RESEARCH PAPER

Combining spatial and phylogenetic eigenvector filtering in trait analysis

Ingolf Kühn[1,2]*, Michael P. Nobis[3] and Walter Durka[1]

[1]*UFZ, Centre for Environmental Research – UFZ, Department of Community Ecology, Theodor-Lieser-Strasse 4, 06120 Halle, Germany,* [2]*Virtual Institute Macroecology, Theodor-Lieser-Strasse 4, 06120 Halle, Germany,* [3]*Swiss Federal Research Institute WSL, Research Unit Biodiversity and Conservation Biology, Zuercherstrasse 111, 8903 Birmensdorf, Switzerland*

ABSTRACT

Aim To analyse the effects of simultaneously using spatial and phylogenetic information in removing spatial autocorrelation of residuals within a multiple regression framework of trait analysis.

Location Switzerland, Europe.

Methods We used an eigenvector filtering approach to analyse the relationship between spatial distribution of a trait (flowering phenology) and environmental covariates in a multiple regression framework. Eigenvector filters were calculated from ordinations of distance matrices. Distance matrices were either based on pure spatial information, pure phylogenetic information or spatially structured phylogenetic information. In the multiple regression, those filters were selected which best reduced Moran's *I* coefficient of residual autocorrelation. These were added as covariates to a regression model of environmental variables explaining trait distribution.

Results The simultaneous provision of spatial and phylogenetic information was effectively able to remove residual autocorrelation in the analysis. Adding phylogenetic information was superior to adding purely spatial information. Applying filters showed altered results, i.e. different environmental predictors were seen to be significant. Nevertheless, mean annual temperature and calcareous substrate remained the most important predictors to explain the onset of flowering in Switzerland; namely, the warmer the temperature and the more calcareous the substrate, the earlier the onset of flowering. A sequential approach, i.e. first removing the phylogenetic signal from traits and then applying a spatial analysis, did not provide more information or yield less autocorrelation than simple or purely spatial models.

Main conclusions The combination of spatial and spatio-phylogenetic information is recommended in the analysis of trait distribution data in a multiple regression framework. This approach is an efficient means for reducing residual autocorrelation and for testing the robustness of results, including the indication of incomplete parameterizations, and can facilitate ecological interpretation.

Keywords

Central Europe, environmental correlates, phenology, phylogenetic autocorrelation, spatial autocorrelation, trait distribution.

*Correspondence: Ingolf Kühn, UFZ, Centre for Environmental Research – UFZ, Department of Community Ecology, Theodor-Lieser-Strasse 4, 06120 Halle, Germany.
E-mail: ingolf.kuehn@ufz.de

INTRODUCTION

Traditionally, ecological studies have focused on the analysis of species and their environment to unravel a variety of ecological problems. As such analyses are often done in a spatial context, they are frequently influenced by the presence of spatial auto-correlation (SAC). This well-known phenomenon, i.e. when the values of variables sampled at nearby locations are not independent of each other, has recently gained much attention at the level of communities or assemblages (e.g. Legendre, 1993; Lennon, 2000; Diniz-Filho *et al.*, 2003; Dormann, 2007; Hawkins *et al.*, 2007; Kühn, 2007). While there is not necessarily

a bias in using non-spatial techniques in a spatial context, the presence of SAC can have severe effects on parameter estimates (Kühn, 2007) and this is only known after applying adequate spatial methods which reduce residual autocorrelation. Spatial autocorrelation is often regarded as 'nuisance' since the specific process that leads to SAC is often unknown because usually only the spatial structure is considered. In addition this can be distorted by artefactual delimitation of study regions, for example by political rather than 'natural' boundaries. Further reasons for SAC (see, for example, Legendre, 1993, and Dormann *et al.*, 2007, for details) might be of a biological nature (e.g. distance-related biological processes such as speciation, extinction, dispersal or species interactions) as well as of an environmental nature (environmental determinants, which in themselves are spatially structured and thus cause spatial structuring of biological response) or incorrect parameterizations. Anyhow, rigorous analyses of the biological reasons behind SAC are so far lacking.

Recently, the importance of analysing trait compositions of communities or assemblages came into focus (Kühn *et al.*, 2006; McGill *et al.*, 2006). Analyses of traits are often used across species, i.e. in a taxonomic or more general phylogenetic context. As with spatial autocorrelation, phylogenetic data are not independent of each other; the more phylogenetic information species share, the more closely they are related. Therefore, the use of phylogenetic methods was recommended for comparative analyses of communities or assemblages a while ago (Harvey & Pagel, 1991). However, in contrast to SAC, with its often 'technical' characteristics, phylogenetic autocorrelation (PAC) is a result of evolutionary history, common adaptations and common selection pressure or niche conservatism (Diniz-Filho & Bini, 2008; Freckleton & Jetz, 2009). Therefore, PAC is often better understood in biological patterns and processes.

The need for joint consideration of space and phylogeny in comparative analyses was recently recognized for the first time (as far as we are aware) by Diniz-Filho *et al.* (2007) and summarized by Freckleton & Jetz (2009): both spatially structured environmental factors and phylogenetic processes may cause variations in traits. Species that live in similar environments may have common adaptations, which should be correlated with spatial proximity. Additionally, closely related species show high similarity because they share a common evolutionary history. This means that species traits can be conserved across space and phylogeny as a consequence of ecological adaptation and evolutionary history. Therefore, it is advisable to disentangle the roles of phylogenetic and spatial processes in the relationship between traits and their environment (Diniz-Filho *et al.*, 2007; Freckleton & Jetz, 2009). Knapp *et al.* (2008) showed that some communities or localities may, for example, be dominated by specific evolutionary lineages while other can be phylogenetically more diverse. In Germany, for example, the northern lowlands are characterized by wind-pollinated grasslands while southern uplands are good habitats for a wide range of insect-pollinated species (Kühn *et al.*, 2006). While the former are dominated by closely related grass species (the species are pseudoreplicated in a single family), the latter are made up of a wider range of families (and hence are true replicates across families). Besides the bio-

logical reasons, it is therefore clear that it is important to consider phylogenetic structure in space in statistical analyses, and this is more biologically sensible than using pure spatial information.

Structurally, the problem to be solved here looks similar to the classical 'fourth-corner problem' (Dolédec *et al.*, 1996; Legendre *et al.*, 1997). In this problem three corners are represented by three matrices: one with species at sites, one with environmental characteristics per site and one with traits of species. The fourth corner is the one to be constructed: the trait by environment matrix. In principle, we also have these three matrices. However, Legendre *et al.* (1997) explicitly state that their fourth-corner method is not a modelling technique and, for example, does not take spatial structure of environmental variation into account. They continue that a multiple interaction form (e.g. multiple regression) remains to be developed. This also holds true for the complementary approach of Dolédec *et al.* (1996). The fourth-corner approach was recently improved by Dray & Legendre (2008). Still the method cannot take into account autocorrelation between species and between sites. Dray & Legendre (2008) claimed this to be an objective of prime interest. We are therefore unable to use this tool but needed to extend some techniques in which accounting for autocorrelation is already implemented.

As we are unable to use the fourth-corner approach, the joint analysis of spatial and phylogenetic information poses problems: spatial analyses have replicates in space (locations) while analyses with phylogenetic information have replicates across species. A method capable of incorporating spatial as well as phylogenetic information is of advantage, as one cannot know a priori whether SAC or PAC or both are present. Recently, new methods were suggested which combine spatial and phylogenetic information: Diniz-Filho *et al.* (2007) used so-called 'phylogenetic eigenvector filtering' (Diniz-Filho *et al.*, 1998) to partition the phylogenetic and ecological components across species and subsequently explain these signals by environmental variables across Europe by using simultaneous autoregressive models. Freckleton & Jetz (2009) extended the use of phylogenetic independent contrasts (PICs) to incorporate spatial distances for trait analyses across species. Here, we use a complementary approach and extend the eigenvector approach since it is available in both a phylogenetic context and a spatial context (Diniz-Filho & Bini, 2005; Tiefelsdorf & Griffith, 2007). This method can be considered as a unifying strategy for such problems (Peres-Neto, 2006). This approach seemed especially suitable for us, since the calculation of eigenvectors from any kind of information is straightforward. By using simple matrix operations, we can easily transform one kind of information in a way that makes both groups of variables comparable. Furthermore, spatial filtering is among the methods recommended in a recent review (Dormann *et al.*, 2007). Hence we present a new method for the analysis of traits in environmental space, jointly accounting for spatial and spatially structured phylogenetic non-independence of trait distribution and based on eigenvector filtering. In a phylogenetic context, this method is considered to have some advantages over the often-used phylogenetic independent contrasts (PIC; Harvey & Pagel, 1991). PIC analyses have been intensely discussed (see Westoby *et al.*, 1995, and the

following debate). One of the arguments was that PICs do not only remove a pure phylogenetic part of variation but also the variation that covaries with ecological processes. Due to the possibility of partialling out these two effects into pure phylogenetic, pure ecological and joint effects (Desdevises *et al.*, 2003), eigenvector filtering is preferred here. Hence, we extend the method already established by Desdevises *et al.* (2003) by spatial components.

One of the traits that has recently gained increased attention in temporal analyses is flowering phenology of vascular plants, especially in the context of climate warming (e.g. Fitter & Fitter, 2002; Badeck *et al.*, 2004). It has been known for a long time that phenology also differs spatially (e.g. south versus north or lowlands versus mountains; Defila & Clot, 2005) with some limited plasticity within species among populations (e.g. Menzel *et al.*, 2001) but also among communities in different regions defined by different environments (e.g. Defila & Clot, 2005; Menzel *et al.*, 2006). It has been shown repeatedly that the evolution of flowering time may be constrained within certain lineages (Levin, 2006). Thus, phenological patterns at the community or landscape level may be affected by phylogeny (e.g. Johnson, 1993), underlining the importance of phylogenetic relationships in analysing macroecological patterns.

To give an example of the new method we analysed the spatial distribution of flowering phenology, i.e. the average onset of flowering in Switzerland, asking the following questions: (1) which environmental gradients can explain macroecological patterns of phenology, (2) what is the effect of spatial autocorrelation, (3) what is the effect of spatially structured phylogenetic autocorrelation, and (4) what is more important in reducing autocorrelation in our analysis, pure spatial information or phylogenetic information in space?

METHODS

Data sources

Plant distribution data

Distribution data for vascular plants were derived from the governmental 'Biodiversity Monitoring' (BDM) programme of Switzerland. This programme was launched by the Federal Office for the Environment (FOEN) in 2001 to monitor Swiss biological diversity (Weber *et al.*, 2004). One of the central indicators of the BDM focuses on surveying the 'species richness in landscapes' (Z7 indicator). In a systematic national grid, 520 plots of 1 km^2 are surveyed by standardized transect sampling providing a species list for each plot (Plattner *et al.*, 2004). For our study we used species lists of 471 plots recorded between 2001 and 2005 (Fig. 1). Eleven plots near to the border of Switzerland had to be excluded because of missing environmental data. While the average number of species per plot was 220.7 ± 66.7 (mean ± SD), one plot with very low species richness ($n = 3$) and an outlying average onset of flowering was excluded from further analyses. Additionally, three isolated plots with missing neighbours in the later applied spatial context were excluded to provide a consistent data basis for all analyses. The final dataset for model selection contained 456 plots with 103,665 occurrences of 1740 vascular plant species.
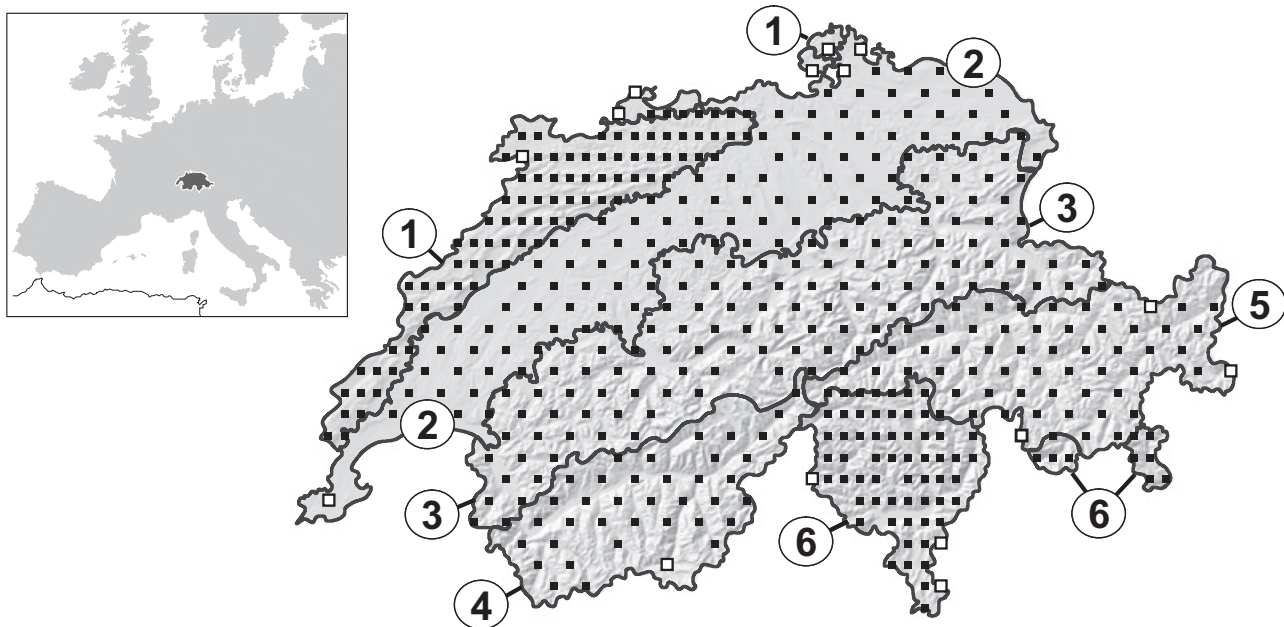

Figure 1 Location of sample plots of vascular plants on the landscape scale (1 km^2) within the framework of Switzerland's Federal Biodiversity Monitoring Programme ($n = 471$) and the six biogeographical regions of Switzerland (1, Jura; 2, Central Plateau; 3, Northern Prealps; 4, Western Central Alps; 5, Eastern Central Alps; 6, Southern Alps) following Gonseth *et al.* (2001). Black squares, used sites; open squares, discarded (marginal) sites.

Table 1 Environmental variables per 1 km² used for models and corresponding to Wohlgemuth *et al.* (2008).

Variable	Description	Derivation
TY	Temperature, mean average (°C)	Zimmermann & Kienast (1999)
PY	Precipitation, mean sum (mm)	ditto
WB7	Water balance, July; sum of precipitation minus potential evapotranspiration	ditto
L.forest	Land cover, closed forest (%)	Bundesamt für Statistik (2001)
L.agrilow	Land cover, agricultural lowlands (%)	ditto
L.urban	Land cover, urban areas (%)	ditto
CALC	Calcareous substrate (%)	De Quervain *et al.* (1963–1967)

Trait data (flowering phenology)

For a species trait, we analysed the onset of flowering. For each species, the month of first flowering was taken from Landolt *et al.* (2009). For each plot, the average onset of flowering was calculated and transformed into the number of days since the beginning of the year. One species (*Helleborus niger*) already started flowering in December (with two locations). We coded this as −1 instead of 12 as the mean of, for example, December and February is in January and not in July. Our recoding provides the same results as using the mean of circular statistics.

Environmental data

All environmental predictors used in this study are available as 1-ha GIS grids and average values were calculated for the each plot of area 1 km². We included seven predictors (Table 1), which are known – at least on the level of single species – to have an impact on the flowering phenology (Roetzer *et al.*, 2000; Defila & Clot, 2005), i.e. mean annual temperature (TY), three land-cover classes, namely forest (L.forest), agricultural areas of the lowlands (L.agrilow) and urban areas (L.urban), calcareous substrate (CALC), mean annual precipitation (PY) and the water balance for July (WB7).

Phylogenetic data

Phylogenetic data were taken from a phylogenetic supertree for the plant species of Germany collated by Durka (2002) in the BiolFlor database. The phylogeny was updated with more recent published data and missing Swiss species were added manually. The phylogeny was purely topological with equal branch lengths and was resolved 78.5%, with 89% of the polytomies at the intrageneric level. Phylogenetic data are provided in comparative analysis by independent contrasts (CAIC) format (Purvis & Rambaut, 1995) by BiolFlor.

Computation
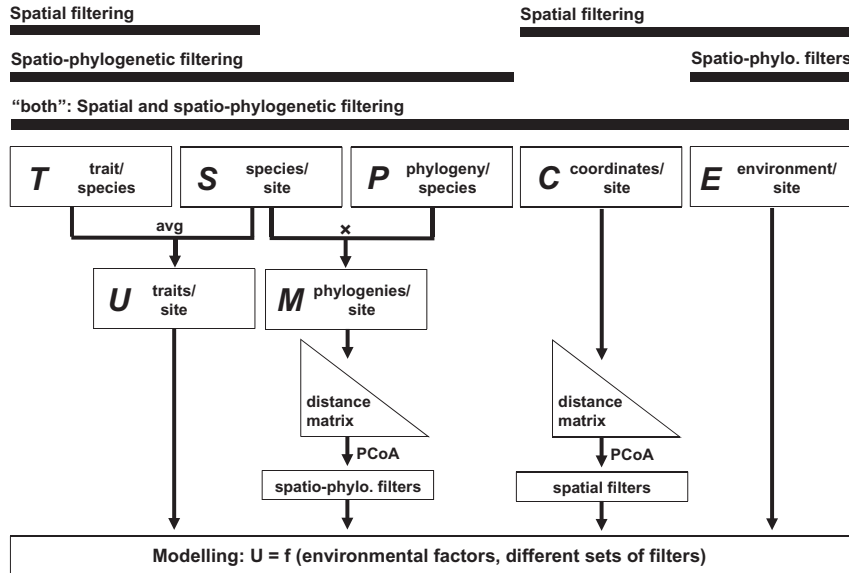
Different filtering approaches

When analysing distribution data, replicates were our spatial locations (i.e. grid cells or plots are rows and species or environmental variables are columns in a matrix) while for phylogenetic analyses, replicates were the species (i.e. species were rows and trait or phylogenetic information were columns in a matrix). Therefore matrix multiplication was used to combine the different sets of matrices (i.e. have trait information or phylogenetic information per location). In doing so, we used four different approaches (Fig. 2): (1) spatial filtering, (2) spatio-phylogenetic filtering, (3) a simultaneous combination of both (all Fig. 2a), and (4) a sequential approach by first using phylogenetic filtering to account for the phylogenetic autocorrelation and then spatial filtering to account for spatial autocorrelation (Fig. 2b). All these approaches will be described in more detail after introducing the basic conventions used. T is the species by trait matrix, S is the site by species matrix, P is the species by phylogeny matrix, C denotes the matrix of coordinates per site and E the environmental variables per site. To construct P, the phylogeny was transformed into a species by phylogenetic branch matrix, recording only presence or absence of a branch per species (see Fig. S1 in Supporting Information). For all analyses (1–3) mentioned above (Fig. 2a), the site by species matrix S was combined with T, resulting in matrix U, such that it contains the information of average trait values per site. The further steps of the three calculations are similar but not equal.

1. In spatial filtering, the coordinates of the sites were transformed into a geographical distance matrix, which then was subjected to a principal coordinates analysis (PCoA) to yield a set of spatial filters (the principal coordinates of the eigenvectors). These spatial filters were then used in a filtering algorithm to reduce spatial autocorrelation (Diniz-Filho & Bini, 2005; Tiefelsdorf & Griffith, 2007) when modelling the relationship between average trait value (i.e. average onset of flowering) per site U and environmental matrix E.
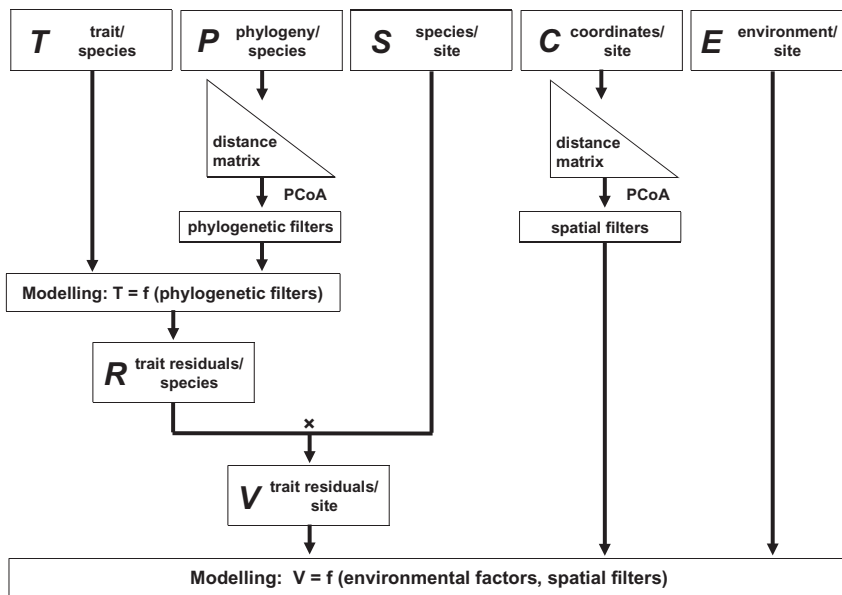
2. In spatio-phylogenetic filtering, S was multiplied with P, resulting in a site by phylogenetic branch matrix M. Thus, the product of the two matrices yielded the phylogenetic branches within each site. This matrix was then subject to a PCoA on a chi-square distance matrix (in fact, we used the computationally faster equivalent, which is a correspondence analysis) to avoid the so-called 'double-zero problem' (Legendre & Legendre, 1998). The resulting eigenvectors represent the spatial structure of the phylogenetic information, called for short spatio-phylogenetic filters. We only entered those with positive eigenvalues into a filtering algorithm (Diniz-Filho *et al.*, 1998; Diniz-Filho & Bini, 2005) to reduce phylogenetic autocorrelation when modelling the relationship between average trait value per site U and environmental matrix E.

3. When simultaneously combining 'both' approaches described above, we used the first half of the spatial eigenvectors

(a): single and simultaneous filtering: (i) spatial, (ii) spatio-phylognetic filtering and (iii) "both"

Spatial filtering

Spatial filtering

Spatio-phylogenetic filtering

Spatio-phylo. filters

"both": Spatial and spatio-phylogenetic filtering

Modelling: U = f (environmental factors, different sets of filters)

(b): sequential filtering: 1. phylogenetic filtering, 2. spatial filtering

Modelling: V = f (environmental factors, spatial filters)

and the first half of the spatio-phylogenetic eigenvectors (so that the number of eigenvectors did not exceed the number of cases) as the initial set of filters to remove autocorrelation when modelling the relationship between U and E. We also checked that pure spatial or spatio-phylogenetic eigenvector filtering did not yield relevant eigenvectors from the latter half of the vectors. Theoretical considerations, however, suggest that information which is relevant for autocorrelation is only represented by the first eigenvectors, whereas the latter eigenvectors represent idiosyncratic information (or noise) of the elements in the analysis. Note that within one group (spatial or spatio-phylogenetic) all filters are orthogonal to each other. In the combined analysis, however, spatial filters are to be correlated with selected spatio-phylogenetic filters (Tables S1 & S2).

4. Additionally, we used a sequential approach to first remove the phylogenetic signal from the data and then use spatial filtering (Fig. 2b). We calculated the patristic distance (Desdevises *et al.*, 2003) from the species by phylogeny matrix P (see Fig. S1). This is a triangular matrix adding up pairwise all branches between two species. This distance matrix was subjected to a PCoA and the positive eigenvectors were regressed on the trait value (onset of flowering) per species to reduce phylogenetic autocorrelation (Diniz-Filho *et al.*, 1998; Desdevises *et al.*, 2003). The residuals of this regression (matrix R) were combined with matrix S to yield average trait values, corrected for phylogeny (matrix V). The computation of spatial eigenvectors is equal to a purely spatial filtering (approach 1). These spatial filters were then used in a filtering algorithm to reduce

Figure 2 General overview on the methodological approaches used in this study. Panel (a) summarizes the first three approaches, namely 'spatial filtering', 'spatio-phylogenetic filtering' and simultaneously filtering 'both' spatial and spatio-phylogenetic information. The matrices and pathways used by each of the approaches are headed by a bold bar on top. In all three approaches, the trait per species matrix *T* and the species per site matrix *S* are combined so that the resulting matrix *U* averages trait values (onset of flowering) per grid cell. In 'spatial filtering' the site coordinates are transformed into a distance matrix which then is subjected to a principal coordinates analysis (PCoA) to receive the spatial filters (i.e. the principal coordinates or eigenvectors of the PCoA). In 'spatio-phylogenetic filtering', the species per site matrix *S* and the phylogeny per species matrix *P* are multiplied to receive a matrix *M* of phylogenetic information per site which is transformed into a distance matrix and subjected to a PCoA to receive spatio-phylogenetic filters. In the approach using both types of filters, the first half of each of the sets was combined. The resulting filters are used as covariates in modelling the relationship between trait distribution and environmental variation so that the selected filters reduce the autocorrelation to a non-significant level (see methods for details). Panel (b) summarizes the fourth approach, which is a sequential filtering of phylogenetic and spatial information. In the first step, pure phylogenetic information from matrix *P* is transformed into a patristic distance matrix and then subjected to a PCoA to receive phylogenetic filters (analogous to the spatial filters). These phylogenetic filters are then used to remove phylogenetic information form the trait (onset of flowering) per species matrix *T* in a multiple regression framework. The residuals matrix *R* is then multiplied with the species per sites matrix *S* to bring the residual information of onset of flowering excluding phylogenetic signals into space (matrix *V*). From matrix *C* spatial filters are computed as depicted in panel (a). Similarly to the other approaches, these are used as covariates in modelling the relationship between residual trait distribution and environmental variation so that the selected filters reduce the autocorrelation to a non-significant level (see methods for details).

spatial autocorrelation (Diniz-Filho & Bini, 2005; Tiefelsdorf & Griffith, 2007) when modelling the relationship between *V* and environmental matrix *E*.

Reducing autocorrelation by filter selection

To reduce residual autocorrelation in all of the approaches mentioned above we used Moran eigenvector filtering (Dray *et al.*, 2006; Griffith & Peres-Neto, 2006). This approach is roughly comparable to the third variant of spatial eigenvector mapping used by Bini *et al.* (2009), which was one of the most robust methods in their comparison. The eigenvectors to be included in the regression models are chosen by calculating the empirical Moran's *I* values for the initial model. Subsequently, all eigenvectors are scrutinized and the first eigenvector is chosen as that with the lowest Moran's *I* value. This procedure is repeated to find sequentially the best set of eigenvectors to reduce residual autocorrelation below the $\alpha = 0.05$ level of significance. All probability values used (for predictors as well as Moran's *I*) are based on using the *z* statistics calculated from the mean and standard deviation of 999 permutations. Calculations were performed using a modification of the R function ME (originally written by Bivand & Peres-Neto in 'spdep'; Bivand *et al.*, 2006).

For this approach, it is necessary to define a neighbourhood matrix, i.e. to define which spatial locations or taxa are in a given neighbourhood and which ones are not. A heuristic approach using several neighbourhood distances was used. For purely phylogenetic eigenvectors, we used distances of 2, 3, 5, 7, 10 and 25 unit branch lengths. Species in a neighbourhood up to 3 unit branch lengths yielded the best results [lowest Moran's *I* and lowest Akaike information criterion (AIC)]. For eigenvectors with a spatial component (purely spatial eigenvectors and spatially structured phylogenetic eigenvectors) we used maximum distances of 10, 15, 25 and 100 km. In accordance with previous analyses on comparable scales (Koellner *et al.*, 2004; Kühn, 2007; Tautenhahn *et al.*, 2008), lag distances above 25 km were not relevant in removing spatial autocorrelation. Best results (i.e. the

lowest Moran's *I* and lowest AIC; Kissling & Carl, 2008) were yielded for the 10 km distance. Hence we used a lag distance of 10 km for all further analyses, i.e. all plots up to 10 km distance are included as neighbours.

Model selection in multiple linear regression

To set up our initial model we used seven environmental variables for which we assumed a strong influence on the trait pattern. Starting with the full model of all environmental predictors considered as linear terms in multiple linear regression, variable selection was carried out by backward selection based on the AIC. The non-filtering approach left six variables in the minimum adequate model – all of them being significant ($P < 0.05$). Since autocorrelation usually causes further variables to become insignificant but not insignificant variables to become significant, these six variables were entered into the filtering approaches. Influential plots and outliers were tested by examining regression diagnostics (residuals versus fitted values, normal *Q*–*Q* plots, and Cook's distance plots). All analyses were performed using R version 2.6.2 (R Development Core Team, 2008).

Partitioning the variation

We partitioned the variation in the final models among environmental information, eigenvectors accounting for spatial autocorrelation and eigenvectors accounting for autocorrelation of spatio-phylogenetic information by using hierarchical partitioning (Chevan & Sutherland, 1991) in R using 'hier.part' (Mac Nally & Walsh, 2004). In this method the explained variance is calculated for generalized linear models using all possible combinations of independent variables. From this, we calculated the proportions of the variance which could exclusively be explained by a particular variable (or group of variables, such as spatial or spatio-phylogenetic eigenvectors, respectively).

Average Onset of Flowering

- ● 1 - 3 May
- ● 4 - 7 May
- ● 8 - 9 May
- ● 10 - 11 May
- ● 12 - 14 May
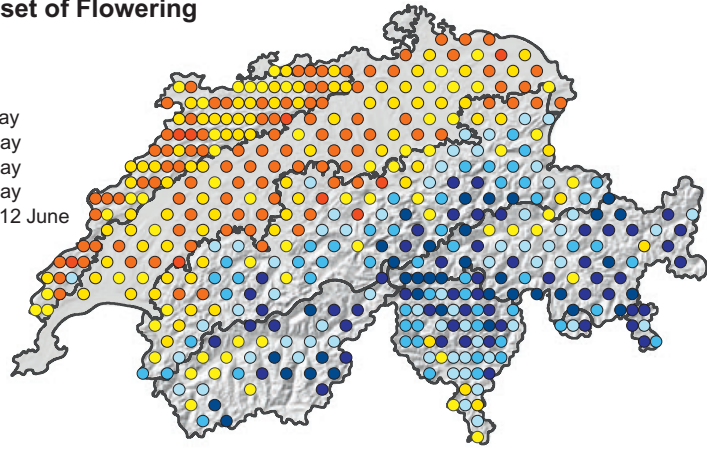- ● 15 - 19 May
- ● 20 - 29 May
- ● 30 May - 12 June

Figure 3 Spatial distribution of average onset of flowering of vascular plant species in Swiss floristic sample locations ($n = 456$).

Table 2 Model coefficients of the relationship between average onset of flowering and environmental predictors in Switzerland: no additional filters (Non), spatial filters (Spatial), spatially structured phylogenetic filters (Spatio-phylo), both filters (Both), as well as the residuals of pure phylogenetic filters on the traits (Filtering (residuals)) without further spatial filtering (Non) and with sequential phylogenetic and spatial filtering (Spatial).

Coefficients	Filtering				Filtering (residuals)	
	Non	Spatial	Spatio-phylo	Both	Non	Spatial
Intercept	131.22***	131.22***	131.22***	131.22***	−5.61***	−5.61***
Annual temperature	−7.81***	−8.00***	−8.92***	−8.13***	−5.31***	−5.48***
Calcareous substrate	−3.04***	−2.40***	−0.69*	−2.51***	−2.22***	−1.74***
Closed forest	−0.52*	−0.52*	0.05 ns	−0.29 ns	−0.27 ns	−0.28 ns
Urban areas	0.82**	0.94***	−0.05 ns	1.15***	0.62**	0.62**
Annual precipitation	−0.69*	−0.16 ns	−0.97***	−0.58(*)	−0.67**	−0.32 ns
July water balance	0.63*	−0.16 ns	1.44***	0.45 ns	0.52*	−0.11 ns
Global Moran's I	0.160***	0.050 ns	0.049 ns	0.045 ns	0.136***	0.042 ns
Filters selected		S3, S49, S8, S35, S57, S10, S40, S51, S17	P4, P27, P3, P5, P94, P147	S3, P32, S35, P45, P4, P27		S3, S10, S18, S57, S49, S8
AIC	2642.4	2586.6	2518.3	2584.5	2413.9	2370.0
R^2	0.819	0.846	0.866	0.845	0.775	0.801
R^2 crossvalidated	0.812	0.833	0.837	0.821	0.767	0.789
MAE	3.342	3.138	2.795	3.038	2.601	2.472

Filter were selected according their ability to reduce Moran's I coefficient of autocorrelation significantly ($\alpha = 0.05$); neighbourhood distance ≤ 10 km; see methods for details. AIC, Akaike information criterion; MAE, mean absolute error.
(*)$0.05 < P \leq 0.1$, *$0.01 < P \leq 0.05$, **$0.001 < P \leq 0.01$, ***$P \leq 0.001$; ns, not significant.

RESULTS

As expected, there is a clear spatial structure of the average onset of flowering across Switzerland (Fig. 3) with a strong gradient of early flowering in the lowlands and a retarded average onset of flowering by up to as much as 6 weeks in the Alps.

Stepwise regression during model selection excluded one land-cover variable for agricultural areas in the lowlands (L.agrilow). All other predictor variables remained significant in the non-filtering model (Table 2) and no outlier or highly influential plot was detected.

All models have quite a high proportion of variation explained. Ignoring the effect of autocorrelation, the most important factors influencing the beginning of flowering were

mean annual temperature (negative), the proportion of calcareous substrate (negative) and the proportion of closed forests (negative) (Fig. 4a, Table 2). Spatial autocorrelation in the first lag distance was considerable (Fig. S2a). The inclusion of spatial filters increased the fit and variance explained and reduced autocorrelation significantly (Table 2, Fig. S2b), but had substantial effects on parameter estimates with mean annual precipitation and July water balance becoming insignificant (Table 2). The inclusion of spatio-phylogenetic instead of spatial filters increased model fit and explained variance further and reduced global autocorrelation. Compared with the non-filtering model, proportions of closed forests and urban areas became insignificant (Table 2). The independent effect of each of the filters was relatively small (Fig. 4c). In the simultaneous model, annual
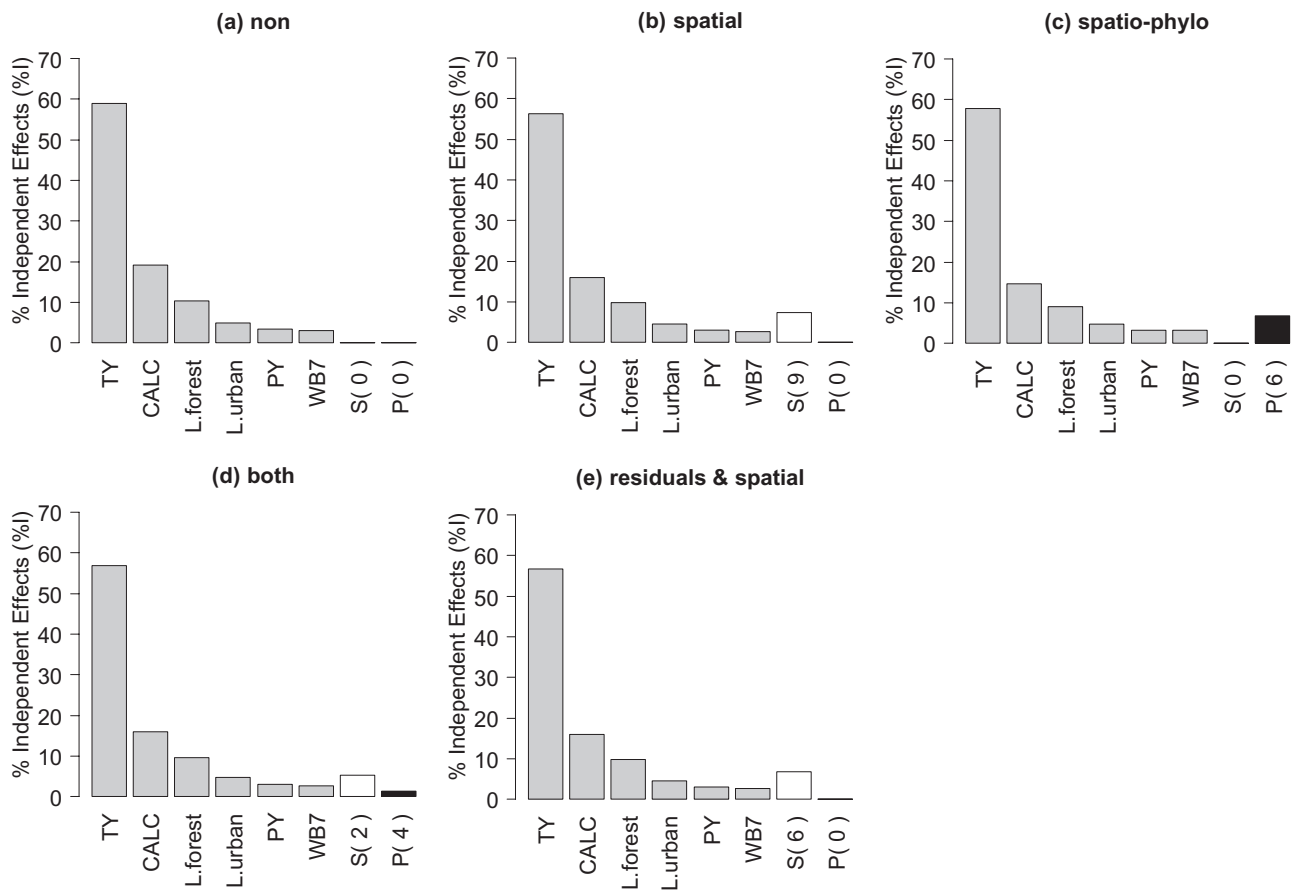
Figure 4 Hierarchical partitioning of the independent effects of environmental variables as well as spatial and phylogenetic filters (being abbreviated S or P, respectively, with the number of filters in parenthesis) of the full set of environmental predictors explaining the variance in average onset of flowering in Switzerland: simple model without filtering (a); spatial filtering (b); spatio-phylogenetic filtering (c); simultaneous filtering of spatial and spatio-phylogenetic filters (d); sequential filtering by first removing the phylogenetic effect of the trait and then apply a spatial filtering approach (e); bars of environmental variables, spatial and phylogenetic filters are shown in grey, white and black, respectively. For abbreviations of environmental variables see Table 1.

precipitation became marginally significant and closed forests as well as water balance became insignificant (Table 2).

The explained variation of the pure phylogenetic model (i.e. regressing filters on traits) was 46.8% ($R^2_{adj} = 0.435$). Using the residuals of this model without further filtering, results and autocorrelation were comparable with those of the simple model without any filters. The sequential approach yielded results comparable to pure spatial filtering, but in addition proportions of closed forests were insignificant. Explained variance is considerably lower than in any of the first three filtering approaches (Table 2). However, the order and effect size of the environmental predictors is remarkably robust among all models (Fig. 4).

To test how the method performs when ignoring an important environmental predictor, we excluded the proportion of calcareous substrate. This had an effect not only on parameter estimates (Table 3) but especially on the autocorrelation structure, largely expanding the lag distance (Fig. S3a,e), with the choice of filters (the first spatio-phylogenetic filter becoming important) and spatio-phylogenetic filters as such becoming

much more important (Fig. 5c,d). This, however, is an effect of some high degree of collinearity between the spatio-phylogenetic filters and the proportion of calcareous substrate.

In both cases, including or ignoring calcareous substrate, more spatial filters ($n = 9$ vs. $n = 15$, respectively) were necessary to reduce Moran's I below the specified level of significance than when using spatio-phylogenetic filters ($n = 6$ vs. $n = 4$) or both simultaneously ($n = 6$ vs. $n = 4$). The highest eigenvectors included in the spatial analysis was no. 57 vs. no. 62, the highest eigenvector of the spatio-phylogenetic analysis was no. 147 vs. no. 22, and for the simultaneous analysis S35 and P45 vs. S1 and P22, respectively. At the same time, global Moran's I was better reduced with simultaneous use of spatial and spatio-phylogenetic filters. Furthermore, the selected filters of the simultaneous approach were usually of lower dimensions than the other ones. The sequential filtering approach needed 101 phylogenetic eigenvectors, with the first selected being P2 and finishing with P570, and six spatial eigenvectors.

There is some correlation between the spatial filter S3 and spatio-phylogenetic filter P5 (Table S1). All correlation coeffi-

Table 3 Model coefficients of the relationship between average onset of flowering and environmental predictors in Switzerland but excluding calcareous substrate: no additional filters (Non), spatial filters (Spatial), spatially structured phylogenetic filters (Spatio-phylo), both filters (Both), as well as the residuals of pure phylogenetic filters on the traits (Filtering (residuals)) without further spatial filtering (Non) and with sequential phylogenetic and spatial filtering (Spatial).

Coefficients	Filtering				Filtering (residuals)	
	Non	Spatial	Spatio-phylo	Both	Non	Spatial
Intercept	131.22***	131.22***	131.22***	131.22***	−5.61***	−5.61***
Annual temperature	−8.57***	−9.03***	2.56***	2.68***	−5.87***	−6.12***
Closed forest	−0.81**	−0.65**	0.31*	0.26(*)	−0.49*	−0.40*
Urban areas	0.67*	0.85***	0.53***	0.56***	0.52*	0.55**
Annual precipitation	−0.47 ns	−0.72*	−0.19 ns	−0.22 ns	0.17 ns	−1.08***
July water balance	−0.19 ns	−0.24 ns	0.20 ns	0.34(*)	0.09 ns	−0.28 ns
Global Moran's I	0.395***	0.045 ns	0.052 ns	0.049 ns	0.346***	0.050 ns
Filters selected		S3, S4, S8, S9, S57, S35, S5, S51, S22, S49, S62, S18, S13, S26, S52	P1, P3, P22, P18	P1, P3, P22, S1		S3, S4, S8, S9, S18, S57, S5, S11
AIC	2792.7	2596.5	2129.0	2140.0	2548.5	2377.3
R^2	0.747	0.846	0.942	0.941	0.697	0.799
R^2 crossvalidated	0.740	0.830	0.934	0.932	0.687	0.784
MAE	4.080	3.129	1.816	1.861	3.061	2.486

Filter were selected according their ability to reduce Moran's I coefficient of autocorrelation significantly ($\alpha = 0.05$); neighbourhood distance ≤ 10 km; see methods for details. AIC, Akaike information criterion; MAE, mean absolute error.
(*)$0.05 < P \leq 0.1$, *$0.01 < P \leq 0.05$, **$0.001 < P \leq 0.01$, ***$P \leq 0.001$; ns, not significant.

cients are, however, below $r = 0.3$. Combining both filters in one model resulted in the selection of negligible correlation between spatial and spatio-phylogenetic filters (Table S2).

The distribution of the first selected spatial and spatio-phylogenetic filters in Switzerland is quite interesting (Fig. 6). Spatial filters exhibit a regular wave-like distribution (Fig. 6a–c) with the lower dimensions (S3, S8; Fig. 6a,c) having lower resolution (longer wavelength) and the higher dimension (S49; Fig. 6b) having higher resolution (shorter wavelength, like rolling hills). Spatio-phylogenetic filters yielded a clear spatial pattern (Fig. 6d–f) which – in contrast to spatial filters – strongly reflects the biogeographical zonation of Switzerland (Fig. 1). Here, the higher dimension (P27) subdivides the regions identified by P3/P4 with an indicated east–west gradient for the Central Plateau and Jura.

DISCUSSION

The key results were: (1) all the methods used were able to reduce spatial autocorrelations and modify the results of the regression analysis compared with a non-filtering approach; (2) spatio-phylogenetic eigenvectors provided more information in terms of explained variation and reduced spatial autocorrelation more efficiently than spatial eigenvectors; (3) sequential filtering provided the least insight into ecological processes and required a large number of phylogenetic filters to remove phylogenetic autocorrelation from the traits; and (4) the spatio-phylogenetic eigenvectors were best for interpreting data in an ecological and biogeographical context.

Spatial phylogenetic filtering

Prior to the discussion of the results, the question is whether our proposed new method is mathematically sound. There is a mathematical foundation for the use eigenvectors of (1) geographical distances (for spatial analyses) (Diniz-Filho & Bini, 2005; Tiefelsdorf & Griffith, 2007) or (2) patristic distance (for phylogenetic analyses) (Diniz-Filho *et al.*, 1998; Desdevises *et al.*, 2003) for filtering approaches. We combined the approaches by multiplying the phylogeny per species matrix with the species per site matrix. The resulting distance matrix provides spatially structured phylogenetic information. Unlike a true spatial distance matrix, in which distance is based on Euclidean (= geographical) distance, distance among objects in our approach is based on the phylogenetic distance among spatial locations. Thus, the approach is conceptually comparable to both strict phylogenetic eigenvector filtering and strict spatial eigenvector filtering.

Tiefelsdorf & Griffith (2007) showed how their spatial filtering approach is mathematically related to genuine autoregressive methods. Hence, their approach is not only capable of reducing Moran's I effectively, but is also mathematically valid. However, spatial and spatio-phylogenetic eigenvectors are not orthogonal to each other, therefore there is the potential for some collinearity to occur. It is therefore advisable to test for collinearity among the filters first to avoid problems caused by multicollinearity (Graham, 2003). Because all eigenvectors are screened sequentially for their ability to reduce Moran's I (and not according to their explained variance or statistical significance), multicollinearity should hence not cause any problem in
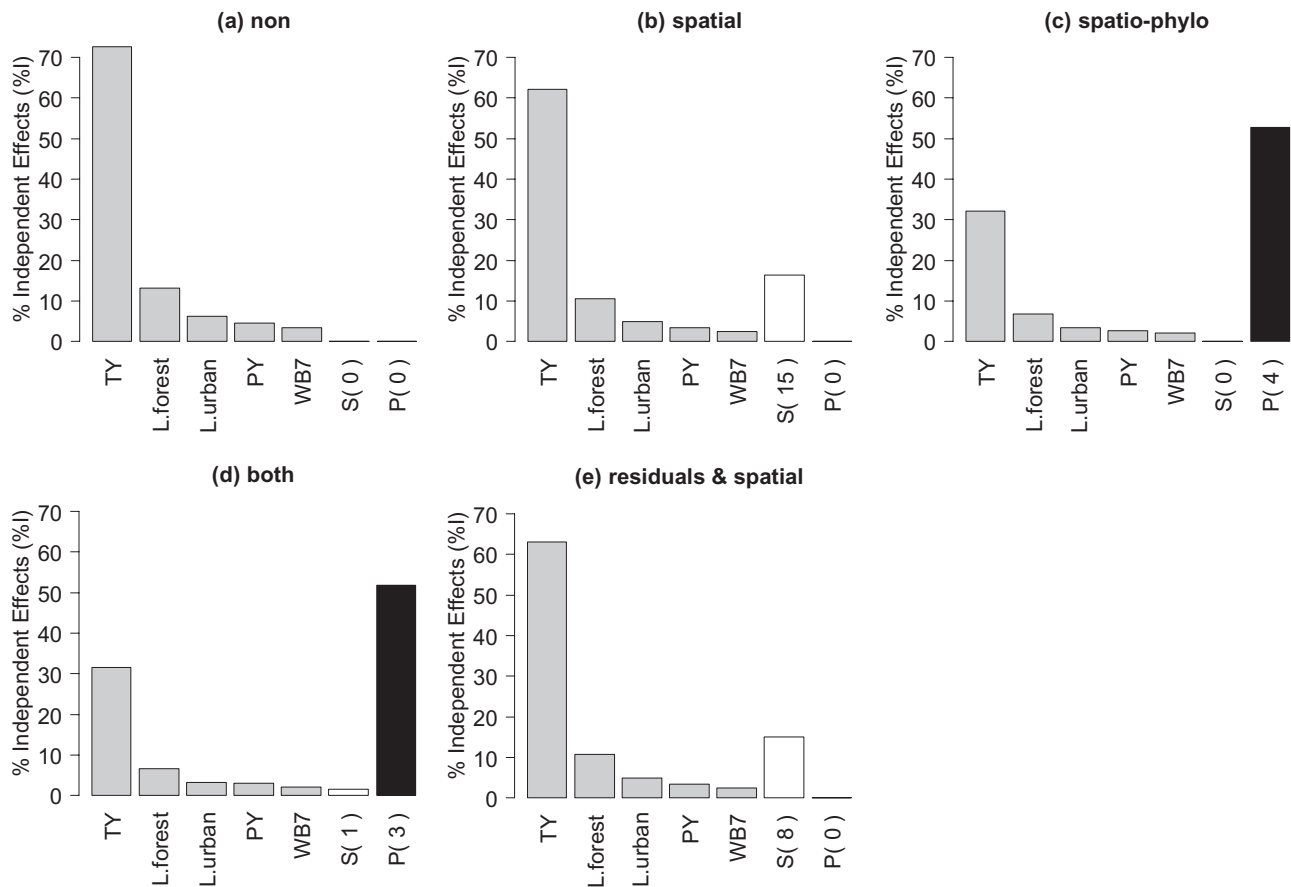
Figure 5 Hierarchical partitioning of the independent effects of environmental variables as well as spatial and phylogenetic filters (being abbreviated S or P, respectively, with the number of filters in parenthesis) of the reduced set of environmental predictors (excluding calcareous substrate) explaining the variance in average onset of flowering in Switzerland: simple model without filtering (a); spatial filtering (b); spatio-phylogenetic filtering (c); simultaneous filtering of spatial and spatio-phylogenetic filters (d); sequential filtering by first removing the phylogenetic effect of the trait and then apply a spatial filtering approach (e); bars of environmental variables, spatial and phylogenetic filters are shown in grey, white and black, respectively. For abbreviations of environmental variables see Table 1.

our type of analysis. Although there was some degree correlation between spatial and spatio-phylogenetic filters, only filters which were hardly correlated were selected in the combined analysis. However, though spatial and spatio-phylogenetic filters were slightly correlated, adding the spatio-phylogenetic context yielded a better statistical model (better fit, fewer filters selected, less residual autocorrelation) and provided more ecological relevant information for the interpretation of the analysis than only the spatial model.

Recently, Freckleton & Jetz (2009) also used a method to simultaneously account for spatial and phylogenetic effects in trait analyses. Their approach is based on the use of phylogenetic contrasts implemented through a generalized least squares (GLS) approach, while we used eigenvector filtering. Both approaches are extensions of conventional methods, yet their approach is based on species as replicates in the analysis while ours is based on using spatial samples as replicates. Their approach averages trait variation within species and environmental variation across sites to a mean value. It is hence unable to account for spatial variation in traits and in environmental factors. In a stepwise approach, Diniz-Filho *et al.* (2007) first

used eigenvector filtering to partition phylogenetic and ecological components. Environmental information per species was therefore averaged for each of the environmental predictors across as in Freckleton & Jetz (2009). Secondly, they averaged the ecological and phylogenetic components again across space, in a similar way to our approach, and explained them in a spatial simultaneous autoregressive model. All three approaches are complementary and designed to answer different questions. Using the approach of Diniz-Filho *et al.* (2007) might potentially result in some tautology, as environmental predictors are used in both steps. However, unlike our approach, their approach clearly distinguishes between purely phylogenetic and purely spatial effects. The approach of Freckleton & Jetz (2009) can help to disentangle whether evolutionary history or environment are more important in determining trait variation across species. Our spatio-phylogenetic filtering approach allows one to account for the role of phylogeny when unravelling the relationship between trait variation and spatially structured, heterogeneous environments. Therefore, it is possible to account for the fact that species which are closely related often share similar environments in a joint analysis of environmental pre-
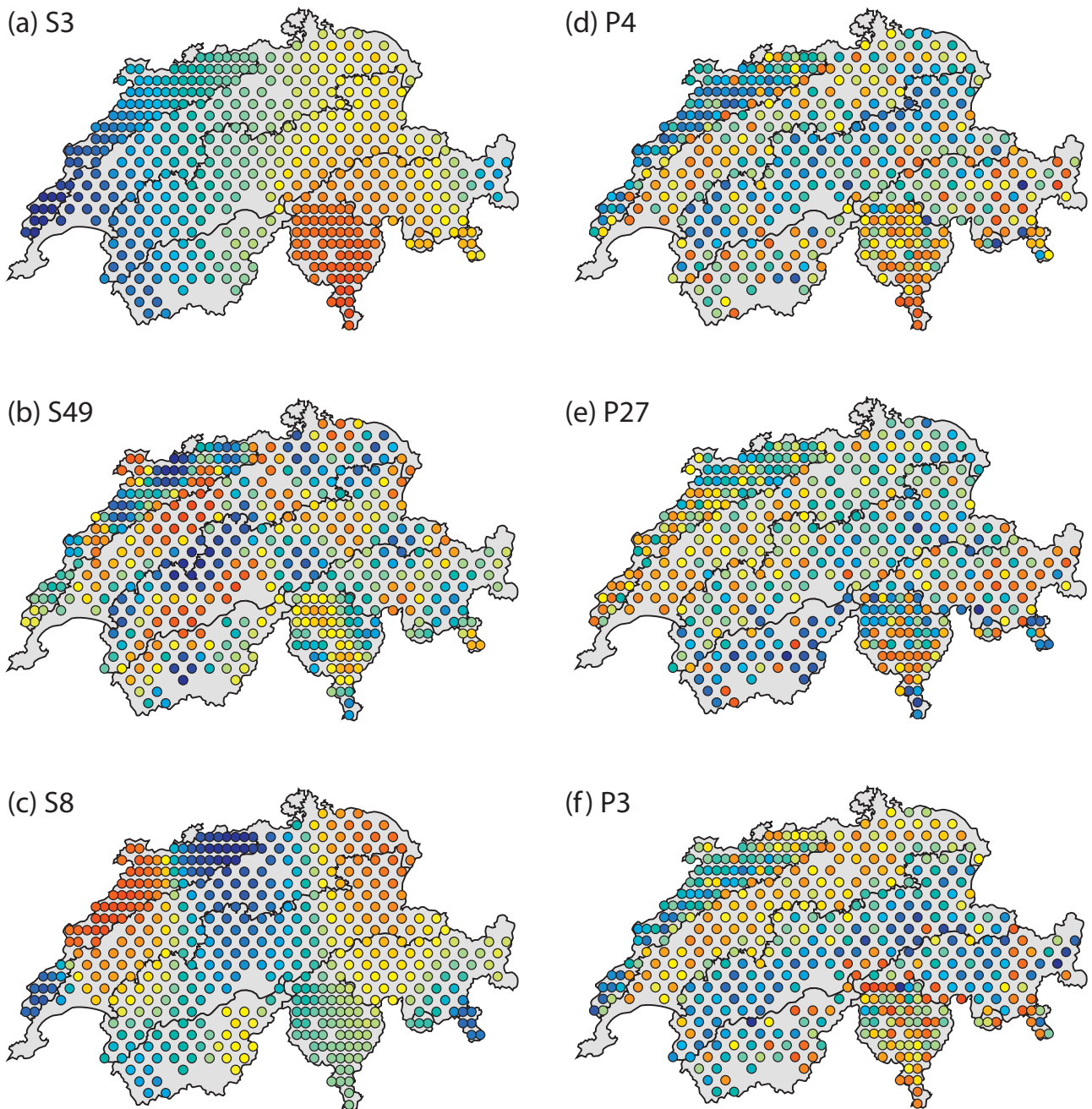
Figure 6 Distribution of the first three selected filters using only spatial filters (a–c) and spatially structured phylogenetic filters (d–f). Cold (blue) to warm (red) colours represent negative to positive axis scores of the plots. Class limits were chosen according 'Natural Breaks (Jenks)' in ArcMap 9.2.

dictors. Unlike Diniz-Filho *et al.* (2007), we combine phylogenetic information and spatial information jointly. Our fourth (sequential) approach, however, is very similar to the one used by Diniz-Filho *et al.* (2007) but only using the equivalent of their ecological components (ignoring environmental predictors, however, at that stage).

Basing analyses on trait variation in space without accounting for phylogenetic effects can be potentially misleading in the interpretation of results, for example by assuming that specific relationships are caused by a response of a species' trait to the

environment whereas in fact this signal might only result from the common evolutionary history of many species at sites with similar environmental conditions.

Spatio-phylogenetic filters unravel ecological patterns of flowering phenology

Flowering phenology and its relationships to the environment are well documented (e.g. Menzel *et al.*, 2001, 2006; Badeck *et al.*, 2004; Defila & Clot, 2005) and we did not intend to

unravel new relationships. The relationship with temperature is obvious. Specific land use and soil characteristics were included due to their straightforward and known relationships with microclimate (open calcareous habitats are usually warmer and many arable weeds need to flower before the crops grow tall or before disturbance occurs).

The selected spatially structured phylogenetic eigenvectors, however, provided ecologically meaningful information, as they mirrored the biogeographical units of Switzerland. These units are defined by their characteristic species inventories, although some regions are also different in the density of sampling locations. Nevertheless, even in regions with the same sampling plot density, the selected spatio-phylogenetic filters capture the transition between biogeographical regions (e.g. Central Plateau and Jura or Northern Prealps). The spatio-phylogenetic eigenvectors hence capture the phylogenetic differences of the different species inventories which are relevant for the species trait under consideration. Furthermore, due to niche conservatism (e.g. Prinzing *et al.*, 2001), phylogenetically related species not only share similar traits with a higher probability but are also more likely to respond in a similar way to environmental pressures. Hence, spatially structured environmental variables will lead to spatially structured species assemblages (or communities) as well as trait assemblages (Kühn *et al.*, 2006; Tautenhahn *et al.*, 2008) sharing a common phylogenetic history. Due to this relationship, spatio-phylogenetic filters can make up for some of the environmental information missing in an analysis in an ecologically sensible way. Therefore, spatio-phylogenetic filters (although selected according to their ability to reduce residual autocorrelation) account for more information than their strictly spatial counterparts and increase even more in explained variability when important key variables (such as calcareous substrate in our example) are missing. This approach can therefore help us to better realize whether important predictors are missed from a parameterized model.

Phylogenetic incomplete knowledge

One problem might be the lack of branch lengths in our phylogeny. Of course, the use of branch lengths is preferred over measures relying solely on topology. However, as in our case, proper branch lengths are often not available, but (super)tree topologies mostly are. Under these circumstances, substituting branch lengths by the number of nodes (i.e. setting branch lengths to unity) can be a fruitful alternative (Schweiger *et al.*, 2008). In principle, the chosen methods would work just the same with correct branch lengths and obviously would provide better (i.e. more realistic) and more differentiated results. Still, incorporating even suboptimal phylogenetic information is better than ignoring such information (Schweiger *et al.*, 2008).

CONCLUSIONS

When analysing traits in spatial context it might be better to use spatially structured phylogenetic information to account for autocorrelation than pure spatial information. The sequential approach as implemented here, however, is not recommended because: (1) it removes the complete phylogenetic signal before being applied in spatial context and related to environmental predictors, and (2) it requires much more phylogenetic information than the simultaneous approach due to the lack of appropriate ecological predictors. An alternative option would be the related approach of Diniz-Filho *et al.* (2007). Eigenvector filtering may also be a way to solve the remaining challenge and account for spatial and phylogenetic autocorrelation in the fourth-corner problem (Dray & Legendre, 2008). A tangible implementation, however, needs to be elaborated. Overall, we were able to show that a combination of spatial and spatio-phylogenetic eigenvector filters is an effective and efficient tool in trait analyses to reduce unwanted spatial autocorrelation, to indicate the absence of important environmental covariates and to provide aids in interpreting the results of trait–environment relationships.

REFERENCES

Badeck, F.W., Bondeau, A., Bottcher, K., Doktor, D., Lucht, W., Schaber, J. & Sitch, S. (2004) Responses of spring phenology to climate change. *New Phytologist*, **162**, 295–309.

Bini, L.M., Diniz-Filho, J.A.F., Rangel, T.F.L.V., Akre, T.S.B., Albaladejo, R.G., Albuquerque, F.S., Aparicio, A., Araújo, M.B., Baselga, A., Beck, J., Bellocq, M.I., Böhning-Gaese, K., Borges, P.A.V., Castro-Parga, I., Chey, V.K., Chown, S.L., de Marco, P., Dobkin, D.S., Ferrer-Castán, D., Field, R., Filloy, J., Fleishman, E., Gómez, J.F., Hortal, J., Iverson, J.B., Kerr, J.T., Kissling, W.D., Kitching, I.J., León-Cortés, J.L., Lobo, J.M., Montoya, D., Morales-Castilla, I., Moreno, J.C., Oberdorff, T., Olalla-Tárraga, M.Á., Pausas, J.G., Qian, H., Rahbek, C., Rodríguez, M.Á., Rueda, M., Ruggiero, A., Sackmann, P., Sanders, N.J., Terribile, L.C., Vetaas, O.R. & Hawkins, B.A. (2009) Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. *Ecography*, **32**, 193–204.

Bivand, R., Anselin, L., Berke, O., Bernat, A., Carvalho, M., Chun, Y., Dormann, C., Dray, S., Halbersma, R., Lewin-Koh, N., Ono, H., Peres-Neto, P., Tiefelsdorf, M. & Yu, D. (2006) *spdep: spatial dependence: weighting schemes, statistics and models*. R package version 0.3-22. http://cran.r-project.org.

Bundesamt für Statistik (2001) *GEOSTAT Benutzerhandbuch*. Bundesamt für Statistik, Neuchâtel, Switzerland.

Chevan, A. & Sutherland, M. (1991) Hierarchical partitioning. *American Statistician*, **45**, 90–96.

Defila, C. & Clot, B. (2005) Phytophenological trends in the Swiss Alps, 1951–2002. *Meteorologische Zeitschrift*, **14**, 191–196.

De Quervain, F., Hofmänner, F., Jenny, V., Köppel, V. & Frey, D. (1963–1967) *Geotechnische Karte der Schweiz, 1 : 200 000*, 2nd edn. Kümmerly & Frei, Bern, Switzerland.

Desdevises, Y., Legendre, P., Azouzi, L. & Morand, S. (2003) Quantifying phylogenetically structured environmental variation. *Evolution*, **57**, 2647–2652.

Diniz-Filho, J.A.F. & Bini, L.M. (2005) Modelling geographical patterns in species richness using eigenvector-based spatial filters. *Global Ecology and Biogeography*, **14**, 177–185.

Diniz-Filho, J.A.F. & Bini, L.M. (2008) Macroecology, global change and the shadow of forgotten ancestors. *Global Ecology and Biogeography*, **17**, 11–17.

Diniz-Filho, J.A.F., De Sant'ana, C.E.R. & Bini, L.M. (1998) An eigenvector method for estimating phylogenetic inertia. *Evolution*, **52**, 1247–1262.

Diniz-Filho, J.A.F., Bini, L.M. & Hawkins, B.A. (2003) Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology and Biogeography*, **12**, 53–64.

Diniz-Filho, J.A.F., Bini, L.M., Rodríguez, M.Á., Rangel, T.F.L.V. & Hawkins, B.A. (2007) Seeing the forest for the trees: partitioning ecological and phylogenetic components of Bergmann's rule in European Carnivora. *Ecography*, **30**, 598–608.

Dolédec, S., Chessel, D., ter Braak, C.J.F. & Champely, S. (1996) Matching species traits to environmental variables: a new three-table ordination method. *Environmental and Ecological Statistics*, **3**, 143–166.

Dormann, C.F. (2007) Effects of incorporating spatial autocorrelation into the analysis of species distribution data. *Global Change Biology*, **16**, 129–138.

Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A.H., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.

Dray, S. & Legendre, P. (2008) Testing the species traits–environment relationships: the fourth-corner problem revisited. *Ecology*, **89**, 3400–3412.

Dray, S., Legendre, P. & Peres-Neto, P.R. (2006) Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, **196**, 483–493.

Durka, W. (2002) Phylogenie der Farn- und Blütenpflanzen Deutschlands. *BIOLFLOR – eine Datenbank zu biologisch-ökologischen Merkmalen der Gefäßpflanzen in Deutschland* (ed. by S. Klotz, I. Kühn and W. Durka), Schriftenreihe für Vegetationskunde, Vol. 38, pp. 75–91. Bundesamt für Naturschutz, Bonn, Germany.

Fitter, A.H. & Fitter, R.S.R. (2002) Rapid changes in flowering time in British plants. *Science*, **296**, 1689–1691.

Freckleton, R.P. & Jetz, W. (2009) Space versus phylogeny: disentangling phylogenetic and spatial signals in comparative data. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 21–30.

Gonseth, Y., Wohlgemuth, T., Sansonnens, B. & Buttler, A. (2001) *Die biogeographischen Regionen der Schweiz. Erläuterungen und Einteilungsstandard/Les régions biogéographiques de la Suisse. Explications et division standard*, Vol. 137. Bundesamt für Umwelt, Wald und Landschaft BUWAL, Bern, Switzerland.

Graham, M.H. (2003) Confronting multicollinearity in ecological multiple regression. *Ecology*, **84**, 2809–2815.

Griffith, D.A. & Peres-Neto, P.R. (2006) Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. *Ecology*, **87**, 2603–2613.

Harvey, P.H. & Pagel, M.D. (1991) *The comparative method in evolutionary biology*. Oxford University Press, Oxford.

Hawkins, B.A., Diniz-Filho, J.A.F., Bini, L.M., De Marco, P. & Blackburn, T.M. (2007) Red herrings revisited: spatial autocorrelation and parameter estimation in geographical ecology. *Ecography*, **30**, 375–384.

Johnson, S.D. (1993) Climatic and phylogenetic determinants of flowering seasonality in the Cape flora. *Journal of Ecology*, **81**, 567–572.

Kissling, W.D. & Carl, G. (2008) Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*, **17**, 59–71.

Knapp, S., Kühn, I., Schweiger, O. & Klotz, S. (2008) Challenging urban species diversity: contrasting phylogenetic patterns across plant functional groups in Germany. *Ecology Letters*, **11**, 1054–1064.

Koellner, T., Hersperger, A.M. & Wohlgemuth, T. (2004) Rarefaction method for assessing plant species diversity on a regional scale. *Ecography*, **27**, 532–544.

Kühn, I. (2007) Incorporating spatial autocorrelation may invert observed patterns. *Diversity and Distributions*, **13**, 66–69.

Kühn, I., Bierman, S.M., Durka, W. & Klotz, S. (2006) Relating geographical variation in pollination types to environmental and spatial factors using novel statistical methods. *New Phytologist*, **172**, 127–139.

Kühn, I., Böhning-Gaese, K., Cramer, W. & Klotz, S. (2008) Macroecology meets global change research. *Global Ecology and Biogeography*, **17**, 3–4.

Landolt, E., Bäumler, B., Erhardt, A., Hegg, O., Klötzli, F., Lämmler, W., Nobis, M., Rudmann-Maurer, K., Schweingruber, F.H., Theurillat, J.-P., Urmi, E., Vust, M. & Wohlgemuth, T. (2009) *Flora indicativa – Ecological indicator values and biological attributes of the flora of Switzerland and the Alps*. Haupt, Bern, Switzerland.

Legendre, P. (1993) Spatial autocorrelation – trouble or new paradigm. *Ecology*, **74**, 1659–1673.

Legendre, P. & Legendre, L. (1998) *Numerical ecology*. Developments in environmental modelling, Vol. 20. Elsevier, Amsterdam.

Legendre, P., Galzin, R. & Harmelin-Vivien, M.L. (1997) Relating behaviour to habitat: solutions to the fourth-corner problem. *Ecology*, **78**, 547–562.

Lennon, J.J. (2000) Red-shifts and red herrings in geographical ecology. *Ecography*, **23**, 101–113.

Levin, D.A. (2006) Flowering phenology in relation to adaptive radiation. *Systematic Botany*, **31**, 239–246.

McGill, B.J., Enquist, B.J., Weiher, E. & Westoby, M. (2006) Rebuilding community ecology from functional traits. *Trends in Ecology and Evolution*, **21**, 178–185.

Mac Nally, R. & Walsh, C.J. (2004) Hierarchical partitioning public-domain software. *Biodiversity and Conservation*, **13**, 659–660.

Menzel, A., Estrella, N. & Fabian, P. (2001) Spatial and temporal variability of the phenological seasons in Germany from 1951 to 1996. *Global Change Biology*, **7**, 657–666.

Menzel, A., Sparks, T.H., Estrella, N. & Roy, D.B. (2006) Altered geographic and temporal variability in phenology in response to climate change. *Global Ecology and Biogeography*, **15**, 498–504.

Peres-Neto, P.R. (2006) A unified strategy for estimating and controlling spatial, temporal and phylogenetic autocorrelation in ecological models. *Oecologia Brasiliensis*, **10**, 105–119.

Plattner, M., Birrer, S. & Weber, D. (2004) Data quality in monitoring plant species richness in Switzerland. *Community Ecology*, **5**, 135–143.

Prinzing, A., Durka, W., Klotz, S. & Brandl, R. (2001) The niche of higher plants: evidence for phylogenetic conservatism. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 2383–2389.

Purvis, A. & Rambaut, A. (1995) Comparative analysis by independent contrasts (CAIC): an Apple Macintosh application for analysing comparative data. *CABIOS*, **11**, 247–251.

R Development Core Team (2008) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Roetzer, T., Wittenzeller, M., Haeckel, H. & Nekovar, J. (2000) Phenology in central Europe – differences and trends of spring phenophases in urban and rural areas. *International Journal of Biometeorology*, **44**, 60–66.

Schweiger, O., Klotz, S., Durka, W. & Kühn, I. (2008) A test of phylogenetic diversity indices. *Oecologia*, **157**, 485–495.

Tautenhahn, S., Heilmeier, H., Götzenberger, L., Klotz, S., Wirth, C. & Kühn, I. (2008) On the biogeography of seed mass in Germany – distribution patterns and environmental correlates. *Ecography*, **31**, 457–468.

Tiefelsdorf, M. & Griffith, D.A. (2007) Semiparametric filtering of spatial autocorrelation: the eigenvector approach. *Environment and Planning A*, **39**, 1193–1221.

Weber, D., Hintermann, U. & Zangger, A. (2004) Scale and trends in species richness: considerations for monitoring biological diversity for political purposes. *Global Ecology and Biogeography*, **13**, 97–104.

Westoby, M., Leishman, M.R. & Lord, J.M. (1995) On misinterpreting the 'phylogenetic correction'. *Journal of Ecology*, **83**, 531–534.

Wohlgemuth, T., Nobis, M.P., Kienast, F. & Plattner, M. (2008) Modelling vascular plant diversity at the landscape scale using systematic samples. *Journal of Biogeography*, **35**, 1226–1240.

Zimmermann, N.E. & Kienast, F. (1999) Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *Journal of Vegetation Science*, **10**, 469–482.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1 Representation of a binary phylogenetic tree in a species by a phylogenetic branch matrix, using phylogenies in the comparative analysis by independent contrasts (CAIC) format.

Figure S2 Moran's *I* correlograms of residual autocorrelation of the different models explaining average onset of flowering in Switzerland.

Figure S3 Moran's *I* correlograms of residual autocorrelation of the different models explaining average onset of flowering in Switzerland excluding calcareous substrate.

Table S1 Correlation matrix of eigenvectors selected by spatial and spatially structured phylogenetic (spatio-phylo) filtering.

Table S2 Correlation matrix of eigenvectors for combining spatial and spatially structured phylogenetic (spatio-phylo) filters in one model (both).

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCHES

Ingolf Kühn is a senior scientist at the Helmholtz Centre for Environmental Research – UFZ and is interested in spatial analyses of plant species and their traits especially in the context of climate change or biological invasions.

Michael P. Nobis heads the spatial ecology group at the Swiss Federal Research Institute WSL. His research concentrates on spatial analyses of species distributions and biodiversity patterns.

Walter Durka is a senior scientist at the Helmholtz Centre for Environmental Research – UFZ. He works on plant population genetics and is interested in traits affecting plant reproduction and in plant phylogeny and their effect on population structure under habitat and global change.

Editor: José Alexandre F. Diniz-Filho