Including community composition in biodiversity–productivity models

Nadine Sandau[1], Rudolf P. Rohr[1,2], Russell E. Naisbit[1], Yvonne Fabian[1], Odile T. Bruggisser[1], Patrik Kehrli[3], Alexandre Aebi[4] and Louis-Félix Bersier[1*]

[1]*Department of Biology, University of Fribourg, Chemin du Musée 10, 1700 Fribourg, Switzerland;* [2]*Integrative Ecology Group, Estacion Biologica de Doñana, EBD-CSIC, 41092 Sevilla, Spain;* [3]*Station de recherche Agroscope Changins – Wädenswil ACW, 1260 Nyon, Switzerland; and* [4]*Laboratory of Soil Biology, University of Neuchâtel, Emile-Argand 11, 2000 Neuchâtel, Switzerland*

Summary

1. Studies on biodiversity and ecosystem functioning (BEF) have elicited debate over the interpretation of the positive relationship between species richness and plant productivity. Manipulating richness cannot be achieved without affecting composition; it is thus essential to consider the latter in statistical models.

2. We firstly review existing approaches that use species richness as an explanatory variable and propose modifications to improve their performance. We use an original data set to illustrate the analyses. The classical method where composition is coded as a factor with a level for each different species mixture can be improved by defining the levels using clustering. Methods based on ordinations reduce the dimensionality of plant composition and use the new coordinates as fixed effects; they provide a much better fit to our observations.

3. Secondly, we develop a new method where composition is included as a similarity matrix affecting the residual variance–covariance. Similarity in composition between plots is treated in the same way as shared evolutionary history between species in phylogenetic regression. We find that it outperforms the other models.

4. We discuss the different approaches and suggest that our method is particularly suited for observational studies or for manipulative studies where plant diversity is not kept constant by weeding. By treating species composition in an intuitive and sensible way, it offers a valuable and powerful complement to existing models.

Key-words: BEF, biodiversity, composition, ecosystem function, ecosystem services, mixed effects models, residual correlation structure, species richness

Introduction

Species are being lost at an unprecedented rate, prompting fears that ecosystem services will suffer. Such concerns have inspired many manipulative biodiversity experiments such as Cedar Creek, BIODEPTH or The Jena Experiment (Tilman, Wedin & Knops 1996; Hector *et al.* 1999; Roscher *et al.* 2004) to examine the causal relationships between biodiversity and ecosystem functioning (BEF) (Hooper *et al.* 2012). Some experiments analyse ecosystem multifunctionality, considering services such as nutrient cycling and decomposition (Gamfeldt, Hillebrand & Jonsson 2008; Zavaleta *et al.* 2010; Mouillot *et al.* 2011), but the majority concentrate on the diversity–productivity relationship and have typically revealed a positive relationship between the two (Hector *et al.* 1999; Van Ruijven & Berendse 2003; Roscher *et al.* 2005). The topic has generated several hypotheses (mass ratio hypothesis, diversity

hypothesis...) and a vast literature, but remains contentious because outcomes from natural ecosystems and synthetic plant communities differ greatly (Chapin *et al.* 1997; Grace *et al.* 2007; Assaf, Beyschlag & Isselstein 2011). The results of the first experiments stirred up heated debate (Kaiser 2000), where critics argued that the results of the randomly assembled biodiversity-ecosystem functioning experiments were consequences of 'hidden treatments'; amongst other issues, the effect of species composition was not considered in most of the statistical analyses (Huston 1997). This has serious consequences, for example, the so-called variance reduction effect arises from the overlap in species composition in higher diversity mixtures created by random draws from a species pool (Huston 1997), showing that community composition is not independent when subplots share species. Furthermore, the idiosyncratic hypothesis predicts that ecosystem functioning changes with changing diversity, but unpredictably; as species contribute unequally to productivity, when members of the community are lost, species identity/composition is important, not merely the resulting number of species (Lawton 1994). In response to

*Correspondence author. Louis-Félix Bersier, E-mail: louis-felix. bersier@unifr.ch

these discussions about the drawbacks of previous experiments, Bell *et al.* (2009) developed a new experimental design that allows separation of species richness and species identity.

Attempts to disentangle diversity effects from compositional effects are further complicated by the very different nature of the two variables: plant species richness is a quantitative (integer) variable that is easy to include in statistical models; composition can be described by the presence/absence or abundance of each species in the experimental plots. Composition is thus inherently multivariate, and each plot can be visualised as a point in an n-dimensional space, where n corresponds to the total number of species in the study. The contrasting nature of richness and composition raises analytical problems when considered together in a statistical model.

Several researchers have proposed statistical procedures to account for community composition while analysing the relationship between species richness and measures of ecosystem function, but each has its limitations. Some are suited for experimental studies where plots are weeded to maintain the sown plant composition. Others were proposed for the analysis of natural systems where richness and composition were not manipulated. Here, we discuss these existing methods, where species richness is explicitly used as explanatory variable, and propose new approaches to incorporate plant composition in the statistical model. We illustrate the methods using data from an original study of plant productivity in experimental wildflower strips in agricultural fields, where repeated diversity levels were sown in blocks across a set of 12 fields, and the plant communities were not weeded, thus creating an intermediate situation between natural and controlled designs. We do not consider alternative models for the study of BEF relationships, for example those based on the Price equation (Fox 2006; Fox & Harpole 2008), the diversity interaction model and its variants (Kirwan *et al.* 2007; Connolly *et al.* 2013), or the method of Bell *et al.* (2009), because they do not apply easily to our experiment (see Discussion section).

The statistical models differ in the way that plant composition is included. (1) We start with methods that reduce the dimensionality of the composition variable, typically using ordination methods like nonmetric multidimensional scaling (NMDS) or correspondence analysis (CA) (Legendre & Legendre 1998). Following Kahmen *et al.* (2005), we use NMDS to transform plant composition into few quantitative variables that are then included as fixed factors in the linear mixed-effects models (LMMs). This method was devised for the analysis of natural experiments. (2) We then discuss the treatment of composition as a multilevel categorical variable (Troumbis *et al.* 2000; Caldeira *et al.* 2005; Hector *et al.* 2011). This approach emerges from manipulative studies where plant species composition was a priori treated as a random effect: a new level is created for each distinct composition, considered to be one realisation drawn from all possible mixtures. However, such an approach does not take into account the fact that some compositions must be very similar and

others very distinct (Fukami, Naeem & Wardle 2001; Schmid *et al.* 2002). (3) To alleviate this limitation, we suggest an approach where levels are the outcome of a clustering analysis of plant species composition in the subplots. For this purpose, we use the partitioning around medoids clustering method (PAM) based on Jaccard dissimilarities. (4) We propose a new approach where we account for the fact that floristic similarity between subplots creates a violation of independence between observations. We used pairwise Jaccard similarities between subplots to capture the floristic correlation and included this similarity matrix in the model as correlation structure. This approach is built on classical methods dealing with auto-correlated data (e.g. in time-series analyses) or nonindependent objects (e.g. in phylogenetic regression, where species are linked by the underlying phylogeny; Freckleton, Harvey & Pagel 2002).

Methods

To illustrate the different methods of incorporating plant composition into statistical models, we use an original data set from an experiment conducted in 12 wildflower strips in Switzerland (Bruggisser *et al.* 2012; Fabian *et al.* 2012). Details are provided in Appendix S1 and Table S1 (Supporting Information). The number of species was manipulated by sowing different mixtures in 2007. A treatment to control the main herbivores and their predators was applied with fences of different mesh size (variable Treat with three levels, *C*, PE, PHE, for control, predator exclusion, and predator and herbivore exclusion, respectively). Each of the 12 experimental fields (random factor Field as an indicator variable) contained the three levels of Treat, and in each level, four 9 m × 6 m subplots were sown with different mixtures (2, 6, 12 and 20 sown plant species in a split-plot design with 12 × 3 blocks). The fields were not weeded, and species richness is expressed as the total number of species *S*, not as the number of sown species. Plant species composition was described by the percentage cover of all plant species in the subplots, evaluated by visual inspection. The response variable is total above-ground biomass TB (dry weight, expressed in the analyses as g/0·45 m²), which was estimated from leaf area index (LAI) measurements. We use here data from 2008.

Our analyses were based on LMMs with *TB* as dependent variable, Treat and *S* as fixed effects, and represented the split-plot design by including random factors for Field and for the Treat block nested within Field. We included the interaction between Treat and *S* in all models. We only considered random intercept models (Zuur *et al.* 2009) as we often encountered convergence problems with models including a random slope for *S*. We used a 'reference' model, which does not include plant composition, as a benchmark for the different approaches treated below. This reference model is given by:

$$
\begin{aligned}
\mathrm{TB}_{ijk} = {} & \beta_1(\mathrm{Treat}=C)_{ijk} + \beta_2(\mathrm{Treat}=C)_{ijk}:(S)_{ijk} \\
& + \beta_3(\mathrm{Treat}=\mathrm{PE})_{ijk} + \beta_4(\mathrm{Treat}=\mathrm{PE})_{ijk}:(S)_{ijk} \\
& + \beta_5(\mathrm{Treat}=\mathrm{PHE})_{ijk} + \beta_6(\mathrm{Treat}=\mathrm{PHE})_{ijk}:(S)_{ijk} \\
& + z_k + z_{j|k} + \varepsilon_{ijk}
\end{aligned}
\qquad \text{eqn 1}
$$

The indices *i* denote the observation (subplot), *j* denote the block and *k* denote the field. The parameters β_1 and β_2 are the intercept and the slope with species richness for the control treatment, β_3 and β_4 are the intercept and the slope for the predator exclusion treatment, and β_5 and β_6 are the intercept and the slope for the predator and herbivore exclusion treatment. The terms z_k and $z_{j|k}$ are the random factors corre-

sponding to the field and blocks nested within field (split-plot design), respectively. The final term ε_{ijk} represents the residuals. For simplicity, we define:

$$
\begin{aligned}
B(\text{Treat} * S)_{ijk} = {} & \beta_1(\text{Treat} = C)_{ijk} + \beta_2(\text{Treat} = C)_{ijk} : (S)_{ijk} \\
& + \beta_3(\text{Treat} = PE)_{ijk} + \beta_4(\text{Treat} = PE)_{ijk} : (S)_{ijk} \\
& + \beta_5(\text{Treat} = PHE)_{ijk} + \beta_6(\text{Treat} = PHE)_{ijk} : (S)_{ijk}
\end{aligned}
$$

for use in the equations that follow.

REDUCING THE DIMENSIONALITY OF COMPOSITIONAL DATA

Transforming the floristic data using ordination methods enables community composition to be represented in few informative dimensions. The coordinates of the subplots along a chosen number of axes form the new variables. We used NMDS for this purpose, based on a Jaccard dissimilarity matrix calculated from the presence–absence of species in each subplot. NMDS uses rank order rather than Euclidian distances and hence is less prone to be affected by non-normally distributed data while also being considered very robust (Minchin 1987). We applied the metaMDS function of the vegan package (Oksanen *et al.* 2013), using 1000 random starting points. We chose to represent our data in two dimensions, which resulted in a stress value of 0·242, indicating that composition may not be perfectly represented by two axes. It is, however, known that stress values increase with greater numbers of observations (Wantzen *et al.* 2008), and we opted for this number of axes to avoid including too many explanatory variables in the LMMs. The statistical model read as follows:

$$
\begin{aligned}
\text{TB}_{ijk} = {} & B(\text{Treat} * S)_{ijk} \\
& + \beta_7(X)_{ijk} + \beta_8(Y)_{ijk} + z_k + z_{j|k} + \varepsilon_{ijk}.
\end{aligned}
\qquad \text{eqn 2}
$$

Parameters β_7 and β_8 are the slopes for the NMDS axes, and the terms z_k, $z_{j|k}$ and ε_{ijk} are the same as in eqn 1.

As far as we are aware, former studies using this approach in natural experiments with plants (Kahmen *et al.* 2005; Schultz *et al.* 2011) or in controlled experiments with lichens (Maestre *et al.* 2012) considered only linear effects of the NMDS variables. The coordinates of the plots on the different axes were used to reflect composition. However, the values of the coordinates (X and Y for the first and second NMDS axes, respectively) reflect some gradient of similarity in plant composition between the subplots, and there is no reason *a priori* to expect these values to be linearly related to a dependent variable (here TB). Consequently, we investigated potential nonlinear relationships using two additional analyses. Firstly, we included higher-order terms (i.e. X^2, Y^2, XY, X^3, Y^3, X^2Y and XY^2) in eqn 2. This approach is similar to using higher-order terms in spatial analyses to capture nonlinear gradients (a trend surface analysis; see Borcard, Legendre & Drapeau 1992 or Legendre & Legendre 1998). Secondly, we used a Generalized Additive Mixed Model (GAMM), which uses a smoothing function to link the response and explanatory variables, thus being very flexible in coping with nonlinearities in the data.

COMPOSITION AS A CATEGORICAL VARIABLE: LEVELS BASED ON EXPERIMENTAL SPECIES MIXTURE

Working with controlled BEF experiments where subplots were weeded, Hector *et al.* (2011) and others (Troumbis *et al.* 2000; Fridley 2003; Caldeira *et al.* 2005; Dimitrakopoulos *et al.* 2006)

included composition as a categorical variable where a new level was defined for each different sown mixture. We took the same approach, resulting in 37 levels. In our case, this does not fully reflect actual composition, because 'invader' plants are not considered (which would otherwise have resulted in 144 levels, leaving no degrees of freedom to test the effect of diversity). In the approach of Hector *et al.* (2011), this categorical variable was included as a random effect, which was crossed with experimental site. In our case, as only four of the 37 levels were present in each field, we nested the composition variable within block, which is in turn nested within field. The statistical model is:

$$
\text{TB}_{ijk} = B(\text{Treat} * S)_{ijk} + z_k + z_{j|k} + z_{c(i)|j|k} + \varepsilon_{ijk}, \qquad \text{eqn 3}
$$

with symbols as for eqn 2, except the term $z_{c(i)|j|k}$ that corresponds to the random factor of the compositional category $c(i)$ of observation i nested within block j and within field k.

COMPOSITION AS A CATEGORICAL VARIABLE: LEVELS BASED ON FLORISTIC SIMILARITY

A disadvantage of the above approach is that similarity between the levels is not taken into account: for example, mixtures sharing all except one species are treated as completely independent. This limitation can be seen as problematic especially in observation studies, and we suggest using *a posteriori* categorisation of subplots by clustering to define groups based on floristic composition. In our case, additional advantages are (1) that the analysis can be performed on all species present, not only those that were sown and (2) that pairwise similarity between subplots can be based on species presence/absence or percentage cover. Firstly, we computed a compositional dissimilarity matrix using Jaccard dissimilarity (Legendre & Legendre 1998) based on presence–absence data with the function vegdist from the vegan package (Oksanen *et al.* 2013) in R. We then used the nonhierarchical PAM algorithm implemented in the cluster package (Maechler *et al.* 2012) to determine group membership. PAM reduces the dissimilarity within and maximises it between clusters. In our case, the first maximum of the average silhouette width equalled 0·656, resulting in three clusters (see Fig. S1, Supporting Information). We transformed these cluster groups into a categorical variable that was nested within block and within field in the random effects part of the LMM.

The equation for this statistical model is the same as eqn 3, except that $c(i)$ refers here to the compositional group of observation i obtained by the clustering.

COMPOSITION AS CORRELATION STRUCTURE

Here, we propose a new approach in which floristic similarity between subplots is considered to create a violation of independence between observations, similar to the way that shared phylogenetic history makes species dependent in phylogenetic regression (Freckleton, Harvey & Pagel 2002). We used the Jaccard coefficient to represent the similarity in plant composition amongst subplots and included this similarity matrix in the residual variance of the model. The model includes a parameter λ that multiplies the off-diagonal elements of the similarity matrix and that can be interpreted as a measure of the importance of floristic composition. A value of 0 would indicate that composition does not account for TB and that the subplots are independent with regard to composition; a value of 1 indicates that the contribution of composition to the error variance

is similar to that generated by the subplots taken as independent. The equation for this model can be written as:

$$\text{TB}_{ijk} = B(\text{Treat} * S)_{ijk} + z_k + z_{j|k} + \epsilon_{ijk},$$

with $\epsilon_{ijk} \sim N(0, \Sigma)$

$$\Sigma = \sigma^2 \left[\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} + \lambda \begin{pmatrix} 0 & \cdots & b_{(i_1 j_1 k_1)(i_2 j_2 k_2)} \\ \vdots & \ddots & \vdots \\ b_{(i_1 j_1 k_1)(i_2 j_2 k_2)} & \cdots & 0 \end{pmatrix} \right], \qquad \text{eqn 4}$$

σ^2 the error variance, λ the parameter capturing the strength of the correlation induced by the plant compositional structure, $b_{(i_1 j_1 k_1)(i_2 j_2 k_2)}$ the Jaccard similarity between pairs of observations i_1 and i_2 in treat blocks j_1 and j_2 within fields k_1 and k_2, and other terms as in eqn 1. Note that to have a well-defined variance–covariance matrix, it is necessary that the eigenvalues of the correlation matrix Σ are all positive, and this for any value of λ between 0 and 1, which can be easily tested by computing them for different values of λ.

The likelihood function for the multivariate normal model with random effects is given by:

$$\mathcal{L}(\text{TB}_{ijk} | \beta_{1-6}, \sigma_{z_k}, \sigma_{z_{j|k}}, \sigma, \lambda) = \frac{1}{(2\pi)^{N/2} \det(\Omega)^{1/2}} e^{-1/2(\overrightarrow{\text{TB}} - \vec{B}(\text{Treat}*\text{S}))^\text{T} \Omega^{-1}(\overrightarrow{\text{TB}} - \vec{B}(\text{Treat}*\text{S}))}$$

where N is the number of observations, $\overrightarrow{\text{TB}}$ is the vector of observation, $\vec{B}(\text{Treat} * S)$ is the vector of the fixed effects and Ω is the variance–covariance matrix. Ω is given by

$$\Omega = \Sigma + (\sigma_{z_k})^2 Z_k Z_k^T + (\sigma_{z_{j|k}})^2 Z_{j|k} Z_{j|k}^T,$$

with Z_k and $Z_{j|k}$ the design matrices for the two random effects, and the σ terms their corresponding standard deviation.

DATA ANALYSIS

Three subplots (sown species richness = 2, Field 8) represented productivity outliers and were excluded from the analyses. The variable *TB* was Box–Cox-transformed using the car package (Fox & Weisberg 2011), with the power parameter $\gamma = 0.40$. Variables expressing composition based on NMDS axes were scaled (zero mean and unit standard deviation) before regression analyses.

Assumptions of normality of residuals were examined with Q–Q plots. For the models with correlation structure, this raises a problem as the residuals have a variance–covariance structure defined by Σ in eqn 4 and that induced by the two random factors. They should be 'decorrelated' before checking for normality (Houseman, Ryan & Coull 2004). This is achieved by the transformation $\tilde{r} = L^T r$, with r and $\tilde{r}$ the vector of residuals and of transformed residuals, respectively; L^T is the transpose of the lower triangular matrix, L, from Cholesky decomposition of Ω^{-1}, the inverse of the matrix Ω (L^T can be thought of as the square root of the matrix Ω^{-1}; see Houseman, Ryan & Coull 2004).

For LMM analyses, we used the lme procedure of the nlme package (Pinheiro & Bates 2010), using maximum likelihood (ML) when model selection was applied, and restricted maximum likelihood

(REML) for parameter estimation once the best model was determined (Zuur *et al.* 2009). For the model including species composition as a similarity matrix in the residual variance, we wrote an R routine (provided in the Supporting Information) following chapter 9·4·2, on linear mixed models in Davison (2003); ML and REML methods were used as above. We also checked that our code yielded results similar to lme for models without correlation, which was the case. We compared the performance of the different models with AIC. Because AIC is known to favour overfitting, we also used more stringent AICc and BIC methods. All analyses were performed with R (R Development Core Team 2013).

Results

Table 1 presents different measures of performance of the models, and Table 2 presents the estimates and relative importance of the parameters of the different models. Descriptive statistics and detailed results for all the models are given in Tables S2–S7 (Supporting information). We observe that the model with composition as correlation in the residual variance outperforms the others and that most provide a better fit than the 'reference' model (eqn 1) lacking any term accounting for composition.

Based on the reference model without composition, total biomass was not significantly different between the three levels of the exclusion treatment (log-likelihood ratio test with a model without *Treat*: $P = 0.14$; note that the *P*-values associated with the factor *Treat* in Table 2 are all significant because the intercept is not included; they simply indicate that the biomasses are different from 0). The weak effect of the variable *Treat* can also be seen in Table 2, where the relative importance of the explanatory variables are evaluated with differences in AIC between the full model and the model without the variable of interest. There was a significant interaction between species richness *S* and the exclusion treatment, with the effect of species richness being very negative and significant in the control treatment (*C*), and more weakly negative and nonsignificant in both the predator exclusion (PE) and predator and herbivore exclusion (PHE) treatments. We will now see that including composition does not alter the significance or the sign of the parameters markedly, but depending on the approach used, can greatly improve the goodness-of-fit of the models.

REDUCING THE DIMENSIONALITY OF COMPOSITION DATA

The results of the mixed effects model (eqn 2) are given in Table 2. The relationship between total biomass and species richness remains negative in the control treatment, but the *P*-value is not anymore significant (see Table S4, Supporting information). Only the first NMDS axis is significant. The inclusion of plant composition as NMDS axes provides a major improvement over the model without composition (ΔAIC = −20·85). Examination of the NMDS plot (Fig. 1), where biomass is indicated by the size of the points, clearly shows that the first axis separates subplots according to biomass; it is also apparent that species richness is strongly negatively correlated with the first NMDS axis (indicated by the colour gradient), which explains why the *Treat:S* term is not significant in this model. Including higher-order terms in the model did not improve the goodness-of-fit, which is confirmed by the results of the GAMMs: the AIC with the smoother is higher than that of the linear model. Thus, there is no support for nonlinearity in our case.

COMPOSITION AS A CATEGORICAL VARIABLE: LEVELS BASED ON EXPERIMENTAL SPECIES MIXTURE

Coding each different sown plant mixture as a new level of a categorical variable for composition yields results for the vari-

Table 1. Comparison of the fives models

Model:	Reference model (without composition)	Composition as NMDS axes	Composition as random variable (based on experimental mixture)	Composition as random variable (based on PAM clustering)	Composition as correlation in the residual variance
Log like	−321·84	−309·05	−321·84	−320·40	−307·78
AIC	661·68	640·09	663·68	660·80	635·55
ΔAIC	26·13	5·28	28·13	25·25	0
Relative AIC	0	0·071	0	0	1
Akaike weight	0	0·067	0	0	0·933
AICc	645·06	620·16	645·38	642·51	617·26
BIC	688·15	672·45	693·09	690·22	664·97

We provide the maximum likelihood estimation, AIC, ΔAIC, Relative AIC, Akaike weight, AICc and the BIC for each of the five models. The ΔAIC$_i$ is defined as AIC$_i$ − min(AIC), the Relative AIC is defined as exp(−0·5* ΔAIC) and the Akaike weight is define as exp(−0·5* ΔAIC$_i$)/sum (exp(−0·5* ΔAIC$_i$)). The Akaike weight gives the probability that a model is the best one for the observed data, given the set of candidate models.

Table 2. Parameters estimates for the fives models

Model:	Reference model (without composition)	Composition as NMDS axes	Composition as random variable (based on experimental mixture)	Composition as random variable (based on PAM clustering)	Composition as correlation in the residual variance–covariance
Estimates of the parameters					
S:Treat *C*	**−0·193**	−0·079	**−0·193**	**−0·175**	−0·151
S:Treat PE	−0·036	0·028	−0·036	−0·026	−0·010
S:Treat PHE	−0·092	0·009	−0·092	−0·079	−0·038
Treat *C*	31·3	28·9	31·3	30·8	30·0
Treat PE	27·8	26·3	27·8	27·5	26·7
Treat PHE	29·7	27·3	29·7	29·5	27·7
λ	–	–	–	–	**0·80**
NMDS *X*	–	**1·45**	–	–	–
NMDS *Y*	–	−0·26	–	–	–
Relative importance of the parameters (difference in AIC between the model without the covariable of interest and the full model)					
S	7·04	−3·27	7·04	4·45	1·88
Treat	−1·68	−4·26	−1·68	−0·17	−1·06
S:Treat	0·41	−1·35	0·41	−0·18	2·56
NMDS	–	21·6	–	–	–
Random var.	–	–	−2·00	0·87	–
λ	–	–	–	–	26·1

S, plant species richness; Treat (*C*, PE, PHE), exclusion treatment (control, predator exclusion, predator and herbivore exclusion, respectively); λ, parameter indicating the strength of the correlation structure for the residual variance; *X* and *Y*, the coordinates of the subplots on the first and second NMDS axes, respectively; interaction between variables is indicated by a colon (:). Bold typeface indicates significant parameters at the 0·05 level (except for Treat levels, which are all significant as the models are run without intercept). Full results of the analyses can be found in Tables S3 to S7, Supporting information.
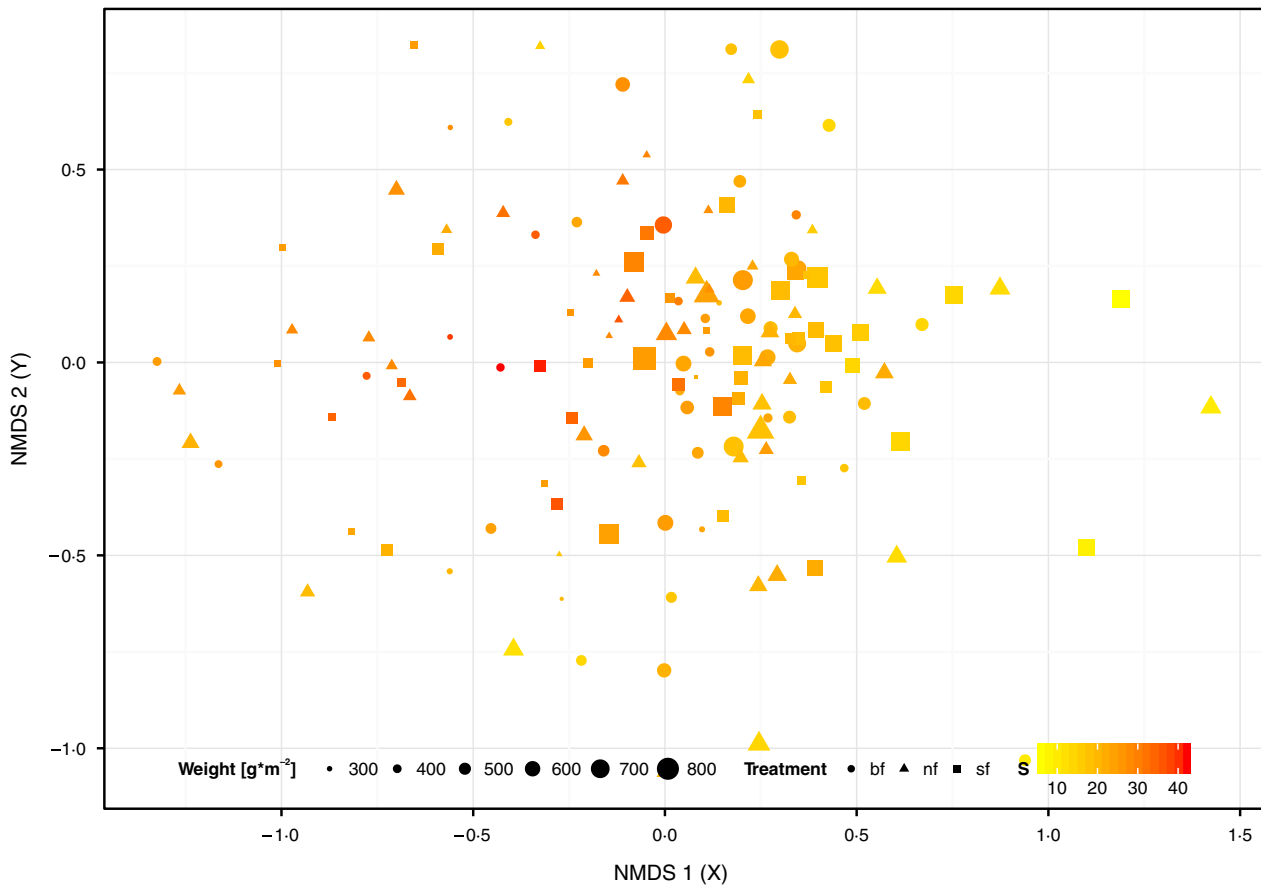
Fig. 1. Nonmetric multidimensional scaling ordination diagram of the subplots. Colour = plant species richness *S*; Symbol size = total biomass TB [g·m^{-2}]; and Symbols = exclusion treatment Treat.

ables *S* and Treat that are very similar to those in the reference model (Table 2). Adding composition in the model as a random effect produces no improvement in goodness-of-fit over the reference model. However, this apparent lack of performance can be explained by the split-plot design of our experiment that, when combined with the 37 levels for composition, leaves few degrees of freedom for this effect.

COMPOSITION AS A CATEGORICAL VARIABLE: LEVELS BASED ON FLORISTIC SIMILARITY

The results from the model with plant composition included as a random variable with three levels derived from PAM cluster analysis are again very similar to the previous results (Table 2). Taking composition into account, using this approach improves the goodness-of-fit in comparison with the reference model, but only moderately (ΔAIC = −0·88). The goodness-of-fit is superior to the model in which the categorical variable was based on sown species mixture, despite the fact that the composition term now has 12 times fewer levels.

COMPOSITION AS CORRELATION STRUCTURE

Including the similarity in floristic composition as correlation in the residual variance outperforms the other approaches,

with a ΔAIC of −26·13 compared to the reference model. The relationships between TB, *S* and Treat are again similar to the previous results (Table 2), but in this analysis, the relationship between total biomass and species richness within the control treatment is less significant. The estimated value of the parameter λ is 0·80, which indicates that the contribution of the floristic similarity between pairs of subplots is equal to 80% of the error variance in the subplots taken as independent observations.

In all, the way composition is included in a model does not change markedly the biological conclusions about the effect of species richness on biomass, but greatly affects goodness-of-fit, and of course the contribution due to plant composition. When the latter is high, as in the NMDS and the correlation models, the importance of species richness decreases. Another interesting result concerns the validation of the models by visual inspection of the residuals Q–Q plots (see Fig. S2 to S6, Supporting information): only the model with composition as correlation nicely fulfils the assumption of normality in the residuals.

Discussion

Including plant composition in models analysing the relationship between biodiversity and productivity generally yielded a

better fit. However, the way in which this factor is incorporated does matter. Clearly, the best fit is obtained with the use of a correlation matrix expressing similarity in plant composition within the residual variance. We also show that the 'classical' method of defining composition as a factor, with levels for each different plant mixture, can be improved simply by taking composition similarity into account. The approach using NMDS axes performs much better; our suggestion of adding higher-order terms does not improve the performance of the model in our case, but may be more useful in other situations.

For our data, the different methods do not alter the general relationship between total biomass TB and species richness S, exclusion treatment Treat and the S:Treat interaction. The general negative relationship between TB and S found in our system matches the results of earlier analyses where plots were not weeded (Pfisterer *et al.* 2004; Petermann *et al.* 2010); note that a negative relationship has also been found in natural systems (Schultz *et al.* 2011). In our case, the negative effect can be explained by a strong negative relationship between the number of sown species and the number of additional subordinate species; the latter with low cover contribute little to biomass production, but much to species richness. The exclusion treatment with fences (PE and PHE) shows weaker negative effects (Table 2); this is not due to a lower number of additional species in these blocks (14·8 and 14·0 respectively, compared with 14·7 in the unfenced C treatment), but possibly to the effect of higher trophic levels, and will be treated in future work.

The incorporation of plant composition as a similarity matrix affecting the residual variance has, to our knowledge, never before been used in BEF studies. This is an intuitive approach similar to the philosophy behind phylogenetic regression, where external information on the observations is used to account for their statistical dependence. It was by far the most powerful model in terms of goodness-of-fit. Treating the observations as nonindependent according to their floristic similarity thus yielded the best results and revealed a very strong effect of composition (see relative AIC in Table 2 for the parameter λ).

We chose Jaccard index measured on presence–absence data as a metric to represent floristic composition between the subplots. There is a palette of metrics to measure ecological resemblance, and their choice is dictated by the nature of the data and the context of the research (see Chapt. 7 of Legendre & Legendre 1998). Jaccard index is a classical metric that does not treat double absences as informative ('asymmetrical' indices *sensu* Legendre & Legendre 1998). Our data set includes quantitative information on species abundance (percentage cover), and including this information may be seen as desirable, especially in natural systems. In our case, computing the similarity matrix with Jaccard index on quantitative data yields a major improvement in term of AIC (AIC = 605·0 compared with 635·6 with presence–absence data; note that using percentage cover in all models does not change our conclusions). However, this approach may be criticised because of a possibility for circularity in the model: particular species may be abundant and have a strong influence on biomass, thus strongly influencing both the similarity and the response variable (total biomass). Nonetheless, it may be contended that two subplots harbouring such species will have a high similarity and thus their common contribution to the model will be downweighted, thus in effect reducing the potential for circularity. Here, we chose a conservative approach using Jaccard index on presence–absence data.

Until now, the incorporation of plant composition with ordination methods like NMDS has considered only linear effects of the ordination axes (Kahmen *et al.* 2005; Zavaleta *et al.* 2010; Schultz *et al.* 2011; Zuo *et al.* 2012). For our data, including higher-order terms does not improve the goodness-of-fit, but we would recommend that they be examined as part of standard analyses. There is indeed no reason not to consider the higher-order terms, because the coordinates on the ordination axes do not bear any meaning *per se*, but it is the pairwise distances between these coordinates that can be ecologically interpreted. Including higher-order terms, as in trend surface analyses to detect spatial trends, allows the identification of global nonlinear patterns in the data.

Despite the low support for this model based on Akaike weights compared with the correlation approach, we see an advantage in the use of ordination methods: they allow visualisation of the data as biplots, where both observations (subplots) and plant species (not shown in Fig. 1) are represented; information on biomass can be easily included, as shown in our case with colour coding. Such graphs may be very useful to explore the combination of species associated with high (or low) values of biomass.

The 'classical' approach of treating plant composition as a categorical factor with a level for each different mixture yielded in our case the worst results in terms of goodness-of-fit. This approach emerges logically from the design in controlled experimental settings, where the different sown mixtures are considered to be randomly drawn from the pool of possible mixtures. In our case, as we defined the levels based on sown rather than actual composition (which would have created as many levels as observations), the poor performance is not surprising. Also, our split-plot design leaves few degrees of freedom to detect an effect of composition. We show that applying a clustering to the plant cover data to create groups according to their species composition markedly improves the fit, despite the use of many fewer levels (3 rather than 37). The use of clustering is perfectly suited for observations in natural environments, representing an analogue to the 'classical' approach.

The manipulation of species richness cannot be performed without influencing plant composition, so it is essential to take the latter into account in analyses of the diversity–productivity relationship, as well as for other ecosystem functions. Indeed, our models including plant composition perform much better than models disregarding its effect. Due to the very different nature of the variables, it is not possible to treat species richness and species composition in an egalitarian way in linear modelling. This difficulty has motivated the formulation of diverse approaches. For example, the Price equation has been used to partition the effects of species richness, species composition and context dependence on ecosystem functioning (Fox

2006), with recent extensions accounting for species traits (Fox & Harpole 2008). Others have suggested representing biodiversity not by the mere number of species, but with measures accounting for species characteristics based on functional diversity (Petchey & Gaston 2002; Petchey, Hector & Gaston 2004) or phylogenetic diversity (Faith 1992; Cadotte, Cardinale & Oakley 2008). Recent developments by Kirwan *et al.* (2007) and Connolly *et al.* (2013) decompose in a sensible and intuitive fashion the contributions to ecosystem functioning that are due to species individually and in their pairwise interactions. Another methodological breakthrough is the method of Bell *et al.* (2009), where an experimental design is proposed to partition the variance between species identities and species richness (Harvey *et al.* 2013). In our case, the presence of sub-plots with high species richness, including additional subordinate species, precludes the use of both latter methods. Our approach is flexible and can accommodate such deviation from strict experimental settings where species richness is maintained by weeding.

Ultimately, the choice of a particular technique should be dictated by the objectives and design of the research. For studies in natural systems or in experiments without weeding, we have shown that existing methods using species richness as explanatory variable can easily be improved, by considering higher-order terms for NMDS axes or using clustering methods when treating composition as a factor. However, our new approach with floristic correlation in the residual variance outcompetes the other models, and we believe that it represents a powerful addition to existing techniques. Also, it can be easily tailored for other contexts, as correlation between observations is not limited to floristic similarity (e.g. Fabian *et al.* 2013). It should find its place in future analyses of biodiversity effects on ecosystem functioning.

Acknowledgements

Data accessibility

The data are freely available at https://datadryad.org/resource/doi:10.5061/dryad.d2v35 (Sandau *et al.* 2014).

Authorship

NS, PK and LFB planned the experiment, NS, AA, YF and OTB conducted the experiment, NS, RPR and LFB conducted the analyses; NS, REN and LFB wrote the article.

References

Assaf, T., Beyschlag, W. & Isselstein, J. (2011) The relationship between plant diversity and productivity in natural and Managed Grassland. *Applied Ecology and Environmental Research*, **9**, 157–166.

Bell, T., Lilley, A.K., Hector, A., Schmid, B., King, L. & Newman, J.A. (2009) A linear model method for biodiversity-ecosystem functioning experiments. *American Naturalist*, **174**, 836–849.

Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.

Bruggisser, O.T., Sandau, N., Blandenier, G., Fabian, Y., Kehrli, P., Aebi, A., Naisbit, R.E. & Bersier, L.F. (2012) Direct and indirect bottom-up and top-down forces shape the abundance of the orb-web spider *Argiope bruennichi*. *Basic and Applied Ecology*, **13**, 706–714.

Cadotte, M.W., Cardinale, B.J. & Oakley, T.H. (2008) Evolutionary history and the effect of biodiversity on plant productivity. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 17012–17017.

Caldeira, M.C., Hector, A., Loreau, M. & Pereira, J.S. (2005) Species richness, temporal variability and resistance of biomass production in a Mediterranean grassland. *Oikos*, **110**, 115–123.

Chapin, F.S., Walker, B.H., Hobbs, R.J., Hooper, D.U., Lawton, J.H., Sala, O.E. & Tilman, D. (1997) Biotic control over the functioning of ecosystems. *Science*, **277**, 500–504.

Connolly, J., Bell, T., Bolger, T., Brophy, C., Carnus, T., Finn, J.A. *et al.* (2013) An improved model to predict the effects of changing biodiversity levels on ecosystem function. *Journal of Ecology*, **101**, 344–355.

Davison, A.C. (2003) *Statistical Models*. Cambridge University Press, Cambridge, UK.

Dimitrakopoulos, P., Siamantziouras, A.-S., Galanidis, A., Mprezetou, I. & Troumbis, A.Y. (2006) The interactive effects of fire and diversity on short-term responses of ecosystem processes in experimental Mediterranean Grasslands. *Environmental Management*, **37**, 826–839.

Fabian, Y., Sandau, N., Bruggisser, O.T., Kehrli, P., Aebi, A., Rohr, R.P., Naisbit, R.E. & Bersier, L.F. (2012) Diversity protects plant communities against generalist molluscan herbivores. *Ecology and Evolution*, **2**, 2460–2473.

Fabian, Y., Sandau, N., Bruggisser, O.T., Aebi, A., Kehrli, P., Rohr, R.P., Naisbit, R.E. & Bersier, L.F. (2013) The importance of landscape and spatial structure for hymenopteran-based food webs in an agro-ecosystem. *Journal of Animal Ecology*, **82**, 1203–1214.

Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**, 1–10.

Fox, J.W. (2006) Using the price equation to partition the effects of biodiversity loss on ecosystem function. *Ecology*, **87**, 2687–2696.

Fox, J.W. & Harpole, W.S. (2008) Revealing how species loss affects ecosystem function: the trait-based price equation partition. *Ecology*, **89**, 269–279.

Fox, J. & Weisberg, S. (2011) *An R Companion to Applied Regression*, 2nd edn. Sage, Thousand Oaks, CA.

Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002) Phylogenetic analysis and comparative data: a test and review of evidence. *American Naturalist*, **160**, 712–726.

Fridley, J.D. (2003) Diversity effects on production in different light and fertility environments: an experiment with communities of annual plants. *Journal of Ecology*, **91**, 396–406.

Fukami, T., Naeem, S. & Wardle, D.A. (2001) On similarity among local communities in biodiversity experiments. *Oikos*, **95**, 340–348.

Gamfeldt, L., Hillebrand, H. & Jonsson, P.R. (2008) Multiple functions increase the importance of biodiversity for overall ecosystem functioning. *Ecology*, **89**, 1223–1231.

Grace, J.B., Anderson, T.M., Smith, M.D., Seabloom, E., Andelman, S.J., Meche, G. *et al.* (2007) Does species diversity limit productivity in natural grassland communities? *Ecology Letters*, **10**, 680–689.

Harvey, E., Séguin, A., Nozais, C., Archambault, P. & Gravel, D. (2013) Identity effects dominate the impacts of multiple species extinctions on the functioning of complex food webs. *Ecology*, **94**, 169–179.

Hector, A., Schmid, B., Beierkuhnlein, C., Caldeira, M.C., Diemer, M., Dimitrakopoulos, P.G. *et al.* (1999) Plant diversity and productivity experiments in European Grasslands. *Science*, **286**, 1123–1127.

Hector, A., Bell, T., Hautier, Y., Isbell, F., Kéry, M., Reich, P.B., van Ruijven, J. & Schmid, B. (2011) BUGS in the analysis of biodiversity experiments: species richness and composition are of similar importance for grassland productivity. *PLoS ONE*, **6**, e17434.

Hooper, D.U., Adair, E.C., Cardinale, B.J., Byrnes, J.E.K., Hungate, B.A., Matulich, K.L. *et al.* (2012) A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature*, **486**, 105–108.

Houseman, E.A., Ryan, L.M. & Coull, B.A. (2004) Cholesky residuals for assessing normal errors in a linear model with correlated outcomes. *Journal of the American Statistical Association*, **99**, 383–394.

Huston, M.A. (1997) Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. *Oecologia*, **110**, 449–460.

Kahmen, A., Perner, J., Audorff, V., Weisser, W. & Buchmann, N. (2005) Effects of plant diversity, community composition and environmental parameters on productivity in montane European grasslands. *Oecologia*, **142**, 606–615.

Kaiser, J. (2000) Rift over biodiversity divides ecologists. *Science*, **289**, 1282–1283.

Kirwan, L., Lüscher, A., Sebastià, M.T., Finn, J.A., Collins, R.P., Porqueddu, C. *et al.* (2007) Evenness drives consistent diversity effects in intensive grassland systems across 28 European Sites. *Journal of Ecology*, **95**, 530–539.

Lawton, J.H. (1994) What do species do in ecosystems? *Oikos*, **71**, 367–374.

Legendre, P. & Legendre, L. (1998) *Numerical Ecology*. Elsevier Science B. V, Amsterdam.

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M. & Hornik, K. (2012). cluster: Cluster Analysis Basics and Extensions. URL http://cranr-projectorg/web/packages/cluster/.

Maestre, F.T., Castillo-Monroy, A.P., Bowker, M.A. & Ochoa-Hueso, R. (2012) Species richness effects on ecosystem multifunctionality depend on evenness, composition and spatial pattern. *Journal of Ecology*, **100**, 317–330.

Minchin, P.R. (1987) An evaluation of the relative robustness of techniques for ecological ordination. *Plant Ecology*, **69**, 89–107.

Mouillot, D., Villéger, S., Scherer-Lorenzen, M. & Mason, N.W.H. (2011) Functional structure of biological communities predicts ecosystem multifunctionality. *PLoS ONE*, **6**, e17476.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B. *et al.* (2013) vegan: Community Ecology Package. URL http://CRAN.R-project.org/package=vegan.

Petchey, O. & Gaston, K. (2002) Functional diversity (FD), species richness and community composition. *Ecology Letters*, **5**, 402–411.

Petchey, O., Hector, A. & Gaston, K. (2004) How do different measures of functional diversity perform? *Ecology*, **85**, 847–857.

Petermann, J.S., Fergus, A.J.F., Roscher, C., Turnbull, L.A., Weigelt, A. & Schmid, B. (2010) Biology, chance, or history? The predictable reassembly of temperate grassland communities. *Ecology*, **91**, 408–421.

Pfisterer, A.B., Joshi, J., Schmid, B. & Fischer, M. (2004) Rapid decay of diversity-productivity relationships after invasion of experimental plant communities. *Basic and Applied Ecology*, **5**, 5–14.

Pinheiro, J. & Bates, D. (2010) nlme: Linear and Nonlinear Mixed Effects Models R package version 3.1-103. URL http://cran.r-project.org/web/packages/nlme/index.html.

R Development Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Roscher, C., Schumacher, J., Baade, J., Wilcke, W., Gleixner, G., Weisser, W.W., Schmid, B. & Schulze, E.D. (2004) The role of biodiversity for element cycling and trophic interactions: an experimental approach in a grassland community. *Basic and Applied Ecology*, **5**, 107–121.

Roscher, C., Temperton, V.M., Scherer-Lorenzen, M., Schmitz, M., Schumacher, J., Schmid, B., Buchmann, N., Weisser, W.W. & Schulze, E.D. (2005) Overyielding in experimental grassland communities - irrespective of species pool or spatial scale. *Ecology Letters*, **8**, 419–429.

Sandau, N., Rohr, R.P., Naisbit, R.E., Fabian, Y., Bruggisser, O.T., Kehrli, P. *et al.* (2014) Data from: Including community composition in biodiversity–productivity models. *Dryad Digital Repository*, https://datadryad.org/resource/doi:10.5061/dryad.d2v35.

Schmid, B., Hector, A., Huston, M.A., Inchausti, P., Nijs, I., Leadley, P.W. & Timan, D. (2002) The design and analysis of biodiversity experiments. *Biodi-versity and Ecosystem Functioning* (eds M. Loreau, S. Naeem & P. Inchausti), pp. 61–78. Oxford University Press, Oxford.

Schultz, R., Andrews, S., O'Reilly, L., Bouchard, V. & Frey, S. (2011) Plant community composition more predictive than diversity of carbon cycling in freshwater wetlands. *Wetlands*, **31**, 965–977.

Tilman, D., Wedin, D. & Knops, J. (1996) Productivity and sustainability influenced by biodiversity in grassland ecosystems. *Nature*, **379**, 718–720.

Troumbis, A.Y., Dimitrakopoulos, P.G., Siamantziouras, A.-S.D. & Memtsas, D. (2000) Hidden diversity and productivity patterns in mixed Mediterranean grasslands. *Oikos*, **90**, 549–559.

Van Ruijven, J. & Berendse, F. (2003) Positive effects of plant species diversity on productivity in the absence of legumes. *Ecology Letters*, **6**, 170–175.

Wantzen, K.M., Rothhaupt, K.-O., Mörtl, M., Cantonati, M., Tóth, L.G. & Fischer, P. (2008) Ecological effects of water-level fluctuations in lakes: an urgent issue. *Hydrobiologia*, **204**, 1–4.

Zavaleta, E.S., Pasari, J.R., Hulvey, K.B. & Tilman, G.D. (2010) Sustaining multiple ecosystem functions in grassland communities requires higher biodiversity. *Proceedings of the National Academy of Sciences of the United States of America*, **107**, 1443–1446.

Zuo, X.A., Knops, J.M.H., Zhao, X.Y., Zhao, H.L., Zhang, T.H., Li, Y.Q. & Guo, Y.R. (2012) Indirect drivers of plant diversity-productivity relationship in semiarid sandy grasslands. *Biogeosciences*, **9**, 1277–1289.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A. & Smith, G.M. (2009) *Mixed Effects Models and Extensions in Ecology With R*, 1st edn. Springer, New York, NY.

Supporting Information

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

Appendix S1. Description of field methods.

Appendix S2. R code for the five models.

Appendix S3. R routine to include correlation structure in the mixed effects model.

Table S1. List of sown plant species.

Table S2. Descriptive statistics for vegetation.

Table S3–S7. Detailed results for the five models.

Fig. S1. Silhouette profile of PAM clustering.

Fig. S2–S6. Q–Q plots for the five models.