

Research

Neutral processes related to regional bee commonness and dispersal distances are important predictors of plant–pollinator networks along gradients of climate and landscape conditions

M. A. K. Sydenham, Z. S. Venter, S. R. Moe, K. Eldegard, M. Kuhlmann, T. Reitan, C. Rasmussen, R. Paxton, Y. L. Dupont, A. B. Skrindo, S. J. Hegland, A. Nielsen, J. M. Olesen, M. Nowell and G. M. Rusch

M. A. K. Sydenham (<https://orcid.org/0000-0002-7711-2399>) ✉ (markus.sydenham@nina.no), Z. S. Venter (<https://orcid.org/0000-0003-2638-7162>), A. B. Skrindo (<https://orcid.org/0000-0003-0065-3434>) and M. Nowell (<https://orcid.org/0000-0002-3882-8546>), The Norwegian Inst. for Nature Research, Oslo, Norway. – S. R. Moe (<https://orcid.org/0000-0003-1005-3192>) and K. Eldegard (<https://orcid.org/0000-0002-3276-8087>), Faculty of Environmental Sciences and Natural Resource Management, Norwegian Univ. of Life Sciences, Ås, Norway. – M. Kuhlmann (<https://orcid.org/0000-0003-3664-6922>), Zoological Museum of Kiel Univ., Kiel, Germany. – T. Reitan (<https://orcid.org/0000-0003-0793-8731>), Dept of Biosciences, Centre for Ecological and Evolutionary Synthesis (CEES), Univ. of Oslo, Oslo, Norway. – C. Rasmussen (<https://orcid.org/0000-0003-1529-6548>), Dept of Agroecology, Aarhus Univ., Foulum, Denmark. – R. Paxton (<https://orcid.org/0000-0003-2517-1351>), General Zoology, Inst. for Biology, Martin Luther Univ. Halle-Wittenberg, Halle (Saale), Germany and German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany. – Y. L. Dupont (<https://orcid.org/0000-0002-8811-2773>) and J. M. Olesen (<https://orcid.org/0000-0002-3294-6234>), Dept of Ecoscience, Aarhus Univ., Rønde, Denmark. – S. J. Hegland (<https://orcid.org/0000-0001-7039-5966>), Dept of Environmental Sciences, Western Univ. of Applied Sciences, Sogndal, Norway. – A. Nielsen (<https://orcid.org/0000-0002-3294-6234>), Dept of Landscape and Biodiversity, Norwegian Inst. of Bioeconomy Research (NIBIO), Ås, Norway. – G. M. Rusch (<https://orcid.org/0000-0003-3769-8345>), The Norwegian Inst. for Nature Research, Trondheim, Norway.

Ecography

2022: e06379

doi: 10.1111/ecog.06379

Subject Editor: Elizabeth Le Roux

Editor-in-Chief: Miguel Araújo

Accepted 31 August 2022



www.ecography.org

Understanding how niche-based and neutral processes contribute to the spatial variation in plant–pollinator interactions is central to designing effective pollination conservation schemes. Such schemes are needed to reverse declines of wild bees and other pollinating insects, and to promote pollination services to wild and cultivated plants. We used data on wild bee interactions with plants belonging to the four tribes Loteae, Trifolieae, Anthemideae and either spring- or summer-flowering Cichorieae, sampled systematically along a 682 km latitudinal gradient to build models that allowed us to 1) predict occurrences of pairwise bee–flower interactions across 115 sampling locations, and 2) estimate the contribution of variables hypothesized to be related to niche-based assembly structuring processes (viz. annual mean temperature, landscape diversity, bee sociality, bee phenology and flower preferences of bees) and neutral processes (viz. regional commonness and dispersal distance to conspecifics). While neutral processes were important predictors of plant–pollinator distributions, niche-based processes were reflected in the contrasting distributions of solitary bee and bumble bees along the temperature gradient, and in the influence of bee flower preferences on the distribution of bee species across plant types. In particular, bee flower preferences separated bees into three main groups, albeit with some overlap: visitors to spring-flowering Cichorieae; visitors to Anthemideae and summer-flowering Cichorieae; and visitors to Trifolieae and Loteae. Our findings suggest that both neutral and niche-based processes are significant contributors to the spatial distribution of plant–pollinator interactions so that conservation actions in our region should be directed towards areas:

© 2022 The Authors. Ecography published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos

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.

near high concentrations of known occurrences of regionally rare bees; in mild climatic conditions; and that are surrounded by heterogeneous landscapes. Given the observed niche-based differences, the proportion of functionally distinct plants in flower-mixes could be chosen to target bee species, or guilds, of conservation concern.

Keywords: ecological networks, machine learning, plant–pollinator interactions, spatial, wild bees

Introduction

Pollinators are vital to the reproduction for most flowering plants (Ollerton et al. 2011), including many self-compatible species (Rodger et al. 2021). Bees are considered the most important pollinators (Ollerton 2017) and wild bee visitation frequencies often correlate with seed set in wild plants (Herbertsson et al. 2021). Fueled by declines in wild bees and other pollinating insects (reviewed in Ollerton 2017), and by evidence that plant reproduction is frequently pollen limited (Bennett et al. 2020, Reilly et al. 2020), a central objective in pollination ecology is to improve our understanding of the processes shaping spatial variation in plant–pollinator interactions (Knight et al. 2018). Such knowledge should allow identification of areas or actions that should be prioritized for the protection or restoration of plant–pollinator interactions (Cariveau et al. 2020, Pereira et al. 2022).

Landscape diversity is considered a central driver of pollinator distributions and plant interaction–partner diversity (Librán-Embú et al. 2021), but the influence of landscape simplification on pollination differs between plant species and regions (Herbertsson et al. 2021). A simple explanation for these differences may be the differential response to landscape structure and climate by different bee groups (Hoiss et al. 2012, Sydenham et al. 2015). Since bumble bees respond to habitat conditions at larger spatial scales than solitary bees (Steffan-Dewenter et al. 2002), pollen limitation in bumble bee pollinated plants could be less correlated with land-use intensity in regions with cool climates, where bumble bees predominate, than in regions with warmer climates where the main pollinators are thought to be solitary bees. Moreover, in temperate regions, bumble bees often form the core of plant–pollinator networks (Maia et al. 2019, Librán-Embú et al. 2021). Determining the relative influence that niche-based processes (viz. flower morphology, wild bee flower preferences, bee sociality, landscape diversity and climatic conditions) have on the spatial distribution of plant–pollinator interactions would be of value in the design of pollination–conservation schemes.

A complicating factor is that plant–pollinator interactions are not driven solely by niche-based processes but also correlate with, for example, the relative abundance of species (Vázquez et al. 2009a, b). The effect of abundance can be interpreted as being ecologically neutral, because it results in random interaction–distributions (Krishna et al. 2008). In addition to abundance, pollinator communities show a considerable amount of species turnover with geographic distance (Trøjelsgaard et al. 2015), and the occurrence and local abundances of pollinators have been found to decrease with distance to the nearest source population (Franzén and Nilsson 2013). If neutral processes are important contributors

to plant–pollinator interactions, then interactions between plants and rare pollinators will, theoretically, be most likely in large habitat patches close to existing populations of the pollinator species. Because local abundance is often correlated with species distributions (Caten et al. 2022), accounting for regional commonness and dispersal limitation is likely to improve predictions of bee pollination potential in plant communities (Burkle and Alarcón 2011, Tylianakis and Morris 2017, Sydenham et al. 2022a), and to guide the selection of habitat patches for pollination-oriented conservation and restoration.

The aim of this study was to assess the relative importance of niche-based species sorting and neutral processes in predicting the spatial distribution of bee–flower interactions in northern regions of Europe. We hypothesized that the probability of bee–flower interactions depends on a series of niche-based spatially structured environmental filters (Keddy 1992) that include trait-matching (Pichler et al. 2020) between plants and pollinators at the smallest spatial scale, as well as neutral processes related to regional commonness, and dispersal limitation. Interpreting the importance that ecological processes have on plant–pollinator interactions requires an understanding of their influence on pollinator distributions and probabilities of interactions. We elaborated our main hypothesis with four testable predictions: 1) bumble bees and solitary bees show contrasting responses to an increase in annual mean temperature (Hoiss et al. 2012), i.e. bumble bees increasingly dominate the community of wild bees along a S–N latitudinal gradient; and 2) the probability of bee–flower interactions, and hence flower–visitor species richness, increase with landscape diversity. We expected this increase to be more pronounced for solitary bees than for bumble bees, because of their stronger response to landscape conditions at small spatial scales (Steffan-Dewenter et al. 2002). Furthermore, we predicted that, 3) because of differences in typical flower preferences between bee species, pairwise interactions between plants and wild bees can be predicted from existing knowledge of associations between bees and plant families as well as bee phenologies; and 4) as an outcome of neutral community assembly, the probability of bee–plant interactions increases with regional commonness of a bee species and decreases with geographic distance to the nearest source population of the bee species.

Methods

Study design and field sampling

In 2017, we sampled flower-visiting bees at 115 sites in northern Germany, western Denmark and southeastern

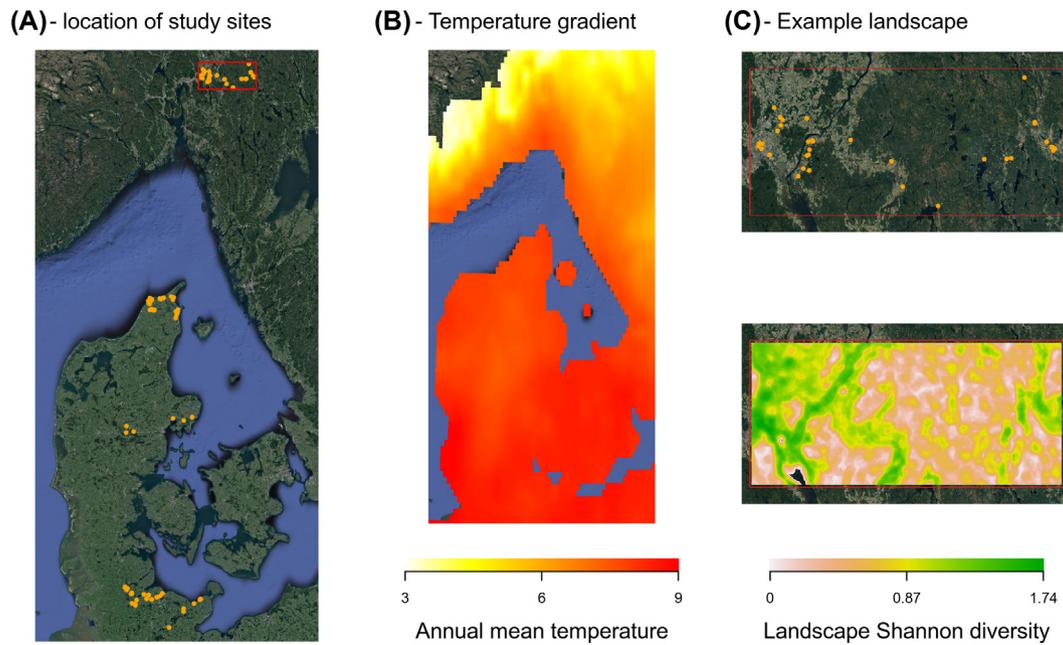


Figure 1. (A) Flower visiting bees were collected at 115 study sites (orange points) distributed across northern Germany, western Denmark and southeastern Norway (B) corresponding to a gradient in annual mean temperature. (C) Shannon landscape diversity at 1000 m radii illustrated for the spatial extent of the Norwegian study sites. Satellite imagery from Map data ©2021 Google via QGIS 2021.

Norway (Fig. 1, Table 1). We sampled bees visiting plant species belonging to four tribes with morphologically distinct flowers and flowering times: white-petalled Anthemideae (Asteraceae) species; yellow-petalled Cichorieae (Asteraceae) species, which also differed in their flowering phenology, Cichorieae spring (*Taraxacum* spp.) or Cichorieae summer (e.g. *Hieracium* spp.); and two tribes of Fabaceae with zygomorphic flowers: Loteae species with yellow flowers and Trifolieae species with white to purple flowers. Species within these plant tribes attract a diverse community of bees (Rasmussen et al. 2016). The number of sampled plant types

varied between sites, with one, two and three types sampled at 92, 17 and 6 sites, respectively. Anthemideae were not sampled in Denmark as we were unable to find suitable locations. Site-pairs with between-site distances < 1000 m could potentially share pollinators from the same populations (Greenleaf et al. 2007). We identified 76 clusters of sites within which between-site distances were < 1000 m and between-cluster-distances were always > 1000 m, i.e. 76 statistically independent clusters. Of the 76 clusters, 50 contained one site, 17 contained two, six contained three, two contained four and one cluster contained five sites.

Table 1. Overview of the number of interactions, bee species and sites sampled with different floral forms in northwestern Germany, western Denmark and southeastern Norway in 2017. Bees were collected on flowers belonging to four tribes of plants: yellow composites from the tribe Cichorieae (e.g. *Taraxacum* spp. in May and *Hieracium* spp. in June/July); white petalled composites from the tribe Anthemideae (e.g. *Leucanthemum vulgare*); yellow flowered legumes within the tribe Loteae (e.g. *Lotus corniculatus*); and white to purple flowered legumes within Trifolieae (i.e. *Trifolium* spp.). A total of 69 wild nest-building bee species, were sampled across all study sites and included in further analyses.

Country	Plant type	Sites	First date	Last date	Interactions	Bee species
Germany	Cichorieae spring	6	29.04.2017	30.04.2017	122	18
Germany	Cichorieae summer	10	01.06.2017	09.06.2017	105	24
Germany	Loteae	12	01.06.2017	10.06.2017	159	11
Germany	Trifolieae	11	01.06.2017	10.06.2017	157	12
Germany	Anthemideae	9	03.06.2017	07.06.2027	76	13
Denmark	Cichorieae spring	12	02.05.2017	06.05.2017	165	20
Denmark	Cichorieae summer	14	14.06.2017	20.06.2017	186	20
Denmark	Loteae	11	14.06.2017	20.06.2017	91	10
Denmark	Trifolieae	10	14.06.2017	20.06.2017	161	9
Norway	Cichorieae spring	7	20.05.2017	21.05.2017	101	17
Norway	Cichorieae summer	8	04.07.2017	08.07.2017	63	20
Norway	Loteae	12	26.06.2017	05.07.2017	101	15
Norway	Trifolieae	13	26.06.2017	08.07.2017	158	18
Norway	Anthemideae	9	29.06.2017	05.07.2017	29	8

At each site, we collected bees within an area of up to 50 by 50 m, depending on the spatial distribution of the focal plant species. All bees were sampled by the first author. At each site, bee sampling was conducted once by collecting flower-visiting bees for 30 min on each plant type, adding 30 s handling time per sampled bee. Sampling took place when the air temperature was $> 15^{\circ}\text{C}$ and with little to no wind (Beaufort scale 0–3). Sampling followed onset of the flowering season, i.e. it started in northern Germany and ended in southeastern Norway (Table 1). All collected bees were stored in 96% ethanol prior to pinning and species identification. Voucher specimens are deposited in the entomological collection at the Norwegian Institute for Nature Research, Oslo.

Predictors of bee–flower interactions

Mean annual temperature (BIO1) at a 10 km resolution was obtained from the WorldClim database (Fick and Hijmans 2017) using the 'Raster' package in R (Hijmans 2018). We focused on annual mean temperature as bee diversity is known to vary along temperature gradients (Hoiss et al. 2012) and mean annual temperature within our region was strongly correlated with other climatic variables such as annual precipitation (BIO12, Pearson's $r = -0.68$) and mean temperature in the warmest quarter (BIO10, Pearson's $r = 0.84$). Land-cover-data were extracted within radii of 250 and 1000 m surrounding each site from a 10 m resolution land cover map with eight land cover classes: built-up land; cropland; woodland; shrubland; grassland; bare land; water bodies; and wetland (Venter and Sydenham 2021). We used the proportion of land-cover types within each radius to calculate a Shannon landscape diversity index (hereafter: SH250m and SH1000m).

To obtain estimates of bee regional commonness and distance to nearest known potential source population of the sampled bee species, we used species occurrence records from gbif.org (GBIF 2022). We used the 'Dismo' package in R (Hijmans et al. 2017) to download all records of each species from the past 20 years (2000–2021) within a spatial extent (longitude = [8.99, 14.6], latitude = [52.4, 60.59]) encompassing our study region (longitude = [9.49, 12.18], latitude = [54.04, 60.09]) to ensure that at least one record of all sampled species was retrieved. To reduce the effect of spatial and taxonomic collector bias on our estimates of regional commonness, we used the number of 10×10 km grid cells occupied as an estimate of species-specific regional commonness. We pooled records of *Bombus terrestris* and *Bombus lucorum* as these cannot be reliably identified without molecular analyses (Carolan et al. 2012). The number of records per bee species ranged from three to 25 784 ($Q_{25} = 323$, $Q_{50} = 687$, $Q_{75} = 1826$). For each study site, we calculated the distance to the nearest GBIF-record of each bee species as an estimate of dispersal distances to the nearest known source population.

As a phenology trait, we used the first month each bee species is active (emerges), as listed in Scheuchl and Willner (2016). We assigned a categorical variable (bee sociality) to each species to distinguish between bumble bees and solitary bees. We

assigned floral association traits to each bee species using the approach in Sydenham et al. (2022a) and available information on plant–bee associations (Rasmussen et al. 2021, Wood et al. 2021) to build a bee species-by-plant family data frame containing the number of genera per plant family each bee species is known to visit. We used a detrended correspondence analysis (DCA) in the 'vegan' package in R (Oksanen et al. 2018) to ordinate the bee-by-plant matrix and used bee species scores (Bee DCA1, Bee DCA2, Bee DCA3, Bee DCA4) as flower-association traits (Supporting information).

Analyses

To predict bee–flower interactions and to assess the relative importance of predictor variables, we used the MetaComNet framework (Table 2, Sydenham et al. 2022a) with random forest regressions and classifications (Breiman 2001). Because different modelling approaches might vary in how predictor variables are ranked we also ran our MetaComNet models using boosted regression trees and binomial generalized linear models (GLMs). Random forest models and boosted regression tree models were fit calling 'ranger' (Wright and Ziegler 2017) or 'gbm' (Greenwell et al. 2020) via 'caret' in R (Kuhn et al. 2018), while GLMs were fit using base R (www.r-project.org). All four models were fit following the general formula (R syntax):

$$\begin{aligned}
 Y \sim & \text{Plant tribe} + \text{bee sociality} + \text{Bee DCA1} + \text{Bee DCA2} \\
 & + \text{Bee DCA3} + \text{Bee DCA4} + \text{bee phenology} + \text{SH250m} \\
 & + \text{SH1000m} + \text{BIO1} + \text{Plant tribe} \times (\text{Bee DCA1} \\
 & + \text{Bee DCA2} + \text{Bee DCA3} + \text{Bee DCA4} + \text{bee phenology}) \\
 & + \text{bee sociality} \times (\text{SH250m} + \text{SH1000m} + \text{BIO1}) \\
 & + \text{Distance to species} + \text{Regional commonness}
 \end{aligned}$$

where Y indicated the presence or absence of an interaction between a bee species and plant at a given site. Y was a binary variable (0 versus 1) for random forest regressions and binomial GLMs and a two-level categorical variable for random forest classification trees and boosted regression trees. Although it is not necessary to specify interaction terms when fitting machine learning models, we did so in order to obtain variable importance scores for the same predictor terms that were included in the GLMs.

For the binomial GLM we used a second order polynomial transformation on bee DCA-scores and bee phenology (month of emergence) to allow for non-linear, bell-shaped, responses. Although machine learning methods are robust to correlated predictors, multicollinearity may still affect the estimated variable importance scores. However, we found no strong correlations between the continuous predictor variables (Supporting information).

We used leave-one-out cross-validation by iteratively removing all sites belonging to one cluster and trained

random forest models on sites from the remaining 75 clusters before predicting occurrence probabilities of bee–flower interactions for plant tribes in the withheld sites and before assessing variable importance. This blocking strategy allowed models to be trained on bee–plant interactions from the sites in the training data. The cluster-oriented, or block, cross-validation ensured that a model would not be trained on sites belonging to the same cluster as the data the model was tested on. Machine learning models are sensitive to user-specified hyper parameters. For random forest models and boosted regression trees we therefore used a nested cross-validation with 10-fold cross-validation on the training data to identify the hyper-parameter settings that resulted in the highest accuracy. For random forest models we used the default tuning parameters from 'caret' (Kuhn et al. 2018) with: *mtry* set to 2, 20 or 38; *splitrule* set to *gini* or *extratrees* for classification trees and variance or *extratrees* for regression trees; and *min.node.size* set to 1 for classification, and 5 for regression trees. For boosted regression tree models we used a tuning grid with: *n.trees* set to 100, 250 or 500; *interaction.depth* set to 3, 4, 5 or 6; *shrinkage* set to 0.1; and *n.minobsinnode* set to 10 or 20. The optimal hyper parameter setting was then used to refit the model on the whole training data before predicting onto the withheld data. We used the *VarImp* function in 'caret' (Kuhn et al. 2018) to obtain permutation based variable importance scores.

For GLMs we fit the full models, including all interactions. We also used the *dredge* function in 'MuMIn' (Barton 2020) to identify the highest ranked models i.e. with $\Delta AICc < 2$. We applied full model averaging across highest ranked models to obtain parameter estimates and effect sizes and for predicting the probability of interactions on the withheld data. Following the approach used in 'caret' (Kuhn et al. 2018), we used the effect sizes (z-values) as a measure of variable importance, scaled between

0 and 100 for comparability with importance scores from the machine learning methods. We calculated variable importance scores for both the full GLM and the AICc selected GLM. For each model family (random forests, boosted regression trees, GLMs) we calculated the average importance for each predictor variable across the 76 models and used this value as an overall estimate of predictor importance.

We tested if the models predicted pairwise interactions between bees and plants in the withheld data. We used three metrics (AUC, regression slopes, R^2). AUC was calculated using the *roc* function from the 'pROC' package in R (Robin et al. 2011). Regression slopes were obtained from binomial GLMs with observed occurrences of interactions as binary response variable as a function of the logit-transformed predicted frequency of interactions (to allow for proportionality). Nagelkerke R^2 values were obtained using the *r.squaredLR* function in 'MuMIn' in R (Barton 2020).

To explore and test our predictions of how the predictor variables might influence the occurrence of pairwise interactions, we refitted the GLM to the whole dataset. In addition to the full model also used the *dredge* function (Barton 2020) to identify the single most probable (final) model based on AICc. Generalized linear mixed models (GLMMs) formulated as the full or the AICc selected model but with site identity as a random intercept term, yielded regression coefficients that were strongly correlated ($r=0.99$) to those from the GLMs, suggesting that the parameter estimates from our GLMs were not biased from site-specific differences in species occurrences. We used 'DHARMA' (Hartig 2021) on the final GLMs for residual diagnostics.

All analyses and data preprocessing were conducted using R ver. 4.1.2 (<www.r-project.org>) on a Rstudio server (x86_64-pc-linux-gnu (64-bit), Ubuntu 18.04.6 LTS). Code and data for reproducing the results is included in the Supporting information.

Table 2. Parameters used to predict interactions between wild bee species and plants.

Parameter	Class	Description
Response		
Interaction occurrence	Binary	Presence (1) or absence (0) of observations of wild bee species on the focal plant tribe within a site
Predictors		
Annual mean temperature	Numeric	Annual mean temperature within 10 × 10 km grids
Shannon landscape diversity 250 m	Numeric	Landscape diversity of land cover classes within 250 m of study sites
Shannon landscape diversity 1000 m	Numeric	Landscape diversity of land cover classes within 1000 m of study sites
Plant type	Factor	Plant tribe of the focal plant on which bees were sampled: 1) Cichorieae sampled in spring, 2) Cichorieae, 3) Anthemideae, 4) Loteae and 5) Trifolieae, sampled in summer
Wild bee regional commonness	Numeric	Number of 10 km grid cells with records of the species
Wild bee distance to conspecific	Numeric	Geographic distance to nearest GBIF occurrence of the species
Wild bee phenology	Numeric	Earliest month of emergence for the species
Wild bee social status	Factor	Solitary versus Eusocial (<i>Bombus</i>)
Wild bee floral preference (DCA1–4)	Numeric	Species scores on DCA axes 1–4 from the detrended correspondence analysis on presence or absence of documented interactions between bee species and plant families obtained from the literature

Results

We sampled a total of 69 non-parasitic wild bee species and recorded 1674 pairwise interactions between bees and wild plants across the 115 study sites. Predicted occurrence probabilities of interactions corresponded to observed occurrences (Fig. 2A–D). Compared to the boosted regression trees (Fig. 2C, $R^2=0.35$) and the binomial GLM (Fig. 2D, $R^2=0.37$), predictions from the random forest classification ($R^2=0.38$) and regression models ($R^2=0.39$) explained slightly more of the variation (R^2) in observed occurrences (Fig. 2A–B). However, the random forest models (slopes = 0.71 and 0.79) tended to overestimate the proportion of occurrences (Fig. 2A–B, slopes < 1) whereas the predicted probabilities of occurrence from boosted regression trees (slope = 0.91) and binomial GLMs (slope = 0.94)

tended to be more closely aligned with the actual proportion of occurrences (Fig. 2C–D, slopes closer to one).

The four prediction models largely agreed upon the ranking of predictor variables in terms of their importance. The ecologically neutral variable; regional commonness of wild bees, was consistently ranked as the most important predictor of pairwise interactions, and distances to conspecifics was also ranked highly (Fig. 2E–F). The ranking of variables related to niche-based species sorting differed slightly among models, but bee flower preferences along DCA1 and its interaction with plant tribes was consistently ranked highly and on par with, or more important than, variable terms that included climatic conditions or landscape diversity. The annual mean temperature and its interaction with wild bee sociality was ranked slightly higher than landscape diversity and its interaction with wild bee sociality by all models except for the

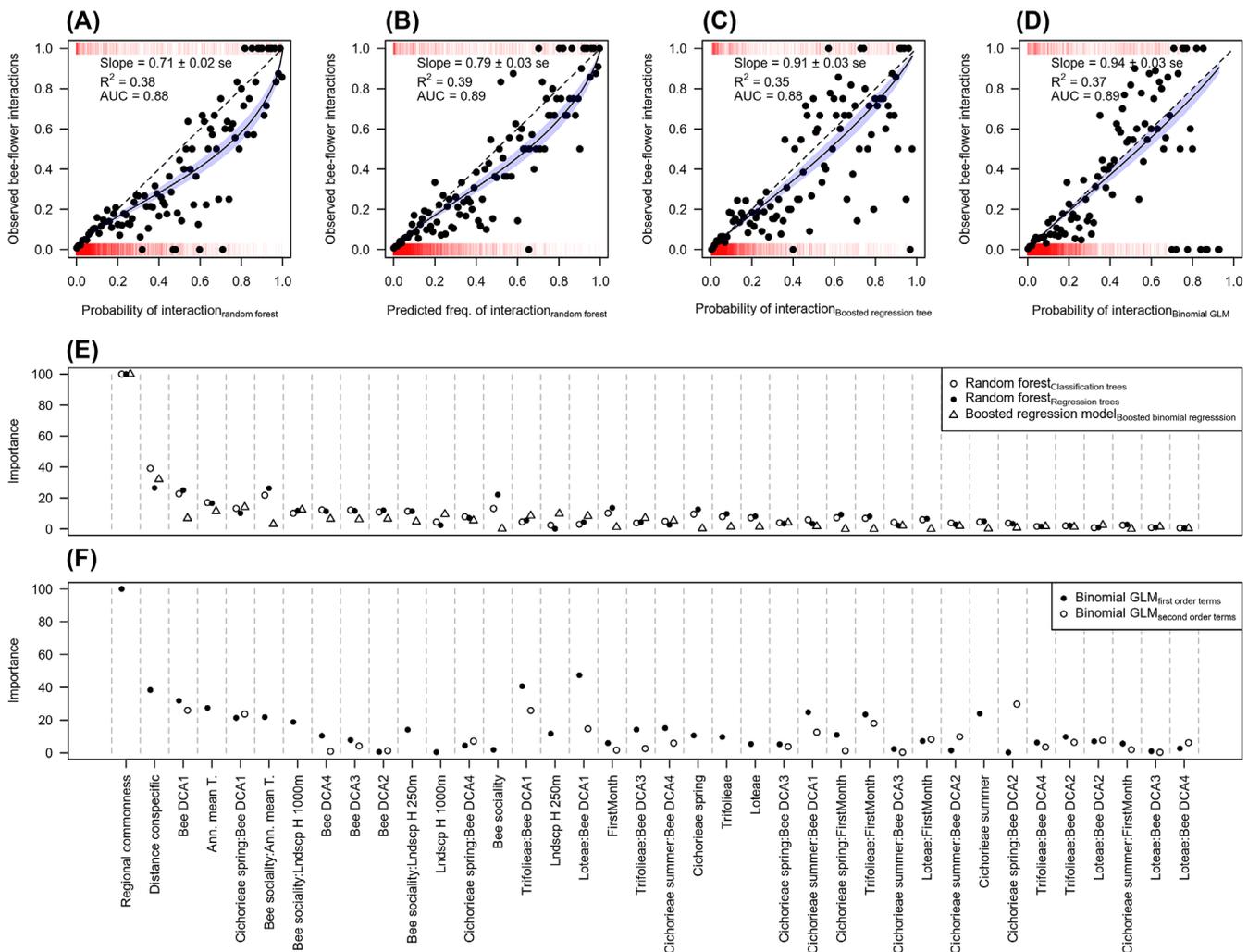


Figure 2. Validation of the ability of the four models to predict (A–D) occurrences of pairwise interactions between 69 wild bee species and plants belonging to four plant tribes. Black points in (A)–(D) show the frequency of occurrences within bins of 0.01 along the x-axis. Regression lines and 95% confidence intervals are from logistic GLMs. Black dashed lines show an ideal 1:1 relationship between observed and predicted occurrences of pairwise interactions. Red ticks show observed occurrences or absences of interactions between a bee species and the target plant. (E) Relative importance of predictor variables according to the machine learning models. (F) Relative importance according to the binomial GLM.

boosted regression trees, which consistently ranked bee sociality and terms with which it interacted lowly. A notable difference between the models was that the binomial GLM ranked terms that included bee DCA scores relatively higher than the machine learning models. Despite these differences, the predicted probability of interactions occurring were strongly correlated ($r > 0.7$) across all four models and were also positively correlated with the number of interactions between wild bees and plants within sites (Table 3).

The binomial GLM, fitted to all the data, showed that the probability of occurrence of interactions increased strongly with the regional commonness of the wild bee species (Fig. 3A) and decreased with distances to conspecifics (Fig. 3B). Among the spatial environmental filters, there was a selection for solitary bees as annual mean temperatures increased, while social species, i.e. bumblebees, were largely unaffected (Fig. 3C). Social bee occurrence increased with landscape diversity at the 1000 m radius (Fig. 3D). While solitary bee occurrence was largely unaffected by landscape diversity at 1000 m, occurrence of solitary bees increased with landscape diversity at the 250 m scale whereas social bees showed a weak decrease (Fig. 3E). These results differed from the AICc selected model where both solitary and social bees showed a weak increase with landscape diversity at the 1000 m radius, with social bees showing a slightly stronger increase, though this was not statistically significant ($df = 1$, $\chi^2 = 3.6$, $p = 0.06$, Supporting information). The floral association traits (DCA-scores) of bees as well as bee phenology were strong predictors of the probability of interactions between bee species and the four plant tribes (Fig. 3F–J). Species scores on DCA1 separated species based their probabilities of interacting with Anthemideae and Cichorieae (low to intermediate DCA1 scores), from those interacting with the Fabaceae tribes: Trifolieae and Loteae (high DCA1 scores). The second, third and fourth DCA axes (Fig. 3G–I) were mainly related to increases in the probability of wild bees interacting with summer flowering Cichorieae plants. Similarly, bees that emerge later in the summer were the most likely visitors on summer flowering Cichorieae, Anthemideae and Loteae (Fig. 3J). By contrast, and perhaps not surprising, the most likely visitors of spring flowering Cichorieae and Trifolieae were bees with early emergence phenologies.

Discussion

Our aim was to assess the relative importance of regional commonness, dispersal limitation and niche-based processes

on the spatial distribution of plant–pollinator interactions along a climatic gradient. All four prediction models of pairwise plant–bee interactions resulted in predictions that were positively related to observed pairwise interactions, and the predictions from the four models were strongly correlated to each other, suggesting a high degree of agreement between models. Taken together, our findings suggest that our models identified well the relative importance of variables for predicting plant–bee interactions.

In line with our predictions and with the findings from previous studies (Hoiss et al. 2012), lower annual mean temperatures corresponded to an increasing dominance of bumble bees in flower–visitor assemblages. The selection for bumble bees with decreasing temperatures is likely because bumble bees evolved in cool climates (Hines 2008), and therefore are generally less thermophilic than solitary bees.

However, in contrast to our predictions, we did not find solitary bee occurrences increased more strongly with landscape diversity than for bumble bees. Rather there was a tendency for bumble bee occurrences to be more responsive to landscape diversity at the 1000 m scale. The GLM did however find a stronger and positive response to increasing landscape diversity for solitary bees than for bumble bees at the 250 m radii scale, but this effect was weak (Bee sociality \times Landsc H 250m, z -value = -1.95) and was not included in the AICc selected model (Supporting information). This was surprising because bee foraging ranges correlate with body sizes (Greenleaf et al. 2007) and because the smaller solitary bees have earlier been found to respond more strongly to landscape factors at small spatial scales than bumble bees (Steffan-Dewenter et al. 2002). However, landscape diversity was not one of the most important predictors in our system and it may be that the effect of landscape diversity on solitary bees was underestimated in our study because of limited sampling intensity. Still, despite its relatively low importance, bee occurrence increased with landscape diversity, as also found by others (Boscolo et al. 2017), suggesting that landscape diversity could be a useful proxy for habitat area.

The relatively high importance of climatic conditions in comparison to landscape diversity, identified by the random forest models and the GLM, is in line with the multi-level environmental filtering hypothesis, which posits that local species assemblages, in our case flower-visiting bees, represent the outcome of a hierarchically nested set of environmental filters (Keddy 1992). The importance of plant type and flower-preference related bee traits was often as high or higher

Table 3. Pearson's correlation between the observed number of interactions between wild bees and plants and the prediction outputs from random forest (RF) classification and regression trees, boosted regression trees (BTR) and binomial generalized linear models (GLM). For GLMs correlation coefficients are shown for both the full model and the AICc selected model.

	Interactions	RF _{Classification}	RF _{Regression}	BTR _{Classification}	GLM
RF _{Classification}	0.50	1			
RF _{Regression tree}	0.52	0.98	1		
BTR _{Classification}	0.49	0.85	0.87	1	
GLM	0.45	0.78	0.81	0.86	1
GLM _{AICc selected}	0.45	0.78	0.81	0.86	0.99

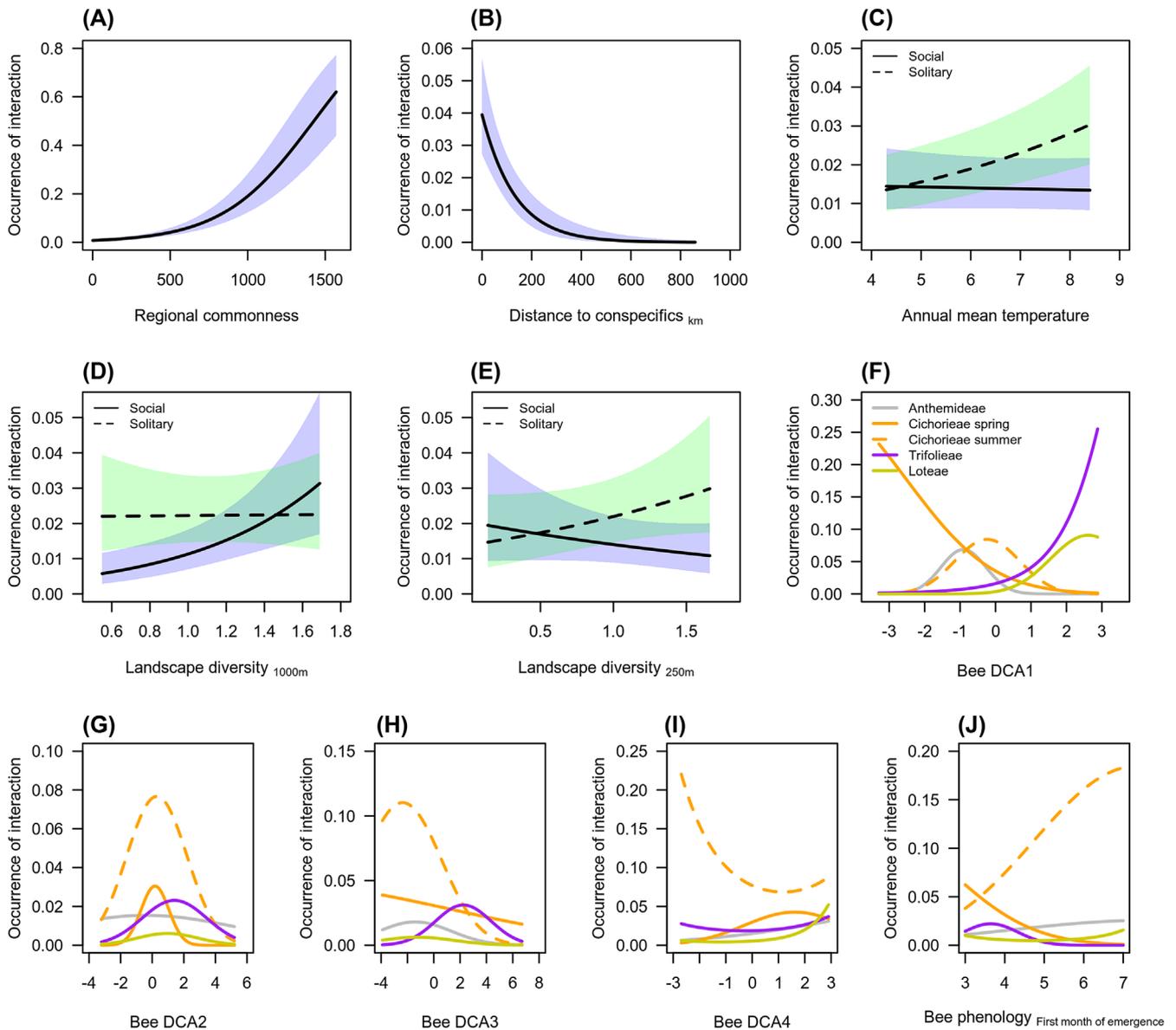


Figure 3. Predictors of pairwise interactions between wild bees and plants illustrated by their marginal effects on occurrences of interactions. Predictor effects were estimated from a binomial GLM. Polygons in (A)–(E) show 95% confidence intervals around estimated effects. Note that the scale on the y-axis varies between subplots. Overall the model explained 41% of the variation in occurrences of pairwise interactions (full model AICc = 2851.6, on 9874 degrees of freedom).

than climatic conditions, which seemingly contradicts the multi-level filtering hypothesis. However, although functionally distinct plants can share generalist pollinator bees, different plants also attract specific bee species (Rasmussen et al. 2021), and the phenology of plants may limit the subset of bee species available to them (Olesen et al. 2011). This complementarity of plants in terms of their attraction to different species of bees likely underlies the frequent finding that richness of both common and rare bee species increases with plant species richness (Sutter et al. 2017), and may explain why plant diversity is often found to be of greater importance in explaining patterns of bee diversity than landscape diversity (Sydenham et al. 2015, Lane et al. 2022, but see

Griffin et al. 2021). In terms of the multi-level filtering hypothesis, plant species or plant type, may therefore operate as a high-level filter by constraining the subset of bee species from the regional species pool that can be filtered by climatic and other environmental conditions.

In line with previous studies showing that species abundances are strongly related to patterns of plant–pollinator interactions within (Krishna et al. 2008, Vázquez et al. 2009a) and among networks (Sydenham et al. 2022a), bee regional commonness was the most important predictor of the occurrences of pairwise interactions across sites. Because we used independent data to estimate regional commonness, its effect on interaction occurrence can not be attributed to

a sampling bias against uncommon species, which might have been the case if relative abundance – or commonness – had been estimated directly from plant–pollinator networks (Vázquez et al. 2009b). Still, the effect of regional commonness needn't be ecologically neutral but could be related to niche-based processes. For instance, compared to solitary bees, bumble bees are regionally common, occur at high densities locally, often visiting a wide range of plant species within our region (Maia et al. 2019; Wood et al. 2021). However, regional commonness, or range, is also a good predictor of the number of local occurrences within bumble bees (Williams 2005), and the positive relationship between species range sizes and distributions or local abundances is a recurring pattern in macroecology (Caten et al. 2022). Despite its generality, the mechanism behind the range size and local abundance relationship is not fully understood and it is often regarded as an outcome of niche-based, neutral and sampling-based processes and their interactions (Gaston et al. 1997, Borregaard and Rahbek 2010). In plant–pollinator networks, species occurrences could reflect an ecologically neutral network-assembly (Krishna et al. 2008), with regional commonness itself reflecting species' tolerances to regional climate conditions, habitat availability at wider spatial scales and the spatially synchronizing effects of high rates of species dispersals.

In addition to regional commonness, we found that distance to the nearest known occurrence of a species was an important predictor of bee–flower interactions in our study system. The decrease in interactions with distance to the nearest conspecific could partly reflect our study sites being located along a biogeographic gradient. Indeed, the phylogenetic diversity of bees decreases with elevation (Hoiss et al. 2012), and some species such as *Halictus quadricinctus* and *Andrena strobilmella* were confined to the southernmost part of our study region. However, the steep decrease in occurrence at distances to conspecifics of up to 200 km suggests that distance to conspecifics not only reflected changes in the species pool along the climatic gradient, but also bee species turnover at smaller spatial scales, which others have found to be pronounced even at modest spatial scales (Morón et al. 2017). Species repositories, such as the GBIF database used in this study to estimate regional commonness and distance to conspecifics, are known to suffer from spatial and taxonomic bias (Rocha-Ortega et al. 2021). Nevertheless, despite these limitations, the neutral variables derived from GBIF-data proved to be among the most important predictors of bee–flower interactions in our system, suggesting that incorporating citizen science data into models of plant–pollinator interactions can improve spatial predictions.

Conclusions

Occurrences of plant–bee interactions seems to be jointly influenced by neutral and niche-based processes. By targeting areas that are near known occurrences of rare bees and that are located in heterogenous landscapes, conservation schemes can account for neutral processes, increasing the

odds of conserving interactions between plants and rare bees. Important niche-based differences that should be considered are that solitary and bumble bees had contrasting responses to annual mean temperature and differed in the plant types they typically visited. Preserving and restoring areas with a high abundance of both zygomorphic and actinomorphic flowers appear to be important strategies for conserving wild bees. Furthermore, because solitary bees seem to be restricted by cool climates, flower community composition with a larger component of actinomorphic plants (e.g. Cichorieae and Anthemideae) could be a possible measure to reduce the impact of environmental (e.g. decreasing temperatures) stress on solitary bee populations.

Acknowledgements – We are thankful to landowners for allowing us to collect bees on their land and to Mikaela E. Olsen for digitizing the collected bee specimens. Two anonymous reviewers kindly provided constructive comments and suggestions.

Funding – The Norwegian University of Life Sciences funded fieldwork and identifications of species through a postdoctoral position awarded to the first author. This research was funded in part by The Research Council of Norway (project no. 302692) as part of the project 'MetaComNet: Linking metacommunity dynamics to the structure of ecological networks and ecosystem functioning', awarded to M. Sydenham.

Ethics statement – Bees in northern Germany were collected under permit from the Schleswig-Holstein Nature Conservation Bureau, issued to M. Kuhlmann.

Author contributions

Markus Arne Kjær Sydenham: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Funding acquisition (lead); Methodology (lead); Project administration (lead); Validation (lead); Visualization (lead); Writing – original draft (lead); Writing – review and editing (lead). **Zander S. Venter:** Data curation (supporting); Formal analysis (supporting); Funding acquisition (supporting); Methodology (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Stein R. Moe:** Funding acquisition (supporting); Supervision (lead); Writing – original draft (supporting); Writing – review and editing (supporting). **Katrine Eldegard:** Funding acquisition (supporting); Supervision (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Michael Kuhlmann:** Methodology (supporting); Supervision (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Trond Reitan:** Formal analysis (supporting); Funding acquisition (supporting); Methodology (supporting); Validation (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Claus Rasmussen:** Data curation (supporting); Funding acquisition (supporting); Methodology (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Robert Paxton:** Data curation (supporting); Writing – original draft (supporting); Writing – review and editing (supporting).

Yoko L. Dupont: Funding acquisition (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Astrid B. Skrindo:** Methodology (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Megan Nowell:** Writing – original draft (supporting); Writing – review and editing (supporting). **Stein J. Hegland:** Funding acquisition (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Anders Nielsen:** Funding acquisition (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Jens M. Olesen:** Funding acquisition (supporting); Writing – original draft (supporting); Writing – review and editing (supporting). **Graciela M. Rusch:** Funding acquisition (supporting); Writing – original draft (supporting); Writing – review and editing (supporting).

Transparent peer review

The peer review history for this article is available at <<https://publons.com/publon/10.1111/ecog.06379>>.

Data availability statement

Data and R code are available from the Dryad Digital Repository: <<https://doi.org/10.5061/dryad.9w0vt4bjr>> (Sydenham et al. 2022b).

Supporting information

The Supporting information associated with this article is available with the online version.

References

- Barton, K. 2020. MuMIn: multi-model inference. – R package ver. 1.43.17, <<https://CRAN.R-project.org/package=MuMIn>>.
- Bennett, J. M. et al. 2020. Land use and pollinator dependency drives global patterns of pollen limitation in the Anthropocene. – *Nat. Commun.* 11: 1–6.
- Borregaard, M. K. and Rahbek, C. 2010. Causality of the relationship between geographic distribution and species abundance. – *Q. Rev. Biol.* 85: 3–25.
- Boscolo, D. et al. 2017. Positive responses of flower visiting bees to landscape heterogeneity depend on functional connectivity levels. – *Perspect. Ecol. Conserv.* 15: 18–24.
- Breiman, L. 2001. Random forests. – *Mach. Learn.* 45: 5–32.
- Burkle, L. A. and Alarcón, R. 2011. The future of plant–pollinator diversity: understanding interaction networks across time, space and global change. – *Am. J. Bot.* 98: 528–538.
- Cariveau, D. P. et al. 2020. A review of the challenges and opportunities for restoring animal-mediated pollination of native plants. – *Emerg. Top. Life Sci.* 4: 99–109.
- Carolan, J. C. et al. 2012. Colour patterns do not diagnose species: quantitative evaluation of a DNA barcoded cryptic bumblebee complex. – *PLoS One* 7: e29251.
- Caten, C. T. et al. 2022. Weak but consistent abundance–occupancy relationships across taxa, space and time. – *Global Ecol. Biogeogr.* 31: 968–977.
- Fick, S. E. and Hijmans, R. J. 2017. WorldClim 2: new 1 km spatial resolution climate surfaces for global land areas. – *Int. J. Climatol.* 37: 4302–4315.
- Franzén, M. and Nilsson, S. G. 2013. High population variability and source–sink dynamics in a solitary bee species. – *Ecology* 94: 1400–1408.
- Gaston, K. J. et al. 1997. Interspecific abundance–range size relationships: an appraisal of mechanisms. – *J. Anim. Ecol.* 66: 579–601.
- GBIF 2022. Filtered export of GBIF occurrence data. – <<https://doi.org/10.15468/dd.5k37kz>>.
- Greenleaf, S. S. et al. 2007. Bee foraging ranges and their relationship to body size. – *Oecologia* 153: 589–596.
- Greenwell, B. et al. 2020. gbm: generalized boosted regression models. – R package ver. 2.1.8, <<https://CRAN.R-project.org/package=gbm>>.
- Griffin, S. R. et al. 2021. Bee communities in restored prairies are structured by landscape and management, not local floral resources. – *Basic Appl. Ecol.* 50: 144–154.
- Hartig, F. 2021. DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. – R package ver. 0.4.1, <<https://CRAN.R-project.org/package=DHARMa>>.
- Herbertsson, L. et al. 2021. Bees increase seed set of wild plants while the proportion of arable land has a variable effect on pollination in European agricultural landscapes. – *Plant Ecol. Evol.* 154: 341–350.
- Hijmans, R. J. 2018. raster: geographic data analysis and modeling. – R package ver. 2.8-4, <<https://CRAN.R-project.org/package=raster>>.
- Hijmans, R. J. et al. 2017. dismo: species distribution modeling. – R package ver. 1.1-4, <<https://CRAN.R-project.org/package=dismo>>.
- Hines, H. M. 2008. Historical biogeography, divergence times and diversification patterns of bumble bees (Hymenoptera: Apidae: *Bombus*). – *Syst. Biol.* 57: 58–75.
- Hoiss, B. et al. 2012. Altitude acts as an environmental filter on phylogenetic composition, traits and diversity in bee communities. – *Proc. R. Soc. B* 279: 4447–4456.
- Keddy, P. A. 1992. Assembly and response rules: two goals for predictive community ecology. – *J. Veg. Sci.* 3: 157–164.
- Knight, T. M. et al. 2018. Reflections on, and visions for, the changing field of pollination ecology. – *Ecol. Lett.* 21: 1282–1295.
- Krishna, A. et al. 2008. A neutral-niche theory of nestedness in mutualistic networks. – *Oikos* 117: 1609–1618.
- Kuhn, M. et al. 2018. caret: classification and regression training. – R package ver. 6.0-81, <<https://CRAN.R-project.org/package=caret>>.
- Lane, I. G. et al. 2022. Differences in bee community composition between restored and remnant prairies are more strongly linked to forb community differences than landscape differences. – *J. Appl. Ecol.* 59: 129–140.
- Librán-Embíd, F. et al. 2021. A plant–pollinator metanetwork along a habitat fragmentation gradient. – *Ecol. Lett.* 24: 2700–2712.
- Maia, K. P. et al. 2019. Does the sociality of pollinators shape the organisation of pollination networks? – *Oikos* 128: 741–752.
- Morón, D. et al. 2017. Railway lines affect spatial turnover of pollinator communities in an agricultural landscape. – *Divers. Distrib.* 23: 1090–1097.
- Oksanen, J. et al. 2018. vegan: community ecology package. – R package ver. 2.5-3, <<https://CRAN.R-project.org/package=vegan>>.

- Olesen, J. M. et al. 2011. Missing and forbidden links in mutualistic networks. – *Proc. R. Soc. B* 278: 725–732.
- Ollerton, J. 2017. Pollinator diversity: distribution, ecological function and conservation. – *Annu. Rev. Ecol. Evol. Syst.* 48: 353–376.
- Ollerton, J. et al. 2011. How many flowering plants are pollinated by animals? – *Oikos* 120: 321–326.
- Pereira, J. et al. 2022. Priority areas for protection of plant–pollinator interaction networks in the Atlantic Forest. – *Ecol. Indic.* 136: 108598.
- Pichler, M. et al. 2020. Machine learning algorithms to infer trait-matching and predict species interactions in ecological networks. – *Methods Ecol. Evol.* 11: 281–293.
- Rasmussen, C. et al. 2016. Distribution, phenology and host plants of Danish bees (Hymenoptera, Apoidea). – *Zootaxa* 4212: 1–100.
- Rasmussen, C. et al. 2021. Evaluating competition for forage plants between honey bees and wild bees in Denmark. – *PLoS One* 16: e0250056.
- Reilly, J. R. et al. 2020. Crop production in the USA is frequently limited by a lack of pollinators. – *Proc. R. Soc. B* 287: 20200922.
- Robin, X. et al. 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. – *BMC Bioinform.* 12: 77.
- Rocha-Ortega, M. et al. 2021. Geographical, temporal and taxonomic biases in insect GBIF data on biodiversity and extinction. – *Ecol. Entomol.* 46: 718–728.
- Rodger, J. G. et al. 2021. Widespread vulnerability of flowering plant seed production to pollinator declines. – *Sci. Adv.* 7: eabd3524.
- Scheuchl, E. and Willner, W. 2016. Taschenlexikon der Wildbienen Mitteleuropas: Alle Arten im Porträt. – Quelle and Meyer Verlag, 917 pp.
- Steffan-Dewenter, I. et al. 2002. Scale-dependent effects of landscape context on three pollinator guilds. – *Ecology* 83: 1421–1432.
- Sutter, L. et al. 2017. Enhancing plant diversity in agricultural landscapes promotes both rare bees and dominant crop-pollinating bees through complementary increase in key floral resources. – *J. Appl. Ecol.* 54: 1856–1864.
- Sydenham, M. A. K. et al. 2015. Does multi-level environmental filtering determine the functional and phylogenetic composition of wild bee species assemblages? – *Ecography* 38: 140–153.
- Sydenham, M. A. K. et al. 2022a. MetaComNet: a random forest-based framework for making spatial predictions of plant–pollinator interactions. – *Methods Ecol. Evol.* 13: 500–513.
- Sydenham, M. A. K. et al. 2022b. Data from: Neutral processes related to regional bee commonness and dispersal distances are important predictors of plant–pollinator networks along gradients of climate and landscape conditions. – *Dryad Digital Repository*, <<https://doi.org/10.5061/dryad.9w0vt4bjr>>.
- Trøjelsgaard, K. et al. 2015. Geographical variation in mutualistic networks: similarity, turnover and partner fidelity. – *Proc. R. Soc. B* 282: 20142925.
- Tylianakis, J. M. and Morris, R. J. 2017. Ecological networks across environmental gradients. – *Annu. Rev. Ecol. Evol. Syst.* 48: 25–48.
- Vázquez, D. P. et al. 2009a. Evaluating multiple determinants of the structure of plant–animal mutualistic networks. – *Ecology* 90: 2039–2046.
- Vázquez, D. P. et al. 2009b. Uniting pattern and process in plant–animal mutualistic networks: a review. – *Ann. Bot.* 103: 1445–1457.
- Venter, Z. S. and Sydenham, M. A. K. 2021. Continental-scale land cover mapping at 10 m resolution over Europe (ELC10). – *Remote Sens.* 13: 2301.
- Williams, P. 2005. Does specialization explain rarity and decline among British bumblebees? A response to Goulson et al. – *Biol. Conserv.* 122: 33–43.
- Wood, T. J. et al. 2021. Global patterns in bumble bee pollen collection show phylogenetic conservation of diet. – *J. Anim. Ecol.* 90: 2421–2430.
- Wright, M. N. and Ziegler, A. 2017. ranger: a fast implementation of random forests for high dimensional data in C++ and R. – *J. Stat. Softw.* 77: 1–17.