Journal of Vegetation Science **23** (2012) 805–821

Assessing species and community functional responses to environmental gradients: which multivariate methods?

Michael Kleyer, Stéphane Dray, Francesco de Bello, Jan Lepš, Robin J. Pakeman, Barbara Strauss, Wilfried Thuiller & Sandra Lavorel

Keywords
Cluster analysis; Functional groups; Functional traits; Functional types; Inference; Multivariate analysis; Plants; Regression tree; Statistical modelling

Abbreviations

CWM = average trait expression across all species of a community, weighted by their abundance; RDA = redundancy analysis; CCA = canonical correspondence analysis; RLQ = a double inertia analysis of two arrays (R and Q) with a link expressed by a contingency table (L); mRegTree = multivariate regression tree analysis; sRegTree = univariate regression tree analysis; OMI = outlying mean index; GAM = general additive model; Cluster regression = a combination of cluster analyses and logistic regression

Nomenclature

Jäger & Werner (2002)

Received 7 February 2011
Accepted 7 February 2012
Co-ordinating Editor: Otto Wildi

Kleyer, M. (corresponding auhor, michael.kleyer@uni-oldenburg.de) & **Strauss, B.** (barbara.strauss@uni-oldenburg.de): Landscape Ecology Group, University of Oldenburg, Oldenburg, 26111, Germany
Dray, S.: Laboratoire de Biométrie et Biologie Evolutive, Université Lyon 1, Lyon, F-69000, France
Dray, S. (stephane.dray@univ-lyon1.fr): CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, Villeurbanne, F-69622, France
Bello, F. (fradebello@ctfc.es), **Thuiller, W.** (wilfried.thuiller@ujf-grenoble.fr) & **Lavorel, S.** (sandra.lavorel@ujf-grenoble.fr): Laboratoire d'Ecologie Alpine (CNRS UMR 5553) and Station Alpine Joseph Fourier (UMS-UJF-CNRS 2925), Université Joseph Fourier, BP 53, Grenoble Cedex 09, F-38042, France

Abstract

Question: How do multivariate methods perform in relating species- and community-level trait responses to the environment?

Location: (1) Field data from grazed semi-natural grasslands, NE Germany; (2) artificial data.

Methods: Research questions associated with trait–environment relationships were briefly reviewed and seven available methods evaluated. The main distinction between research questions is whether trait–environment relationships should be addressed at community or species level. A redundancy analysis (RDA) of mean trait values of species in a plot weighted by their abundances (CWM-RDA) is exclusively suitable for the community level. The other six methods address the species level. A double inertia analysis of two arrays (RLQ) and double canonical correspondence analysis (double CCA) use combinations of ordinations to simultaneously analyse species and trait responses to the environment. A combination of the outlying mean index with generalized additive models (OMI-GAM) predicts the response of species to environmental variables on trait gradients. RDA-RegTree first analyses species responses to the environment with RDA and then uses a regression tree to classify trait expressions according to scores of species responses on the ordination axes. Cluster regression uses cluster analyses and logistic regression to search for trait combinations with the best response to the environmental variables. This method models the distribution of functional groups on environmental gradients. All methods and data are available as R scripts.

Results: All methods consistently revealed the main trait responses to environment in the field data set, namely that life history was associated with available phosphorus while grazing intensity was related to leaf C:N ratio and canopy height. At community level, CWM-RDA gave a good overview of trait–environment relationships, as also provided by the species-based methods RLQ and double CCA. OMI-GAM revealed non-linear relationships in the field data set. Field and artificial data gave that the number and stability of functional groups produced by Cluster regression and RDA-RegTree varied more strongly than RLQ, double CCA and OMI-GAM.

Conclusions: Each method addresses particular ecological concepts and research questions. If a user asks for the response of average trait expressions of communities to environmental gradients, CWM-RDA may be the first choice. However, species-based methods should be applied to address questions regarding co-existence of different life histories or to assess how groups of species respond to environmental changes. The artificial data set revealed that the methods differed in sensitivity to gradient lengths and random data.

Doi: 10.1111/j.1654-1103.2012.01402.x

Lepš, J. (suspa@prf.jcu.cz): Faculty of Biological Sciences, University of South Bohemia, Branisovska 31, Ceske Budejovice, Czech Republic

Lepš, J.: Biology Centre of ASCR, Institute of Entomology, Branišovská 31, Ceské Budejovice, CZ-37005, Czech Republic

Pakeman, R.J. (robin.pakeman@hutton.ac.uk): The James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK

Introduction

Over the last decade, the use of functional traits of organisms, rather than taxonomic specification, as a method to produce generic rules on community dynamics in response to environmental change, has gained increasing momentum (McGill et al. 2006; Suding et al. 2008). Trait–environment relationships are of interest for several important purposes in ecology and environmental management. This is particularly evident in vegetation studies. First, for example, the overwhelming majority of dynamic vegetation models use (explicitly or not) plant functional types, i.e. groups of plants with similar suites of co-occurring functional traits. However, there is a consensus that future models will need to use more sophisticated, and less subjective, functional classification approaches (Lavorel et al. 2007). Second, the analysis of responses of traits associated with persistence, regeneration and dispersal to environmental gradients should assist in a more mechanistic understanding of community ecology and species niches (McGill et al. 2006). Third, functional groups can be used to aggregate the overwhelmingly large diversity of organisms into groups of species that share similar trait expressions. This approach has been prevalent across many disciplines (e.g. guilds, Fox & Brown 1993; Wilson 1999; biogeochemical functional groups, Hood et al. 2006). Fourth, environmental management can strongly benefit from a functional approach that identifies key plant or animal functional groups with specific environmental sensitivities or ecosystem effects (Bonada et al. 2007).

With increasing availability of trait data, the methodological challenge of linking species and community trait responses to the environment has gained prominence (Legendre et al. 1997), and in spite of many years of debate, it remains an open question (Pakeman 2004; Dray & Legendre 2008). The methods should directly link candidate traits to environmental data via the performance of the species bearing expressions of these traits, thereby separating responsive from non-responsive traits. Ultimately, techniques for functional group identification should group organisms with similar responses to environmental gradients, and with similar expressions of multiple traits underlying such response (response groups characterized by their 'response traits' *sensu* Lavorel & Garnier 2002).

How does a user, such as a field ecologist with basic expertise in the intricacies of multivariate methods, choose which method to apply for a given objective and data set? To our knowledge, no study has presented different methods to relate species- and community-level trait response to the environment and evaluated their advantages and drawbacks. To guide users in choosing the most appropriate methods, we briefly review the range of research questions that require the identification of response traits or functional response groups. Second, using field and artificial data, we present seven different statistical procedures to address some of these research questions and examine their advantages and constraints. We compare species- and community-based functional classifications for plants, but the methods are also applicable to other organisms (e.g. Dolédec et al. 2007; Moretti et al. 2009). R scripts and functions are provided in the Appendix that contains a full analysis of the field data set organized as a tutorial that can be readily applied to any other similar data set.

Which ecological questions and which methodological issues do they imply?

In most situations, considering species traits in an ecological study implies the need to analyse three tables: a table containing the abundances (or presence/absence) of species in plots, a second table with the measurements of environmental variables (e.g. climate, disturbance, land use) for the plots, and a third table describing traits for the species (e.g. plant height, animal body mass). Several methods have been developed to study the relationships between traits and environmental variables using the information contained in these three tables. We identified a set of key generic questions that can be addressed with these different methods, and classified them according to the level of analysis and related statistical properties (Table 1).

The main distinction between the questions in Table 1 is whether trait–environment relationships should be addressed at the community level (question 1) or the species level (question 2; Ackerly et al. 2002). In the community-level approach the community is considered as the observation unit. In this case, a plot by trait matrix is constructed by combining the abundance and trait tables and a value for each trait is computed for each plot. Numerical traits can be averaged over all species present in the plot, weighted (or not) with their abundance or frequency. Categorical traits can be expressed as proportions or frequencies.

Doi: 10.1111/j.1654-1103.2012.01402.x

Table 1. Research questions and criteria for choosing an appropriate method.

Ecological questions	Unit of analysis	Dependent variable (yr)	Phylogeny	Criteria			Methods
				Within species trait variability	Species frequency vs occurrence	Single vs multiple traits	
1 How do average trait expressions of communities respond to environmental gradients?							
1a How do average trait expressions of communities change along environmental gradients?	Average species or individuals	Community trait composition	Not relevant	Preferable	Both possible	Both possible	CWM-RDA, RLQ, double CCA
1b Do average trait expressions of *a priori* groups of species (e.g. dominants, invasives) respond differently than the rest of the community?	Average species and individuals	Community trait composition	Not relevant	Possible	Abundance needed for dominance	Both possible	CWM-RDA, RLQ, double CCA
1c To what extent is community functional response driven by species replacement or phenotypic variation?	Average individuals	Community trait composition	Not relevant	Inter-treatment variability	Both possible	Both possible	
1d How do patterns of trait correlation at the community levels change with environmental conditions?	Average species or individuals	Community trait composition	Not relevant	Preferable	Both possible	Pairs of traits	Violle et al. (2007)
2 How do trait expressions of species respond to environmental gradients?							
2a Which traits predict species response to environmental gradients?	Species	Species position	Possible	Can be used as predictor	Both possible	Both possible	Cluster regression, RDA-sRegTree, RDA-mRegTree, OMI-GAM, RLQ, double CCA
2b Which traits predict species niche breadth?	Species	Species range	Possible	Average trait values per species	Both possible	Both possible	OMI-GAM, RLQ, double CCA
2c How do groups of species with similar trait expressions respond to environmental gradients?	Species		Possible	Possible (ecotypes)	Both possible	Both possible	Cluster regression, RDA-sRegTree, RDA-mRegTree, OMI-GAM, RLQ, double CCA
2d Is the response to the environment of *a priori* groups of species (e.g. dominant vs non-dominant or invasive vs non-invasive) related to the same traits?	Species	Species properties (invasiveness, dominance)	Not applicable	Possible	Both possible	Both possible	Cluster regression, RDA-sRegTree, RDA-mRegTree, OMI-GAM, RLQ, double CCA

CCA, canonical correspondence analysis; RDA, redundancy analysis; RLQ, a double inertia analysis of two arrays (R and Q) with a link expressed by a contingency table (L).

The trait table is then analysed against the environment table by, for example, constrained ordination methods.

The ecological concept behind the community-based approach is that the environment exerts strong selective pressure on a trait and sieves only a small subset of the total range of trait expressions in the geographical species pool (‘habitat filtering’; Woodward & Diament 1991; Grime 2006). Thus, the distribution of numerical trait values in a community should converge around a mean describing this subset (Grime 2006).

Overall, analyses at the community level address questions on which trait expressions will become more likely in a community under a given environmental change, e.g. whether vegetation stature will be shorter under intensive grazing or mowing (Nygaard & Ejrnaes 2004; Table 1, question 1a). Communities can be split into different ecological groups (e.g. dominants vs subordinates, Cingolani et al. 2007; invaders vs resident species, Thuiller et al. 2006; frequent vs rare species, de Bello et al. 2007) and the average trait composition of these groups can be compared with respect to their position on environmental gradients (question 1b). An open research question concerns the extent to which trait variation along a gradient is mostly due to intra-species trait variability (phenotypic plasticity and/or genetic variability; Crutsinger et al. 2007; Whitlock et al. 2007; question 1c).

In the species-level approach the statistical units are the species (except when phylogenetic independent contrasts are applied, see below). The first set of species-level questions relates to the traits that determine individualistic species distributions (Table 1, question 2a, 2b). Much effort has been devoted to assess which combinations of traits determine species adaptations to different environmental conditions. For example, which traits are needed for plant species to adapt to cold alpine conditions? The answer is that plants in these conditions need a low stature, often with a cushion shape, a perennial life cycle and a set of ecophysiological adaptations to cope with high light intensity (Körner 2003).

Although the community-level approach assumes that the mean trait expression of a community is predictable, functionally different species and groups of species can co-occur in the same community. Such a pattern points to different survival strategies or different ways of resource exploitation in a community. A trivial example is the co-occurrence of trees, shrubs and herbs in a forest. Question 2c asks whether these functionally different groups show a different or similar response to the environment. Likewise, differences in ecological specialization may or may not depend on different responsive traits (e.g. generalist vs specialist strategies; question 2d). The species-level approach can also be used to assess several more specific questions, which will not be dealt with in our review of methods. These include the question, to what extent the adaptations to a given habitat are phylogenetically-driven or independent of phylogeny (Westoby et al. 1995).

Methods

Field and artificial data

We used two complementary data sets; first original field data as would have been collected by a typical user of the methods analysed here, and second artificial data in order to support the robustness of our analysis. Data collected in the field introduce real-life variation by plot selection, fuzzy ecological patterns or legacies of past histories in trait–environment relationships. The performance of each method can thus be assessed under realistic conditions. Field data were collected in a marginal grassland, grazed year-round by free-ranging cattle, horses and sheep in northeast Germany. The data set comprised plant species abundances, grazing intensity, soil water-holding capacity, extractable phosphorus and the following traits: canopy height, specific leaf area (SLA), seed mass (log transformed), leaf C:N ratio, onset of flowering date, flowering mode (polycarpic/monocarpic). See Appendix S1 and Garnier et al. (2007) for a detailed description of the study site and the sampling methods.

In artificial data, criteria for assessing methods can explicitly be incorporated. Our criteria were that (1) each method separated traits responding to environmental conditions from non-responsive traits, (2) classification of functional groups was stable, (3) strong environmental gradients produced stronger responses than absent gradients, and (4) random species distributions produced no responses. We generated an artificial data set consisting of two orthogonal environmental gradients, named disturbance and soil resources. Niches for 50 species were constructed on the two gradients with Gaussian response curve formulas, including a random component. These niches were built according to three scenarios: (1) at the corners and the centre of the disturbance–resource template, producing two strong orthogonal gradients, (2) at the intermediate position between centre and corner as well as centre, i.e. two intermediate gradients, (3) at the centre only, i.e. no gradient (Fig. S3.1, Appendix S3). A fourth scenario consisted of random distributions of species occurrences and abundances. Finally, trait values were generated for canopy height, SLA, leaf N, seed mass, leaf N content, onset of flowering date, based on linear and non-linear functions simulating trait–environment relationships (e.g. disturbance decreased canopy height) or trait–trait correlations, including a random component. To separate responsive from non-responsive traits, a trait with a completely random distribution was introduced [terminal velocity (TV)]. The artificial data set was organized in

Doi: 10.1111/j.1654-1103.2012.01402.x

three tables, comprising species abundances, environmental data and traits, as described above. We ran 200 repetitions for each scenario, applied all methods to each repetition and compared the results. Appendix S3 gives a more detailed description of the artificial data set and associated results, including tables and figures. Here, we present the results of the field data with illustrations, and summarize the results of the artificial data set.

Statistical methods

All methods combine the trait table with the environment table via the abundance table. To achieve this task, the methods combine ordination, regression or classification (cluster analysis, regression tree) in different order. We labelled the methods so that acronyms indicate the main components. Here we briefly present general features of the seven methods (see Table 2). They are described in detail in a tutorial (Appendix S2), comprising an introduction to each method, R code, results and graphical outputs of the field data set. All methods are available as Rscripts (R Foundation for Statistical Computing, Vienna, Austria) in Appendix S4, which also provides the field data set.

Several other methods are available which we will not address in this paper. A well-known method is the fourth-corner method of Legendre et al. (Legendre et al. 1997, see also Dray & Legendre 2008; Lehsten et al. 2009). Other authors have developed comparable methods that use similar components as those mentioned above (see e.g. Pillar & Sosinski 2003; Pakeman 2004). Pillar et al. (2009) have recently proposed a method using partial Mantel correlations and a null model to evaluate correlations between trait-based described communities and ecological gradients and to sort out trait convergence and trait divergence.

CWM-RDA: a community-based approach (Method A)

Method A (CWM-RDA) is a community-based analysis to assess the response of functional traits to environmental variables (Nygaard & Ejrnaes 2004), which can be applied to type 1 questions (Table 1). From the trait and abundance tables, a plot by trait matrix was built by averaging the trait expressions of all species weighted by their abundance in each plot (CWMs = community weighted means of traits; Garnier et al. 2007). Then, a redundancy analysis (RDA; Rao 1964) of the plot by trait matrix constrained by the environment table was performed.

The species-based approaches (Methods B–G)

The six other approaches all involved three components (Table 2): (1) the responses of the species to the environmental variables, (2) identification of responsive traits, (3) grouping of the species based on similarity in the trait table. Methods B–F investigate species responses with multivariate ordination methods constrained by environmental variables. Species responses were expressed as species scores on the ordination axes. Method G started by constructing clusters of species with similar trait expressions; then the distributions of these groups were linked to the environmental gradients (McIntyre & Lavorel 2001; Kühner & Kleyer 2008). All methods took into account species abundances.

RDA-RegTree: combination of RDA and regression trees (Methods B and C)

To determine species responses to environmental gradients (component 1), RDA-RegTree used RDA, because the gradients were short. Traits that predict the RDA scores were identified with a regression tree analysis (RegTree; Breiman et al. 1984; component 2). In the case of RDA-sRegTree (Method B), the functional groups (component 3) resulted from two single, univariate regression trees (sRegTree), where the response variables were the species scores on the first axes of two separate RDA analyses of the environment and abundance tables, the first with soil phosphorus as the main environmental gradient and disturbance as co-variable and the second vice versa. The traits were used as explanatory variables, which provided a separate classification for each environmental gradient. In RDA-mRegTree (Method C), the functional groups resulted from a multivariate regression tree (mRegTree; De'Ath 2002,) which used the axes 1 and 2 of a RDA. This time, the RDA included both phosphorus and disturbance as environmental gradients and the traits were used as explanatory variables again. In the following, RDA-RegTree refers to methods B and C pooled.

OMI-GAM: outlying mean index followed by generalized additive modelling (Method D)

OMI-GAM first determined species responses to environmental conditions, using the outlying mean index ordination procedure (OMI; Dolédec et al. 2000; Dray et al. 2003), and then explained these responses using generalized additive models (GAM) with the traits as explanatory variables. OMI determines species niche positions and niche breadths by measuring the distance between the mean environmental conditions used by each species and the mean environmental conditions of the study area (Dolédec et al. 2000; see Appendix S2). It makes no hypothesis on the length of the gradients and gives equal weight to species-poor and species-rich sites, unlike canonical correspondence analysis (CCA; ter Braak 1986).

Table 2. The statistical procedures and their components. In component 3, two ways of grouping are possible, depending on questions 2c or 2d in Table 1.

Analysis family	Name	Component 1 species response to environment (R × L)	Component 2 identification of responsive trait combinations	Component 3 grouping of species based on responsive traits	Steps	Trait–environment model	Functional groups	Multi-gradient analysis
Averaging – ordination	A: CWM – RDA	Ordination of community weighted averages of traits			2	Linear	No	Yes
Ordination – regression	B: RDA – sRegTree	Ordination of species constrained by single environmental gradients (RDA of L × R)	Regression tree with species RDA scores as dependent and traits as explanatory variables		1, 2, 3	Any	Yes	No
	C: RDA mRegTree	Ordination of species constrained by multiple environmental gradients (RDA of L × R)	Regression tree with species RDA scores as dependent and traits as explanatory variables		1, 2, 3	Any	Yes	Yes
	D: OMI – GAM	Ordination of species constrained by environmental gradients (OMI of L × R)	GAM regression using information theory framework with species scores on OMI axes as dependent and traits as explanatory variables	2c: cluster analysis 2d: classification *a priori* of species based on specific properties	1, 2, 3	Any	Yes	Yes
Double ordination	E: RLQ	Generalization of co-inertia analysis to study the link between species traits Q and environmental gradients L through the species table L		2c: cluster analysis 2d: classification *a priori* of species based on specific properties	1–2, 3	Unimodal	Yes	Yes
	F: double CCA	Ordination of species table L constrained by both environmental gradients R and species traits Q		2c: cluster analysis 2d: classification *a priori* of species based on specific properties	1–2, 3	Unimodal	Yes	Yes
Clustering-regression	G: Cluster regression	Logistic regression of all trait combinations with stable clusters. Cumulated frequency of species in clusters as dependent variables, env. gradients as explanatory variables.	Trait combination with stable clusters and best regression model in terms of R^2.	2c: cluster analyses of Q. Selection of stable clusters via bootstrapping. 2d: classification *a priori* of species based on specific properties (invasiveness, dominance)	3, 1, 2	Linear, sigmoidal or unimodal	Yes	Yes

L, abundance table; R, environment table; Q, trait table; RDA, redundancy analysis; RLQ, a double inertia analysis of two arrays (R and Q) with a link expressed by a contingency table (L).

Doi: 10.1111/j.1654-1103.2012.01402.x

Species scores on OMI axes 1 and 2 were used as response variables in the GAMs in an information theory framework (model averaging; Burnham & Anderson 2002). To group species according to their traits into functional groups, OMI-GAM applied Euclidean distance and Ward's hierarchical clustering (Everitt et al. 2011) to produce clusters with similar numbers of species (see Appendix S2). Other forms of clustering may also be appropriate. The optimal number of clusters was determined via the Calinski and Harabasz index (Gordon 1999).

RLQ and double CCA: double ordination methods (Methods E and F)

A double inertia analysis of two arrays (R and Q) with a link expressed by a contingency table (L) (RQ) and double CCA analyse the three tables simultaneously and thus consider components 1 and 2 in a single step (Table 2). RLQ (Method E; Dolédec et al. 1996) is an extension of co-inertia analysis that searches for a combination of traits and environmental variables of maximal co-variance, which is weighted by the abundances of species in plots. Double CCA (Method F; Lavorel et al. 1999) performs an ordination of the abundance table that is constrained by both traits and environmental descriptors. In both analyses (E and F), species responses to environmental gradients are based on correspondence analysis (CA) of the abundance table, which treats the species and plots in a symmetric manner. To determine functional groups, both methods used the same procedure as OMI-GAM.

Cluster regression: combination of iterative cluster analyses and logistic regression modelling (Method G)

In contrast to the previous methods, Cluster regression started by clustering the traits table to obtain species groups with similar trait values. For each group, the number of species occurring in each plot was counted, weighted by species abundances. These counts were then used as a dependent variable to calculate the response of the clusters to environmental variables with logistic regression and model averaging (Burnham & Anderson 2002).

To separate responsive from non-responsive trait combinations, clusterings were conducted based both on single traits and based on all possible combinations of single traits (63 clusterings in total). To reduce this data set, only trait combinations with stable clusters were considered for further analyses. Cluster stability was assessed by bootstrapping the clusters and the co-phenetic correlation (Legendre & Legendre 1998; see Appendix S2). The response of each cluster of each trait combination to the environment was determined by the procedure described above. The most responsive trait combination was identified as the one with the best response to the environmental variables, expressed as average goodness of fit of the models of all clusters of a certain trait combination (Kühner & Kleyer 2008).

The order of the components

While the components are quite similar in the different methods, their order is different. The order also determines the questions that can be addressed with these methods (Table 1). RDA-RegTree and OMI-GAM start with component 1, the response of species to environmental variables, then identify the predictive traits (component 2) and finally group the species according their similarity in traits (component 3, Fig. 1). This order reflects the assumption that the environmental conditions act as a filter on species co-existence and that traits are an expression of the adaptive selection made by this filter. It also applies to CWM-RDA (Method A).

In contrast, Cluster regression starts to build species groups from their traits (component 3), then finds the most responsive trait combination in components 1 and 2. This method assumes that the landscape species pool can be classified based on correlations between traits that indicate underlying trade-offs or allometries, and that these emergent groups (*sensu* Lavorel et al. 1997) can have consistent responses to the environment (Hérault & Honnay 2005; Kühner & Kleyer 2008).

Thus, in RDA-RegTree and OMI-GAM, the functional groups are built *post hoc*, i.e. after exploring the relationships between traits and environment, whereas Cluster regression begins with groupings from all possible trait combinations and separates responsive trait combinations from neutral combinations in subsequent steps. As a consequence, Cluster regression produces only response groups, but no single response traits. RLQ and double CCA perform components 1 and 2 in a single step and therefore provide a more general co-variation pattern between traits and environment without any *a priori* assumption regarding dependent and explanatory variables. The identification of functional groups is a major result of methods B, C and G, but not necessarily the aim of methods D–F. For the latter, it is an additional step that can be performed *post hoc* and was carried out here to help assess consistency between methods.

Results

Because it was not feasible to address all the questions listed in Table 1 to illustrate the use of the different methods presented here, we decided to focus on three

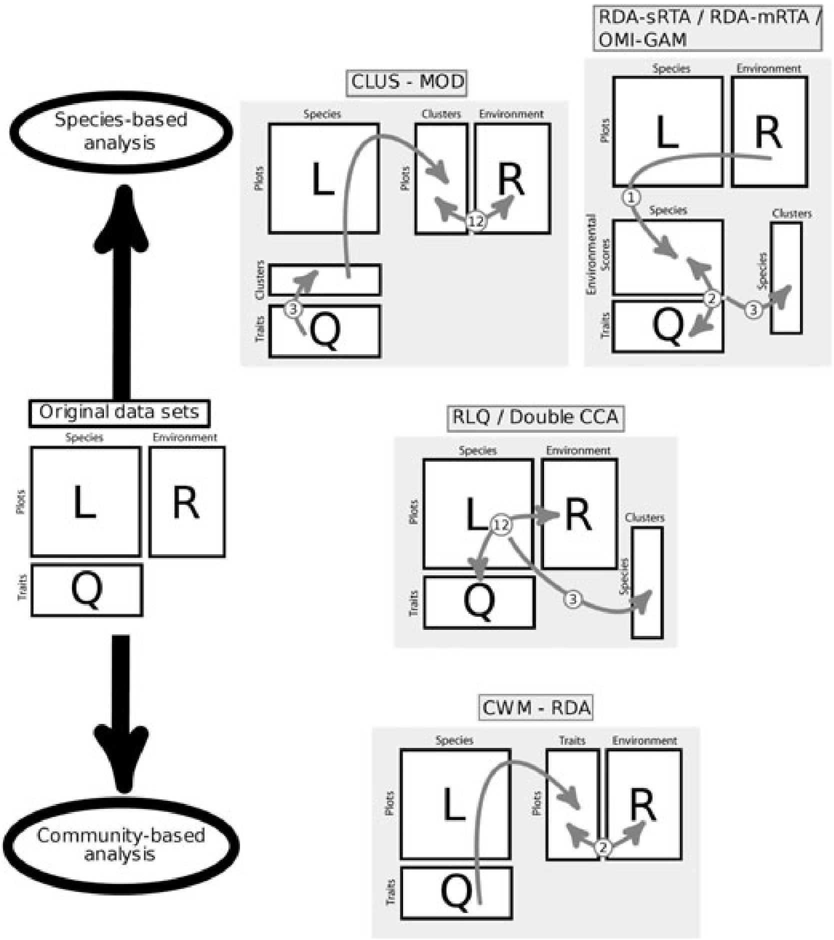

Fig. 1. The order of components 1, 2 and 3 (in circles) in the different methods when dealing with the abundance table (L), the environment table (R) and the traits table (Q) (see Table 2 for further explanations). Arrows denote the sequence of the components.

main questions (1a, 2a and 2c in Table 1). These are, to the best of our knowledge, the most frequently investigated in the ecological literature. Question 1a focuses on determining how the aggregated mean traits of selected communities respond to environmental gradients. Question 2a focuses on species instead of communities. The idea is to investigate which candidate traits are mostly accounting for by differences in species distribution along environmental gradients, the form and direction of the relationship being of interest. Question 2c has been the focus of 20 yrs of investigation in functional ecology, i.e. how groups of species with similar trait expressions respond to environmental gradients, a necessary understanding for any modelling at a global scale (Woodward & Cramer 1996).

Species responses to the environmental gradients of the semi-natural grassland are presented in Appendix S1 as a baseline against which to assess the trait–environment relationships. Here we examine the applicability of methods A–G (Table 2) to determine these relationships with regard to questions 1a, 2a and 2c from Table 1 (see Appendix S2 for more detailed results of the individual methods).

Question 1a: how do average trait expressions of communities change along environmental gradients?

Three methods addressed question 1a of Table 1, i.e. change of trait expressions of communities along environmental gradients. Although CWM-RDA uses trait averages per plot whereas RLQ and double CCA are species-based, the results of the three methods were similar. Only the rotation of the axes was different (Fig. 2). In the field data set, grazing intensity and phosphorus explained a high degree of variation in all three methods, whereas soil water-holding capacity was relatively unimportant (Table 3). SLA increased significantly with increasing disturbance intensity, whereas

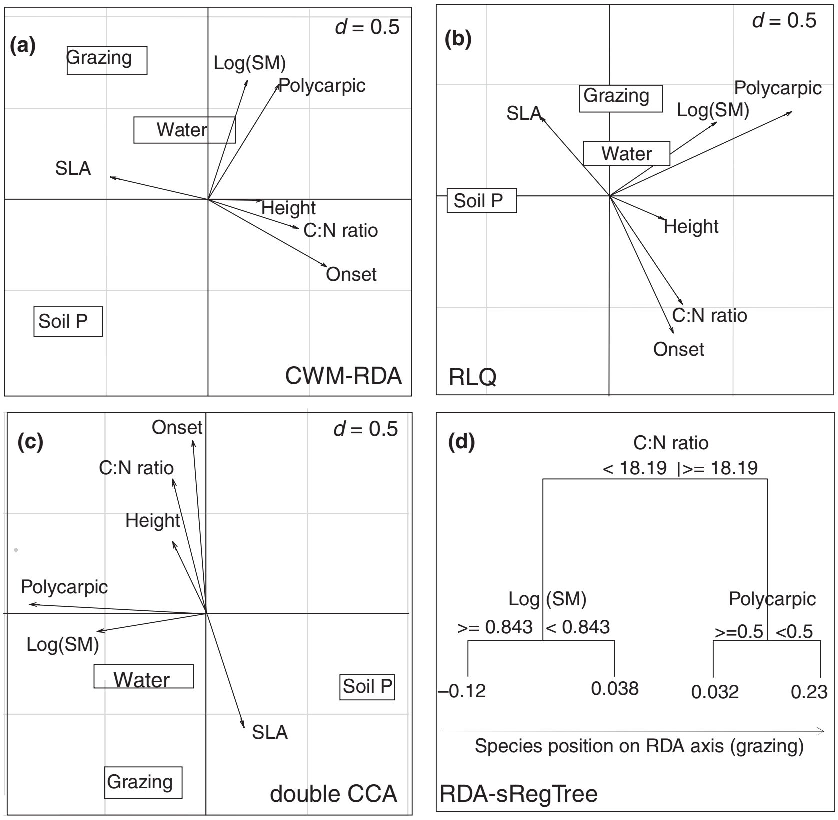

Fig. 2. Projections of traits and environmental variables in (**a**) redundancy analysis (RDA), (**b**) RLQ, (**c**) double canonical correspondence analysis (CCA), (**d**) RDA-RegTree. The percentage variance explained by each axis and the correlations of the environmental variables in each of the analyses are given in Table 3. Abbrviations: polycarpic, species monocarpic (0) or polycarpic (1); log (SM), log seed mass; C:N ratio, C:N ratio in leaves; height, canopy height; SLA, specific leaf area; onset, onset of flowering; grazing, grazing intensity; soil P, available phosphorus; water, soil water content.

onset of flowering and C:N ratio decreased. Phosphorus was positively related to SLA and onset of flowering and negatively related to C:N ratio, log seed mass and polycarpic life history.

Question 2a: which traits predict species response to environmental gradients?

In general, all species-based methods provided similar results regarding the traits that predict species responses to environmental gradients. OMI-GAM showed this directly by the weight of evidence of the traits in explaining species responses (Fig. 3a), whereas RLQ and double CCA used the correlation of traits to environmental gradients. In Cluster regression and RDA-RegTree, the most relevant traits were those selected for grouping. In the field data set, life history (polycarpic/monocarpic) was the most segregating trait in all methods (Table 3), followed by C:N ratio. Onset of flowering, SLA, seed mass and canopy height differed in importance among the methods. In general, results produced by the ordination-based methods (RDA-RegTree, RLQ, double CCA) were very similar to the pattern produced by CWM-RDA (Fig. 2, Table 3). OMI-GAM revealed that the relationship between continuous traits and species responses to environmental gradients was sigmoidal or slightly unimodal (Fig. 3b).

In the artificial data set, the built-in trait–environment relationships were generally identified by all methods (see Appendix S3 for details). The explained variation decreased strongly from the 'strong' and 'intermediate' to the 'absent gradients' scenario. Regarding Cluster regression and RDA-mRegTree, the degree of explained variation varied across repetitions in the 'absent gradients' scenario. Finally, in the 'random' scenario all methods failed to explain any variation, as expected.

For the 'strong', 'intermediate' and 'absent gradients' scenarios, trait–environment relationships incorporated in the artificial data set were well reproduced by CWM-RDA and RLQ. Regarding OMI-GAM and double CCA, associations of the ordination axes to the two environmental gradients shifted strongly between the repetitions, resulting in high variation of trait responses to the two gradients. RDA-mRegTree gave more weight to the traits related to

Table 3. Trait responses to the environmental gradients and composition of functional groups. The degree of association (canonical correlation) of each of the environmental variables to the two main axes (with their percentage contribution to the explained variance) is presented for each method. In CWM-RDA, the three environmental variables explained 32.1% of the total variation and the values presented are percentages of this explained variability. In case of Cluster regression, the average weight of evidence across all groups is given. Functional groups are described by their trait combinations.

Method	Groups	Response to gradients		
		Grazing	*P*	Water
A. CWM – RDA	Axis 1 (63%)	0.27	0.36	NS
	Axis 2 (35%)	−0.27	0.20	NS
B. RDA-singleRegTree response to phosphorus	Axis constrained by P (5.5%)			
	1 poly (–)		+	
	2 poly (+)		–	
B. RDA-singleRegTree response to grazing	Axis constrained by Grazing (5.6%)			
	1 C:N (+) poly (–)	–		
	2 C:N (+) poly (+)	+/–		
	3 C:N (–) SM (–)	+		
	4 C:N (–) SM (+)	+/–		
C. RDA-multiRegTree	Axis 1 (72%)			
	Axis 2 (28%)			
	1 poly (–)	–	+	
	2 poly (+)	+	–	
D. OMI–GAM	Axis 1 (61%)	0.81	−0.41	0.70
	Axis 2 (28%)	0.55	0.79	0.35
	1 poly (–) C:N (–) height (–) SM (–) SLA (+) onset (–)	+/–	+	–
	2 poly (+) C:N (–) height (–) SM (+) SLA (+/–) onset (+/–)	+	–	+
	3 poly (+) C:N (+) height (+) SM (+/–) SLA (–) onset (+)	–	–	+/–
E. RLQ	Axis 1 (57%)	0.09	−0.99	0.17
	Axis 2 (42%)	0.96	−0.11	0.76
	1 poly (–) C:N (–) height (–) SM (+) SLA (+) onset (–)	+	+	
	2 poly (+) C:N (–) height (–) SM (+) SLA (+) onset (+)	+	–	
	3 poly (+) C:N (+) height (+) SM (+) SLA (–) onset (+)	+/–	–	
	4 poly (–) C:N (–) height (–) SM (–) SLA (–) onset (+)	–	+	
F. double CCA	Axis 1 (58%)	−0.42	0.94	−0.31
	Axis 2 (38%)	−0.90	−0.32	−0.36
	1 poly (–) C:N (–) height (–) SM (+) SLA (+) onset (–)	+	+	+
	2 poly (+) C:N (–) height (–) SM (+) SLA (+/–) onset (+)	+	–	+
	3 poly (+) C:N (+) height (+) SM (+) SLA (–) onset (+)	+/–	–	+
	4 poly (–) C:N (–) height (+/–) SM (–) SLA (–) onset (+)	–	+	–
G. Cluster regression	Average weight of evidence [%]	39	28	33
	1 poly (–) C:N (–) height (–);R^2 = 0.23		+	–
	2 poly (+) C:N (–) height (–);R^2 = 0.27	+		+/–
	3 poly (+) C:N (+) height (–);R^2 = 0.49	–	–	
	4 poly (+) C:N (–) height (+);R^2 = 0.32	+/–	+/–	+

Trait expressions are denoted as (+) indicating a high level, or as (–) indicating a low level. For example, 'poly (–) C:N (–) height (–)' indicates a group consisting of small monocarpic species with low C:N ratio in leaves. The response to the gradients is indicated with a (+) in case that a functional group increases on the gradient, with a (–) when the group decreases, and (+/–) in case of a unimodal response. Poly (–), monocarpic; poly (+), polycarpic; SM, seed mass; C:N, C:N ratio in leaves; height, canopy height; SLA, specific leaf area; onset, onset of flowering. For abbreviation of methods see text.

gradients associated with the first RDA axis. In contrast, Cluster regression favoured traits that depended on both gradients. All methods correctly neglected the random trait TV in the three scenarios.

As for the random scenario, all methods except RLQ produced either no or minor responses that did not differ between traits, including TV. In contrast, RLQ produced responses with high variance between the repetitions. However, the sum of eigenvalues (i.e. a global measure of the environment–trait relationships) for the random scenario was almost zero, as mentioned above.

Question 2c: how do groups of species with similar trait expressions respond to environmental gradients?

Apart from CWM-RDA, all methods can be used to produce functional groups and to link them to environmental gradients (Table 4). One important difference was that RDA-RegTree and Cluster regression formed groups from

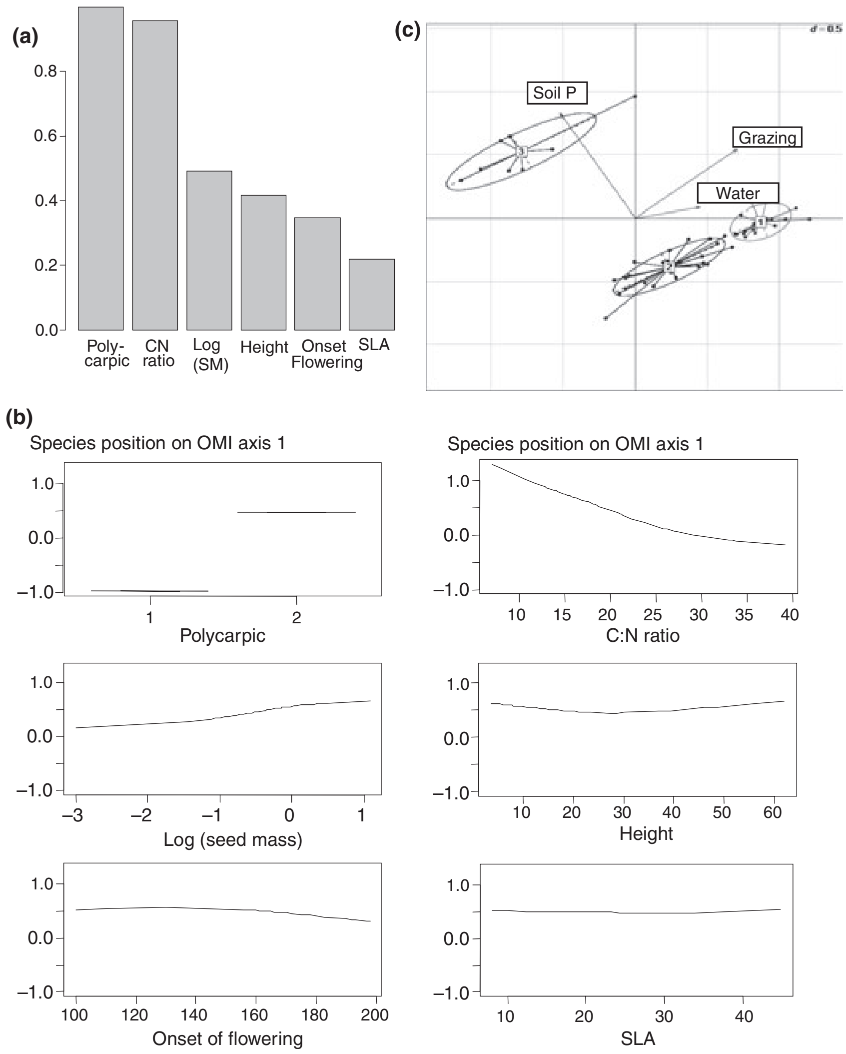

Fig. 3. (**a**) Weight of evidence of the six selected traits to explain the position of species on OMI axis 1 as estimated by the generalized additive models; (**b**) variation of the six selected traits along the OMI axis 1 as estimated by the generalized additive models; (**c**) position of derived functional groups onto the trait–environment space. The percentage variance explained by axis 1 and the correlations of the environmental variables with this axis are given in Table 3. Abbreviations see Fig. 2.

the most responsive traits, whereas OMI-GAM, RLQ and double CCA used all traits for clustering.

Regarding the field data set, all methods primarily divided the total species set into polycarpic and monocarpic species. Monocarps always had a low C:N ratio, while polycarpic species were further split into those with high or low C:N ratio. For further division, onset of flowering and seed mass were relevant for RLQ and double CCA, whereas in Cluster regression height was important, specifically to split polycarpic species with low C:N ratio further into those with short vs tall canopies.

Of the two groups resulting from RDA-mRegTree, the monocarpic plants responded positively to phosphorus and negatively to disturbance (Table 3). Across the other methods, small monocarps with low C:N ratio (Group1) responded positively to phosphorus, while small polycarpic species with low C:N ratio (Group 2) increased with grazing (disturbance) intensity. Conversely, polycarpic species with high C:N ratio decreased with increasing grazing intensity and displayed a unimodal or decreasing distribution with respect to phosphorus (Group 3 in Cluster regression, RLQ, double CCA, OMI-GAM). Group 4 was dissimilar among the methods, either being a mixture of many monocarps and some polycarpic species with low C:N ratio (in RLQ and double CCA), or larger polycarpic species with low C:N ratio in Cluster regression. OMI-GAM produced no fourth group.

The artificial data set showed that the number and stability of groups produced by Cluster regression and RDA-mRegTree varied more strongly than those of RLQ, double CCA and OMI-GAM, as already detected for the field data. The reason for this was that the latter always used all six traits for building functional groups, whereas the former used only responsive traits that could change over the repetitions.

Discussion

We developed a comparative study motivated by the need to produce guidelines for choosing statistical methods suitable to address research questions concerning functional responses of organisms to environmental gradients. Overall, the results produced by the seven methods are concordant regarding the key trait–environment relationships and the species groups they determined. When examined in more detail, however, results differ in the finer detail of selected traits and functional classification. We discuss the main ecological implications of using these different methods in various contexts. A detailed discussion of the statistical features of each procedure is beyond the scope of this paper.

How do traits modulate species response to environmental gradients?

In case of strong habitat filtering of single responsive traits, all methods, whether at the community or species level, should provide similar results. In the field study, all methods consistently revealed that life history was associated with available phosphorus while grazing intensity was related to leaf C:N ratio and canopy height. Other studies have suggested opposite relationships, i.e. changes in plant life span are linked to disturbance intensity whereas changes in traits associated with C gain and nutrient economy (e.g. leaf C:N ratio, height) are linked to soil nutrients (e.g. Westoby et al. 2002; Grime et al. 1997). Interactions between soil properties and the effects of free-ranging herbivores were responsible for the obvious discrepancy between the results reported here and elsewhere. During the night, animals usually rested on a dry, sandy hilltop and defecated here, leading to phosphorus enrichment. The hilltop was also strongly grazed, which, together with dryness and phosphorus enrichment, favoured monocarpic species. Perennial species occurred in other grazed pastures located on former lake sediments in lower parts of the study area. These sites were very rich in calcium, which meant that available phosphorus was almost absent, probably due to immobilization. These site peculiarities explain why the abundance of monocarpic species was more related to phosphorus than to grazing intensity. Increasing canopy height and leaf C:N ratio with decreasing grazing intensity corresponds to results found in many other studies (Quetier et al. 2007; Louault et al. 2005; Nygaard & Ejrnaes 2004; McIntyre et al. 1999).

The broad trait–environment relationships discussed above were revealed in a similar way by the community-based CWM-RDA, the species-based RLQ and double CCA, and were found by the other methods as well. CWM-RDA is a relatively easy method; its ordination techniques are well rooted in the ecological community. It allows assessment of the response of multiple traits and deals with either frequency/abundance or presence/absence data (e.g. Nygaard & Ejrnaes 2004; Table 1). Moreover, community-based methods can easily take into account intraspecific variation in trait expression across populations (Garnier et al. 2007). In this case, CWM-RDA simply uses population- rather than species-level trait values for calculations of community weighted mean trait values. Species-based methods, however, will suffer from inflated trait and abundance tables. In the extreme case, when each plot represents an individual population of a species and traits were measured for all populations of each species, each population occurs only once in the abundance table, all other cells being zero. This extremely sparse matrix might be intractable for regression or ordination methods.

The species-based approach can by definition account for trait variability among species within a community. It therefore allows examination of the response to environmental changes of trait distributions among individual species (Louault et al. 2005), and co-existence or specialization of different life-history strategies along environmental gradients (Grime et al. 1988). Additionally, the species' view is relevant for the (1) functional comparison of biogeographic distributions, regardless of their affiliation to communities (Thuiller et al. 2004), and (2) when phylogenetic constraints on trait–environment relationships are to be analysed (see Table 1; Lord et al. 1995). However, Pillar & Duarte (2010) have recently shown that the phylogenetic signal in trait–environment relationships can also be revealed at the community level.

How do traits relate to environmental gradients at the species level?

To answer question 2, RLQ and double CCA both analyse species–trait and species– environment relationships simultaneously. RLQ has been frequently used to describe trait–environment relationships, for instance the interactions between environment, species traits and human uses to describe patterns of plant invasions (Thuiller et al. 2006) or co-variations between traits of alpine plants and the abiotic environment (Choler 2005).

As an alternative to the simultaneous analysis provided by RLQ and double CCA, question 2 can be addressed by either explaining the species position in environmental space by their traits or the species position in trait space by environmental variables. Conceptually, the first alternative

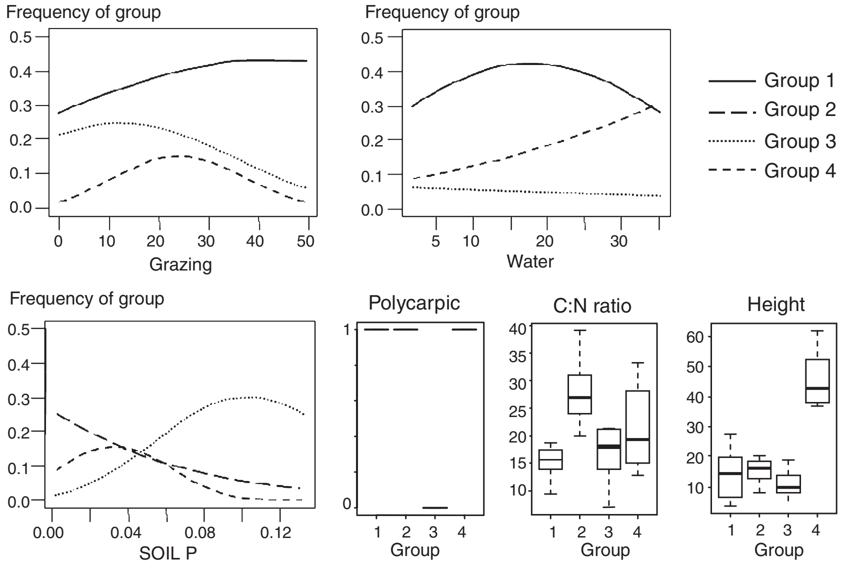

Fig. 4. Response of four functional groups composed with Cluster regression from the traits life cycle, C:N ratio in leaves and canopy height to environmental gradients. The trait expressions characterizing each group are indicated by the boxplots. For instance, group 1 consists of low-growing monocarpic species with intermediate to low C:N ratio. The weight of evidence of the environmental variables in the averaged models is given in Table 3. Abbreviations see Fig. 2.

predicts the species niche from the functional traits whereas the second predicts the functional niche from the environment. In contrast to the species niche where the axes are environmental gradients, a functional niche can be defined as an *n*-dimensional hypervolume in functional space, where the axes are functions (e.g. persistence, regeneration) associated with specific response traits (e.g. SLA, seed mass; Rosenfeld 2002). OMI-GAM and Cluster regression are exponents of both alternatives. OMI-GAM predicts the species environmental niche on trait gradients (Table 1, question 2a). It can thus be used to investigate the form and direction of the relationship between a trait and the environment. In our case study, the relationships are sigmoidal or even slightly unimodal, probably reflecting trade-offs in plant ecophysiology (Fig. 3). In contrast, Cluster regression predicts the distribution of functional groups defined by combinations of multiple traits (i.e. emergent groups; Lavorel et al. 1997) on environmental gradients (question 2c, Fig. 4; see Kühner & Kleyer 2008). Functional groups may be conceived as groups of species with similar functional niches. Functional group distribution models allow the mapping of functional niches to the landscape to infer biological constraints on species ranges and habitat suitability (Kearney & Porter 2009).

How do groups of species with similar trait expressions respond to environmental gradients?

Building functional groups requires analysing combinations of traits rather than single traits. These combinations could give better insight than CWM-RDA into the functional relevance of trait combinations, or syndromes, in the adaptation of species to multiple environmental conditions. For instance, the species-based methods showed that small polycarpic species with low C:N ratio responded positively to grazing, whereas those with high C:N ratio decreased. These different syndromes could not be revealed by CWM-RDA because this method aggregates the traits across co-occurring species.

Between procedures, there were differences in the trait combinations used for the grouping and the number of groups produced from them. While most methods were consistent regarding the importance of life cycle and C:N ratio, the other traits varied in their contribution to the clustering. The regression tree provided a straightforward classification that included only responsive traits. RDA-RegTree selected in a forward hierarchical way the traits that were best predictors of species response. If other traits were correlated with the selected traits they were not included in the model, although they may also be good predictors. Cluster regression needed to sift through all possible combinations of traits and assessed the responsiveness of each combination by using the R^2 and two cluster stability criteria as an objective function. Sifting through all possible combinations of traits to identify the most responsive trait combination can also be done using RLQ (Bernhardt-Römermann et al. 2008). Likewise, Pillar & Sosinski (2003) used a recursive algorithm that searches for the trait combination and number of groups that maximize the response to the environment. In contrast, RLQ, double CCA and OMI-GAM used all traits to produce clusters of species with similar traits and similar response to given environmental factors. Responsive traits had greater weight in the clusters than other traits, although, as the final species scores used for clustering were those maximizing species trait–environment relationships.

In general, differences in combinations of responsive traits among methods should not be over-emphasized. Due to correlations among traits, different trait combinations may still result in similar species groups. For instance, if canopy height is correlated with seed mass (Moles et al. 2004; Leishman et al. 1995), clustering based on canopy height may yield the same species groups as clustering based on seed mass.

However, the analyses of the artificial data set showed that gradient length in both environmental and functional terms requires consideration when choosing the right method. All methods, particularly Cluster regression and RDA-mRegTree, required environmental and corresponding trait heterogeneity to produce correct results. When species became less well separated in functional and environmental space because gradients were shorter or even absent, random effects in species distributions, traits and environmental conditions had an increasing influence on the result of a single run, and hence decreased the consistency of the results across all repetitions. This seemingly trivial fact may however be difficult to assess in real landscapes where gradients are not always strong, or even evident.

Using only responsive traits for the classification of functional groups introduces an additional source of variation that decreased classification stability among repetitions for Cluster regression and RDA-mRegTree. The ordination-based methods extracted some trait–environment correlations even under random conditions. Users should carefully scrutinize their results in situations where chance effects and random fluctuations are likely, such as in early successional or climax stages, and use the fourth-corner testing procedure (Legendre et al. 1997; Dray & Legendre 2008) to evaluate the statistical significance of the trait–environment relationships.

Conclusion

McGill et al. (2006) suggested rebuilding community ecology from functional traits, with emphasis on environmental gradients. Here we present and compare multivariate methods to assist in this attempt. These methods combine several statistical components to analyse three matrices (1) species abundances by plots, (2) plots by environmental factors and (3) species by traits. All methods reproduced the broad trait–environment relationships that were present in the data sets. To yield stable results, all methods require non-random heterogeneity in the environment and corresponding trait values. The most important difference is that each method addresses particular ecological concepts and research questions that cannot be answered by the other methods. If the response of average trait expressions of communities to environmental gradients is the question, a community-level method such as CWM-RDA may be the first choice. However, CWM-RDA does not account for trait variability among species in a community, because trait values are aggregated to a single average. Therefore, this approach cannot be applied to address questions regarding the co-existence of different life histories or to assess how single species respond to environmental changes. In these cases, species-based methods should be applied. RLQ and double CCA provide an excellent overview of the trait–environment relationships at the species level, allow *post-hoc* building of functional groups, and produce repeatable results when gradients are strong. If the user asks whether individual traits determine the response of species to the environment in a linear or non-linear form, OMI-GAM may be chosen. RDA-RegTree and Cluster regression both determine the response of functional groups to the environment. However, the number and identity of traits selected for functional group classification proved to be sensitive to random components in the artificial data set. RDA-RegTree allows assessment of the hierarchical combination of trait expressions that determine the environmental response of species. If the user wants to predict the distribution of functional groups on environmental gradients, Cluster regression is the appropriate choice. All in all, these methods cover different aspects in the assessment of functional responses to environmental gradients, in terms of ecological concepts, usability and application. Thus, once a user has clarified the research question, the appropriate method can be chosen and adapted for any similar data by using the scripts in Appendix S4 and the tutorial in Appendix S2.

Acknowledgements

The authors thank the Müritz National Park Authority for permission to sample vegetation data in the park and Tonia Meier for fieldwork. We also thank several anonymous reviewers for help in improving the manuscript. This work resulted from the EU project VISTA (Vulnerability of Ecosystem Services to Land Use Change in Traditional Agricultural Landscapes, contract no. EVK2-2001-15 000356). BS acknowledges support from the ESF EuroDIVERSITY 040 project ASSEMBLE (DFG KL756/2-1). FdB was funded through the ANR DIVERSITALP project (contract No ANR 07 BDIV 014). WT acknowledges support from EU funded EcoChange project (FP6 European Integrated project 2007 –2011, contract No066866 GOCE). JL acknowledges support of GACR 206/09/1471. SD acknowledges support of ANR 07 BDIV 014.

References

Ackerly, D.D., Knight, C.A., Weiss, S.B., Barton, K. & Starmer, K. P. 2002. Leaf size, specific leaf area and microhabitat distribution of chaparral woody plants: contrasting patterns in species level and community level analyses. *Oecologia* 130: 449–457.

de Bello, F., Leps, J., Lavorel, S. & Moretti, M. 2007. Importance of species abundance for assessment of trait composition: an example based on pollinator communities. *Community Ecology* 8: 163–170.

Bernhardt-Römermann, M., Römermann, C., Nuske, R., Parth, A., Klotz, S., Schmidt, W. & Stadler, J. 2008. On the identification of the most suitable traits for plant functional trait analyses. *Oikos* 117: 1533–1541.

Bonada, N., Dolédec, S. & Statzner, B. 2007. Taxonomic and biological trait differences of stream macroinvertebrate communities between mediterranean and temperate regions: implications for future climatic scenarios. *Global Change Biology* 13: 1658–1671.

ter Braak, C.J.F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. *Ecology* 67: 1167–1179.

Breiman, L., Friedman, J.H., Olshen, R.A. & Stone, C.J. 1984. *Classification and regression trees*. Chapman & Hall, New York, NY, US.

Burnham, K.P. & Anderson, D.R. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd edn. Springer, New York, NY, US.

Choler, P. 2005. Consistent shifts in Alpine plant traits along a mesotopographical gradient. *Arctic, Antarctic and Alpine Research* 37: 444–453.

Cingolani, A.M., Cabido, M., Gurvich, D.E., Renison, D. & Diaz, S. 2007. Filtering processes in the assembly of plant communities: are species presence and abundance driven by the same traits? *Journal of Vegetation Science* 18: 911–920.

Crutsinger, G.M., Souza, L. & Sanders, N.J. 2007. Intraspecific diversity and dominant genotypes resist plant invasions. *Ecology Letters* 11: 16–23.

De'Ath, G. 2002. Multivariate regression trees: a new technique for modeling species–environment relationships. *Ecology* 83: 1105–1117.

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.

Dolédec, S., Chessel, D. & Gimaret-Carpentier, C. 2000. Niche separation in community analysis: a new method. *Ecology* 81: 2914–2927.

Dolédec, S., Lamouroux, N., Fuchs, U. & Mérigoux, S. 2007. Modelling the hydraulic preferences of benthic macroinvertebrates in small European streams. *Freshwater Biology* 52: 145–164.

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

Dray, S., Chessel, D. & Thioulouse, J. 2003. Co-inertia analysis and the linking of ecological data tables. *Ecology* 84: 3078–3089.

Everitt, B.S., Landau, S., Leese, M. & Stahl, D. 2011. *Cluster analysis*, 5th edn. Wiley, Chichester, UK.

Fox, B.J. & Brown, J.H. 1993. Assembly rules for functional groups in North American desert rodent communities. *Oikos* 67: 358–370.

Garnier, E., Lavorel, S., Ansquer, P., Castro, H., Cruz, P., Dolezal, J., Eriksson, O., Fortunel, C., Freitas, H., Golodets, C., Grigulis, K., Jouany, C., Kazakou, E., Kigel, J., Kleyer, M., Lehsten, V., Leps, J., Meier, T., Pakeman, R., Papadimitriou, M., Papanastasis, V.P., Quested, H., Quetier, F., Robson, M., Roumet, C., Rusch, G., Skarpe, C., Sternberg, M., Theau, J. P., Thebault, A., Vile, D. & Zarovali, M.P. 2007. Assessing the effects of land-use change on plant traits, communities and ecosystem functioning in grasslands: a standardized methodology and lessons from an application to 11 European sites. *Annals of Botany* 99: 967–985.

Gordon, A.D. 1999. *Classification*, 2nd edn. Chapman & Hall, Boca Raton, FL, US.

Grime, J.P. 2006. Trait convergence and trait divergence in herbaceous plant communities: mechanisms and consequences. *Journal of Vegetation Science* 17: 255–260.

Grime, J.P., Hodgson, J.G. & Hunt, R. 1988. *Comparative plant ecology*. Unwin-Hyman, London, UK.

Grime, J.P., Thompson, K., Hunt, R., Hodgson, J.G., Cornelissen, J.H.C., Rorison, I.H., Hendry, G.A.F., Ashenden, T.W., Askew, A.P., Band, S.R., Booth, R.E., Bossard, C.C., Campbell, B.D., Cooper, J.E.L., Davison, A.W., Gupta, P.L., Hall, W., Hand, D.W., Hannah, M.A., Hillier, S.H., Hodkinson, D. J., Jalili, A., Liu, Z., Mackey, J.L., Matthews, N, Mowforth, M.A., Neal, A.M., Reader, R.J., Reiling, K., Ross-Fraser, A. M., Spencer, R.E., Sutton, F., Tasker, D.E., Thorpe, P.C. & Whitehouse, J. 1997. Integrated screening validates a primary axis of specialisation in plants. *Oikos* 79: 259–281.

Hérault, B. & Honnay, O. 2005. The relative importance of local, regional and historical factors determining the distribution of plants in fragmented riverine forests: an emergent group approach. *Journal of Biogeography* 32: 2069–2081.

Hood, R. R., Laws, E.A., Armstrong, R.A., Bates, N.R., Brown, C. W., Carlson, C.A., Chai, F., Doney, S.C., Falkowski, P.G., Feely, R.A., Friedrichs, M.A.M., Landry, M.R., Moore, J.K., Nelson, D.M., Richardson, T.L., Salihoglu, B., Schartau, M., Toole, D.A. & Wiggert, J. D. 2006. Pelagic functional group modeling: progress, challenges and prospects. *Deep-Sea Research Part II-Topical Studies in Oceanography* 53: 459–512.

Jäger, E.J. & Werner, K. 2002. *Exkursionsflora von Deutschland*, Vol. 4, 9th edn. Spektrum, Berlin, DE.

Kearney, M. & Porter, W. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters* 12: 334–350.

Körner, C. 2003. *Alpine plant life: functional plant ecology of high mountain ecosystems*, 2nd edn. Springer, Berlin and Heidelberg, DE.

Kühner, A. & Kleyer, M. 2008. A parsimonious combination of functional traits predicting plant response to disturbance and soil fertility. *Journal of Vegetation Science* 19: 681–642.

Lavorel, S. & Garnier, E. 2002. Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Functional Ecology* 16: 545–556.

Lavorel, S., McIntyre, S., Landsberg, J. & Forbes, T.D.A. 1997. Plant functional classifications: from general groups to specific groups based on response to disturbance. *Trends in Ecology and Evolution* 12: 474–478.

Lavorel, S., Rochette, C. & Lebreton, J. 1999. Functional groups for response to disturbance in Mediterranean old fields. *Oikos* 84: 480–498.

Lavorel, S., Díaz, S., Cornelissen, J.H.C., Garnier, E., Harriso, S.P., McIntyr, S., Pausas, J.G., Pérez-Harguindeguy, N., Roume, C. & Urcelay, C. 2007. Plant functional types: are we getting any closer to the holy grail? In: Canadell, J.G, Pataki, D. & Pitelk, L. (eds.) *Terrestrial ecosystems in a changing world*. pp. 149–160. The IGBP Series, Springer, Berlin, Heidelberg, DE.

Legendre, P. & Legendre, L. 1998. *Numerical ecology*, 2nd edn. Elsevier, Amsterdam, NL.

Legendre, P., Galzin, R. & HarmelinVivien, M.L. 1997. Relating behavior to habitat: solutions to the fourth-corner problem. *Ecology* 78: 547–562.

Lehsten, V., Harmand, P. & Kleyer, M. 2009. Fourth Corner Generation of Plant Functional Response Groups. *Environmental and Ecological Statistics* 16: 561–584.

Leishman, M.R., Westoby, M. & Jurado, E. 1995. Correlates of seed size variation: a comparison among five temperate floras. *Journal of Ecology* 83: 517–530.

Lord, J., Westoby, M. & Leishman, M. 1995. Seed size and phylogeny in 6 temperate floras – constraints, niche conservatism, and adaptation. *American Naturalist* 146: 349–364.

Louault, F., Pillard, V D., Aufrère, J., Garnier, E. & Soussana, J.-F. 2005. Plant traits and functional types in response to reduced disturbance in a semi-natural grassland. *Journal of Vegetation Science*, 16: 151–160.

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.

McIntyre, S. & Lavorel, S. 2001. Livestock grazing in subtropical pastures: steps in the analysis of attribute response and plant functional types. *Journal of Ecology* 89: 209–226.

McIntyre, S., Lavorel, S., Landsberg, J. & Forbes, T.D.A. 1999. Disturbance response in vegetation – towards a global perspective on functional traits. *Journal of Vegetation Science* 10: 621–630.

Moles, A.T., Falster, D.S., Leishman, M.R. & Westoby, M. 2004. Small-seeded species produce more seeds per square metre of canopy per year, but not per individual per lifetime. *Journal of Ecology* 92: 384–396.

Moretti, M., de Bello, F., Roberts, S.P.M. & Potts, S.G. 2009. Taxonomical vs. functional responses of bee communities to fire in two contrasting climatic regions. *Journal of Animal Ecology* 78: 98–108.

Nygaard, B. & Ejrnaes, R. 2004. A new approach to functional interpretation of vegetation data. *Journal of Vegetation Science* 15: 49–56.

Pakeman, R.J. 2004. Consistency of plant species and trait responses to grazing along a productivity gradient: a multi-site analysis. *Journal of Ecology* 92: 893–905.

Pillar, V. D. & Duarte, L. D. 2010. A framework for metacommunity analysis of phylogenetic structure. *Ecology Letters* 13: 587–596.

Pillar, V.D. & Sosinski, E.E. Jr 2003. An improved method for searching plant functional types by numerical analysis. *Journal of Vegetation Science* 14: 232–332.

Pillar, V. D., Duarte, L. D., Sosinski, E. E. & Joner, F. 2009. Discriminating trait-convergence and trait-divergence assembly patterns in ecological community gradients. *Journal of Vegetation Science* 20: 334–348.

Quetier, F., Thebault, A. & Lavorel, S. 2007. Plant traits in a state and transition framework as markers of ecosystem response to land-use change. *Ecological Monographs* 77: 33–52.

Rao, C.R. 1964. The use and interpretation of principal correspondence analysis in applied research. *Sankhya A* 26: 329–359.

Rosenfeld, J.S. 2002. Functional redundancy in ecology and conservation. *Oikos* 98: 156–162.

Suding, K.N., Lavorel, S., Chapin, F.S., Cornelissen, J.H.C., Diaz, S., Garnier, E., Goldberg, D., Hooper, D.U., Jackson, S.T. & Navas, M.L. 2008. Scaling environmental change through the community-level: a trait-based response-and-effect framework for plants. *Global Change Biology* 14: 1125–1140.

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.

Thuiller, W., Richardson, D.M., Rouget, M., Procheş, Ş. & Wilson, J.R.U. 2006. Interactions between environment, species traits, and human uses describe patterns of plant invasions. *Ecology* 87: 1755–1769.

Violle, C., Navas, M.L., Vile, D., Kazakou, E., Fortunel, C., Hummel, I. & Garnier, E. 2007. Let the concept of trait be functional! *Oikos* 116: 882–892.

Westoby, M., Leishman, M. & Lord, J. 1995. Issues of interpretation after relating comparative datasets to phylogeny. *Journal of Ecology* 83: 892–893.

Westoby, M., Falster, D.S., Moles, A.T., Vesk, P.A. & Wright, I.J. 2002. Plant ecological strategies: some leading dimensions of variation between species. *Annual Review of Ecology and Systematics* 33: 125–159.

Whitlock, R., Grime, J.P., Booth, R. & T. Burke, T. 2007. The role of genotypic diversity in determining grassland community structure under constant environmental conditions. *Journal of Ecology* 95: 895–907.

Wilson, J.B. 1999. Guilds, functional types and ecological groups. *Oikos* 86: 507–522.

Woodward, F.I. & Cramer, W. 1996. Plant functional types and climatic changes: introduction. *Journal of Vegetation Science* 7: 306–308.

Woodward, F.I. & Diament, A.D. 1991. Functional approaches to predicting the ecological effects of global change. *Functional Ecology* 5: 202–212.

Supporting Information

Additional supporting information may be found in the online version of this article:

Appendix S1. Detailed description of the case study site, sampling methods and species′ responses to environmental gradients.

Appendix S2. Tutorial: methods for assessing functional responses to environmental gradients.

Appendix S3. The artificial data set.

Appendix S4. R scripts and field data set.