of great frustration! I repeatedly found myself cursing the computer because I was sure I did everything correctly – only to find I did not. There is a glitch in `make_tapnet`, if you have a single trait and only for one of the groups. In this case, `make_tapnet` will complain of a missing trait in the other group. Please simply invent a trait for that other group, e.g. just assign a value of 1 to each species and call the trait "fakeTrait" or so.

C.1 Details in R

C.1.1 Phylogeny, taxonomy

For plants, a global phylogeny can be produced using "U.PhyloMaker" (which has to be installed from github) like so:

```
#devtools::install_github(”jinyizju/U.PhyloMaker”)
library(”U.PhyloMaker”)
# plant megatree and genus/family file from here:
# https://github.com/megatrees/plant_20221117/blob/main/plant_megatree.tre
megatreePlants <- read.tree(”../plant_megatree.tre”)
genlistPlants <- read.csv(”../plant_genus_list.csv”)
plantResult <- phylo.maker(plantSpeciesList, megatreePlants, genlistPlants,
```

```
      nodes.type = 1, scenario = 3)    # takes a minute or two
str(plantResult)
```

You have to download the megatree and the genus/family tree (or use the one downloaded on googleDrive).

This process may identify species not in the phylogeny, e.g. due to spelling mistakes. Please fix this!

Plot the resulting tree (e.g. into a long PDF) and check that all species were correctly placed!

```
library(ape)
library(phytools)
plantPhylo <- plantResult$phylo

pdf(file="MyData_plantPhylogeny.pdf", height=15, width=8)
par(mar=c(1,1,1,1))
plot(plantPhylo, cex=0.5) # node labels uninformative
dev.off()
```

For pollinators, no global phylogeny is available, so we resort to using only their taxonomy to make a pseudo-phylogenetic tree.

```
library(taxize)
pollinatorsUnique <- sort(unique(pollSpeciesList$Species))

ac <- classification(pollinatorsUnique, db = "gbif", return_id = TRUE, rows = 1 )
# ac = Animal classification (list).
(labelsMissingInDatabase <- which(!sapply(ac, is.data.frame)))
```

Again, this will lead to many unidentified species, which you have to handle one by one manually! There are many lines of code in the dataPrep4Weiner2011.R-file demonstrating what such manual adaptations can look like.

```
polltree <- class2tree(ac, check = T) # rewrites labels!
# correct names: class2tree uses the lowest level label in ac;
# this line puts the name of ac back into the tip label
polltree$phylo$tip.label <- polltree$names

pdf(file="myData_pollTaxonomy.pdf", height=50, width=8)
par(mar=c(0,0,0,0))
plot(polltree, show.node.label=T, cex=0.5)
dev.off()
```

Check the resulting pseudo-phylogeny! You probably have to go back and re-assign species to different families, if the classification got it wrong, as in these examples:

```
# Checks based on taxonomy plot:
#Zygoptera !! outgroup
ac$"Zygoptera" <- classification("Libellulidae", db = "gbif", return_id = TRUE, rows = NA)[[1]]
#Cryptinae_cf._Glyphicnemis, Acari !! Bivalvia
ac$"Cryptinae_cf._Glyphicnemis" <- classification("Ichneumonidae", db = "gbif",
        return_id = TRUE, rows = 1)[[1]]
ac$"Cryptinae" <- classification("Ichneumonidae", db = "gbif", return_id = TRUE, rows = 1)[[1]]
ac$"Acari" <-  classification("Trombidiformes", db = "gbif", return_id = TRUE, rows = NA)[[1]]
```

C.1.2 Interactions into list of networks

This process is greatly facilitated by the `frame2webs`-function in bipartite. If your interactions are in a long list of plant A interacting with pollinator B x times in site s, this looks like:

```
library(bipartite)
network.list <- frame2webs(dats, varnames=c(”A”, ”B”, ”site”, ”x”))
```

That is surprisingly painless!

C.1.3 External abundances

Here you have to write rather case-specific code! Below is an example for a list of networks for a site called “AEG”. It loops through all networks, then checks in the abundance table `blub` for the plant species with the right name for that row and puts it into a named vector. That vector is then put into a list itself. If a species is in the network, but not the species list, it sets its abundance to 1.

```
blub <- rbind(plantSpeciesList1, plantSpeciesList2, plantSpeciesList3)
options(”warn”=0) # turns all warnings into errors!
abunsListAEG <- list()
for (i in seq_along(names(network.list.AEG))){
  plantNames <- attr(network.list.AEG[[i]], ”dimnames”)$Plant
  # get all species for that plot:
  perPlot <- blub[blub$EP_ID == names(network.list.AEG)[i]  & blub$Species %in% plantNames,]
  # get maximal floral coverage for each specis
  abunsPlot <- tapply(perPlot$Flowering_unit, perPlot$Species, max)
  if (length(rownames(network.list.AEG[[i]])) != length(abunsPlot)) {
    abunsPlot <- c(abunsPlot, rep(1, length(setdiff(plantNames, names(abunsPlot)))))
    #adds abundance of 1 (equivalent to 1% cover) for plants missing in external abundances
    nAP <- length(abunsPlot)
    nNew <- length(setdiff(plantNames, names(abunsPlot)))
    names(abunsPlot)[(nAP-nNew+1):nAP] <- setdiff(plantNames, names(abunsPlot))
  }
  abunsListAEG[[i]] <- abunsPlot
}
abunsListAEG
```

C.1.4 Check using `make_tapnet`

With all the above successfully sorted (after days of hard work and tearing out your hair for all the mistakes in the tables), only two steps remain: check that everything is nicely aligned and named, and convert it to the data actually required for the analysis.

```
AEG_tapnet <- make_tapnet(tree_high = pollPhylo,
                          tree_low=plantPhylo,
                          networks = network.list.AEG,
                          abun_low = NULL,
                          abun_high=NULL,
                          use.all.pems = T)

# check when error:
for (i in 1:39){
  if (sum(!(rownames(network.list.AEG[[i]]) %in% sort(plantPhylo$tip.label)))) stop()
  if (sum(!(colnames(network.list.AEG[[i]]) %in% sort(pollPhylo$tip.label)))) stop()
```

```
}
rownames(network.list.AEG[[i]]) %in% sort(plantPhylo$tip.label)
colnames(network.list.AEG[[i]]) %in% sort(pollPhylo$tip.label)

save(AEG_tapnet, HEG_tapnet, SEG_tapnet, file=”Weiner_networksAHSEG.Rdata”)

# AND FINALLY:
WeinerAEG_full <- tapnet2df(AEG_tapnet)
#...
save(WeinerAEG_full, WeinerHEG_full, WeinerSEG_full, file=”Weiner_tapnets_full.RData”)
```

The analysis will use both the actual tapnet object (e.g. to get the networks as such) and the "full" data frame created in the last step. (This is because the methods require different formats, e.g. base 1 and 2 and tapnet want the tapnet-object, while machine learning is fine with the table.)

C.2 Predictive analysis

The analysis itself is then carried out in a different file in the "analysis" folder, e.g. `Weineranalysis.R`.