

Relating geographical variation in pollination types to environmental and spatial factors using novel statistical methods

Ingolf Kühn^{1,2}, Stijn Martinus Bierman³, Walter Durka¹ and Stefan Klotz^{1,2}

¹UFZ – Centre for Environmental Research Leipzig-Halle, Department of Community Ecology (BZF), Theodor-Lieser-Strasse 4, D–06120 Halle, Germany;

²Virtual Institute Macroecology, Theodor-Lieser-Strasse 4, D–06120 Halle, Germany; ³BioSS – Biomathematics and Statistics Scotland, James Clerk Maxwell Building, The King's Building, Edinburgh EH9 3JZ, UK

Summary

Author for correspondence:

Ingolf Kühn

Tel: +49 345 5585311

Fax: +49 345 5585329

Email: ingolf.kuehn@ufz.de

Received: 24 March 2006

Accepted: 8 May 2006

- The relative frequencies of functional traits of plant species show notable spatial variation, which is often related to environmental factors. Pollination type (insect-, wind- or self-pollination) is a critical trait for plant reproduction and provision of ecosystem services.

- Here, we mapped the distribution of pollination types across Germany by combining databases on plant distribution and plant pollination types. Applying a new method, we modelled the composition of pollination types using a set of 12 environmental variables as predictors within a Bayesian framework which allows for the analysis of compositional data in the presence of spatial autocorrelation.

- A clear biogeographical pattern in the distribution of pollination types was revealed which was adequately captured by our model. The most striking relationship was a relative increase in insect-pollination and a corresponding decrease of selfing with increasing altitude. Further important factors were wind speed, geology and land use.

- We present a powerful tool to analyse the distribution patterns of plant functional types such as pollination types and their relationship with environmental parameters in a spatially explicit framework.

Key words: Bayesian methods, central Europe, compositional data, environmental correlates, plant functional types, pollination types, spatial autocorrelation.

New Phytologist (2006) **172**: 127–139

© The Authors (2006). Journal compilation © *New Phytologist* (2006)

doi: 10.1111/j.1469-8137.2006.01811.x

Introduction

Traits of plant species have been used in numerous studies to reveal general ecological patterns irrespective of species identity (McIntyre *et al.*, 1999; Rusch *et al.*, 2003). The relationship between plant functional traits and the environment at large geographic scales has recently received considerable attention (MacGillivray *et al.*, 1995; Skarpe, 1996; Díaz *et al.*, 1999; Wright *et al.*, 2004; Traiser *et al.*, 2005; Wright *et al.*, 2005; Mayfield *et al.*, 2006). Although pollination is a key process in ecosystem functioning, as it plays an important role in plant reproduction, it is often

neglected (e.g. Weiher *et al.*, 1999). There are three pollination types which account for the majority of plant species in temperate ecosystems: wind pollination (anemophily), insect pollination (entomophily) and self pollination (autogamy).

Many studies focused on the relationships among different pollination types, mainly from an evolutionary viewpoint, albeit often in an ecological context (Midgley & Bond, 1991a, 1991b; Kevan *et al.*, 1993). Often, such studies focused on either few species within a genus or family (Berry & Calvo, 1989; Tamura & Kudo, 2000; Dupont, 2002; Kalisz & Vogler, 2003) or specific groups of plants such as trees and shrubs (Regal, 1982). Some studies investigated

biogeographic trends or relationships between specific pollination types and environmental parameters. In moist temperate forests wind pollination has been found to increase with latitude and altitude and to decrease with plant species richness (Regal, 1982; Whitehead, 1983). Wind pollination also depends on factors other than just wind, such as humidity, rainfall and temperature (Culley *et al.*, 2002). Rainfall has a negative effect as it washes the pollen away (Regal, 1982). The optimum conditions for wind pollination are low to moderate wind speed, low humidity and infrequent precipitation (Whitehead, 1983; Culley *et al.*, 2002). Too high wind speeds may hinder stigmatic pollen capture (Niklas, 1985). Insect pollination is typically associated with zero to low wind speed, medium to high humidity and infrequent to common precipitation (Regal, 1982). Obviously, insect pollination is restricted to regions where insects could thrive. However, it is difficult to get data on insect abundance on biogeographic scales. Furthermore, entomophily increases with plant species richness (Whitehead, 1968; Regal, 1982).

Selfing is regarded as a method of reproductive assurance (Baker, 1955; Schoen *et al.*, 1996; Kalisz *et al.*, 2004). Selfing should be especially favoured under variable pollination environments (Kalisz *et al.*, 2004) or poor climatic conditions where pollinators or mates are absent (Baker, 1955).

There are other factors in addition to climatic variables or the direct physical environment that may have an influence on the composition of pollination types. Wind pollination is favoured by open vegetation while insect pollination occurs in open to closed vegetation (Culley *et al.*, 2002). Conversely, many tree species of temperate forests are wind pollinated (Regal, 1982).

The relationships and biogeographic trends reported in these studies suggest that pollination types are differentially selected for by different ecological and environmental conditions. If this is a general rule, one would expect to find different pollination types displaying different spatial patterns, based on the geographical variation of the underlying environmental factors. This in turn should allow the variation in pollination types not only to be mapped and their spatial structure to be analysed, but also to develop proper statistical models that could explain and predict such patterns. Such an approach, which aims at identifying the underlying environmental drivers of the geographic distribution of traits, is different from the approach generally applied in studies of functional traits (i.e. focusing on dynamic processes such as responses of trait composition to management). Hitherto, actual geographical distribution of selected traits was taken rarely into account (e.g. Traiser *et al.*, 2005) and we are not aware of any such analysis on the distribution of pollination types.

Here, we used a biogeographic approach and analysed the distribution patterns of a trait (pollination) that we derived from a comprehensive mapping scheme and which can be related to environmental data at the same spatial scale. The scale of our study is different from previous analyses of

pollination types (Regal, 1982; Whitehead, 1983) because our analysis is based on a greater number of species (several thousand species) and their occurrence patterns over a large geographical extent and at a coarse spatial resolution. We analysed the distribution patterns of the relative frequencies of different pollination types in Germany at a 10' longitude \times 6' latitude resolution (*c.* 130 km²) in a spatially explicit statistical framework.

Analysing spatial maps of relative frequencies (compositions) of species or traits poses certain critical statistical challenges. First, the proportions of traits or species in different groups add up to 100% (unit sum constraint of compositions). Therefore, an increase in relative frequency of one group results in the decrease in relative frequency of one or more other groups (Aitchison, 1982, 1986; Billheimer & Guttorp, 1995; Billheimer *et al.*, 2001). A second statistical challenge is to deal with the potential spatial autocorrelation structure in the data or model residuals. The presence of spatial autocorrelation in a data set may lead to several problems. In the presence of positive spatial autocorrelation, errors are not independently distributed, which violates the basic assumption of usual linear modelling techniques (Haining, 2003). This will lead to an overestimation of degrees of freedom and Type I errors may strongly be inflated (Legendre, 1993). Furthermore, the effects of the explanatory variables may be estimated incorrectly (Cressie, 1993; Anselin & Bera, 1998). In this paper we present a novel statistical approach that enables us to relate maps of trait compositions with maps of the environment, based upon the breaking of the unit sum constraint using log-ratios of proportions. Equally importantly, the approach applied here enables us also to account for the potential biases in the pollination traits models stemming from spatial autocorrelation by the spatial smoothing of model residuals using a conditional autoregressive model. While a similar methodology was introduced by Billheimer *et al.* (2001) for the modelling of species composition data gathered at a number of sampling stations, this is to our knowledge the first application of these techniques in the analysis of geographical maps of species' trait compositions. Such maps will become more commonly available as species atlases are coupled with databases of species traits.

The aims of this study can be summarized as (1) exploring whether pollination types will yield spatially structured distribution patterns; and (2) testing whether mapped environmental variables can account for those patterns by modelling the distribution of distribution of pollination type composition using a Bayesian framework in the presence of spatial autocorrelation.

As suggested by other studies (Whitehead, 1968, 1983; Regal, 1982; Niklas, 1985; Culley *et al.*, 2002; see earlier) we expect that altitude, temperature, precipitation, wind speed and specific geological substrates associated with species richness (Kühn *et al.*, 2003) may influence the composition of pollination types. Specifically, we hypothesize that: (1) the

proportion of insect-pollination increases with increasing temperature and area of lime and loess subsoil, and decreases with increasing precipitation and wind speed; and (2) the proportion of wind pollination increases with altitude, open vegetation (e.g. grasslands and arable fields) and moderately with wind speed, and decreases with increasing precipitation. We expect that (3) the spatial variation in the proportion of self-pollination will be least well explained, because selfing results from a lack of other opportunities of pollination or as a reaction to unpredictable or highly varying environmental conditions, (i.e. it should increase with altitude and in disturbed regions).

Materials and Methods

Data sources

Data on plant species distribution was taken from the database FLORKART provided by the German Federal Agency for Nature Conservation. The data in this database was collated with the help of thousands of volunteers. The spatial resolution is defined by grid cells with a size of $10'$ longitude \times $6'$ latitude (c. 130 km^2), totalling 2995 cells for Germany. Mapping was organized at a regional level, which resulted in areas that were very well mapped and those that were insufficiently mapped. To correct for heterogeneity in mapping intensity throughout Germany, we designated 50 'control species'. These are the 45 most frequent species according to Krause (1998) plus five species that are either inconspicuous or taxonomically difficult. We could reasonably assume that the control species should occur in every grid cell in Germany. Having > 15 yr experience with volunteer botanists (I.K. and S.K.), we were aware of the fact that some species that are considered to be 'difficult' were often overlooked or mapped at a higher aggregate level by less experienced botanists or in less frequently mapped cells. This knowledge is used here to identify such less well mapped regions. Therefore, only grid cells having at least 45 of the 50 control species were analysed. Borderline cells (with parts in other countries or the sea) had > 50% of their area inside Germany ($n = 176$ out of 394 borderline cells). This resulted in 2733 grid cells for analysis.

Pollination types were taken from the 'floral biology' chapter (Durka, 2002) of the BiolFlor database on biological and ecological traits of the German flora (Klotz *et al.*, 2002; Kühn *et al.*, 2004; <http://www.ufz.de/biolflor>). We restricted the analyses to the major pollination types 'wind-pollination', 'insect-pollination' and 'self-pollination' which account for c. 95–98% of each grid cell's species with known pollination types. Pollination modes were considered if reported to be 'always', 'the rule' or 'often' but not if reported as 'rare' or 'possible' (see Durka, 2002).

Pollination types were available for 3503 species (95.7% of all German species). In total, we had information on the

required pollination type and on distribution for 2678 species (73.2%) in 2733 grid cells totalling to two million records for our analyses. Of these species, 1571 species (58.7%) were insect-pollinated, 541 spp. (20.2%) were wind-pollinated, 397 spp. (14.8%) were self-pollinated, 133 spp. (5%) were insect- or self-pollinated, 33 spp. (1.2%) were wind- or self-pollinated, 2 spp. (0.001%) were insect- or wind-pollinated and one species was recorded as being pollinated by all three vectors. The proportion of each of the pollination types of the total flora with information on pollination available was calculated per grid cell. Here, we were primarily interested in trait compositions, and therefore computed the frequencies of the occurrences of the three main pollination types (wind, selfing and insect-pollination) in each grid cell. For example, the total number of the pollination state 'insect-pollination' in a grid cell was computed as the sum of the number of insect pollinated and insect-or self-pollinated plant species. We therefore have, per grid cell, a composition consisting of the relative proportions of the three pollination types out of the total number of pollination types recorded in that grid cell. Relative proportions, as used in this analysis, however, do not provide any information on absolute frequencies. An increase (e.g. in the proportion of traits in one group) does not tell us whether the absolute number of plants in this group tended to increase.

For each cell, we compiled data on altitude, geology, land use, and climate. Information on relief was derived from the ARCDDeutschland500 dataset from ESRI (ESRI Geoinformatik GmbH, Kranzberg, Germany) and we calculated the average altitude per grid cell. Geological data was aggregated from the Geological Survey Map of Germany (Bundesanstalt für Geowissenschaften & Rohstoffe, 1993). We used the area of geological substrate classes lime, sand and loess per grid cell. We provided the following land use data as explanatory variables: area of agricultural fields, agricultural grasslands (semi)natural grasslands and deciduous forests, as provided by the Corine Land Cover data sets (Statistisches Bundesamt, 1997). For climate data, we used the mean annual temperature (1960–90), mean annual precipitation (1950–80) and mean wind speed (10 m above ground, 1960–90) provided by the German Meteorological Service (Deutscher Wetterdienst, Department Klima und Umwelt, Offenbach, Germany). We intentionally did not use latitude as an independent predictor, because it is known that there is a south–north gradient (in the northern hemisphere) in species richness (Francis & Currie, 2003; Hawkins *et al.*, 2003; Kier *et al.*, 2005). However, this is caused by a number of underlying physical and climatic parameters that covary with latitude which are included in our data set.

Modeling spatially referenced compositional data

A general introduction to the statistical framework A generally suitable statistical method for handling compositional data is to 'break' the unit sum constraint by replacing the

observed proportions by the logarithms of ratios of proportions (log-ratios hereafter) (Aitchison, 1986). Log-ratios have some desirable properties, most importantly the invariance property, which ensures that the choice of numerator and denominator in forming the ratios is unimportant. Moreover, a version of the central limit theorem exists which provides an explanation of why random variation in log-ratios can often be assumed to be normally distributed (Aitchison, 1982). Aitchison (1982, 1986) models the log-ratios using the multivariate normal distribution:

$$\phi(\vec{u}) \sim MVN(\vec{\mu}, \Sigma)$$

The vector $\vec{u} = (u_1, u_2, \dots, u_k)$ consists of the proportions in the k different groups (for example $k = 3$, in a composition consisting of proportions of wind, insect and self-pollination trait states). The function $\phi(\cdot)$ is the log-ratio transformation, so that

$$\phi(\vec{u}) = \left(\log\left(\frac{u_1}{u_k}\right), \log\left(\frac{u_2}{u_k}\right), \dots, \log\left(\frac{u_{k-1}}{u_k}\right) \right)$$

are the $k-1$ log-ratios. The vector $\vec{\mu} = (\mu_1, \mu_2, \dots, \mu_{k-1})$ is the multivariate mean on the log-ratio scale, and Σ is a $(k-1) \times (k-1)$ variance-covariance matrix. The key benefits of assuming multivariate normality are that inference tools developed for the multivariate normal can be applied, and that its covariance structure can be used to model the dependencies between the components of the composition (Billheimer *et al.*, 2001). The estimated parameters $\vec{\mu}$ can be back-transformed to proportions using the inverse of the additive log-ratio transform:

$$\vec{\mu} = \phi^{-1} \left(\frac{e^{\mu_1}}{1 + e^{\sum_{j=1}^{k-1} \mu_j}}, \frac{e^{\mu_2}}{1 + e^{\sum_{j=1}^{k-1} \mu_j}}, \dots, \frac{e^{\mu_{k-1}}}{1 + e^{\sum_{j=1}^{k-1} \mu_j}} \right)$$

It is possible to interpret $\phi^{-1}(\vec{\mu})$ as a point estimate of the multivariate median for the composition (Billheimer & Guttorp, 1995; Billheimer *et al.*, 2001). Within a linear (regression) framework, $\vec{\mu}$ may be related to covariates (i.e. explanatory variables) using a vector of $k-1$ slopes (\vec{a}) and intercepts (\vec{a}_0):

$$(\vec{\mu}_0) = \vec{a}_0 + \vec{a}_1(x_i - \bar{x}_i)$$

(x_i is a covariate, indexed by $i = 1, 2, \dots, n$ observations; \bar{x}_i is the mean of the observed covariate values).

Spatial dependence of the compositions can be incorporated into the modelling structure using a conditional autoregressive model (CAR) (Besag, 1974; Billheimer *et al.*, 1997). Here, we use an intrinsic version of the CAR, as proposed by Besag *et al.* (1991). Assume a set of multivariate spatially correlated Gaussian random effects $\vec{S}_i = (S_i^1, S_i^2, \dots, S_i^{k-1})$, where k is the number of components in the composition, and $i = 1, 2, \dots, N$, and N is the number of grid cells. We model the distribution of these random effects as multivariate normal,

with conditional location vector given by the vector of average values of random effects in the neighbourhood (denoted by δ_i) of grid cell i , \vec{M}_{δ_i} , and conditional covariance matrix inversely proportional to the number of grid cells contained in this neighbourhood, n_i : $\vec{S}_i^1 | S_{-i} \sim MNV(\vec{M}_{\delta_i}, \Omega/n_i)$, where

$$\vec{M}_{\delta_i} = \left(\sum_{j \in \delta_i} S_j^1/n_i, \sum_{j \in \delta_i} S_j^2/n_i, \dots, \sum_{j \in \delta_i} S_j^{k-1}/n_i \right)$$

and Ω is a $(k-1) \times (k-1)$ variance-covariance matrix. The neighbourhood δ_i is defined as the grid cells directly adjacent and diagonal to grid cell i . The random effects are restricted to have zero mean, such that $\sum_{i=1}^N S_i^l = 0$ for all components $l = (1, 2, \dots, k-1)$ (Besag & Kooperberg, 1995).

Modelling pollination types

Let $\vec{u}_i = (u_i^1, u_i^2, u_i^3)$ denote the composition of pollination types in grid cell i , with u_i^1, u_i^2 , and u_i^3 , the relative proportions of the occurrences of the trait-states insect-, wind- and self-pollination, respectively. Here, we have chosen to use the proportion of trait-state wind-pollination as the denominator in the log-ratios (the choice of the denominator is irrelevant as described above). As described above, we model the log-ratios of pollination types in grid cell i ,

$$\phi(\vec{u}_i) = \left(\log\left(\frac{u_i^1}{u_i^2}\right), \log\left(\frac{u_i^3}{u_i^2}\right) \right)$$

using the bivariate normal distribution, with multivariate mean \vec{u}_i , and variance-covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_1 \\ \rho\sigma_1\sigma_1 & \sigma_2^2 \end{pmatrix}$$

where σ_1 and σ_2 are the variances of the log-ratios, and ρ is the correlation between the log-ratios. We model $\vec{\mu}_i$ as a linear function of covariates and account for spatial autocorrelation using the CAR model as described above:

$$\vec{\mu}_i = \vec{a}_0 + \vec{a}_1 x_{1j} + \vec{a}_2 x_{2j} + \dots + \vec{a}_n x_{nj} + \vec{S}_i \tag{Eqn 1}$$

where \vec{a}_0 are the intercepts, $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ vectors for the slopes for the effect of covariates, and \vec{S}_i spatial random effects drawn from a bivariate normal distribution with location vector $(0, 0)$ and variance-covariance matrix

$$\Omega = \begin{pmatrix} v_1^2 & \kappa v_1 v_1 \\ \kappa v_1 v_1 & v_2^2 \end{pmatrix}$$

(with v_1 and v_2 the variances of the spatial random effects and κ the correlation between the spatial random effects of the different log-ratio components).

Estimation of the model parameters using WinBugs

Because the model is high-dimensional, i.e. has many parameters (including the elements of two 2×2 variance-covariance matrices), we chose to use Markov chain Monte Carlo (MCMC) methods for model fitting (Gilks *et al.*, 1996; Brooks & Roberts, 1998). We used the GeoBugs module (Thomas *et al.*, 2004) of the WINBUGS package to implement the model. The multivariate conditional autoregressive model, as described above, is available in GeoBugs, which greatly facilitates the implementation of the model.

The model (Eqn 1) must be completed by specifying prior distributions for the parameters. For all parameters, we chose uninformative (vague) priors, indicating that we assumed no knowledge on the parameters prior to fitting the model to the data. We used location invariant uniform distributions $U(-\infty, +\infty)$ for both components of the location parameter vector \vec{a}_0 , as required in the implementation of the multivariate CAR in GeoBugs (Thomas *et al.*, 2004). As prior distributions for the components of all parameter vectors of slopes $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ we used normal distributions centred on zero and with a variance of 100. As priors for the inverses of the variance-covariance matrices Σ and Ω (WinBugs typically parameterizes normal distributions in terms of precisions, not variances) we used Wishart distributions $\Sigma^{-1} \sim \text{Wishart}(A, h)$, and $\Omega^{-1} \sim \text{Wishart}(B, h)$, with degrees of freedom $h = 2$, and $A = B = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix}$ with $a = 0.01$. The specification of the previous distributions for the precision of the spatial random effect may be influential. We checked this by using different sets of prior values of a set to 0.1 and 0.005. The estimated means and standard deviations of the variances and covariances of the spatial random effects and all other model parameters were similar for the different prior specifications. Therefore, we report results here only from the fitting of the model with a prior set to $a = 0.01$.

When using MCMC to estimate model parameters, one has to determine if the Markov chains have settled to stationary behaviour (convergence), and how many iterations the MCMC algorithm have to be run for, before it is safe to assume that the set of samples can be used to represent the posterior distribution of model parameters. Many diagnostic and graphical techniques are available to do this, but experience and a good understanding of MCMC techniques remains crucial (Gilks *et al.*, 1996; Brooks & Roberts, 1998). We initialized several independent Markov chains with widely dispersed starting values for the parameters $\vec{a}_0, \vec{a}_1, \vec{a}_2, \dots, \vec{a}_n, \Omega$ and Σ (initial values for the spatial random effects were always initialised at $\vec{S}_i = 0$). Convergence of these Markov chains was assumed when the obtained medians and the 2.5% and 97.5% quantiles of the posterior distributions of the individual chains were similar. Furthermore, we checked convergence of the Markov chains using visual inspection of time-series of Markov chains, autocorrelations of these

time-series, and the Brooks–Gelman–Rubin statistic (Brooks & Gelman, 1998). We ran our MCMC chains for 200,000 iterations, discarded the first 20,000 as burn-in, and saved one in every 10 simulated values of the remaining chains to reduce computational overheads. We based our inferences on the remaining 18,000 draws of the posterior distribution.

Assessing the significance of covariates

Initial graphical investigation indicated that altitude and temperature were strongly correlated with proportions of pollination types. Furthermore, these covariates were strongly correlated themselves, which was to be expected given the physical relationship between them. In addition, the relationship between altitude and temperature and the proportions was nonlinear. In comparison, proportions appeared to be weakly and linearly related to the other variables. Variable selection is impractical using the Bayesian CAR model because of the long computing times involved in estimating the parameters of these models. We therefore build up a picture of the importance of the covariates in two steps. First, we modelled each of the log-ratios of pollination types independently within a linear regression framework. Within this framework, models can be fitted quickly at the expense of ignoring dependencies between the log-ratios and potential spatial autocorrelation. For each of the log-ratios, a subset of the variables from a full model including all variables as regressors (and second- and third-order polynomials for altitude and temperature to allow for nonlinearities) was chosen using backwards stepwise selection based on AIC values, as implemented by the stepAIC function in R (Venables & Ripley, 2002; R Development Core Team, 2004). Second, all covariates that were selected in the final models of at least one of the log-ratios were chosen as covariates in the Bayesian CAR model, where their significance was further assessed. The expected effect of ignoring spatial autocorrelation and dependencies between log-ratios is that we are likely to pick up on more variables that are deemed to be important in the Bayesian CAR model, for the following reasons: (1) the effect of covariates tends to be overestimated when spatial autocorrelation is ignored (Anselin & Bera, 1998), and (2) the effect of ignoring correlations between log-ratios will be that explanatory variables that are significantly related to one of the log-ratios, but not the other, may be found not to be important in explaining the composition as a whole.

Results

Overview on trait composition and spatial distribution

The proportions of each of the pollination types are depicted in a ternary plot (Fig. 1). On average (± 1 SD),

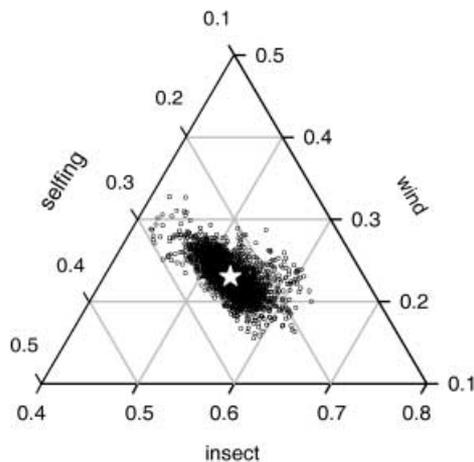


Fig. 1 Ternary plot of proportions of pollination types in Germany. The asterisk marks the average (gravity centre) of the pollination type proportions.

insect-pollination made up 53.7% ($\pm 3.1\%$) of all pollination types present in a grid cell, wind-pollination 22.7% ($\pm 2.3\%$) and selfing 23.6% ($\pm 1.6\%$). This resulted from an average species number per grid cell of 395.1 (± 91.1) species that can be insect pollinated, 164.9 (± 31.9) species that can be wind pollinated and 172 (± 37.8) species that can be self-pollinated. The geographical distribution of the different pollination types is given in Fig. 2(a–c). Insect pollination is less frequent in northern Germany, especially along the coasts and the river valleys of Elbe and Ems, the Upper Rhine valley and the Danube valley. Insect pollination is most frequent in the mid to high altitude mountain ranges of central and southern Germany. Wind pollination is most frequent in the northern German lowlands (where the highest wind speeds occur) and the Alpine upland in southern Germany; it is least frequent along the mid-altitude mountain ranges. Selfing is most frequent along the North Sea coast, the Elbe and North-western Rhine valley and least frequent in southern Germany. However, the general pattern of proportions of selfing is much more scattered than that of insect- or wind-pollination.

Modelling trait composition and environmental predictors

Visual comparison of the distribution of pollination type compositions with physical properties of Germany, suggests a relationship with altitude (Fig. 2d) and wind speed (Fig. 2e), but this requires a formal statistical test. The following set of variables was retained as predictors for the log-ratios using the stepwise backwards regression: altitude, altitude², area of lime subsoil, area of sand subsoil, area of loess subsoil, area of deciduous forest, wind speed, area of agricultural fields and area of agricultural grassland. The following variables were excluded: altitude³, precipitation, precipitation², temperature, temperature² and area of natural grassland.

The 95% credible intervals (the Bayesian equivalent of confidence intervals) of the estimated regression coefficients of the log ratio analyses (Table 1) revealed that the estimated posterior distributions of the intercepts, and parameters for the slopes of altitude², area of agricultural grasslands, area of agricultural land (i.e. arable fields), area of lime, area of sand and wind speed did not include zero (i.e. they could be regarded as significant).

The deviations of the proportions of pollination types from the mean composition along the observed gradients of the environmental predictors in Germany, as predicted by the model, are visualized in Fig. 3. The largest gradient in the composition of pollination types was predicted by the covariate altitude (Fig. 3a). The second largest gradient was predicted by the estimated slope for the covariate wind speed (Fig. 3b). In comparison, the predicted gradients using the estimated slopes for the covariates lime, sand, grassland and agricultural land were relatively small (Fig. 3c–f).

The predicted shifts in the compositions of pollination types along the observed gradients in the environmental predictors in Germany were different for each pollination type (Fig. 3). The relative proportions of insect pollination were nonlinearly related to altitude, increasing sharply from low to mid-altitudes, and remained relatively constant from mid to high altitudes. Furthermore the proportions of insect pollination increased with increasing area of lime and arable fields and decreased with increasing wind speed and area of sand and grasslands. Almost the opposite pattern could be observed for wind pollination, with the smallest proportions at mid-altitudes, and higher proportions in the lowlands and at higher altitudes. The relative proportions of wind pollination increased with increasing wind speed and (to a lesser degree) areas of sand and grassland. The proportions of wind-pollination decreased with increasing area of lime and arable fields. The proportions of self-pollination increased with increasing altitude and to a lesser degree with increasing wind speed.

We compared the observed log-ratios with the log-ratios as predicted using the estimated means of the model parameters $\bar{a}_0, \bar{a}_1, \dots, \bar{a}_n$ (predictions without the spatial random effect) to compute an estimate of the amount of variation explained. This can be compared with the conventional R^2 statistic in classical statistical methods, and may be used as an indication of the amount of variability explained by the environmental predictors. With an R^2 of 34.5% for the complete model (46.7% for log(entomophily/anemophily), and 22% for log(autogamy/anemophily)), the spatial variation in the environmental predictors explained a considerable amount of the spatial variation in the composition of pollination types.

Nevertheless, the fit of the model without a conditional autoregressive term for the spatial smoothing of residuals was inadequate because of the large spatial autocorrelation in residuals (Fig. 4a,b). The predicted 95% credible intervals of the predicted log-ratios made without the spatial random effects were too narrow, and only contained 50% of the data points (Fig. 4a,b).

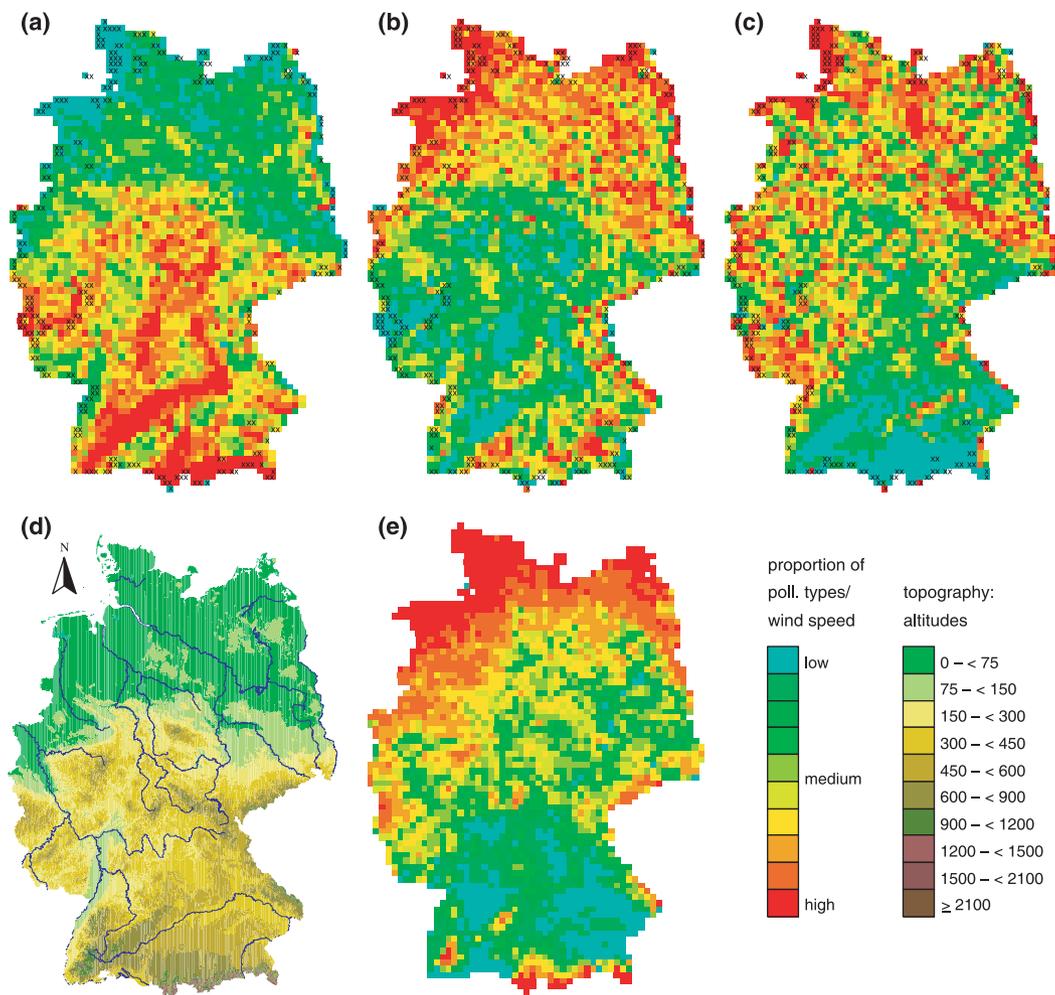


Fig. 2 Distribution of proportions of pollination types throughout Germany. (a) Insect pollination, range 41.9–63.1%; (b) wind pollination, range 15.5–32.7%; (c) self pollination, range 16.1–29.9%; (d) an overview over the topography of Germany; (e) wind speed, range 1.85–8.89 m s⁻¹. For proportions of pollination types see also Fig. 1. To better compare the maps, classes for proportions of pollination types and wind speed are decantiles (i.e. 10 classes of equal size with different class limits for each map). Cells not used for modelling (borderline cells and cells regarded as insufficiently mapped) are marked by an 'x'. Legends for pollination distribution and wind speed are in the Supplementary material Fig. S1.

Table 1 Estimates of the model parameters (the mean and 95% credible interval (c.i.) of the posterior distributions of the parameters)

	Log(entomophily/ anemophily)			Log(autogamy/ anemophily)		
	Mean	95% c.i.		Mean	95% c.i.	
Intercept	0.85	0.84	0.85	0.05	0.04	0.05
Altitude (×10 ⁻⁵)	1.88	1.16	2.64	-2.25	-9.00	4.10
Altitude ² (×10 ⁻⁷)	-1.62	-2.78	-4.91	-1.85	-2.82	-8.31
Forest	0.07	-0.005	0.14	-0.07	-0.14	-0.01
Grassland	-0.12	-0.17	-0.07	-0.15	-0.20	-0.11
Loess	0.07	-0.03	0.12	0.10	0.05	0.14
Lime	0.17	0.12	0.21	0.09	0.03	0.13
Sand	-0.07	-0.09	-0.05	-0.05	-0.06	-0.03
Wind	-0.05	-0.07	-0.04	-0.02	-0.03	-0.01
Agriculture	0.21	0.13	0.28	0.10	0.03	0.17

Parameters that are significantly different from zero are given in bold type.

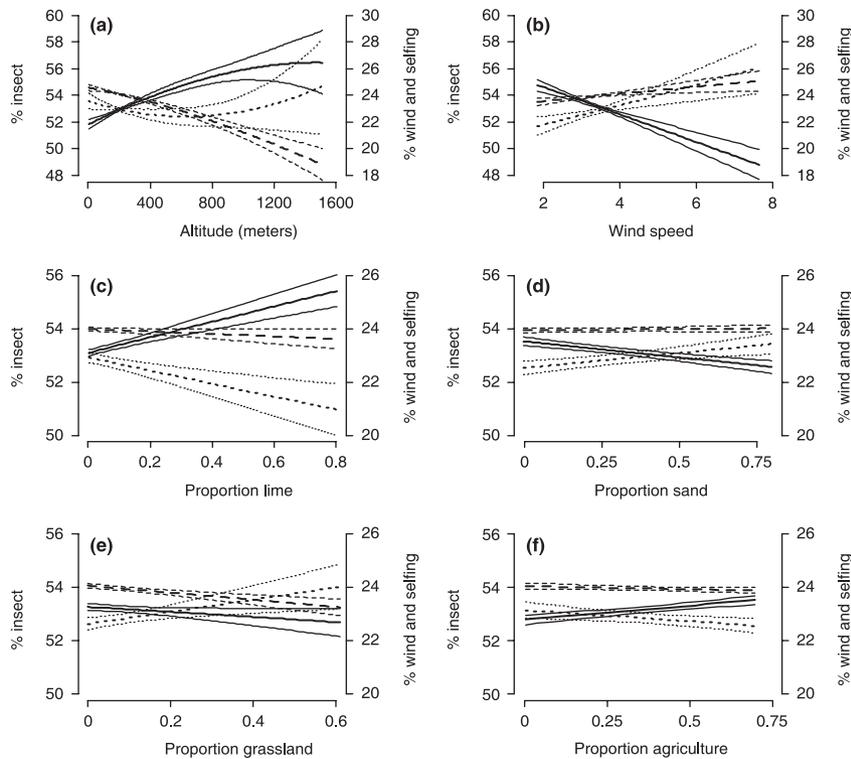


Fig. 3 The estimated deviations in the composition of pollination types from the mean composition, along the gradients in environmental predictors as observed in Germany. Gradients in the percentages of insect pollination (solid lines; left axis), self pollination (dashed lines; right axis) and wind-pollination (dotted lines; right axis) are shown for the six covariates regarded as significant: (a) altitude, (b) wind speed, (c) proportion of lime subsoil, (d) proportion of sand subsoil, (e) proportion of grassland and (f) proportion of agricultural land. The thick lines are the medians and the thin lines the 2.5% and 97.5% quantiles, of the predictions made using 5000 draws of the posterior distributions of the parameters. The predicted gradients of proportions of pollination types were predicted separately for each covariate, with the value of the other covariates set at their mean values.

The presence of strong spatial autocorrelation in the model residuals was reflected by the small estimated values (mean \pm SD) of the variances of the spatial random effects, with $v_1^2 = 0.027 (\pm 0.0014)$ and $v_2^2 = 0.018 (\pm 0.0013)$, indicating that residuals were restricted to vary smoothly over the grid. The model residuals of the log-ratios in the same grid cells were estimated to be strongly correlated ($\kappa = 0.72 (\pm 0.02)$). Ignoring spatial autocorrelation resulted in unrealistically small confidence bounds on the estimates of the model parameters. Furthermore, predictions of the model without spatial smoothing were biased with small and large values of log-ratios over- and under-estimated, respectively. The use of a conditional autoregressive term resulted in an adequate fit of the model to the data, with wider predicted credible intervals that reflected the uncertainty in model predictions more realistically (Fig. 4c,d). The complete model with autoregressive term had an R^2 of 95% (97.3% for $\log(\text{entomophily}/\text{anemophily})$, and 92.7% for $\log(\text{autogamy}/\text{anemophily})$). The proportion of data points contained by the 95% credible intervals of log-ratios as fitted by the model with spatial smoothing was 98% (Fig. 4c,d). Predictions of this model appeared only slightly biased.

The large improvement in fit by including a spatial term, \bar{S}_p in the model indicates that there are spatial patterns that

are not captured by the covariates. This is reflected by the patterns in the maps of the spatial random effects (Supplementary material Fig. S2) which indicate the existence of gradients in the composition of pollination types that are not captured by the environmental predictors in our model. This can be seen by a pattern of positive values of \bar{S}_i along the Swabian-Frankonian Alb (in the south of Germany), the mid-German hills and the very eastern Valley of the Oder River and negative values of \bar{S}_i in the Alpine uplands (Fig. S2). These regions are characterized by specific geological or climatic conditions that were not explicitly incorporated into the model. The maps of the spatial random effects can thus be used to highlight such unexplained spatial patterns, and suggest potential environmental predictors that may be included into the model.

Discussion

The modelling approach

There are several methods to analytically combine data on species distributions, traits and the environment. The combinations of these three elements are the classical domain

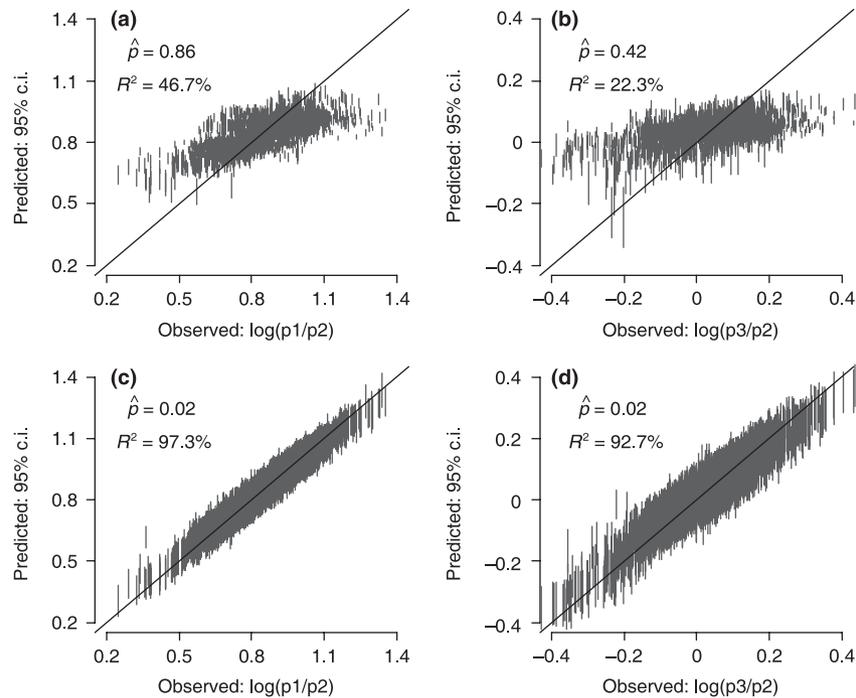


Fig. 4 Observed values of log-ratios (x-axes) against predicted 95% credible intervals of log-ratios (y-axes). Predictions were made using the covariates only (leaving the spatial random effect out) (a,b) and with the full model with both covariates and spatial random effect (c,d). (a,c) $\log(p1/p2)$ (i.e. $\log(\text{entomophily}/\text{anemophily})$); (b,c) $\log(p3/p2)$, (i.e. $\log(\text{selfing}/\text{anemophily})$). The log-ratios have been sorted from their smallest to largest observed values along the x-axes, and their corresponding 95% credible intervals are indicated by the vertical segments. The straight lines through the origin indicate a perfect fit where predictions equal observations. The proportion of data points that fall outside of their corresponding 95% predicted credible intervals, \hat{p} , are given for each log-ratio. In addition, as an indication of the amount of variation explained by the full model (with random effect: c,d) and the covariates only (a,b), R^2 values have been computed using the means of the predicted intervals.

of ‘4th corner problems’ (Legendre *et al.*, 1997; Legendre & Legendre, 1998), where an environment by trait matrix (the fourth corner) is estimated by three other matrices: (1) a species-by-locality matrix, (2) a species-by-traits matrix and (3) an environment-by-locality matrix. Legendre *et al.* (1997) solve the problem by combining these three matrices through multiplication. Other possibilities are, for example, the combination of the three matrices through multivariate approaches such as coinertia analysis (Dray *et al.*, 2003) or the use of linear models with traits as explanatory variables to explain species niche breadth (Thuiller *et al.*, 2004) or species responses (Lososova *et al.*, 2004) of prior multivariate analyses. However, we analysed the spatial distribution of the relative frequencies of traits, and thus have to: (1) treat the grid cells as the unit of observation, since the composition in each grid cell is one observation on the relative frequencies of the pollination types; (2) allow for the unit sum constraint of compositional data; (3) allow for spatial autocorrelation in model residuals, since the trait compositions vary smoothly over space. The statistical framework as proposed in this paper meets the above requirements.

Allowing for the spatial dependency of the compositions by means of a conditional autoregressive model proved necessary to obtain an adequate fit of the model to the data. Models that

did not allow for spatial autocorrelation in the compositions proved too optimistic in the uncertainty of the model predictions, as reflected by the narrow credible intervals that excluded most of the observed data points. This illustrates the fact that models that ignore spatial autocorrelation are susceptible to type I errors because they are overoptimistic about the precision of the estimated values of the parameters (Cressie, 1993; Legendre, 1993; Anselin & Bera, 1998). By contrast, the estimated uncertainty in the parameter estimates and predictions from the model with spatial smoothing were realistic, making any inferences on the effect of covariates on the composition of plant communities more credible. The composition of plant assemblages, as derived from species atlases with a coarse spatial resolution, will often vary smoothly over large geographical areas. However, since the spatial distribution of most covariates and the response variable will not overlap entirely, allowing for spatial autocorrelation in the model residuals is crucial when modelling plant community compositions. Few examples exist in the literature on applications of statistical methods that are suitable for spatially correlated compositional data, except the work by Billheimer *et al.* (1998) and Billheimer *et al.* (2001). In this paper we demonstrate the application of these methods in the analysis of community composition data as derived from

species atlases on a coarse spatial grid. In addition, our study demonstrates the usefulness of GeoBugs (Thomas *et al.*, 2004), which greatly facilitates the implementation of these relatively nonstandard geostatistical models.

Biogeographic patterns of pollination types on meso-scale

The shifts in compositions of pollination types revealed in our study were rather small across Germany (Fig. 1). In addition, the resolution of our analysis was rather coarse (*c.* 130 km²) and pollination effects are most likely to operate on small scales (Kunin, 1992). Nevertheless, we still were able to detect a clear distributional pattern of pollination types and a response to environmental factors. Though the pattern is striking, it is worthwhile to note that across Germany entomophily is still almost twice as frequent as anemophily or autogamy; there is only a relative shift of the pollination types in space.

At this spatial scale and resolution, many of the previously published results on relationships between the spatial distribution of the environment and pollination types cannot be confirmed. The estimated increase in insect pollination, and decrease in wind and self-pollination, from low to high altitudes is even the complete opposite of earlier hypotheses or observations (Regal, 1982; Whitehead, 1983; Richards, 1997).

A number of processes may be involved in determining the distribution of pollination types: evolution, weather and climate, physical properties of vegetation types, species interactions and history including human influence. We will not discuss the evolution of pollination modes since most of the German flora evolved outside of the area and colonized it after the last glaciation.

Insect pollination is typically associated with zero to low wind speed, medium to high humidity and infrequent to common precipitation (Regal, 1982). In our model, distribution of entomophily may best be explained by human land use as a consequence of topography, geology and the absence of strong winds. The most species-rich areas, with a high number of colourful flowering entomophilous species, are calcareous pastures and meadows which are extensively used in mid-altitudes and naturally open vegetation above the timberline in the Alps (Ellenberg, 1996). Thus they are situated in the mid- to high-altitude mountain areas (Ssymank *et al.*, 1998) where agriculture could not be intensified as much as in the lowlands or is characterized by extensively used seasonal mountain pastures. Unfortunately, only proxy variables are available, but the combination of altitude and area of limestone support the ideas mentioned above, as does the distribution of the respective habitats (see Ssymank *et al.*, 1998). Clearly, because high wind speeds impede insect flight, this covariate is negative correlated with insect pollination. It is not easy to explain the positive relationship between the area of arable fields (which are usually species poor) with proportion of insect pollination, as many weed species of

fields are self- or wind-pollinated (Baker, 1974). However, the predicted gradient is small with relatively large uncertainty.

One of the best predictors for the spatial distribution of proportions of wind pollination is wind speed, although it is known that the optimum for wind-pollination is at low to moderate wind speed (Whitehead, 1983). Terminal velocities of pollen for most anemophilous species range from 0.02 to 0.06 m s⁻¹ (Whitehead, 1968). Wind speed in the studied area ranges from 1.8 to 8.9 m s⁻¹ at 10 m above ground level. The average wind speed within plant communities is about the same as those in our data at 10 m above ground level (1–10 m s⁻¹, Whitehead, 1983 citing Tauber, 1965, Geiger, 1966). Wind speeds are therefore *c.* 100 times faster than needed for pollen dispersal and should not be a limiting factor. However, pollinators do have difficulties in flying from flower to flower and pollinating when wind speed is high, so that the observed pattern could result from a decrease in insect pollination. Furthermore, it seems that the distribution of anemophilous plant species may also be explained by factors (covarying) other than just wind. An altitudinal increase in anemophily seems to be a wide spread pattern across many regions of the world and across several taxa, as a result of low pollinator availability (Regal, 1982; Berry & Calvo, 1989; Anderson *et al.*, 2001). In Germany, however, there is also an increase of anemophily in the lowest areas which are the northernmost ones. Although this seems to fit in with the latitudinal trend of increase in anemophily (Regal, 1982), which is understood as a result of unfavourable climatic conditions, the climatic variables – when exchanging altitude with temperature in our model – do not support this idea.

Wind pollination is facilitated by open vegetation (Culley *et al.*, 2002). Thus, broader-scale vegetation patterns may also influence trait composition, especially if open, graminoid-dominated vegetation types exist. Most of Germany is part of the temperate forest biome, particularly beech forest. However, grasslands of various types exist mainly as secondary anthropogenic vegetation throughout the country. The most species-rich types of grassland on limestone occur mainly in the mid-altitudinal ranges of Central and Southern Germany. These grasslands are, however, especially rich in insect-pollinated species, which may partly account for the high level of insect pollination at mid- to high-altitude ranges found in our analyses (see above). Vegetation types which strongly differ in their abundance across Germany are bogs and fens that occur both in the climatically humid areas in the northern lowlands and in the South in the peri-alpine and alpine zone but are rare or absent in the central mid elevation parts (Ellenberg, 1996; Succow & Joosten, 2001). In bogs, a high species richness of mostly wind-pollinated Cyperaceae is found, at least part of which are biogeographically restricted within Germany to these two areas and thus may contribute to the relative minimum of wind-pollination in the central, mid-altitude parts of Germany. In addition, the distribution of these bogs across Germany is partly caused by and partly

covaries with the last Pleistocene glaciation (Liedtke & Marcinek, 2002). This historical factor can potentially contribute to the mid-altitude minimum of wind-pollination as both glaciated regions (the northern lowlands and southern alpine areas) still may have imprints on trait composition owing to glacial relict species.

Self pollination is the pollination mode that was most scattered across Germany, only being less frequent in the south-east. The reproductive assurance hypothesis (Baker, 1955) states that selfing is a selective advantage when pollinators are absent (e.g. owing to poor or unpredictable climatic conditions, frequent disturbance or during colonization; see also Schoen *et al.*, 1996; Kalisz & Vogler, 2003; Kalisz *et al.*, 2004). The areas of highest selfing frequencies are along the North Sea coast, the Pleistocene lake areas in the north-east and some parts of the Elbe and Rhine valley. All these are regions where natural disturbance is high (e.g. owing to flooding events and/or storms). The large river valleys are also known to be especially rich in alien species (Platy-Tabacchi *et al.*, 1996; Deuschewitz *et al.*, 2003) which are more frequently self-pollinated than native species (Klotz *et al.*, 2002).

The percentage of self-pollination was clearly found to decline with increasing altitude (Fig. 3a). This finding strongly contrasts with the classical expectation based on the reproductive assurance hypothesis that selfing should be selected for under unfavourable environmental conditions such as high altitudes or latitudes in which pollinator service may be uncertain (Bliss, 1962; Richards, 1997). This view, however, seems to have been based on the premature adoption for alpine floras of the suggested role of autogamy and apomixis in arctic floras (Packer, 1974). Empirical evidence for a declining role of insect pollination may be biased towards particularly sensitive and taxonomically narrow groups of insect pollinated species, like orchids pollinated mainly by Lepidoptera (Jacquemyn *et al.*, 2005). The decline of selfing species with increasing altitude may not be related to the breeding system itself, but may result from the predominant annual life cycle of selfing species. Thus, they depend on the successful completion of the life cycle within one season. This makes them more susceptible to unfavourable and variable climatic conditions than the preferentially outcrossing perennials which can accumulate scarce resources over time and endure unfavourable conditions.

Indeed, a number of studies point to the maintenance or predominance of outcrossing breeding systems at high altitudes of temperate regions (Gugerli, 1998; Körner, 2003). Unlike selfing, outcrossing breeding systems ensure the maintenance of high genetic variability at population level which is considered a prerequisite for long-term persistence in stochastic environments (Lande & Shannon, 1996). Indeed, many high-elevation plants combine outcrossing breeding systems that ensure the maintenance of genetic variability with clonal propagation, allowing persistence and reproduction of successful genotypes under harsh environmental conditions. Selfing was also negatively affected by percentage

grassland. Species-rich grasslands are characterized by strong competition among species resulting in low numbers of weak competitors like annuals which mostly are selfing species (Aarssen, 2000). Wind speed was weakly positively correlated with selfing. This, however, is probably a result of the strong decrease of insect pollination.

A number of factors and processes that may influence the distributional patterns of pollination modes cannot be tackled by our analysis. First, pollination is not static within a species but may vary both in space and time in adaptation to local conditions. Thus there is not only a shift between species and in species composition but also a shift within species and populations from outcrossing to selfing under adverse environmental conditions (Kalisz *et al.*, 2004); this is also known for a number of European species (Couderc, 1978). However, such processes could not be recognized at the scale of our analysis and the species data at hand. Second, pollination is not necessarily needed for reproduction if species can reproduce clonally. Thus, despite having a specific pollination mode and breeding system, the distribution of clonal species may be independent of functioning of their sexual system (Hollingsworth *et al.*, 1998). Third, and most important, the patterns of pollination type distribution may be biased by phylogenetic effects (i.e. closely related species of one grid cell may dominate a pattern against species from distant clades of a phylogeny). Unfortunately, we are not aware of any way of incorporating comparative methods (Harvey & Pagel, 1991) for compositional data in our context and that could be applied to > 2700 'samples'.

There was some caution raised by Quinn *et al.* (1994) when interpreting spatial patterns in the abundance of pollination types that are extrapolated to larger scales (e.g. 10×10 km resolution). Nevertheless, at a scale of $10'$ longitude \times $6'$ latitude, we were successfully able to show a distinct pattern in the spatial distribution of the composition of pollination types across Germany, which we were able to explain using a set of environmental variables by effectively employing a novel statistical method which is applied (to our knowledge) for the first time to species distribution atlases.

Acknowledgements

We thank Isabell Augenstein for transforming wind speed into the projection needed, and thousand of volunteers mapping the flora of Germany which is incorporated in FLORKART, provided by the Federal Agency for Nature Conservation, namely Rudolf May and Hans Fink, on behalf of the German Network for Phytodiversity (NetPhyD). Risto Heikkinen, Ralf Ohlemüller, Erik Welk, Glenn Marion and David Ackerly commented earlier versions and kindly improved the text. This work was funded by the European Commission within Framework Programme 6 Integrated Project 'ALARM: Assessing LARge scale environmental Risks with tested Methods' (GOCE-CT-2003-506675) (Settele *et al.*, 2005).

References

- Aarssen LW. 2000. Why are most selfers annuals? A new hypothesis for the fitness benefit of selfing. *Oikos* 89: 606–612.
- Aitchison J. 1982. The statistical analysis of compositional data. *Journal of the Royal Statistical Society Series B – Methodological* 44: 139–177.
- Aitchison J. 1986. *The statistical analysis of compositional data*. New York, NY, USA: Chapman & Hall.
- Anderson GJ, Bernardello G, Stuessy TF, Crawford DJ. 2001. Breeding system and pollination of selected plants endemic to Juan Fernandez Islands. *American Journal of Botany* 88: 220–233.
- Anselin L, Bera AK. 1998. Spatial dependence in linear regression models with an introduction to spatial econometrics. In: Ullah A, Giles DEA, eds. *Handbook of applied economic statistics*. New York, NY, USA: Marcel Dekker, 237–289.
- Baker HG. 1955. Self-compatibility and establishment after 'long-distance' dispersal. *Evolution* 9: 347–349.
- Baker HG. 1974. The evolution of weeds. *Annual Review of Ecology and Systematics* 5: 1–24.
- Berry PE, Calvo RN. 1989. Wind pollination, self-incompatibility, and altitudinal shifts in pollination systems in the high Andean genus *Espeletia* (Asteraceae). *American Journal of Botany* 76: 1602–1614.
- Besag J. 1974. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society Series B – Statistical Methodology* 36: 192–236.
- Besag J, Kooperberg CL. 1995. On conditional and intrinsic autoregressions. *Biometrika* 82: 733–746.
- Besag J, York J, Mollie A. 1991. Bayesian image restoration, with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics* 43: 1–59.
- Billheimer D, Guttorp P. 1995. *Spatial models for discrete compositional data. Technical report*. Seattle, WA, USA: University of Washington, Department of Statistics.
- Billheimer D, Cardoso T, Freeman E, Guttorp P, Ko H, Silkey M. 1997. Natural variability of benthic species in the Delaware Bay. *Journal of Environmental and Ecological Statistics* 4: 9–115.
- Billheimer D, Guttorp P, Fagan WF. 1998. *Statistical analysis and interpretation of discrete compositional data*. NRCSE Technical Report Series 011. Seattle, WA, USA: University of Washington.
- Billheimer D, Guttorp P, Fagan WF. 2001. Statistical interpretation of species composition. *Journal of the American Statistical Association* 96: 1205–1214.
- Bliss LC. 1962. Adaptations of arctic and alpine plants to environmental conditions. *Arctic* 15: 117–144.
- Brooks SP, Gelman A. 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7: 434–455.
- Brooks SP, Roberts GO. 1998. Assessing convergence of Markov chain Monte Carlo algorithms. *Statistics and Computing* 8: 319–335.
- Bundesanstalt für Geowissenschaften und Rohstoffe. 1993. *Geologische Karte der Bundesrepublik Deutschland 1 : 1 000 000*. Hanover, Germany: Bundesanstalt für Geowissenschaften und Rohstoffe.
- Couderc H. 1978. Adaptation de la fleur entomophile de l'*Anthyllis vulneraria* L. à l'autogamy. *Bulletin de la Société Botanique de France* 125: 369–378.
- Cressie NAC. 1993. *Statistics for spatial data*. Cambridge, MA, USA: Wiley.
- Culley TM, Weller SG, Sakai AK. 2002. The evolution of wind pollination in angiosperms. *Trends in Ecology and Evolution* 17: 361–369.
- Deuschewitz K, Lausch A, Kühn I, Klotz S. 2003. Native and alien plant species richness in relation to spatial heterogeneity on a regional scale in Germany. *Global Ecology and Biogeography* 12: 299–311.
- Díaz S, Cabido M, Zak M, Carretero EM, Aranibar J. 1999. Plant functional traits, ecosystem structure and land-use history along a climatic gradient in central-western Argentina. *Journal of Vegetation Science* 10: 651–660.
- Dray S, Chessel D, Thioulouse J. 2003. Co-inertia analysis and the linking of ecological data tables. *Ecology* 84: 3078–3089.
- Dupont YL. 2002. Evolution of apomixis as a strategy of colonization in the dioecious species *Lindera glauca* (Lauraceae). *Population Ecology* 44: 293–297.
- Durka W. 2002. Blüten- und Reproduktionsbiologie. In: Klotz S, Kühn I, Durka W, eds. *BIOLFLOR – eine Datenbank zu biologisch-ökologischen Merkmalen der Gefäßpflanzen in Deutschland. Schriftenreihe für Vegetationskunde*, 38. Bonn, Germany: Bundesamt für Naturschutz, 133–175.
- Ellenberg H. 1996. *Vegetation Mitteleuropas mit den Alpen*, 5th edn. Stuttgart, Germany: Ulmer.
- Francis AP, Currie DJ. 2003. A globally consistent richness-climate relationship for angiosperms. *American Naturalist* 161: 523–536.
- Geiger R. 1966. *The climate near the ground*. Cambridge, MA, USA: Harvard University Press.
- Gilks WR, Richardson S, Spiegelhalter DJ. 1996. *Markov chain Monte Carlo in practice*. London, UK: Chapman & Hall.
- Gugerli F. 1998. Effect of elevation on sexual reproduction in alpine populations of *Saxifraga oppositifolia* (Saxifragaceae). *Oecologia (Berlin)* 114: 60–66.
- Haining RP. 2003. *Spatial data analysis: theory and practice*. Cambridge, UK: Cambridge University Press.
- Harvey PH, Pagel MD. 1991. *The comparative method in evolutionary biology*. Oxford, UK: Oxford University Press.
- Hawkins BA, Field R, Cornell HV, Currie DJ, Guegan JF, Kaufman DM, Kerr JT, Mittelbach GG, Oberdorff T, O'Brien EM, Porter EE, Turner JRG. 2003. Energy, water, and broad-scale geographic patterns of species richness. *Ecology* 84: 3105–3117.
- Hollingsworth ML, Hollingsworth PM, Jenkins GI, Bailey JP, Ferris C. 1998. The use of molecular markers to study patterns of genotypic diversity in some invasive alien *Fallopia* spp. (Polygonaceae). *Molecular Ecology* 7: 1681–1691.
- Jacquemyn H, Micheneau C, Roberts DL, Paillet T. 2005. Elevational gradients of species diversity, breeding system and floral traits of orchid species on Reunion Island. *Journal of Biogeography* 32: 1751–1761.
- Kalisz S, Vogler DW. 2003. Benefits of autonomous selfing under unpredictable pollinator environments. *Ecology* 84: 2928–2942.
- Kalisz S, Vogler DW, Hanley KM. 2004. Context-dependent autonomous self-fertilization yields reproductive assurance and mixed mating. *Nature* 430: 884–887.
- Kevan PG, Tikhmenev EA, Usui M. 1993. Insects and Plants in the Pollination Ecology of the Boreal Zone. *Ecological Research* 8: 247–267.
- Kier G, Mutke J, Dinerstein E, Ricketts TH, Kuper W, Kreft H, Barthlott W. 2005. Global patterns of plant diversity and floristic knowledge. *Journal of Biogeography* 32: 1107–1116.
- Klotz S, Kühn I, Durka W. 2002. *BIOLFLOR – eine Datenbank zu biologisch-ökologischen Merkmalen der Gefäßpflanzen in Deutschland. Schriftenreihe für Vegetationskunde*, 38. Bonn, Germany: Bundesamt für Naturschutz.
- Körner C. 2003. *Alpine plant life. functional plant ecology of high mountain ecosystems*. Berlin, Germany: Springer.
- Krause A. 1998. Floras Alltagskleid oder Deutschlands 100 häufigste Pflanzenarten. *Natur und Landschaft* 73: 486–491.
- Kühn I, May R, Brandl R, Klotz S. 2003. Plant distribution patterns in Germany – will aliens match natives? *Feddes Repertorium* 114: 559–573.
- Kühn I, Durka W, Klotz S. 2004. BioFlor – a new plant-trait database as a tool for plant invasion ecology. *Diversity and Distributions* 10: 363–365.
- Kunin WE. 1992. Density and reproductive success in wild populations of *Diploaxis-Erucoides* (Brassicaceae). *Oecologia* 91: 129–133.
- Lande R, Shannon S. 1996. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution* 50: 434–437.
- Legendre P. 1993. Spatial autocorrelation – trouble or new paradigm. *Ecology* 74: 1659–1673.

- Legendre P, Legendre L. 1998. *Numerical ecology. Developments in environmental modelling* 20. Amsterdam, The Netherlands: Elsevier.
- Legendre P, Galzin R, Harmelin-Vivien ML. 1997. Relating behaviour to habitat: solutions to the fourth-corner problem. *Ecology* 78: 547–562.
- Liedtke H, Marcinek J. 2002. *Physische Geographie Deutschlands*. Stuttgart, Germany: Klett-Perthes.
- Lososova Z, Chytrý M, Cimalova S, Kropac Z, Otypkova Z, Pyšek P, Tichý L. 2004. Weed vegetation of arable land in Central Europe: Gradients of diversity and species composition. *Journal of Vegetation Science* 15: 415–422.
- MacGillivray CW, Grime JP, The Integrated Screening Programme (ISP) Team. 1995. Testing predictions of the resistance and resilience of vegetation subjected to extreme events. *Functional Ecology* 9: 640–649.
- Mayfield MM, Ackerly D, Daily GC. 2006. The diversity and conservation of plant reproductive and dispersal functional traits in human-dominated tropical landscapes. *Journal of Ecology* 94: 522–536.
- McIntyre S, Díaz S, Lavorel S, Cramer W. 1999. Plant functional types and disturbance dynamics – introduction. *Journal of Vegetation Science* 10: 604–608.
- Midgley JJ, Bond WJ. 1991a. Ecological aspects of the rise of angiosperms – a challenge to the reproductive superiority hypotheses. *Biological Journal of the Linnean Society* 44: 81–92.
- Midgley JJ, Bond WJ. 1991b. How important is biotic pollination and dispersal to the success of the angiosperms. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* 333: 209–215.
- Niklas KJ. 1985. The aerodynamics of wind pollination. *Botanical Review* 51: 328–386.
- Packer JG. 1974. Differentiation and dispersal in alpine floras. *Arctic and Alpine Research* 6: 117–128.
- Planty-Tabacchi AM, Tabacchi E, Naiman RJ, DeFerrari C, Decamps H. 1996. Invasibility of species rich communities in riparian zones. *Conservation Biology* 10: 598–607.
- Quinn RM, Lawton JH, Eversham BC, Wood SN. 1994. The biogeography of scarce vascular plants in Britain with respect to habitat preference, dispersal ability and reproductive biology. *Biological Conservation* 70: 149–157.
- R Development Core Team. 2004. *Royal: a language and environment for statistical computing*. Vienna, Austria: Royal Foundation for Statistical Computing.
- Regal PJ. 1982. Pollination by wind and animals – ecology of geographic patterns. *Annual Review of Ecology and Systematics* 13: 497–524.
- Richards AJ. 1997. *Plant breeding systems*, 2nd edn. Cheltenham, UK: Stanley Thornes.
- Rusch GM, Pausas JG, Leps J. 2003. Plant functional types in relation to disturbance and land use: Introduction. *Journal of Vegetation Science* 14: 307–310.
- Schoen DJ, Morgan MT, Bataillon T. 1996. How does self-pollination evolve? Inferences from floral ecology and molecular genetic variation. *Philosophical Transactions of the Royal Society of London Series B – Biological Sciences* 351: 1281–1290.
- Settele J, Hammen V, Hulme PE, Karlson U, Klotz S, Kotarac M, Kunin WE, Marion G, O'Connor M, Petanidou T, Peterseon K, Potts S, Pritchard H, Pyšek P, Rounsevell M, Spangenberg J, Steffan-Dewenter I, Sykes MT, Vighi M, Zobel M, Kühn I. 2005. ALARM: Assessing LArge scale environmental Risks for biodiversity with tested Methods. *GIA – Ecological Perspectives in Science, Humanities, and Economics* 14: 69–72.
- Skarpe C. 1996. Plant functional types and climate in a southern African savanna. *Journal of Vegetation Science* 7: 397–404.
- Symank A, Hauke U, Rückriem U, Schröder E. 1998. *Das europäische Schutzgebietssystem Natura 2000. Schriftenreihe für Landschaftspflege und Naturschutz* 53. Bonn–Bad Godesberg, Germany: Landwirtschaftsverlag Hiltrup.
- Statistisches Bundesamt. 1997. *Daten zur Bodenbedeckung für die Bundesrepublik Deutschland 1 : 100.000*. Wiesbaden, Germany: Statistisches Bundesamt.
- Succow M, Joosten H. 2001. *Landschaftsökologische Moorkunde*. Stuttgart: Schweizerbart.
- Tamura S, Kudo G. 2000. Wind pollination and insect pollination of two temperate willow species, *Salix miyabeana* and *Salix sachalinensis*. *Plant Ecology* 147: 185–192.
- Tauber H. 1965. Differential pollen dispersion and the interpretation of pollen diagrams. *Danmarks Geologiske Underøgelse, Raecke* 2 89: 1–70.
- Thomas A, Best N, Lunn D, Arnold R, Spiegelhalter DJ. 2004. *GeoBugs user manual*. <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/geobugs.shtml>
- Thuiller W, Lavorel S, Midgley G, Lavergne S, Rebelo T. 2004. Relating plant traits and species distributions along bioclimatic gradients for 88 *Leucadendron* taxa. *Ecology* 85: 1688–1699.
- Traiser C, Klotz S, Uhl D, Mosbrugger V. 2005. Environmental signals from leaves – a physiognomic analysis of European vegetation. *New Phytologist* 166: 465–484.
- Venables WN, Ripley BD. 2002. *Modern applied statistics with S*, 4th edn. New York, NY, USA: Springer.
- Weihner E, van der Werf A, Thompson K, Roderick M, Garnier E, Eriksson O. 1999. Challenging *Theophrastus*: a common core list of plant traits for functional ecology. *Journal of Vegetation Science* 10: 609–620.
- Whitehead DR. 1968. Wind pollination in the angiosperms: evolutionary and environmental considerations. *Evolution* 23: 28–35.
- Whitehead DR. 1983. Wind pollination: some ecological and evolutionary perspectives. In: Real L, ed. *Pollination biology*. Orlando, FL, USA: Academic Press, 97–108.
- Wright IJ, Reich PB, Westoby M, Ackerly DD, Baruch Z, Bongers F, Cavender-Bares J, Chapin T, Cornelissen JHC, Diemer M, Flexas J, Garnier E, Groom PK, Gulias J, Hikosaka K, Lamont BB, Lee T, Lee W, Lusk C, Midgley JJ, Navas ML, Niinemets U, Oleksyn J, Osada N, Poorter H, Poot P, Prior L, Pyankov VI, Roumet C, Thomas SC, Tjoelker MG, Veneklaas EJ, Villar R. 2004. The worldwide leaf economics spectrum. *Nature* 428: 821–827.
- Wright IJ, Reich PB, Cornelissen JHC, Falster DS, Garnier E, Hikosaka K, Lamont BB, Lee W, Oleksyn J, Osada N, Poorter H, Villar R, Warton DI, Westoby M. 2005. Assessing the generality of global leaf trait relationships. *New Phytologist* 166: 485–496.

Supplementary material

The following supplementary material is available for this article online:

Fig. S1 Legend to the distribution of proportions of pollination types (insect-pollination, wind-pollination and self-pollination), Fig. 2a–c, e.

Fig. S2 Maps of the spatial random effects (spatially correlated model residuals) for the log-ratios: (a) log(entomophily/anemophily) and (b) log(autogamy/anemophily). These maps indicate that there is spatial autocorrelation in the trait compositions that could not be explained by the environmental predictors. Spatial patterns in these maps (areas with relatively large negative or positive values for the random effects) may therefore be used to suggest underlying environmental factors that have not been included in the model.

This materials is available as part of the online article from <http://www.blackwell-synergy.com>