Journal of Ecology

MR ØYSTEIN OPEDAL (Orcid ID : 0000-0002-7841-6933)

Using hierarchical joint models to study reproductive interactions in plant communities

Øystein H. Opedal[1,*] & Stein Joar Hegland[2]

[1]Research Centre for Ecological Change, Faculty of Biological and Environmental Sciences, University of Helsinki, Helsinki, Finland.

[2]Department of Environmental Sciences, Western Norway University of Applied Sciences, P.O. Box 133, N-5801 Sogndal, Norway.

*Corresponding author e-mail: ohopedal@gmail.com

Running head: Studying reproductive interactions with joint models

Abstract

1. Pollinator-mediated reproductive interactions among coflowering plant species are prime examples of how species interactions may affect fitness and community assembly. Despite considerable interest in these issues, statistical methods for assessing signal of reproductive interactions in observational data on coflowering species are currently lacking.

2. We propose a flexible method for quantifying potential reproductive interactions among coflowering plant species using the hierarchical latent-variable joint models implemented in the Hierarchical Modelling of Species Communities (HMSC) framework. The method accommodates any measure of reproductive success, including pollinator visitation, stigma pollen loads, and seed set. We demonstrate the method by analysing a dataset on bumblebee visitation to a set of coflowering plant species in a species-rich meadow in Norway, and provide R tutorials for this and additional data types.

3. The example analysis revealed both positive and negative effects of heterospecific flower abundances on visitation to coflowering species, which we interpret as potential reproductive interactions.

4. Hierarchical joint models provide a flexible approach to analysing patterns of covariation in the reproductive success of coflowering species, thus identifying potential species interactions. Important strengths include explicit consideration of community-level effects and the assessment of residual fitness correlations after controlling for covariates such as flower abundances and phenotypic traits, yielding more complete insights into pollinator-mediated reproductive interactions.

Key-words: competition, facilitation, hierarchical modelling of species communities, natural selection, plant-pollinator interactions, selection gradient

Introduction

Species interactions within and among trophic levels play key roles in the assembly and dynamics of plant and insect communities. Most research addressing species interactions and coexistence in plant communities has focused on vegetative processes such as competition for light or nutrients, yet several authors have pointed out that plant coexistence may also depend on partitioning of biotic resources such as pollinators (Armbruster 1995; Sargent & Ackerly 2008; Pauw 2013; Bartomeus, Godoy & Rees 2018 and references therein). Variation in pollinator visitation can directly affect plant fitness, and access to pollination services can be seen as a limiting resource for animal-pollinated plants (the vast majority of flowering plants; Ollerton, Winfree & Tarrant 2011). Consequently, competition for access to pollinators may play an important role in long-term species coexistence in plant communities (Pauw 2013), and can also affect the probability of rare immigrants establishing in a community (Runquist & Stanton 2013). Understanding the role of pollinators in community assembly and species coexistence is key to predicting the consequences of ongoing changes in biotic communities, including the impact of species invasions (Schweiger *et al.* 2010) and declines in pollinator populations (Potts *et al.* 2010). As a step in this direction, we here propose and demonstrate how to use recently developed hierarchical joint models to probe data on the reproductive success of coflowering species with the aim of detecting potential reproductive interactions.

Pollination ecologists have long recognized that coflowering species may affect each other's reproductive success negatively if they compete for the same limiting pollinator resource, and positively if larger total floral abundances attract proportionally more pollinators (e.g. Waser 1978; Feinsinger 1987; Moeller 2004; Hegland, Grytnes & Totland 2009; Mitchell *et al.* 2009). We will refer to these interactions as reproductive interactions, to keep them separate from those interactions mediated e.g. by competition for access to nutrients (Armbruster 1995). Most flowering plants are visited by multiple pollinator species, and most pollinator species visit multiple plant species (e.g. Waser *et al.* 1996; Bascompte & Jordano 2007; Lázaro, Lundgren & Totland 2009). Negative reproductive interactions (competition) may therefore arise, for example, if coflowering species differ in abundance and shared pollinators prefer the more abundant species, or if flowers of different species differ in attractiveness. Positive (facilitative) reproductive interactions may arise through joint attraction of shared pollinators when the flower density of each individual species is too low to efficiently attract pollinators (Schemske 1981; Moeller 2004; Mesgaran *et al.* 2017; Losapio *et al.* 2019), or more generally when plant species

have mechanisms in place that alleviate the negative effects of pollinator sharing through avoidance or tolerance of heterospecific pollination (Hegland, Grytnes & Totland 2009; Hegland & Totland 2012; Ashman & Arceo-Gomez 2013).

If pollinator-mediated reproductive interactions affect plant fitness and the assembly of communities, their signatures should be detectable in patterns of reproductive success among coflowering plant species. Despite the long-standing interest in these issues, there have been relatively few attempts at detecting a signal of reproductive interactions in observational data from coflowering plant communities (but see Hegland, Grytnes & Totland 2009; Benadi & Pauw 2018). Here, we suggest that methods recently developed for analysing patterns of species cooccurrence can also be used to quantify potential reproductive interactions among coflowering plant species. Such approaches can enhance our understanding of general patterns of, and explanations for, ecological interactions. In the following, we first introduce the principles of the modelling approach, and discusses its advantages in studies of pollinator-mediated plant-plant interactions. We then demonstrate a specific application of the method by analysing a dataset on bumblebee visitation to a set of coflowering plant species in a species-rich meadow in Norway. Note that applications of the proposed methods are not limited to this data type, and we therefore provide tutorials for this and additional data types (Fig. 1) in the appendices. Finally, while our focus here is on pollinator-mediated interactions in plant communities, we note that our approaches can equally well be applied to analysing fitness data for other organisms.

Materials and Methods

Statistical analyses of reproductive interactions among coflowering plant species

Community ecologists have long sought to infer patterns of species interactions from observational data on species cooccurrence patterns (e.g. Diamond 1975; Connor & Simberloff 1979). One of the core ideas of these analyses is that if two species occur together more, or less, often than would be expected from their shared response to the abiotic environment, this can be taken as an indication of a facilitative or competitive process. Several model-based approaches have recently been developed to facilitate such analyses (e.g. Pollock *et al.* 2014; Hui 2016; Ovaskainen *et al.* 2017b). Following a similar logic, we can use data on the reproductive performance (fitness) of several species to ask whether their performance when occurring together differ from expectations. In the following, we will demonstrate the use of a recently developed statistical framework, Hierarchical Modelling of Species Communities (HMSC; Ovaskainen *et al.*

2017b), to analyze data on the reproductive success of coflowering species with the aim of detecting potential pollinator-mediated reproductive interactions.

The minimum data requirement for applying these methods is some measure of reproductive performance of at least two species occurring together in a set of observation units (e.g. study plots or pollinator-observation periods). Measures of reproductive performance may include number of pollinator visits, number of conspecific pollen grains deposited onto stigmas, and number of fruits or seeds set. The kind of data available will naturally dictate the kinds of questions that can be asked, and the kinds of inferences that can be made (see further discussion below).

The Hierarchical Modelling of Species Communities framework

The Hierarchical Modelling of Species Communities (HMSC) framework belongs to the class of 'joint models' (Warton *et al.* 2015), and allows modelling the effects of a set of covariates on multiple response variables. An advantage of these models is that they allow estimating shared responses to covariates among response variables, and to leverage this information to refine parameter estimation. In analyses of pollinator-mediated reproductive interactions, this allows asking, for example, whether plant species respond similarly to the abundance of a specific coflowering species, or whether pollinators systematically prefer specific phenotypic traits (e.g. larger flowers). Furthermore, joint models can be used to assess and quantify residual covariance among response variables after accounting for the effects of covariates (Ovaskainen *et al.* 2016a; Ovaskainen *et al.* 2017b). In the context of community analyses these residual associations may represent potential reproductive interactions among species, such as greater attractiveness of the flowers of some species over others.

In HMSC, the response to covariates (fixed part of linear predictor, L_{ij}^F) is modelled as the regression $L_{ij}^F = \sum_k x_{ik}\beta_{kj}$, where x_{ik} is the value of covariate k for observation i, and β_{kj} is the regression slope of response variable j on covariate k. To explicitly model joint responses to the covariates, the vector of regression coefficients for each response variable is assumed to adhere to a multivariate normal distribution as $\beta_j \sim \mathcal{N}(\mu, V)$, where μ is a column vector of mean (expected) regression coefficients and V is a variance-covariance matrix describing covariation in regression coefficients among response variables (Ovaskainen & Soininen 2011). The diagonal of the matrix V thus describes variation among species in their response to the covariates.

Residual associations among response variables are quantified through a latent-variable approach, where latent variables (random part of linear predictor, L_{ij}^R) can be specified at multiple hierarchical levels, e.g. observations within plots within sites (Ovaskainen *et al.* 2016a). The latent variables can be thought about as unobserved variables representing either unmeasured environmental variation or the influence of species interactions, and are modelled as $L_{ij}^R = \sum_h \eta_{ih} \lambda_{hj}$, where η_{ih} is the value of the unmeasured latent factor h for observation i, and λ_{hj} is the loading of response variable j onto latent factor h. The residual associations are quantified by variance-covariance matrices, which for latent variable R is denoted as $\boldsymbol{\Omega}^R = (\boldsymbol{\lambda}^R)^T \boldsymbol{\lambda}^R$, where $\boldsymbol{\lambda}^R$ represents the loadings of each response variable onto the latent variable R, and ^{T} denotes transposition. The strength of the associations among response variables (species in the analyses that follow) can be quantified by translating the covariances into correlations.

The HMSC model has been implemented in a Bayesian framework, has been shown to perform well in terms of parameter estimation and prediction (Tikhonov *et al.* 2019), and was ranked first in a recent comparison among similar models (Norberg *et al.* 2019). In the following we will demonstrate applications of the HMSC framework to data on flower visitation to a set of plant species at a plot level recorded during multiple censuses at each plot (Fig. 1a). We implemented the analyses with the HMSC-R 3.0 R package (Tikhonov *et al.* 2019) in R 3.5.0 (R Core Team 2018). A detailed tutorial including the R code reproducing the analyses reported in the main text, along with additional details, is available in Appendix S1. In Appendix S2 we use simulated data to demonstrate how to set up a HMSC model for data collected at the level of individual plants, including the measurement of phenotypic selection gradients (covariance of phenotypic traits and relative fitness; Lande & Arnold 1983) for multiple species simultaneously (Fig. 1b).

Case study: Plot-level analysis of pollinator visitation

As an example of an application of HMSC to data on pollinator visitation at a plot level, we reanalysed the data collected by Hegland, Grytnes and Totland (2009). Pollinator observations were made at 1.5×1.5 m plots during 10-min censuses ($n = 201$ censuses in 20 plots) in a species-rich meadow in Norway. During each census, an observer recorded the number of pollinator visits to each plant species flowering in the plot, and counted the total number of flowers of each species. Observations were made during a single growing-season (late May to August) largely under good weather conditions (no rain, little wind). Temperature measurements were made at the

beginning of each census. We focus here on visits by bumblebees to those species visited primarily (>50% of visits, often >80%) by bumblebees (9 species).

In these data, pollinator-mediated reproductive interactions may arise via an abundance effect, where more abundant species are more attractive to pollinators, or via a preference effect, where individual flowers of some species are more attractive to pollinators than are flowers of other species. A simple correlation analysis of pollinator visitation to each species would confound these processes. As an attempt to disentangle these mechanisms, we will fit a series of models analysing both the direct effect of the flower abundance of each species (including the focal species) on the number of pollinator visits to the focal species, and the residual covariance of pollinator visitation among species after controlling for any effects of flower abundances and other covariates. The residual covariances should capture covariation in pollinator visitation after controlling for the floral abundance of each species, and could represent, for example, differential attractiveness of individual flowers to pollinators.

We will consider two kinds of matrices describing species-to-species associations (Fig. 1a). The first, denoted in HMSC as $\mathbf{\Omega}$, is a symmetrical matrix describing the residual correlations of pollinator visitation to each species. Positive residual correlations indicate that both species tend to be visited within the same sampling units (pollinator censuses or plots), after accounting for the effects of covariates included in the model. The second, which we denote as $\mathbf{A}$, is a square matrix with elements $\mathbf{A}_{ij}$ describing the effect of the floral abundance of species j on the reproductive success of species i, with intraspecific effects on the diagonal, and interspecific effects on the off-diagonals.

Model 1: Latent variables and environmental covariates only

To illustrate the meaning of the latent variables in joint models, we started by fitting a model with temperature as the only covariate ($\mathbf{X}$, Fig. 1) so that the latent variables represent associations among species after controlling for effects of temperature on insect activity. In the HMSC model, we included in the response matrix ($\mathbf{Y}$, Fig. 1) the number of pollinator visits to each species, with *NA* for species not flowering in the focal plot during a census. We $\log_e(x+1)$-transformed the number of pollinator visits to place the response variables on a proportional scale and to reduce the leverage of large values. Plot ($n = 20$) and census ($n = 176$) were treated as hierarchical random levels represented by latent variables, allowing us to assess residual species associations at each of

these scales. In addition to potential reproductive interactions, these associations include, for example, any tendency for species to flower in the same plots or during the same censuses.

For this and the following models, we assumed the default priors in HMSC (Ovaskainen *et al.* 2017b). We sampled the posterior distribution with 2 replicate chains of 300,000 iterations each with the first 100,000 discarded as burnin, and confirmed convergence by computing effective sample sizes, potential scale reduction factors, and by visual assessment of posterior trace plots (see Appendix 1 for details). Even though the number of pollinator visits is a count variable, the residuals of the Gaussian models conformed to the assumption of normality.

Model 2: Conspecific flower abundances

Here, we assessed to what extent the number of conspecific flowers in a plot during a census predicts the overall number of visits to those flowers. To do so, we added to the model the number of flowers of the focal species in each plot during each census as a species-specific covariate. This corresponds to estimating the diagonal of the matrix **A** (Fig. 1a)**.** We $\log_e(x+1)$-transformed the flower abundances so that, because pollinator visits are also on a natural log scale, the regression parameters are on a proportional scale and are interpretable roughly as elasticities describing the percent change in pollinator visitation per percent change in flower abundance.

Model 3: Conspecific + heterospecific flower abundances

To assess whether and how the number of pollinator visits to a focal species depends on the flower abundances of the focal species as well as all coflowering species, and thus estimate the full matrix **A** (Fig. 1a), we included the $\log_e(x+1)$-transformed flower abundances of all species as covariates. To assess to what extent including all flower abundances led to overfitting of the model, we performed ten-fold cross-validation by sequentially re-training the model on 90% of the data, and making predictions for the remaining 10% of the data.

Model 4: Visitation rates, conspecific + heterospecific flower abundances

Finally, plant-level fitness may depend on the number of visits per individual flower, and it may therefore be of interest to know whether the number of visits to individual flowers depends on the abundances of conspecific and heterospecific flowers. Thus, we fitted a model similar to Model 3, but with the number of visits translated into visitation rates (i.e. number of visits divided by number of flowers). The intraspecific effects in this model represent changes in the per-flower visitation rate with increasing flower abundance, thus positive effects can be interpreted as positive density dependence, and negative effects as negative density dependence.

Results

Model 1: Latent variables and environmental covariates only

On average across species, the temperature-only model explained 27% of the variation in pollinator visitation, of which temperature explained 41%, and more variation was explained by the plot-level latent variable than by the census-level latent variable (Fig. 2). We detected both positive and negative associations for pollinator visitation at the plot level (Fig. 3a), but not at the census level (Fig. 3b). For example, pollinator visitation to *Campanula rotundifolia* was positively associated with pollinator visitation to *Euphrasia stricta* and *Prunella vulgaris*, and negatively associated with pollinator visitation to *Hypericum maculatum*.

Model 2: Conspecific flower abundances

Adding the conspecific flower abundances as covariates led to a limited increase in the explanatory power of the model ($r^2 = 0.34$ vs. 0.27, Fig. 2), indicating that the latent factors in the temperature-only model (Model 1) explained a substantial amount of the variance attributed to conspecific flower abundances. Pollinator visitation to all species increased with increasing conspecific flower abundance (Fig. 4), with the strongest effect for *Centaurea jacea* (beta = 0.74 log visits per log flower, 95% CI = 0.28 – 1.25), and the weakest effect for *E. stricta* (beta = 0.27 log visits per log flower, 95% CI = 0.16 – 0.40).

When controlling for conspecific flower abundances as well as temperature we detected no strong residual associations for pollinator visitation. This is also evident from the limited variation in pollinator visitation explained by the latent variables in Model 2 (Fig. 2).

Model 3: Conspecific + heterospecific flower abundances

The full model explained 43% of the variation in pollinator visitation, of which heterospecific flower abundances explained 33.8% (Fig. 2). The relative contributions of conspecific vs. heterospecific flower abundances differed among species, with the greatest contribution of heterospecific flower abundances observed for *Knautia arvensis*. As expected for models including many explanatory variables, predictive power as quantified by ten-fold cross-validation was lower than the explanatory power (mean $r^2_{CV} = 0.19$).

While the covariance matrix Ω is symmetrical the **A** matrix is not, allowing some insights into the directionality of observed associations. Among the nine focal species, we detected cases of mutual positive or negative effects, unidirectional effects, and opposite effects (Fig. 5). Importantly, while we have chosen for visual clarity to plot a subset of well-supported parameters,

the interpretation of potential interactions for specific species pairs should be based on the parameter estimates with their associated uncertainty (Table 1). The previously observed associations of pollinator visitation involving *C. rotundifolia* appears to occur in both directions at similar strength. For example, visitation to *H. maculatum* tended to decrease with greater *C. rotundifolia* flower abundance (beta = -0.10 log visits per log flower, 95% CI = -0.30 – 0.09), and *vice versa* (beta = -0.11 log visits per log flower, 95% CI = -0.27 – 0.06). In contrast, visitation to *E. stricta* tended to decrease with greater *H. maculatum* abundance (beta = -0.11 log visits per log flower, 95% CI = -0.26 – 0.05), but the opposite was not true. At the community level, co-flowering species responded rather similarly to variation in the flower abundance of *Euphrasia stricta*, while responses to variation in the flower abundance of *Knautia arvensis* were the most variable (Table 1). Visitation to all species but one tended to decrease when *Trifolium repens* was more abundant.

Model 4: Visitation rates, conspecific + heterospecific flower abundances
The visitation-rate model explained 26% of the variance in visitation rates, of which 62.2% was explained by heterospecific flower abundances, and 11.2% by conspecific flower abundances. Compared to Model 3, the effects of conspecific flower abundances largely disappeared for seven species, remained positive for *C. rotundifolia* (suggesting positive density dependence), and became negative for *Trifolium pratense* (suggesting negative density dependence; Fig. 6).

Interspecific effects remained qualitatively similar (i.e. had the same sign) between the two models in most cases (79.2%), and the parameter estimates were strongly correlated between models (*r* = 0.82). However, the posterior support for individual parameters (i.e. the proportion of posterior samples that were positive, or negative in the case of negative parameter estimates) changed in some cases, qualitatively supporting a partly different set of potential interactions (Fig. 6). This underlines the importance of interpreting parameter estimates rather than assigning positive and negative interactions based on qualitative criteria.

Discussion

Detecting reproductive interactions from observational data
Pollinator-mediated reproductive interactions among coflowering plant species are challenging to detect from observational data. An important reason for this is that the effect of one species on another may depend in complex ways on the entire coflowering community. While most analyses of pollinator-mediated reproductive interactions have focused on individual focal species or

species pairs (e.g. Waser 1978; Armbruster & McGuire 1991; Moeller 2004; Ghazoul 2006; Runquist & Stanton 2013), the complex nature of plant-pollinator interactions suggests that deeper insights may arise from joint analyses of multiple interacting species (Hegland, Grytnes & Totland 2009). Importantly, variation in the flower abundance of a species favoured by pollinators may affect multiple coflowering species in a similar way. One advantage of hierarchical joint models when analysing potential reproductive interactions among multiple species is that joint models explicitly consider similarity among species in their responses to covariates, thus improving parameter estimation and predictive power especially for less common species (Ovaskainen & Soininen 2011; Warton *et al.* 2015; Ovaskainen *et al.* 2017b; Norberg *et al.* 2019). Explicit modelling of species' joint responses to their coflowering community can therefore increase the power to detect community-wide patterns of reproductive interactions by leveraging information across all species, such as the consistent negative effect of *Trifolium repens* on most co-flowering species observed in the case study. Conversely, some species may have variable, species-specific effects on coflowering species, as observed for *Knautia arvensis*, suggesting additional factors affecting pollinator behaviour (e.g. differences in floral traits). Finally, we note that while joint models generally perform well for detecting community-level patterns, they do not necessarily increase the power to detect particular associations for particular species (see Norberg et al. 2019 for a similar discussion of single vs. joint species distribution models).

Hegland, Grytnes and Totland (2009) analysed the data used in the case study by fitting univariate models to visitation rates of each species separately. Quantitative comparison between the two studies is difficult due to several differences in the modelling approach and pre-treatment of data, yet we can compare qualitatively the results obtained for the effects of flower abundances on visitation rates (Model 4 in the case study). Of the 17 detectable associations reported by Hegland, Grytnes and Totland (2009), 13 (76.5%) were positive. In the present analysis, eight of the 14 associations (57.1%) with at least 85% posterior support were positive. There are also some differences in the identities of the associated species (compare Fig. 6 vs. Table 1 in Hegland, Grytnes and Totland 2009). However, our reanalysis supports the original conclusion that not only negative but also positive reproductive interactions through pollinator attraction may be rather common among coflowering plant species.

A second advantage of joint models over multiple single-species analyses of reproductive performance is that these models allow estimating residual covariances after controlling for covariates such as flower abundances or floral traits (see Appendix S2). Residual covariation in

pollinator visitation may arise, for example, if pollinators prefer some species over others due to differences in floral advertisements and rewards. In the case study, the residual associations observed in the simple temperature-only model largely disappeared when adding conspecific and heterospecific flower abundances to the model, suggesting that such mechanisms are of limited importance in the focal community. However, this insight could not have been directly obtained from univariate analyses.

Residual species associations such as those quantified by the Ω matrix of HMSC are essentially correlations and cannot directly be interpreted as causal. For example, a positive residual association may represent a joint response to unmeasured aspects of the environment, such as variation in soil nutrients affecting nectar production and quality. Our **A** matrix is based on regressions rather than correlations and therefore goes some way towards disentangling the directionality of raw associations, as demonstrated in the case study. However, the input data is still observational, and we urge researchers to consider any associations detected using these methods as hypotheses to be tested, rather than strong causal results. A natural next step will be to test these hypotheses using experimental approaches. For example, a hypothesized negative association could be tested by generating experimental arrays containing various proportions of two species (Ghazoul 2006; Runquist & Stanton 2013), or by experimentally removing species and assessing how patterns of species associations change (Hegland & Totland 2012; Biella *et al.* 2019; Losapio *et al.* 2019).

Community-level phenotypic-selection analyses
The case study illustrates one application of joint models to study reproductive interactions, yet the approach is not limited to this specific data type. Data on reproductive success (e.g. number of seeds produced) combined with data on phenotypic traits are commonly used to estimate phenotypic selection gradients following the methods of Lande and Arnold (1983). A few studies have assessed how the presence of coflowering species affects pollinator-mediated selection on a focal species (Caruso 2000; Wassink & Caruso 2013; Parachnowitsch, Cook-Patton & McArt 2014). By similar arguments as those made above, we suggest that joint models provide a powerful approach for studying selection on multiple species simultaneously. In Appendix S2, we demonstrate how to set up a HMSC model to simulated data on the fitness of individual plants of multiple species (Fig. 1b). This method allows estimating selection gradients for each species in

addition to species-to-species associations, and would be a promising approach towards understanding patterns of selection at the community level.

Considerations for choice of response variable type

Data on visitation to the flowers of coflowering species such as those analysed in the case study are relatively easy to collect in many systems. However, while visitation data allow analyses of interactions mediated by floral-visitor attraction, they do not necessarily yield insights into consequences for reproductive success. A crucial point here is the separation between visitation and pollination, or floral visitors and pollinators, where pollinators are only those visitors that effect pollen transfer (e.g. Armbruster 1988).

Data on pollen arrival onto stigmas allows further insights into the dynamics of pollen transfer, thus overcoming uncertainty in the efficiency of pollen transfer. Such analyses can also be informative about the possible consequences of heterospecific pollination, that may substantially reduce the fitness advantage of joint pollinator attraction (Ashman & Arceo-Gomez 2013; Arceo-Gomez *et al.* 2019). For example, a rare species may benefit from large heterospecific flower displays in its surroundings in terms of number of pollinator visits, but this advantage may be greatly reduced if most pollen grains arriving onto stigmas are heterospecific ('quality' component of competition for pollination, see Mitchell *et al.* 2009). Analyses of stigma pollen loads may become increasingly feasible through the ongoing development of state-of-the-art approaches such as pollen metabarcoding (Bell *et al.* 2016) and labelling of pollen grains with quantum dots (Minnaar & Anderson 2019).

Finally, data on fruit or seed set might yield even deeper insights into the fitness consequences of pollinator sharing. Note, however, that interactions detected from fruit- or seed-set data alone cannot necessarily be attributed to pollination processes, because differences in seed set can arise from other processes such as variation in maternal resource levels and thus represent e.g. species interactions through vegetative processes. This problem is analogous to the issue of separating pollinator-mediated phenotypic selection from other sources of variation in seed set, and can be resolved by combining the analytical approaches demonstrated here with experimental approaches such as hand-pollination treatments (e.g. Sletvold & Ågren 2014), or by combining analyses of pollinator visitation, pollen loads, and seed set (e.g. Pérez-Barrales *et al.* 2013).

Extensions of model structure

The flexible HMSC framework makes it straightforward to extend the methods demonstrated here by, for example, adding additional hierarchical random levels (say, plots nested within sites nested with landscapes). Such analyses can be informative about the spatial scale of reproductive interactions, because a given species pair could, for example, be positively associated at the level of sites due to joint attraction of more pollinators, but negatively associated at the plot-scale due to pollinator preference for one species over the other (see also Albor *et al.* 2019). If geographic coordinates of plots are available, the model can be fitted with spatially explicit latent factors (see Ovaskainen *et al.* 2016b for discussion and technical details), which would allow asking whether species associations are scale-dependent (Hegland 2014). If data are collected in several environments, such as along altitudinal gradients, HMSC includes methods for estimating how residual associations vary along environmental gradients (Tikhonov *et al.* 2017). Finally, it is technically straightforward to include multiple fitness proxies for each plant species, e.g. both pollinator visitation and seed set, or visitation frequencies by multiple pollinator taxa.

Error distributions, link functions and regression equations can be easily adjusted to accommodate more complex relationships between covariates (e.g. flower abundances or phenotypic traits) and response variables. For example, adding square terms would allow assessing frequency-dependent effects of flower abundances on pollinator visitation (Feinsinger 1987; Benadi & Pauw 2018), and estimating quadratic selection gradients (Lande & Arnold 1983). Similarly, adding interaction terms could allow assessing more complex patterns of multispecies interactions, and estimating correlational selection. Because our method for estimating the **A** matrix is based on multiple-regressions with all species included as predictors, complex models fitted to limited data from large communities may suffer from overfitting. While the main aim in many cases may be to explain patterns of reproductive interactions, predictions based on such models will be less accurate, as demonstrated by cross-validating Model 3 in the case study. The number of species and covariates included in the model should therefore be limited in analyses of small data sets if a main aim of the analysis is to make predictions for novel sites. Note that parameter estimation for the species associations ($\mathbf{\Omega}$ matrix) are based on dimension-reduction implemented through a latent-variable approach (Ovaskainen *et al.* 2016a), and overfitting should not be an issue for simpler models such as Model 1 and 2 in the case study. When modelling large communities, a similar approach could be taken for the fixed covariates. Dimension reduction for fixed covariates (reduced-rank regression) in HMSC has been developed in the context of time-series analyses (Ovaskainen *et al.* 2017a), and could also be applied to the kinds of analyses

discussed here. Such techniques could also be useful for measuring phenotypic selection on high-dimensional phenotypes (Chong, Fung & Stinchcombe 2018).

Conclusions

Ongoing development of statistical methods in community ecology now allows students of biotic communities to ask more nuanced questions than have been possible in the past. Here we have shown how community-level analyses can yield a more complete understanding of pollinator-mediated reproductive interactions through explicit consideration of community-level responses to coflowering communities, and through assessment of residual fitness correlations. The flexible nature of these methods makes them widely applicable to related questions in other study systems and to different data types.

Acknowledgements

Authors' contributions

ØO conceived the ideas, developed the methods, analysed data, and wrote the first draft of the manuscript. SJH collected data and contributed to the writing of the manuscript.

Data Accessibility

All data and R code is available from GitHub (github.com/oysteiop/PollinatorInteractions). DOI: 10.5281/zenodo.3462048

References

Albor, C., García-Franco, J.G., Parra-Tabla, V., Díaz-Castelazo, C. & Arceo-Gómez, G. (2019) Taxonomic and functional diversity of the co-flowering community differentially affect *Cakile edentula* pollination at different spatial scales. *Journal of Ecology,* **107,** 2167-2181.

Arceo-Gomez, G., Kaczorowski, R.L., Patel, C. & Ashman, T.L. (2019) Interactive effects between donor and recipient species mediate fitness costs of heterospecific pollen receipt in a co-flowering community. *Oecologia,* **189,** 1041-1047.

Armbruster, W.S. (1988) Multilevel comparative analysis of the morphology, function, and evolution of *Dalechampia* blossoms. *Ecology,* **69,** 1746-1761.

Armbruster, W.S. (1995) The origins and detection of plant community structure: reproductive versus vegetative processes. *Folia Geobotanica & Phytotaxonomica,* **30,** 483-497.

Armbruster, W.S. & McGuire, A.D. (1991) Experimental assessment of reproductive interactions between sympatric *Aster* and *Erigeron* (Asteraceae) in interior Alaska. *American Journal of Botany,* **78,** 1449-1457.

Ashman, T.L. & Arceo-Gomez, G. (2013) Toward a predictive understanding of the fitness costs of heterospecific pollen receipt and its importance in co-flowering communities. *Am J Bot,* **100,** 1061-1070.

Bartomeus, I., Godoy, O. & Rees, M. (2018) Biotic controls of plant coexistence. *Journal of Ecology,* **106,** 1767-1772.

Bascompte, J. & Jordano, P. (2007) Plant-Animal Mutualistic Networks: The Architecture of Biodiversity. *Annual Review of Ecology, Evolution, and Systematics,* **38,** 567-593.

Bell, K.L., de Vere, N., Keller, A., Richardson, R.T., Gous, A., Burgess, K.S. & Brosi, B.J. (2016) Pollen DNA barcoding: current applications and future prospects. *Genome,* **59,** 629-640.

Benadi, G. & Pauw, A. (2018) Frequency dependence of pollinator visitation rates suggests that pollination niches can allow plant species coexistence. *Journal of Ecology,* **106,** 1892-1901.

Biella, P., Akter, A., Ollerton, J., Tarrant, S., Janeček, Š., Jersáková, J. & Klecka, J. (2019) Experimental loss of generalist plants reveals alterations in plant-pollinator interactions and a constrained flexibility of foraging. *Sci Rep,* **9**.

Caruso, C.M. (2000) Competition for pollination influences selection on floral traits of *Ipomopsis aggregata*. *Evolution,* **54,** 1546-1557.

Chong, V.K., Fung, H.F. & Stinchcombe, J.R. (2018) A note on measuring natural selection on principal component scores. *Evol Lett,* **2,** 272-280.

Connor, E.F. & Simberloff, D. (1979) The assembly of species communities - chance or competition. *Ecology,* **60,** 1132-1140.

Diamond, J.M. (1975) Assembly of species communities. *Ecology and evolution of communities* (eds M.L. Cody & J.M. Diamond), pp. 342–444. Harvard University Press, Cambridge, MA.

Feinsinger, P. (1987) Effects of plant species on each other's pollination: is community structure influenced? *Trends in Ecology & Evolution,* **2,** 123-126.

Ghazoul, J. (2006) Floral diversity and the facilitation of pollination. *Journal of Ecology,* **94,** 295-304.

Hegland, S.J. (2014) Floral neighbourhood effects on pollination success in red clover are scale-dependent. *Functional Ecology,* **28,** 561-568.

Hegland, S.J., Grytnes, J.A. & Totland, Ø. (2009) The relative importance of positive and negative interactions for pollinator attraction in a plant community. *Ecological Research,* **24,** 929-936.

Hegland, S.J. & Totland, Ø. (2012) Interactions for pollinator visitation and their consequences for reproduction in a plant community. *Acta Oecologica,* **43,** 95-103.

Hui, F.K.C. (2016) BORAL - Bayesian ordination and regression analysis of multivariate abundance data in R. *Methods in Ecology and Evolution,* **7,** 744-750.

Lande, R. & Arnold, S.J. (1983) The measurement of selection on correlated characters. *Evolution,* **37,** 1210-1226.

Lázaro, A., Lundgren, R. & Totland, Ø. (2009) Co-flowering neighbors influence the diversity and identity of pollinator groups visiting plant species. *Oikos,* **118,** 691-702.

Losapio, G., Fortuna, M.A., Bascompte, J., Schmid, B., Michalet, R., Neumeyer, R., Castro, L., Cerretti, P., Germann, C., Haenni, J.-P., Klopfstein, S., Ortiz-Sanches, F.J., Pont, A.C., Rousse, P., Schmid, J., Sommaggio, D. & Schöb, C. (2019) Plant interactions shape pollination networks via nonadditive effects. *Ecology,* **100,** e02619.

Mesgaran, M.B., Bouhours, J., Lewis, M.A. & Cousens, R.D. (2017) How to be a good neighbour: facilitation and competition between two co-flowering species. *Journal of Theoretical Biology,* **422,** 72-83.

Minnaar, C. & Anderson, B. (2019) Using quantum dots as pollen labels to track the fates of individual pollen grains. *Methods in Ecology and Evolution,* **10,** 604-614.

Mitchell, R.J., Flanagan, R.J., Brown, B.J., Waser, N.M. & Karron, J.D. (2009) New frontiers in competition for pollination. *Ann Bot,* **103,** 1403-1413.

Moeller, D.A. (2004) Facilitative interactions among plants via shared pollinators. *Ecology,* **85,** 3289-3301.

Norberg, A., Abrego, N., Blanchet, F.G., Adler, F.R., Anderson, B.J., Anttila, J., Araújo, M.B., Dallas, T., Dunson, D., Elith, J., Foster, S.D., Fox, R., Franklin, J., Godsoe, W., Guisan, A., O'Hara, B., Hill, N.A., Holt, R.D., Hui, F.K.C., Husby, M., Kålås, J.A., Lehikoinen, A., Luoto, M., Mod, H.K., Newell, G., Renner, I., Roslin, T., Soininen, J., Thuiller, W., Vanhatalo, J., Warton, D., White, M., Zimmermann, N.E., Gravel, D. & Ovaskainen, O. (2019) A comprehensive evaluation of predictive performance of 33 species distribution models at species and community levels. *Ecological Monographs,* **89**.

Ollerton, J., Winfree, R. & Tarrant, S. (2011) How many flowering plants are pollinated by animals? *Oikos,* **120,** 321-326.

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

Ovaskainen, O., Roy, D.B., Fox, R. & Anderson, B.J. (2016b) Uncovering hidden spatial structure in species communities with spatially explicit joint species distribution models. *Methods in Ecology and Evolution,* **7,** 428-436.

Ovaskainen, O. & Soininen, J. (2011) Making more out of sparse data: hierarchical modeling of species communities. *Ecology,* **92,** 289-295.

Ovaskainen, O., Tikhonov, G., Dunson, D., Grotan, V., Engen, S., Saether, B.E. & Abrego, N. (2017a) How are species interactions structured in species-rich communities? A new method for analysing time-series data. *Proceedings of the Royal Society B-Biological Sciences,* **284,** 20170768.

Ovaskainen, O., Tikhonov, G., Norberg, A., Guillaume Blanchet, F., Duan, L., Dunson, D., Roslin, T. & Abrego, N. (2017b) How to make more out of community data? A conceptual framework and its implementation as models and software. *Ecol Lett,* **20,** 561-576.

Parachnowitsch, A.L., Cook-Patton, S.C. & McArt, S.H. (2014) Neighbours matter: natural selection on plant size depends on the identity and diversity of the surrounding community. *Evolutionary Ecology,* **28,** 1139-1153.

Pauw, A. (2013) Can pollination niches facilitate plant coexistence? *Trends Ecol Evol,* **28,** 30-37.

Pérez-Barrales, R., Bolstad, G.H., Pélabon, C., Hansen, T.F. & Armbruster, W.S. (2013) Pollinators and seed predators generate conflicting selection on *Dalechampia* blossoms. *Oikos,* **122,** 1411-1428.

Pollock, L.J., Tingley, R., Morris, W.K., Golding, N., O'Hara, R.B., Parris, K.M., Vesk, P.A., McCarthy, M.A. & McPherson, J. (2014) Understanding co-occurrence by modelling species simultaneously with a joint species distribution model (JSDM). *Methods in Ecology and Evolution,* **5,** 397-406.

Potts, S.G., Biesmeijer, J.C., Kremen, C., Neumann, P., Schweiger, O. & Kunin, W.E. (2010) Global pollinator declines: trends, impacts and drivers. *Trends in Ecology & Evolution,* **25,** 345-353.

R Core Team (2018) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Runquist, R.B. & Stanton, M.L. (2013) Asymmetric and frequency-dependent pollinator-mediated interactions may influence competitive displacement in two vernal pool plants. *Ecol Lett,* **16,** 183-190.

Sargent, R.D. & Ackerly, D.D. (2008) Plant-pollinator interactions and the assembly of plant communities. *Trends Ecol Evol,* **23,** 123-130.

Schemske, D.W. (1981) Floral convergence and pollinator sharing in 2 bee-pollinated tropical herbs. *Ecology,* **62,** 946-954.

Schweiger, O., Biesmeijer, J.C., Bommarco, R., Hickler, T., Hulme, P.E., Klotz, S., Kuhn, I., Moora, M., Nielsen, A., Ohlemuller, R., Petanidou, T., Potts, S.G., Pysek, P., Stout, J.C., Sykes, M.T., Tscheulin, T., Vila, M., Walther, G.R., Westphal, C., Winter, M., Zobel, M. & Settele, J. (2010) Multiple stressors on biotic interactions: how climate change and alien species interact to affect pollination. *Biol Rev Camb Philos Soc,* **85,** 777-795.

Sletvold, N. & Ågren, J. (2014) There is more to pollinator-mediated selection than pollen limitation. *Evolution,* **68,** 1907-1918.

Tikhonov, G., Abrego, N., Dunson, D., Ovaskainen, O. & Warton, D. (2017) Using joint species distribution models for evaluating how species-to-species associations depend on the environmental context. *Methods in Ecology and Evolution,* **8,** 443-452.

Tikhonov, G., Opedal, Ø.H., Abrego, N., Lehikoinen, A. & Ovaskainen, O. (2019) Joint species distribution modelling with HMSC-R. *bioRxiv,* **doi: https://doi.org/10.1101/603217**.

Warton, D.I., Blanchet, F.G., O'Hara, R.B., Ovaskainen, O., Taskinen, S., Walker, S.C. & Hui, F.K. (2015) So many variables: joint modeling in community ecology. *Trends Ecol Evol,* **30,** 766-779.

Waser, N.M. (1978) Competition for hummingbird pollination and sequential flowering in two Colorado wildflowers. *Ecology,* **59,** 934-944.

Waser, N.M., Chittka, L., Price, M.V., Williams, N.M. & Ollerton, J. (1996) Generalization in pollination systems, and why it matters. *Ecology,* **77,** 1043-1060.

Wassink, E. & Caruso, C.M. (2013) Effect of coflowering Mimulus ringens on phenotypic selection on floral traits of gynodioecious Lobelia siphilitica. *Botany-Botanique,* **91,** 745-751.

Table 1. Parameter estimates with posterior support values in parentheses for the effect of the log(flower abundance) of the species given in rows on the log(pollinator visits) to the species given in columns. Posterior support is the proportion of posterior samples that are positive or negative (for positive and negative parameter estimates, respectively).

	S1	S2	S3	S4	S5	S6	S7	S8	S9
Campanula rotundifolia (S1)	0.51	0.01	-0.02	0.11	-0.1	0.08	0.11	0.12	0.09
	(1.00)	(0.50)	(0.60)	(0.93)	(0.86)	(0.74)	(0.86)	(0.94)	(0.77)
Centaurea jacea (S2)	-0.05	0.70	-0.13	-0.01	-0.03	0.53	0.17	0.08	0.11
	(0.66)	(1.00)	(0.75)	(0.54)	(0.57)	(0.98)	(0.78)	(0.70)	(0.63)
Clinopodium vulgare (S3)	0.07	-0.06	0.63	0.00	-0.04	-0.04	-0.01	0.00	-0.05
	(0.85)	(0.70)	(1.00)	(0.52)	(0.69)	(0.65)	(0.56)	(0.50)	(0.67)
Euphrasia stricta (S4)	0.07	0.02	-0.07	0.31	0.01	0.06	0.01	0.06	0.03
	(0.93)	(0.59)	(0.92)	(1.00)	(0.60)	(0.74)	(0.54)	(0.90)	(0.63)
Hypericum maculatum (S5)	-0.11	-0.17	0.00	-0.11	0.49	-0.09	-0.05	0.03	0.07
	(0.91)	(0.89)	(0.52)	(0.91)	(1.00)	(0.73)	(0.66)	(0.67)	(0.72)
Knautia arvensis (S6)	0.04	-0.41	-0.31	0.10	-0.10	0.47	-0.33	-0.25	-0.11
	(0.59)	(0.85)	(0.84)	(0.65)	(0.62)	(0.98)	(0.82)	(0.88)	(0.64)
Prunella vulgaris (S7)	0.11	-0.14	0.10	0.07	-0.07	0.06	0.52	-0.01	0.02
	(0.95)	(0.78)	(0.90)	(0.82)	(0.78)	(0.62)	(1.00)	(0.58)	(0.58)
Trifolium pratense (S8)	-0.03	0.08	0.16	-0.10	-0.04	0.05	0.10	0.45	-0.04
	(0.66)	(0.68)	(0.96)	(0.89)	(0.68)	(0.65)	(0.81)	(1.00)	(0.63)
Trifolium repens (S9)	0.00	-0.38	-0.04	-0.11	-0.03	-0.12	-0.09	-0.03	0.34
	(0.51)	(0.90)	(0.63)	(0.87)	(0.59)	(0.78)	(0.82)	(0.63)	(1.00)

Figure legends

Fig. 1. Examples of how hierarchical joint models can be applied to study pollinator-mediated reproductive interactions among coflowering plant species. Pollinator foraging choices will often depend on the relative abundances of plant species (a). These effects can be studied by modelling the reproductive success ('fitness', as measured e.g. by the number of pollinator visits, pollen deposition onto stigmas, or seed set) of multiple species as a function of the flower abundances of coflowering species within an observation unit such as a plot. Here, direct effects of flower abundances can be quantified by regression coefficients summarized in matrix **A**, and residual effects through e.g. differential attractiveness of individual flowers through the residual association matrix **Ω**. Pollinator foraging choices can also depend on the phenotypic traits of individuals, such as plant height and flower size (b). These effects can be studied by modelling fitness as a function of the phenotypic traits of individual plants, thus estimating phenotypic selection gradients (β). In this case, residual correlations quantify potential effects of coflowering species on the fitness of the focal species, after controlling for any selection acting on traits included in the model.

Fig. 2. Variance partitioning for bumblebee visitation to each of nine coflowering plant species. The contribution of each variance component is given as a proportion of the total variance explained by the model, given by r^2 values. Model 1 (mean $r^2 = 27\%$) includes only the linear and square terms for temperature as fixed effects, model 2 (mean $r^2 = 34\%$) includes in addition the number of conspecific flowers in each observation unit (census), and model 3 (mean $r^2 = 43\%$) includes also the number of flowers of all coflowering species.

Fig. 3: Residual correlations for bumblebee visitation to nine coflowering plant species after controlling for the effect of temperature. Redder colours indicate positive associations, i.e. that species tend to be visited in the same plots, and bluer colours indicate negative associations, i.e. that species tend to be visited in different plots. For visual clarity, associations with less than 75% posterior support are set to zero.

Fig. 4: Effects of conspecific flower abundance on bumblebee visitation. Colours indicate effect sizes interpretable roughly as elasticities, where a value of 1 means that an increase in flower abundance of 1% increases pollinator visitation by 1%. All parameter estimates have >95% posterior support.

Fig. 5. Summary of inferred effects of intra- (diagonal) and interspecific (off-diagonals) flower abundances on bumblebee visitation, as indicated by posterior support for positive or negative effects. Colours indicate the posterior support for effects of flower abundance of the species on the *x*-axis on bumblebee visitation to the species on the *y*-axis. For visual clarity, effects with less than 85% posterior support are set to zero. See Table 1 for parameter estimates.

Fig. 6. Summary of inferred effects of intra- (diagonal) and interspecific (off-diagonals) flower abundances on bumblebee visitation rate, as indicated by posterior support for positive or negative effects. Colours indicate the posterior support for effects of flower abundance of the species on the *x*-axis on the rate of bumblebee visitation to flowers of the species on the *y*-axis. For visual clarity, effects with less than 85% posterior support are set to zero.

Supporting information

Appendix 1. R tutorial including code for reproducing the analyses of plot-level visitation data reported in the main text of the paper.

Appendix 2. R tutorial including example analysis of simulated data on individual fitness of coflowering plants.

Ecological context

(a) Variation in coflowering community   jec_13301_f1.pdf

(b) Variation in phenotypic traits
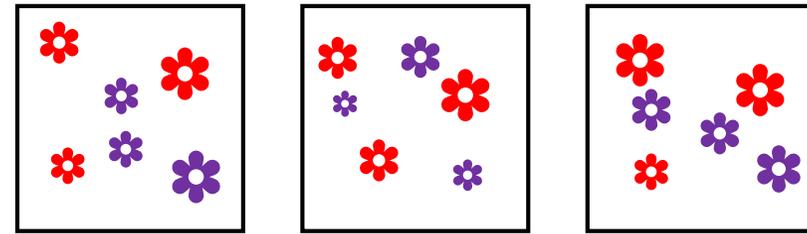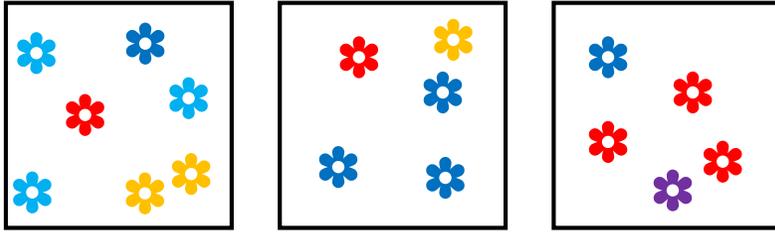
HMSC input data

Response matrix

Focal species fitness — Y (Observation unit)

Covariate matrix

Flower abundances — X (Observation unit)

Response matrix

Fitness — Y (Focal individual)

Covariate matrix

Phenotypic trait — X (Focal individual)
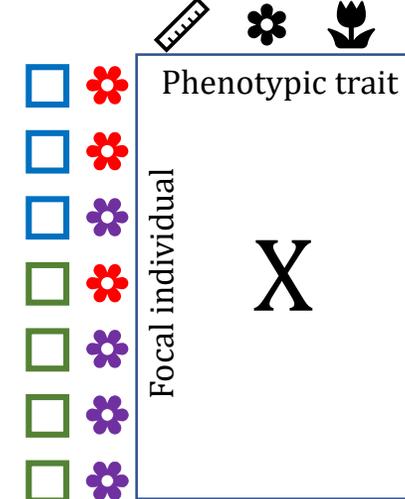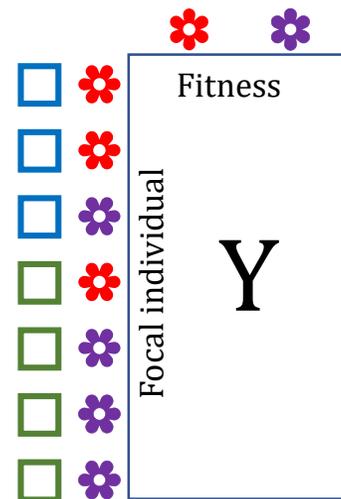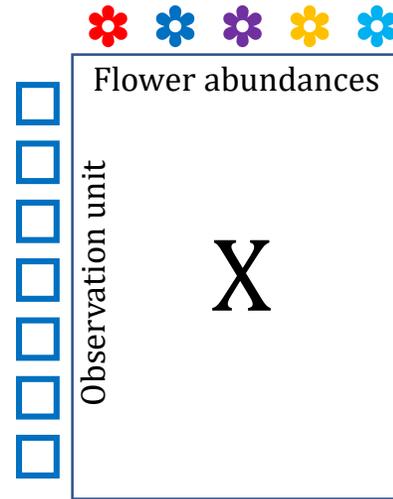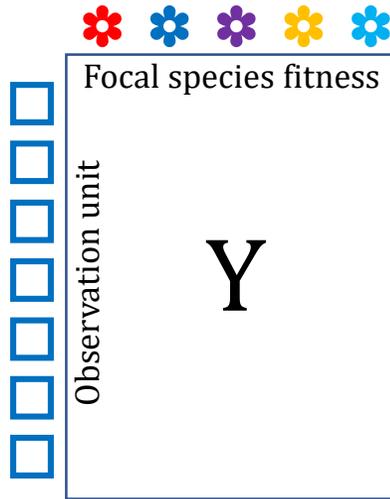
HMSC output

Regression coefficients

Focal species fitness — A (Flower abund.)

Residual correlations

Ω

Regression coefficients

Fitness — β (Phenotypic trait)

Residual correlations

Ω
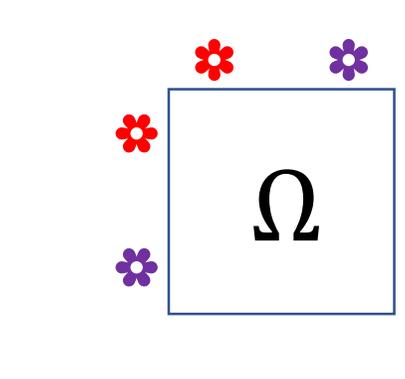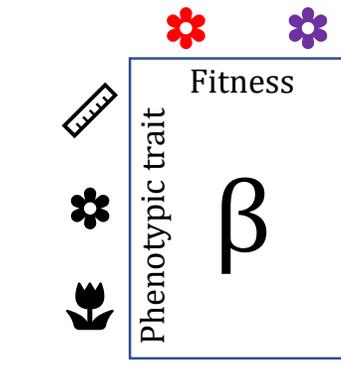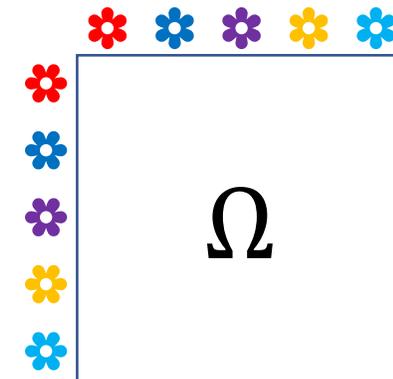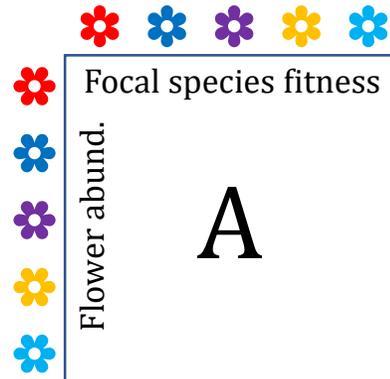
Model 1: Temperature only

Legend:
- Random: Census (mean = 24)
- Random: Plot (mean = 35)
- Temperature (mean = 41)

Species (x-axis): Campanula rotundifolia, Centaurea jacea, Clinopodium vulgare, Euphrasia stricta, Hypericum maculatum, Knautia arvensis, Prunella vulgaris, Trifolium pratense, Trifolium repens

Y-axis: Variance proportion

Model 2: + Conspecific flowers

Legend:
- Random: Census (mean = 7.3)
- Random: Plot (mean = 6.6)
- Temperature (mean = 12)
- Conspecific flowers (mean = 74.1)

Species (x-axis): Campanula rotundifolia, Centaurea jacea, Clinopodium vulgare, Euphrasia stricta, Hypericum maculatum, Knautia arvensis, Prunella vulgaris, Trifolium pratense, Trifolium repens

Y-axis: Variance proportion

Model 3: + Heterospecific flowers

Legend:
- Random: Census (mean = 4.1)
- Random: Plot (mean = 3.9)
- Temperature (mean = 6.5)
- Conspecific flowers (mean = 51.7)
- Heterospecific flowers (mean = 33.8)

Species (x-axis): Campanula rotundifolia, Centaurea jacea, Clinopodium vulgare, Euphrasia stricta, Hypericum maculatum, Knautia arvensis, Prunella vulgaris, Trifolium pratense, Trifolium repens

Y-axis: Variance proportion
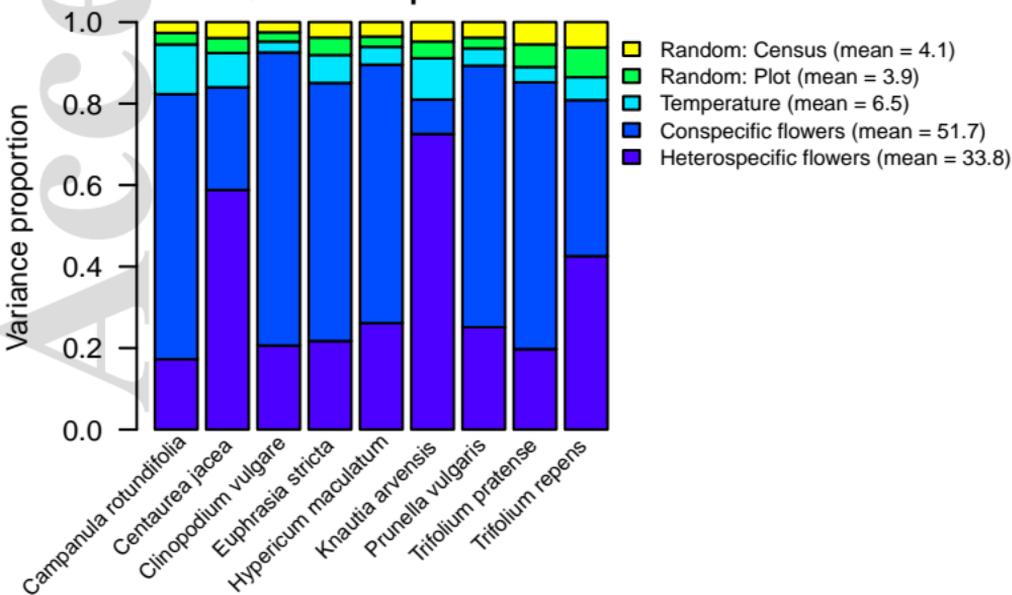
Plot level   jec_13301_f3.pdf   Census level