Article type : Statistical Reports

Running title: Interactions in multi-species ecosystems

Biodiversity and ecosystem function: Making sense of numerous species interactions in multi-species communities

CAROLINE BROPHY[1,8], ÁINE DOOLEY[1], LAURA KIRWAN[2], JOHN A. FINN[3], JACK MCDONNELL[1,4], THOMAS BELL[5], MARC W. CADOTTE[6] AND JOHN CONNOLLY[7]

[1]*Department of Mathematics and Statistics, Maynooth University, Maynooth, Co. Kildare, Ireland*

[2] *UCD School of Agriculture and Food Science, University College Dublin, Belfield, Dublin 4, Ireland*

[3]*Teagasc Environment Research Centre, Johnstown Castle, Co. Wexford, Ireland*

[4]*Animal and Grassland Research and Innovation Centre, Teagasc, Fermoy, Co Cork, Ireland*

[5]*Department of Life Sciences, Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot, Berkshire, SL5 7PY, UK*

[6]*Department of Biological Sciences, University of Toronto–Scarborough, 1265 Military Trail, Toronto, Ontario M1C 1A4, Canada*

[7]*School of Mathematics and Statistics, Ecological and Environmental Modelling Group, University College Dublin, Dublin 4, Ireland*

[8]*Email: caroline.brophy@nuim.ie*

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ecy.1872

Abstract. Understanding the biodiversity and ecosystem function relationship can be challenging in species-rich ecosystems. Traditionally, species richness has been relied on heavily to explain changes in ecosystem function across diversity gradients. Diversity-Interactions models can test how ecosystem function is affected by species identity, species interactions and evenness, in addition to richness. However, in a species-rich system, there may be too many species interactions to allow estimation of each coefficient, and if all interaction coefficients are estimable, they may be devoid of any sensible biological meaning. Parsimonious descriptions using constraints among interaction coefficients have been developed but important variability may still remain unexplained. Here, we extend Diversity-Interactions models to describe the effects of diversity on ecosystem function using a combination of fixed coefficients and random effects. Our approach provides improved standard errors for testing fixed coefficients and incorporates lack-of-fit tests for diversity effects. We illustrate our methods using data from a grassland and a microbial experiment. Our framework considerably reduces the complexities associated with understanding how species interactions contribute to ecosystem function in species-rich ecosystems.

Key words: biodiversity and ecosystem function relationship; community structure; evenness; Diversity-Interactions model; mixed model; variability; random diversity effects; random effects; richness; species interactions; species rich; variance components.

INTRODUCTION

Widespread study of the biodiversity and ecosystem function (BEF) relationship has led to broad consensus that increasing the biodiversity of a system improves its ability to maintain and/or increase functionality (Bell et al. 2005, Hooper et al. 2005, Duffy 2009, Finn et al. 2013). The benefits of biodiversity to ecosystem function are frequently quantified using species richness (e.g. Spehn et

al. 2005) but modelling interactions among species and/or evenness, in addition to richness, can lead to enhanced understanding of diversity driven improvements to ecosystem function (Wilsey and Polley 2004, Connolly et al. 2011, Finn et al. 2013). However, modelling species interactions becomes increasingly difficult as species richness, and hence the number of interactions, increases.

The Diversity-Interactions (DI) (Kirwan et al. 2007, Kirwan et al. 2009, Dooley et al. 2015) and Generalised Diversity-Interactions (GDI) (Connolly et al. 2013) modelling approaches estimate the combined contributions of species-specific and pairwise species interaction effects to total ecosystem functioning. These models have successfully assessed the impact of variables that determine community structure such as species identity, species initial proportions, species interactions, species richness and evenness on ecosystem function. Any two species may interact to affect ecosystem function in a positive, negative or neutral way and the combined effect of all interactions in a multi-species community is the 'diversity effect'. However, the 'full' pairwise interaction DI or GDI model requires a coefficient for every possible pairwise interaction in the system and when there is a large species pool there may be too many coefficients to estimate (due to study design), or if coefficients are estimable, they become uninformative or difficult to interpret due to their large number. Combining a small number of fixed coefficients (biologically motivated where possible, Kirwan et al. 2009) with random effects, to capture remaining variability in ecosystem function due to species interactions, could provide a more parsimonious and biologically informative description of species interaction effects than estimating all individual pairwise interactions as fixed coefficients.

The relationship between ecosystem function and diversity can assume many forms, generally an increasing response to diversity (often measured as richness) that may saturate at higher diversity levels (Tilman et al. 1997), as in Figure 1. Variation in ecosystem function among communities across levels of richness may be somewhat constant (Fig. 1a) or may vary (Fig. 1b); this variation is likely caused by factors such as species identities, community composition, species

relative abundances, species interactions or evenness (Connolly et al. 2013). A process of random selection of species is commonly used to determine the composition of communities at each level of richness in designed factorial experiments and the associated variability in ecosystem function can be measured as a variance component in BEF models (Schmid et al. 2002). However, with DI and GDI models, differences in communities at a given level of richness can be attributed to the identity of the species, how species interact and the evenness of the community rather than modelling it as a single 'random selection' variance component.

In this paper, we extend the DI and GDI models (Kirwan et al. 2009, Connolly et al. 2013) to a mixed modelling framework to develop a parsimonious solution to modelling diversity effects in species-rich ecosystems that will sufficiently explain the variability in ecosystem function due to numerous species interactions. We model species interactions using a small number of fixed coefficients combined with random effects to capture remaining differences among pairwise species interactions. This approach provides a lack-of-fit test for the fixed component of the diversity effect and improves model inference by using an appropriate variance structure. We illustrate our modelling framework using data from two experiments, one grassland ("Jena" with nine species) and one microbial ("Bell" with 72 species); these data sets lead to a high number of pairwise species interactions, 36 and 2556 respectively. Estimating a unique coefficient for each pairwise interaction in the case of the Jena data is possible but 36 coefficients is a high number to interpret and likely to be biologically uninformative as a result. Estimating a unique coefficient for each interaction with the Bell data is not possible since there are more interactions than data points, nor is it desirable since 2556 fixed coefficient estimates would be devoid of useful biological information. A motivating question for our work is: How can the effects on ecosystem function of the numerous species interactions in these experiments be captured using a small number of coefficients without missing out on important variability?

MATERIALS AND METHODS

Example data sets

We introduce the two illustrative data sets: the "Jena" and the "Bell" data sets. The Jena data set was from one year of a nine-species grassland experiment in Jena, Germany (Roscher et al. 2004, Roscher et al. 2005). There were 206 communities assembled with various levels of species richness (1, 2, 3, 4, 6 or 9 species) across four blocks. Each pair of species appeared together in exactly 30 communities. The species were classified into three functional groups (grasses, non-legume herbs and legumes), and dry aboveground biomass in 2003 (the year after establishment) was measured. The Bell data set was from a 72-species microbial experiment (Bell et al. 2005). There were 1,374 microcosm communities inoculated with species of bacteria across varying richness levels (1, 2, 3, 4, 6, 8, 9, 12, 18, 24, 36 and 72 species). Each pair of species appeared together in 26 communities on average. The average daily respiration rate (over a period of 28 days) of the bacterial community was recorded. Additional information on both experiments is in Appendix S1.

Model descriptions

Diversity-Interactions (DI) models (Kirwan et al. 2007, Kirwan et al. 2009) can be expressed in the general form of

$$y = \mathrm{ID} + \mathrm{DE}_{\mathrm{fixed}} + \varepsilon \quad (1)$$

where ID stands for 'identity effects' and quantifies expected species monoculture behaviour, and includes treatment or block effects; $\mathrm{DE}_{\mathrm{fixed}}$ stands for 'diversity effect' and is comprised of a number of fixed coefficients representing interactions among species. For example:

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \sum_{\substack{i,j=1 \\ i<j}}^{s} \delta_{ij} P_i P_j + \varepsilon \quad (2)$$

where the first two terms comprise the ID component, the third term comprises the DE_{fixed} component and $\varepsilon \sim N(0, \sigma_1^2)$. The ecosystem function is *y*, P_i is the initial relative abundance of the *i*th species in a pool of *s* species (*i*=1,...,*s*), and *A* can include a block and/or a treatment factor. The coefficient β_i is the expected performance of species *i* in monoculture and δ_{ij} measures the potential interactive effect of species *i* with species *j* (for *i*,*j*=1,…,*s* and *i*<*j*) on the ecosystem function (*y*). The diversity effect ($\Sigma\ \delta_{ij}\ P_iP_j$ in equation (2)) is the difference between the expected mixture response based solely on the species monoculture responses ($\Sigma\beta_iP_i+\alpha A$) and that including mixing effects. An additional coefficient, θ_1, that enters the model as a power to P_iP_j can be included to allow for non-linearity in the interaction terms; this model is known as the Generalised Diversity-Interactions (GDI) model (Connolly et al. 2013). The DE_{fixed} component can take many forms; in equation (2), the 'full' pairwise interaction Diversity-Interactions model is specified and the diversity effect requires estimating $s(s-1)/2$ δ_{ij} coefficients, which is, for example, six coefficients in a four-species system but 190 coefficients in a 20-species system. Kirwan et al. (2009) provide several alternatives for the DE_{fixed} component including the average pairwise model where interactions among species are all assumed to be equal, the functional group model where interactions among species are dictated by functional group membership, and the additive species model where each species contributes a constant additive amount to its interaction with any another species.

Here, we propose using a description of the diversity effect with a small number of fixed coefficients, and augmenting it with random effects to fully capture the variability in ecosystem function due to species interactions. We extend equation (1) to include random effects to measure the additional variability due to pairwise species interactions:

$$y = ID + DE_{fixed} + DE_{random} + \varepsilon \qquad (3)$$

In this model, DE_{fixed} contains fewer coefficients than the full model which has an interaction coefficient for each pair of species (equation (2)). For example, it might be assumed that all interaction coefficients are equal, leading to the realisation of equation (3):

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} P_i P_j + \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij} P_i P_j + \varepsilon \quad (4)$$

where $d_{ij} \sim N(0, \sigma_2^2)$ and $\varepsilon \sim N(0, \sigma_1^2)$. The DE_{fixed} component here assumes that all δ_{ij} coefficients are equal to δ_{av}, while the DE_{random} component recognises that there may be variability in the true δ_{ij} around δ_{av} and σ_2^2 measures this variability. The purpose of the DE_{fixed} component is to elicit as much information as possible in as low a number of fixed coefficients as possible. The DE_{random} component is constructed by adding each P_iP_j term to the model as a random effect and constraining them to have the same variability and with zero covariances. This necessitates the inclusion of a large number of random effects, one for each pair of interactions; for our illustrative data sets this is 36 and 2556 random effects as they had 9 and 72 species respectively, but only one additional model parameter (σ_2^2) is required. The inclusion of these many random effects with a common variance is quite different to the typical use of random effects (which are usually indexed by a community level factor, e.g. block) but is a statistical technique to allow estimation of the variability across the δ_{ij} coefficients. Both variance parameters (σ_1^2 and σ_2^2) will feed into the standard errors for all the fixed coefficients in the model. Including a power coefficient, θ_2, on the P_iP_j's in the DE_{random} component, allows for flexibility in how the marginal variance of the response (y) varies across community structures.

Model estimation and comparisons

We used least squares to fit models with no random effects, profile likelihood to estimate power coefficients (θ_1 and θ_2), and restricted maximum likelihood (REML) to fit models with random effects; the software package SAS version 9.3 (SAS Institute Inc) was used to fit all models.

We tested various forms of the DE_{fixed} component using F-tests, or likelihood ratio tests (LRT) for comparisons involving the non-linear power coefficient on interaction terms. We tested inclusion of the DE_{random} component using LRT. One argument against the use of LRT to test random effects is that it is overly conservative on account of variances being bounded below by 0 (Self and Liang 1987, Stram and Lee 1994, Mc Culloch and Searle 2001). To overcome this we recommend divided the LRT *P*-value by 2 (Littell et al. 2006, pages 752-3). There are two possible outcomes for the likelihood ratio test of DE_{random}:

1. The test is significant, showing that the random effects (DE_{random}) are needed. This indicates that the DE_{fixed} component isn't sufficient to explain the variability in *y* caused by species interactions. The random effects will account for this additional variability and the variance term (σ_2^2) will be incorporated into the standard errors for fixed coefficients providing more reliable tests for them. The random effects can be estimated using empirical best linear unbiased predictors (eBLUPs) and explored for further information.
2. Alternatively, the test is not significant, indicating no need to include the random effects. The DE_{random} component should be omitted and it can be assumed that the DE_{fixed} component is sufficiently explaining variability attributed to species interactions. This outcome shows no evidence of lack of fit in the DE_{fixed} component, validating the results and inference provided by it.

Thus, regardless of the outcome of the test for any specific data set, testing the DE_{random} component plays an important role in the analysis.

The residual error term in equation (3) is assumed to have a constant variance (σ_1^2) but it may be affected by community structure. Using LRT, we tested whether or not the residual variance differs for monocultures and mixtures by assuming ε to be normally distributed with mean 0 and with variance σ_{1a}^2 for monocultures and σ_{1b}^2 for mixtures; under the null hypothesis $\sigma_{1a}^2 = \sigma_{1b}^2 =$

σ_1^2. Allowing the residual variance to differ depending on community structure could be further explored, for example, by allowing the residual error term to vary by richness or evenness. Splitting of the residual variance in this way can be included in any BEF model (e.g. the presence / absence method of Bell et al. 2009), not just one, such as ours, that has random pairwise interactions built in.

Application of our approach to the two data sets

We used the following procedures to select our model for each data set:

1. We selected a 'baseline' model which involved exploring a range of options for the DE_{fixed} component (e.g. the functional group or additive species models, see Appendix S2 for the full list of models tested), each of which contained a relatively small number of fixed coefficients (but no random effects were included at this stage).
2. The baseline model was extended to test inclusion of the DE_{random} component and the residual error variance was tested for a difference among monocultures and mixtures.
3. Significance of each DE_{fixed} coefficient was re-evaluated using the new variance structure (if applicable).

RESULTS

The baseline model selected for the Jena grassland data set included interaction terms that were functional group specific (Appendix S2: Table S1, model 3). In the ID component of this model was an identity coefficient for each species, and block effects; the DE_{fixed} component included six interaction coefficients: three 'within functional group' coefficients where any pair of species from a functional group were assumed to interact in the same way and three 'between functional group' coefficients where any species from one functional group was assumed to interact in the same way

with any species from the second functional group. It was possible to fit the full pairwise interaction models (i.e. 36 δ_{ij} coefficients) via fixed coefficients (equation 2). This model was a better fit than the functional group model (Appendix S2: Table S1, *P*=0.012), but in practice, 36 fixed coefficients is a high number to elicit useful biological meaning from, and with many other data sets it will not be possible to fit the full pairwise interactions model.

Extending the baseline model to include the random interaction terms was significant (Table 1a, model J1 *vs*. J2, *P* = 0.008). This means that using the six functional group interaction coefficients was not sufficient to explain the variability caused by all 36 pairwise species interactions and fitting the d_{ij} random terms bridged the gap between the six fitted coefficients and the possible 36 fixed coefficients with a single variance component. A power coefficient on the P_iP_j in the DE_{random} component was estimated but did not improve the model fit further (tested for a difference from 1 using a likelihood ratio test, *P* = 0.237) and was kept at 1. Fitting different residual error variances to monocultures and mixtures did not improve the model fit (Table 1a, J2 *vs*. J3). Thus, the final selected model for the Jena data set was one that included within- and between-functional group interactions and had random effects for pairwise interactions:

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \mathrm{DE}_{\mathrm{fixed}} + \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij}\left(P_i P_j\right) + \varepsilon \qquad (5)$$

where $d_{ij} \sim N(0, \sigma_2^2)$ and $\varepsilon \sim N(0, \sigma_1^2)$; model J2 in Appendix S3 gives the full specification. The variance estimates were $\hat{\sigma}_1^2$ =15,311 and $\hat{\sigma}_2^2$ =90,101, all other coefficient estimates are in Appendix S4, Table S1. Interaction between a grass and herb was the highest among the six estimated interaction coefficients (Fig. 2a). We estimated the 36 random effects (d_{ij}'s) for the Jena data and added each to the corresponding functional group estimate (Fig. 2a); variability among individual interaction estimates within each group was highest for the grass-grass and grass-herb interaction groups. See Appendix S4 for model diagnostics using the estimated random effects.

There are no 'functional groupings' with the Bell microbial data, as there are with the Jena data, therefore the baseline model fitting included testing of the average pairwise species model and the additive species model. The Generalised Diversity-Interaction model, with an average pairwise interaction effect and power coefficient on the interactions, was selected as the baseline model (Appendix S2: Table S2). The fit of the baseline model was not improved by including the DE_{random} component or by splitting the residual variance by monoculture and mixture (Table 1b). Thus, the final model for the Bell data included the average interaction fixed effect with θ_1 power coefficient and no random interaction effects:

$$y = \sum_{i=1}^{S} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{S} (P_i P_j)^{\theta_1} + \varepsilon$$

where $\varepsilon \sim N(0, \sigma_1^2)$, and $\hat{\sigma}_1^2$ =7.55, $\hat{\delta}_{av}$ = 2.12 and $\hat{\theta}_1$ =0.79. A histogram of the estimated identity effect (β_i) coefficients is shown in Appendix S5. It was not possible to fit the full pairwise interactions model here (that would require the estimation of 2556 coefficients for which there is not enough data); our result is therefore quite powerful. First, it shows that two coefficients δ_{av}, and ϑ_1 were sufficient to define the diversity effect in the average BEF relationship (i.e. there was no evidence of lack of fit in the two coefficient explanation of diversity effects). Second, variation of community responses around this relationship was determined solely by the residual error variation, the contribution of the variation of individual pairwise interaction terms was negligible.

Model predictions for each data set and the raw data are shown in Figure 2 (b and c). The Jena data set is given in Data S1 and is also available in Connolly et al. (2011), and SAS code to fit the models in Table 1a is provided in Data S1. The models can also be fitted using the ASREML-R package in R (code in Data S1) this package is only available by trial or by purchase from VSN International (Butler et al. 2009). In Appendix S6, we compare and discuss the two software options used for fitting our method.

DISCUSSION

The main aim of our framework was to model the effects of multiple species' interactions on ecosystem function using only a small number of coefficients. We achieved this through eliciting as much information as possible on species interactions via a small number of fixed coefficients, and supplementing this with random effects to explain any further variability in ecosystem function attributed to species interactions. Our extended Diversity-Interactions modelling framework is particularly useful for species-rich ecosystems which can be complex, with potentially a large number of interactions affecting ecosystem function. A major benefit of including random interaction effects is the ability to test for lack of fit in the fixed effect coefficients of the diversity effect. The inclusion of random interaction effects also feeds into the estimation of standard errors of the model fixed coefficients, thus improving inference. Our modelling approach is suited to data from ecosystems that generate more interactions than can be estimated or that permit sensible interpretation; this may be a five-species pool or higher.

The concept of random pairwise interactions is grounded in both statistical and biological motivations. From a practical perspective, it may not be possible to fit the large number of pairwise interactions as fixed terms in a species-rich system and therefore using a small number of fixed coefficients combined with random effects is a statistical convenience that bridges any gap in unexplained variability between a 'reduced' model compared to the elusive 'full' model with all pairwise interactions fitted as fixed. However, it is not just a statistical convenience; even if the study design permits estimation of all fixed pairwise interactions, it is unlikely that they will be biologically informative due to the their large number as was the case with the Jena data set which had 36 species interactions. Also, the extra random variation will automatically be built into the standard errors of predicted mean responses from the model and so will reflect the extra uncertainty in prediction due to the extra variation. If the functional group model was fitted to the Jena data without the random effects, the standard errors on the model estimates would be incorrect since

important pairwise species variability would have been omitted. For the Jena data, it is evident that a large portion of the variability in biomass is due to how species within and between functional groups interact; for example, there is almost no overlap in the estimated individual pairwise species interactions for grass-grass compared with grass-herb interactions (Fig. 2a). This means that information from the functional group interaction estimates, combined with the identity effect estimates (Appendix S4, Table S1), can influence management practices aimed at maximising yield, whilst the model still acknowledges that there is variability caused by species interactions in addition to the functional group explanation.

For the Bell data, it was surprising that only two coefficients were needed to describe pairwise interactions given that 72 species were investigated. The non-significant random effects test is powerful as it validates the inference from the parsimonious description of the diversity effects: there was no additional variability in species interactions beyond the estimated average effect. When the random effects are significant, we recommend estimation and examination of the random effects; doing so allows assessment of the relative importance of the DE_{fixed} and DE_{random} components (e.g. Fig. 2a). If exploration of the random effects indicates that variability in the DE_{random} component is considerably more important than that explained by the DE_{fixed} component, practical information derived from the fixed diversity effect coefficients is less influential. However, clustering techniques applied to the estimated random effects could identify patterns among the interactions that may inform future biological hypotheses; this may be particularly relevant when *a priori* species groupings are not available.

The identity effects in our model could be assumed to follow a random distribution (Lipowsky et al. 2015) with the fixed identity effects (β_i in equations 2 and 4) constrained in a biologically sensible manner. In a species rich system, many degrees of freedom are used in estimating the species identity effects and a benefit of introducing random identity effects is to reduce the number of coefficients. A drawback however, is that predictive ability of the model is

reduced as the estimates of species behaviour in monoculture feed into predictions of the behaviour of a mixture. As such, we recommend fitting all identity effects as fixed coefficients when possible, but if study design limits this or if multiple experiments are being analysed together, introducing random identity effects may help improve model parsimony.

Our framework offers a modelling approach that is parsimonious and informative. The method has the ability to greatly reduce the number of coefficients required to model the effects of species' interactions on ecosystem functioning, thereby simplifying the description of diversity effects without ignoring potentially important ecosystem function variability. This is an improvement on current DI models (Kirwan et al. 2009, Connolly et al. 2013), particularly pertinent for species-rich systems, and still retains all the benefits of Diversity-Interactions models. These benefits include understanding how species interact (Kirwan et al. 2009, Connolly et al. 2013), the ability to predict ecosystem function(s) for any set of species at any relative abundances across richness and evenness gradients and the ability to identify combinations of species that lead to a strong (or weak) performance of ecosystem function(s) (Dooley et al. 2015), thus increasing our knowledge of complex biodiversity and ecosystem function relationships.

ACKNOWLEDGMENTS

This work was supported by the Science Foundation Ireland, Grant Number 09/RFP/EOB2546.

LITERATURE CITED

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

Bell, T., J. A. Newman, B. W. Silverman, S. L. Turner, and A. K. Lilley. 2005. The contribution of species

richness and composition to bacterial services. Nature 436:1157-1160.

Butler, D. 2009. asreml: asreml() fits the linear mixed model. R package version 3.0. www.vsni.co.uk.

Connolly, J., et al. 2013. An improved model to predict the effects of changing biodiversity levels on ecosystem function. Journal of Ecology 101:344-355.

Connolly, J., M. W. Cadotte, C. Brophy, Á. Dooley, J. Finn, L. Kirwan, C. Roscher, and A. Weigelt. 2011. Phylogenetically diverse grasslands are associated with pairwise interspecific processes that increase biomass. Ecology 92:1385-1392.

Dooley, Á., F. Isbell, L. Kirwan, J. Connolly, J. A. Finn, and C. Brophy. 2015. Testing the effects of diversity on ecosystem multifunctionality using a multivariate model. Ecology Letters 18:1242-1251.

Duffy, J. E. 2009. Why biodiversity is important to functioning of real-world ecosystems. Frontiers in Ecology and the Environment 7:437-444.

Finn, J. A., et al. 2013. Ecosystem function enhanced by combining four functional types of plant species in intensively managed grassland mixtures: a 3-year continental-scale field experiment. Journal of Applied Ecology 50:365–375.

Hooper, D. U., et al. 2005. Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. Ecological Monographs 75:3-35.

Kirwan, L., J. Connolly, J. A. Finn, C. Brophy, A. Luscher, D. Nyfeler, and M. T. Sebastia. 2009. Diversity-interaction modeling: estimating contributions of species identities and interactions to ecosystem function. Ecology 90:2032-2038.

Kirwan, L., et al. 2007. Evenness drives consistent diversity effects in intensive grassland systems across 28 European sites. Journal of Ecology 95:530-539.

Lipowsky, A., C. Roscher, J. Schumacher, S. G. Michalski, M. Gubsch, N. Buchmann, E.-D. Schulze, and B. Schmid. 2015. Plasticity of functional traits of forb species in response to biodiversity. Perspectives in Plant Ecology, Evolution and Systematics 17:66-77.

Littell, R. C., G. A. Milliken, W. R. Stroup, R. D. Wolfinger, and O. Schabenberger. 2006. SAS for mixed models. 2nd edition. SAS Press.

Mc Culloch, C. E., and S. R. Searle. 2001. Generalised, Linear, and Mixed Models. Pages 222-223 *in* W. A. Shewhart and S. S. Wilks, editors. Probability and Statistics. John Wiley & Sons, Inc, New York.

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

SAS Institute Inc. Version 9.3 of the SAS System for Windows. Copyright © 2002-2010 SAS Institute Inc. SAS and all other SAS Institute Inc. product or service names are registered trademarks of SAS Institute Inc., Cary, NC, USA.

Schmid, B., A. Hector, M. A. Huston, P. Inchausti, I. Nijs, P. W. Leadley, and D. Tilman. 2002. The design and analysis of biodiversity experiments. *In* M. Loreau, S. Naeem, and P. Inchausti, editors. Biodiversity and Ecosystem Functioning: Systhesis and Perspectives. Oxford University Press, Oxford.

Self, S. G., and K. Y. Liang. 1987. Asymptotic properties of maximum-likelihood estimators and likelihood ratio tests under nonstandard conditions. Journal of the American Statistical

Association 82:605-610.

Spehn, E. M., et al. 2005. Ecosystem effects of biodiversity manipulations in European grasslands. Ecological Monographs 75:37-63.

Stram, D. O., and J. W. Lee. 1994. Variance component testing in the longitudinal mixed effects model. Biometrics 50:1171-1177.

Tilman, D., C. L. Lehman, and K. T. Thomson. 1997. Plant diversity and ecosystem productivity: Theoretical considerations. Proceedings of the National Academy of Sciences of the United States of America 94:1857-1861.

Wilsey, B. J., and H. W. Polley. 2004. Realistically low species evenness does not alter grassland species-richness-productivity relationships. Ecology 85:2693-2700.

TABLE 1. Model comparisons for (a) the Jena and (b) the Bell data sets. The full algebraic specification of each model is in Appendix S3. Abbreviations: # c = the number of fixed coefficients + variance parameters, -2LL = -2 log likelihood (from REML estimation), LRT = likelihood ratio test statistic.

Model #	Model terms	# c	-2LL	Comparison	Testing	LRT	*P*-value
(a) Jena data set							
J1	ID + DE_{fixed}	18 + 1	2394.5				
J2	ID + DE_{fixed} + DE_{random}	18 + 2	2388.8	J1 *vs*. J2	$\sigma_2^2 = 0$	5.7	0.008
J3	ID + DE_{fixed} + DE_{random}, resid var split	18 + 3	2385.3	J2 *vs*. J3	$\sigma_{1a}^2 = \sigma_{1b}^2$	3.5	0.061
(b) Bell data set							
B1	ID + DE_{fixed}	74 + 1	6464.1				
B2	ID + DE_{fixed} + DE_{random}	74 + 2	6463.2	B1 *vs*. B2	$\sigma_2^2 = 0$	0.9	0.171
B3	ID + DE_{fixed} + DE_{random}, resid var split	74 + 3	6463.1	B2 *vs*. B3	$\sigma_{1a}^2 = \sigma_{1b}^2$	0.1	0.752

FIG. 1. Illustration of how the variation of community responses (◆) around the mean response (——) may be (a) constant or (b) may change across the richness axis.

FIG. 2. (a) Estimated fixed effect functional group interactions (-) combined with estimated random effects for each pair of species (◇) for the Jena data. Predicted ecosystem function for the average community at each level of richness (solid line) with raw data superimposed for (b) the Jena and (c) the Bell data sets.