

RESEARCH ARTICLE

Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks

Maximilian Pichler¹  | Virginie Boreux² | Alexandra-Maria Klein²  |
Matthias Schleuning³  | Florian Hartig¹ ¹Theoretical Ecology, University of Regensburg, Regensburg, Germany²Nature Conservation and Landscape Ecology, University of Freiburg, Freiburg, Germany³Senckenberg Biodiversity and Climate Research Centre (SBiK-F), Frankfurt (Main), Germany

Correspondence

Maximilian Pichler
Email: maximilian.pichler@biologie.uni-regensburg.de

Handling Editor: Luisa Carvalheiro

Abstract

1. Ecologists have long suspected that species are more likely to interact if their traits match in a particular way. For example, a pollination interaction may be more likely if the proportions of a bee's tongue fit a plant's flower shape. Empirical estimates of the importance of trait-matching for determining species interactions, however, vary significantly among different types of ecological networks.
2. Here, we show that ambiguity among empirical trait-matching studies may have arisen at least in parts from using overly simple statistical models. Using simulated and real data, we contrast conventional generalized linear models (GLM) with more flexible Machine Learning (ML) models (Random Forest, Boosted Regression Trees, Deep Neural Networks, Convolutional Neural Networks, Support Vector Machines, naïve Bayes, and k-Nearest-Neighbor), testing their ability to predict species interactions based on traits, and infer trait combinations causally responsible for species interactions.
3. We found that the best ML models can successfully predict species interactions in plant-pollinator networks, outperforming GLMs by a substantial margin. Our results also demonstrate that ML models can better identify the causally responsible trait-matching combinations than GLMs. In two case studies, the best ML models successfully predicted species interactions in a global plant-pollinator database and inferred ecologically plausible trait-matching rules for a plant-hummingbird network from Costa Rica, without any prior assumptions about the system.
4. We conclude that flexible ML models offer many advantages over traditional regression models for understanding interaction networks. We anticipate that these results extrapolate to other ecological network types. More generally, our results highlight the potential of machine learning and artificial intelligence for inference in ecology, beyond standard tasks such as image or pattern recognition.

KEYWORDS

bipartite networks, causal inference, deep learning, hummingbirds, insect pollinators, machine learning, pollination syndromes, predictive modelling

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

1 | INTRODUCTION

The understanding and analysis of species interactions in ecological networks has become a central building block of modern ecology. Research in this field, however, has concentrated in particular on analyzing observed network structures (e.g. Galiana et al., 2018; González, Dalsgaard, & Olesen, 2010; Mora, Gravel, Gilarranz, Poisot, & Stouffer, 2018; Poisot, Stouffer, & Gravel, 2015). Our understanding of why particular species interact, and others not, is comparatively less developed (cf. Bartomeus et al., 2016; Poisot et al., 2015). A key hypothesis regarding this question is that species interact when their functional properties (traits) make an interaction possible (e.g. Eklöf et al., 2013; Jordano, Bascompte, & Olesen, 2003). In plant–pollinator networks, for example, one would imagine that an interaction is easier to achieve when the tongue or body of the bee matches with the shape and size of the flower (Garibaldi et al., 2015; Stang, Klinkhamer, & van der Meijden, 2007). The idea that interactions will occur when traits are compatible is known as trait-matching (e.g. Schleuning, Fründ, & García, 2015, see also Figure 1).

The assumption that trait-matching is important for species interactions is engraved in many other ecological ideas and hypotheses. For example, trait-matching is a prerequisite for the idea of pollination syndromes (i.e. the hypothesis that flower and pollinator traits co-evolve, Faegri & van der Pijl, 1979; see also Fenster, Armbruster, Wilson, Dudash, & Thomson, 2004; Ollerton et al., 2009, and Rosas-Guerrero et al., 2014). Moreover, it has been suggested that trait-matching occurs also in other mutualistic ecological networks, for example fruit–frugivore interactions (e.g. Dehling et al., 2014), or antagonistic ecological networks, for example host–predator or host–parasitoid networks (Gravel, Poisot, Albouy, Velez, & Mouillot, 2013; see also Eklöf et al., 2013). Trait-matching between species has ample consequences for fundamental research, such as the identification and prediction of species interactions (Bartomeus et al., 2016, see Valdovinos, 2019), but also impacts ecosystem management. For example, it could be used for identifying effective pollinators to optimize production of pollinator-dependent crops (Garibaldi et al., 2015; Bailes, Ollerton, Patrick, & Glover, 2015; see

Potts et al., 2016). Finally, explaining and predicting links between interaction partners from information about their properties has applications far beyond ecology. An example is molecular medicine, where analogue concepts are used to study gene association (e.g. van Laarhoven & Marchiori, 2013; Menden et al., 2013; Yamanishi, Araki, Gutteridge, Honda, & Kanehisa, 2008; Zhang, Wang, Xi, Yang, & Li, 2018) or harmful drug–drug interactions (e.g. Cheng & Zhao, 2014; Tari, Anwar, Liang, Cai, & Baral, 2010).

While the idea of trait-matching itself is intuitive, it is less clear how important this mechanism is for determining species interactions (Bartomeus et al., 2016; Eklöf et al., 2013). On the one hand, recent findings in plant–pollinator networks support the concept of pollination syndromes (Rosas-Guerrero et al., 2014) and the utility of syndromes for predicting or understanding species interactions (Danieli-Silva et al., 2012; Murúa & Espíndola, 2015; Fenster, Reynolds, Williams, Makowsky, & Dudash, 2015; see Garibaldi et al., 2015). Recent studies also demonstrate that species interactions can be reasonably well predicted with phylogenetic predictors (Brousseau, Gravel, & Handa, 2018; Pearse & Altermatt, 2013; Pomeranz, Thompson, Poisot, & Harding, 2019), which supports the idea of trait-matching when assuming that traits are phylogenetically conserved. Similarly, studies of mutualistic pollination and seed-dispersal networks have accumulated evidence for strong signals of trait-matching, in particular in diverse tropical ecosystems (Dehling, Jordano, Schaefer, Böhning-Gaese, & Schleuning, 2016; Maglianesi, Blüthgen, Böhning-Gaese, & Schleuning, 2014). On the other hand, many ecological networks show low to moderate levels of specialization (Blüthgen, Menzel, Hovestadt, Fiala, & Blüthgen, 2007) and high flexibility in partner choice (Bender et al., 2017), questioning the idea of strong co-evolutionary feedback loops between plants and animals (Janzen, 1985; Ollerton et al., 2009). Moreover, while there is some direct evidence for trait–trait relationships as predictors for trophic interactions in simple prey–predator networks (Gravel et al., 2013), recent models that relied solely on trait–trait predictors (without phylogenetic predictors) showed only moderate performance in predicting species interactions (Brousseau et al., 2018; Pomeranz et al., 2019).

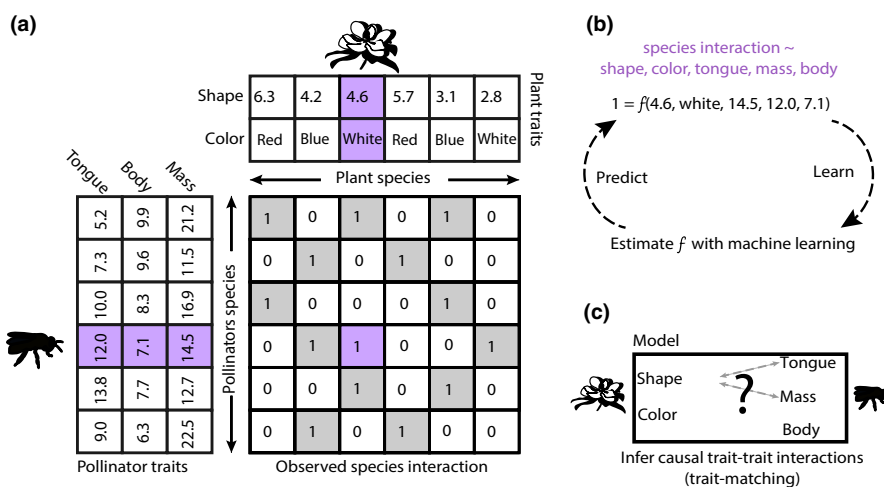


FIGURE 1 An illustration of the trait-matching concept. (a) Two classes of organisms, each with their own traits, interact in a bipartite network. (b) The goal of the statistical algorithm is to predict the probability of a plant–pollinator interaction, based on their trait values and (c) to infer the trait–trait interaction structure (trait-matching) that is causally responsible for those interactions

Progress on these questions is complicated by the fact that, until very recently, analyses of empirical networks relied almost exclusively on conventional regression models and phylogenetic predictors (Brousseau et al., 2018; Pearse & Altermatt, 2013; Pomeranz et al., 2019), or on simple regression trees (e.g. Berlow et al., 2009). Reasonable doubts exist as to whether these models are flexible enough to capture the way traits give rise to interactions (see e.g. Mayfield & Stouffer, 2017). Machine Learning (ML) models could be a solution to this problem. Modern ML models can flexibly detect interactions between predictors (trait-trait interactions), depend on fewer a-priori assumptions and usually achieve higher predictive performance than traditional regression techniques (e.g. Breiman, 2001b). State-of-the-art deep learning algorithms can detect complex pattern (e.g. LeCun, Bengio, & Hinton, 2015) and excel in tasks such as image or species recognition (e.g. Gray et al., 2019; Tabak et al., 2019). In food webs, recent findings demonstrate the potential of ML models in predicting species interactions. For example, Desjardins-Proulx et al. (2017) report that both a k-nearest neighbor and random forest (based on phylogenetic relationships and traits) can successfully predict food web interactions. It therefore seems promising to further explore the performance of machine learning algorithms for predicting species interactions from measurable traits, and whether those more flexible models change our view on the importance of trait matching for plant-pollinator interactions.

When assessing the suitability of ML algorithms for this problem, it is important to note that, while ML models tend to excel in predictive performance, their interpretation is often challenging (e.g. Ribeiro, Singh, & Guestrin, 2016). Ecologists, however, would likely not be satisfied with predicting species interactions, but would also want to know which traits are causally responsible for those interactions, for instance due to their importance as essential biodiversity variables (see Kissling et al., 2018). Unlike for statistical models, however, fitted ML models typically provide no direct information about how they generate their predictions. In recent years, also in response to issues such as fairness and discrimination (see Olhede & Wolfe, 2018), techniques aiming at interpreting fitted ML models have emerged (e.g. Guidotti et al., 2018). For example, permutation techniques (Fisher, Rudin, & Dominici, 2018) allow estimating the importance of predictors for any kind of model, similar to the variable importance in tree-based models (Breiman, 2001a). In this case, however, we are not primarily interested in the effects of a single predictor, but we want to know how interactions between predictors (trait-trait matching) influence interaction probabilities. A suggested solution to this problem is the H-statistic (Friedman & Popescu, 2008), which uses partial dependencies to estimate feature-feature (trait-trait) interactions from fitted ML models. Assuming that networks emerge due to a few important trait-trait interactions (Eklöf et al., 2013), the H-statistic should be able to identify those from a fitted ML, but to our knowledge, the efficacy of this or similar techniques in inferring causal traits has not yet been demonstrated.

The purpose of this study is to (a) systematically assess the predictive performance of different ML models for the identification

of trait-matching in plant-pollinator networks and (b) to investigate if causal traits can be extracted from the fitted models with the H-statistics. We consider the most common ML models (k-nearest neighbor, random forest, boosted regression trees, deep neural networks, support vector machine, naïve Bayes, and convolutional neural networks), with standard generalized linear model (GLM) as a benchmark. We apply all models to simulated and empirical plant-pollinator networks to establish how networks properties influence their predictive performance, and to test if the causally responsible trait-trait interactions be inferred from the fitted models. We ask the following questions: (1) Which algorithms display the highest predictive performance for simulated plant-pollinator networks, varying network sizes, observation times, and species abundances? (2) Can we retrieve the true underlying trait-trait interaction structure (trait-matching) in the simulated plant-pollinator networks from the fitted ML models? We demonstrate the practical utility of the developed methods by predicting interactions in a global crop-pollinator interaction database, and by inferring the causal trait-trait interaction structure in a Costa Rican plant-hummingbird network.

2 | MATERIALS AND METHODS

2.1 | Machine learning models for predicting species interactions from trait-matching

Throughout this paper, we consider that empirical observations of species interactions may be available either as binary (presence-absence) or weighted (counts, intensity, interaction frequencies) data. The objective for the models is to predict those plant-pollinator interactions based on the species' traits. We selected seven classes of ML models, either because they were previously used for trait-matching, or because the general ML literature suggests that they should perform well for this task (Table 1). For more details on the respective models, see the column 'Design principle' and the cited literature in Table 1, and the Supporting Information S1 in the Appendix.

Each of the ML models in Table 1 includes model-specific tuning parameters (so-called hyperparameters, for instance to control the model's learning behaviour) that can be adjusted by hand or optimized. To factor out idiosyncrasies due to the choice of these parameters, we optimized each models' hyperparameters with a random search in 30 (20 for empirical data) steps (see also Bergstra & Bengio, 2012), with nested cross-validations to avoid overfitting (for details see Appendix S1). Furthermore, ML models often perform poorly with imbalanced classes (proportion of plant-pollinator interactions to no plant-pollinator interactions is extremely low/high, Krawczyk, 2016). To address this, we applied the standard approach of oversampling observed plant-pollinator interactions when their proportion (compared to plant-pollinator pairs without an interaction) was lower than 20%. To compare ML with traditional regression models, we fitted GLMs (binomial GLM for presence-absence

TABLE 1 Machine learning models and their usage for trait-matching

ML models	Type	Design principle	Applied with trait-matching
Random forest (RF)	Tree-based	Ensemble of a finite number of regression trees (see Breiman, 2001a).	Desjardins-Proulx et al. (2017), Ryo and Rillig (2017) and Hu, Li, Yang, Shen, and Yu (2016)
Boosted regression trees (BRT)	Tree-based	After fitting the first weak regression tree to the response, subsequent regression trees are fitted on the previous residuals (see Friedman, 2001).	He, Heidemeyer, Ban, Cherkasov, and Ester (2017) and Rayhan et al. (2017)
k-nearest-neighbor (kNN)	Distance-based	Given new point X, nearest k neighbors determine response.	Desjardins-Proulx et al. (2017) (as recommender system) and Rodgers, Zhu, Fourches, Rusyn, and Tropsha (2010)
Support vector machines (SVM)	Distance-based	In the n -dimensional feature space, a hyperplane to separate the classes is fitted (see Cristianini & Shawe-Taylor, 2000).	Fang et al. (2013)
Deep neural networks (DNN)	Neural networks	By learning to represent the input over several hidden layers, they are able to identify the patterns in the data for the task	Wen et al. (2017)
Convolutional neural networks (CNN)	Neural networks	Topological patterns in the input space (images, sequences) are preserved and processed by a number of kernels to extract features (see LeCun et al., 2015).	Liu, Tang, Chen, and Wang (2016)
Naive Bayes	Probabilistic classifier	The model learns the probability belonging to a class given a specific input vector.	Fang et al. (2013)
GLM	Parametric	A specific theory or model is fitted to the data	Pomeranz et al. (2019)

plant–pollinator interactions and Poisson GLM for plant–pollinator interaction counts), using all traits and all their possible two-way interactions as predictors and plant–pollinator interactions as response. Analyses were conducted with the statistical software R (R Core Team, 2019). The R package MLR (Bischof et al., 2016, version 2.12) was used for hyperparameter tuning and cross validation of our ML models.

2.2 | Simulating plant–pollinator interactions

To assess predictive and inferential performance of the models, we created a minimal simulation model for plant–pollinator interactions. The model assumes that the interaction probability between individuals of plants (group A) and pollinators (group B) arises from a Gaussian niche, matching the logarithmic ratio of the plant and pollinator traits. The logarithmic ratio ensures a symmetrically shaped interaction niche, see Figure S1. The niche value is multiplied by a weight to allow modifying the interaction strength independent of the niche width, and thus to control the overall trait-matching effect signal. Plant and pollinator abundances can either be drawn from an exponential distribution or a uniform distribution, to examine the effects of uneven abundance distributions and rare species. The expected number of observed interactions (i.e. their probability, $P_{\text{interaction}}$) was then calculated as the interaction probability times the interaction partner's abundances times the observation time. Observation times were adjusted to standardize the proportion of plant–pollinator interactions to no plant–pollinator interactions. To create the final interaction counts, we sampled from a Poisson distribution with $\lambda = P_{\text{interaction}}$. For presence–absence

species interactions (1 = interaction, 0 = no interaction), we set all counts >0 to 1.

Our default simulation scenario used 50*100 (plants*pollinators) for the simulated plant–pollinator networks. To remove obstacles such as class imbalance, we adjusted the observation duration to have a class proportion of $\approx 40\%$ for plant–pollinator interactions to no plant–pollinator interactions. The absence of interactions cannot be observed explicitly, and we speculate that most empirical datasets consist of observed species interactions (and possible non-interactions are inferred afterwards), thus we removed species with no observed plant–pollinator interaction.

2.3 | Comparison of predictive performance

2.3.1 | Predicting species interactions in simulated plant–pollinator networks

To assess predictive performance, we simulated reference data with six traits for each plant and pollinator. A possible issue with measuring predictive performance is that hidden correlations or structure in the data can lead to seemingly higher-than-random predictive performance even on random data (e.g. Roberts et al., 2017). To check that this is not the case, we created a first baseline scenario, consisting of equal species abundances and no trait–trait interactions (no trait-matching, the latter was achieved by setting the trait–trait interaction niche extremely wide). A second issue is that interactions of rare species will be less frequent than those of abundant species. As a result, models can achieve higher-than-random performance even without any trait–trait interactions when species abundances are

uneven. To ensure that the performance of our models exceeds these trivial performance levels, we created a second baseline scenario with exponential abundance distributions, but without trait-matching.

For the trait-matching scenario, we simulated networks with even abundance distributions and three trait-trait interactions (A1-B1, A2-B2, and A3-B3), each with a weight of 10. The scale parameter controlling the niche width was randomly sampled between 0.5 and 1.2 for simulating varying degrees of specialization in ecological networks (cf. Blüthgen et al., 2007). The even abundance distributions assumed here are unrealistic to some extent, but allow a better contrast between the models (because abundance effects are removed). In the case studies, we consider real abundance distributions. Other than that, the trait-matching scenario used the same parameter settings as the baseline scenarios (network size 50*100, $\approx 40\%$ class balance). To test additionally for the effect of network sizes and observation time, we also varied network size to 25*50 and 100*200 (plants*pollinators) setting and proportions plant-pollinator interactions to $\approx 10\%$, $\approx 25\%$, and $\approx 40\%$ one-factor-at-a-time from the base setting.

2.3.2 | Case study 1 - Predicting plant-pollinator interactions

Our first case study uses data from a global database of crop-pollinator interactions, assembled from 1607 published studies from 77 countries worldwide (details see Data availability statement). Of these, we selected only crops that appeared at least two times at different geographical locations, resulting in 80 crops with 256 entries for pollinators.

The database lists five pollinator traits: guild (bumbees, butterflies etc.), tongue length, body size, sociality (yes or no), and feeding behaviour (oligolectic, polylectic, or parasitic). In case of sexual dimorphism, the female measures were taken. Plants are described by 10 traits: type of plant (arboreous or herbaceous), flowering season, flower diameter, corolla shape (open, campanulate, or tubular), flower colour, nectar (yes or no), bloom system (type of pollination: insects, insects/wind, or insects/birds), self-pollination (yes or no), inflorescence (yes, solitary, solitary/pairs, solitary/clusters), and composite flowers (yes or no). Flower diameter, body size, and tongue length were provided as continuous traits (see Tables S1 and S2 for detailed information). When traits for a species were available from different sources, they were averaged. We filled missing trait values with a multiple imputation algorithm based on random forest (Stekhoven & Bühlmann, 2012). We used all available traits as predictors in our models.

2.3.3 | Measures of predictive performance

To assess the models' predictive performance on the simulated plant-pollinator networks, we used the area under the receiver operating characteristic curve (AUC, measures how well the models are able to distinguish between plant-pollinator interaction and no plant-pollinator interaction regardless of classification threshold)

and true skill statistic (TSS, which assess the predictive performance under a specific classification threshold, see Allouche, Tsoar, & Kadmon, 2006) for presence-absence, and spearman's correlation for interaction frequencies. Because the TSS for the empirical plant-pollinator database (case study 1) was similar, we additionally calculated classification threshold-dependent performance measurements: accuracy (proportion of correct predicted labels), sensitivity (recall), precision, and specificity (true negative rate). Classification thresholds were optimized with TSS. The interpretation of these statistics is as follows: if our focus is to detect plant-pollinator interactions, we want to achieve a high true positive rate (sensitivity) with an acceptable rate of false positives in the as true predicted labels (precision). Specificity estimates the rate of true negatives of all predicted negatives (no plant-pollinator interaction).

2.4 | Measuring accuracy for inferring causal traits

2.4.1 | H-statistics for inferring causal traits

We used the H-statistic (Friedman & Popescu, 2008) to infer causally responsible trait-trait interactions from the fitted ML models. The idea of this algorithm is similar to the principle of partial dependence plots. The H-statistic estimates the variance of the model's response caused by two traits separately (main effects) compared to the variance caused by the two traits combined partial function (trait-trait interaction). The H-statistic is scaled to [0,1]. A high value indicates that the interaction is the main reason for the variance in the response (probability for plant-pollinator interactions and counts for plant-pollinator interaction counts).

2.4.2 | Inferential performance in simulated plant-pollinator networks

To assess the accuracy with which causal trait combinations can be identified from the fitted models via Friedman's H-statistic, we considered 25*50 (plants*pollinators) species networks with one, two, three and four trait-trait interactions (always six traits for each group, but varying number of trait-trait interactions that correspond to trait-matching), and equal interaction strength. We replicated the simulated plant-pollinator networks eight to ten times. The reason for choosing a smaller network size than for the predictive analysis was the computational cost of the H-statistics, which made applying a large number of replicates to larger networks computationally prohibitive.

The resulting networks had a 'real' observed size of 800–1,200 data points (we removed two networks with four true trait-trait interactions, because they had under 20 remaining samples after removing species with no plant-pollinator interactions at all). We fitted RF, BRT, DNN and kNN (the top predictive models) on the 76 simulated networks, 38 for presence-absence plant-pollinator interactions and 38 for plant-pollinator interaction counts (with uniform species abundances). For each sample, we calculated the H-statistic for all possible trait-trait

interactions between the two species' groups. We calculated for each, the averaged true positive rate (true trait–trait interaction in found interactions with highest H-statistic) over the eight/ten repetitions. In a second step, based on our interim results (see results), we repeated the procedure with BRT and DNN for 50*100 (plants*pollinators) simulated networks (see Appendix S1 for details regarding model fitting).

For GLMs, we selected the n (n = number of true trait–trait interactions) predictors with lowest p -value to calculate the true positive rate.

2.4.3 | Case study 2—Inferring trait-matching in a plant-hummingbird network

As a case study for inferring causally responsible traits, we used a dataset of plant–hummingbird interactions from Costa Rica. Plant–hummingbird networks are characterized by particularly strong signals of trait-matching (Vizentin-Bugoni, Maruyama, & Sazima, 2014). Maglianesi et al. (2014) filmed and analyzed plant–hummingbird interactions at three elevations in Costa Rica (700 hr of observations on 50 m a.s.l.; 695 hr of observation on 1,000 m a.s.l.; 727 hr of observations on 2000 m a.s.l.). The resulting network consisted of 21*8, 24*8 and 20*9 plant and hummingbird species, respectively. To predict plant–pollinator interactions, we used bill length, bill curvature, body mass, wing length, and tail length of hummingbirds, and corolla length, corolla curvature, inner corolla diameter width, and external corolla diameter width of plants. Flower volume was calculated by corolla length and external diameter (Maglianesi et al., 2014). We used all available traits because the ML models should automatically learn trait–trait interactions.

We fitted the BRT with a Poisson maximum likelihood estimator and RF with a root mean squared error (RMSE) objective function

(we did not log count data). We optimized DNNs with Poisson and negative binomial likelihood loss functions. We trained models on each elevation and on combined elevations (e.g. Low, Mid, High, Low-Mid-High, for details see Appendix S1). We calculated for the Low, Mid, High and Low-Mid-High models interaction strengths (H-statistics) for all possible trait–trait interactions (with trait–trait interactions within hummingbird/plant group). We checked the eight trait–trait interactions with highest interaction strengths for their biological plausibility by reviewing relevant literature.

3 | RESULTS

3.1 | Predictive performance

3.1.1 | Predictive performance in simulated plant–pollinator networks

In the first baseline scenario (no trait-matching and equal species abundances), all models performed as expected for random plant–pollinator interactions, with AUC ≈ 0.5 , TSS ≈ 0 , and Spearman Rho factor ≈ 0 for both for presence–absence data and count data (Figure 2), indicating that our cross-validation setup is accurate. In the second baseline scenario (no trait-matching and networks with uneven species abundances), models achieved a TSS between 0.0–0.38, AUC between 0.64–0.76, and Spearman Rho factor of between 0.26–0.5 (Figure 2). The latter provides an indication, also with respect to existing literature, of what performance values can be achieved through imbalance of the data alone, even if there is no trait-matching.

For simulated networks with strong trait-matching and even abundances, all ML models except SVMs achieved higher TSS,

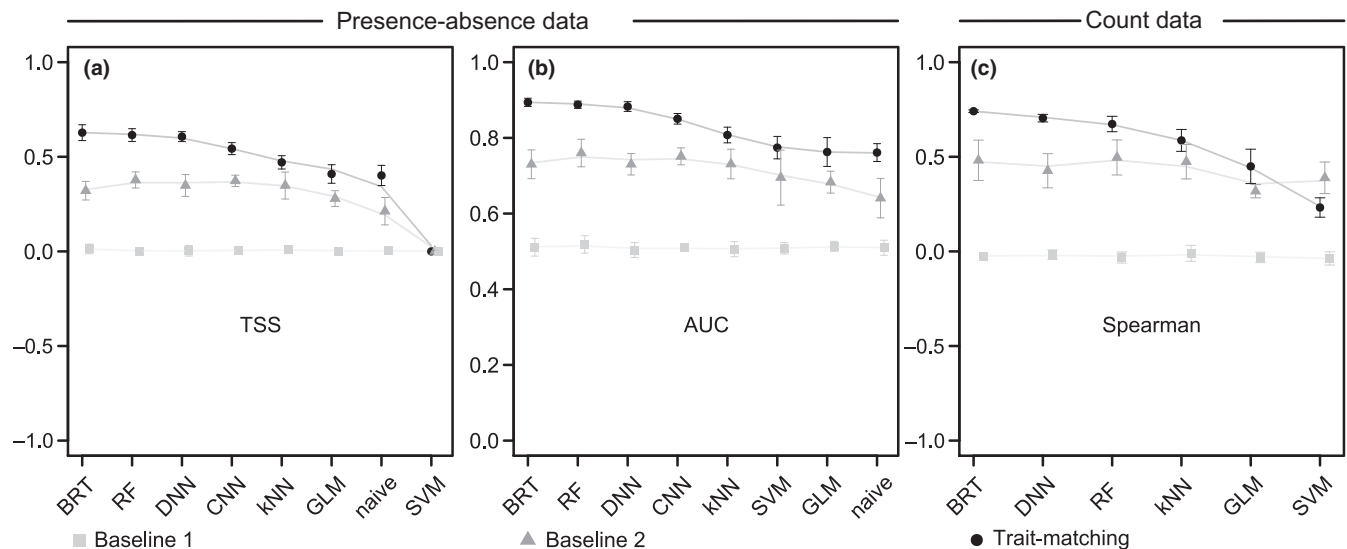


FIGURE 2 Predictive performance of kNN, CNN, DNN, RF, BRT, naive Bayes, GLM and SVM with simulated plant–pollinator networks (50 plants * 100 pollinators) for baseline scenarios with random interactions and even (baseline 1, squares) or uneven species abundances (baseline2, triangles, respectively), and trait-based interactions with even species abundances (circles). Predictive performance was measured by TSS (a) and AUC (b) for binary interaction data; and Spearman Rho factor (c) for interaction counts. Lowest predictive performance corresponds to zero for TSS, AUC, and Spearman Rho factor

AUC, and Spearman Rho than for the baseline scenarios (Figure 2). Moreover, DNN, RF, and BRT achieved a higher TSS (0.61–0.63) than GLMs (0.41). SVM, naïve Bayes, kNN were around GLM's performance or lower (Figure 2).

While all models improved their predictive performance with increasing network sizes with count data (Figure S2c), only DNN, RF, and BRT improved their performance with increasing network sizes with presence–absence plant–pollinator interactions (Figure S2a,b). Prolonging the observation time (i.e. creating more plant–pollinator interactions and thus reducing data imbalance) generally increased the models' performances (Figure S2d–f).

3.1.2 | Predicting species interactions in a global crop–pollination database

After fitting the models to real data from a global crop–pollination database, we calculated AUC, TSS and additional performance

measures (Figure 3, Table S4) on the left-out samples. kNN achieved the highest TSS (0.36), RF achieved the highest AUC (0.73), and naïve Bayes achieved highest TPR, followed by CNN. Overall, RF achieved the overall best predictive performance with highest AUC and second highest TSS (Figure 3, Table S4).

3.2 | Inference of causal trait–trait interactions

3.2.1 | Inference of causal trait–trait interactions in simulated networks

In the second analysis step, we tested the ability of the H-statistics to infer the trait–trait interactions causally responsible for plant–pollinator interactions from the fitted models. In simulated networks, RF and BRT achieved highest true positive rates (Figure 4). For presence–absence plant–pollinator interactions, RF, DNN and BRT exceeded GLM performance with an averaged true positive rate

FIGURE 3 Predictive performance of different ML methods (naïve Bayes, SVM, BRT, kNN, DNN, CNN, RF) and GLM in a global database of plant–pollinator interactions. Dotted lines depict training and solid lines validation performances. Models were sorted from left to right with increasing true skill statistic. The central figure compares directly the models' performances. Sen = Sensitivity (recall, true positive rate); Spec = Specificity (true negative rate); Prec = Precision; Acc = Accuracy; AUC = Area under the receiver operating characteristic curve (AUC); TSS in % = True skill statistic rescaled to 0–1

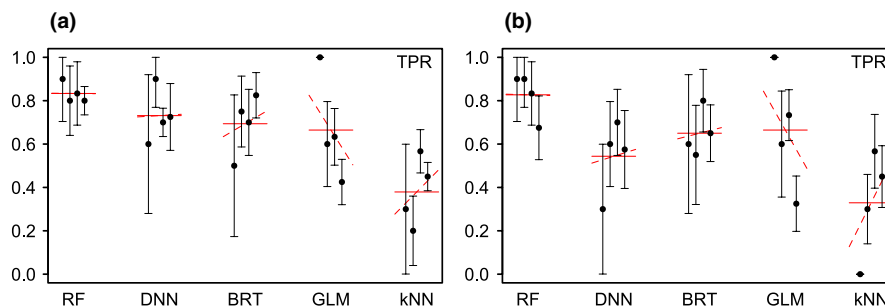
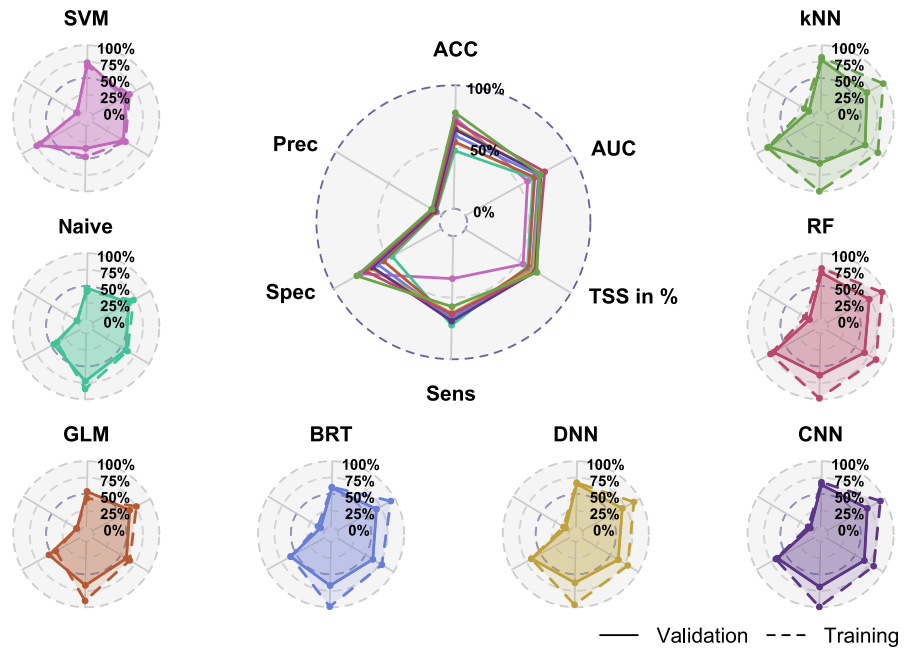


FIGURE 4 Comparison of the top predictive models' (RF, DNN, BRT, kNN, and GLM) abilities to infer the causal trait–trait interaction structure in simulated networks, using presence–absence data (a) and count data (b). The four values associated with each algorithm represent the mean true positive rate (TPR, dot) and its standard error (error bar) for the four interaction scenarios (one to four true trait–trait interactions in the simulations). The values were calculated based on 8–10 replicate simulations each. Solid red lines display the mean TPR across all four scenarios, dotted red lines show a linear regression estimate of TPR against the number of true trait–trait interactions

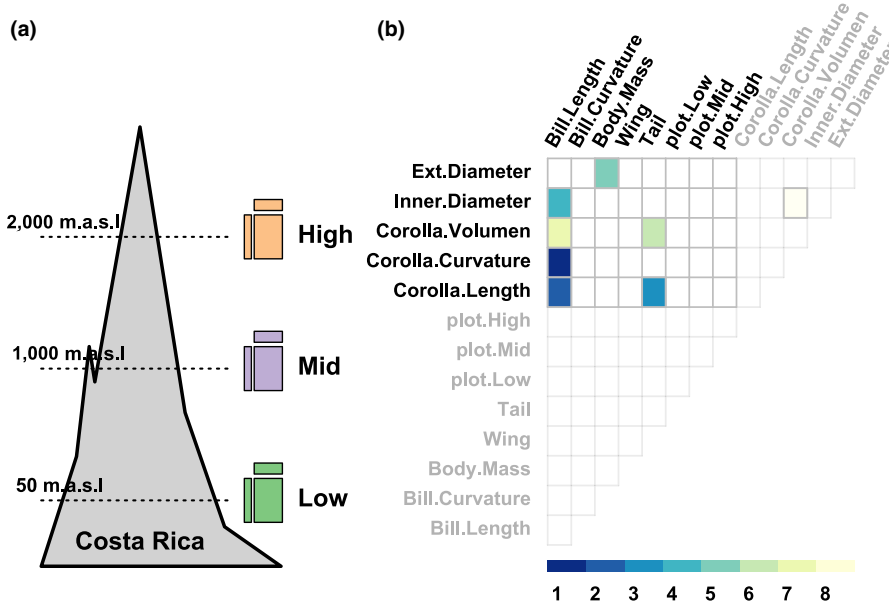


FIGURE 5 (a) Elevation profile for the three plant-hummingbird networks in Costa Rica (details see Maglianesi et al., 2014). (b) The eight strongest trait-trait interactions (blue–yellow gradient) inferred with the H-statistic from RF models fitted to the combined plant-hummingbird network (colors code the ranking of strengths). Corolla length–bill length and corolla curvature–bill length had the highest interaction strength

of 70% to 80% over one to four true trait–trait matches (Figure 4a, the models were able to identify most of the true trait–trait interactions). For plant–pollinator interaction count data, only RF achieved a higher true positive rate than GLM (Figure 4b). However, it should be noted that the good GLM performance hinged on simulations with 1–3 trait–trait interactions and decreased most strongly of all algorithms with the number of trait–trait interactions (Figure 4).

When increasing network size (from 25*50 to 50*100), DNN and BRT improved their overall performance to 70%–95% and 87%–98% for presence–absence networks (Figure S3a), but showed a lower TPR for count data (Figure S3b).

3.2.2 | Inference of causal trait–trait interactions in a plant–hummingbird network

In a second case study, we computed interaction strength (H-statistic) for all possible trait–trait interactions in plant–hummingbird networks (Figure S5). The seven trait–trait interactions with highest interaction strength were identified by RF (Figure 5b). These interactions also achieved highest predictive performance (Figure S4). The four trait–trait interactions with highest interaction strength identified by BRT were in accordance with the ones that RF identified (Figure S5).

RF and BRT identified corolla length–bill length, corolla curvature–bill length, inner diameter–bill length, and external diameter–body mass as the most important trait–trait interactions (Figure 5b, Figure S5). The models identified varying trait–trait interactions for networks at different elevations, but corolla and bill associations tended to be most important across elevations (Figure S5).

4 | DISCUSSION

We assessed the ability of seven ML models, plus GLM as a reference, to predict plant–pollinator interactions based on their traits.

In a second step, we tested whether it is possible to identify the causally responsible trait–trait interaction structure (trait-matching) from the fitted models. Our main results are that the best ML models (RF, BRT, and DNN) outperform GLMs to a substantial degree in predicting plant–pollinator interactions from their traits, and that it is possible to identify the trait–trait interactions causally responsible for plant–pollinator interactions from the fitted models with satisfying accuracy. The best ML models outperformed the simpler GLMs particularly for more complex trait–trait interaction structures, for which GLM performance dropped sharply.

4.1 | Comparison of performance in predicting species interactions

In our analysis of predictive performance, we found that ML models such as RF, BRT and DNNs exceeded GLM performance for predicting plant–pollinator interactions from trait-matching data. They also worked surprisingly well with small network sizes (25*50, Figure S2a), such that performance did not increase substantially for larger networks (50*100, 100*200, Figure S2a–c).

An important point, also for comparing our performance indicators to the literature, is that all algorithms can achieve higher than naïve random performance (e.g. AUC of 0.5) when species distributions are uneven, even when plant–pollinator interactions are not tied to traits (Figure 2). These results, in line with earlier findings (e.g. Aderhold, Husmeier, Lennon, Beale, & Smith, 2012; Canard et al., 2014), highlight the importance of considering abundance when analyzing network structures: frequent species tend to have more observed interactions, and this effect might interfere with the trait-matching signal (e.g. Olito & Fox, 2015). While the trait-matching effect may influence which plant–pollinator interaction is feasible, the species abundance effect determine the actual observed plant–pollinator interactions. Without adjusting observed

plant–pollinator interactions for species abundances, it is difficult to separate the contributions of abundance and trait-matching to predictive performance (Olito & Fox, 2015).

Observation time and type are further critical factors in ecological network analysis. Short observation times often lead to sparse networks with many unobserved plant–pollinator interactions, potentially creating biases in the analysis. Moreover, few plant–pollinator interactions result in data with imbalanced class distributions, presenting challenges for many ML methods (Krawczyk, 2016), which is also reflected in our results (Figure S2d–f). On the other hand, too long observation times could also negatively affect predictive performance, in particular when using binary links. The reason is that, given sufficient time, even weak links will be included in the network, potentially reducing the models' ability to identify the essential traits. Count data are more robust to these problems, and as our approach is equally applicable with count data, this data type seems generally preferable (see also Dormann & Strauss, 2014).

While the ML models detected the important trait–trait interactions automatically, GLMs were pre-specified with all possible two-way trait–trait interactions. To check that the resulting high complexity did not disadvantage them unduly, we additionally confirmed that AIC selection on their interaction structure did not increase their predictive performance. We therefore believe that their lower performance is either explained by the fact that GLMs are not flexible enough to capture the complex form of the trait-matching structures (see also Mayfield & Stouffer, 2017), or that ML methods are more successful than AIC variable selection in addressing overfitting induced by the high combinatorial number of possible trait–trait interactions. These results mirror findings in the literature: while a few studies showed that GLM can predict species interactions based on trait-matching (e.g. Gravel et al., 2013), most studies struggled in predicting species interactions with the trait-matching signal alone (e.g. Brousseau et al., 2018; Pearse & Altermatt, 2013; Pomeranz et al., 2019). We speculate based on our results that previous studies based on GLMs may have underestimated the importance of trait-matching considerably, unless a very small number of trait–trait interactions (1–2) is dominantly responsible for the structure of the networks.

Previous studies often showed improved performance in predicting species interactions by using phylogenetic predictors, serving as proxies for unobserved traits (see Morales-Castilla, Matias, Gravel, & Araújo, 2015). The drawback, however, is that such phylogenetic proxies can be hard to interpret in the context of specific ecological hypothesis of why species interact (see Díaz et al., 2013). For example, a phylogenetic signal could arise both as a result of trait-matching (because traits tend to be phylogenetically conserved), or as a result other genetically coded preferences for particular interactions that are not accessible as traits. Based on our results, we expect that the relative importance of phylogenetic proxies will decrease when using appropriate ML models, which could help to better explore to what extent species interactions are determined by measurable functional traits.

We found that the models' predictive performance was lower for the empirical plant–pollinator database than for the simulated

networks. There are several plausible reasons for this. Firstly, trait-matching rules may change over scales (Poisot et al., 2015). As the database consists of globally observed plant–pollinator interactions, this may complicate the identification of a common trait-matching signal. Secondly, the high share of discrete predictors and high-class imbalance is likely to negatively affect the predictive performance. Despite these obstacles, kNN, RF, and CNN achieved >0.3 TSS, and CNN and RF >70% AUC (Figure 2, Table S4), much higher than null expectation, and consistent with results from the simulated networks. While it may be possible to improve GLM performance by manual selection of predictors, we also find that the case study highlights that algorithms such as RF and BRT are more parsimonious and robust in their use than a GLM which further suffered convergence problems.

4.2 | Causal inference of trait-matching

To infer trait–trait interactions causally responsible for species interactions, we used the H-statistics. We found that this method, coupled with RF, DNN and BRT, could identify around 90% of the true trait–trait interactions in simulated plant–pollinator networks (Figure 4, Figure S5). Increasing the network size improved the detection accuracy of true trait-matches for BRT and DNN (Figure S3). When increasing the number of trait–trait interactions, the approach outperformed GLMs (Figure 2).

Our results demonstrate that identifying trait-matching from fitted models with the H-statistic works, but it also comes with drawbacks. The H-statistic depends on partial dependencies (Friedman & Popescu, 2008) and is therefore sensitive to collinearity (see Apley, 2016). Other alternative approaches (e.g. Apley, 2016) might overcome this limitation. Moreover, the H-statistic is extremely computationally expensive, which is the reason why we tested it only on small network sizes (25*50 species). Neither of these issues, however, would change the balance in favour of GLMs, which are prone to collinearity issues, too. To make sure that GLMs are not unjustly disfavoured, we additionally tested if AIC selection or choosing causal traits based on regression estimates instead of p-values would change the results, but neither improved inferential performance. In summary, we think that ML models are the better choice, not only for predictions, but also for causal inference in this setting. Future research should, however, focus on testing and advancing methods for the causal analysis of fitted models.

Analyzing plant–hummingbird networks with RF, we highlighted the seven trait–trait interactions with highest interaction strength (Figure 5b). The inferred trait–trait interactions are highly plausible for the following reasons: (a) RF showed high accuracy with low consistent errors in the simulated networks (Figure 4). (b) The identified trait–trait interactions are ecologically plausible (Figure 5b): Trait matches with highest interaction strength (corolla length–bill length and corolla curvature–bill length) are in line with previous findings that emphasize their importance in plant–hummingbird networks (Temeles, Koulouris, Sander, &

Kress, 2009; Maglianesi et al., 2014; Vizentin-Bugoni, Maruyama, & Sazima, 2014; Weinstein & Graham, 2017). Collinearity of traits likely explains other matches. For instance, body mass is positively correlated with tail length, explaining why corolla volume was associated with tail length. These results further support the view that it is possible to infer trait-matching with ML in ecologically realistic settings, without a priori assumptions.

Estimated trait-trait interactions in the plant-hummingbird networks differed for the three elevations, but the match of corolla length–bill length was generally most important (Figure S5). Maglianesi et al. (2014) and Maglianesi, Blüthgen, Böhning-Gaese, and Schleuning (2015) reported similarly varying trait-trait interactions in plant-hummingbird networks across elevations, consistent with our results. While interactions in ecological networks vary over scales (Poisot et al., 2015), a common backbone is assumed (Mora et al., 2018). With corolla length–bill length, identified by RF and BRT with highest interaction strength (Figure 5b, Figure S5), we speculate that we identified with ML the central trait-matching phenomenon in plant-hummingbird networks.

5 | CONCLUSIONS

In conclusion, our study demonstrates that RF, BRT, and DNN exceeded GLM performance in predicting plant–pollinator interactions from trait information. ML models could also identify causally responsible trait-trait interactions with a higher accuracy than GLMs. The ability to automatically extract species interactions from observed networks and traits, and causally interpreting the underlying trait-trait interactions, makes our approach, which we provide in an R package, a powerful new tool for ecologists.

While we considered only plant–pollinator networks in this study, our method could be applied to other types of species interaction networks such as any mutualistic and antagonistic interactions in complex food webs (this is also supported by Desjardins-Proulx et al., 2017). In either of these ecological network types, there are ample opportunities for further analyses, for example how species interactions will change under global change or how species interactions will rewire in novel communities with reshuffled species and trait composition (Bailes et al., 2015; see Kissling & Schleuning, 2015). By identifying crucial rules of trait-matching between species, our approach can give insights into how biotic interactions shape community assembly and also contribute to the identification of Essential Biodiversity Variables in the context of global change (Kissling et al., 2018).

ACKNOWLEDGEMENTS

Maria A. Maglianesi recorded interaction and trait data of plants and hummingbirds in Costa Rica. We would like to thank Johannes Oberpriller and Lukas Heiland, as well as Roozbeh Valavi and

an anonymous reviewer for their valuable comments and suggestions. V.B. acknowledges funding for the assembly of the global plant–pollinator database (case study 1) by Bayer CropScience.

AUTHORS' CONTRIBUTIONS

M.P. and F.H. conceived the ideas and designed methodology; V.B., A.-M.K. and M.S. provided data; M.P. performed the analyses. M.P. and F.H. wrote the first draft of the manuscript. All authors contributed critically to the completion and revision of the manuscript.

DATA AVAILABILITY STATEMENT

The plant-hummingbird data associated with this study is available at <https://doi.org/10.6084/m9.figshare.3560895.v1> (Maglianesi, Blüthgen, Böhning-Gaese, & Schleuning, 2016). The global plant–pollinator database used with this study is available at <https://doi.org/10.6084/m9.figshare.9980471.v1> (Boreux & Klein, 2019). The analysis and the Trait-matching R package is available at <https://doi.org/10.5281/zenodo.3522854> (<https://github.com/TheoreticalEcology/Pichler-et-al-2019>) (Pichler & Hartig, 2019).

ORCID

Maximilian Pichler  <https://orcid.org/0000-0003-2252-8327>

Alexandra-Maria Klein  <https://orcid.org/0000-0003-2139-8575>

Matthias Schleuning  <https://orcid.org/0000-0001-9426-045X>

Florian Hartig  <https://orcid.org/0000-0002-6255-9059>

REFERENCES

- Aderhold, A., Husmeier, D., Lennon, J. J., Beale, C. M., & Smith, V. A. (2012). Hierarchical Bayesian models in ecology: Reconstructing species interaction networks from non-homogeneous species abundance data. *Ecological Informatics*, 11, 55–64. <https://doi.org/10.1016/j.ecoinf.2012.05.002>
- Allouche, O., Tsoar, A., & Kadmon, R. (2006). Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, 43, 1223–1232. <https://doi.org/10.1111/j.1365-2664.2006.01214.x>
- Apley, D. W. (2016). Visualizing the effects of predictor variables in black box supervised learning models. *arXiv:1612.08468 [stat]*.
- Bailes, E. J., Ollerton, J., Patrick, J. G., & Glover, B. J. (2015). How can an understanding of plant–pollinator interactions contribute to global food security? *Current Opinion in Plant Biology*, 26, 72–79. <https://doi.org/10.1016/j.pbi.2015.06.002>
- Bartomeus, I., Gravel, D., Tylianakis, J. M., Aizen, M. A., Dickie, I. A., & Bernard-Verdier, M. (2016). A common framework for identifying linkage rules across different types of interactions. *Functional Ecology*, 30, 1894–1903. <https://doi.org/10.1111/1365-2435.12666>
- Bender, I. M. A., Kissling, W. D., Böhning-Gaese, K., Hensen, I., Kühn, I., Wiegand, T., ... Schleuning, M. (2017). Functionally specialised birds respond flexibly to seasonal changes in fruit availability. *Journal of Animal Ecology*, 86, 800–811. <https://doi.org/10.1111/1365-2656.12683>
- Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of Machine Learning Research*, 13, 281–305.
- Berlow, E. L., Dunne, J. A., Martinez, N. D., Stark, P. B., Williams, R. J., & Brose, U. (2009). Simple prediction of interaction strengths in

- complex food webs. *Proceedings of the National Academy of Sciences*, 106, 187–191. <https://doi.org/10.1073/pnas.0806823106>
- Bischi, B., Lang, M., Kotthoff, L., Schiffner, J., Richter, J., Studerus, E., ... Jones, Z. M. (2016). MLR: Machine learning in R. *Journal of Machine Learning Research*, 17, 1–5.
- Blüthgen, N., Menzel, F., Hovestadt, T., Fiala, B., & Blüthgen, N. (2007). Specialization, constraints, and conflicting interests in mutualistic networks. *Current Biology*, 17, 341–346. <https://doi.org/10.1016/j.cub.2006.12.039>
- Boreux, V., & Klein, A.-M. (2019). Global pollinator database (Version 1). figshare. <http://doi.org/10.6084/m9.figshare.9980471.v1>
- Breiman, L. (2001a). Random forests. *Machine Learning*, 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L. (2001b). Statistical Modeling: The Two Cultures (with comments and a rejoinder by the author). *Statistical Science*, 16, 199–231. <https://doi.org/10.1214/ss/1009213726>
- 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://doi.org/10.1111/1365-2435.12943>
- Canard, E. F., Mouquet, N., Mouillot, D., Stanko, M., Miklisova, D., & Gravel, D. (2014). Empirical evaluation of neutral interactions in host–parasite networks. *The American Naturalist*, 183, 468–479. <https://doi.org/10.1086/675363>
- Cheng, F., & Zhao, Z. (2014). Machine learning-based prediction of drug–drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties. *Journal of the American Medical Informatics Association*, 21, e278–e286. <https://doi.org/10.1136/amiajnl-2013-002512>
- Cristianini, N., & Shawe-Taylor, J. (2000). *An introduction to support vector machines*. Cambridge, UK: Cambridge University Press.
- Danieli-Silva, A., de Souza, J. M. T., Donatti, A. J., Campos, R. P., Vicente-Silva, J., Freitas, L., & Varassin, I. G. (2012). Do pollination syndromes cause modularity and predict interactions in a pollination network in tropical high-altitude grasslands? *Oikos*, 121, 35–43. <https://doi.org/10.1111/j.1600-0706.2011.19089.x>
- Dehling, D. M., Jordano, P., Schaefer, H. M., Böhning-Gaese, K., & Schleuning, M. (2016). Morphology predicts species' functional roles and their degree of specialization in plant–frugivore interactions. *Proceedings of the Royal Society B: Biological Sciences*, 283, 20152444. <https://doi.org/10.1098/rspb.2015.2444>
- Dehling, D. M., Töpfer, T., Schaefer, H. M., Jordano, P., Böhning-Gaese, K., & Schleuning, M. (2014). Functional relationships beyond species richness patterns: Trait-matching in plant–bird mutualisms across scales. *Global Ecology and Biogeography*, 23, 1085–1093. <https://doi.org/10.1111/geb.12193>
- Desjardins-Proulx, P., Laigle, I., Poisot, T., & Gravel, D. (2017). Ecological interactions and the Netflix problem. *PeerJ*, 5, e3644. <https://doi.org/10.7717/peerj.3644>
- Díaz, S., Purvis, A., Cornelissen, J. H. C., Mace, G. M., Donoghue, M. J., Ewers, R. M., ... Pearce, W. D. (2013). Functional traits, the phylogeny of function, and ecosystem service vulnerability. *Ecology and Evolution*, 3, 2958–2975. <https://doi.org/10.1002/ece3.601>
- Dormann, C. F., & Strauss, R. (2014). A method for detecting modules in quantitative bipartite networks. *Methods in Ecology and Evolution*, 5, 90–98. <https://doi.org/10.1111/2041-210X.12139>
- Eklöf, A., Jacob, U., Kopp, J., Bosch, J., Castro-Urgal, R., Chacoff, N. P., ... Allesina, S. (2013). The dimensionality of ecological networks. *Ecology Letters*, 16, 577–583. <https://doi.org/10.1111/ele.12081>
- Faegri, K., & van der Pijl, L. (1979). *The principles of pollination ecology*. ID - 19790561579. Oxford: Pergamon Press.
- Fang, J., Yang, R., Gao, L. I., Zhou, D., Yang, S., Liu, A.-L., & Du, G.-H. (2013). Predictions of BuChE inhibitors using support vector machine and naive bayesian classification techniques in drug discovery. *Journal of Chemical Information and Modeling*, 53, 3009–3020. <https://doi.org/10.1021/ci400331p>
- Fenster, C. B., Armbruster, W. S., Wilson, P., Dudash, M. R., & Thomson, J. D. (2004). Pollination syndromes and floral specialization. *Annual Review of Ecology and Systematics*, 35, 375–403. <https://doi.org/10.1146/annurev.ecolsys.34.011802.132347>
- Fenster, C. B., Reynolds, R. J., Williams, C. W., Makowsky, R., & Dudash, M. R. (2015). Quantifying hummingbird preference for floral trait combinations: The role of selection on trait interactions in the evolution of pollination syndromes. *Evolution*, 69, 1113–1127. <https://doi.org/10.1111/evo.12639>
- Fisher, A., Rudin, C., & Dominici, F. (2018). All models are wrong but many are useful: Variable importance for black-box, proprietary, or misspecified prediction models, using model class reliance. *ArXiv e-prints*.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29, 1189–1232. <https://doi.org/10.1214/aos/1013203451>
- Friedman, J. H., & Popescu, B. E. (2008). Predictive learning via rule ensembles. *The Annals of Applied Statistics*, 2, 916–954. <https://doi.org/10.1214/07-AOAS148>
- Galiana, N., Lurgi, M., Claramunt-López, B., Fortin, M.-J., Leroux, S., Cazelles, K., ... Montoya, J. M. (2018). The spatial scaling of species interaction networks. *Nature Ecology & Evolution*, 2, 782. <https://doi.org/10.1038/s41559-018-0517-3>
- Garibaldi, L. A., Bartomeus, I., Bommarco, R., Klein, A. M., Cunningham, S. A., Aizen, M. A., ... Woyciechowski, M. (2015). Trait-matching of flower visitors and crops predicts fruit set better than trait diversity. *Journal of Applied Ecology*, 52, 1436–1444. <https://doi.org/10.1111/1365-2664.12530>
- González, M. A. M., Dalsgaard, B., & Olesen, J. M. (2010). Centrality measures and the importance of generalist species in pollination networks. *Ecological Complexity*, 7, 36–43. <https://doi.org/10.1016/j.ecocom.2009.03.008>
- Gravel, D., Poisot, T., Albouy, C., Velez, L., & Mouillot, D. (2013). Inferring food web structure from predator–prey body size relationships. *Methods in Ecology and Evolution*, 4, 1083–1090. <https://doi.org/10.1111/2041-210X.12103>
- Gray, P. C., Fleishman, A. B., Klein, D. J., McKown, M. W., Bézy, V. S., Lohmann, K. J., & Johnston, D. W. (2019). A convolutional neural network for detecting sea turtles in drone imagery. *Methods in Ecology and Evolution*, 10, 345–355. <https://doi.org/10.1111/2041-210X.13132>
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. *ACM Computing Surveys*, 51, 93:1–93:42. <https://doi.org/10.1145/3236009>
- He, T., Heidemeyer, M., Ban, F., Cherkasov, A., & Ester, M. (2017). SimBoost: A read-across approach for predicting drug–target binding affinities using gradient boosting machines. *Journal of Cheminformatics*, 9, 24. <https://doi.org/10.1186/s13321-017-0209-z>
- Hu, J., Li, Y., Yang, J.-Y., Shen, H.-B., & Yu, D.-J. (2016). GPCR–drug interactions prediction using random forest with drug–association-matrix-based post-processing procedure. *Computational Biology and Chemistry*, 60, 59–71. <https://doi.org/10.1016/j.compbiolchem.2015.11.007>
- Janzen, D. H. (1985). On ecological fitting. *Oikos*, 45(3), 308. <https://doi.org/10.2307/3565565>
- Jordano, P., Bascompte, J., & Olesen, J. M. (2003). Invariant properties in coevolutionary networks of plant–animal interactions. *Ecology Letters*, 6, 69–81. <https://doi.org/10.1046/j.1461-0248.2003.00403.x>
- Kissling, W. D., & Schleuning, M. (2015). Multispecies interactions across trophic levels at macroscales: Retrospective and future directions. *Ecography*, 38, 346–357. <https://doi.org/10.1111/ecog.00819>

- Kissling, W. D., Walls, R., Bowser, A., Jones, M. O., Kattge, J., Agosti, D., ... Guralnick, R. P. (2018). Towards global data products of essential biodiversity variables on species traits. *Nature Ecology & Evolution*, 2, 1531. <https://doi.org/10.1038/s41559-018-0667-3>
- Krawczyk, B. (2016). Learning from imbalanced data: Open challenges and future directions. *Progress in Artificial Intelligence*, 5, 221–232. <https://doi.org/10.1007/s13748-016-0094-0>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521, 436–444. <https://doi.org/10.1038/nature14539>
- Liu, S., Tang, B., Chen, Q., & Wang, X. (2016). Drug-drug interaction extraction via convolutional neural networks. *Computational and Mathematical Methods in Medicine*. Retrieved from <https://www.hindawi.com/journals/cmmm/2016/6918381/abs/>. Last accessed 29 May 2019.
- Maglianesi, M. A., Blüthgen, N., Böhning-Gaese, K., & Schleuning, M. (2014). Morphological traits determine specialization and resource use in plant-hummingbird networks in the neotropics. *Ecology*, 95, 3325–3334. <https://doi.org/10.1890/13-2261.1>
- Maglianesi, M. A., Blüthgen, N., Böhning-Gaese, K., & Schleuning, M. (2015). Functional structure and specialization in three tropical plant-hummingbird interaction networks across an elevational gradient in Costa Rica. *Ecography*, 38, 1119–1128. <https://doi.org/10.1111/ecog.01538>
- Maglianesi, M. A., Blüthgen, N., Böhning-Gaese, K., & Schleuning, M. (2016). Appendix A. Tables showing mean values of morphological traits of plant species and of hummingbird species, hummingbird abundance and the number of observed legitimate visits of each hummingbird species to plant individuals. and a list of plant species and the number of legitimate visits by hummingbirds to each plant species. (Version 1). Wiley. <https://doi.org/10.6084/m9.figshare.3560895.v1>
- Mayfield, M. M., & Stouffer, D. B. (2017). Higher-order interactions capture unexplained complexity in diverse communities. *Nature Ecology & Evolution*, 1, 0062. <https://doi.org/10.1038/s41559-016-0062>
- Menden, M. P., Iorio, F., Garnett, M., McDermott, U., Benes, C. H., Ballester, P. J., & Saez-Rodriguez, J. (2013). Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties. *PLoS ONE*, 8, e61318. <https://doi.org/10.1371/journal.pone.0061318>
- Mora, B. B., Gravel, D., Gilarranz, L. J., Poisot, T., & Stouffer, D. B. (2018). Identifying a common backbone of interactions underlying food webs from different ecosystems. *Nature Communications*, 9, 2603.
- Morales-Castilla, I., Matias, M. G., Gravel, D., & Araújo, M. B. (2015). Inferring biotic interactions from proxies. *Trends in Ecology & Evolution*, 30, 347–356. <https://doi.org/10.1016/j.tree.2015.03.014>
- Murúa, M., & Espíndola, A. (2015). Pollination syndromes in a specialised plant-pollinator interaction: Does floral morphology predict pollinators in Calceolaria? *Plant Biology*, 17, 551–557. <https://doi.org/10.1111/plb.12225>
- Olhede, S. C., & Wolfe, P. J. (2018). The growing ubiquity of algorithms in society: Implications, impacts and innovations. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376, 20170364.
- 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://doi.org/10.1111/oik.01439>
- Ollerton, J., Alarcón, R., Waser, N. M., Price, M. V., Watts, S., Cranmer, L., ... Rotenberry, J. (2009). A global test of the pollination syndrome hypothesis. *Annals of Botany*, 103, 1471–1480. <https://doi.org/10.1093/aob/mcp031>
- Pearse, I. S., & Altermatt, F. (2013). Predicting novel trophic interactions in a non-native world. *Ecology Letters*, 16, 1088–1094. <https://doi.org/10.1111/ele.12143>
- Pichler, M., & Hartig, F. (2019). Theoretical Ecology/Pichler-et-al-2019 v1.0 (Version v1.0). *Zenodo*. <http://doi.org/10.5281/zenodo.3522854>
- Poisot, T., Stouffer, D. B., & Gravel, D. (2015). Beyond species: Why ecological interaction networks vary through space and time. *Oikos*, 124, 243–251. <https://doi.org/10.1111/oik.01719>
- Pomeranz, J. P. F., Thompson, R. M., Poisot, T., & Harding, J. S. (2019). Inferring predator-prey interactions in food webs. *Methods in Ecology and Evolution*, 10, 356–367.
- Potts, S. G., Imperatriz-Fonseca, V., Ngo, H. T., Aizen, M. A., Biesmeijer, J. C., Breeze, T. D., ... Vanbergen, A. J. (2016). Safeguarding pollinators and their values to human well-being. *Nature*, 540, 220–229. <https://doi.org/10.1038/nature20588>
- R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rayhan, F., Ahmed, S., Shatabda, S., Farid, D. M., Mousavian, Z., Dehngangi, A., & Rahman, M. S. (2017). iDTI-ESBoost: Identification of drug target interaction using evolutionary and structural features with boosting. *Scientific Reports*, 7, 17731. <https://doi.org/10.1038/s41598-017-18025-2>
- Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?": Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*. ACM, New York, NY, USA, pp. 1135–1144.
- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Aroita, G., ... Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40, 913–929. <https://doi.org/10.1111/ecog.02881>
- Rodgers, A. D., Zhu, H., Fourches, D., Rusyn, I., & Tropsha, A. (2010). Modeling liver-related adverse effects of drugs using k nearest neighbor quantitative structure-activity relationship method. *Chemical Research in Toxicology*, 23, 724–732. <https://doi.org/10.1021/tx900451r>
- Rosas-Guerrero, V., Aguilar, R., Martín-Rodríguez, S., Ashworth, L., Lopezaraiza-Mikel, M., Bastida, J. M., & Quesada, M. (2014). A quantitative review of pollination syndromes: Do floral traits predict effective pollinators? *Ecology Letters*, 17, 388–400. <https://doi.org/10.1111/ele.12224>
- Ryo, M., & Rillig, M. C. (2017). Statistically reinforced machine learning for nonlinear patterns and variable interactions. *Ecosphere*, 8(11), e01976. <https://doi.org/10.1002/ecs2.1976>
- Schleuning, M., Fründ, J., & García, D. (2015). Predicting ecosystem functions from biodiversity and mutualistic networks: An extension of trait-based concepts to plant-animal interactions. *Ecography*, 38, 380–392. <https://doi.org/10.1111/ecog.00983>
- 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. <https://doi.org/10.1007/s00442-006-0585-y>
- Stekhoven, D. J., & Bühlmann, P. (2012). MissForest – Non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28, 112–118. <https://doi.org/10.1093/bioinformatics/btr597>
- Tabak, M. A., Norouzzadeh, M. S., Wolfson, D. W., Sweeney, S. J., Vercauteren, K. C., Snow, N. P., ... Miller, R. S. (2019). Machine learning to classify animal species in camera trap images: Applications in ecology. *Methods in Ecology and Evolution*, 10, 585–590. <https://doi.org/10.1111/2041-210X.13120>
- Tari, L., Anwar, S., Liang, S., Cai, J., & Baral, C. (2010). Discovering drug-drug interactions: A text-mining and reasoning approach based on properties of drug metabolism. *Bioinformatics*, 26, i547–i553. <https://doi.org/10.1093/bioinformatics/btq382>
- Temeles, E. J., Koulouris, C. R., Sander, S. E., & Kress, W. J. (2009). Effect of flower shape and size on foraging performance and trade-offs in a tropical hummingbird. *Ecology*, 90, 1147–1161. <https://doi.org/10.1890/08-0695.1>
- Valdovinos, F. S. (2019). Mutualistic networks: Moving closer to a predictive theory. *Ecology Letters*, 22, 1517–1534. <https://doi.org/10.1111/ele.13279>

- van Laarhoven, T., & Marchiori, E. (2013). Predicting drug-target interactions for new drug compounds using a weighted nearest neighbor profile. *PLoS ONE*, 8, e66952. <https://doi.org/10.1371/journal.pone.0066952>
- 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 B: Biological Sciences*, 281, 20132397. <https://doi.org/10.1098/rspb.2013.2397>
- Weinstein, B. G., & Graham, C. H. (2017). Persistent bill and corolla matching despite shifting temporal resources in tropical hummingbird-plant interactions. *Ecology Letters*, 20, 326–335. <https://doi.org/10.1111/ele.12730>
- Wen, M., Zhang, Z., Niu, S., Sha, H., Yang, R., Yun, Y., & Lu, H. (2017). Deep-learning-based drug-target interaction prediction. *Journal of Proteome Research*, 16, 1401–1409. <https://doi.org/10.1021/acs.jproteome.6b00618>
- Yamanishi, Y., Araki, M., Gutteridge, A., Honda, W., & Kanehisa, M. (2008). Prediction of drug-target interaction networks from the integration of chemical and genomic spaces. *Bioinformatics*, 24, i232–i240. <https://doi.org/10.1093/bioinformatics/btn162>

- Zhang, F., Wang, M., Xi, J., Yang, J., & Li, A. (2018). A novel heterogeneous network-based method for drug response prediction in cancer cell lines. *Scientific Reports*, 8, 3355. <https://doi.org/10.1038/s41598-018-21622-4>

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Pichler M, Boreux V, Klein A-M, Schleuning M, Hartig F. Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. *Methods Ecol Evol.* 2020;11:281–293. <https://doi.org/10.1111/2041-210X.13329>