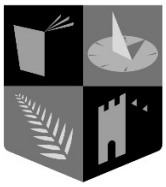

Modelling techniques for biodiversity and ecosystem multifunctionality - theoretical development and application.

Áine Dooley

A thesis presented for the degree of

Doctor of Philosophy

Department of Mathematics and Statistics,

Maynooth University

February 2018

Department Head: Professor Stephen Buckley

Supervisor: Dr. Caroline Brophy

Table of contents

Note that sections 2.1, 3.1, 4.2 and 4.3 are written as stand-alone papers, thus there is some variation in their form of presentation and they contain some duplicate conceptual explanations.

Acknowledgements

I'd like to take this opportunity to thank the many people who have helped me complete this work. Firstly my supervisor Dr. Caroline Brophy, who has not only been a great mentor throughout my PhD but has also gone above and beyond in the effort she's put in to helping me. I'd like to say thank you to my friends and family for helping me relax when I got stressed and for not asking when I would be finished too often, special thanks go to my housemates who didn't ask at all.

Throughout my PhD I received great support from the Mathematics and Statistics department in Maynooth University, not only with the PhD work but with a warm, friendly atmosphere. Gráinne O' Rourke, Janice Love and Tony Waldron were vital in helping me get here and so I'd like to say thank you specifically to them. The entire department made me feel welcomed so thank you to you all.

My PhD was part of a group project between a number of collaborators. I'd like to thank John Connolly, Laura Kirwan and John Finn for working with me and helping me throughout the PhD. Finally I would also like to thank the Science Foundation of Ireland for providing the financing for my PhD and a trip to America which allowed me to work with Forest Isbell, a great ecologist.

Publications

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

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

Summary

This thesis extends current methods for analysing the biodiversity and ecosystem function relationship, focusing on complexities arising from examining species rich ecosystems and examining multiple ecosystem functions. I developed a parsimonious model for assessing diversity effects on ecosystem function in species rich ecosystems, by including random effects in current univariate analysis methods. Using multivariate techniques, I created a method for analysing multiple ecosystem functions simultaneously. The work presented in this thesis allows for a greater ability to model and understand the biodiversity and ecosystem function relationship.

Chapter 1

Introduction

Biodiversity and ecosystem functioning

An ecosystem is a community containing a variety of living species which coexist and interact with one another and their local environmental factors. The term ecosystem covers a wide variety of communities of different types and sizes and developing an understanding of how ecosystems operate has been of great scientific interest. Ecosystems (such as agronomic, marine, microbial and grassland) have been widely studied because understanding how a system works can lead to improved management techniques and increased system outputs (Loreau *et al.* 2001; Tilman *et al.* 2006; Naeem *et al.* 2009; Weigelt *et al.* 2009; Finn *et al.* 2013).

Ecosystem functions are the measurable outputs of an ecosystem. What is classed as an ecosystem function depends on the ecosystem in question and its overall purpose. For agronomic ecosystems, some examples of ecosystem functions are the biomass produced by the crops grown or the fodder quality of the biomass produced. Ecosystem functions may be influenced by a number of different factors within the ecosystem, such as the number of species living in the system or abiotic factors such as soil type.

The biodiversity of an ecosystem is the number and variety of species within the ecosystem. Discussions on the importance of biodiversity to the maintenance and understanding of ecosystems date back as far as Charles Darwin (1859), who advocated more diverse ecosystems for improved outputs, although it wasn't until

the early 1990s that people began to understand the significance of biodiversity for ecosystem processes and outputs (Ehrlich & Wilson 1991; Chapin *et al.* 1992). In recent decades it has been established that declining biodiversity frequently has a negative effect on ecosystem functioning (Tilman *et al.* 1996; Hector *et al.* 1999; Loreau *et al.* 2001; Hooper *et al.* 2005; Cardinale *et al.* 2006; Cardinale *et al.* 2007; Cardinale *et al.* 2012). Developing a better understanding of how biodiversity loss is impacting on the ability of ecosystems to produce services has become an important research topic for the maintenance, development and sustainability of ecosystems globally. Studies across the globe have identified biodiversity loss across multiple ecosystems from areas such as marine life (Roger 2013) to forestry (Oldfield & Eastwood 2008). This global loss of species has strong effects on the ability of the planet to maintain life, through the loss of sustainable food supplies or the ability of forests to provide oxygen. There is currently an international drive, formed at the Convention of Biological Diversity in 2010, to maintain and improve biodiversity in over 160 countries around the globe; in Ireland this drive is headed by the Department of Arts, Heritage and the Gealtacht (2010).

Community characteristics and modelling the biodiversity and ecosystem function relationship

Community characteristics are aspects of the ecosystem, such as the identity of the species present, species richness, species relative abundances or species trait information which can illustrate and quantify the biodiversity of the ecosystem. The biodiversity and ecosystem function (BEF) relationship is the relationship between these community characteristics and the ability of the ecosystem to provide functions

such as the biomass produced or the nitrate content of the biomass produced. Here, several approaches that have been used in modelling the BEF relationship are introduced.

The Diversity-Interactions (DI) model

The Diversity-Interactions (DI) model (Kirwan *et al.* 2007; 2009) seeks to explain ecosystem function using the identity of the species in the ecosystem, the relative abundances of those species and how the species interact with one another. The DI model identifies how each species will perform in monoculture (a single species) and, for mixtures (more than one species), it separates the ecosystem function into two components: a component based solely on monoculture performances and the additional interaction effect caused by mixing species, known as the diversity effect. The use of species proportions allows for the simultaneous examination of the effect on ecosystem function of the richness (number of species) and evenness (measure of the distribution of the relative abundances of the species) of the ecosystem. Further details of this method and potential for its further development will be given later in the introduction.

The presence/absence model

Presence/absence modelling (Bell *et al.* 2009; Naeem *et al.* 2009) involves modelling the ecosystem function based on which species are present within the community using their species identities. The presence/absence model allows for the modelling of the relationship between the species present and the ecosystem

function, however, the model does not use the relative abundances of the species present. As such, the model cannot distinguish between communities that have the same composition (species identities) but that differ greatly in relative abundances. For example, a community where all species are equally present versus one where the same species are present, but one species is highly dominant. Modelling the BEF relationship using the species identities can provide a simple summary of the BEF relationship and identification of the effects of changing biodiversity levels. However, as discussed by Manel *et al.* (2001) interpretations should be treated cautiously as the prevalence of a species can strongly affect the BEF relationship.

Trait analysis

The physical and historical traits of species within an ecosystem have often been used in understanding the BEF relationship (for example Engelhardt & Kadlec 2000; Walters & Reich 2000; Cornelissen *et al.* 2003; Diaz *et al.* 2004; Cadotte *et al.* 2009). In grassland communities it has been shown that many traits such as plant growth rates or the plant specific leaf area can prove to be useful predictors of ecosystem function (Wright *et al.* 2004; Poorter *et al.* 2009). The phylogenetic distance trait measures how ancestrally distant two species are genetically with larger phylogenetic distances implying that species are genetically further apart. Previous studies have shown that the phylogenetic diversity of an ecosystem can be a strong predictor of ecosystem function (Cadotte *et al.* 2009; Connolly *et al.* 2011; Cadotte *et al.* 2012). Using a community level metric of the phylogenetic diversity, Connolly *et al.* (2011) have shown that higher phylogenetic diversity corresponds to increased ecosystem function.

Overyielding

Overyielding occurs when a mixture outperforms the average performance of the monocultures (Cardinale *et al.* 2006; Cardinale *et al.* 2007; Kirwan *et al.* 2007). Testing for overyielding involves a comparison of the ecosystem function (typically biomass in a grassland ecosystem) of mixture communities to monoculture (a single species) communities (Schmid *et al.* 2002; Cardinale *et al.* 2006; Cardinale *et al.* 2007). Transgressive overyielding is achieved when the mixture outperforms the best performing monoculture. A number of different methods can be used to evaluate overyielding. Kirwan *et al.* (2007) used a two-sided permutation test whereas Cardinale *et al.* (2006; 2007) and Schmid *et al.* (2002) used log ratios to examine the proportional differences in the communities. These studies have shown evidence of overyielding (Cardinale *et al.* 2006; 2007) and transgressive overyielding (Schmid *et al.* 2002; Kirwan *et al.* 2007).

Additive Partitioning

The effect of mixing species, or the diversity effect, can be positive (negative) if the mixture ecosystem function is higher (lower) than the expected performance based on monocultures, or zero if they equal. Loreau and Hector (2001) developed the additive partitioning method to divide the diversity effect in mixture communities into two components: the complementarity effect and the selection effect. Selection effects occur when a species which performed well in monoculture also performs well in mixture, i.e. that species contributes largely to the diversity

effect. The complementarity effect models any remaining diversity effect once the selection effects are accounted for. Loreau and Hector (2001) identified that, even accounting for selection effects, there was evidence that plant biodiversity significantly affected ecosystem function.

Ecosystem multifunctionality

The methods discussed so far focus on modelling a single ecosystem function however ecosystems provide multiple functions simultaneously (multifunctionality). The effect of biodiversity loss when analysing a single ecosystem function may underestimate how important the loss truly is when considered in the context of multiple functions. Multifunctionality is an emerging area of ecological research; initial work examining the BEF relationship using multifunctionality has led to the development of a number of multifunctional methods include the threshold method (Gamfeldt *et al.* 2008), the overlap method (Hector & Bagchi 2007), the averaging method (Maestre *et al.* 2012a) and multiple univariate analyses (Cardinale *et al.* 2013). These methods have shown that communities with higher biodiversity were more likely to be able to maintain multiple ecosystem functions.

Challenges that remain in modelling the BEF relationship

Examining the BEF relationship is key to understanding the true effect of biodiversity loss on ecosystem function however there are a number of challenges remaining. One such challenge is that the complexity of models for the BEF relationship can greatly increase as the number of species in the ecosystem increases.

Ecosystems can have very high species richness (e.g. Bell *et al.* 2005) and as the number of species present in an ecosystem increases it becomes increasingly difficult to create a parsimonious model that fits the data well.

The challenge of high species richness is magnified when considering multiple ecosystem functions. Current multifunctional methods often try to reduce the dimensionality of multiple function responses to a single measurement of multifunctionality, such as a threshold of their maximum (Zavaleta *et al.* 2010; Byrnes *et al.* 2014) or an average functioning metric (Maestre *et al.* 2012a; 2012b), however, this dimension reduction can cause a serious loss of information about the relationship between the biodiversity and the individual ecosystem functions. Ideally a BEF multifunctional model can simultaneously assess the effects of species identities, their relative abundances, community level richness and evenness for multiple ecosystem functions and test the relative importance of these effects across functions.

Details on the Diversity-Interactions model

The Diversity-Interactions (DI) model (Kirwan *et al*. 2007, 2009) is

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

where y is the ecosystem function, s is the number of species in the pool, A is a treatment or block factor, α is the effect associated with the treatment or block factor, P_i is the sown proportion of species i, β_i is the expected response of the i^{th} species in monoculture (i.e. when $P_i = 1$) known as the identity effect, δ_{ij} is the interaction effect between species i and j and $\varepsilon \sim N(0, \sigma^2)$. This modelling approach has two

main components, the first identifies the expected performance based on monoculture performances: $\sum_{i=1}^{s} \beta_i P_i + \alpha A$ and the second is the diversity effect (DE): $\sum_{\substack{i,j=1 \\ i<j}}^{s} \delta_{ij} (P_i P_j)$. This DI model has *s* identity parameters and *s(s-1)/2* interaction parameters. For species rich ecosystems the number of interaction parameters required for the full DI model can be difficult to interpret or impossible to fit due to lack of data. Kirwan *et al.* (2009) developed a number of possible ways to reduce the number of interaction parameters required to model the diversity effect, such as assuming all pairwise species interactions were equal ($\delta_{ij} = \delta_{av}$ for all i, j), to combat this drawback. Connolly *et al.* (2013) extended the DI model to develop the Generalised Diversity-Interactions (GDI) model by adding an additional parameter θ to allow for a nonlinear relationship between the ecosystem function and the pairwise species interactions.

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

Applications of the Diversity-Interactions model have shown that species abundances, as well as their identities, have a significant effect on the BEF relationship (Kirwan *et al.* 2007; Sheehan *et al.* 2008; Finn *et al.* 2013).

The goals of my PhD work

My research aims to extend the current statistical and biological understanding of the BEF relationship, primarily focusing on potential extensions to the Diversity-Interactions modelling framework. I aim to focus on two main

challenges in understanding the BEF relationship; firstly, the modelling of high species richness and secondly, modelling multiple ecosystem functions (multifunctionality). To address the first challenge, I will make use of species biological trait information and statistical techniques involving random effects to parsimoniously describe diversity effects and to provide validation for fixed effects models which have made assumptions about the diversity effects due to high species richness. To address the second challenge, I aim to examine and modify current multifunctionality methods. I also intend to build a multifunctional BEF model which will provide knowledge on how the biodiversity of the ecosystem affects each ecosystem function and allows for comparison of the relationship across ecosystem functions.

For species rich ecosystems, the use of underlying species traits, such as the phylogenetic distance, could potentially be useful to aid understanding a large number of interactions between species. Connolly *et al.* (2011) found that including a measure of community phylogenetic diversity improved model fit for two datasets in a model containing an average diversity effect, in lieu of estimating all pairwise interactions. Their results concluded that ecosystems with higher phylogenetic diversity had stronger diversity effects. In chapter 2, section 1, I aim to test whether the findings of Connolly *et al.* (2011) hold across multiple grassland ecosystems. If they do, then this is further evidence that the phylogenetic diversity is strongly linked to how species interact within an ecosystem.

In chapter 3, I aim to advance current modelling techniques for a single ecosystem function to address the challenge of building parsimonious models for complex ecosystems with high species numbers by using random effects alongside fixed effects. The use of random effects provides additional information about the

diversity effects in the ecosystem without requiring the estimation of high numbers of interaction coefficients. The inclusion of the random effects can also be used to assess lack of fit in the fixed diversity effects component of the model.

Chapter 4 is concerned with examining the multifunctional BEF relationship. Ecosystems provide multiple functions simultaneously and full understanding of the effects of biodiversity loss on ecosystem function requires multifunctional models. Some current multifunctional methods (overviewed in chapter 4, section 1) either examine each function individually or try to reduce the multiple responses into a binary or metric value. These reduction techniques can cause a serious loss of information, especially in methods such as the averaging method (Maestre *et al.* 2012a; 2012b) which analyses only the average of all ecosystem functions for each community. I aim to modify and improve upon the current averaging method (chapter 4, section 2). The current averaging metric may not differentiate between two communities that are functioning very differently (e.g. in one community all functions could be performing at similar levels, whereas in a different community one function could be strongly outperforming the others but both communities have the same average metric value). I will modify the average metric to penalise communities where functions are not performing similarly to each other. However, the improved metric will likely still suffer from many of the problems associated with the current averaging metric, such as interpretation. Finally, I aim to develop a more advanced method for analysing multiple functions by extending the DI model to a multivariate framework (chapter 4, section 3). By extending the DI model in this way I can gain information about how the community characteristics affect each ecosystem function simultaneously and how the functions correlate with one another. I can also compare the effects of community characteristics across functions. The

Multivariate DI model allows for the examination and prediction of ecosystem function responses across the full range of species proportions. The model can provide a full assessment of the multifunctional BEF relationship, which can be used for identification of areas where the ecosystem is performing well for all ecosystem functions or where trade-offs are occurring among functions.

To summarise, the main four goals of my PhD are

1. To explore the use of community phylogenetic diversity information to help improve models for the BEF relationship for species rich communities.
2. To develop a random effects Diversity-Interactions model to increase the understanding of the BEF relationship for a single function.
3. To review and improve upon current multifunctionality metrics focusing on the averaging metric.
4. To develop a Multivariate Diversity-Interactions model to analyse the multifunctional BEF relationship.

References

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.

Bell, T., Newman, J., Silverman, B., Turner, S. & Lilley, A. (2005). The contribution of species richness and composition to bacterial services. *Nature*, 436, 1157-1160.

Byrnes, J., Gamfeldt, L., Isbell, F., Lefcheck, J., Griffin, J., Hector, A. *et al.* (2014). Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions. *Methods in Ecology and Evolution*, 5, 111-124.

Cadotte, M., Cavender-Bares, J., Tilman, D. & Oakley, T. (2009). Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. *PLoS ONE*, 4, e5695.

Cadotte, M., Dinnage, R. & Tilman, D. (2012). Phylogenetic diversity promotes ecosystem stability. *Ecology*, 93, 223-233.

Cardinale, B., Gross, K., Fritschie, K., Flombaum, P., Fox, J., Rixen, C. *et al.* (2013). Biodiversity simultaneously enhances the production and stability of community biomass, but the effects are independent. *Ecology*, 94, 1697-1707.

Cardinale, B., Srivastava, D., Duffy, J., Wright, J., Downing, A., Sankaran, M. *et al.* (2006). Effects of biodiversity on the functioning of trophic groups and ecosystems. *Nature*, 443, 989-992.

Cardinale, B., Wright, J., Cadotte, M., Carroll, I., Hector, A., Srivastava, D. *et al.* (2007). Impacts of plant diversity on biomass production increase through time because of species complementarity. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 18123-18128.

Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P. *et al.* (2012). Biodiversity loss and its impact on humanity. *Nature*, 486, 59-67.

Chapin, F.S., Schulze, E.D. & Mooney, H.A. (1992). Biodiversity and ecosystem processes. *Trends in Ecology & Evolution*, 7, 107-108.

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.

Connolly, J., Cadotte, M., Brophy, C., Dooley, Á., Finn, J.A., Kirwan, L. *et al.* (2011). Phylogenetically diverse grasslands are associated with pairwise interspecific processes that increase biomass. *Ecology*, 92, 1385-1392.

Cornelissen, J., Lavorel, S., Garnier, E., Diaz, S., Buchmann, N., Gurvich, D. *et al.* (2003). A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany*, 51, 335-380.

Darwin, C. (1859). *On the origin of species by means of natural selection,or the perservation of favoured races in the struggle for life*. John Murray, London, England.

Department of Arts, Heritage and the Gealtacht, Ireland (2010). Actions for biodiversity 2011-2016. http://www.ahg.gov.ie/en/Publications/HeritagePublications/NatureConservationPublications/Actions%20for%20Biodiversity%202011%20-%202016.pdf.

Diaz, S., Hodgson, J., Thompson, K., Cabido, M., Cornelissen, J., Jalili, A. *et al.* (2004). The plant traits that drive ecosystems: Evidence from three continents. *Journal of Vegetation Science*, 15, 295-304.

Ehrlich, P.R. & Wilson, E.O. (1991). Biodiversity studies - science and policy. *Science*, 253, 758-762.

Engelhardt, K.A.M. & Kadlec, J.A. (2000). Species traits, species richness and the resilience of wetlands after disturbance. In: *40th Annual Meeting of the Aquatic-Plant-Management-Society* San Diego, Ca, pp. 36-39.

Finn, J.A., Kirwan, L., Connolly, J., Sebastia, M., Helgadottir, A., Baadshaug, O. *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.

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

Hector, A. & Bagchi, R. (2007). Biodiversity and ecosystem multifunctionality. *Nature*, 448, 188-191.

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

Hooper, D., Chapin, F., Ewel, J., Hector, A., Inchausti, P., Lavorel, S. *et al.* (2005). Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs*, 75, 3-35.

Kirwan, L., Connolly, J., Finn, J.A., Brophy, C., Luscher, A., Nyfeler, D. *et al.* (2009). Diversity-Interaction modeling: estimating contributions of species identities and interactions to ecosystem function. *Ecology*, 90, 2032-2038.

Kirwan, L., Luescher, A., Sebastia, M., Finn, J.A., Collins, R., Porqueddu, C. *et al.* (2007). Evenness drives consistent diversity effects in intensive grassland systems across 28 European sites. *Journal of Ecology*, 95, 530-539.

Loreau, M. & Hector, A. (2001). Partitioning selection and complementarity in biodiversity experiments. *Nature*, 412, 72-76.

Loreau, M., Naeem, S., Inchausti, P., Bengtsson, J., Grime, J., Hector, A. *et al.* (2001). Ecology - Biodiversity and ecosystem functioning: Current knowledge and future challenges. *Science*, 294, 804-808.

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

Maestre, F.T., Quero, J.L., Gotelli, N.J., Escudero, A., Ochoa, V., Delgado-Baquerizo, M. *et al.* (2012b). Plant species richness and ecosystem multifunctionality in global drylands. *Science*, 335, 214-218.

Manel, S., Williams, H.C. & Ormerod, S.J. (2001). Evaluating presence–absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, 38, 921-931.

Naeem, S., Bunker, D.E., Hector, A., Loreau, M. & Perrings, C. (2009). *Biodiversity, Ecosystem Functioning and Human Wellbeing*. Oxford University Press.

Oldfield, S. & Eastwood, A. (2008). The red list of Oaks. http://www.bgci.org/files/Worldwide/Publications/the_red_list_of_oaks.pdf.

Poorter, H., Niinemets, U., Poorter, L., Wright, I. & Villar, R. (2009). Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. *New Phytologist*, 182, 565-588.

Roger, A. (2013). International Programme on the state of the ocean press release. http://www.stateoftheocean.org/pdfs/IPSO-PR-2013-FINAL.pdf.

Schmid, B., Hector, A., Huston, M.A., Inchausti, P., Nijs, I., Leadley, P.W. *et al.* (2002). *The design and analysis of biodiversity experiments*.

Sheehan, C., Kirwan, L., Connolly, J. & Bolger, T. (2008). The effects of earthworm functional diversity on microbial biomass and the microbial community level physiological profile of soils. *European Journal of Soil Biology*, 44, 65-70.

Tilman, D., Reich, P. & Knops, J. (2006). Biodiversity and ecosystem stability in a decade-long grassland experiment. *Nature*, 441, 629-632.

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

Walters, M. & Reich, P. (2000). Seed size, nitrogen supply, and growth rate affect tree seedling survival in deep shade. *Ecology*, 81, 1887-1901.

Weigelt, A., Weisser, W., Buchmann, N. & Scherer-Lorenzen, M. (2009). Biodiversity for multifunctional grasslands: equal productivity in high-diversity low-input and low-diversity high-input systems. *Biogeosciences*, 6, 1695-1706.

Wright, I., Groom, P., Lamont, B., Poot, P., Prior, L., Reich, P. *et al.* (2004). Leaf trait relationships in Australian plant species. *Functional Plant Biology*, 31, 551-558.

Zavaleta, E., Pasari, J., Hulvey, K. & Tilman, G. (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.

Chapter 2

Phylogenetic distance and its relationship to the diversity effect

Collaborators: Caroline Brophy, John Connolly, Laura Kirwan, John A. Finn and Marc W. Cadotte.

Introduction

In grassland ecosystems the underlying physical traits of the plants within the community have often been used to examine the biodiversity and ecosystem function (BEF) relationship. The focus of this chapter is the use of the phylogenetic distance between plant species to examine patterns in the BEF relationship. The phylogenetic distance between two plant species is a measure of how genetically related the species are, based on their ancestry. A small phylogenetic distance implies that the species shared a common genetic ancestor more recently than those with larger phylogenetic distances.

In previous work, discussed in section 2.1, it was established that a community phylogenetic distance measurement is a useful predictor in BEF relationship models in addition to species relative abundances and an average species interaction effect, and that communities with higher phylogenetic distances are more likely to have larger diversity effects. In section 2.1 we test whether these results hold across a range of different datasets and examine possible reasons as to why they may or may not hold.

Section 2.1

Testing the association between phylogenetic distance and ecosystem functioning across eight grassland experiments.

Introduction

Reduced biodiversity in ecosystems is known to have a negative effect on the sustainability and productivity of grassland ecosystems (Naeem *et al.* 1994; Tilman *et al.* 1996; Hector *et al.* 1999), however, less is known about the role played by evolutionary associations among species in ecosystems. Evolutionary relationships among species in an ecosystem can be described by phylogenetic information and there has been a recent increase in the availability of species phylogenetic information which may be useful in examining the biodiversity and ecosystem function (BEF) relationship. Webb *et al.* (2002) provide a review of phylogenetic biology and the development of an increased understanding of how it is linked to community characteristics such as species identities, diversity and species relative abundances within an ecosystem. They also discuss possible metrics for including phylogenetic information in a BEF model, such as the phylogenetic distance metric net relatedness index which measures the mean phylogenetic distance between pairs of species within an ecosystem and the nearest taxon index which examines the phylogenetic distance of a species to the nearest taxon. Taxa are groups of species which have similar genetic characteristics. Phylogenetic distance is a trait measure of how genetically related two plant species are, based on their ancestry. A smaller phylogenetic distance

between two species implies that they had a common ancestor more recently than those with a larger phylogenetic distance. Cadotte *et al.* (2008; 2009; 2012) showed that a community measure of the phylogenetic distance was a more useful predictor for biomass produced in grasslands than the species richness (count of species) and functional groups classifications (grouping of species that perform similarly, such as grasses and legumes). Cavender-Bares *et al.* (2009) discuss various studies which have examined the relationship between phylogenetic distance and community characteristics within the ecosystem such as the species abundances present or other trait information of the species. These studies have shown mixed results as to the relationship between phylogenetic distance and the community characteristics in the ecosystem. However, Cavender-Bares *et al.* (2009) discuss how these mixed results may be due to the differing methods applied in the studies and points out a need for a more rigorous approach.

The Diversity-Interactions model (Kirwan *et al.* 2007; 2009) has been used to model the BEF relationship with explanatory variables including species relative abundances and species interactions. In our previous work in Connolly *et al*. (2011), we extended the Diversity-Interactions model to include a combined measurement of the phylogenetic distances of all species in a community. This allowed us to determine whether the community phylogenetic diversity could explain patterns in ecosystem function in addition to species relative abundances and an average species interaction effect. We found (in two data sets) that the community phylogenetic distance was a significant predictor for the ecosystem response, with more phylogenetically diverse communities yielding higher than expected ecosystem function. In this current work we apply the Connolly *et al*. (2011) approach to a further eight datasets to examine whether community

phylogenetic distance is consistently a useful predictor of ecosystem function across multiple datasets. The aim of this work is to test how robust the results found in Connolly *et al*. (2011) are to varying species and ecosystem conditions.

Methods

The Diversity-Interactions model including community phylogenetic distance

The Diversity-Interactions model (Kirwan *et al.* 2007; 2009) is

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

where y is the ecosystem function, P_i is the sown proportion of species i, for i = 1,…,s, s is the number of species present in the ecosystem, A is a block or treatment factor, α is change in the response for the changing levels of A, β_i is the identity effect or, equivalently, the expected response of the i^{th} species in monoculture, δ_{ij} is the interaction effect between species i and j and $\varepsilon \sim N(0, \sigma^2)$. In this model the diversity effect (DE) of sowing species together in a community is modelled through the interaction terms, DE: $\sum_{\substack{i,j=1 \\ i<j}}^{s} \delta_{ij} (P_i P_j)$. Through various assumptions about patterns among the species interaction coefficients, this model can be made more parsimonious (Kirwan *et al.* 2009). For example, by assuming each species interaction effect is equal ($\delta_{ij} = \delta_{av}$ for all i, j) we can reduce the number of parameters required to model the diversity effect from *s(s-1)/2* down to one.

Connolly *et al.* (2011) expanded the Diversity-Interactions model by including a community measure of phylogenetic diversity giving the model

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

where C_D is a combined measure of the phylogenetic distances of all species in the community. Specifically

$$C_D = \sum_{\substack{i,j=1 \\ i<j}}^{s} (D_{ij} - \overline{D}) P_i P_j$$

where D_{ij} is the phylogenetic distance between species i and j, $\overline{D}$ is the average phylogenetic distance of all species present in the species pool (i.e. species in the ecosystem) and P_i represents the sown proportion of the i^{th} species present in the plot. Positive values of C_D indicate greater than average phylogenetic diversity and vice versa for negative. Using this modelling approach, Connolly *et al.* (2011) found that, for two grassland datasets, the community phylogenetic distance explained significant variability in ecosystem function in addition to the average pairwise effect and that the greater the community phylogenetic diversity the larger the diversity effect.

Application to multiple datasets

We applied the methodology of Connolly *et al*. (2011) to eight datasets from field and greenhouse grassland biodiversity experiments. The datasets used were a subset of the database compiled by Cadotte *et al.* (2008), for which the phylogenetic distances between all species in the experiment were available (Table 2.1.1). In each of these experiments the species richness was manipulated

to examine the effect of species diversity on ecosystem function. In some cases additional treatments were also tested (Table 2.1.1; additional dataset details are available in Appendix 2.1.1).

Table 2.1.1: The eight datasets with reference to their source, the number of species in the species pool, the number of plots analysed, any additional treatments applied and whether the data was from a greenhouse or field experiment.

Dataset	Reference	Number of species	Number of plots / pots	Additional treatment	Type
1	Dimitrakopoulos & Schmid (2004)	10	90	Soil depth	Greenhouse
2	Fridley (2002)	9	233	Soil fertility	Field
3	Fridley (2003)	7	252	Soil fertility, light	Field
4	Lanta & Leps (2006)	16	178	Soil fertility	Greenhouse
5	Naeem (1999)	6	360	None	Greenhouse
6	Naeem *et al.* (1996)	13	90	None	Field
7	Craine *et al.* (2003)	11	56	CO_2, light	Field
8	Tilman (1997)	12	22	None	Field

The following models were fitted to each of the eight datasets:

Model 1: $y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \varepsilon$

Model 2: $y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} P_i P_j + \varepsilon$

Model 3: $y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} P_i P_j + \kappa C_D + \varepsilon$

Model 1 has no diversity effect and assumes $\delta_{ij} = 0$ for all i, j, Model 2 is the average Diversity-Interactions model which assumes $\delta_{ij} = \delta_{av}$ for all i, j and Model 3 includes the average interaction effect and the community phylogenetic distance effect. All models were fitted using SAS 9.3 (*SAS* Institute Inc.) software and model comparisons (2 vs. 1 and 3 vs. 2) were made using F-tests. To examine how the variability of the phylogenetic effect parameter changes across datasets the response variable (the plant biomass produced) was standardized for each dataset by subtracting the mean and dividing by its standard deviation. Model 3 was then refitted to the standardized response. Model assumptions, i.e. $\varepsilon \sim N(0, \sigma^2)$ are independent and identically distributed, were tested for the final model selected for each dataset. Residual diagnostics plots from model 3 for each dataset are given in Appendices 2.1.2 to 2.1.9 respectively.

We also fitted the full pairwise Diversity-Interactions model to dataset 5:

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

This model could not be fully fitted to the other datasets because of insufficient data for estimating the *s(s-1)/2* pairwise interaction terms or because of confounding of the interactions terms due to experimental design. For dataset 5,

we compared models 3 and 4 using an F-text to test for lack of fit in the two-parameter (i.e. δ_{av} and κ) explanation of diversity effects in model 3. This is a nested comparison since model 4 can be reparametrized as:

Model 4*: $y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \kappa C_D + \sum_{\substack{i,j=1 \\ i<j}}^{s} \delta_{ij} P_i P_j + \varepsilon$

Results

The inclusion of an average pairwise interaction effect δ_{av} (Model 2) significantly improved the model fit over the identity effects model (Model 1) for seven of the eight datasets (Table 2.1.2, Model 2 vs. 1). The inclusion of the community phylogenetic distance variable further improved the model fit for two of the datasets (datasets 4 and 5, Table 2.1.2, Model 3 vs. 2). For both of these datasets the estimated community phylogenetic distance measure κ was positive (Table 2.1.2, estimated κ), implying that communities with higher phylogenetic diversity have increased expected biomass produced. These two datasets also had the smallest associated standard errors of all the datasets (Table 2.1.2, Standardized Response Estimated κ). The residual diagnostic plots for model 3 for datasets 4 and 5 generally indicated that model assumptions were satisfied although there were some indications of issues with assumptions for some of the other datasets (Appendices 2.1.2 to 2.1.9). For dataset 5, the F-test comparing models 3 and 4 showed a significant lack of fit (F=2.12, p-value =0.012).

Table 2.1.2: Model comparisons using F-tests for the eight datasets and the estimated phylogenetic coefficient (κ) on the raw data scale and on a standardised scale. The p-value for each F test is given in brackets after the test value. Standard errors are given in brackets after the coefficient estimates. Significant tests are highlighted in bold (α=0.05).

Dataset	F-tests			Estimated κ	
	Model 2 vs. 1 (δ_{av})	Model 3 vs. 2 (κ)	Model 4 vs. 3 (lof)	Scale of data	Standardized
1	0.35 (0.554)	0.06 (0.802)		70.03 (277.709)	0.26 (1.047)
2	**5.72 (0.018)**	2.62 (0.107)		509.54 (314.821)	5.47 (3.377)
3	**27.29 (<0.001)**	0.38 (0.541)		-206.86 (337.644)	-2.46 (4.014)
4	**87.01 (<0.001)**	**8.56 (0.004)**		**2.68 (0.916)**	**1.69 (0.579)**
5	**11.08 (0.001)**	**4.86 (0.028)**	**2.12 (0.013)**	**6.30 (2.856)**	**1.16 (0.527)**
6	**9.22 (0.003)**	1.47 (0.229)		-10.08 (8.315)	-1.55 (1.274)
7	**4.34 (0.043)**	0.23 (0.631)		286.26 (590.652)	1.60 (3.297)
8	**7.17 (0.032)**	0.27 (0.623)		181.48 (350.494)	1.63 (3.143)

Discussion

Connolly *et al.* (2011) found that the community phylogenetic distance was a useful predictor for ecosystem function. The aim of this current work was to examine whether this result held in general. To do this, we fitted a Diversity-Interactions model with an average pairwise interaction effect and a community measure of phylogenetic diversity to a range of datasets and, where possible, tested for lack of fit in the two-parameter description of diversity effects.

Of the eight different grassland experiments tested, we found that for two datasets (datasets 4 and 5) the community phylogenetic distance measure contributed significantly to the ecosystem response, in addition to the average pairwise interaction effect. For both of these datasets the phylogenetic effect κ was positive, agreeing with the conclusion found in Connolly *et al.* (2011) that more phylogenetically diverse communities tend to have higher ecosystem function values. Thus we provide some further evidence of the patterns shown in Connolly *et al*. (2011). For dataset 5 we found there was evidence of lack of fit between the model containing the phylogenetic effect and an average diversity effect (model 3) and the model with all pairwise interaction effects (model 4 or 4*). For this dataset, it may be useful to explore additional patterns (e.g. related to functional groups) to the average pairwise interaction term alongside the phylogenetic description. For dataset 4, it may be possible to test for lack of fit using a random effects approach to modelling pairwise interactions (see thesis Section 3.1).

However, we also found that in six datasets there were no significant community phylogenetic diversity effects. This lack of datasets where the

community phylogenetic distance was a significant predictor could be due to a number of reasons. The studies we examined were not originally designed to examine phylogenetic distance. As such it may be that we do not have a wide enough range of phylogenetic diversity in our datasets to truly examine its effect. There was considerable overlap and a range of widths in the confidence intervals for the phylogenetic effects constructed on a standardised scale (Figure 2.1.1).

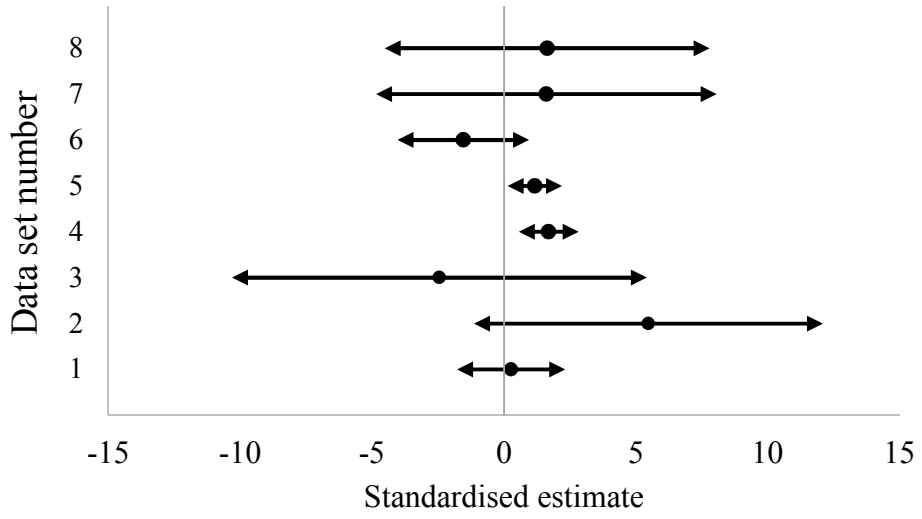

Figure 2.1.1: Estimated κ on a standardised scale for each dataset with 95% confidence intervals.

Examining the individual datasets showed that datasets 4, 5 (where κ was significant) and 6 were among those with the greatest range in community phylogenetic distance measures of the eight datasets (Appendix 2.1.10). In addition, the datasets come from differing environments, namely some datasets are greenhouse and some are field experiments and seven of the eight experiments

have additional treatments applied. While we included additive effects of treatments, there could be interaction effects between these factors and community phylogenetic distance which we did not account for but which might influence our ability to detect phylogenetic effects.

Although overall we only found two out of eight datasets agreed with the results found by Connolly *et al*. (2011) there are a number of possible reasons as to why the community phylogenetic distance may not have explained significant variability for the other datasets. Therefore further research is needed to test the robustness of community level phylogenetic diversity on ecosystem function. Of the datasets where the community phylogenetic distance was significant we found that our conclusions agreed with those of Connolly *et al*. (2011), i.e. increased community phylogenetic diversity had a positive effect on ecosystem function.

References

Cadotte, M., Cardinale, B. & Oakley, T. (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.

Cadotte, M., Cavender-Bares, J., Tilman, D. & Oakley, T. (2009). Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. *PLoS ONE*, 4, e5695.

Cadotte, M., Dinnage, R. & Tilman, D. (2012). Phylogenetic diversity promotes ecosystem stability. *Ecology*, 93, 223-233.

Cavender-Bares, J., Kozak, K.H., Fine, P.V.A. & Kembel, S.W. (2009). The merging of community ecology and phylogenetic biology. *Ecology Letters*, 12, 693-715.

Connolly, J., Cadotte, M., Brophy, C., Dooley, Á., Finn, J.A., Kirwan, L. *et al.* (2011). Phylogenetically diverse grasslands are associated with pairwise interspecific processes that increase biomass. *Ecology*, 92, 1385-1392.

Craine, J., Reich, P., Tilman, G., Ellsworth, D., Fargione, J., Knops, J. *et al.* (2003). The role of plant species in biomass production and response to elevated CO_2 and N. *Ecology Letters*, 6, 623-630.

Dimitrakopoulos, P. & Schmid, B. (2004). Biodiversity effects increase linearly with biotope space. *Ecology Letters*, 7, 574-583.

Fridley, J. (2002). Resource availability dominates and alters the relationship between species diversity and ecosystem productivity in experimental plant communities. *Oecologia*, 132, 271-277.

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

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

Kirwan, L., Connolly, J., Finn, J.A., Brophy, C., Luscher, A., Nyfeler, D. *et al.* (2009). Diversity-Interaction modeling: estimating contributions of species identities and interactions to ecosystem function. *Ecology*, 90, 2032-2038.

Kirwan, L., Luescher, A., Sebastia, M., Finn, J.A., Collins, R., Porqueddu, C. *et al.* (2007). Evenness drives consistent diversity effects in intensive

grassland systems across 28 European sites. *Journal of Ecology*, 95, 530-539.

Lanta, V. & Leps, J. (2006). Effect of functional group richness and species richness in manipulated productivity-diversity studies: a glasshouse pot experiment. *Acta Oecologica*, 29, 85-96.

Naeem, S., Hakansson, K., Lawton, J., Crawley, M. & Thompson, L. (1996). Biodiversity and plant productivity in a model assemblage of plant species. *Oikos*, 76, 259-264.

Naeem, S., Thompson, L., Lawler, S., Lawton, J. & Woodfin, R. (1994). Declining biodiversity can alter the performance of ecosystems. *Nature*, 368, 734-737.

Naeem, S., Tjossem, S., Byers, D., Bristow, C. & Li, S. (1999). Plant neighborhood diversity and production. *Ecoscience*, 6, 355-365.

SAS Institute Inc., C., NC, USA. *SAS* System for Windows. Copyright © 2002-2008 *SAS* Institute Inc.

Tilman, D. (1997). Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology*, 78, 81-92.

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

Webb, C., Ackerly, D., McPeek, M. & Donoghue, M. (2002). Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, 33, 475-505.

Chapter 3

Modelling the biodiversity and ecosystem function relationship in species rich systems

Collaborators: Caroline Brophy, John Connolly, Laura Kirwan, John A. Finn, Thomas Bell and Marc W. Cadotte.

INTRODUCTION

Modelling the biodiversity and ecosystem function (BEF) relationship in species rich systems can be challenging because the large number of parameters required may be difficult to interpret or impossible to estimate. Previously, assumptions have been made to reduce the number of parameters required to model the BEF relationship. The focus of this chapter is to create a mixed model which is parsimonious for species rich ecosystems and can test whether there is evidence of lack of fit in models with reduced numbers of parameters. This model will provide additional information about the relationship between ecosystem functional response and species interactions that does not rely on a fixed effect estimate of each pairwise interaction, which will be useful in a species rich ecosystem in particular.

Section 3.1

The use of random effects for modelling the biodiversity and ecosystem function relationship in diverse species rich communities.

Áine Dooley[1], John Connolly[2], Laura Kirwan[3,7], John A. Finn[4], Thomas Bell[5], Marc W. Cadotte[6] and Caroline Brophy[1].

[1]Department of Mathematics and Statistics, Maynooth University, Maynooth, Co. Kildare, Ireland; [2]School of Mathematical Sciences, Ecological and Environmental Modelling Group, University College Dublin, Dublin 4, Ireland; [3]Waterford Institute of Technology, Cork Road, Waterford, Ireland; [4]Teagasc Environment Research Centre, Johnstown Castle, Co. Wexford, Ireland; [5]Faculty of Natural 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]Present address: UCD Institute of Food and Health, UCD, Belfield, Dublin 4, Ireland.

Summary

1. Biodiversity research has shown that ecosystem function can be improved by increasing community biodiversity such as species richness. In addition to richness, evenness and interactions among species may also play important roles. Disentangling the various diversity impacts on the biodiversity and ecosystem function (BEF) relationship can be complex, particularly in species rich ecosystems.
2. Generalised Diversity-Interactions models have been used for testing how ecosystem function is affected by a range of community characteristics including species identity, species interactions, richness and evenness. However, the number of coefficients required to describe species interactions in a species rich system may be difficult to interpret or impossible to estimate. Parsimonious descriptions using constraints among fixed coefficients have been developed but a combination of fixed and random coefficients may provide further explanatory power.
3. We develop the Generalised Diversity-Interactions Mixed model to model the biodiversity and ecosystem function relationship across a wide range of community characteristics using a combination of fixed and random terms, resulting in a relatively small number of coefficients to describe diversity effects. If the random effects are not needed, it provides validation for the fixed effect explanation of the diversity effect. If they are needed, the additional variability will feed into standard errors for fixed effects, improving inference.
4. We apply the methods to two data sets from a grassland and a bacterial experiment. The random effects were significant in the final model fitted to the data from the grassland experiment, while in the bacterial experiment, the random effects

were not needed, providing reassurance for the inference provided by the fixed effect component.

5. The Generalised Diversity-Interactions Mixed model provides a parsimonious description of how species interact in a community and can address a wide range of questions related to which community characteristics affect ecosystem function. It provides a platform for assessing species interactions that does not rely on a fixed effect estimate of each pairwise interaction, which is particularly useful in a species rich ecosystem.

Keywords: community characteristics, evenness, mixed model, random diversity effects, random effects, richness, species interactions, species rich.

Introduction

The biodiversity and ecosystem function (BEF) relationship has been widely studied (Tilman et al. 1996, Hector et al. 1999, Tilman 1999, Loreau et al. 2001, Cardinale et al. 2002, Petchey and Gaston 2006, Tilman et al. 2006, Kirwan et al. 2007, Duffy 2009) and it is often concluded that increasing the biodiversity of a system improves its ability to maintain and/or increase functionality (Hooper et al. 2005, Duffy 2009, Hillebrand and Matthiessen 2009). Models of the BEF relationship often seek to explain the conditions under which ecosystem function will be maximised (e.g. biomass yield in agronomy) or minimised (e.g. invasion by exotic species in natural systems) using species richness as the main driver (Tilman et al. 1996, Hooper et al. 2005, Lanta and Lepš 2006), however species evenness or species interactions may also contribute significantly to the relationship. In a species rich ecosystem

quantifying species interactions may present analytical difficulties due to the potentially large number of influential interactions.

The Diversity-Interactions (Kirwan et al. 2007, Kirwan et al. 2009) and Generalised Diversity-Interactions (GDI) (Connolly et al. 2013) modelling approaches estimate the contributions of species-specific and pairwise species interaction effects to total ecosystem functioning. These models have successfully assessed the impact of community characteristics such as species identity, species initial proportions, species interactions, species richness and evenness on ecosystem function. When there is a large species pool, interpretation of the high number of pairwise interaction coefficients in GDI models may be difficult, or estimation of all pairwise coefficients may not be possible due to study design. Biologically motivated constraints among the interaction coefficients can however lead to meaningful and parsimonious model variants (Kirwan et al. 2009), *i.e.* assuming patterns in species interactions that can be represented by a few coefficients. While these models involving fixed effect solutions are useful, their explanatory power could be improved by modelling the remaining variability among the constrained interaction coefficients using variance components. This would provide a more parsimonious description of species interaction effects than estimating all individual pairwise interactions.

The relationship between ecosystem function and richness has been shown to be a positive saturating curve in many systems (e.g. Hector et al. 1999), as in Figure 3.1.1. The spread of communities or variability around the line may be somewhat constant at each level of richness (Fig. 3.1.1a) or it may vary depending on richness (Fig. 3.1.1b). This spread (constant or not) is not pure replicate variability and is

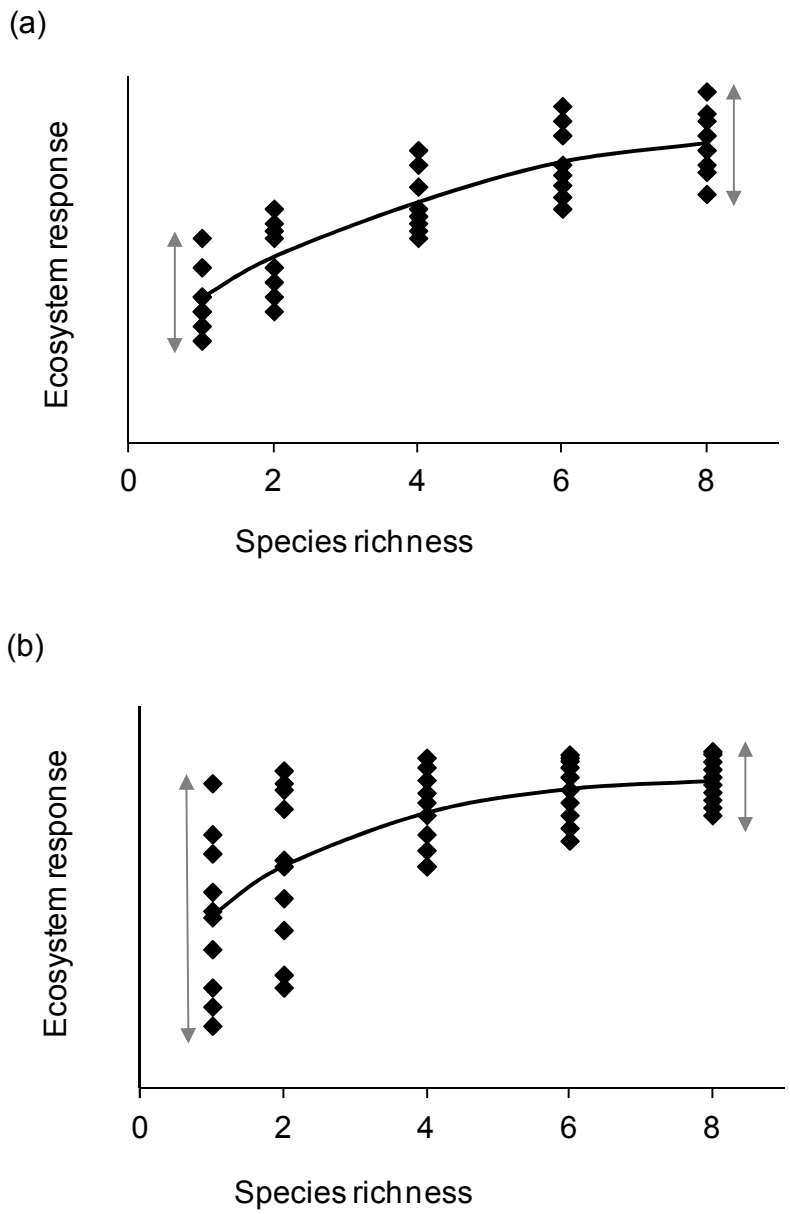

Fig. 3.1.1. Hypothetical illustration of how the spread of community responses (◆) around the mean response (——) may be (a) constant or (b) may change across the richness axis.

likely caused by factors such as species identities, species relative abundances, specific pairwise interactions or community evenness, each of which can be tested for using a GDI model (Connolly et al. 2013). In this paper, we extend the GDI

model to a mixed modelling framework. We assume a random distribution for the pairwise interaction coefficients and test various assumptions about the error terms. The benefits of our proposed Generalised Diversity-Interactions Mixed (GDIM) model are two-fold; 1. Large numbers of fixed effect species pairwise interaction coefficients can be replaced by a smaller number of fixed coefficients combined with variance components providing a parsimonious but powerful description of diversity effects. 2. The inclusion of variance components for interaction coefficients provides a means to test for lack-of-fit in the fixed effect description of the diversity effect. We apply the method to data sets from two experiments, one grassland and one bacterial. Our approach provides new methodological tools to assess the relationship between biodiversity and ecosystem function that is particularly useful for species rich ecosystems.

Materials and Methods

MODELS

The Generalised Diversity-Interactions (GDI) model (Connolly et al. 2013) is of the form

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

where $\varepsilon \sim N(0, \sigma_1^2)$ (Model 1a)

The community ecosystem function is y, P_i is the initial relative abundance of the i^{th} species (i=1,...,s), A is a block or treatment factor, α is the block/treatment effect and there is a pool of s species. The GDI model is a generalised version of the Diversity-

Interactions (DI) model (Kirwan et al. 2007, Kirwan et al. 2009); if $\theta_1 = 1$ then the GDI model reverts to the DI model. In model 1a, 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). Each δ_{ij} coefficient is scaled by the product of the initial relative abundances of the two species (P_i and P_j) to the power of θ_1 to compute the expected interactive contribution of those two species to ecosystem function. The value of θ_1 determines the nature of this contribution (see Figures 2 and 3 in Connolly et al. 2013), for example if $\theta_1 = 0$ then the δ_{ij} pairwise interaction coefficients are not scaled regardless of the sown species proportions, while if $\theta_1=1$, the scaling is exactly the product of the sown proportions. In the absence of any species interactions (i.e. δ_{ij} = 0 for all i, j) then $\Sigma\beta_iP_i+\alpha A$ is the expected ecosystem response. The diversity effect is an additional effect on the expected response caused by mixing of species, i.e. the difference between the expected mixture response and what would be expected based solely on the species monoculture responses; in model 1a the diversity effect is $\sum\delta_{ij}(P_iP_j)^{\theta_1}$. For the full pairwise Generalised Diversity-Interactions model (model 1a), the diversity effect requires estimating θ_1 and $s(s-1)/2$ δ_{ij} coefficients, which is, for example, seven coefficients in a four-species system but 191 coefficients in a 20-species system. This number of coefficients can be substantially reduced by testing for various patterns among the interactions coefficients (Kirwan et al. 2009).

Here we propose the Generalised Diversity-Interactions Mixed (GDIM) model which assumes that the pairwise interaction coefficients follow a random normal distribution as opposed to being fixed:

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} (P_i P_j)^{\theta_1} + \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij} (P_i P_j)^{\theta_2} + \varepsilon$$

where $d_{ij} \sim N(0, \sigma_2^2)$ and $\varepsilon \sim N(0, \sigma_1^2)$, (independent of each other) (Model 1b)

This model requires only four coefficients (δ_{av}, θ_1, θ_2 and σ_2^2) to describe the diversity effect, regardless of the species pool size, which is a more parsimonious description than with model 1a. The coefficient δ_{av} is the average of the true pairwise interactions δ_{ij}. In addition to reducing the number of coefficients required to model the ecosystem function, the GDIM model allows us to test for lack-of-fit in the fixed effect description of the diversity effect. Specifically, testing if $\sigma_2^2 = 0$ allows us to test whether, after accounting for fixed effects, there is additional variability among the true δ_{ij} pairwise interaction coefficients; if the test is non-significant, it can be assumed that the fixed effects sufficiently capture the variability across the δ_{ij} coefficients.

The residual error variance, σ_1^2, in model 1b is assumed to be constant across all communities; however it may be related to community characteristics. To explore this, we first fitted model 1c which is as stated in model 1b but with different residual error variance for monoculture (σ_{1a}^2) and mixture (σ_{1b}^2) communities. We then also fitted model 1d which is as per model 1c but with σ_{1b}^2 allowed to vary according to some mixture community measurement (*e.g.* richness), i.e. the residual error variance for monocultures was σ_{1a}^2 while that for mixtures was $f(z) * \sigma_{1b}^2$, where *f(z)* was a function of some community characteristic measurement *z*. There are many forms *f(z)* could take; we let $f(z) = z^{\gamma}$, where *z* is either a measure of community species richness (species number) or evenness ($E = (2s/(s-1)) * \Sigma_{i<j} P_i P_j$) and γ is a coefficient whose value determines whether *f(z)* is an increasing or decreasing function of the community characteristic. If $\gamma = 0$ then *f(z)*=1 and σ_{1b}^2 is constant across mixtures.

The Generalised Diversity-Interactions model (before any random assumptions are added) can take a variety of different forms (Kirwan et al. 2007, Kirwan et al. 2009, Connolly et al. 2013), for example,

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} P_i P_j^{\theta_1} + \varepsilon \qquad \text{(Model 2a)}$$

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{wfg1} \sum_{\substack{i,j=1 \\ i<j}}^{t} P_i P_j^{\theta_1} + \delta_{wfg2} \sum_{\substack{i,j=t+1 \\ i<j}}^{s} P_i P_j^{\theta_1} + \delta_{bfg} \sum_{\substack{i\in(1,..,t) \\ j\in(t+1,..,s)}} P_i P_j^{\theta_1} + \varepsilon \qquad \text{(Model 3a)}$$

Models 2a and 3a are versions of model 1a with some coefficient constraints applied (Kirwan et al. 2009); model 2a constrains all δ_{ij} to equal δ_{av}, model 3a assumes two functional groupings of species and constrains the δ_{ij} among the t (s-t) species from group 1 (2) to equal δ_{wfg1} (δ_{wfg2}), and the δ_{ij} for pairs of species with one from each group to equal δ_{bfg}, where *wfg* and *bfg* stand for 'within functional group' and 'between functional groups' respectively. For simplicity, model 3a is specified for two functional groups but can be modified for more functional groups as required. Each of these models can be extended to a GDIM model as described for model 1a; the coefficients and their descriptions are listed in Table 3.1.1 with full algebraic specifications in Appendix 3.1.1. Note while models 1a and 2a differ, models 1b-1d are equivalent to models 2b-2d respectively.

DATA SETS

We tested our methods on two data sets. The first data set, referred to as the Jena data set, was from a nine-species grassland experiment in Jena, Germany (Roscher et al. 2004). There were 206 communities assembled with various levels of species richness (1, 2, 3, 4, 6 or 9 species) and across four blocks based on soil

Table 3.1.1: List and description of the fixed and random coefficients in each model 1a-d, 2a-d and 3a-d. Full model specifications are in Appendix 3.1.1. Note that models 1a and 2a differ but models 1b, 2b are the same (as are models 1c and 2c and models 1d and 2d).

Model	1	2	3
a	*Fixed* Identity effects (β_i) Treatment effects (α) Pairwise interactions (δ_{ij}) Power on P_iP_j (θ_1)	*Fixed* Identity effects (β_i) Treatment effects (α) Average pairwise interactions (δ_{av}) Power on P_iP_j (θ_1)	*Fixed* Identity effects (β_i) Treatment effects (α) Functional group pairwise interactions (δ_{wfg1}, δ_{wfg2}, δ_{bfg}) Power on P_iP_j (θ_1)
b	*Fixed* Identity effects (β_i) Treatment effects (α) Average pairwise interactions (δ_{av}) Powers on P_iP_j (θ_1,θ_2) *Random* Pairwise interactions (d_{ij}) *Assumptions* $\varepsilon \sim N(0, \sigma_1^2)$ $d_{ij} \sim N(0, \sigma_2^2)$		*Fixed* Identity effects (β_i) Treatment effects (α) Functional group pairwise interactions (δ_{wfg1}, δ_{wfg2}, δ_{bfg}) Powers on P_iP_j (θ_1,θ_2) *Random* Pairwise interactions (d_{ij}) *Assumptions* $\varepsilon \sim N(0, \sigma_1^2)$ $d_{ij} \sim N(0, \sigma_2^2)$
c	*Additional assumptions* $\varepsilon \sim N(0, \sigma_{1a}^2)$ for monocultures $\varepsilon \sim N(0, \sigma_{1b}^2)$ for mixtures		
d	*Additional assumptions* $\varepsilon \sim N(0, f(z)*\sigma_{1b}^2)$ for mixtures		

characteristics. The species were classified into three functional groups (grasses, legumes and non-legume herbs), and aboveground biomass was the ecosystem function measured. The second data set, referred to as the Bell data set, was from a 72-species bacterial 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). The ecosystem function measured was the average daily respiration rate (over a period of 28 days) of the bacterial community. Additional information on both experiments can be found in Appendix 3.1.2.

ANALYSIS

The first step in the analysis was to select a 'baseline' model for the data sets. A set of candidate Generalised Diversity-Interactions models (including models 1a, 2a and 3a) were tested and the best was selected using likelihood ratio tests for comparisons involving the non-linear coefficient θ_1 and F-tests otherwise. If the model with the estimate of θ_1 was not a significant improvement over the model with its value set to 1, then the simpler model with θ_1=1 was used. These models were fitted using least squares, maximum likelihood or profile maximum likelihood as appropriate using the software package SAS 9.3 (SAS Institute Inc.). Each chosen baseline model was then extended to a Generalised Diversity-Interactions Mixed (GDIM) model to test for the inclusion of the pairwise interaction random effects and for an effect of community structure on the residual error term using likelihood ratio tests. When testing a variance term against zero, p-values were divided by 2 (Littell et al. 2006, pages 752-3) to avoid issues associated with hypothesis testing close to a boundary space (Self

and Liang 1987). These models were fitted using restricted maximum likelihood (see, for example, Pawitan 2001) using the software package SAS 9.3 (SAS Institute Inc). For those models that included θ_1, θ_2 and/or γ, these coefficients were estimated using profile likelihood independently of each other. Using the final models, we predicted ecosystem function across a range of characteristics for each data set.

Results

The baseline model selected for the Jena grassland data set (Roscher et al. 2004) was model 3a, the functional group effect model, with $\theta_1 = 1$ (Appendix 3.1.3). Extending to the GDIM model 3b provided a significant improvement over model 3a (Table 3.1.2a, M3a versus M3b, p=0.008). Including a profiled estimate of θ_2 did not improve the model fit further ($\hat{\theta}_2$=0.65, p=0.234, testing for a difference from 1 using a likelihood ratio test) and so θ_2 was set to 1. Fitting different residual error variances to monocultures and mixtures (model 3c) did not improve the model fit further nor did allowing the residual error variance to differ across mixtures (model 3d) (Table 3.1.2a), thus the finally selected model for the Jena data set was the GDIM model that included within and between functional group interactions, and included random pairwise interactions:

$$y = \alpha_k + \sum_{i=1}^{9} \beta_i P_i + \delta_{wfg1} \sum_{\substack{i,j=1 \\ i<j}}^{5} P_i P_j + \delta_{wfg2} P_6 P_7 + \delta_{wfg3} P_8 P_9$$
$$+ \delta_{bfg12} \sum_{\substack{i=1,\ldots,5 \\ j=6,7}} P_i P_j + \delta_{bfg13} \sum_{\substack{i=1,\ldots,5 \\ j=8,9}} P_i P_j + \delta_{bfg23} \sum_{\substack{i=6,7 \\ j=8,9}} P_i P_j + \sum_{\substack{i,j=1 \\ i<j}}^{9} d_{ij} P_i P_j + \varepsilon$$

where the α_k are block effects, $d_{ij} \sim N(0, \sigma_2^2)$ and $\varepsilon \sim N(0, \sigma_1^2)$. (Model 3b)

Table 3.1.2. Generalised Diversity-Interactions Mixed (GDIM) model fits and tests for (a) the Jena and (b) the Bell data sets.

	# c	-2LL	Comparison	Testing	LRT	p-value
(a) Jena data set						
Model 3a (for three functional groups)	18	2394.5				
Model 3b (θ_1=1, θ_2=1)	19	2388.8	M3a vs M3b	$\sigma_2 = 0$	5.7	0.008
Model 3c	20	2385.3	M3b vs M3c	$\sigma_{1a} = \sigma_{1b}$	3.5	0.061
Model 3d_richness (γ profiled)	21	2383.8	M3c vs M3d_r	*f(z)* for richness	1.5	0.221
Model 3d_evenness (γ profiled)	21	2384.1	M3c vs M3d_e	*f(z)* for evenness	1.2	0.273
(b) Bell data set						
Model 2a	75	6464.1				
Model 2b (θ_1 profiled, θ_2=1)	76	6463.2	M2a vs M2b	$\sigma_2 = 0$	0.9	0.171
Model 2c	77	6463.1	M2b vs M2c	$\sigma_{1a} = \sigma_{1b}$	0.1	0.752
Model 2d_richness (γ profiled)	78	6463.2	M2c vs M2d_r	*f(z)* for richness	0	1.000
Model 2d_evenness (γ profiled)	78	6463.2	M2c vs M2d_e	*f(z)* for evenness	0	1.000

Footnote: # c = number of coefficients in model, -2LL = -2 Log likelihood value from REML model fit, LRT=likelihood ratio test value. For the Jena data set, the profiled estimate of the γ coefficient for Model 3d_richness was -0.3 and for Model 3d_evenness was -0.6. For the Bell data set, the profiled estimate of the θ_1 coefficient was 0.79 in models 2b to 2d and the profiled estimate of the γ coefficient for Model 2d_richness was 0 and for Model 3d_evenness was 0.025.

The GDI model 2a with an average pairwise interaction effect and $\hat{\theta}_1 = 0.79$, was selected as the baseline model for the Bell bacterial data set (Appendix 3.1.4) The fit of the baseline model was not improved by extending it to any of the GDIM models (Table 3.1.2b). A profiled estimate of θ_2 was tested in model 2b and the value with the smallest likelihood was $\theta_2 = 1$, therefore all models 2b-2d had θ_2 set to 1 (Table 3.1.2b). The final model selected for the Bell data was the GDI model 2a which included the average interaction effect and the power coefficient θ_1, but no random interaction terms:

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} (P_i P_j)^{0.79} + \varepsilon \text{ , where } \varepsilon \sim N(0, \sigma_1^2),$$

thus, there was no evidence that the residual error variance changed across richness or community composition. It would not have been possible to fit the full pairwise interactions model here (that would require the estimation of 2557 coefficients for which there is not enough data). Our result is therefore quite powerful because it provides evidence that it was not necessary to fit a full pairwise interactions model since all significant variability among the true δ_{ij} terms was captured by the two coefficients δ_{av}, and θ_1 without the need to actually fit the full δ_{ij} model.

Figure 3.1.2 shows ecosystem function predictions with raw data superimposed for each dataset. An example of each of the GDIM models fitted to the Jena data set (models 3b-d) is given in Appendix 3.1.5, the Jena data set is detailed in Appendix 3.1.6 and SAS code to fit each model is in Appendix 3.1.7.

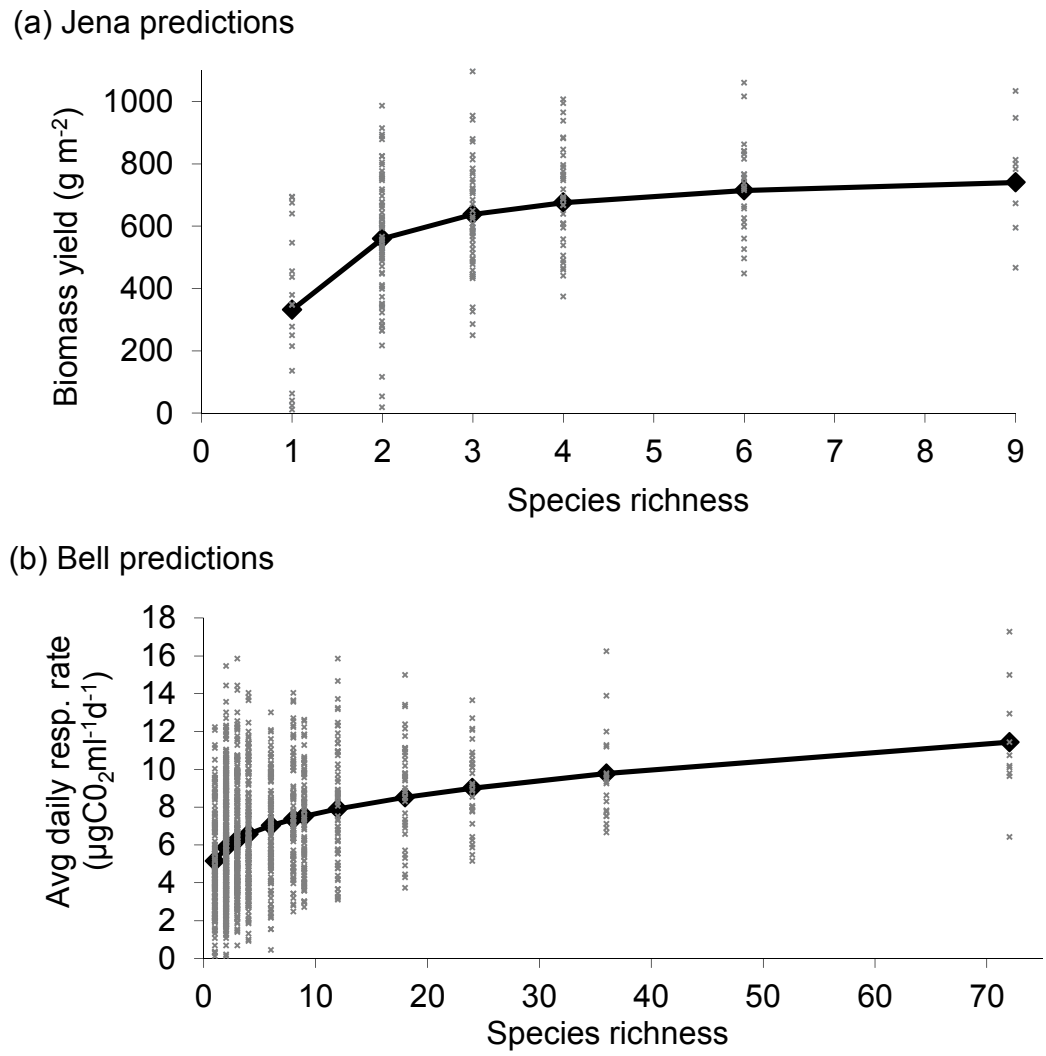

Fig. 3.1.2. Predicted ecosystem response (—◆—) and the raw data (x) versus richness for (a) the Jena and (b) the Bell data sets. The predicted mean response is averaged across all possible community types at each level of richness.

Discussion

The purpose of developing the Generalised Diversity-Interactions Mixed (GDIM) model was twofold; the first aim was to create a parsimonious model which could potentially replace a large number of fixed coefficients for describing diversity

effects with a smaller number of fixed and random coefficients combined. For data sets where fitting a full Diversity-Interactions model estimating all species interactions is impossible (*e.g.* the Bell dataset), or not desirable due to the difficulty extracting biological information from a large number of coefficients (*e.g.* the Jena dataset), the random coefficients in the GDIM model may facilitate using only a small number of fixed coefficients to describe diversity effects but still ensure that standard errors include any remaining uncertainty due to individual pairwise interactions. The second aim was to provide a lack of fit test for the fixed effects models where a reduced number of parameters are used to describe the diversity effects. In the event that the random effects are not needed, this lack of fit test can validate the inference from the reduced fixed effects model without the need to test against the full pairwise Diversity-Interactions model with all interactions fitted as fixed coefficients. We also provide a means to test if the residual error variance differs across varying community conditions. Specifically, it can be tested whether there is a difference between the residual error variances for monocultures and mixtures (Model 1c, is $\sigma^2_{1a} = \sigma^2_{1b}$?) and if the residual error variation for mixtures is dependent on some community characteristic such as richness or evenness (Model 1d, is $f(z)$=1?).

As with previous models for the BEF relationship (Mulder et al. 2002, Mulder et al. 2004, Hooper et al. 2005, Lanta and Lepš 2006, Kirwan et al. 2007), the GDIM model allows exploring of the conditions under which ecosystem function will be maximised or minimized. Current methods for modelling the BEF relationship use many different community characteristics, such as species richness (Hooper et al. 2005, Spehn et al. 2005), functional grouping (Mulder et al. 2002, Cadotte et al. 2009), evenness (Cardinale et al. 2002, Finn et al. 2013) and the

presence / absence of individual species (Bell et al. 2009) to examine the relationship between species and ecosystem functioning. The Diversity-Interactions model (Kirwan et al. 2007, Kirwan et al. 2009), the Generalised Diversity-Interactions model (Connolly et al. 2013) and the GDIM model presented here implicitly test for the effects of a range of different community characteristics, such as species identity, species initial sown proportions, species interactions, species richness and evenness to examine the BEF relationship. Thus, when using our GDIM model to develop a more complete understanding of the BEF relationship, the benefits of each of the above methods are included, with the added benefit of providing a lack of fit test for a small number of coefficients describing diversity effects, and when the random effects are significant, ensuring that the extra uncertainty is built into standard errors improving inference.

The GDIM models presented here investigated including random terms for species' interactions. It would also be possible to assume that the identity effects follow a random probability distribution (Appendix 3.1.8). This extension could be useful in a particularly species-rich ecosystem as there may be difficulty estimating all identity effect coefficients (β_i) but there may also be biologically motivated fixed effects solutions that would also reduce the number of coefficients that need to be estimated in a sensible manner.

The GDIM model offers a modelling approach that is parsimonious, versatile and informative. The method has the ability to greatly reduce the number of coefficients required to model the effects of species' interactions on ecosystem function, thereby simplifying the description of species-rich systems in particular. Our approach also allows us to test various assumptions as to how the residual error variance may be related to community structure. Ensuring the correct residual error

variance structure, along with the inclusion of random effects to capture variation in species interactions additional to the fixed effects, provides improved standard errors with which to test fixed effects, thus improving inference.

References

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.

Cadotte, M. W., J. Cavender-Bares, D. Tilman, and T. H. Oakley. 2009. Using phylogenetic, functional and trait diversity to understand patterns of plant community productivity. PLoS One **4**:e5695.

Cardinale, B. J., M. A. Palmer, and S. L. Collins. 2002. Species diversity enhances ecosystem functioning through interspecific facilitation. Nature **415**:426-429.

Connolly, J., T. Bell, T. Bolger, C. Brophy, T. Carnus, J. A. Finn, L. Kirwan, F. Isbell, J. Levine, A. Lüscher, V. Picasso, C. Roscher, M. T. Sebastia, M. Suter, and A. Weigelt. 2013. An improved model to predict the effects of changing biodiversity levels on ecosystem function. Journal of Ecology **101**:344-355.

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., L. Kirwan, J. Connolly, M. T. Sebastià, A. Helgadóttir, O. H. Baadshaug, G. Bélanger, A. Black, C. Brophy, R. P. Collins, J. Čop, S. Dalmannsdóttir, I. Delgado, A. Elgersma, M. Fothergill, B. E. Frankow-Lindberg, A. Ghesquiere, B. Golinska, P. Golinski, P. Grieu, A. M. Gustavsson, M. Höglind, O. Huguenin-Elie, M. Jørgensen, Z. Kadziuliene, P. Kurki, R. Llurba, T. Lunnan, C. Porqueddu, M. Suter, U. Thumm, and A. Lüscher. 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.

Hector, A., B. Schmid, C. Beierkuhnlein, M. C. Caldeira, M. Diemer, P. G. Dimitrakopoulos, J. A. Finn, H. Freitas, P. S. Giller, J. Good, R. Harris, P. Hogberg, K. Huss-Danell, J. Joshi, A. Jumpponen, C. Korner, P. W. Leadley, M. Loreau, A. Minns, C. P. H. Mulder, G. O'Donovan, S. J. Otway, J. S. Pereira, A. Prinz, D. J. Read, M. Scherer-Lorenzen, E. D. Schulze, A. S. D. Siamantziouras, E. M. Spehn, A. C. Terry, A. Y. Troumbis, F. I. Woodward, S. Yachi, and J. H. Lawton. 1999. Plant diversity and productivity experiments in European grasslands. Science **286**:1123-1127.

Hillebrand, H., and B. Matthiessen. 2009. Biodiversity in a complex world: consolidation and progress in functional biodiversity research. Ecology Letters **12**:1405-1419.

Hooper, D. U., F. S. Chapin, J. J. Ewel, A. Hector, P. Inchausti, S. Lavorel, J. H. Lawton, D. M. Lodge, M. Loreau, S. Naeem, B. Schmid, H. Setala, A. J. Symstad, J. Vandermeer, and D. A. Wardle. 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. Sebastià. 2009. Diversity-interaction modeling: estimating contributions of species identities and interactions to ecosystem function. Ecology **90**:2032-2038.

Kirwan, L., A. Lüscher, M. T. Sebastià, J. A. Finn, R. P. Collins, C. Porqueddu, A. Helgadóttir, O. H. Baadshaug, C. Brophy, C. Coran, S. Dalmannsdóttir, I. Delgado, A. Elgersma, M. Fothergill, B. E. Frankow-Lindberg, P. Golinski, P. Grieu, A. M. Gustavsson, M. Höglind, O. Huguenin-Elie, C. Iliadis, M. Jørgensen, Z. Kadziuliene, T. Karyotis, T. Lunnan, M. Malengier, S. Maltoni, V. Meyer, D. Nyfeler, P. Nykanen-Kurki, J. Parente, H. J. Smit, U. Thumm, and J. Connolly. 2007. Evenness drives consistent diversity effects in intensive grassland systems across 28 European sites. Journal of Ecology **95**:530-539.

Lanta, V., and J. Lepš. 2006. Effect of functional group richness and species richness in manipulated productivity–diversity studies: a glasshouse pot experiment. Acta Oecologica **29**:85-96.

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

Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, M. A. Huston, D. Raffaelli, B. Schmid, D. Tilman, and D. A. Wardle. 2001. Ecology - Biodiversity and ecosystem functioning: Current knowledge and future challenges. Science **294**:804-808.

Mulder, C., A. Jumpponen, P. Högberg, and K. Huss-Danell. 2002. How plant diversity and legumes affect nitrogen dynamics in experimental grassland communities. Oecologia **133**:412-421.

Mulder, C. P. H., E. Bazeley-White, P. G. Dimitrakopoulos, A. Hector, M. Scherer-Lorenzen, and B. Schmid. 2004. Species evenness and productivity in experimental plant communities. Oikos **107**:50-63.

Pawitan, Y. 2001. In all likelihood: Statistical modelling and inference using likelihood. Clarendon Press, Oxford.

Petchey, O. L., and K. J. Gaston. 2006. Functional diversity: back to basics and looking forward. Ecology letters **9**:741-758.

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.

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.

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., A. Hector, J. Joshi, M. Scherer-Lorenzen, B. Schmid, E. Bazeley-White, C. Beierkuhnlein, M. C. Caldeira, M. Diemer, P. G. Dimitrakopoulos, J. A. Finn, H. Freitas, P. S. Giller, J. Good, R. Harris, P. Hogberg, K. Huss-Danell, A. Jumpponen, J. Koricheva, P. W. Leadley, M. Loreau, A. Minns, C. P. H. Mulder, G. O'Donovan, S. J. Otway, C. Palmborg, J. S. Pereira, A. B. Pfisterer, A. Prinz, D. J. Read, E. D. Schulze, A. S. D. Siamantziouras, A. C. Terry, A. Y. Troumbis, F. I. Woodward, S. Yachi, and J. H. Lawton. 2005.

Ecosystem effects of biodiversity manipulations in European grasslands. Ecological Monographs **75**:37-63.

Tilman, D. 1999. Diversity and production in European grasslands. *Science* **286**:1099-1100.

Tilman, D., P. B. Reich, and J. M. H. Knops. 2006. Biodiversity and ecosystem stability in a decade-long grassland experiment. Nature **441**:629-632.

Tilman, D., D. Wedin, and J. Knops. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. Nature **379**:718-720.

Chapter 4

Modelling the multifunctional biodiversity and ecosystem function relationship

Collaborators: Caroline Brophy, John Connolly, Laura Kirwan, John A. Finn and Forest Isbell.

INTRODUCTION

Analysis of a single ecosystem function may not provide a full assessment of the biodiversity and ecosystem function (BEF) relationship. Different functions may require different species and so, by observing only a single function, we may be underestimating the ecosystem requirements and the effects of biodiversity loss. Recent research has examined the BEF relationship for multiple ecosystem functions (multifunctionality). The focus of this chapter is to discuss current multifunctional methods, to highlight potential improvements and to develop a multivariate model for analysing the multifunctional BEF relationship.

Current multifunctional methods focus on reducing the complexity of analysing multiple functions, through methods such as multiple separate univariate analyses or the use of metrics. In section 4.1 we discuss the current multifunctional methods available, how each method is implemented and any potential difficulties.

The aim of this section is to highlight areas where we wish to contribute to multifunctional BEF research.

In section 4.2 we aim to address some of the problems we highlighted in section 4.1 for the averaging approach, one of the current methods for analysing the multifunctional BEF relationship. The averaging approach averages all ecosystem functions of interest into a single metric value. Although this simplifies analysis it does not account for how the functions within the ecosystem are behaving, i.e. whether all functions are performing similarly of whether one or more functions are outperforming the others. In section 4.2 we develop a scaled averaging metric that penalises ecosystems where the function responses are more variable. This scaled metric improves upon the averaging metric and provides greater information about the underlying ecosystem functions.

Many of the current multifunctional methods, including the scaled averaging metric developed in section 4.2, suffer loss of information about the ecosystem functions. In section 4.3, we develop the multivariate Diversity-Interactions model which allows for the analysis of multiple ecosystem functions simultaneously without the loss of information about the individual functions. By carrying out a multivariate analysis of the ecosystem functions, the model allows us to examine the effect of community characteristics on each function, the correlation between functions and how the effects of community characteristics change across functions.

Section 4.1

Biodiversity and ecosystem multifunctionality: a review of current statistical methods.

Introduction

Ecosystems provide multiple functions simultaneously (multifunctionality) which may interrelate and interact with one another. Many current methods for modelling the biodiversity and ecosystem function (BEF) relationship analyse a single ecosystem function (Tilman & Downing 1994; Cardinale *et al.* 2002; Hector & Bagchi 2007; Fox & Harpole 2008; Hillebrand & Matthiessen 2009; Kirwan *et al.* 2009; Hector *et al.* 2010), such as the aboveground biomass produced in grassland experiments. Methods analysing a single ecosystem function may not provide a full assessment of the BEF relationship. For example, when analysing a single ecosystem function, such as biomass produced, a saturation effect as species richness increases has been observed in many cases (Hector *et al.* 1999; Cardinale *et al.* 2002), implying that additional species contribute less to the ecosystem function as richness increases. Although this is often the case when analysing a single function this may not be the case when multiple functions are considered simultaneously; different functions may require different species, reducing the possibility of observing a saturation effect as richness increases. Analysis based on multiple functions is important to the understanding of the BEF relationship as it allows for a fuller analysis and better estimation of the effect of changing community characteristics on the ecosystem, as well as providing us with a better ability to predict ecosystem

responses (Bennett *et al.* 2009). In addition, given that most ecosystems provide multiple functions simultaneously and in the face of serious global declines of biodiversity, understanding how to maintain the provisioning of multiple functions may be critical to human welfare.

Using multifunctional analysis to examine the biodiversity and ecosystem function relationship is a novel area of research in ecology that is gaining much interest. Initial work examining the multifunctional BEF relationship has examined how the ability of an ecosystem to maintain multiple functions is affected by species richness (Hector & Bagchi 2007; Gamfeldt *et al.* 2008; Mouillot *et al.* 2011; Maestre *et al.* 2012a; Maestre *et al.* 2012b; Byrnes *et al.* 2014) and temporal and spatial factors (Zavaleta *et al.* 2010; Isbell *et al.* 2011). A number of key methods for examining the BEF relationship for multiple functions have been developed, namely:

1. The threshold method (Gamfeldt *et al.* 2008; Zavaleta *et al.* 2010; Byrnes *et al.* 2014).
2. The overlap method (Hector & Bagchi 2007; Isbell *et al.* 2011).
3. The averaging method (Mouillot *et al.* 2011; Maestre *et al.* 2012a; Maestre *et al.* 2012b).
4. Multiple univariate analyses (Allan *et al.* 2013; Cardinale *et al.* 2013; Orwin *et al.* 2014).

Byrnes *et al.* (2014) provide a discussion of the methods currently in use for examining the multifunctional BEF relationship. In this section we aim to briefly review the methodological aspects of multifunctionality research to date, discussing each of the above methods, how they are implemented and any potential benefits or drawbacks the methods may have.

The threshold method

The threshold method was developed by Gamfeldt *et al.* (2008) to examine how species loss affects multiple ecosystem functions. The method was also used by Zavaleta *et al.* (2010) to estimate the number of species required to maintain a minimum threshold of functionality across multiple functions in a long-term grassland experiment. Byrnes *et al.* (2014) then extended the method to increase the range of thresholds considered during analysis. To implement the threshold method a maximum level of functioning is chosen for each ecosystem function, for example Gamfeldt *et al*. (2008) choose their level to be the maximum observed monoculture response for each individual function. A threshold of this maximum value is then chosen, *e.g.* to exceed a 50% threshold a community must have an ecosystem function value of at least 50% of the maximum level for that function. The community is said to be able to maintain the function if the community's value is at the threshold or greater. Gamfeldt *et al.* (2008) used the threshold method to examine multifunctionality and the BEF relationship for five data sets, one containing grassland plant data, one containing bacteria data and three containing marine plant data. They used the monoculture responses to construct mixture communities and, having set a threshold of 50%, they randomly deleted a number of species and calculated the probability the ecosystem could sustain the functions given the species loss. The probability was calculated by simulating random species loss at each combination of species richness and ecosystem function 10,000 times and calculating the proportion of these simulations where the ecosystem maintained the functions to at least the threshold level. Zavaleta *et al.* (2010) used the threshold method to calculate the proportion of communities which could achieve a given

threshold of ecosystem function across a range of thresholds (40%, 50% and 60% of the maximum level), number of functions (1 to 8 functions) and species richness (1-16 species) for three separate years (1998, 2000 and 2002) of a long-term biodiversity grassland experiment. They then used these proportions to estimate minimum species requirements such that 50% of communities at a given threshold and year could maintain a given number of functions. To maintain the multiple functions simultaneously both studies required that the ecosystem be able to maintain each individual function separately and found that high species richness increased the probability an ecosystem could maintain multiple functions. Byrnes *et al.* (2014) developed the multiple threshold approach which involves analysing the effect of species diversity on the ability of the system to maintain multiple functions across the full range of thresholds (*i.e.* for 0% to 100% of the maximum value) to help reduce the information loss that may occur from analysing the relationship based on a single or small number of thresholds. Byrnes *et al.* (2014) found that by using the full range of thresholds they were able to develop a better understanding of the changing effect of species diversity on ecosystem multifunctionality as the threshold value changed. Gamfeldt *et al.* (2008) found that there was a higher probability that the species rich communities could maintain the multiple functions after the simulated species loss whereas Zavaleta *et al.* (2010) found that as the number of functions being maintained increased, the minimum number of species required to maintain them also increased. Byrnes *et al.* (2014) found that the effect of species diversity on ecosystem multifunctionality was dependent on the threshold chosen.

The threshold method is a simple method to implement across different types of ecosystems, as in Gamfeldt *et al.* (2008), and different ecosystem variables, as in

Zavaleta *et al.* (2010). It is a very useful method for performing examinations of the ability of the ecosystem to maintain multiple functions however the method has some disadvantages such as a loss of information. The threshold method converts the ecosystem function to a binary response, *i.e.* is the functioning above or below the threshold level; because of this, some information about the ecosystem function is lost, for example the threshold method does not provide information about by how much the ecosystem function exceeded or failed to reach the threshold. Zavaleta *et al.* (2010) used multiple threshold levels and Byrnes *et al.* (2014) further extended the number of thresholds used to help to compensate for this information loss. Their work showed that the choice of threshold can greatly influence the outcome of the study.

The overlap method

The overlap method (Hector & Bagchi 2007; Isbell *et al.* 2011) identifies the species which affect each ecosystem function, focusing on the species which have desirable effects on the function, *i.e.* those that increase (decrease) ecosystem function where high (low) ecosystem function is desirable. The method initially identifies which species have a significant effect on each function, then subsets these species lists to only those species which have desirable effects on the ecosystem function. The overlap between functions, *i.e.* the number of species which desirably affect any pair of functions, can then be calculated as

$$O_{ij} = \frac{\left|E_i \cap E_j\right|}{0.5(\left|E_i\right| + \left|E_j\right|)}$$

where E_i is the number of species which desirably affect the i^{th} function. E_i, E_j are values greater than 0 as functions must be affected by at least one of the species present in the ecosystem. The mean overlap between functions is then used to calculate the predicted value for the number of species required to maintain a number of functions:

$$S_E = \sum_{i=1}^{E} \binom{E}{i} \bar{x}(-\bar{o})^{(i-1)}$$

where S_E is the predicted species number, E is the number of ecosystem functions, $\bar{x}$ is the mean number of influential species per function and $\bar{o}$ is the mean overlap between functions. Hector and Bagchi (2007) found that, as the number of functions increased, the number of species required to maintain multifunctionality increased. Isbell *et al.* (2011) extended the overlap method to examine how the species affecting the ecosystem functioning change across time, location, environmental changes and number of functions being maintained. By examining the ecosystems across these four factors Isbell *et al.* (2011) were able to identify that the species maintaining ecosystem functioning did in fact change and so greater species richness would be required to consistently maintain ecosystem functioning.

The overlap method allows for the identification of which species drive ecosystem functioning and whether different species are required to maintain multiple functions. The method also allows for the calculation of the strength and direction (*i.e.* a positive, neutral or negative effect) of the species effect. A drawback of the method however is that for a species to be considered important to two or more functions, the direction of the species effect must be the same for all functions under consideration. This means that the overlap method cannot currently interpret when a species has a positive effect on one function and a negative effect on another.

The averaging method

The averaging method (Mouillot *et al.* 2011; Maestre *et al.* 2012a; Maestre *et al.* 2012b) combines a number of standardised functions into a single average value which allows for the application of well known univariate methods to explore the multifunctionality of the system. Maestre *et al.* (2012a; 2012b) found that the species richness of the ecosystem had a significant positive correlation with the average metric, *i.e.* the average metric increased as the species richness increased, implying that the ability of the ecosystem to maintain higher ecosystem function values increased with species richness. Wagg *et al.* (2014) found a similar result, *i.e.* that higher species richness in the soil community corresponded to a higher average functional metric value.

Although the averaging method is simple to calculate it has a number of significant drawbacks. The average metric is not easily interpreted in terms of the multifunctionality of the ecosystem. A high metric value means high functional values for the ecosystem on average. However, since the metric is calculated as the average of the ecosystem functions, conclusions cannot be drawn from the metric as to how the ecosystem is maintaining each individual function. If only some functions are performing very well, taking the average may hide functions for which the ecosystem is performing poorly. As it is currently implemented the averaging method also does not consider whether high or low functional values are desirable for each function, for example, it may be desirable to have high biomass produced but low nitrate leaching in a grassland system.

Analysis using univariate methods

The final method most commonly used is a univariate analysis to analyse the effect of community characteristics on each function individually and then combining the resulting information to compare the effects across functions in a qualitative manner. This method has no exact definitive steps to follow as the analyst chooses which univariate methods to use and how to collate the resulting information. This method for analysing the multifunctional BEF relationship has been used for large data sets by Allan *et al.* (2013) and Cardinale *et al.* (2013) to examine the effect of species richness across large temporal and spatial scales. Both studies found that biodiversity had a significant effect on the ability of the ecosystem to maintain multiple functions.

Although univariate analysis offers a method for highlighting how each individual function is affected by community characteristics the method requires multiple analyses and does not measure the correlations among functions or allow for a quantitative analysis or formal test of the effect of a species across functions.

Final remarks

The four methods presented here offer different approaches to analysing the complex problem of multifunctional ecosystems. We have briefly discussed the main benefits and drawbacks of each of the current methods, similar to what was presented by Byrnes *et al.* (2014). The purpose of this discussion of current multifunctional ecosystem analysis methods was to highlight some areas which we hope to address in the sections that follow in this chapter, namely ways to address the information

loss which is suffered by most multifunctional methods. Byrnes *et al.* (2014) have attempted to address the information loss for the threshold method by extended the method to cover all possible thresholds but other methods, such as the averaging method, suffer from a serious loss of information which has yet to be examined. Additionally, no method has so far built the correlations among functions into an analysis that can test specific species or other effects across multiple functions. Multifunctional ecosystem analysis is a new and expanding area of research and, as such, has great potential for the development of new concepts.

References

Allan, E., Weisser, W., Fischer, M., Schulze, E., Weigelt, A., Roscher, C. *et al.* (2013). A comparison of the strength of biodiversity effects across multiple functions. *Oecologia*, 173, 223-237.

Bennett, E., Peterson, G. & Gordon, L. (2009). Understanding relationships among multiple ecosystem services. *Ecology Letters*, 1394-1404.

Byrnes, J., Gamfeldt, L., Isbell, F., Lefcheck, J., Griffin, J., Hector, A. *et al.* (2014). Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions. *Methods in Ecology and Evolution*, 5, 111-124.

Cardinale, B., Gross, K., Fritschie, K., Flombaum, P., Fox, J., Rixen, C. *et al.* (2013). Biodiversity simultaneously enhances the production and stability of community biomass, but the effects are independent. *Ecology*, 94, 1697-1707.

Cardinale, B., Palmer, M. & Collins, S. (2002). Species diversity enhances ecosystem functioning through interspecific facilitation. *Nature*, 415, 426-429.

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

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

Hector, A. & Bagchi, R. (2007). Biodiversity and ecosystem multifunctionality. *Nature*, 448, 188-191.

Hector, A., Hautier, Y., Saner, P., Wacker, L., Bagchi, R., Joshi, J. *et al.* (2010). General stabilizing effects of plant diversity on grassland productivity through population asynchrony and overyielding. *Ecology*, 91, 2213-2220.

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

Hillebrand, H. & Matthiessen, B. (2009). Biodiversity in a complex world: consolidation and progress in functional biodiversity research. *Ecology Letters*, 12, 1405-1419.

Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W., Reich, P. *et al.* (2011). High plant diversity is needed to maintain ecosystem services. *Nature*, 477, 199-202.

Kirwan, L., Connolly, J., Finn, J.A., Brophy, C., Luscher, A., Nyfeler, D. *et al.* (2009). Diversity-Interaction modeling: estimating contributions of species identities and interactions to ecosystem function. *Ecology*, 90, 2032-2038.

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

Maestre, F.T., Quero, J.L., Gotelli, N.J., Escudero, A., Ochoa, V., Delgado-Baquerizo, M. *et al.* (2012b). Plant species richness and ecosystem multifunctionality in global drylands. *Science*, 335, 214-218.

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

Orwin, K., Ostle, N., Wilby, A. & Bardgett, R. (2014). Effects of species evenness and dominant species identity on multiple ecosystem functions in model grassland communities. *Oecologia*, 174, 979-992.

Tilman, D. & Downing, J. (1994). Biodiversity and stability in grasslands. *Nature*, 367, 363-365.

Wagg, C., Bender, S., Widmer, F. & van der Heijden, M. (2014). Soil biodiversity and soil community composition determine ecosystem multifunctionality. *Proceedings of the National Academy of Sciences of the United States of America*, 111, 5266-5270.

Zavaleta, E., Pasari, J., Hulvey, K. & Tilman, G. (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.

Section 4.2

This section has been formatted for PLoS ONE.

Testing multifunctionality in the biodiversity and ecosystem functioning relationship: Critique of the averaging method and development of an improved metric.

Áine Dooley[1], Forest Isbell[2], John Connolly[3], Laura Kirwan[4,6], John A. Finn[5], Caroline Brophy[1].

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

2. Department of Ecology, Evolution, and Behavior, University of Minnesota, St Paul, Minnesota 55108, USA

3. School of Mathematical Sciences, Ecological and Environmental Modelling Group, University College Dublin, Dublin 4, Ireland

4. Department of Chemical and Life Science, Waterford Institute of Technology, Cork Road, Waterford, Ireland

5. Teagasc Environment Research Centre, Johnstown Castle, Co. Wexford, Ireland

6. Present address: UCD Institute of Food and Health, UCD, Belfield, Dublin 4, Ireland

Abstract

Analysing multiple ecosystem functions simultaneously (multifunctionality) has become an area of great interest in the ecological community. A number of methods have been developed to examine the biodiversity and ecosystem function relationship for multiple functions, however, some of these have major conceptual issues. Here we critique the averaging method, which analyses multifunctionality by averaging multiple ecosystem function responses into a single metric value. The issues we highlight include loss of information at the individual function level and the possibility of two communities that differ greatly yielding the same average metric value. We also introduce the SAM metric, an improvement on the average metric that includes information about the variability between ecosystem function responses in its calculation. We analyse the average and the SAM metric for data from a grassland biodiversity experiment to demonstrate how the SAM metric improves on the average metric.

Introduction

Multifunctionality is an emerging area for exploring the biodiversity and ecosystem function (BEF) relationship which assesses multiple ecosystem functions simultaneously [1-6] as opposed to a single function as has previously been typical [7-11]. These recent multifunctionality studies have proposed various new analytical methods for examining the multifunctional BEF relationship, e.g. the threshold method [4, 6, 12], the averaging method [13, 14] and the overlap method [1]. These analytical methods have been reviewed and associated pros and cons have been

discussed by Byrnes *et al*. [12]. The purpose of this paper is to further critique the average metric that is used in the averaging method; we will discuss problems with how the method is currently implemented and interpreted. We will also develop an improved metric, called the Scaled Average Multifunctionality or SAM metric. The SAM metric will be illustrated and compared to the average metric using data from the Irish site of the grassland BIODEPTH project [15].

Critique of the averaging method

The averaging method [13, 14, 16, 17] combines a number of standardised ecosystem functions into a single average metric which can then be analysed using univariate techniques. Reducing the multivariate nature of multiple ecosystem function responses to a single dimension reduces the complexity of analysis considerably; however, there are a number of drawbacks to this approach, from both technical and interpretational aspects, that should be carefully considered before using the averaging method in multifunctional BEF analysis. These drawbacks are summarised as follows:

1. Two communities with the same average metric value could have greatly different individual ecosystem function responses. Figure 4.2.1 shows a hypothetical example of three standardised ecosystem function responses for two different communities. For the first community (Figure 4.2.1a) the three standardised ecosystem functions are performing similarly whereas in the second community (Figure 4.2.1b) function A is performing at a much higher rate than either functions B or C. The average metric for both hypothetical

communities is 0.4 and thus they are treated as equivalent by the metric despite individual values varying greatly.

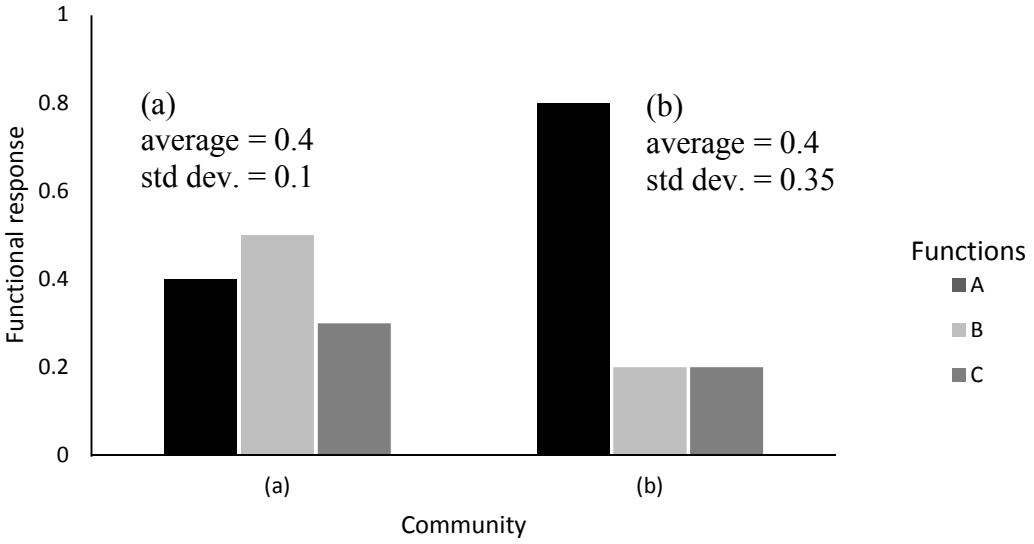

Figure 4.2.1. Two sets of hypothetical community measurements for three functions (A, B and C). In community (a) the functions each have a similar standardised value, while in community (b) function A has a much larger standardised value than B and C. The two communities yield the same average metric value.

2. Removing the multivariate nature of multiple ecosystem function responses means that it is not possible to describe the effect of the community characteristics on any of the individual functions, merely their effect on the average across all functions being considered in the analysis. This results in loss of information on how individual functions respond to varying

community characteristics. Information is also lost on how the functions relate to one another as correlations among them are ignored.

3. When taking an average of multiple functions, it is assumed that each function is equally important, which may not be the case. For example in agronomic grasslands the aboveground biomass produced may be considered more important than other functions, such as the belowground biomass. The averaging method currently calculates the average metric by giving all functions equal weight, regardless of the level of interest in the functions.
4. While it is often desirable to maximise a function, this is not always the case. All functions included in the average metric are assumed to be desirable in the same direction which limits what functions can or cannot be included when calculating it. For example, in a grassland system, high ecosystem functioning for functions such as aboveground biomass or plant nitrogen content is usually considered desirable whereas with other functions, such as nitrate leaching, low ecosystem functioning is usually desirable.

While the averaging method is easy to implement, these four criticisms of it highlight the loss of information that occurs in its practical use and identify how its interpretations may be misleading.

Development of the Scaled Average Multifunctionality (SAM) metric, an improvement on the average metric

Let $(y_{i1},..., y_{ik})$ be the observed ecosystem function responses for the i^{th} community for the k functions recorded. The responses across communities are then transformed to a comparable scale, separately for each function, giving the vector of

responses for the i^{th} community (z_{i1},..., z_{ik}). The average metric for the i^{th} community is the average of its z values. The process of transformation can take many different forms; Maestre *et al.* [13, 14] transformed the functions by converting each response to a percentage of the maximum five percent of responses for the function whereas Wagg *et al.* [17] standardised the responses to have a mean of 0 and standard deviation of 1. We transformed functions where higher output was considered desirable by expressing each value as a percentage of the average of the top 5% of values as in Maestre *et al.* [13, 14]. For functions where a lower output was considered desirable, we transformed the function by computing the maximum value minus the current value, then converting each new response value to a percentage of its new maximum, as presented by Byrnes *et al.* [12]. The Scaled Average Multifunctionality (SAM) metric was then calculated for the i^{th} community by

$$\text{SAM} = \bar{z}_i \Big/ s_i$$

where s_i is the standard deviation across the k transformed responses for the i^{th} community. This is the reciprocal of the coefficient of variation across functions. For communities where the transformed functions are behaving similarly to each other, the standard deviation among functions will be low, leading to a higher SAM metric value, while communities that have a lot of variability among the k responses will be penalised. We assume that the higher the SAM metric value, the higher the ecosystem multifunctionality.

Application of the SAM metric and comparison to the average metric

We computed and analysed the average metric and the SAM metric for data from the Irish site of the grassland BIODEPTH (BIODiversity and Ecosystem

Processes in Terrestrial Herbaceous systems: experimental manipulations of plant communities) project [15]. The experiment consisted of 31 experimental plant assemblages (10 monoculture communities and 21 mixture communities) each planted in two blocks across a range of species richness (1, 2, 3, 4, or 8 species) and functional group richness (1, 2 or 3 groups) levels giving 66 plots in total (one four species and one eight species assemblage appeared in four rather than two plots). The species pool contained 10 species and the five ecosystem functions, recorded in the third year of the experiment, were aboveground plant biomass, belowground plant biomass, unconsumed soil nitrogen, aboveground nitrogen pool and cotton decomposition. Although the experiment originally had a 12 species pool, communities containing two particular species (*Cerastium fontanum* and *Taraxacum officinale*) were omitted from the analysis as the species were only in four plots and did not appear in monoculture. We assumed aboveground plant biomass, belowground plant biomass, aboveground nitrogen pool and cotton decomposition were functions where high output was considered desirable and unconsumed soil nitrogen was a function where low output was considered desirable.

We computed the average and the SAM metric for the data as described above with one alteration; when computing the standard deviation for each community, we used the pooled standard deviation from plots which had the same sown composition, i.e. we computed the variance across the five functions for each of the two communities with the same sown composition, averaged the two variances and computed the square root. The reason for this was to reduce the possibility of an extremely low standard deviation value which may occur by chance if the five standardised values were very close. An initial examination of the plot of the average and SAM metric values for the communities from block 1 (Figure 4.2.2)

showed that the SAM metric can be used to distinguish between communities where the five functions are behaving similarly and those where the ecosystem function responses are more varied. For example community compositions 21 and 24 (highlighted by a dotted line in Figure 4.2.2); the average values for these compositions are almost identical but the SAM metric values differ.

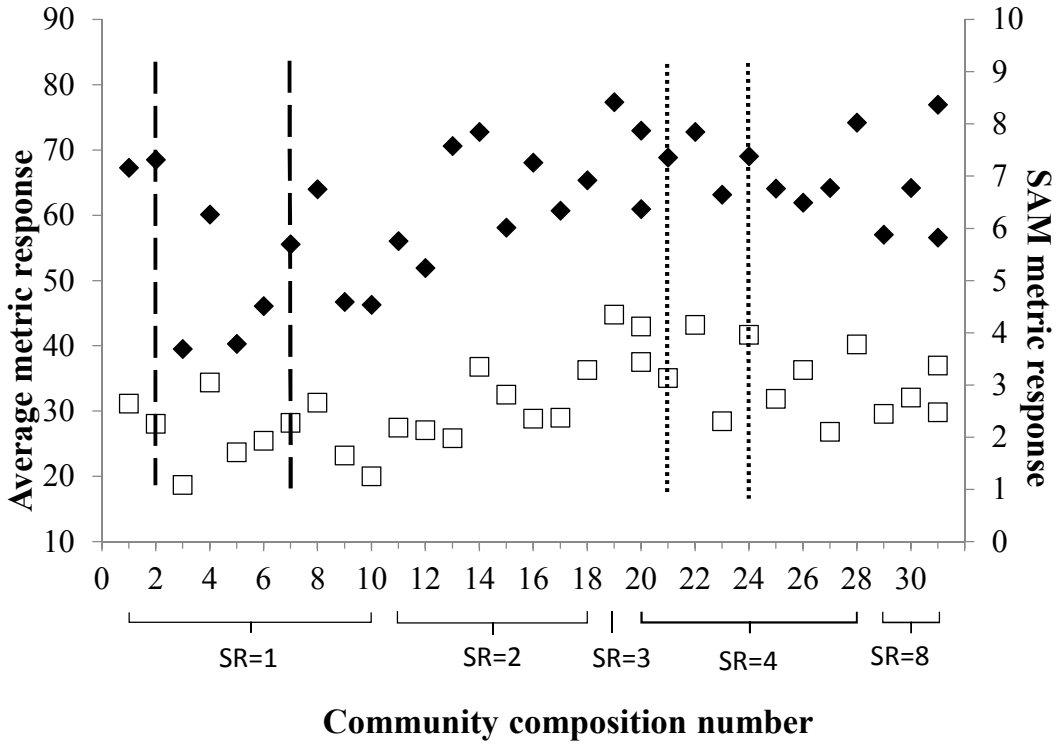

Figure 4.2.2. Plot to compare the average metric and SAM metric responses across 31 community compositions. The plot shows the average (♦, left y-axis) and the SAM (□, right y-axis) metric responses for block 1 of the dataset. The x-axis represents the composition number of the communities. Species richness (SR) is given for each community composition. Plot compositions 21 and 24 are highlighted by a dotted line and compositions 2 and 7 are highlighted by a dashed line.

This is because there is a higher spread of standardised responses for community 21 (values range from 38 to 81) than for community 24 (values range from 57 to 84) (Appendix 4.2.1). The average metric permits trade-offs between functions where we consider a trade-off to occur when one or more functions performing strongly compensates for other functions performing poorly. The SAM metric does not permit trade-offs to the same degree, instead it penalises against variability among the individual responses. For example, the average metric response of community composition 2 (a grass monoculture) is higher than for community composition 7 (highlighted by a dashed line in Figure 4.2.2) which is partly attributable to a high standardised aboveground plant biomass value in composition 2 compensating for low standardised belowground plant biomass, however, when the variability among the five responses is taken into consideration, the SAM metric values for these two communities are very similar (Figure 4.2.2 and Appendix 4.2.1).

We analysed the average metric and the SAM metric using the Diversity-Interactions model [10, 18]

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

where y is the metric value response, P_i is the sown relative abundance of the i^{th} species, α_b is the block effect where b=1, 2, ε is an i.i.d. normally distributed error term and there is a species pool of s species. β_i is the identity effect of the i^{th} species, δ_{ij} is the interaction effect of species i with species j, for j=1, …,s and i<j. The full Diversity-Interactions model has s identity parameters and $s*(s-1)/2$ interaction parameters to be estimated but assumptions about the parameters can reduce these numbers [18], which can be particularly useful for large s. Some examples of these reduced models are the model which assumes no diversity effects (δ_{ij}=0 for all i,j);

the average Diversity-Interactions model, which assumes that each δ_{ij} is equal to an average diversity effect δ_{av}; and the functional group Diversity-Interactions model. The functional group Diversity-Interactions model assumes that the functional group of a species dictates how it interacts with other species. For example, if a system has species from two functional groups, the functional group model estimates two parameters for interactions within each functional group (δ_{wfg1} and δ_{wfg2}) and a third parameter for interactions between species from different functional groups (δ_{bfg}).

We fitted a range of Diversity-Interactions models to both the average and SAM metrics and the set of models for each metric were compared using F-tests. We found that for both the average metric and SAM metric the average Diversity-Interactions model:

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha_b + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} P_i P_j + \varepsilon \qquad (2)$$

provided the best fitting model (see Appendix 4.2.2 for model fitting details). For ease of interpretation of the interaction coefficient, we rescaled the sum of the pairwise interactions to be $E = \frac{2s}{s-1} \sum P_i P_j = \frac{20}{9} \sum P_i P_j$, where E lies between 0 for monocultures and 1 for the centroid community (all 10 species equally present) [10]. This gives the model

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha_b + \delta E + \varepsilon \qquad (3)$$

where δ is the expected diversity effect for the 10-species centroid community.

The estimated parameters for the SAM metric (Table 4.2.1) and the average metric (Table 4.2.2) show that all species identity effects and the average diversity effect were significant for both metrics. As species evenness (measured here by E) increases across communities there is a linear increase in the predicted diversity

Table 4.2.1. Table of parameter estimates from model 3 for the SAM metric.

Parameter	Estimate	Standard Error
β_1	**2.64**	**0.309**
β_2	**2.48**	**0.314**
β_3	**2.05**	**0.310**
β_4	**1.89**	**0.313**
β_5	**2.39**	**0.327**
β_6	**3.25**	**0.304**
β_7	**1.59**	**0.357**
β_8	**1.32**	**0.333**
β_9	**2.50**	**0.307**
β_{10}	**1.34**	**0.307**
α_1	2.43	1.807
α_2	0.00	-
δ	**2.54**	**0.397**

Significant parameter values are highlighted in bold.

Table 4.2.2. Table of parameter estimates from model 3 for the average metric.

Parameter	Estimate	Standard Error
β_1	**57.23**	**4.726**
β_2	**60.52**	**4.802**
β_3	**56.26**	**4.741**
β_4	**66.17**	**4.780**
β_5	**63.51**	**4.992**
β_6	**67.40**	**4.650**
β_7	**51.91**	**5.430**
β_8	**59.45**	**5.080**
β_9	**63.36**	**4.690**
β_{10}	**52.73**	**4.699**
α_1	**-5.62**	**1.966**
α_2	0.00	-
δ	**30.54**	**5.937**

Significant parameter values are highlighted in bold.

effect for the SAM metric (Figure 4.2.3, ■). The observed diversity effects (Figure 4.2.3, ○) were calculated as the difference between the observed SAM metric value for each plot and the expected SAM metric value based on combined monoculture effects. There was no evidence that a quadratic term for evenness was needed for either response metric.

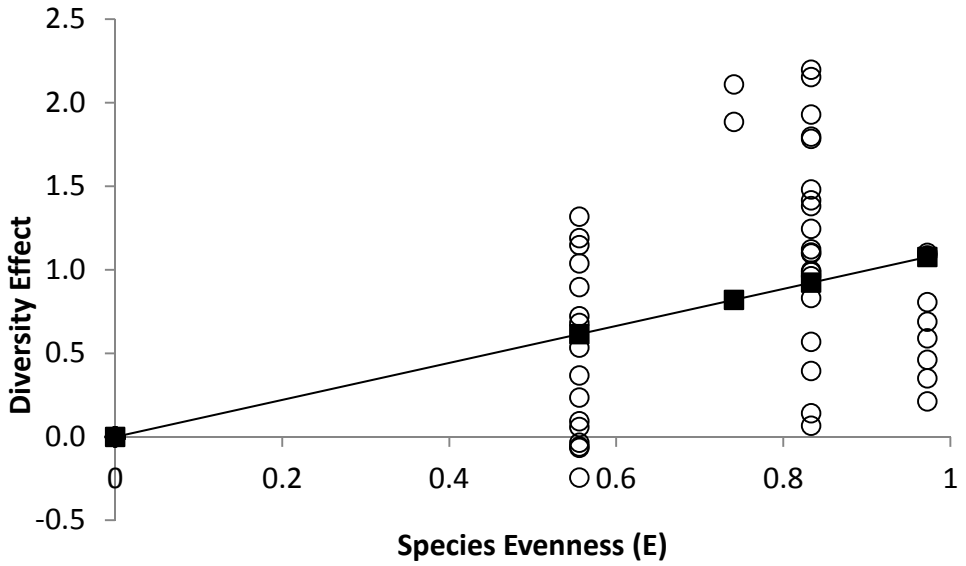

Figure 4.2.3. Plot of the predicted diversity effect (■) and the observed diversity effects (○) for the SAM metric across species evenness. The observed diversity effects were calculated as the difference between the observed SAM metric value for each plot and the expected SAM metric value based on combined monoculture effects.

Discussion

The averaging method [13, 14, 16] is a method which simplifies the complexity caused by analysing multiple functions simultaneously into a simple, single metric value which can then be analysed using univariate methods. Care must be taken however, when deciding to implement the averaging method in practice because there are numerous issues that arise with its interpretation. In this paper we highlighted a number of disadvantages of the method and addressed some with an extension to the Scaled Average Multifunctionality (SAM) metric.

The SAM metric was developed to address the drawback that communities with very different individual ecosystem function responses could have the same average metric value (Figure 4.2.1) which is seen in the case of the observed metric responses for community compositions 21 and 24 in block one of our data (Figure 4.2.2). Here we have two communities where the average metric is approximately equal but where the individual functions are performing differently. In community 24 the individual functions are each performing more similarly to one another whereas in community 21 the individual responses are more varied (Appendix 4.2.1). The SAM metric allows for this variability among the functions by dividing the average by the standard deviation between the ecosystem function values, thus the SAM metric penalises against communities where individual ecosystem functions are more varied compared to a more stable performance across functions.

Another drawback of the averaging method is that the averaging metric does not currently account for the individual function desirability. Calculating a metric value from functions where some have positive desirability and some have negative desirability does not make logical sense. For the work presented by Maestre *et al.* [13, 14] all functions examined were assumed to be functions where higher functioning is considered favourable. Byrnes *et al.* [12] addressed the problem by

incorporating the ecosystem function desirability when scaling the functions to create scaled functions with positive desirability for each function. The method presented by Byrnes *et al.* [12] was used in the calculation of the SAM metric to transform the unconsumed soil nitrogen content response. Higher values of unconsumed soil nitrogen content were deemed to be less desirable than lower values as high values can lead to higher nitrate leaching from the system. A second possible method to incorporate the response desirability is to split the functions by desirability and then analyse the functions with positive and negative desirability separately. This method allows for the desirability but yields two metric values instead of one. Care should be taken, however, as the desirability of an ecosystem function is subjective to the stakeholder's opinion and therefore may change between stakeholders.

The SAM metric does not currently address two of the four problems we identified with the averaging metric. Firstly, by combining the multiple ecosystem function responses into a single metric measurement we lose the ability to examine the effect of community characteristics on any individual function, thus reducing the amount of information we can gain about the multifunctional BEF relationship. Secondly, each ecosystem function used to calculate the SAM metric is assumed to be equally important. One possible way to incorporate the relative importance of the ecosystem functions into the SAM metric is to introduce a weighting for each function within the metric so that functions which are considered more important are given a heavier weight than the less important functions. This would allow the metric to focus more on the important functions whilst still allowing for the other functions being analysed. However, again, the importance of functions is dependent on each stakeholder's views and, as such, any proposed weighting system would be highly

subjective. Hill [19] discusses a number of indices which allow for different weighting among the components of the index for species richness. These ideas could be extended to consider multiple functions rather than species.

Although the SAM metric is an improvement on the averaging method care must be taken when implementing this method as it still carries a number of the drawbacks of the averaging method and requires additional care when using the standard deviation between the functions. A high SAM metric value could be caused by high ecosystem function values or by ecosystem functions performing similarly to one another. Functions performing similarly to one another will yield a small standard deviation which in turn will increase the SAM metric value. A strong assumption of the method is that it is desirable to have all functions functioning at a similar level, thus the metric penalises communities when functions have widely varying values. The SAM metric is designed to improve on the averaging method and should only be used in appropriate situations, i.e. when interested in examining the average response to changing community characteristics of a system where it is required that all functions are performing similarly. The SAM metric does not allow for trade-offs, *i.e.* one or more function which is performing well compensating for other functions performing poorly. Community composition 2 (Figure 4.2.2) showed that, by taking the variability between the functions into account, the SAM metric negated the trade-off between a high biomass yield and poorer performing functions such as root biomass.

In cases where we wish to examine the multifunctional BEF relationship without such a loss of information, other multifunctional methods [1-6, 20-22] are currently available. However, none of these methods fully deal with the multivariate nature of multifunctional BEF data. The use of a multivariate analysis would allow

for a more complete examination of the relationship between multiple ecosystem functions and changing community characteristics. Such a method is developed in the final section of this chapter.

References

1. Hector A, Bagchi R. Biodiversity and ecosystem multifunctionality. Nature. 2007;448(7150):188-91. doi: 10.1038/nature05947. PubMed PMID: ISI:000247934500041.

2. Orwin K, Ostle N, Wilby A, Bardgett R. Effects of species evenness and dominant species identity on multiple ecosystem functions in model grassland communities. Oecologia. 2014;174(3):979-92. doi: 10.1007/s00442-013-2814-5. PubMed PMID: WOS:000332002300033.

3. Hillebrand H, Matthiessen B. Biodiversity in a complex world: consolidation and progress in functional biodiversity research. Ecology Letters. 2009;12(12):1405-19. doi: 10.1111/j.1461-0248.2009.01388.x. PubMed PMID: ISI:000271631500016.

4. Zavaleta E, Pasari J, Hulvey K, Tilman G. Sustaining multiple ecosystem functions in grassland communities requires higher biodiversity. Proceedings of the National Academy of Sciences of the United States of America. 2010;107(4):1443-6. doi: DOI 10.1073/pnas.0906829107. PubMed PMID: ISI:000273974600040.

5. Gamfeldt L, Hillebrand H, Jonsson PR. Multiple functions increase the importance of biodiversity for overall ecosystem functioning. Ecology. 2008;89(5):1223-31. PubMed PMID: ISI:000256101200007.

6. Isbell F, Calcagno V, Hector A, Connolly J, Harpole W, Reich P, et al. High plant diversity is needed to maintain ecosystem services. Nature.

2011;477(7363):199-202. doi: DOI 10.1038/nature10282. PubMed PMID: ISI:000294603900033.

7. Loreau M, Naeem S, Inchausti P, Bengtsson J, Grime J, Hector A, et al. Ecology - Biodiversity and ecosystem functioning: Current knowledge and future challenges. Science. 2001;294(5543):804-8. PubMed PMID: ISI:000171851800033.

8. Schmid B, Hector A. The value of biodiversity experiments. Basic and Applied Ecology. 2004;5(6):535-42. doi: 10.1016/j.baae.2004.07.001. PubMed PMID: WOS:000226824700006.

9. Tilman D, Reich P, Knops J. Biodiversity and ecosystem stability in a decade-long grassland experiment. Nature. 2006;441(7093):629-32. doi: DOI 10.1038/nature04742. PubMed PMID: ISI:000237920800046.

10. Kirwan L, Luescher A, Sebastia M, Finn JA, Collins R, Porqueddu C, et al. Evenness drives consistent diversity effects in intensive grassland systems across 28 European sites. Journal of Ecology. 2007;95(3):530-9. doi: DOI 10.1111/j.1365-2745.2007.01225.x. PubMed PMID: ISI:000245693800013.

11. Naeem S, Bunker DE, Hector A, Loreau M, Perrings C. Biodiversity, Ecosystem Functioning and Human Wellbeing: Oxford University Press; 2009.

12. Byrnes J, Gamfeldt L, Isbell F, Lefcheck J, Griffin J, Hector A, et al. Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions. Methods in Ecology and Evolution. 2014;5(2):111-24. PubMed PMID: WOS:000331402000002.

13. Maestre FT, Quero JL, Gotelli NJ, Escudero A, Ochoa V, Delgado-Baquerizo M, et al. Plant species richness and ecosystem multifunctionality in global drylands. Science. 2012;335(6065):214-8. doi: 10.1126/science.1215442. PubMed PMID: WOS:000299033100052.

14. Maestre FT, Castillo-Monroy AP, Bowker MA, Ochoa-Hueso R. Species richness effects on ecosystem multifunctionality depend on evenness, composition and spatial pattern. Journal of Ecology. 2012;100(2):317-30. doi: 10.1111/j.1365-2745.2011.01918.x. PubMed PMID: WOS:000300500800003.

15. Hector A, Schmid B, Beierkuhnlein C, Caldeira M, Diemer M, Dimitrakopoulos P, et al. Plant diversity and productivity experiments in European grasslands. Science. 1999;286(5442):1123-7. PubMed PMID: ISI:000083534200033.

16. Mouillot D, Villeger S, Scherer-Lorenzen M, Mason N. Functional structure of biological communities predicts ecosystem multifunctionality. PLoS ONE. 2011;6(3). Epub e17476. doi: 10.1371/journal.pone.0017476. PubMed PMID: WOS:000288219100009.

17. Wagg C, Bender S, Widmer F, van der Heijden M. Soil biodiversity and soil community composition determine ecosystem multifunctionality. Proceedings of the National Academy of Sciences of the United States of America. 2014;111(14):5266-70. doi: 10.1073/pnas.1320054111. PubMed PMID: WOS:000333985200055.

18. Kirwan L, Connolly J, Finn JA, Brophy C, Luscher A, Nyfeler D, et al. Diversity-Interaction modeling: estimating contributions of species identities and interactions to ecosystem function. Ecology. 2009;90(8):2032-8. PubMed PMID: ISI:000268293000002.

19. Hill, MO. Diversity and evenness: a unifying notation and its consequences. Ecology. 1973;54(2): 427-432. doi: 10.2307/1934352.

20. Cardinale B, Gross K, Fritschie K, Flombaum P, Fox J, Rixen C, et al. Biodiversity simultaneously enhances the production and stability of community

biomass, but the effects are independent. Ecology. 2013;94(8):1697-707. doi: 10.1890/12-1334.1. PubMed PMID: WOS:000322336600005.

21. Allan E, Weisser W, Fischer M, Schulze E, Weigelt A, Roscher C, et al. A comparison of the strength of biodiversity effects across multiple functions. Oecologia. 2013;173(1):223-37. doi: 10.1007/s00442-012-2589-0. PubMed PMID: WOS:000323504300021.

22. He J-Z, Ge Y, Xu Z, Chen C. Linking soil bacterial diversity to ecosystem multifunctionality using backward-elimination boosted trees analysis. Journal of Soils and Sediments. 2009;9(6):547-54. doi: 10.1007/s11368-009-0120-y. PubMed PMID: WOS:000271643900005.

Section 4.3

This section has been formatted for Ecology Letters. The manuscript was first submitted to Ecology Letters in March 2015 and was published in September 2015.

Testing the effects of diversity on ecosystem multifunctionality using a multivariate model.

Áine Dooley[1], Forest Isbell[2], Laura Kirwan[3,6], John Connolly[4], John A. Finn[5], Caroline Brophy[1].

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

[2]*Department of Ecology, Evolution, and Behavior, University of Minnesota, St Paul, Minnesota 55108, USA.*

[3]*Waterford Institute of Technology, Cork Road, Waterford, Ireland.*

[4]*School of Mathematical Sciences, Ecological and Environmental Modelling Group, University College Dublin, Dublin 4, Ireland.*

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

[6]*Present address: UCD Institute of Food and Health, UCD, Belfield, Dublin 4, Ireland.*

Key words: Biodiversity, Diversity-Interactions model, ecosystem function, evenness, multifunctionality, multivariate, species interaction, species richness, trade-offs.

ABSTRACT

Most ecosystems provide multiple services, thus the impact of biodiversity losses on ecosystem functions may be considerably underestimated by studies that only address single functions. We propose a multivariate modelling framework for quantifying the relationship between biodiversity and multiple ecosystem functions (multifunctionality). Our framework consolidates the strengths of previous approaches to analysing ecosystem multifunctionality and contributes several advances. It simultaneously assesses the drivers of multifunctionality, such as species relative abundances, richness, evenness, and other manipulated treatments. It also tests the relative importance of these drivers across functions, incorporates correlations among functions and identifies conditions where all functions perform well and where trade-offs occur among functions. We illustrate our framework using data from three ecosystem functions (sown biomass, weed suppression and nitrogen yield) in a four-species grassland experiment. We found high variability in performance across the functions in monocultures, but as community diversity increased, performance increased and variability across functions decreased.

INTRODUCTION

The biodiversity and ecosystem function (BEF) relationship has been widely researched over the past few decades and ecosystem functions such as biomass production or resistance to weed invasion are generally reduced as biodiversity is lost (Hector *et al.* 1999; Cardinale *et al.* 2011; Finn *et al.* 2013). Since most investigations of the BEF relationship have focused on a single ecosystem function, the impact of biodiversity losses on the delivery of ecosystem services may be underestimated, however, several recent studies have explored the BEF relationship for multiple ecosystem functions (multifunctionality) (Hector & Bagchi 2007; Gamfeldt *et al.* 2008; Mouillot *et al.* 2011; Allan *et al.* 2013; Byrnes *et al.* 2014a). These studies have generally shown that the number of species required to maintain multifunctionality increases with the number of functions being considered, partly because different sets of species control different ecosystem functions (Hector & Bagchi 2007; Isbell *et al.* 2011).

Statistical methods for analysing the multifunctional BEF relationship include (1) qualitatively combining univariate models for each function (Allan *et al.* 2013), (2) the averaging approach (Mouillot *et al.* 2011), (3) the overlap method (Hector & Bagchi 2007) (4) the single threshold method (Gamfeldt *et al.* 2008) and (5) the multiple threshold method (Byrnes *et al.* 2014a). These methods are summarised in Appendix 4.3.1 and have been reviewed and critiqued in Byrnes *et al.* (2014a). Although these previous methods provide useful insights, each suffers from loss of information through simplifying the multivariate nature of the data (see Box 1). This information loss includes: reduced information on individual functions, correlations among functions

Box 1 Summary of the information loss associated with previous multifunctionality approaches (each described in Appendix 4.3.1) and description of the consolidation of the strengths of those approaches and the added benefits that the Multivariate Diversity-Interactions modelling framework provides.

Approach	Issues and information loss	Strengths that are included in the Multivariate Diversity-Interactions framework	Additional value of the Multivariate Diversity-Interactions framework
(1) Combining univariate models	• No information on correlations among functions. • Only qualitative information on multifunctionality.	• Understanding the drivers of each individual function.	• Tests the relative importance of the drivers across functions. • Quantitative information on single functions *and* on multifunctionality. • Incorporates correlations among functions into the assessment of drivers of multifunctionality.
(2) The averaging approach	• Loss of information at the individual ecosystem function level. • Two communities with very different ecosystem functions can yield the same average metric value (*e.g.*, with two functions, the two functions could be equal or one function could be very high and the other very low, but the two communities yield the same average) therefore it is an incomplete description of the underlying multivariate distribution.		• Tests the drivers of individual functions. • Tests the relative importance of the drivers across functions. • Utilises correlations among functions in inference.

Approach	Issues and information loss	Strengths that are included in the Multivariate Diversity-Interactions framework	Additional value of the Multivariate Diversity-Interactions framework
(3) The overlap method	• Ignores how sets of species that positively influence some ecosystem functions might reduce other functions.	• Quantifies the species that positively influence pairs of ecosystem functions.	• Tests how all species and pairwise interactions positively or negatively affect all functions (not just pairs of functions), *i.e.* identifies conditions under which multiple functions all perform well, but will also identify trade-offs among functions.
(4) The single threshold method	• Converts quantitative measurements to categorical thus there is loss of information on the amount by which a function exceeds or falls below a threshold. • Subjective to the choice of threshold. • Ignores effects of correlations among functions.	• Identifies combinations of species that will achieve, for example, 70% of the maximum performance.	• Quantitative predictions on how each function performs under varying diversity characteristics. • Identifies the combinations of species *and their relative abundances* that will attain, for example, 70% of the maximum.
(5) The multiple threshold method	• Requires carrying out the same tests repeatedly (at each threshold) but provides no statistical adjustment for the multiple comparisons. • Ignores effects of correlations among functions.	• Identifies combinations of species that will achieve a certain threshold of the maximum performance.	• Quantitative predictions on how each function performs under varying diversity characteristics. • Provides the combinations of species *and their relative abundances* that will attain a certain percentage of the maximum. • Provides an adjustment for the multiple tests of comparison that are needed in any multifunctionality analysis giving statistical reassurance on the reliability of conclusions.

not being measured and being ignored in analysis, species abundance being summarized as a binary variable (presence or absence) and continuous information being converted to categorical thresholds. While reducing the multivariate nature of data can sometimes be useful, it may lead to misconceptions at the individual ecosystem function level particularly when functions differ markedly in their responses to changing diversity (Bradford *et al.* 2014a, b; Byrnes *et al.* 2014b). These previous methods also focus strongly on species richness as the main driver of multifunctionality, ignoring other potentially highly influential aspects of diversity, such as the relative abundances of species or the ability of pairs of species to interact (Wilsey & Potvin 2000; Wilsey & Polley 2004; Kirwan *et al.* 2007; Finn *et al.* 2013).

The Diversity-Interactions approach (Kirwan *et al.* 2009; Connolly *et al.* 2013) models the BEF relationship for a single ecosystem function as a function of species identities and interactions among pairs of species. Here we develop the Multivariate Diversity-Interactions model to analyse the multifunctional BEF relationship by extending the univariate Diversity-Interactions approach to a multivariate framework. In this framework, comparisons of the model components across ecosystem functions allow testing of the relative performance of each function across diversity characteristics such as species identities, species interactions, evenness, richness and manipulated treatments or environmental variables, and it automatically allows for correlations among functions. Thus, we can identify conditions (if they exist) where all functions perform well relative to each other or identify where trade-offs occur among functions. We illustrate our Multivariate Diversity-Interactions framework with data for three ecosystem functions

from a four-species grassland biodiversity experiment. We investigate the following aspects of ecosystem multifunctionality:

(1) What diversity characteristics (*e.g.*, species abundances, species identities, species interactions, composition and evenness) affect each individual ecosystem function?

(2) How should correlations among ecosystem functions be incorporated in assessing drivers of multifunctionality?

(3) What is the relative importance of the various aspects of diversity and environment (species identities, species interactions and treatments) across ecosystem functions?

(4) Are there conditions under which all ecosystem functions perform well? Are there trade-offs occurring among ecosystem functions?

MATERIALS AND METHODS

The Multivariate Diversity-Interactions framework

The Diversity-Interactions model (Kirwan *et al.* 2009) is:

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

where y is a single ecosystem function, P_i (P_j) is the initial relative abundance of the i^{th} (j^{th}) species with i, j=1,...,s and A can include a measure of community abundance and/or block and/or treatment effects and so α may be a vector including several coefficients.

The coefficient β_i is the expected performance of the i^{th} species in monoculture and is called the species identity effect, δ_{ij} is the interaction effect between species i and j, and $\varepsilon \sim N(0,\sigma^2)$. Further interpretations are in Kirwan *et al.* (2009). Additional interactions can be tested among the model terms, including those between interaction (δ_{ij}) and treatment (α) coefficients. Model (1) addresses a single ecosystem function and here we extend it to a multivariate framework to simultaneously model the relationship between biodiversity and ecosystem multifunctionality.

For the Multivariate Diversity-Interactions model of k functions, the equation for the k^{th} function is of the form:

$$y_k = \sum_{i=1}^{s} \beta_{ik} P_i + \alpha_k A + \sum_{\substack{i,j=1 \\ i<j}}^{s} \delta_{ijk} P_i P_j + \varepsilon_k \qquad (2)$$

where β_{ik} is the identity effect for species i for ecosystem function k, and δ_{ijk} is the species interaction effect between species i and j for ecosystem function k. In matrix notation, the Multivariate Diversity-Interactions model is

$$\underline{Y} = \underline{X}\underline{\beta} + \underline{\varepsilon}$$
$$\underline{\varepsilon} \sim \text{MVN}(\underline{0}, \underline{\Sigma}), \qquad \underline{\Sigma} = \begin{bmatrix} \sigma_1^2 & \Lambda & \sigma_{1k} & & & \\ \text{M} & \text{O} & \text{M} & & 0 & \\ \sigma_{1k} & \text{K} & \sigma_k^2 & & & \\ & & & \text{O} & & \\ & & & & \sigma_1^2 & \Lambda & \sigma_{1k} \\ & 0 & & & \text{M} & \text{O} & \text{M} \\ & & & & \sigma_{1k} & \text{K} & \sigma_k^2 \end{bmatrix} \qquad (3)$$

$\underline{Y}$ represents the observed matrix of ecosystem functions, $\underline{X}$ represents the matrix of explanatory variables which includes all terms shown in equation (2) for the k^{th} function, $\underline{\beta}$ represents the matrix of model coefficients, MVN stands for multivariate normal and $\underline{0}$ is a matrix of zeros corresponding in size to the matrix $\underline{Y}$. The variance-covariance

matrix $\underline{\Sigma}$ is a block diagonal matrix with a k x k block for each plot; within each block, the diagonal entries are the ecosystem function variances and off diagonal entries are the covariances between the errors of each pair of ecosystem functions. There are ks identity effects and $ks(s-1)/2$ interaction effects to be estimated. This number can be reduced by making biologically meaningful assumptions about the patterns among the δ_{ijk} interaction coefficients (for each k) using the techniques outlined in Kirwan *et al*. (2009). For example, it might be assumed that all species interact in the same way ($\delta_{ijk} = \delta_{avk}$ for all i, j) or that all species from a particular functional group interact in the same way (for two functional groups, $\delta_{ijk} = \delta_{wfg1k}$ if i, j are both from functional group 1, $\delta_{ijk} = \delta_{wfg2k}$ if i, j are both from functional group 2, $\delta_{ijk} = \delta_{bfgk}$ if i, j are from different functional groups, where *wfg* represents 'within functional group' and *bfg* represents 'between functional group').

The data set

A four-species grassland biodiversity field experiment was established in 2002 at Merelbeke in Belgium as part of a larger agro-diversity experiment (Kirwan *et al.* 2007; Finn *et al.* 2013) and the data is publicly available as 'site 1' in Kirwan *et al.* (2014). The species sown were two grasses (*Lolium perenne,* denoted G1, and *Phleum pratense*, G2) and two legumes (*Trifolium pratense*, L1 and *Trifolium repens*, L2). Both G1 and L1 were fast-establishing species while G2 and L2 were temporally-persistent species. Thus there were two possible functional group classifications among the four species: grass / legume and fast-establishing / temporally-persistent. A monoculture for each

species and 11 four-species mixture communities were established at two (high and low) seed density levels giving a total of 30 plots each 8.4m^2 in size. The relative abundances in the mixture communities were systematically varied at sowing; at each seed density level, there were four monocultures, a community where the four species were sown in equal abundance (0.25, 0.25, 0.25, 0.25), four communities which were each dominated by one of the species (*e.g.*, (0.7, 0.1, 0.1, 0.1)) and six communities which were co-dominated by two species (*e.g.*, (0.4, 0.4, 0.1, 0.1)). Each community can be described using an evenness metric (Kirwan *et al.* 2007):

$E=(2s/(s-1))*\Sigma_{i<j}P_iP_j=(8/3)*\Sigma_{i<j}P_iP_j$.

The evenness values for the experiment are E=0 for monocultures, E=0.64 for one species dominant, E=0.88 for two species dominant and E=1 for all species equally present. Inorganic nitrogen fertilizer was applied to all plots at a rate of 150 kg N ha^{-1} annum^{-1}. Further details are available in Kirwan *et al.* (2014). Three ecosystem functions were recorded: (1) aboveground biomass of sown species (sown biomass) (t DM ha^{-1}) (2) aboveground biomass of weed species (weed biomass) (t DM ha^{-1}), and (3) the total annual yield of nitrogen in harvested aboveground biomass (N yield) (t DM ha^{-1}) for each plot and each harvest in 2003, the first year of the experiment following establishment. There were four harvests during the year and the annual values for each plot were computed by summing the four values for each ecosystem function. The experiment continued for a further two years, but only results from the first year are considered here to illustrate the new methodological developments.

Analysis

The three ecosystem functions were linearly transformed to a comparable scale allowing direct comparisons of the relative effects of the model terms (species relative abundances, species interactions and seed density effects) across the functions. High values of the functions sown biomass and N yield, and low values of weed biomass are preferred in agronomic practice; to align the direction of desirability for all functions (*i.e.* make higher positive values for all functions desirable), we first multiplied each weed biomass value by -1 and added the maximum (on the original scale) weed biomass value (Byrnes *et al.* 2014a) and called this new variable weed suppression. To linearly transform the data to a common scale, each ecosystem function (sown biomass, weed suppression and N yield) was then converted to a percentage of the average of the highest three values (top 10% of values from 30 plots) for that function (Appendix 4.3.2). From here on, these transformed variables are referred to as sown biomass, weed suppression and N yield. We did not apply any weighting to quantify differences in importance, which implicitly assumes that each function has equal importance (Appendix 4.3.2).

A range of Multivariate Diversity-Interactions models were fitted to the three transformed ecosystem functions, sown biomass, weed suppression and N yield, to explore reductions in the dimensionality of the diversity effect explanation. The data rescaling ensured that model predictions for each ecosystem function were on the same scale, which enabled us to test specific predictions across functions to identify conditions (if they existed) under which all functions performed relatively well (*e.g.*, when all ecosystem functions performed above an *a priori* specified level) and to

determine if trade-offs occurred among functions under other conditions (*e.g.*, when one or more functions performed above a specified level but others fell below). These comparisons were made using t-tests.

All models were estimated with either maximum likelihood (ML) or restricted maximum likelihood (REML) using SAS software version 9.3 (SAS Institute Inc., Cary, NC, USA); model comparisons for testing fixed effects were made using likelihood ratio tests where the models were fitted using ML, while final models were estimated and comparisons among coefficients and predictions were performed using REML. Multivariate normality of the residuals from the final model was tested using Mardia's multivariate normality test in the MVN package (Korkmaz *et al*. 2014) in the software R version 2.15.1 (R Core Team 2014). When testing model terms across functions (for example the comparisons among the coefficients $\beta_{11,}\beta_{12}$ and β_{13}), there were three pairwise t-tests of comparison (one comparison for each pair of ecosystem functions), thus a Bonferroni correction was applied to each set of three tests to avoid the issues associated with multiple comparisons, giving the adjusted $\alpha^{*}=0.05/3=0.017$. Note that the Multivariate Diversity-Interactions model could be fitted to the raw data and inference would be unchanged since only a linear transformation has been applied. However, the benefit of modelling the transformed ecosystem functions is the comparative ability across functions which would be meaningless with raw data modelling. Model predictions could be back-transformed to the original scale of each ecosystem function without affecting inference should this be desired. Note also that the ecosystem function that requires the most complex interaction structure may dictate the form of the final model since the same covariates are included for each ecosystem

function; this is the case with any multivariate regression model. Further information on fitting and interpreting multivariate regression models is available (for example) in Johnson and Wichern (2007). Appendices 4.3.3, 4.3.4 and 4.3.5 provide the data, SAS and R code, and some interpretations of output to assist readers wishing to fit the framework themselves.

RESULTS

Fitting the Multivariate Diversity-Interactions models

Summary statistics for the three ecosystem functions are given in Appendix 4.3.6. After model comparisons (Appendix 4.3.7), the final parsimonious model selected for the k^{th} transformed function was

$$y_k = \beta_{G1k}P_{G1} + \beta_{G2k}P_{G2} + \beta_{L1k}P_{L1} + \beta_{L2k}P_{L2} + \alpha_k \text{Dens} + \delta_{wfg1k}P_{G1}P_{G2} + \delta_{wfg2k}P_{L1}P_{L2} + \delta_{bfgk}\left(P_{G1}P_{L1} + P_{G1}P_{L2} + P_{G2}P_{L1} + P_{G2}P_{L2}\right) + \varepsilon_k \quad (4)$$

where P_{G1}, P_{G2}, P_{L1} and P_{L2} are the sown proportions of G1, G2, L1 and L2 respectively, Dens is coded -1 and 1 for low and high seed density. The β_{G1k} coefficient (for example) is the expected performance of G1 in monoculture for ecosystem function k at average density. The coefficients δ_{wfg1k} and δ_{wfg2k} are the pairwise interaction coefficients for the k^{th} function for the pair of grasses and pair of legumes respectively (*wfg*: within functional group). The interaction coefficient between any grass and any legume is δ_{bfgk} for the k^{th} function and all such interactions are assumed equal (*bfg*: between functional

group). The model residuals showed no evidence of a deviation from the multivariate normal distribution based on the Mardia's multivariate normality test.

Figure 4.3.1 and Table 4.3.1a show how positive species interactions both within and between functional groups were strong drivers of positive diversity effects for each individual ecosystem function (addressing question (1) as laid out in the introduction). There were no significant seed density effects for any function (Table 4.3.1a). There was a positive correlation among the residuals from sown biomass with the other two functions (Table 4.3.1b) and no evidence of a correlation among the residuals from nitrogen yield with weed suppression. The estimated covariances feed directly into the tests of comparison and allow for correct inference when comparing effects across functions (addressing question (2)).

Comparisons of multifunctionality across monocultures and multispecies communities

No one species in monoculture performed best across the three ecosystem functions (Fig. 4.3.2, the first set of clusters of bars). There was also no monoculture for which all three ecosystem functions performed poorly, rather there was considerable variability in performance across the functions for each monoculture. Comparisons of the estimated monoculture performances across ecosystem functions (Table 4.3.1a, comparison of each β coefficient across functions) showed that the performance of *Lolium perenne* (G1) was better for sown biomass and weed suppression than for N yield, and the performance of *Phleum pratense* (G2) was better for weed suppression than both sown

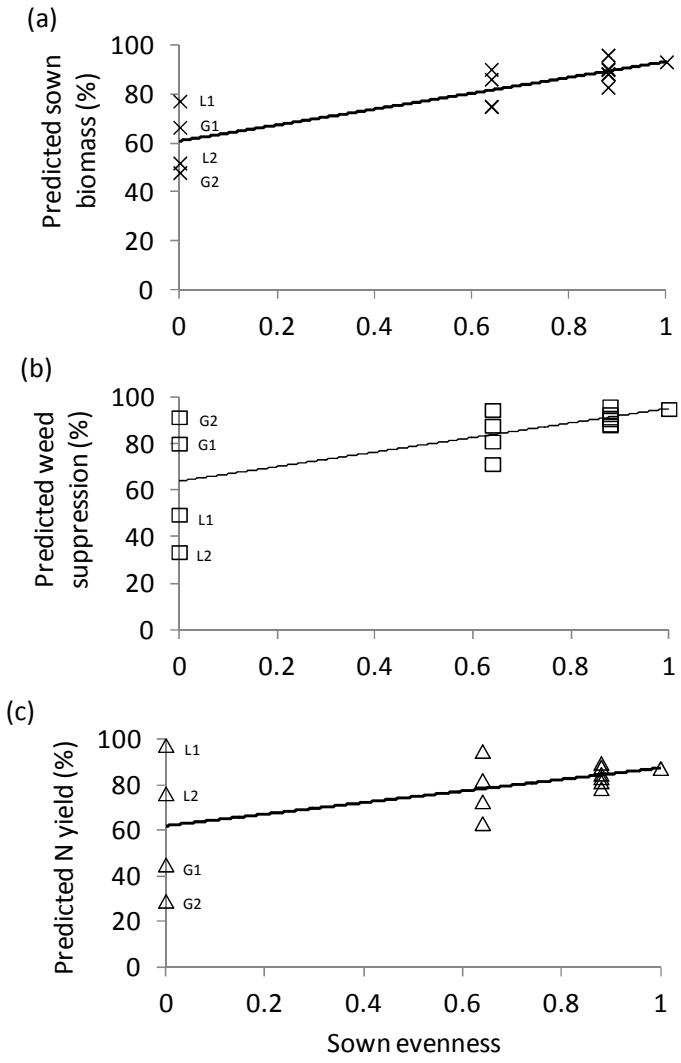

Figure 4.3.1 Predicted (a) sown biomass, (b) weed suppression and (c) N yield at average seed density for each community structure, monoculture (E=0), one dominant species (E=0.64), two dominant species (E=0.88) and all species equally abundant (E=1). The multiple points at each value of evenness represent the varying community types in the experimental design. Linear trendlines are added to indicate patterns as evenness increases and monocultures are labelled to indicate *Lolium perenne* (G1), *Phleum pratense* (G2), *Trifolium pratense* (L1) and *Trifolium repens* (L2).

Table 4.3.1 Estimated model terms for the transformed ecosystem functions, (a) fixed coefficients (b) the variance covariance matrix (left) and correlations (right). Significant ($\alpha < 0.05$) coefficients in (a) are highlighted in bold. Within each row (*i.e.* across ecosystem functions), coefficients that are not significantly different have a letter in common, where the level of significance determined by the Bonferroni correction is $\alpha^* = 0.05/3 = 0.017$.

(a)		Ecosystem function								
		Sown biomass (%)			Weed suppression (%)			N yield (%)		
Term	Coefficient	Est	SE		Est	SE		Est	SE	
G1	β_{G1k}	**66.48**	**4.50**	a	**80.29**	**8.47**	a	**45.02**	**4.60**	b
G2	β_{G2k}	**47.95**	**4.50**	a	**91.57**	**8.47**	b	**29.08**	**4.60**	c
L1	β_{L1k}	**77.22**	**4.50**	a	**49.75**	**8.47**	b	**97.43**	**4.60**	c
L2	β_{L2k}	**51.88**	**4.50**	a	**33.67**	**8.47**	a	**76.26**	**4.60**	b
Dens	α_k	1.15	1.31	a	0.50	2.47	a	-0.63	1.34	a
G1*G2	δ_{wfg1k}	**105.37**	**41.94**	a	-31.99	78.92	a	**150.46**	**42.82**	a
L1*L2	δ_{wfg2k}	64.64	41.94	a	159.97	78.92	ab	-5.32	42.82	b
ΣG*L (bfg)	δ_{bfgk}	**87.24**	**18.81**	a	**92.95**	**35.39**	a	**65.24**	**19.21**	a

(b)	Variances and covariances			Correlations	
	Sown biomass	Weed suppression	N yield	Weed suppression	N yield
Sown biomass	51.6			0.51	0.82
Weed suppression	49.1	182.7			0.07
N yield	43.1	6.6	53.8		

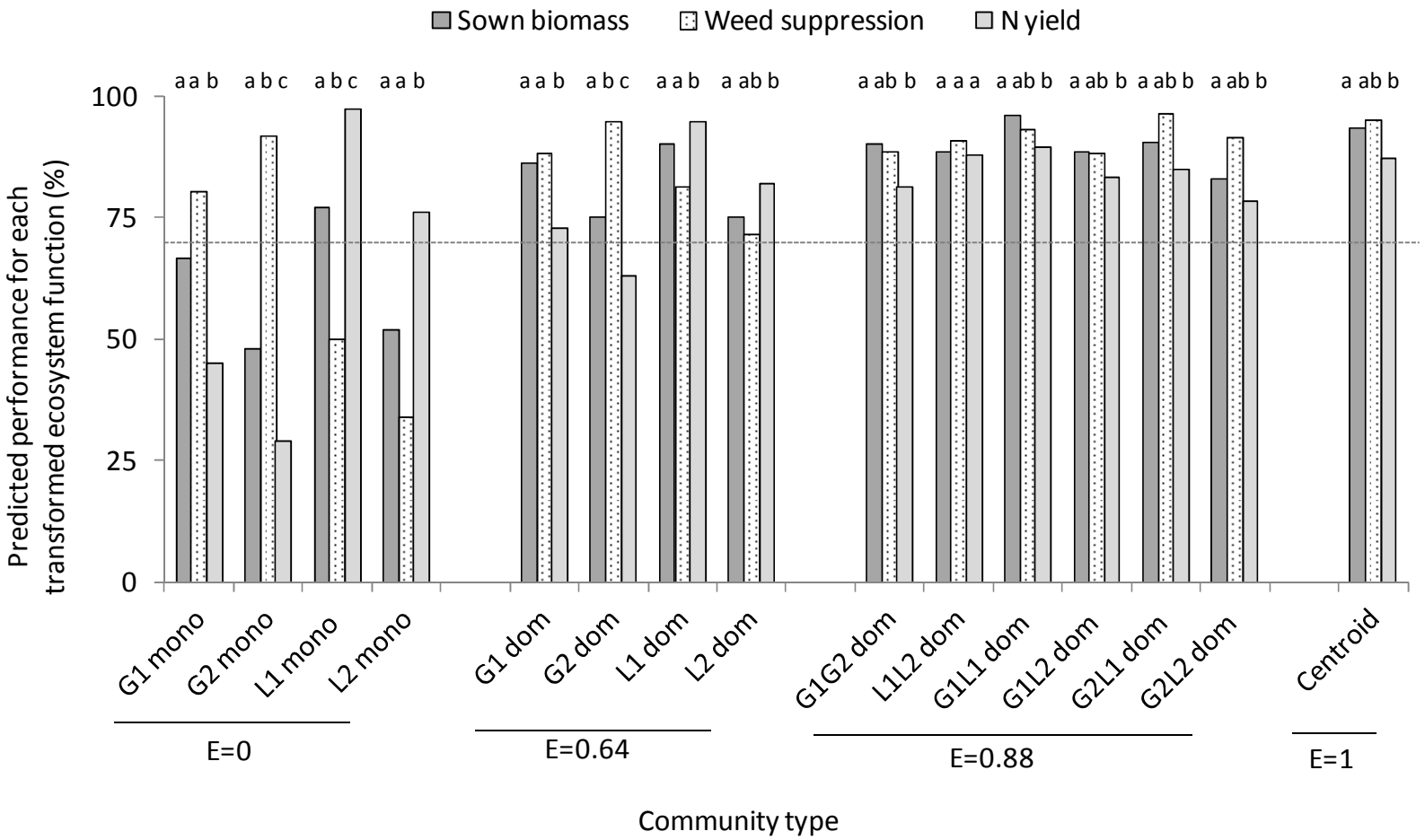

Figure 4.3.2 Predicted response for ecosystem functions sown biomass, weed suppression and N yield for each design community type (monocultures, one species dominant, two species co-dominant and all species equally abundant) at average seed density. Bars

within a cluster that share a letter do not differ significantly. The level of significance for all tests of comparison is determined by the Bonferroni correction, $\alpha^*=0.05/3=0.017$. Note that, *e.g.*, G1 mono is a grass 1 monoculture, G1 dom is (0.7,0.1,0.1,0.1), G1G2 dom is (0.4,0.4,0.1,0.1) and the centroid is (0.25,0.25,0.25,0.25). The species are *Lolium perenne* (G1), *Phleum pratense* (G2), *Trifolium pratense* (L1) and *Trifolium repens* (L2). A horizontal line is included at $y = 70\%$ to aid comparisons.

biomass and N yield. Not surprisingly, given their nitrogen fixing abilities, the performances of *Trifolium pratense* and *Trifolium repens* (L1 and L2) in monoculture were far better for N yield than for either sown biomass or weed suppression (addressing question (3)). The tests displayed in Fig. 4.3.2 show that choosing either of the grass monocultures (over other monocultures) to optimise weed suppression results in relatively poorer performances of sown biomass (G2 only) and N yield (both G1 and G2) while choosing either of the legume monocultures to optimise N yield results in lower relative performances of sown biomass and weed suppression (addressing question (4)). The details of the tests illustrated in Fig. 4.3.2 are shown in Appendix 4.3.8.

The predicted performance of ecosystem functions in community types with evenness equal to 0.64 (one species dominant) varied depending on which species was dominant (Fig. 4.3.2, the second set of clustered bars); the performance of N yield was better relative to the other two functions when *Trifolium pratense* (L1) was dominant, while the performance of weed suppression was better relative to the two other functions when *Phleum pratense* (G2) was dominant. At evenness levels 0.88 (two species co-dominant) and 1 (centroid), each function performed at a high level; predictions for each ecosystem function and all community types at E=0.88 or 1 were higher than 70% ($p<0.05$ for each test). Note that 70% has been chosen arbitrarily for illustration here but should be chosen *a priori* in a practical application. There were still some small (but significant) differences within each cluster at the higher levels of evenness with sown biomass generally outperforming N yield (Fig. 4.3.2).

Despite there being significant differences among the three responses for each of the 15 community types presented in Fig. 4.3.2, the magnitude of the

differences decreased as evenness increased. For example, the estimated difference between sown biomass and N yield was 21% for *Lolium perenne* (G1) monoculture (E=0), 14% for a four-species community dominated by G1 (E=0.64), 9% for a four-species community co-dominated by G1 and G2 (E=0.88) and 6% for the centroid community (E=1), a significant difference in each case but the effect size (*i.e.* the differences) decreased as evenness increased. On average, performance across the three functions was higher and more stable in the communities with evenness equal to 0.88 or 1 when compared to the lower and more variable responses in monoculture and at E=0.64. Thus we show that the ecosystem functions in this experiment strongly trade-off against one another at low levels of evenness but all exhibited desired levels of performance (>70%) at higher levels of evenness (addressing question (4)).

DISCUSSION

The Multivariate Diversity-Interactions framework developed here provides quantitative tools to enhance our understanding of ecosystem multifunctionality. Our framework can test how multiple ecosystem functions are simultaneously driven by species abundances, species identities, species interactions, composition, richness and evenness. It can also test the relative importance of those drivers and identify key species and influential pairwise species interactions across multiple ecosystem functions. The framework provides quantitative information on individual as well as multiple functions and can aid decision-making to support the management of ecosystems in which the high performance of several functions is desired, such as in the agronomic communities in our example.

Our framework integrates the analytical outputs and insights formerly obtained from several separate multifunctionality approaches, including species-level information provided by the overlap approach and community-level information provided by the averaging and multiple threshold approaches. By combining these types of information, our framework is uniquely able to identify combinations of species and relative abundances that produce desirable levels of multiple ecosystem functions. For example, we found that four-species mixtures that were co-dominated by *Lolium perenne* (G1) and *Trifolium pratense* (L1) provided nearly maximal levels of all three ecosystem functions (Fig. 4.3.2). As manipulated evenness increased, we also showed that ecosystem functions were higher on average and that the variability among the three ecosystem functions decreased (Fig. 4.3.2). Other studies have examined ecosystem multifunctionality over time (Isbell *et al.* 2011; Cardinale *et al.* 2013; Pasari *et al.* 2013), trophic levels and ecosystem types (Lefcheck *et al.* 2015) but not variability among the levels of multiple functions across a manipulated treatment. Our agronomic example provides further evidence of the benefits of increased diversity on ecosystem multifunctionality.

A key strength of the Multivariate Diversity-Interactions framework is its comparative ability whereby the model coefficients and predictions from the model under varying diversity conditions can be tested for differences across functions. This ability is directly enabled by the estimation of the variance covariance matrix (Table 4.3.1b). Had three separate univariate Diversity-Interactions models been fitted instead of a multivariate model, the coefficient estimates and their standard errors (Table 4.3.1a) would be no different, but the univariate approach would not have estimated the variance covariance matrix (Table 4.3.1b) and thus it would not have been possible to correctly make comparisons across functions. For example, the

t-test statistic for comparing β_{G11} and β_{G13} (the expected *Lolium perenne* (G1) monoculture performance for sown biomass and N yield respectively) was 7.83 with p-value <0.0001. This test and its inference are valid since the covariance between the two functions contributes to the test statistic calculation. If, however, a zero covariance between the estimates had been assumed, the test statistic would be calculated (incorrectly) as 3.33 with p-value=0.002 resulting in approximately a halving of the test statistic and any inference from this incorrect test would not be valid. This comparative ability of the Multivariate Diversity-Interactions framework allows (1) the identification of compositions and relative abundances where all ecosystem functions perform well or, (2) the identification of how functions may trade off against one another and (3) understanding of how optimisation of one function impacts other functions. In our example, the G2 monoculture attained 92% in weed suppression but only 48% and 29% in sown biomass and N yield respectively, illustrating trade-offs among functions in this monoculture (and others). There were no significant differences among the ecosystem functions for the community co-dominated by L1 and L2 and each function was higher than 70%, illustrating conditions where all functions had similarly high levels of performance (Fig. 4.3.2).

The Multivariate Diversity-Interactions framework includes the benefits and addressees the losses of information that are inherent in other methods for analysing multifunctionality. Our framework estimates the relationship between individual ecosystem functions and manipulated diversity or treatment variables (as in the univariate approach in Allan *et al*. (2013)), quantifies which species positively influence ecosystem function (as in the overlap method, Hector & Bagchi (2007)) and can identify what combination of species will yield a certain percentage of the

maximum of ecosystem function performance (as in the single and multiple threshold methods in Gamfeldt *et al*. (2008) and Byrnes *et al*. (2014a)). In addition, our framework also measures correlations among functions, provides a means for statistical tests of comparisons across multiple functions, provides quantitative estimates on multifunctionality across varying compositions and relative abundances, and identifies important species and species interactions for individual functions and tests their relative importance across functions, which other approaches cannot do. Analysing each ecosystem function individually allows only for qualitative inference on multifunctionality (Byrnes *et al.* 2014a), while dimension reducing indices which quantify multifunctionality may omit important information at the individual ecosystem function level (Bradford *et al.* 2014a, b; Byrnes *et al.* 2014b); the ability of our framework to assess individual ecosystem functions in conjunction with multifunctionality is therefore highly desirable. We thus present our framework as a consolidation of the strengths of previous approaches that also provides several additional advances in the quantification of ecosystem multifunctionality (Box 1).

The rich information available from using our framework goes beyond what is achievable with other approaches used to analyse the biodiversity and ecosystem multifunctionality relationship. In our experiment, the four-dimensional simplex design space was well represented, therefore we can use our model to predict each ecosystem function for any set of relative abundances and compositions of these four species. For example, we can estimate each ecosystem function for the community compositions (0.5 0.5, 0, 0) and (0.8, 0.1, 0.05, 0.05), even though these are not represented by any specific design point; this predictive power reflects an important added advantage of the approach. Generally, when a traditional linear regression

model with log(richness) as a covariate is fitted, the model can predict at each level of richness but cannot distinguish between communities with differing relative abundances at the same level of richness; for example, the two markedly different communities (0.25 0.25, 0.25, 0.25) and (0.85, 0.05, 0.05, 0.05) would yield the same prediction in the traditional model but our framework would provide unique predictions. This distinctive trait is exclusive to our approach and is not provided by other ecosystem multifunctionality approaches. Our framework can still test richness effects by predicting each ecosystem function for equi-proportional communities at each level of richness and comparing across functions within each level. We can also use the Multivariate Diversity-Interactions framework to identify zones in the simplex space when all or most functions perform well or at close to their maximum value.

We found that the most parsimonious model was one of intermediate complexity, which included functional group interactions, rather than unique interactions between all pairs of species (Appendix 4.3.7). The between grass-legume functional group interaction coefficients were strong and positive for each function highlighting the benefits of mixing these functional groups for multifunctionality in grassland systems (Table 4.3.1). This benefit is well documented for individual functions (Ledgard & Steele 1992; Spehn *et al.* 2002; Nyfeler *et al.* 2011) but is shown here for the first time for ecosystem multifunctionality. The two grasses also interacted strongly and positively for both sown biomass and N yield perhaps reflecting the fast-establishing and temporally-persistent traits of G1 and G2 respectively.

The intricacies involved in ecosystem multifunctionality research questions are compounded when the ecosystem is more complex. It is therefore not surprising

that difficulties can arise with our multivariate approach when the numbers of species and / or ecosystem functions increase. These difficulties are a natural consequence of the increasing complexity of the system that can be handled by our framework. We summarise the complexities and outline solutions in the following three points.

1. When the number of species (s) increases, the number of coefficients per ecosystem function also increases. Kirwan *et al*. (2009) suggested constraints among interaction coefficients to alleviate this problem and here we constrained interaction coefficients according to functional groupings. Kirwan *et al*. (2009), Connolly *et al*. (2011; 2013) each provide alternative solutions to reduce the dimensionality of the diversity effect description which readily apply to our multivariate setting. In our experience with single ecosystem functions, it is frequently possible to model the diversity effect using a small number of coefficients even with high species richness, for example a 10-species grassland system (Connolly *et al.* 2011) and a 72-species bacterial system in (Connolly *et al.* 2013) were both modelled with just two diversity coefficients. It is also possible to test for biologically meaningful patterns among the identity effect (β_i) coefficients.
2. When the number of ecosystem functions increases, so too do the overall number of coefficients. Our method maintains individual function information and if this is desirable then there is no option but to increase the number of equations and hence number of coefficients used to describe the system. If individual function information is not required, then alternative multifunctionality approaches (Appendix 4.3.1) may be more useful and we encourage their usage.

3. We used a Bonferroni correction to adjust for multiple comparisons but if the number of ecosystem functions were to increase so too would the number of comparisons resulting in the criterion for a significant result becoming stricter and Bonferroni adjustments would likely be unduly conservative (Gotelli & Ellison 2004). This issue of multiple comparisons arises in other approaches developed for analysing multifunctionality (*e.g.*, Hector & Bagchi 2007; Gamfeldt *et al.* 2008; Isbell *et al.* 2011; Byrnes *et al.* 2014a) but has not been dealt with in any of those methods. Here we show that adjusting for multiple comparisons can be relatively straight-forward, at least for a small number of functions. For a larger number of functions, alternative more powerful large scale methods for adjusting for multiple comparisons to the Bonferroni correction should be used (*e.g.* Donoghue 2004; Verhoeven *et al.* 2005).

The Multivariate Diversity-Interactions framework is applicable to data from many types of designed experiments although sometimes it is not appropriate. For example, it is not recommended to fit a Diversity-Interactions model to an experiment with monocultures of each species and replicates of only one mixture type containing all species in equal relative abundances (*e.g.* Griffin *et al.* 2009). This is because there is inadequate coverage of the simplex space in this design and all mixtures are equal in respect of diversity manipulations therefore it is not possible to estimate pairwise interactions. Many biodiversity experiments have equi-proportional mixtures across a manipulated gradient of richness (*e.g.* Hector *et al.* 1999; Roscher *et al.* 2004) and a smaller number of studies manipulated evenness at a single level of richness (*e.g.* Wilsey & Potvin 2000; Finn *et al.* 2013). Our framework is fully suited to the analysis of such data as has been shown in previous

work in the univariate setting (*e.g.* Connolly *et al.* 2011) for richness manipulations and in our example here for evenness manipulations. A design with both evenness and richness manipulations combined with our modelling approach would provide even further predictive power but both manipulations are not a requirement. Note that the estimation of pairwise interaction terms does not specifically require two-species mixtures in the design. It is also possible to apply the Multivariate Diversity-Interactions framework to observational data although reliability would highly depend on the data in question as the usual regression models caveats would apply; these include ensuring there is sufficient representation in the design space and that caution is exercised in inferring causation from observed correlations.

The Multivariate Diversity-Interactions framework is flexible and can be extended in several directions, four of which we highlight here. (1) The model can analyse multiple ecosystem functions across a range of treatments or environments. Here we presented data with two sown seed densities; however, other treatments, such as different levels of applied nitrogen, can easily be incorporated into the model (*e.g.*, see Kirwan *et al.* 2009). (2) The framework can be extended for the analysis of multiple functions across temporal and spatial variables (Isbell *et al.* 2011), as has already been done for the univariate Diversity-Interactions modelling approach (Kirwan *et al.* 2007; Finn *et al.* 2013). (3) It would be possible to allow for non-linearity in the relationship between the ecosystem functions and the species interactions (see the Generalised Diversity-Interactions approach by Connolly *et al.* 2013). (4) The model here assumes a constant variance across plots for each ecosystem function but could easily be adjusted if this were not the case, for example, the variance for an ecosystem function could differ between monocultures and mixture communities (Schmid *et al.* 2008). These potential extensions further

illustrate the benefits of our framework. Structural equation models have been used to assess the biodiversity and ecosystem function relationship for single functions (*e.g.* Grace *et al.* 2007; Bowker *et al.* 2010). These models may also have a useful role in understanding ecosystem multifunctionality, however, initial attempts to do so may not be valid due to the questionable model selection process used (see comments on Mouillot *et al.* 2011).

The Multivariate Diversity-Interactions framework examines the multifunctional BEF relationship through a multivariate model fit that does not suffer from the loss of information inherent in other approaches. The framework consolidates the strengths and improves on the weaknesses of previous approaches for analysing ecosystem multifunctionality. It can identify the drivers of multiple ecosystem functions and test the relative performances across functions. The Multivariate Diversity-Interactions framework can be adapted to suit varying experimental conditions and is a valuable tool to improve understanding of ecosystem multifunctionality.

REFERENCES

1.

Allan, E., Weisser, W., Fischer, M., Schulze, E.-D., Weigelt, A., Roscher, C. *et al.* (2013). A comparison of the strength of biodiversity effects across multiple functions. *Oecologia*, 173, 223-237.

2.

Bowker, M.A., Maestre, F.T. & Escolar, C. (2010). Biological crusts as a model system for examining the biodiversity–ecosystem function relationship in soils. *Soil Biol. Biochem.*, 42, 405-417.

3.

Bradford, M.A., Wood, S.A., Bardgett, R.D., Black, H.I.J., Bonkowski, M., Eggers, T. *et al.* (2014a). Discontinuity in the responses of ecosystem processes and multifunctionality to altered soil community composition. *P. Natl. Acad. Sci. USA*, 111, 14478-14483.

4.

Bradford, M.A., Wood, S.A., Bardgett, R.D., Black, H.I.J., Bonkowski, M., Eggers, T. *et al.* (2014b). Reply to Byrnes et al.: Aggregation can obscure understanding of ecosystem multifunctionality. *P. Natl. Acad. Sci. USA*, 111, E5491.

5.

Byrnes, J.E.K., Gamfeldt, L., Isbell, F., Lefcheck, J.S., Griffin, J.N., Hector, A. *et al.* (2014a). Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions. *Methods Ecol. Evol.*, 5, 111-124.

6.

Byrnes, J.E.K., Lefcheck, J.S., Gamfeldt, L., Griffin, J.N., Isbell, F. & Hector, A. (2014b). Multifunctionality does not imply that all functions are positively correlated. *P. Natl. Acad. Sci. USA*, 111, E5490.

7.

Cardinale, B.J., Gross, K., Fritschie, K., Flombaum, P., Fox, J.W., Rixen, C. *et al.* (2013). Biodiversity simultaneously enhances the production and stability of

community biomass, but the effects are independent. *Ecology*, 94, 1697-1707.

8.

Cardinale, B.J., Matulich, K.L., Hooper, D.U., Byrnes, J.E., Duffy, E., Gamfeldt, L. *et al.* (2011). The functional role of producer diversity in ecosystems. *Am. J. Bot.*, 98, 572–592.

9.

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. *J. Ecol.*, 101, 344-355.

10.

Connolly, J., Cadotte, M., Brophy, C., Dooley, Á., Finn, J.A., Kirwan, L. *et al.* (2011). Phylogenetically diverse grasslands are associated with pairwise interspecific processes that increase biomass. *Ecology*, 92, 1385-1392.

11.

Donoghue, J.R. (2004). Implementing Shaffer's multiple comparison procedure for a large number of groups. In: *Recent Developments in Multiple Comparison Procedures* (eds. Benjamini, Y, Bretz, F & Sarkar, S). Institute of Mathematical Statistics Beachwood, Ohio, USA, pp. 1-23.

12.

Finn, J.A., Kirwan, L., Connolly, J., Sebastià, M.T., Helgadottir, A., Baadshaug, O.H. *et al.* (2013). Ecosystem function enhanced by combining four functional types of plant species in intensively-managed grassland mixtures: a three-year continental-scale field experiment. *J. Appl. Ecol.*, 50, 365–375.

13.

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

14.

Gotelli, N. & Ellison, A. (2004). A primer of ecological studies. Sunderland, MA. Sinauer Associates, Inc.

15.

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

16.

Griffin, J.N., Jenkins, S.R., Gamfeldt, L., Jones, D., Hawkins, S.J. & Thompson, R.C. (2009). Spatial heterogeneity increases the importance of species richness for an ecosystem process. *Oikos*, 118, 1335-1342.

17.

Hector, A. & Bagchi, R. (2007). Biodiversity and ecosystem multifunctionality. *Nature*, 448, 188-191.

18.

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.

19.

Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W.S., Reich, P.B. *et al.* (2011). High plant diversity is needed to maintain ecosystem services. *Nature*, 477, 199-202.

20.

Johnson, R.A. & Wichern, D.W. (2007). *Applied multivariate stastistical analysis*. Sixth edn. Pearson Eduction, Inc., NJ, USA.

21.

Kirwan, L., Connolly, J., Brophy, C., Baadshaug, O., Belanger, G., Black, A. *et al.* (2014). The Agrodiversity Experiment: three years of data from a multisite study in intensively managed grasslands. *Ecology*, 95, 2680.

22.

Kirwan, L., Connolly, J., Finn, J.A., Brophy, C., Luscher, A., Nyfeler, D. *et al.* (2009). Diversity-Interaction modeling: estimating contributions of species identities and interactions to ecosystem function. *Ecology*, 90, 2032-2038.

23.

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. *J. Ecol.*, 95, 530-539.

24.

Korkmaz, S., Goksuluk, D. & Zararsiz, G. (2014). MVN: An R Package for Assessing Multivariate Normality. The R Journal, 6, 151–162.24.

25.

Ledgard, S.F. & Steele, K.W. (1992). Biological nitrogen fixation in mixed legume/grass pastures. *Plant and Soil*, 141, 137-153.

26.

Lefcheck, J.S., Byrnes, J.E.K., Isbell, F., Gamfeldt, L., Griffin, J.N., Eisenhauer, N. *et al.* (2015). Biodiversity enhances ecosystem multifunctionality across trophic levels and habitats. *Nat Commun*, 6.

27.

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

28.

Nyfeler, D., Huguenin-Elie, O., Suter, M., Frossard, E. & Lüscher, A. (2011). Grass-legume mixtures can yield more nitrogen than legume pure stands due to mutual stimulation of nitrogen uptake from symbiotic and non-symbiotic sources. *Agriculture, Ecosystems and Environment*, 140, 155-163.

29.

Pasari, J.R., Levi, T., Zavaleta, E.S. & Tilman, D. (2013). Several scales of biodiversity affect ecosystem multifunctionality. *P. Natl. Acad. Sci. USA*, 110, 10219-10222.

30.

R Development Core Team (2014) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/.

31.

Roscher, C., Schumacher, J., Baade, J., Wilcke, W., Gleixner, G., Weisser, W.W. *et al.* (2004). The role of biodiversity for element cycling and trophic interactions: an experimental approach in a grassland community. *Basic Appl. Ecol.*, 5, 107-121.

32.

Schmid, B., Hector, A., Saha, P. & Loreau, M. (2008). Biodiversity effects and transgressive overyielding. *J. Plant Ecol.*, 1, 95-102.

33.

Spehn, E.M., Scherer-Lorenzen, M., Schmid, B., Hector, A., Caldeira, M.C., Dimitrakopoulos, P.G. *et al.* (2002). The role of legumes as a component of biodiversity in a cross-European study of grassland biomass nitrogen. *Oikos*, 98, 205-218.

34.

Verhoeven, K.J.F., Simonsen, K.L. & McIntyre, L.M. (2005). Implementing false discovery rate control: increasing your power. *Oikos*, 108, 643-647.

35.

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

36.

Wilsey, B.J. & Potvin, C. (2000). Biodiversity and ecosystem functioning: Importance of species evenness in an old field. *Ecology*, 81, 887-892.

Chapter 5

Conclusion

The aim of my research was to extend the current knowledge and modelling techniques used for examining the biodiversity and ecosystem function (BEF) relationship. My thesis has primarily focused on two difficulties faced when modelling the BEF relationship, namely how to model ecosystems which contain large amounts of species and how to model multiple ecosystem functions simultaneously. I approached these challenges with four main goals to achieve (as laid out in the thesis introduction):

1. To explore the use of community phylogenetic diversity information to help improve models for the BEF relationship for species rich communities.
2. To develop a random effects Diversity-Interactions model to increase the understanding of the BEF relationship for a single function.
3. To review and improve upon current multifunctionality metrics focusing on the averaging metric.
4. To develop a multivariate Diversity-Interactions model to analyse the multifunctional BEF relationship.

In each of the previous chapters I have addressed one or more of these goals; goal 1 was achieved by chapter 2, goal 2 by chapter 3, and goals 3 and 4 by chapter 4.

Chapter 2

The work in chapter 2 focused on the first goal of my research. The work in Connolly *et al.* (2011) identified that a measure of community phylogenetic diversity was a useful predictor of ecosystem function in two datasets and that communities with more phylogenetic diversity had higher diversity effects. Testing this result across multiple different grassland ecosystems (section 2.1), showed that the community phylogenetic diversity significantly added to the model for the ecosystem response in two out of eight grassland datasets. For these two datasets the results agreed with Connolly *et al.* (2011), i.e. that communities with more phylogenetic diversity had higher diversity effects. The reason that the other datasets tested did not show a significant effect of community phylogenetic diversity could be due to the fact that these datasets were not established to be phylogenetically diverse; the two datasets which showed a significant community phylogenetic effect were among the most phylogenetically diverse experiments. This implies that a reasonably wide range of phylogenetic diversity is needed for community phylogenetic diversity effects to be detected. Using this trait measure to explain changes in ecosystem function could prove particularly useful in species rich ecosystems where estimating large numbers of pairwise interactions among species may be difficult.

Chapter 3

In chapter 3, I focused on the second goal by incorporating random effects into the Generalised Diversity-Interactions model (Kirwan *et al.* 2007; Kirwan *et al.* 2009; Connolly *et al.* 2013). For species rich ecosystems, assumptions are often

made (or tested) to reduce the number of fixed model parameters needed to describe diversity effects (Kirwan *et al.* 2009). While this can often be useful, it can also lead to a loss of information, compared to modelling all pairwise interactions. Assuming a random variance component for pairwise interactions in conjunction with a small number of fixed effect coefficients may provide better explanatory power than using fixed effect solutions alone. The developed Generalised Diversity-Interactions Mixed (GDIM) models can also be used to examine whether there is evidence of lack of fit for models where assumptions have been made to reduce the number of fixed parameters needed. The various GDIM models allow us to test multiple hypotheses about the residual error variance across community characteristics without having to complete a separate variance analysis (such as in Tilman *et al.* 1996; Hooper 1998; Ives & Carpenter 2007). The GDIM modelling approach provides a parsimonious solution to modelling species rich ecosystems and can be used to make improved inference about the biodiversity ecosystem function relationship. It provides information about the mean ecosystem function, the residual error variance structure and it improves the standard errors associated with the fixed effects tests by capturing variation in species interactions additional to the fixed effects model used.

Chapter 4

The work in chapter 4 focused on the last two goals of my research. Analysing multiple ecosystem functions simultaneously (multifunctionality) is an emerging research area in ecology (*e.g.* Hector & Bagchi 2007; Gamfeldt *et al.* 2008; Zavaleta *et al.* 2010; Maestre *et al.* 2012a; Byrnes *et al.* 2014) as ecosystems

typically provide multiple functions which interact and affect one another. Section 4.1 reviewed the current methods for analysing multifunctional data and highlighted a number of key issues, namely the loss of information that often occurs with current multifunctional methods. In section 4.2, I aimed to build an improved metric for combining the multiple ecosystem functions into a single metric. The Scaled Average Multifunctional (SAM) metric was developed as an extension of the current averaging metric (Mouillot *et al.* 2011; Maestre *et al.* 2012a; Maestre *et al.* 2012b) by taking the variability among the ecosystem functions into account in the scaling of the average of the functions. By adjusting for the variability between the functions, the SAM metric addressed the potential difficulty of two systems with different functioning levels yielding similar average functional values.

The SAM metric, however, still suffers from loss of information. By combining multiple functions into a single metric value we still lose information about how the individual functions are affected by community characteristics. Also the metric is dependent on the stakeholders' interest in the functions. Different stakeholders may have differing opinions of the importance and desirability of each function and so analysis using the SAM metric should only be done when appropriate, i.e. when interested in examining the average response to changing community characteristics of a system where it is required that all functions are performing similarly.

A more universally appropriate method for analysing multifunctional data is to use the Multivariate Diversity-Interactions framework developed in section 4.3. The Multivariate Diversity-Interactions framework extended the Diversity-Interactions model to a multivariate framework to allow for the simultaneous modelling of multiple functions using the community characteristics of the

ecosystem. The Multivariate Diversity-Interactions framework maintains the flexibility of the Diversity-Interactions model whilst not suffering from the loss of information that other multifunctional methods suffer from. The framework can examine the relationship between the functions and identify compositions and relative species abundances where all ecosystems functions are performing well.

Future work

The aim for my research work in the future is to continue extending the current understanding of the BEF relationship. I intend to do this by improving upon existing models for analysing single ecosystem functions and continuing the development of multifunctional techniques.

There are still some remaining questions as to how robust the use of the phylogenetic diversity is in modelling ecosystem function. Out of the eight datasets tested in section 2.1 the full Diversity-Interactions model could only be fit to one. If a relationship between the phylogenetic diversity and the species interactions could be established, this would prove useful for systems where estimation of the full pairwise interactions Diversity-Interactions model is impossible to fit but the phylogenetic information for the system is available. I would like to establish an experiment or work with data where phylogenetic diversity was built into the experimental design to examine the relationship between phylogenetic diversity and species pairwise interactions more thoroughly.

Most of the Diversity-Interactions models presented in my thesis were fitted to grassland datasets (excluding the bacterial Bell dataset in chapter 3). In addition, the models presented primarily focused on single locations and single time points. I

am interested in continuing to explore the interpretations of the models I have developed (the Generalised Diversity-Interactions Mixed model and the Multivariate Diversity-Interactions model) across different ecosystem types, and across temporal and spatial effects.

The main area I intend to continue working in is in the modelling of the multifunctional BEF relationship. Initially I would like to examine additional properties of the Multivariate Diversity-Interactions model, for example by exploring the variance covariance structure more thoroughly. The model as fitted in my thesis assumes the same variance covariance matrix across sown density and proportion. To test the assumptions I would create a simulation study with replicate data across sown density and proportion to allow for the testing of different structures. The Multivariate Diversity-Interactions framework was tested on a dataset that had four species and three ecosystem functions at a single location and single growing season (section 4.3). I would like to test the model using data for higher numbers of species and higher numbers of ecosystem functions. Particular challenges that remain here are in the understanding of a higher number of effects, and in summarising and visualising the results. I also wish to examine the extension of the Multivariate Diversity-Interactions model to allow for a non-linear relationship (as in the Generalised Diversity-Interactions approach by Connolly et al. (2013)) between the ecosystem functions and the interactions between the species across a number of different datasets.

References

Alperovitch, A., Hesse, C., Lazar, P., Hayat, M. & Flamant, R. (1974). Temporal-spatial distributions of leukaemia and haematosarcoma in seven French regions. *International Journal of Epidemiology*, 3, 209-218.

Byrnes, J., Gamfeldt, L., Isbell, F., Lefcheck, J., Griffin, J., Hector, A. *et al.* (2014). Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions. *Methods in Ecology and Evolution*, 5, 111-124.

Cheverud, J., Wagner, G. & Dow, M. (1989). Methods for the comparative analysis of variation patterns. *Systematic Zoology*, 38, 201-213.

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.

Connolly, J., Cadotte, M., Brophy, C., Dooley, A., Finn, J., Kirwan, L. *et al.* (2011). Phylogenetically diverse grasslands are associated with pairwise interspecific processes that increase biomass. *Ecology*, 92, 1385-1392.

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

Hector, A. & Bagchi, R. (2007). Biodiversity and ecosystem multifunctionality. *Nature*, 448, 188-191.

Holm, L. & Sander, C. (1993). Protein-structure comparison by alignment of distance matrices. *Journal of Molecular Biology*, 233, 123-138.

Hooper, D. (1998). The role of complementarity and competition in ecosystem responses to variation in plant diversity. *Ecology*, 79, 704-719.

Ives, A.R. & Carpenter, S.R. (2007). Stability and diversity of ecosystems. *Science*, 317, 58-62.

Kirwan, L., Connolly, J., Finn, J., Brophy, C., Luscher, A., Nyfeler, D. *et al.* (2009). Diversity-Interaction modeling: estimating contributions of species identities and interactions to ecosystem function. *Ecology*, 90, 2032-2038.

Kirwan, L., Luescher, A., Sebastia, M., Finn, J., Collins, R., Porqueddu, C. *et al.* (2007). Evenness drives consistent diversity effects in intensive grassland systems across 28 European sites. *Journal of Ecology*, 95, 530-539.

Klauber, M.R. & Mustacch, P. (1970). Space-time clustering of childhood Leukaemia in San-Francisco. *Cancer Research*, 30, 1969-1973.

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

Maestre, F.T., Quero, J.L., Gotelli, N.J., Escudero, A., Ochoa, V., Delgado-Baquerizo, M. *et al.* (2012b). Plant species richness and ecosystem multifunctionality in global drylands. *Science*, 335, 214-218.

Manly, B. (1997). *Randomization, bootstrap and monte carlo methods in biology*. 2nd edn.

Mouillot, D., Villeger, S., Scherer-Lorenzen, M. & Mason, N. (2011). Functional structure of biological communities predicts ecosystem multifunctionality. *Plos One*, 6.

Piepho, H.P. (1998). Methods for comparing the yield stability of cropping systems. *Journal of Agronomy and Crop Science*, 180, 193-213.

Roscher, C., Weigelt, A., Proulx, R., Marquard, E., Schumacher, J., Weisser, W.W. *et al.* (2011). Identifying population- and community-level mechanisms of

diversity-stability relationships in experimental grasslands. *Journal of Ecology*, 99, 1460-1469.

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

Zavaleta, E., Pasari, J., Hulvey, K. & Tilman, G. (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.

Appendices

Appendix 2.1.1

Details of the experiments

The descriptions below are from the full experiments conducted. In some cases, only a subset of data was analysed because some plots did not establish and phylogenetic information was not available for all plots.

Dataset 1

Reference

Dimitrakopoulos, P. & Schmid, B. (2004). Biodiversity effects increase linearly with biotope space. *Ecology Letters,* 7, 574-583.

Species

The ten species sown were *Achillea millefolium*, *Arrhenatherum elatius*, *Dactylis glomerata*, *Festuca rubra*, *Holcus lanatus*, *Leucanthemum vulgare*, *Lotus corniculatus*, *Lychnis flos-cuculi*, *Plantago lanceolata* and *Trifolium pratense*.

Experimental design

There were 90 greenhouse pots used in the experiment; 30 were sown at each of three soil depths (5, 10, 15cm). At each depth the 30 pots were filled with 10

monocultures, 10 samples of 3 species communities and 10 samples of 6 species communities.

Dataset 2

Reference

Fridley, J. (2002). Resource availability dominates and alters the relationship between species diversity and ecosystem productivity in experimental plant communities. *Oecologia*, 132, 271-277.

Species

The nine species sown were *Achillea millefolium, Amaranthus hypochondriacus, Avena sativa, Borago officinalis, Calendula officinalis, Fagopyrum esculentum, Hypericum perforatum , Linum usitatissimum* and *Satureja hortensis*.

Experimental design

There were 360 grassland plots used in the experiment; 120 were sown at each of three levels of soil fertility (low, ambient and high fertility). At each soil fertility level 6 replicates of the 10 monocultures (60 plots), three replicates of ten 2-species communities (30 plots) and three replicates of ten 8-species communities were sown (30 plots).

Dataset 3

Reference

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

Species

The seven species sown were *Amaranthus hypochondriacus, Achillea millefolium, Borago officinalis, Calendula officinalis, Fagopyrum esculentum, Linum usitatissimum* and *Satureja hortensis.*

Experimental design

There were 252 grassland plots used in the experiment; 63 were sown across four treatment levels (low and high fertility soil crossed with a low and high shade light treatment). At each treatment level three replicates of seven monoculture, seven 2-species and seven 6-species communities were sown.

Dataset 4

Reference

Lanta, V. & Leps, J. (2006). Effect of functional group richness and species richness in manipulated productivity-diversity studies: a glasshouse pot experiment. *Acta Oecologica*, 29, 85-96.

Species

The sixteen species sown were *Festuca rubra, Trisetum flavescens, Alopecurus pratensis, Holcus lanatus, Lychnis flos-cuculi, Hypochaeris radicata, Plantago*

media, Leontodon autumnalis, Veronica officinalis, Glechoma hederacea, Fragaria vesca, Prunella vulgaris, Lotus corniculatus, Anthyllis vulneraria, Trifolium pratense and *Lathyrus pratensis*.

Experimental design

There were 200 greenhouse pots sown in the experiment. At each of two soil fertility levels (low and high) 100 pots were sown. At each treatment level two replicates of each of the 16 monocultures, 16 2-species communities, 24 4-species communities, 20 8-species communities and eight replicates of the 16-species communities were sown. Six pots were lost during the experiment.

Dataset 5

Reference

Naeem, S., Tjossem, S., Byers, D., Bristow, C. & Li, S. 1999. Plant neighborhood diversity and production. *Ecoscience,* 6**,** 355-365.

Species

The six species sown were *Vicia Villosa, Astragalus Canadenius, Panicum Virgatum, Bouteloua Gracilis, Rudbeckia Hirta* and *Achillea Millefolium*.

Experimental design

There were 360 greenhouse pots used in the experiment. Five replicates of the six monocultures, 34 2-species communities, 12 3-species communities and 20 4-species communities were sown. No additional treatments were applied to the pots.

Dataset 6

Reference

Naeem, S., Hakansson, K., Lawton, J., Crawley, M. & Thompson, L. 1996. Biodiversity and plant productivity in a model assemblage of plant species. *Oikos,* 76**,** 259-264.

Species

The 16 species sown were *Aphanes arvensis Rosaceae, Arabidopsis thaliana Cruciferae, Capsella bursa-pastoris Cruciferae, Cardamine hirsuta Cruciferae, Chenopodium album Chenopodiacae, Conyza canadensis Compositae, Lamium purpureum Labiatae, Poa annua Graminae, Senecio vulgaris Compositae, Sinapis arvensis Cruciferae, Sonchus oleraceus Compositae, Spergula arvensis Caryophyllaceae, Stellaria media Caryophyllaceae, Tripleurospermum inodorum Compositae, Veronica arvensis Scrophulariaceae* and *Veronica persica*.

Experimental design

There were 164 grassland plots sown in the experiment. Four replicates of each of the 16 monoculture, 20 2-species communities, 30 4-species communities, 40 8-species communities and 10 replicates of the 16-species community were sown. No additional treatments were applied to the plots.

Dataset 7

Reference

Craine, J., Reich, P., Tilman, G., Ellsworth, D., Fargione, J., Knops, J. & Naeem, S. 2003. The role of plant species in biomass production and response to elevated CO_2 and N. *Ecology Letters,* 6**,** 623-630.

Species

The 16 species sown were *Andropogon gerardii, Bouteloua gracilis, Schizachyrium scoparium, Sorghastrum nutans, Agropyron repens, Bromus inermis, Koeleria cristata, Poa pratensis, Amorpha canescens, Lespedeza capitata, Lupinus perennis, Petalostemum villosum, Achillea millefolium, Anemone cylindrica, Asclepias tuberosa* and *Solidago rigida*.

Experimental design

There were 232 grassland plots sown across four treatment levels (low and high applied CO_2 treatment crossed with a low and high shade light treatment). At each treatment level monocultures, 4-species, 9-species and 16-species grassland communities were sown.

Dataset 8

Reference

Tilman, D. 1997. Community invasibility, recruitment limitation, and grassland biodiversity. *Ecology,* 78, 81-92.

Species

The 21 species sown were *Achillea millefolium, Agropyron smithii, Amorpha canescens, Andropogon gerardi, Asclepias tuberosa, Elymus canadensis, Koeleria cristata, Lespedeza capitata, Liatris aspera, Lupinus perennis, Monarda fistulosa, Panicum virgatum, Petalostemum candidum, Petalostemum purpureum, Petalostemum villosum, Poa pratensis, Quercus ellipsoidalis, Quercus macrocarpa, Schizachyrium scoparium, Solidago rigida* and *Sorghastrum nutans*.

Experimental design

There were 150 grassland plots sown for the experiment. Thirty monoculture, 30 2-species, 30 4-species, 30 8-species and 30 16-species grassland communities were sown. Two plots were lost during the experiment.

Appendix 2.1.2

Residual diagnostics plots of model 3 for dataset 1

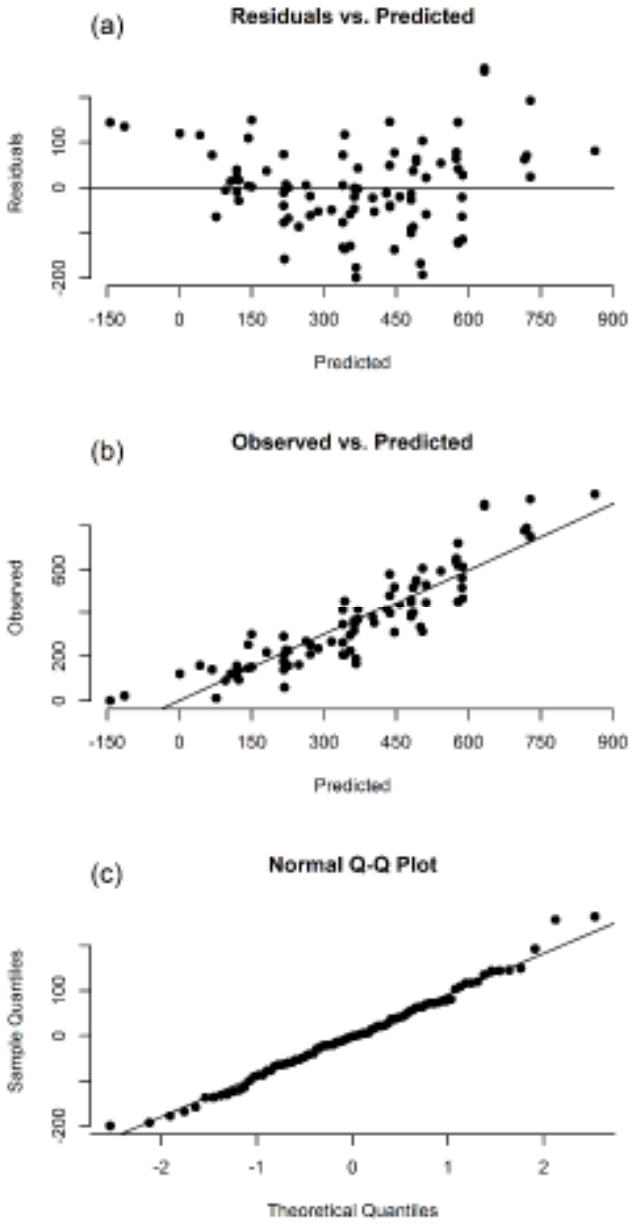

Fig A2.1.2-1: Residual diagnostics plots for dataset 1. The plots are (a) the residuals versus the predicted values; (b) the observed versus the predicted values and (c) the normal QQ plot of the residuals.

Appendix 2.1.3

Residual diagnostics plots of model 3 for dataset 2

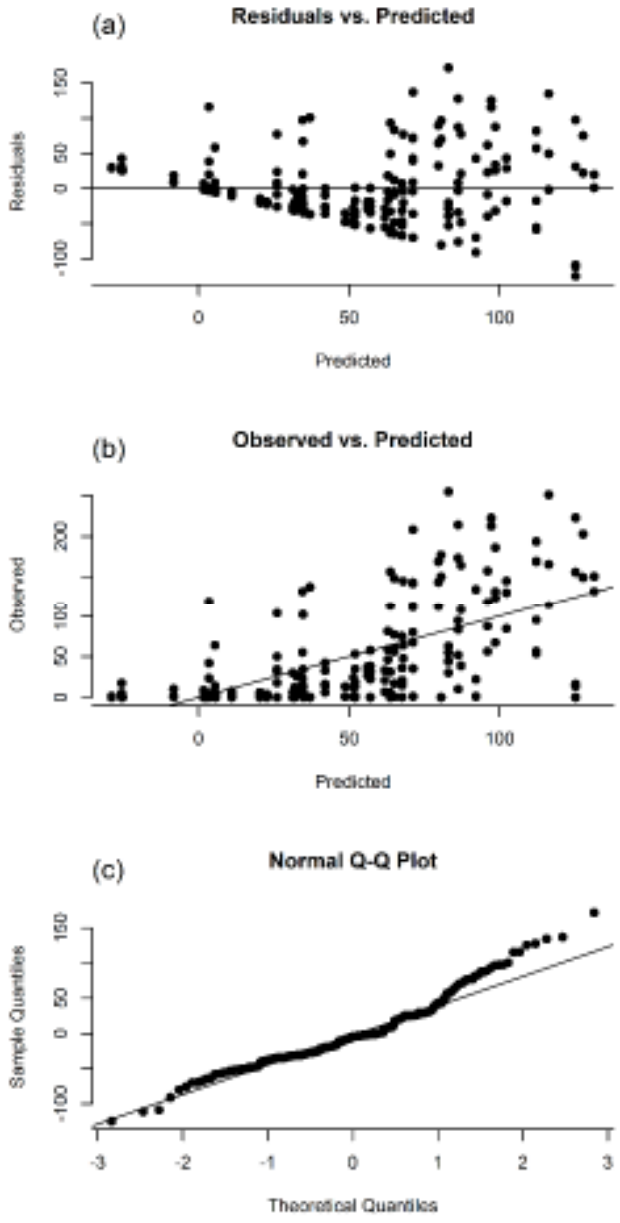

Fig A2.1.3-1: Residual diagnostics plots for dataset 2. The plots are (a) the residuals versus the predicted values; (b) the observed versus the predicted values and (c) the normal QQ plot of the residuals.

Appendix 2.1.4

Residual diagnostics plots of model 3 for dataset 3

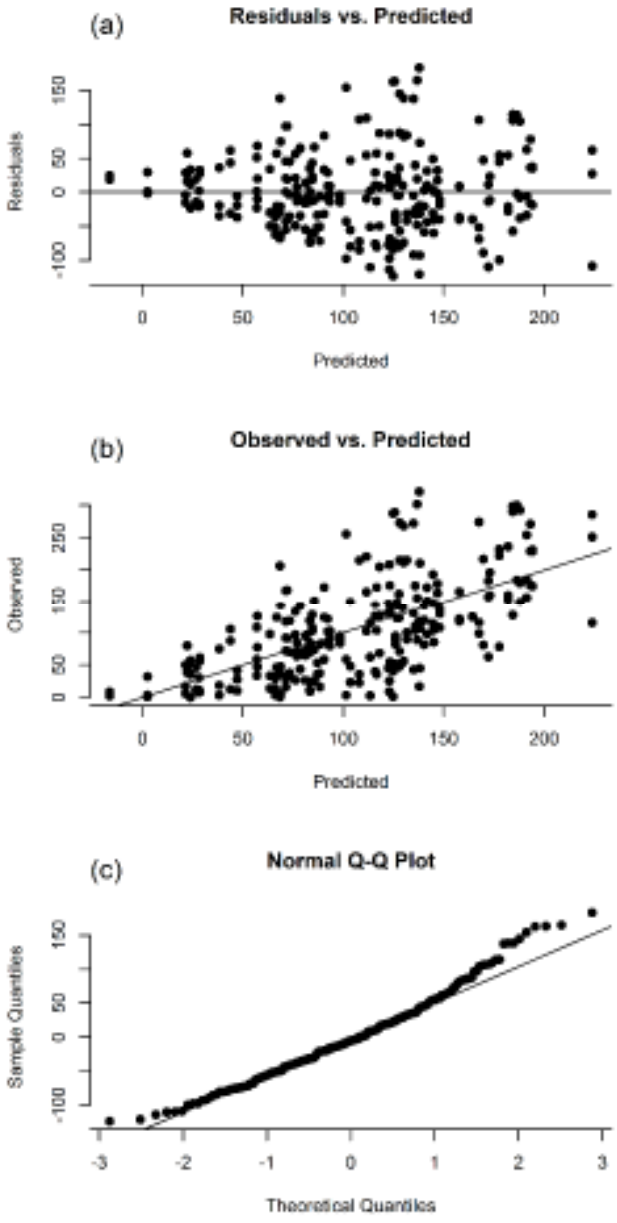

Fig A2.1.4-1: Residual diagnostics plots for dataset 3. The plots are (a) the residuals versus the predicted values; (b) the observed versus the predicted values and (c) the normal QQ plot of the residuals.

Appendix 2.1.5

Residual diagnostics plots of model 3 for dataset 4

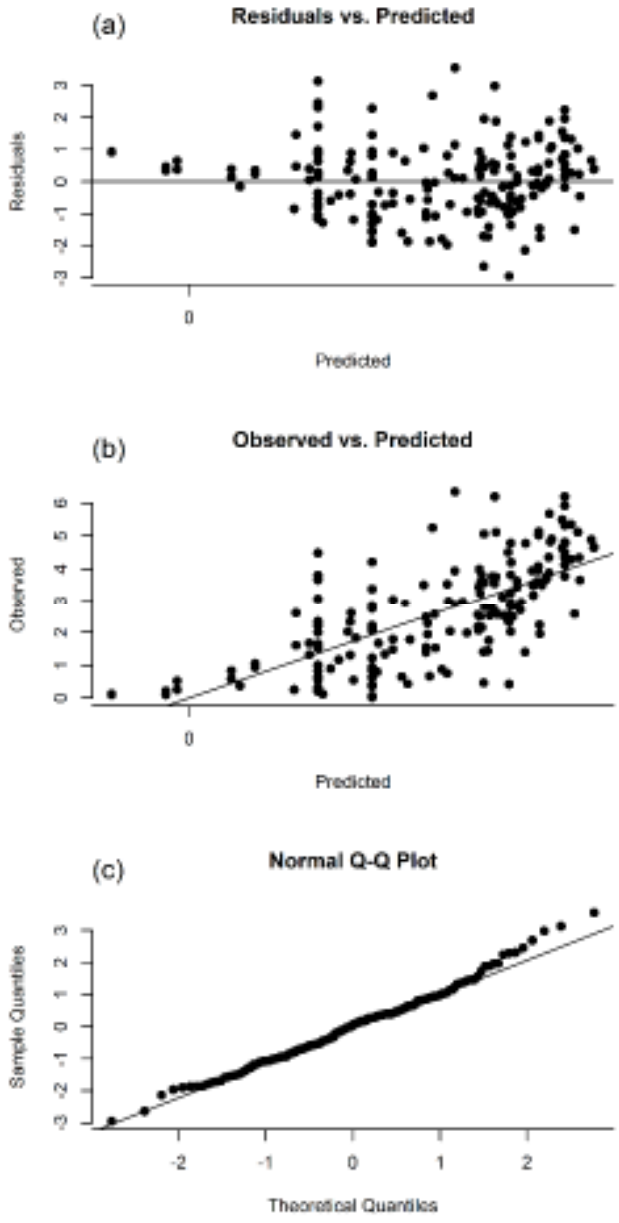

Fig A2.1.5-1: Residual diagnostics plots for dataset 4. The plots are (a) the residuals versus the predicted values; (b) the observed versus the predicted values and (c) the normal QQ plot of the residuals.

Appendix 2.1.6

Residual diagnostics plots of model 3 for dataset 5

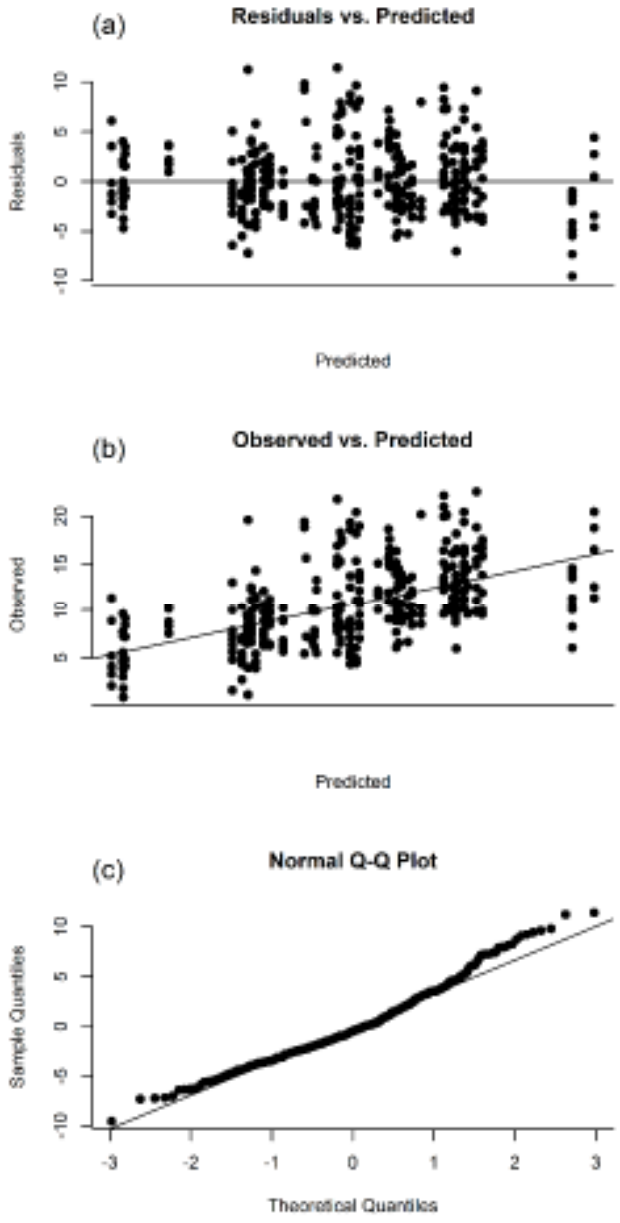

Fig A2.1.6-1: Residual diagnostics plots for dataset 5. The plots are (a) the residuals versus the predicted values; (b) the observed versus the predicted value and (c) the normal QQ plot of the residuals.

Appendix 2.1.7

Residual diagnostics plots of model 3 for dataset 6

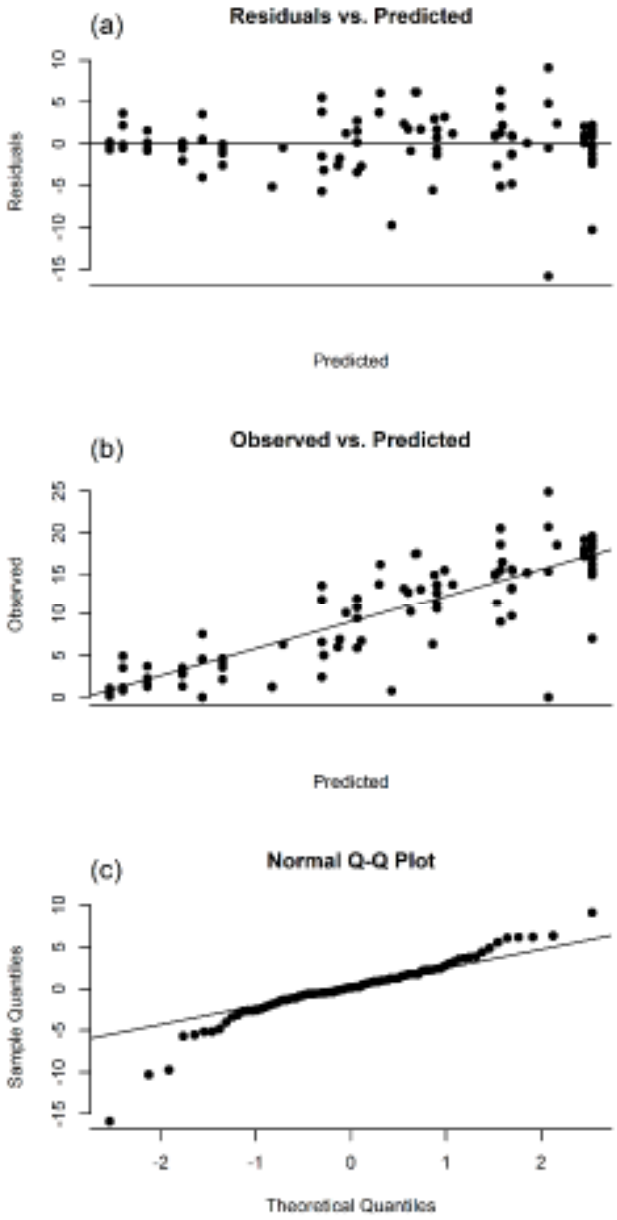

Fig A2.1.7-1: Residual diagnostics plots for dataset 6. The plots are (a) the residuals versus the predicted values; (b) the observed versus the predicted value and (c) the normal QQ plot of the residuals.

Appendix 2.1.8

Residual diagnostics plots of model 3 for dataset 7

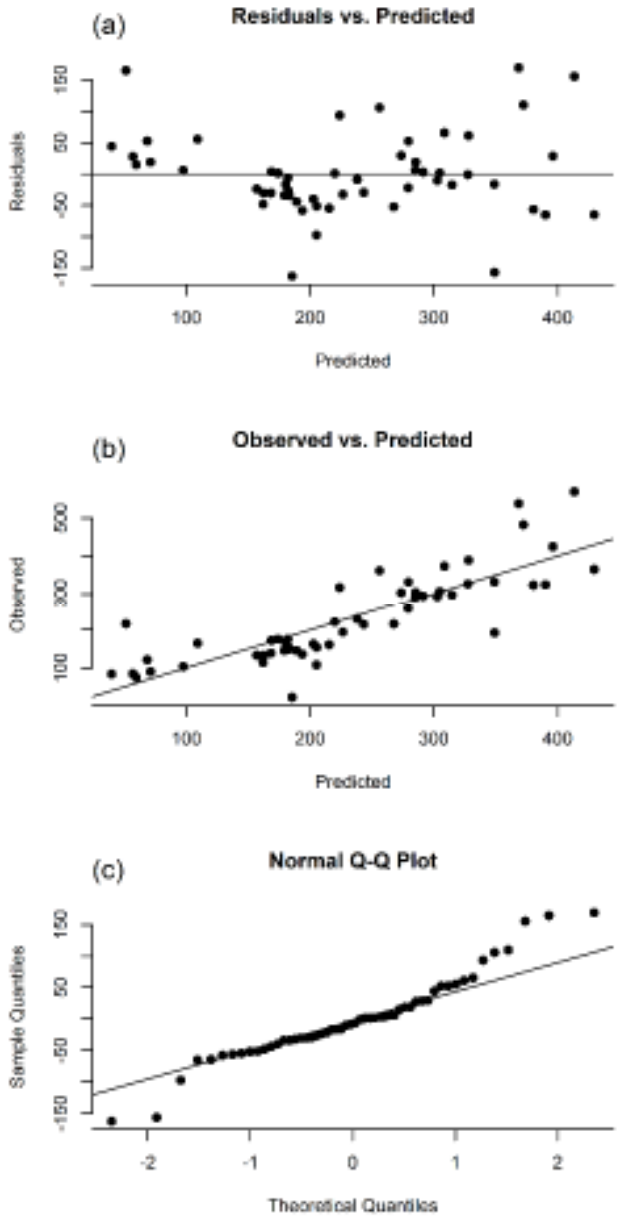

Fig A2.1.8-1: Residual diagnostics plots for dataset 7. The plots are (a) the residuals versus the predicted values; (b) the observed versus the predicted value and (c) the normal QQ plot of the residuals.

Appendix 2.1.9

Residual diagnostics plots of model 3 for dataset 8

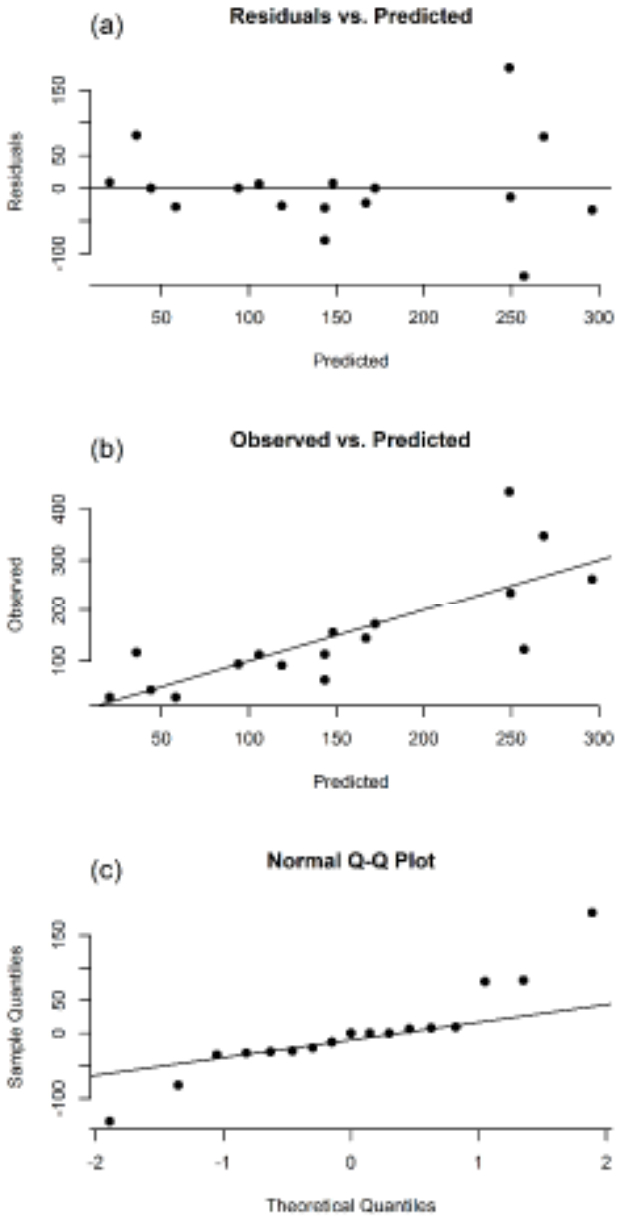

Fig A2.1.9-1: Residual diagnostics plots for dataset 8. The plots are (a) the residuals versus the predicted values; (b) the observed versus the predicted value and (c) the normal QQ plot of the residuals.

Appendix 2.1.10

The range of community-level phylogenetic diversities (C_D) for the eight datasets

		Phylogenetic diversity C_D		
Dataset	Reference	Min	Max	Range
1	Dimitrakopoulos & Schmid (2004)	-0.137	0.089	0.226
2	Fridley (2002)	-0.018	0.063	0.081
3	Fridley (2003)	-0.016	0.024	0.040
4	Lanta & Leps (2006)	-0.498	0.079	0.577
5	Naeem (1999)	-0.283	0.084	0.367
6	Naeem *et al.* (1996)	-0.334	0.111	0.445
7	Craine *et al.* (2003)	-0.080	0.103	0.183
8	Tilman (1997)	-0.248	0.083	0.331

Appendix 3.1.1

Algebraic specification of some Generalised Diversity-Interactions Mixed (GDIM) models.

P_i is the initial relative abundance of the i^{th} species and A is a treatment term.

Model 1a

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \sum_{\substack{i,j=1 \\ i<j}}^{s} \delta_{ij} (P_i P_j)^{\theta_1} + \varepsilon, \text{ where } \varepsilon \sim N(0, \sigma_1^2)$$

Model 2a

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} (P_i P_j)^{\theta_1} + \varepsilon, \text{ where } \varepsilon \sim N(0, \sigma_1^2).$$

Model 1b and 2b

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} (P_i P_j)^{\theta_1} + \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij} (P_i P_j)^{\theta_2} + \varepsilon, \text{ where } d_{ij} \sim N(0, \sigma_2^2) \text{ and}$$

$\varepsilon \sim N(0, \sigma_1^2)$.

Model 1c and 2c

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} (P_i P_j)^{\theta_1} + \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij} (P_i P_j)^{\theta_2} + \varepsilon, \text{ where } d_{ij} \sim N(0, \sigma_2^2),$$

$\varepsilon \sim N(0, \sigma_{1a}^2)$ for monocultures and $\varepsilon \sim N(0, \sigma_{1b}^2)$ for mixtures.

Model 1d and 2d

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} (P_i P_j)^{\theta_1} + \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij} (P_i P_j)^{\theta_2} + \varepsilon \text{, where } d_{ij} \sim N(0, \sigma_2^2),$$

$\varepsilon \sim N(0, \sigma_{1a}^2)$ for monocultures, $\varepsilon \sim N(0, f(z) * \sigma_{1b}^2)$ for mixtures, *z* is a community characteristic (*e.g.* richness) and *f(z)* is a function of *z*.

Assuming two functional groups of species, with species *1*,...,*t* in group 1 and species *t+1*,...,*s* in group 2.

Model 3a

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{wfg1} \sum_{\substack{i,j=1 \\ i<j}}^{t} (P_i P_j)^{\theta_1} + \delta_{wfg2} \sum_{\substack{i,j=t+1 \\ i<j}}^{s} (P_i P_j)^{\theta_1} + \delta_{bfg} \sum_{\substack{i \in (1,..,t) \\ j \in (t+1,..,s)}} (P_i P_j)^{\theta_1} + \varepsilon ,$$

where $\varepsilon \sim N(0, \sigma_1^2)$

Model 3b

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{wfg1} \sum_{\substack{i,j=1 \\ i<j}}^{t} (P_i P_j)^{\theta_1} + \delta_{wfg2} \sum_{\substack{i,j=t+1 \\ i<j}}^{s} (P_i P_j)^{\theta_1} + \delta_{bfg} \sum_{\substack{i \in (1,..,t) \\ j \in (t+1,..,s)}} (P_i P_j)^{\theta_1}$$

$$+ \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij} (P_i P_j)^{\theta_2} + \varepsilon$$

where $d_{ij} \sim N(0, \sigma_2^2)$ and $\varepsilon \sim N(0, \sigma_1^2)$.

Model 3c

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{wfg1} \sum_{\substack{i,j=1 \\ i<j}}^{t} (P_i P_j)^{\theta_1} + \delta_{wfg2} \sum_{\substack{i,j=t+1 \\ i<j}}^{s} (P_i P_j)^{\theta_1} + \delta_{bfg} \sum_{\substack{i\in(1,..,t) \\ j\in(t+1,..,s)}} (P_i P_j)^{\theta_1}$$

$$+ \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij} (P_i P_j)^{\theta_2} + \varepsilon$$

where $d_{ij} \sim N(0, \sigma_2^2)$, $\varepsilon \sim N(0, \sigma_{1a}^2)$ for monocultures and $\varepsilon \sim N(0, \sigma_{1b}^2)$ for mixtures.

Model 3d

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha A + \delta_{wfg1} \sum_{\substack{i,j=1 \\ i<j}}^{t} (P_i P_j)^{\theta_1} + \delta_{wfg2} \sum_{\substack{i,j=t+1 \\ i<j}}^{s} (P_i P_j)^{\theta_1} + \delta_{bfg} \sum_{\substack{i\in(1,..,t) \\ j\in(t+1,..,s)}} (P_i P_j)^{\theta_1}$$

$$+ \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij} (P_i P_j)^{\theta_2} + \varepsilon$$

where $d_{ij} \sim N(0, \sigma_2^2)$, $\varepsilon \sim N(0, \sigma_{1a}^2)$ for monocultures and $\varepsilon \sim N(0, f(z)*\sigma_{1b}^2)$ for mixtures.

Appendix 3.1.2

Additional information on the experiments for the two data sets.

Jena data set

There were 206 communities, each with one of six levels of species richness (1, 2, 3, 4, 6 and 9), established. The species pool for the Jena experiment data set consisted of nine species from three functional groups; five grasses (*Dactylis glomerata*, *Phleum pratense*, *Alopecurus pratensis*, *Poa trivialis*, *Arrhenatherum elatius*), two non-legume herbs (*Geranium pratense* and *Anthriscus sylvestris*), and two legumes (*Trifolium repens* and *Trifolium pratense*). At each species-richness level, each species appeared the same number of times and all possible 2-species combinations were present with the same frequency. Each community was replicated twice. The experimental area was partitioned into four blocks following a gradient of soil characteristics. In all communities, species present were equally represented at sowing. All plots were weeded regularly. The ecosystem function was yield (total aboveground biomass (g m^{-2})) in the year following establishment.

Bell data set

The bacterial ecosystems used were from semi-permanent rainpools that form in bark-lined depressions near the base of large European beech trees (*Fagus sylvatica*). These natural microcosms house an array of heterotrophic organisms, the energy for which is derived principally from beech leaf litter. Microcosms consisting of sterile

beech leaf disks and 10 ml of liquid (phosphate buffer) were inoculated with random combinations of 72 bacterial species isolated from these ecosystems. A total of 1,374 microcosms were constructed at richness levels of r = 1, 2, 3, 4, 6, 8, 9, 12, 18, 24, 36 and 72 species. For a given richness level (*r*) the 72 species were assigned to 72/*r* communities, each with *r* species, by randomly sampling without replacement from the 72 species, *e.g.* for *r* = 4 the 72 species are randomly partitioned into 18 communities of 4 species. This process was repeated five times and each selected composition was replicated twice. The daily respiration rate of the bacterial community in each microcosm was measured over three time intervals (days 0–7, 7–14 and 14–28) and the ecosystem function analyses here was the average over the three time intervals.

Appendix 3.1.3

Baseline model fitting for the Jena data set

(a) residual mean square error and residual degrees of freedom for each model and (b) model comparisons. Model 4 was used as the denominator for the F-tests.

(a)

Model		RMS	df
Model 0	$y = \alpha_k + \sum_{i=1}^{9} \beta_i P_i + \varepsilon$	30341	194
Model 1a	$y = \alpha_k + \sum_{i=1}^{9} \beta_i P_i + \sum_{\substack{i,j=1 \\ i<j}}^{9} \delta_{ij} P_i P_j + \varepsilon$	15241	158
Model 2a	$y = \alpha_k + \sum_{i=1}^{9} \beta_i P_i + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{9} P_i P_j + \varepsilon$	19204	193
Model 3a ($\theta = 1$)	$y = \alpha_k + \sum_{i=1}^{9} \beta_i P_i + \delta_{wfg1} \sum_{i<j}^{5} P_i P_j + \delta_{wfg2} P_6 P_7 + \delta_{wfg3} P_8 P_9 + \delta_{bfg12} \sum_{\substack{i=1,\ldots,5 \\ j=6,7}} P_i P_j + \delta_{bfg13} \sum_{\substack{i=1,\ldots,5 \\ j=8,9}} P_i P_j + \delta_{bfg23} \sum_{\substack{i=6,7 \\ j=8,9}} P_i P_j + \varepsilon$	17437	188
Model 4	$y = \alpha_k + \lambda_c + \varepsilon$	15710	103

Footnote: α_k are the terms for the block effects (k=1,..,4), P_i is the sown proportion of species i and λ_c is a term for each unique community composition.

(b)

Model comparison	F	p-value
Model 0 vs. 2a	138.74	<0.001
Model 2a vs. 1a	2.36	<0.001

Model 2a vs. 3a	5.45	<0.001
Model 3a vs. 1a	1.85	0.012

Footnote: Model 3a was also fitted with θ_I estimated using profile likelihood. The estimate was 0.96 which did not differ significantly from 1 (p=0.294, tested using a likelihood ratio test). Although the full pairwise interaction model 1a was a better fit than the functional group model (M3a vs 1a, p=0.012), in practice it will frequently not be possible to fit the full pairwise interaction model and so we chose the functional group model 3a (with θ_I=1) as the baseline model for comparison to the Generalised Diversity-Interactions Mixed (GDIM) models.

Appendix 3.1.4

Baseline model fitting for the Bell data set

(a) residual mean square error and residual degrees of freedom for each model and (b) model comparisons. Model 4 was used as the denominator for the F-test.

(a)

Model		RMS	df
Model 0	$y = \sum_{i=1}^{72} \beta_i P_i + \varepsilon$	8.72	1302
Model 2a ($\hat{\theta}_1 = 0.79$)	$y = \sum_{i=1}^{72} \beta_i P_i + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{72} (P_i P_j)^{0.79} + \varepsilon$	7.55	1300
Model 4	$y = \lambda_c + \varepsilon$	7.46	691

Footnote: P_i is the initial proportion of bacteria i and λ_c is a term for each unique community composition.

(b)

Model comparison	F	p-value
Model 0 vs. 2a	103.1	<0.001

Footnote: In model 2a, θ_1 was estimated to be 0.79 using profile likelihood and this provided a significant improvement over the model with θ_1 set to 1 ($p<0.001$, tested using a likelihood ratio test) therefore this model was chosen as the baseline model.

Appendix 3.1.5

Example of each of the Generalised Diversity-Interactions Mixed (GDIM) models (with $\theta_1 = 1$ and $\theta_2 = 1$) fitted to the Jena data set.

The Jena data set is detailed in Appendix 3.1.6 and the SAS code to fit the models illustrated in this appendix are in Appendix 3.1.7.

Model 3b

The model is

$$y = \alpha_k + \sum_{i=1}^{9} \beta_i P_i + \delta_{wfg1} \sum_{\substack{i,j=1 \\ i<j}}^{5} P_i P_j + \delta_{wfg2} P_6 P_7 + \delta_{wfg3} P_8 P_9$$
$$+ \delta_{bfg12} \sum_{\substack{i=1,..,5 \\ j=6,7}} P_i P_j + \delta_{bfg13} \sum_{\substack{i=1,..,5 \\ j=8,9}} P_i P_j + \delta_{bfg23} \sum_{\substack{i=6,7 \\ j=8,9}} P_i P_j + \sum_{\substack{i,j=1 \\ i<j}}^{9} d_{ij} P_i P_j + \varepsilon$$

where $d_{ij} \sim N(0, \sigma_2^2)$ and $\varepsilon \sim N(0, \sigma_1^2)$ and where α_k are the terms for the block effect (k=1,..,4) and P_i is the sown proportion of species i, (i=1,..,9). The β_i are the identity effects for each species, δ_{wfg1}, δ_{wfg2} and δ_{wfg3} are the within functional group interaction coefficients for the functional groups grasses, non-legume herbs and legume respectively and δ_{bfg12}, δ_{bfg13} and δ_{bfg23} are the between functional group coefficient. The model coefficient estimates are presented in Table A3.1.5-1.

Table A3.1.5-1. Coefficient estimates and standard errors (SE) for fixed estimates for model 3b fitted to the Jena data set. Variance estimates are also shown.

Fixed coefficients			Variance estimates	
Coefficient	Estimate	SE	Coefficient	Estimate
α_1	168.69	72.298	σ_1^2	15311
α_2	278.91	74.059	σ_2^2	90101
α_3	253.91	75.206		
α_4	216.56	74.085		
β_1	313.92	98.756		
β_2	268.13	99.576		
β_3	97.80	98.956		
β_4	1.28	98.688		
β_5	536.20	98.972		
β_6	-184.22	103.430		
β_7	-234.74	103.670		
β_8	97.20	85.063		
β_9	0.00	.		
δ_{wfg1}	621.75	165.180		
δ_{wfg2}	605.54	460.100		
δ_{wfg3}	936.38	457.860		
δ_{bfg12}	1328.34	158.890		
δ_{bfg13}	913.01	159.330		
δ_{bfg23}	789.30	240.980		

A community sown with 50:50 proportions of species 3 and 4 in block 2 has a predicted response

$$\begin{aligned}\hat{y} &= \hat{\alpha}_2 + \hat{\beta}_3 P_3 + \hat{\beta}_4 P_4 + \hat{\delta}_{wfg1} P_3 P_4 \\ &= 278.91 + (97.80 * 0.5) + (1.29 * 0.5) + 621.75 * (0.5 * 0.5) \\ &= 483.89\end{aligned}$$

Model 3c

The model is

$$y = \alpha_k + \sum_{i=1}^{9} \beta_i P_i + \delta_{wfg1} \sum_{\substack{i,j=1 \\ i<j}}^{5} P_i P_j + \delta_{wfg2} P_6 P_7 + \delta_{wfg3} P_8 P_9$$

$$+ \delta_{bfg12} \sum_{\substack{i=1,\ldots,5 \\ j=6,7}} P_i P_j + \delta_{bfg13} \sum_{\substack{i=1,\ldots,5 \\ j=8,9}} P_i P_j + \delta_{bfg23} \sum_{\substack{i=6,7 \\ j=8,9}} P_i P_j + \sum_{\substack{i,j=1 \\ i<j}}^{9} d_{ij} P_i P_j + \varepsilon$$

where $d_{ij} \sim N(0, \sigma_2^2)$, $\varepsilon \sim N(0, \sigma_{1a}^2)$ for monocultures and $\varepsilon \sim N(0, \sigma_{1b}^2)$ for mixtures. The model coefficient estimates are in Table A3.1.5-2.

Table A3.1.5-2. Parameter estimates with standard errors (SE) for fixed estimates for model 3c fitted to the Jena data set. Variance estimates are also shown.

Fixed coefficients			Variance estimates	
Coefficient	Estimate	SE	Coefficient	Estimate
α_1	166.61	50.029	σ_{1a}^2	5746
α_2	279.52	51.193	σ_{1b}^2	15850
α_3	258.44	53.439	σ_2^2	103918
α_4	215.35	52.663		
β_1	308.40	69.104		
β_2	306.43	69.679		
β_3	125.19	69.729		
β_4	5.39	68.407		
β_5	492.96	69.745		
β_6	-152.07	70.271		
β_7	-229.19	70.483		
β_8	109.81	64.885		
β_9	0.00	.		
δ_{wfg1}	607.80	147.360		
δ_{wfg2}	547.82	454.530		
δ_{wfg3}	916.98	452.480		
δ_{bfg12}	1291.05	144.460		
δ_{bfg13}	900.74	145.330		
δ_{bfg23}	744.00	229.170		

Model 3d_richness

The model is

$$y = \alpha_k + \sum_{i=1}^{9} \beta_i P_i + \delta_{wfg1} \sum_{\substack{i,j=1 \\ i<j}}^{5} P_i P_j + \delta_{wfg2} P_6 P_7 + \delta_{wfg3} P_8 P_9$$

$$+ \delta_{bfg12} \sum_{\substack{i=1,\ldots,5 \\ j=6,7}} P_i P_j + \delta_{bfg13} \sum_{\substack{i=1,\ldots,5 \\ j=8,9}} P_i P_j + \delta_{bfg23} \sum_{\substack{i=6,7 \\ j=8,9}} P_i P_j + \sum_{\substack{i,j=1 \\ i<j}}^{9} d_{ij} P_i P_j + \varepsilon$$

where $d_{ij} \sim N(0, \sigma_2^2)$, $\varepsilon \sim N(0, \sigma_{1a}^2)$ for monocultures and $\varepsilon \sim N(0, f(z) * \sigma_{1b}^2)$ for mixtures. Here *f(z)*=richness$^{\gamma}$ where γ was estimated by profile likelihood as $\hat{\gamma} = -0.3$. The model coefficient estimates are in Table A3.1.5-3.

Table A3.1.5-3. Parameter estimates and standard errors (SE) for fixed estimates for model 3d_richness fitted to the Jena data set. Variance estimates are also shown.

Fixed coefficients			Variance estimates	
Coefficient	Estimate	SE	Coefficient	Estimate
α_1	167.93	49.949	σ_{1a}^2	5758
α_2	277.86	51.138	σ_{1b}^2	10989
α_3	255.92	53.283	σ_2^2	109778
α_4	214.28	52.516		
β_1	312.60	68.999		
β_2	308.07	69.582		
β_3	127.20	69.583		
β_4	6.93	68.324		
β_5	492.38	69.588		
β_6	-151.32	70.185		
β_7	-227.14	70.402		
β_8	109.54	64.694		
β_9	0.00	.		
δ_{wfg1}	600.07	148.620		
δ_{wfg2}	512.63	450.760		
δ_{wfg3}	922.18	449.180		
δ_{bfg12}	1285.73	145.570		
δ_{bfg13}	904.28	146.080		
δ_{bfg23}	730.08	228.500		

Model 3d_evenness

The model is

$$y = \alpha_k + \sum_{i=1}^{9} \beta_i P_i + \delta_{wfg1} \sum_{\substack{i,j=1 \\ i<j}}^{5} P_i P_j + \delta_{wfg2} P_6 P_7 + \delta_{wfg3} P_8 P_9$$

$$+ \delta_{bfg12} \sum_{\substack{i=1,\ldots,5 \\ j=6,7}} P_i P_j + \delta_{bfg13} \sum_{\substack{i=1,\ldots,5 \\ j=8,9}} P_i P_j + \delta_{bfg23} \sum_{\substack{i=6,7 \\ j=8,9}} P_i P_j + \sum_{\substack{i,j=1 \\ i<j}}^{9} d_{ij} P_i P_j + \varepsilon$$

Where $d_{ij} \sim N(0, \sigma_2^2)$, $\varepsilon \sim N(0, \sigma_{1a}^2)$ for monocultures and $\varepsilon \sim N(0, f(z) * \sigma_{1b}^2)$ for mixtures. Here *f(z)*=evenness$^{\gamma}$ where γ was estimated by profile likelihood as $\hat{\gamma} = -0.6$ and evenness = *(2s/(s-1))*$*\Sigma_{i<j}P_iP_j$. The model coefficient estimates are in Table A3.1.5-4.

Table A3.1.5-4. Parameter estimates and standard errors (SE) for fixed estimates for model 3d_evenness fitted to the Jena data set. Variance estimates are also shown.

Fixed coefficients			Variance estimates	
Coefficient	Estimate	SE	Coefficient	Estimate
α_1	168.06	49.949	σ_{1a}^2	5763
α_2	277.70	51.141	σ_{1b}^2	18908
α_3	255.85	53.292	σ_2^2	109541
α_4	213.67	52.524		
β_1	313.12	69.007		
β_2	307.89	69.591		
β_3	127.31	69.589		
β_4	6.92	68.329		
β_5	492.51	69.594		
β_6	-151.60	70.196		
β_7	-227.10	70.411		
β_8	109.57	64.680		
β_9	0.00	.		
δ_{wfg1}	600.50	148.590		
δ_{wfg2}	515.99	450.460		
δ_{wfg3}	928.15	448.950		
δ_{bfg12}	1289.06	145.560		
δ_{bfg13}	905.12	146.040		
δ_{bfg23}	728.16	228.420		

Appendix 3.1.6

Jena dataset

The Jena dataset has been given in the zipped folder ecy1872-sup-0007-DataS1.zip at http://onlinelibrary.wiley.com/doi/10.1002/ecy.1872/suppinfo

Variable descriptions

Community : community number (each community is sown at least twice)

Block : soil gradient blocking B1-B4

Richness : sown richness

Biomass : total aboveground biomass (g m^{-2}) for the plot

p1 - p9 : Sown proportions for species 1 to 9

monomix : 1 if the plot is a sown monoculture, 0 otherwise

int1 - int36 : pairwise species interaction (P_iP_j)

PPsum : sum of int1-int36

wfg1 – wfg3 : within functional group, $\sum_{\substack{i,j=1 \\ i<j}}^{s} P_iP_j$ where *i,j* are both from one functional group

bfg12 - bfg23 : between functional group, $\sum_{\substack{i,j=1 \\ i<j}}^{s} P_iP_j$ where *i,j* are from different functional groups

The following variables need to be added manually to the dataset:

E : Evenness = $(2s/(s-1))*\Sigma P_i P_j$, where $s = 9$

z_richness : $\text{Richness}^{-0.3}$

z_evenness : $E^{-0.6}$

Appendix 3.1.7

SAS code to fit the Generalised Diversity-Interactions Mixed models 3b, 3c and 3d to the Jena data set.

```
/*********************************************************/
/* READING IN THE DATA SET TO BE USED IN ANALYSIS */
PROC IMPORT OUT= jena
      DATAFILE= "C:\...\ Appendix316_Jena.csv"
      DBMS=CSV REPLACE;
    GETNAMES=YES;
RUN;
/* CREATING A DATASET NEEDED FOR THE RANDOM EFFECTS SPECIFICATION*/
data Jena_pairwise;
      do i=1 to 36;
            parm=1;
            row=i;
            col=i;
            value=1;
            output;
      end;
      drop i;
run;
/*********************************************************/
/*********************************************************/
```

```
/*********************************************************/
/*********************************************************/
/* FITTING THE RANGE OF GDIM MODELS */
*MODEL 3B WITH THETA1=1 AND THETA2=1;
proc mixed data= jena method=reml;
      class block;
      model Biomass=block p1-p9 wfg1 wfg2 wfg3 bfg12 bfg13 bfg23/ noint
s;
      random int1-int36/ type=lin(1) ldata=Jena_pairwise;
run;
/* MODEL 3C WITH THETA1=1 AND THETA2=1 */
proc mixed data= jena method=reml;
      class block monomix;
      model Biomass=block p1-p9 wfg1 wfg2 wfg3 bfg12 bfg13 bfg23/ noint
s;
      random int1-int36/ type=lin(1) ldata=Jena_pairwise;
      repeated /group=monomix;
run;
/* MODEL 3D RICHNESS WITH THETA1=1 AND THETA2=1 */
/* GAMMA WAS ESTIMATED USING PROFILE LIKELIHOOD AS -0.3, GIVING
Z_RICHNESS = RICHNESS^-0.3 */
proc mixed data= jena method=reml;
      class block monomix;
      model Biomass=block p1-p9 wfg1 wfg2 wfg3 bfg12 bfg13 bfg23/ noint
s;
      random int1-int36/ type=lin(1) ldata=Jena_pairwise;
      repeated /group=monomix;
      weight z_richness;
run;
```

```
/* MODEL 3D EVENNESS WITH THETA1=1 AND THETA2=1 */
/* GAMMA WAS ESTIMATED USING PROFILE LIKELIHOOD AS -0.6, GIVING
Z_EVENNESS = EVENNESS^-0.6 */
proc mixed data=jena method=reml;
      class block monomix;
      model Biomass=block p1-p9 wfg1 wfg2 wfg3 bfg12 bfg13 bfg23/ noint
s;
      random int1-int36/ type=lin(1) ldata=Jena_pairwise;
      repeated /group=monomix;
      weight z_evenness;
run;

/*************************************************************/
/*************************************************************/
```

Appendix 3.1.8

The Generalised Diversity-Interactions Mixed model with identity effects as random.

Additional to the assumptions in models 1b to 1d, it is possible to assume that the species identity effect coefficients vary randomly around the average species identity effect. This assumption may be useful in cases where the individual species identity effects are not the main interest of the study or where the high number of species make it difficult or impossible to estimate the individual species identity effects. Including this assumption in model 1c gives

$$y = \beta_{av} + \sum_{i=1}^{s} b_i P_i + \alpha A + \delta_{av} \sum_{\substack{i,j=1 \\ i<j}}^{s} (P_i P_j)^{\theta_1} + \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij} (P_i P_j)^{\theta_2} + \varepsilon$$

where $b_i \sim N(0, \sigma_3^2)$, $d_{ij} \sim N(0, \sigma_2^2)$, $\varepsilon \sim N(0, \sigma_{1a}^2)$ for monocultures and $\varepsilon \sim N(0, \sigma_{1b}^2)$ for mixtures.

Appendix 4.2.1

Raw and standardised ecosystem function values for compositions 2,7,21 and 24 in block 1 that are highlighted in Figure 4.2.2.

	Raw values				
Comp	Biomass	Root biomass	N pool	Soil N	Cotton Decomp
2	1374.4	204.91	937.2	7.18	0.52
7	654.6	359.69	505.4	8.10	0.67
21	1075.20	623.09	877.20	7.13	0.53
24	880.20	961.69	619.00	5.29	0.59

	Standardised values				
Comp	Yield	Root	Nveg	Nsoil	Decomp
2	98.80	12.62	86.14	71.77	73.42
7	47.05	22.15	46.45	67.06	95.27
21	77.29	38.37	80.62	72.03	75.96
24	63.27	59.22	56.89	81.45	84.46

Comp	Average metric	Pooled std. dev	SAM metric
2	68.55	30.389	2.26
7	55.60	24.460	2.27
21	68.85	21.969	3.13
24	69.06	17.439	3.96

Footnote: Comp=Composition number. The five functions are: Aboveground biomass (Biomass), Root biomass (Root biomass), Aboveground nitrogen pool (N pool), Unconsumed soil nitrogen (Soil N) and Cotton decomposition (Cotton Decomp). Pooled standard deviations were computed using the data from each pair of communities with the same composition as described in text.

Appendix 4.2.2

Model fitting for the SAM and average metrics

(a) Specification of the models fitted, (b) model fitting for the SAM metric and (c) model fitting for the average metric.

(a) Models fitted

Model #	Details

1 Identity effects Diversity-Interaction model with block effect

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha_b + \varepsilon$$

2 Average diversity effects Diversity-Interaction model

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha_b + \delta E + \varepsilon$$

3 Functional group diversity effects Diversity-Interaction model

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha_b + \delta_{wfg1} \sum_{i<j}^{4} P_i P_j + \delta_{wfg2} \sum_{\substack{i<j \\ i=5}}^{8} P_i P_j + \delta_{wfg3} P_9 P_{10}$$
$$+ \delta_{bfg12} \sum_{\substack{i=1,\ldots,4 \\ j=5,\ldots,8}}^{7} P_i P_j + \delta_{bfg13} \sum_{\substack{i=1,\ldots,4 \\ j=9,10}} P_i P_j + \delta_{bfg23} \sum_{\substack{i=5,\ldots,8 \\ j=9,10}} P_i P_j + \varepsilon$$

4 Average quadratic diversity effects Diversity-Interaction model

$$y = \sum_{i=1}^{s} \beta_i P_i + \alpha_b + \delta E + \delta_{quad} (E)^2 + \varepsilon$$

(b)

Comparison	Terms tested	F	p
Model 2 vs. 1	δ	**40.84**	**<0.001**
Model 3 vs. 2	δ_{wfg1}, δ_{wfg2}, δ_{wfg3}, δ_{bfg1}, δ_{bfg2}, δ_{bfg3}	0.83	0.533
Model 4 vs. 2	δ_{quad}	0.17	0.679

(c)

Comparison	Terms tested	F	p
Model 2 vs. 1	δ	**26.46**	**<0.001**
Model 3 vs. 2	δ_{wfg1}, δ_{wfg2}, δ_{wfg3}, δ_{bfg1}, δ_{bfg2}, δ_{bfg3}	2.36	0.054
Model 4 vs. 2	δ_{quad}	3.00	0.089

Significant F tests in parts (b) and (c) are highlighted in bold.

Appendix 4.3.1

A brief description of multifunctionality methods previously used and the results that have been found in their applications.

(1) The approach that combines univariate models for each function (Allan *et al*. 2013; Cardinale *et al*. 2013; Orwin *et al*. 2014) maintains quantitative information at the individual ecosystem function level and qualitatively discusses ecosystem multifunctionality. Studies have shown that levels of biodiversity had a significant effect on the ability of the ecosystem to maintain multiple functions in grassland (Allan *et al*. 2013; Cardinale *et al*. 2013; Orwin *et al*. 2014) and freshwater algae (Cardinale *et al*. 2013) ecosystems.

(2) The averaging approach (Mouillot *et al*. 2011; Maestre *et al*. 2012a; Maestre *et al*. 2012b) standardises all ecosystem functions, computes the average metric (*i.e.* the average of all ecosystem functions for each community) and uses univariate techniques to analyse it. It has shown that species loss tends to reduce the average levels of multiple ecosystem functions in grassland (Mouillot *et al*. 2011), dryland (Maestre *et al*. 2012b) and lichen (Maestre *et al*. 2012a) ecosystems.

(3) The overlap method (Hector & Bagchi 2007; Isbell *et al*. 2011) identifies a set of species that affect each individual ecosystem function and then quantifies the overlap between sets of species influencing pairs of ecosystem functions. It has

shown that different sets of species can affect different ecosystem functions in grassland ecosystems.

(4) The single threshold method (Gamfeldt *et al*. 2008; Zavaleta *et al*. 2010) examines whether multiple ecosystem functions surpass a threshold at each level of diversity. It has shown that higher richness levels are necessary to achieve high values of multiple ecosystem functions in grassland (Gamfeldt *et al*. 2008; Zavaleta *et al*. 2010), bacterial (Gamfeldt *et al*. 2008) and algae (Gamfeldt *et al*. 2008) ecosystems.

(5) The multiple threshold method is an extension of the single threshold approach that systematically explores all possible thresholds, rather than an arbitrary subset of thresholds (Byrnes *et al*. 2014). It has shown the thresholds at which diversity yields high or medium values of multiple ecosystem functions in grassland ecosystems.

References (for Appendix 4.3.1)

1.

Allan, E., Weisser, W., Fischer, M., Schulze, E.-D., Weigelt, A., Roscher, C. *et al.* (2013). A comparison of the strength of biodiversity effects across multiple functions. *Oecologia*, 173, 223-237.

2.

Byrnes, J.E.K., Gamfeldt, L., Isbell, F., Lefcheck, J.S., Griffin, J.N., Hector, A. *et al.* (2014). Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions. *Methods Ecol. Evol.*, 5, 111-124.

3.

Cardinale, B.J., Gross, K., Fritschie, K., Flombaum, P., Fox, J.W., Rixen, C. *et al.* (2013). Biodiversity simultaneously enhances the production and stability of community biomass, but the effects are independent. *Ecology*, 94, 1697-1707.

4.

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

5.

Hector, A. & Bagchi, R. (2007). Biodiversity and ecosystem multifunctionality. *Nature*, 448, 188-191.

6.

Isbell, F., Calcagno, V., Hector, A., Connolly, J., Harpole, W.S., Reich, P.B. *et al.* (2011). High plant diversity is needed to maintain ecosystem services. *Nature*, 477, 199-202.

7.

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

8.

Maestre, F.T., Quero, J.L., Gotelli, N.J., Escudero, A., Ochoa, V., Delgado-Baquerizo, M. *et al.* (2012b). Plant species richness and ecosystem multifunctionality in global drylands. *Science*, 335, 214-218.

9.

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

10.

Orwin, K., Ostle, N., Wilby, A. & Bardgett, R. (2014). Effects of species evenness and dominant species identity on multiple ecosystem functions in model grassland communities. *Oecologia*, 174, 979-992.

11.

Zavaleta, E.S., Pasari, J.R., Hulvey, K.B. & Tilman, G.D. (2010). Sustaining multiple ecosystem functions in grassland communities requires higher biodiversity. *P. Natl. Acad. Sci. USA*, 107, 1443-1446.

Appendix 4.3.2

Issues associated with transforming ecosystem functions.

We transformed our raw data on ecosystem function variables by computing current value*100 / average of maximum three values (top 10%) for each response. We note that despite the averaging of the top 10% of values, this transformation could still be strongly influenced by outlier values which could have knock on consequences for model estimation and interpretation. In our case outlier values were not an issue but this should be considered when using this type of transformation. Our re-scaling implicitly assumes that each function has equal importance which in practice may not be true. Various ecosystem functions could easily be weighted by relative importance, when known, for particular applications (Alsterberg *et al*. 2014). One way to do this is to compute the average of a different percentile (lower than the top 10^{th}) of the data for a function that was deemed to be less important, with the choice of what percentile decided by how much 'less important' the function was considered to be. Subsequent graphical bar chart presentations (*e.g.*, as in Fig. 4.3.2 in the main text) could be adjusted so that the less important function did not have an artificially higher response by widening its bar while still maintaining the appropriate area for the bar.

An alternative transformation is to standardise each function to have the same mean and standard deviation but this option still contains the issues associated with weighting as discussed above. It also forces all functions to have the same variability which may not be desirable to carry out true comparisons.

Our choice to firstly invert weed biomass but not sown biomass or N yield was subjective as the favoured direction of any given ecosystem function could vary depending on the stakeholder's interest. Generally from an agronomic perspective, weed biomass suppression but high values of sown biomass and N yield are desirable hence why we choice to invert weed biomass prior to transformation.

Reference (for Appendix 4.3.2)

1.

Alsterberg, C., Sundbäck, K. & Gamfeldt, L. (2014). Multiple stressors and multifunctionality: limited effects on an illuminated benthic system. *Biology Letters*, 10.

Appendix 4.3.3

Belgium dataset

The Belgium data set has been given in ele12504-sup-0002-AppendixS3.txt at

http://onlinelibrary.wiley.com/doi/10.1111/ele.12504/full

Variable descriptions

Plot: unique plot identification

G1: sown proportion of G1

G2: sown proportion of G2

L1: sown proportion of L1

L2: sown proportion of L2

E: sown evenness

Density: -1 for low, +1 for high

G1G2: product of G1 and G2

(Similar for the product of each pair of species)

PiPj_sum = G1G2+G1L1+G1L2+G2L1+G2L2+L1L2

Wfg1 = G1G2

Wfg2 = L1L2

Bfg = G1L1+G1L2+G2L1+G2L2

Var: factor indicating ecosystem function type, Sown, Weed and N

Var_num: 1=Sown, 2=Weed, 3=N

Y: ecosystem function value (%)

Appendix 4.3.4

SAS and R code for fitting the framework.

SAS code

```
/*************************************************************************
/
/* READING IN THE DATA */
proc import out = BelData
      Datafile= "C:\Appendix433_Beldata.txt"
      Dbms=dlm replace;
    Delimiter='09'x;
    Datarow=2;
run;
/*************************************************************************
/


/*************************************************************************
/
/* FULL PAIRWISE INTERACTIONS MULTIVARIATE DIVERSITY-INTERACTIONS MODEL
FITTED USING ML*/
proc mixed data=BelData method=ml;
      class Var;
      model Y = Var*G1 Var*G2 Var*L1 Var*L2 Var*Density
                  Var*G1G2 Var*G1L1 Var*G1L2 Var*G2L1 Var*G2L2 Var*L1L2
                  / noint solution;
      repeated Var / subject=Plot type=un r;
run;
/*************************************************************************
/


/*************************************************************************
/
/* FUNCTIONAL GROUP INTERACTIONS MULTIVARIATE DIVERSITY-INTERACTIONS
MODEL
FITTED USING ML*/
proc mixed data=BelData method=ml;
      class Var;
      model Y = Var*G1 Var*G2 Var*L1 Var*L2 Var*Density
                  Var*Wfg1 Var*Wfg2 Var*Bfg
                  / noint solution;
      repeated Var / subject=Plot type=un r;
run;
/*************************************************************************
/
```

```
/*************************************************************************
/
/* FUNCTIONAL GROUP INTERACTIONS MULTIVARIATE DIVERSITY-INTERACTIONS
MODEL
FITTED USING REML */
proc mixed data=BelData method=reml;
      class Var;
      model Y = Var*G1 Var*G2 Var*L1 Var*L2 Var*density
            Var*wfg1 Var*wfg2 Var*bfg /noint solution;
      repeated Var/subject=Plot type=un r;

      /* Testing for differences between each pair of ecosystem
functions
      in the Beta_1 coefficient of G1*/
      estimate 'G1 Sown vs Weed' Var*G1 0 1 -1;
      estimate 'G1 Sown vs N ' Var*G1 -1 1 0;
      estimate 'G1 Weed vs N ' Var*G1 -1 0 1;

      /* Predicting each ecosystem function at the centroid community*/
      estimate 'Pred y at centroid Sown'
      Var*G1 0 0.25 0 Var*G2 0 0.25 0 Var*L1 0 0.25 0 Var*L2 0 0.25 0
      Var*wfg1 0 0.0625 0 Var*wfg2  0 0.0625 0 Var*bfg 0 0.25 0;
      estimate 'Pred y at centroid Weed'
      Var*G1 0 0 0.25 Var*G2 0 0 0.25 Var*L1 0 0 0.25 Var*L2 0 0 0.25
      Var*wfg1 0 0 0.0625 Var*wfg2  0 0 0.0625 Var*bfg 0 0 0.25;
      estimate 'Pred y at centroid N '
      Var*G1 0.25 0 0 Var*G2 0.25 0 0 Var*L1 0.25 0 0 Var*L2 0.25 0 0
      Var*wfg1 0.0625 0 0 Var*wfg2  0.0625 0 0 Var*bfg 0.25 0 0;

      /* Testing for a difference between each pair of ecosystem
functions
      in the predicted response at the centroid community
      (0.25, 0.25, 0.25, 0.25) */
      estimate 'Pred y at centroid Sown vs Weed'
      Var*G1 0 0.25 -0.25 Var*G2 0 0.25 -0.25
      Var*L1 0 0.25 -0.25 Var*L2 0 0.25 -0.25
      Var*wfg1 0 0.0625 -0.0625
      Var*wfg2  0 0.0625 -0.0625
      Var*bfg 0 0.25 -0.25;
      estimate 'Pred y at centroid Sown vs N '
      Var*G1 -0.25 0.25 0 Var*G2 -0.25 0.25 0
      Var*L1 -0.25 0.25 0 Var*L2 -0.25 0.25 0
      Var*wfg1 -0.0625 0.0625 0
      Var*wfg2  -0.0625 0.0625 0
      Var*bfg -0.25 0.25 0;
      estimate 'Pred y at centroid Weed vs N '
      Var*G1 -0.25 0 0.25 Var*G2 -0.25 0 0.25
      Var*L1 -0.25 0 0.25 Var*L2 -0.25 0 0.25
      Var*wfg1 -0.0625 0 0.0625 Var*wfg2  -0.0625 0 0.0625
      Var*bfg -0.25 0 0.25;
run;
/*************************************************************************
/
```

R code

```
#####################################################################
####PACKAGES TO LOAD
library(nlme)
#####################################################################


#####################################################################
#### READING IN THE DATA
BelData<-read.table("C:/Appendix433_Beldata.txt", header=TRUE)
summary(BelData)
#####################################################################


#####################################################################
####FULL PAIRWISE INTERACTIONS MULTIVARIATE DIVERSITY-
####INTERACTIONS MODEL
####FITTED USING ML
FULL <- gls(Y ~ G1:Var+G2:Var+L1:Var+L2:Var+Density:Var+G1G2:Var+
                G1L1:Var+G1L2:Var+G2L1:Var+G2L2:Var+L1L2:Var-1,
                data=BelData,
                correlation  = corSymm(form = ~ -1| Plot),
                weights = varIdent(form = ~ -1 | VarNum), method="ML")
summary(FULL)
logLik(FULL)
-2*logLik(FULL)
#####################################################################


#####################################################################
####FITTING OF THE FUNCTIONAL GROUP MULTIVARIATE DIVERSITY-
####INTERACTIONS MODEL
####FITTED USING ML
FG <- gls(Y ~ G1:Var+G2:Var+L1:Var+L2:Var+Density:Var+
                Wfg1:Var+Wfg2:Var+Bfg:Var-1,
                data=BelData,
                correlation  = corSymm(form = ~ -1| Plot),
                weights = varIdent(form = ~ -1 | VarNum), method="ML")
summary(FG)
logLik(FG)
-2*logLik(FG)
#####################################################################

#####################################################################
```

```
####FUNCTIONAL GROUP INTERACTIONS MULTIVARIATE DIVERSITY-
####INTERACTIONS MODEL
####FITTED USING REML
FG_REML <- gls(Y ~ G1:Var+G2:Var+L1:Var+L2:Var+Density:Var+
                  Wfg1:Var+Wfg2:Var+Bfg:Var-1,
                  data=BelData,
                  correlation  = corSymm(form = ~ -1| Plot),
                  weights = varIdent(form = ~ -1 | VarNum), method="REML")
summary(FG_REML)
getVarCov(FG_REML)
#########################################################################
```

Appendix 4.3.5

Interpretation of selected output from Appendix 4.3.4

Here we provide some examples to illustrate the fitting procedure and aid the interpretation of the final model and associated tests of comparison.

Model selection using likelihood ratio tests

The comparison of the full pairwise interaction Multivariate Diversity-Interactions (FULL) model with the functional group Multivariate Diversity-Interactions (FG) model (Appendix 4.3.7, Model 4 versus Model 6) is provided as an example to illustrate the model fitting procedure used in selecting the final model.

The FULL model fitted using maximum likelihood gives: -2 Log Likelihood = 548.5
The FG model fitted using maximum likelihood gives: -2 Log Likelihood = 559.2
The likelihood ratio test statistic is constructed as: LRT = 559.2 - 548.5 = 10.7.

There were 3 x 6 = 18 interactions terms in the FULL model and 3 x 3 = 9 interactions terms in the FG model giving a difference of 9 df between the two models. Under the H_0 that the FG model is the correct model, the LRT comes from an approximate χ_9^2 distribution and the corresponding p-value is $P(\chi_9^2 > 10.7) = 0.297$. The null hypothesis is not rejected and it is concluded that there is no evidence that FULL model is required.

Interpretation of the final FG model fitted using restricted maximum likelihood (REML)

This is output from SAS.

Solution for Fixed Effects

Effect	Var	Estimate	Standard Error	DF	t Value	Pr > |t|
G1*Var	N	45.0164	4.5987	30	9.79	<.0001
G1*Var	Sown	66.4769	4.5035	30	14.76	<.0001
G1*Var	Weed	80.2891	8.4750	30	9.47	<.0001
G2*Var	N	29.0759	4.5987	30	6.32	<.0001
G2*Var	Sown	47.9474	4.5035	30	10.65	<.0001
G2*Var	Weed	91.5714	8.4750	30	10.80	<.0001
L1*Var	N	97.4344	4.5987	30	21.19	<.0001
L1*Var	Sown	77.2160	4.5035	30	17.15	<.0001
L1*Var	Weed	49.7543	8.4750	30	5.87	<.0001
L2*Var	N	76.2554	4.5987	30	16.58	<.0001
L2*Var	Sown	51.8816	4.5035	30	11.52	<.0001
L2*Var	Weed	33.6734	8.4750	30	3.97	0.0004
Density*Var	N	-0.6304	1.3390	30	-0.47	0.6412
Density*Var	Sown	1.1450	1.3113	30	0.87	0.3895
Density*Var	Weed	0.4956	2.4676	30	0.20	0.8422
Wfg1*Var	N	150.46	42.8224	30	3.51	0.0014
Wfg1*Var	Sown	105.37	41.9361	30	2.51	0.0176
Wfg1*Var	Weed	-31.9931	78.9184	30	-0.41	0.6881
Wfg2*Var	N	-5.3187	42.8224	30	-0.12	0.9020
Wfg2*Var	Sown	64.6390	41.9361	30	1.54	0.1337
Wfg2*Var	Weed	159.97	78.9184	30	2.03	0.0516
Bfg*Var	N	65.2422	19.2051	30	3.40	0.0019
Bfg*Var	Sown	87.2396	18.8076	30	4.64	<.0001
Bfg*Var	Weed	92.9492	35.3935	30	2.63	0.0135

In Figure 4.3.2 in the main text, the three bars for G1 stand at 66.5%, 80.3% and 45.0% for sown biomass, weed suppression and N yield respectively. These are the estimated coefficients for G1*Var highlighted by a box above and are the predicted performances of the ecosystem functions in a G1 monoculture. The first twelve rows of the above

output estimate monoculture performances for each species and ecosystem function and these values are graphed in the first twelve bars in Figure 4.3.2 in the main text. Shown also are the estimated density and functional group interaction coefficients for each ecosystem function. This output is displayed in Table 4.3.1 in the main text.

Interpretation of some predictions and tests of comparison among the estimated coefficients of the final FG model

This is output from SAS.

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|
G1 Sown vs Weed	-13.8122	7.3159	30	-1.89	0.0687
G1 Sown vs N	21.4605	2.7414	30	7.83	<.0001
G1 Weed vs N	35.2727	9.3681	30	3.77	0.0007
Pred y at centroid Sown	93.3161	1.9533	30	47.77	<.0001
Pred y at centroid Weed	95.0580	3.6759	30	25.86	<.0001
Pred y at centroid N	87.3275	1.9946	30	43.78	<.0001
Pred y at centroid Sown vs Weed	-1.7419	3.1732	30	-0.55	0.5871
Pred y at centroid Sown vs N	5.9887	1.1891	30	5.04	<.0001
Pred y at centroid Weed vs N	7.7305	4.0633	30	1.90	0.0667

The letters on top of the first cluster of three bars in Figure 4.3.2 (in the main text) are based on whether or not significant differences were identified among the estimated G1 monoculture performances for each ecosystem function and the associated tests for this are highlighted by a solid black box above. Significant differences in the relative performances in G1 monoculture ($p<0.017$ according to the the Bonferroni adjusted alpha level) were found between sown biomass and N yield and between weed suppression and N yield but not between sown yield and weed suppression. For

example, the predicted performance of sown yield in G1 monoculture is 66.5%, the predicted performance of weed suppression in G1 monoculture is 80.3% and the difference between the two performances is -13.8% but this difference is non-significant with p=0.0687.

The predicted performances shown in Figure 4.3.2 (in the main text) at the centroid community (all species equally present) were 93.3%, 95.1% and 87.3% for sown biomass, weed suppression and N yield respectively. These are highlighted by the dotted line box above while the tests of differences between each pair of these three predictions are highlighted by a dashed line box. For example, the predicted performance of sown yield in the centroid community is 93.3%, the predicted performance of weed suppression in the centroid community is 95.1% and the difference between the two performances is -1.7% but this difference is non-significant with p=0.5871. Further tests of comparison are shown in Appendix 4.3.8.

Appendix 4.3.6

Average monoculture and mixture performance with standard deviations for each ecosystem function for (a) the raw data and (b) the transformed data.

(a) Raw data means	Sown biomass (t DM ha^{-1})		Weed biomass (t DM ha^{-1})		N yield (t DM ha^{-1})	
	Mean	St dev	Mean	St dev	Mean	St dev
Monoculture (8 plots)	10.7	2.48	0.92	0.674	0.28	0.138
Mixture (22 plots)	15.8	1.31	0.33	0.187	0.39	0.042

(b) Transformed data means	Sown biomass (%)		Weed suppression (%)		N yield (%)	
	Mean	St dev	Mean	St dev	Mean	St dev
Monoculture (8 plots)	59.2	13.75	61.9	31.52	60.3	29.59
Mixture (22 plots)	87.6	7.29	89.6	8.77	82.9	8.93

Appendix 4.3.7

Model comparisons

#	(Description of diversity effects) Model terms	# DE co-efficients	-2LL	Comp	LRT	df	P-value
1	(No diversity effects) P_{G1} P_{G2} P_{L1} P_{L2} Dens	0	625.1				
2	(Average pairwise interactions) P_{G1} P_{G2} P_{L1} P_{L2} Dens $\Sigma P_i P_j$	3	577.5	1 vs 2	47.6	3	<.001
3	(Average pairwise interactions squared) P_{G1} P_{G2} P_{L1} P_{L2} Dens $\Sigma P_i P_j$ $\Sigma P_i P_j * \Sigma P_i P_j$	3	566.8	2 vs 3	10.7	3	0.013
4	(Grass - legume functional group interactions) P_{G1} P_{G2} P_{L1} P_{L2} Dens PP_{Wfg1} PP_{Wfg2} PP_{Bfg}	9	559.2	2 vs 4	18.3	6	0.006

5	(Fast establishing - temporally persistent functional group interactions) P_{G1} P_{G2} P_{L1} P_{L2} Dens PP_{Wfg1_F} PP_{Wfg2_P} PP_{Bfg_FP}	9	575.6	2 vs 5	1.9	6	0.929
6	(All pairwise interactions) P_{G1} P_{G2} P_{L1} P_{L2} Dens $P_{G1}P_{G2}$ $P_{G1}P_{L1}$ $P_{G1}P_{L2}$ $P_{G2}P_{L1}$ $P_{G2}P_{L2}$ $P_{L1}P_{L2}$	18	548.5	4 vs. 6	10.7	9	0.297

Footnote: DE= diversity effect, -2LL = -2 Log Likelihood, Comp = model comparison, LRT = Likelihood ratio test, df=degrees of freedom, P_{G1}, P_{G2}, P_{L1}, P_{L2} are the sown proportions of species G1, G2, L1 and L2 respectively, $\Sigma P_i P_j$ = sum over the product of each pair of sown proportions (*i.e.*, its coefficient is the expected average interaction), $PP_{Wfg1} = P_{G1}P_{G2}$, $PP_{Wfg2} = P_{L1}P_{L2}$, $PP_{Bfg} = P_{G1}P_{L1} + P_{G1}P_{L2} + P_{G2}P_{L1} + P_{G2}P_{L2}$, $PP_{Wfg1\_F} = P_{G1}P_{L1}$, $PP_{Wfg2\_P} = P_{G2}P_{L2}$, $PP_{Bfg\_FP} = P_{G1}P_{G2} + P_{G1}P_{L2} + P_{G2}P_{L1} + P_{L1}P_{L2}$. All model terms were crossed with ecosystem function and all models here were fitted using maximum likelihood. The finally selected model was model 4. The two grass species and the two legumes species in model 4 were additionally tested for functional redundancy (see Kirwan *et al.* 2009 for details of this test) but neither pair of species were found to functionally redundant ($P < 0.001$ in each test).

Appendix 4.3.8

Estimates, standard errors, t-values and p-values for each test illustrated in Fig. 4.3.2. Degrees of freedom are 30 in each test. The level of significance for each test is determined by the Bonferroni corrected $\alpha^*=0.017$.

	Sown biomass vs. Weed supp				Sown biomass vs. N yield				Weed suppression vs N yield			
Community	Est	Std Err	t-value	P-value	Est	Std Err	t-value	P-value	Est	Std Err	t-value	P-value
G1 mono	-13.8	7.32	-1.89	0.0687	21.5	2.74	7.83	<.0001	35.3	9.37	3.77	0.0007
G2 mono	-43.6	7.32	-5.96	<.0001	18.9	2.74	6.88	<.0001	62.5	9.37	6.67	<.0001
L1 mono	27.5	7.32	3.75	0.0007	-20.2	2.74	-7.37	<.0001	-47.7	9.37	-5.09	<.0001
L2 mono	18.2	7.32	2.49	0.0186	-24.4	2.74	-8.89	<.0001	-42.6	9.37	-4.55	<.0001
G1 dom	-1.7	4.23	-0.41	0.6877	13.5	1.58	8.53	<.0001	15.2	5.41	2.81	0.0085
G2 dom	-19.6	4.23	-4.64	<.0001	12.0	1.58	7.55	<.0001	31.6	5.41	5.83	<.0001
L1 dom	9.1	4.23	2.15	0.0397	-4.6	1.58	-2.90	0.0069	-13.7	5.41	-2.53	0.0170
L2 dom	3.5	4.23	0.84	0.4094	-7.1	1.58	-4.47	0.0001	-10.6	5.41	-1.96	0.0590

G1G2 dom	1.7	7.52	0.23	0.8222	8.7	2.82	3.08	0.0044	7.0	9.63	0.72	0.4744
L1L2 dom	-2.3	7.52	-0.30	0.7649	0.5	2.82	0.16	0.8719	2.7	9.63	0.28	0.7789
G1L1 dom	3.2	4.12	0.77	0.4471	6.4	1.54	4.17	0.0002	3.3	5.27	0.62	0.5400
G1L2 dom	0.4	4.12	0.10	0.9240	5.2	1.54	3.37	0.0021	4.8	5.27	0.91	0.3701
G2L1 dom	-5.8	4.12	-1.40	0.1713	5.7	1.54	3.67	0.0009	11.4	5.27	2.17	0.0382
G2L2 dom	-8.5	4.12	-2.08	0.0466	4.4	1.54	2.86	0.0076	13.0	5.27	2.46	0.0199
Centroid	-1.7	3.17	-0.55	0.5871	6.0	1.19	5.04	<.0001	7.7	4.06	1.90	0.0667