Development of grassland modelling techniques with weather forecasts.

By Jack McDonnell

Submitted for the qualification of Philosophiae Doctor Degree Statistics

Maynooth University

Department of Mathematics and Statistics

Submitted October 2018

Department Head Professor Stephen Buckley

Supervisors Dr. Caroline Brophy and Dr. Deirdre Hennessy

Maynooth University and Teagasc

Table of contents

Page

Appendices

Electronic Files (EFs)

EF1: sim_data_FG_monomix.csv

EF2: sim_data_R_SAS.csv

EF3: space_data.csv

EF4: Experimental_design.csv

Acknowledgements

I would like to thank my supervisors Caroline Brophy and Deirdre Hennessy who always gave excellent advice and helped me to thoroughly enjoy the experience. I really appreciated Caroline's positivity and outstanding supervision. I learned so much from her, both academically and personally. Thanks to Deirdre for her encouragement and valuable comments, and for teaching me lots about agriculture and scientific writing. I am grateful to everyone in the Maynooth University Mathematics and Statistics Department. I couldn't have done it without my fellow students and tutors.

Thanks to everyone in Teagasc Moorepark for a fantastic year. I learned a great deal from all of you and had great fun. To think that I was dreading moving to Fermoy! A special gratitude goes to Teagasc for providing the funding for the work.

I really appreciate my parents' support. I would like to thank them for chatting every day and being patient when I didn't want to tell them how work was going. Keith Lambkin taught me lots and always made me feel very welcome in Met Éireann. Rowan Fealy and Elodie Ruelle gave me invaluable help in learning all about climate and grass growth. Sincere thanks to Rafael de Andrade Moral for his insightful comments and coding expertise.

And finally, special thanks to Alice for her love and unwavering support, and for being there every step of the way.

Summary

This thesis investigates different aspects of weather, grass growth and statistical modelling. First, the accuracy of weather forecasts is assessed and improved using bias correction methods at twenty-five Irish locations for weather variables that influence grass growth. For the first time, soil temperature observations measured at six depths are verified and bias corrected. Next, the weather forecasts are included in an Irish grass growth model to investigate how the model accuracy is affected. The model predictions at an Irish farm are compared for weather observations and forecasts in multiple years, as well as to on-farm grass growth observations. These studies show that forecasts can be used in place of observations, and model predictions generally describe weekly grass growth accurately. Finally, grass growth modelling methods for experiments involving multiple species are developed for the analysis of a weed invasion study involving a large number of species. These developments include fitting novel random effects over multiple years to describe pairwise interactions between species parsimoniously and incorporating spatial planting pattern treatment into modelling methods.

Glossary

Biodiversity	variety of animal and plant species life
Centroid community	community in which all species are in equal proportion
Forb	a flowering plant that is not a grass
Forecast period	the number of days in advance of the observation that the forecast was produced
Functional group	group containing species that fulfil similar functional traits, for example nitrogen fixation
Legume	a family of plants known for their nitrogen-fixing traits
Met Éireann	Irish meteorological service
Mixture	community with more than one species
Monoculture	community with one species
Mean climatological forecasts	forecast the monthly mean climatological value for each day in the month
PastureBase Ireland	a national Irish database and management tool for grassland management
Persistence forecasts	forecasting the observation from the day before the forecast was generated
Richness	number of species
Stocking rate	animal units per unit of land area
Visual assessment method	a method of grass measurement which involves estimating grass covers visually

List of abbreviations

BC	Bias corrected
BEF	Biodiversity and ecosystem function
DE	Diversity effect
DI	Diversity-Interactions
DM	Dry matter
ECMWF	European Centre for Medium-Range Weather Forecasts
FG	Functional group
GGM	Grass growth model
MOS	Model output statistics
MoSt	Moorepark St Gilles
N	Nitrogen
NWP	Numerical weather prediction
SPaCE	Species Pattern and Community Ecology
UTC	Coordinated Universal Time

Chapter 1

Introduction

1.1 Thesis objectives

The overall aim of this thesis is to extend existing Irish grass growth modelling work by including weather forecasts, so grass growth can be predicted in advance, and therefore grass growth models (GGMs) can be used as farm management tools. The forecasts must first be assessed for accuracy, and then included in the grass growth model to investigate their influence. This work also aims to extend existing methods in the modelling of species-rich grassland experiments.

1.2 Background

Importance of grass growth in Ireland

Grazed grass is the cheapest feed source available to Irish farmers (Dillon et al., 2005). On dairy farms, every extra tonne of dry matter (DM) per hectare utilised can result in an increase of between €161 and €267/ha in farm profit (French et al., 2015; Shalloo, 2009). To maximise utilisation of this valuable resource, farmers must make grassland management decisions daily. For example, when there is a surplus of grass on-farm they should remove excess herbage to maintain feed quality in grazed swards. They must also decide when to apply nitrogen (N) fertiliser, and plan when to feed supplements if there will not be sufficient grass to meet demand. Current best practice in Irish grassland management involves using the visual assessment method, in which farmers visually estimate the amount of grass available to them on each paddock on their farm (grass cover) (Hanrahan et al., 2017; O'Donovan et al., 2002).

PastureBase Ireland is a national grass growth database and management tool . It provides farmers with information on the grass available to them on their farm based on farm cover estimation, stocking rate and weekly management (Hanrahan et al., 2017). Farmers can then make management decisions based on the information and resources available to them in PastureBase Ireland. Currently within PastureBase Ireland, a grass growth estimate for the current week is given based on the usual grass growth at that time of year. An operational grass growth model predicting growth in the coming week to ten days based on current farm conditions and weather forecasts would enhance farmers' abilities to make informed decisions and increase grass utilisation and farm profits (Ruelle et al., 2018).

History of grass growth modelling internationally and in Ireland

An operational grass growth model should give reasonable predictions for the short-term growth on-farm by accounting for local conditions such as soil type and weather conditions, and incorporating factors within the farmer's control such as grazing management and N fertiliser application. Site-specific grass growth models have been developed for European sites (Johnson and Thornley, 1983; Jouven et al., 2006), as well as some models that predict for multiple sites, and different land types (Schapendonk et al., 1998). All of these grass growth models are mechanistic, except Brereton et al. (1996), which is deterministic. Hurtado-Uria et al. (2013b) examined the accuracy of three grass growth models (Brereton et al., 1996; Johnson and Thornley, 1983; Jouven et al., 2006) for an Irish site and found that the Jouven et al. (2006) model performed best. Ruelle et al. (2018) developed the Moorepark St. Gilles (MoSt) grass growth model by equipping the Jouven model with soil water and soil nitrogen (N) sub-models to describe the movement of water and N through the soil. The MoSt GGM is a mechanistic model developed in C++, and is designed to be able to predict growth in a perennial ryegrass (*Lolium perenne* L.) sward on any Irish dairy farm.

Background on link between weather and grass growth

One of the main factors influencing grass growth is the weather conditions. The weather variables that most affect grass growth are rainfall, solar radiation, and air and soil temperatures. Ireland’s temperate climate gives the country’s farmers a competitive advantage because it provides favourable conditions for almost year-round growth (Hurtado-Uria et al., 2013a). The persistent, low-intensity rainfall and the generally mild temperatures are almost optimal grass growth conditions, since grass growth begins at 5°C and ceases between 20°C and 25°C (Frame, 1992; Hopkins, 2000). Rainfall during the Irish growing season is usually persistent and low-intensity, which favours grass growth, although a lack of water in summer months can sometimes reduce grass growth. Excess rainfall can also be problematic, particularly during spring and autumn, and can prevent grazing due to poor ground conditions and reduced growth (Burke et al., 2004). Soil temperatures at depths of less than 10 cm below the surface affect grass growth, with a particular influence on determining the length of the grass growth season (Hurtado-Uria et al., 2013a). Solar radiation is essential for photosynthesis to happen, and thus convert carbon dioxide into biomass (Laidlaw and Frame, 2013).

Weather forecasts for grass growth modelling

To be used in practice, a grass growth model requires weather forecasts. Previous grass growth models developed in Ireland have used retrospective weather observations and yielded reasonable estimates (Hurtado-Uria, 2013; Ruelle et al., 2018). The forecasts used by Met Éireann (Ireland's national meteorological service) up to ten days in advance are generated by the European Centre for Medium-Range Weather Forecasts (ECMWF). The variables relevant for grass growth are air temperature, rainfall, soil temperature and solar radiation. Observations for these weather variables are available at the 25 Met Éireann synoptic stations, which are distributed around Ireland. Because weather forecasts must describe large areas and be homogeneous, local biases can exist in forecasts (Auligne et al., 2007; Roberts, 2008). Before the inclusion of weather forecasts in a grass growth model, it is necessary to assess their accuracy, and attempt to remove any local systematic biases present in them. This can be done using bias-correction techniques such as those described in Joliffe and Stephenson (2011).

Grass growth modelling of mixed species swards

The GGM work in this thesis only examines predictions for perennial ryegrass systems, as the MoSt GGM currently only predicts growth for such systems. Mixed species swards have been shown to have environmental and economical benefits (Finn et al., 2013). Legumes, in particular white clover, are the most common species included with perennial ryegrass in Irish grasslands (Egan et al., 2018). In future, it will be important to adapt the MoSt GGM so that it can predict for such systems, which when managed well usually show increased dry matter production (Guy et al., 2018) and decreased weed invasion (Connolly et al., 2018), without affecting forage quality (Sturludottir et al., 2014). The inclusion of nitrogen-fixing legumes also reduces the amount of fertiliser needed. Although the inclusion of legumes is common when increasing species diversity in grasslands, herbs such as plantain and chicory are also being considered (Hofer et al., 2016). Diversity-Interactions (DI) models have been developed to be used for the analysis of multi-species experiments, including investigating the effects of multi-species swards in grasslands (Kirwan et al., 2009).

DI models and weed invasion

DI models use the proportions of each species in the plot as predictors, and give estimates of each species in monoculture (a single species), as well as estimating the contribution of the interaction between each pair of species (Kirwan et al., 2009). The DI model can be expressed as:

$$y_m = \sum_{i=1}^{s} \beta_i P_{im} + \sum_{\substack{i,j=1 \\ i<j}}^{s} \delta_{ij} P_{im} P_{jm} + \varepsilon_m \qquad (1)$$

where P_{im} is the proportion of species i in experimental unit m, s is the total number of species in the species pool, and the $\varepsilon_m \sim N(0, \sigma^2)$ are assumed independent and identically distributed for a single year, single site setting. β_i is the expected response for species i in monoculture, and δ_{ij} is the contribution of the interaction between species i and j. To simplify the model, the pairwise interactions can be grouped if they are similar (Kirwan et al., 2009), for example, it might be assumed that any pair of species from a group that have similar biological traits will interact in the same way. They can also be included as random pairwise interactions with the same variance to make a more parsimonious model, or if there is insufficient data to fit the model that includes all pairwise interactions as fixed effects (Brophy et al., 2017). DI models allow for prediction at proportions of species that were not sown provided the experimental design appropriately describes the simplex space (Connolly et al., 2013). Weed resistance is an important property of multi-species swards (Connolly et al., 2018), and weed biomass can be analysed as a response using DI modelling methods. The Species Pattern and Community Ecology (SPaCE) experiment investigated the effects of 16 grassland species on weed invasion and grass biomass in a natural system in North Dakota between 2012 and 2014. It incorporated a spatial sowing pattern treatment where species within plots were either aggregated or dispersed. However, DI models have not previously dealt with spatial pattern treatment, and the spatial pattern treatment in this case only applies to mixtures. Nor have DI models been developed that can model a large number of pairwise species interactions over multiple years.

1.3 Summary of each section / goals of the thesis

The aims of my research are to 1) to assess and improve the accuracy of weather forecasts in Ireland, 2) to extend the abilities of grass growth modelling in Ireland by including weather forecasts as inputs, and 3) to improve existing statistical methods for the analysis of multi-year grassland biodiversity experiments with large numbers of species.

In chapter 2, the accuracy of weather forecasts in Ireland is assessed. ECMWF forecasts up to ten days in advance for weather variables relevant to grass growth: maximum, minimum and mean two metre air temperature, rainfall and soil temperature at six depths, are compared to observations at the 25 Met Éireann synoptic weather stations. A variety of correction techniques are employed to attempt to remove any biases in ECMWF forecasts. ECMWF and bias-corrected forecasts are compared to easily generated low-skill forecasts: mean climatological forecasts, where the mean monthly observation is the forecast for every day in the month, and persistence forecasts, where the observation from the day before the forecast was generated is used.

It was found that all air and soil temperature forecasts were predicted accurately up to a week in advance, and were improved by bias corrections, particularly one using a regression model-based approach. Rainfall forecasts were shown to be accurate up to 5 or 6 days in advance. However, after this ECMWF rainfall forecasts generally gave higher RMSE values than equivalent mean climatology forecasts. Although climatology forecasts would not predict extreme rainfall events, it was found that ECMWF forecasts also did not predict these events accurately more than 6 days in advance.

In chapter 3, weather forecasts are included as an input for an Irish grass growth model. The Moorepark StGilles grass growth model (MoSt GGM) is the model used (Ruelle et al., 2018). It can be adjusted to account for local conditions and different farm management systems, but here is only presented at one site: Teagasc Curtin’s research farm. The weather inputs used in the model are rainfall, solar radiation and air temperatures. Two studies are conducted: the first compares predictions from the MoSt GGM from 2008 to 2016 at four different fertiliser levels at Curtin’s farm using observed weather, with corresponding predictions using weather forecasts. The second study replicates the farm management from 2013 to 2016 at Curtin’s in the MoSt GGM, and predictions using observed weather and various weather forecasts are compared to weekly on-farm grass growth observations.

The first study showed that forecasts could be interchanged with observations in the grass growth model and give similar grass growth predictions, and that the ECMWF forecasts, best bias-corrected forecasts, and mean climatological forecasts could all be important at varying forecast periods. There was no apparent interaction between fertiliser application levels and the weather input used, observed or forecast. The study comparing MoSt GGM predictions to on-farm grass growth observations showed agreement in most weeks during the growing season between the weekly grass growth observations and the predictions from the MoSt grass growth model for all weather inputs. However, there were consistent under-predictions by all weather inputs early and late in the growing season. Some poor weekly predictions showed the influence that weather inputs could have in the model, highlighting improvements that could be made.

The final section, which aims to improve existing statistical methods for the analysis of plot-based grassland experiments, uses the weed biomass yield in the three years of the SPaCE biodiversity experiment (2012 to 2014). The study investigated how functional group (groups of species with similar traits) and spatial sowing pattern (aggregated in species-specific clusters or randomly dispersed) influenced weed invasion. There were four species from each of four functional groups: warm-season and cool-season grasses, forbs and legumes. DI models are fitted to the weed biomass over three years. Random pairwise interactions with a common variance in each of the three years are included to help incorporate variation not described by the fixed effects, because there are many pairwise interactions. They do this in a parsimonious way, since each year only requires one extra model parameter. Complex covariance structures are included to account for the different variances for different functional groups, as well as including spatial pattern treatment as a fixed effect. A simulation study explores the ability of the modelling approach to detect when the random pairwise interactions are needed in the model under varying conditions.

Weed biomass was generally lower for mixtures after the year of establishment. The final model showed that each pair of species interacted depending on their functional group membership and the spatial planting pattern, as well as by the proportion of legumes present. Random pairwise interactions were needed in the model in year three of the study to describe additional variation not picked up by the fixed effects. A covariance structure with different blocks for monocultures from each functional group, and another for mixtures was used to satisfy the model assumptions. Model predictions suggested that species aggregation suppressed weed invasion better at establishment but had higher weed biomass in years 2 and 3. The simulation study showed that when the variance of the random pairwise interactions is sufficiently large, they are almost always detected, while when no random pairwise interaction variance is present, the model often does not fit.

1.4 Publications

McDonnell, J., K. Lambkin, R. Fealy, D. Hennessy, L. Shalloo and C. Brophy "Verification and bias correction of ECMWF forecasts for Irish weather stations to evaluate their potential usefulness in grass growth modelling", *Meteorological Applications, Volume 25, April 2018, Pages 292-301*.

(*This journal paper forms the basis for Chapter 2.*)

McDonnell, J., C. Brophy, E. Ruelle, L. Shalloo K. Lambkin and D. Hennessy "Weather forecasts to enhance an Irish grass growth model", *European Journal of Agronomy, Volume 105, April 2019, Pages 168-175.*

(*This journal paper forms the basis for Chapter 3.*)

Brophy, C., A. Dooley, L. Kirwan, J. A. Finn, **J. McDonnell**, T. Bell, M. W. Cadotte and J. Connolly "Biodiversity and ecosystem function: Making sense of numerous species interactions in multi-species communities", *Ecology, Volume 98, Issue 7, July 2017, Pages 1771-1778.*

(*This journal paper outlines new Diversity-Interactions modelling methods that are built upon in Chapter 4.*)

McDonnell, J., C. Brophy, E. Ruelle, K. Lambkin, R. Fealy, L. Shalloo and D. Hennessy (2018) "Evaluation of ECMWF weather forecasts and their inclusion in an Irish grass growth model", In Sustainable meat and milk production from grasslands. *Proceedings of the 27th General Meeting of the European Grassland Federation, Cork, Ireland, 17-21 June 2018 (pp. 823-825)*

(*This peer reviewed conference paper was presented at the European Grassland Federation meeting (EGF 2018). It is based on work from chapters 2 and 3.*)

The work from the thesis has been presented at several national and international conferences, including the Conference on Applied Statistics in Ireland (CASI 2015, 2018), Royal Meteorological Society Student Conference (RMetS 2016), Research Students Conference in Probability and Statistics (RSC 2016), European Meteorological Society meeting (EMS 2017), Ecological Society of America meeting (ESA 2017). After the presentation at the RMetS conference, I was invited to speak at the Scottish meeting of RMetS in 2017.

In all chapters, I was primarily responsible for data management, analysis and presentation of results. While I did most of the writing, all of the authors contributed to it in each chapter they are credited on. Caroline Brophy and Deirdre Hennessy offered advice on all aspects of the thesis. In Chapters 2, Keith Lambkin assisted with the data collection, and advised on meteorological and climatological topics along with Rowan Fealy. In chapter 3, Elodie Ruelle ran the simulations from the MoSt GGM and gave details about the model, while Keith Lambkin assisted with the data collection, and advised on meteorological and climatological topics. Laurence Shalloo assisted with the planning of the research in chapters 2 and 3. In Chapter 4, Thomas McKenna and Kathryn Yurkonis designed and conducted the experiment, collected the data, and advised us on ecological matters, while Rafael de Andrade Moral assisted with the statistical analysis, in particular the execution of DI models in R.

References

Auligne, T., McNally, A. and Dee, D., 2007. Adaptive bias correction for satellite data in a numerical weather prediction system. Quarterly Journal of the Royal Meteorological Society, 133(624): 631-642.

Brereton, A.J., Danielov, S.A. and Scott, D., 1996. Agrometeorology of Grass and Grasslands for Middle Latitudes. World Meteorological Organisation, Technical Note No. 197. Geneva.

Brophy, C. et al., 2017. Biodiversity and ecosystem function: making sense of numerous species interactions in multi-species communities. Ecology, 98(7): 1771-1778.

Burke, J.I., Brereton, A.J., O'Kiely, P. and Schulte, R., 2004. Weather and Crop Production. In Climate, Weather and Irish Agriculture, 2nd. Edition. AGMET, Weather and Crop Production. In Climate, Weather and Irish Agriculture, 2nd. Edition.

Connolly, J. et al., 2013. An improved model to predict the effects of changing biodiversity levels on ecosystem function. Journal of Ecology, 101(2): 344-355.

Connolly, J. et al., 2018. Weed suppression greatly increased by plant diversity in intensively managed grasslands: A continental-scale experiment. Journal of Applied Ecology, 55(2): 852-862.

Dillon, P., Roche, J., Shalloo, L. and Horan, B., 2005. Optimising financial return from grazing in temperate pastures. In: M. JJ (Editor), Utilisation of grazed grass in temperate animal systems. Wageningen Academic, Cork, Ireland.

Egan, M., Galvin, N. and Hennessy, D., 2018. Incorporating white clover (Trifolium repens L.) into perennial ryegrass (Lolium perenne L.) swards receiving varying levels of nitrogen fertilizer: Effects on milk and herbage production. Journal of Dairy Science, 101(4): 3412-3427.

Finn, J. 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(2): 365-375.

Frame, J., 1992. Improved Grassland Management. Farming Press Books, Wharfedale Road, Ipswich IP1 4 LG, United Kingdom.

French, P., O'Brien, B. and Shalloo, L., 2015. Development and adoption of new technologies to increase the efficiency and sustainability of pasture-based systems. Animal Production Science, 55(7): 931-935.

Guy, C., Hennessy, D., Gilliland, T., Coughlan, F. and McCarthy, B., 2018. Growth, morphology and biological nitrogen fixation potential of perennial ryegrass-white clover swards throughout the grazing season. Journal of Agricultural Science, 156(2): 188-199.

Hanrahan, L. et al., 2017. PastureBase Ireland: A grassland decision support system and national database. Computers and Electronics in Agriculture, 136: 193-201.

Hofer, D. et al., 2016. Yield of temperate forage grassland species is either largely resistant or resilient to experimental summer drought. Journal of Applied Ecology, 53(4): 1023-1034.

Hopkins, A., 2000. Grass : its production and utilization. Published for the British Grassland Society by Blackwell Science, Oxford ; Malden, MA, xiii, 440 p. pp.

Hurtado-Uria, C., 2013. Evaluation, adaptation and validation of a model to predict grass growth in Ireland, Cork Institute of Technology.

Hurtado-Uria, C., Hennessy, D., Shalloo, L., O'Connor, D. and Delaby, L., 2013a. Relationships between meteorological data and grass growth over time in the south of Ireland, Irish Geography, pp. 175-201.

Hurtado-Uria, C. et al., 2013b. Evaluation of three grass growth models to predict grass growth in Ireland. Journal of Agricultural Science, 151(1): 91-104.

Johnson, I. and Thornley, J., 1983. Vegetative crop growth-model incorporating leaf-area expansion and senescence, and applied to grass. Plant Cell and Environment, 6(9): 721-729.

Joliffe, I.T. and Stephenson, D.B., 2011. Forecast Verification: A Practitioner's Guide in Atmospheric Science, 2nd Edition. Wiley.

Jouven, M., Carrere, P. and Baumont, R., 2006. Model predicting dynamics of biomass, structure and digestibility of herbage in managed permanent pastures. 1. Model description. Grass and Forage Science, 61(2): 112-124.

Kirwan, L. et al., 2009. Diversity-interaction modeling: estimating contributions of species identities and interactions to ecosystem function. Ecology, 90(8): 2032-2038.

Laidlaw, A. and Frame, J., 2013. Improved grassland management. Crowood, 2013, Ipswich, U.K., 352 pp.

O'Donovan, M., Connolly, I., Dillon, P., Rath, M. and Stakelum, G., 2002. Visual assessment of herbage mass. Irish Journal of Agricultural and Food Research, 41(2): 201-211.

Roberts, N., 2008. Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. Meteorological Applications, 15(1): 163-169.

Ruelle, E., Delaby, L. and Hennessy, D., 2018. Development of the Moorepark St Gilles grass growth model (MoSt GG model): A predictive model for grass growth for pasture based systems. European Journal of Agronomy, 99(September 2018): 80-91.

Schapendonk, A., Stol, W., van Kraalingen, D. and Bouman, B., 1998. LINGRA, a sink/source model to simulate grassland productivity in Europe. European Journal of Agronomy, 9(2-3): 87-100.

Shalloo, L., 2009. Milk production costs - Can we compete? Pages 19-38, Teagasc National Dairy Conference, Cork, Ireland, http://www.teagasc.ie/media/website/publications/2000/national_dairy_proceedings_2009.pdf.

Sturludottir, E. et al., 2014. Benefits of mixing grasses and legumes for herbage yield and nutritive value in Northern Europe and Canada. Grass and Forage Science, 69(2): 229-240.

Chapter 2

Verification and bias correction of ECMWF forecasts for Irish weather stations to evaluate their potential usefulness in grass growth modelling

Collaborators on this paper were Keith Lambkin, Rowan Fealy, Deirdre Hennessy, Laurence Shalloo and Caroline Brophy. The paper is published in Meteorological Applications.

ABSTRACT: Typical weather in Ireland provides conditions favourable for sustaining grass growth throughout most of the year. This affords grass based farming a significant economic advantage due to the low input costs associated with grass production. To optimise the productivity of grass based systems, farmers must manage the resource over short time scales. While research has been conducted into developing predictive grass growth models for Ireland to support on-farm decision making, short-term weather forecasts have not yet been incorporated into these models. To assess their potential for use in predictive grass growth models, deterministic forecasts from the European Centre for Medium-Range Weather Forecasting (ECMWF) were verified for lead times up to ten days using observations from 25 Irish weather stations. Forecasts of air temperature variables were generally precise at all lead times, particularly up to seven days. Verification of ECMWF soil temperature forecasts is limited, but here they were shown to be accurate at all depths and most precise at greater depths such as 50 cm. Rainfall forecasts performed well up to approximately five days. Seven bias correction techniques were assessed to minimise systematic biases in the forecasts. Based on the root mean squared error values, no large improvement was identified for rainfall forecasts on equivalent ECMWF forecasts, but the optimum bias corrections improved air and soil temperature forecasts greatly. Overall, the results demonstrated that forecasts predict observations accurately up to approximately a week in advance and, therefore, could prove valuable in grass growth prediction at farm level in Ireland.

KEYWORDS: forecast verification, bias correction, Ireland, air temperature, rainfall, soil temperature, grass growth, agriculture.

2.1 Introduction

Agriculture is the largest indigenous industry in Ireland; specifically, the agri-food industry accounts for approximately €26 billion of total turnover (~ 7% GDP) and 8.4% of employment (DAFM 2015). At the primary agricultural production level, the grass based beef and dairy sectors account for almost 70% of production (DAFM 2015). In recognition of the 2015 abolition of EU quotas restricting milk production, the Irish Government established targets to increase the value of primary production by 65% and agri-food exports by 85% by 2025, compared to 2012-2014 levels (DAFM 2015). To meet these targets and the growing demands of the international community, farming practices in Ireland will need to optimise the utilisation of valuable natural resources such as grass. Shalloo (2009) and Dillon (2011) previously identified a strong positive linear association between grass utilisation (tonnes dry matter per hectare (t DM/ha)) and on-farm net profitability (€/ha) in Ireland. Due to the mild, maritime climate, with mean annual temperatures from 9-11 ^{o}C and the typically low intensity, long duration of rainfall throughout the year, optimum conditions for grass growth are achieved during most of the year (Hurtado-Uria et al. 2013a). The production of grass, with low input cost requirements, also minimises the need for costly alternative feed supplements. Consequently, primary agricultural production has largely developed around grass/pasture based systems, making it a key element of Irish agricultural productivity (Finneran et al. 2010), and profitability.

Many factors influence grass productivity, some of which are within the farmers control (e.g. stocking rate, fertiliser application), while others are outside (e.g. meteorological conditions, soil type). Frame (1992) has previously highlighted the influence of weather on grass production and utilisation. In particular, air and soil temperatures as well as rainfall, are critical factors determining both growing season length and rate of growth (Brereton 1995, Thorvaldsson et al. 2004). Perennial ryegrass (*Lolium perenne* L.) is the most widely sown grass species in Ireland (DAFM 2016), with leaf growth beginning at 5 °C, and reaching its peak between 20 °C and 25 °C (Hopkins 2000). Temperatures during spring and autumn typically vary between 5-10 ^{o}C, and therefore determine the growing season length (Burke et al. 2004). Soil temperatures between 0 cm and 10 cm below the surface are particularly important for grass growth in Ireland (Hurtado-Uria et al. 2013a). The year-round rainfall and the low permeability of many Irish soil types means that excessive moisture is usually more problematic for grass growth in Ireland than insufficient moisture availability (Burke et al. 2004). However, adequate rainfall during the growing season is essential for grass growth, with winter rainfall necessary to establish sufficient soil water levels in spring for growth to occur (Frame 1992).

To ensure sufficient supply and the optimised utilisation of grass throughout the growing season, grasslands are required to be managed on short timescales, optimally on a daily to weekly basis. The timing of management decisions, such as removing excess herbage at times of peak growth to maintain grass quality, directly affects farm efficiency and ultimately profitability (O'Donovan 2000, O'Donovan et al. 2011). While tools currently exist to assist on-farm decision making for grass budgeting/accounting (eg. grazing planners; PastureBase Ireland – a national grassland database), the use of an operational grass growth model, employing ‘local’ weather conditions, would assist farmers with management decisions influencing grass growth such as N fertiliser application, stocking rate and rotation length. Grass growth models using weather observations retrospectively have previously been developed (Hurtado-Uria et al. 2013b). The research outlined here seeks to build on these developments through the evaluation of outputs from a numerical weather prediction (NWP) model for use in forecasting meteorological parameters that influence grass growth, over time scales of relevance to improving on-farm decision making.

Over the last decade, NWP forecasts have become more skilful largely due to significant advances in computational resources, resolution and improved parameterisation schemes; but model skill varies depending on the meteorological parameter and the forecast lead time being evaluated. Despite these improvements, they also contain systematic biases (Ebert and McBride 2000, Sun et al. 2003, Auligne et al. 2007, Roberts 2008). These biases are most apparent at the surface-atmosphere boundary where errors in fluxes arising from the land surface and atmospheric model interact (Galanis and Anadranistakis 2002, Mass et al. 2008). Another bias arises due to the model grid representation, an integrated value representative of an area, which is typically evaluated against a proximal weather station.

The presence of systematic biases within NWP outputs can be minimised using appropriate post-processing methods (for example Harrison et al. 2000, Boi 2004, Sweeney et al. 2011, Vannitsem and Hagedorn 2011). A variety of approaches, including the moving window technique and Model Output Statistics (MOS), have been found to improve the accuracy of rainfall and 2 m air temperature forecasts (see Yussouf and Stensrud 2007, Huang et al. 2012 for example). Many studies attempt to correct forecasts spatially over a domain of interest (Louka et al. 2008, Vrac and Friederichs 2015 for example), and applying bias correction techniques at individual station locations can yield improvements in forecast accuracy (Taylor and Leslie 2005).

At long forecast lead times, rainfall is more difficult to forecast, and consequently bias correct, than air temperature. This is largely due to the fact that the processes that give rise to rainfall can occur over small space and time scales (Hamill et al. 2008, Fan and van den Dool 2011), requiring parameterisations rather than being resolved dynamically. Verification and bias correction of soil temperature forecasts remains limited internationally (but see Albergel et al. (2015)), in part due to a lack of suitable databases recording soil temperature observations. However, Met Éireann, the Irish National Meteorological Service, maintains a comprehensive soil temperature observation database for Ireland, with observations taken at six depths for 23 locations. This database is analysed in this paper, providing a detailed case study for soil forecast verification at multiple depths.

The purpose of this case study is to identify the accuracy of ECMWF weather forecasts in Ireland and to improve forecast accuracy where possible using bias correction for potential future inclusion in predictive grass growth models. This objective is achieved using data from a distributed network of 25 weather stations in Ireland over a period of seven years (2007 to 2013). The quality of ECMWF forecasts of rainfall, soil temperature and maximum, minimum and mean 2 m air temperature for lead times from 1 to 10 days is assessed. Various bias correction techniques are compared and the resulting forecasts are verified. Rainfall, air and soil temperature are highly influential in grass growth, but investigating and bias correcting their forecasts is an essential prior step to their inclusion in grass growth models.

2.2 Data and Methods

2.2.1 Data collation

Weather observations were collated for each of the 25 Met Éireann synoptic weather stations in Ireland for the period from 2007 to 2013 (Figure 2.1). Hourly observations of air temperature were obtained, while soil temperature (^{o}C) at six depths (5, 10, 20, 30, 50 and 100 cm) were available at 6 hourly intervals (0300, 0900, 1500 and 2100) at 5, 10, 20 cm and at once per day (0900) for soil depths at 30, 50 and 100 cm. Rainfall data (mm) were available daily. Corresponding forecasts for the seven years were obtained from the ECWMF operational forecasting model, for model grids matching the locations of the surface weather stations, for forecast lead times from day-1 to day-10. For example, the day-1 forecast for 10 January 2007 was run at 0000 on 10 January 2007, while the day-10 forecast was run at 0000 on 1 January 2007. Day-1 to day-10 are hereafter referred to as ‘forecast periods’. As eight of the weather stations became operational after 1 January 2007, observations were not available at all locations for the entire seven year period examined

(Appendix 2.1). There was also a small number of missing observations at some stations due to servicing, calibrations or instrument outages. Persistence and mean climatological forecasts were also obtained, with the exception that mean climatological observations were unavailable for soil temperatures.

Figure 2.1. Map of Ireland showing the locations of the 25 Met Éireann synoptic stations. The observations are taken manually at the manual stations which are located at airports.

2.2.1.1 Observations

The daily mean 2 m air temperature was the mean of the 24 hourly 2 m air temperature values recorded on each day starting at 0000 Coordinated Universal Time (UTC). If less than 13 hourly temperatures were available for a day, the daily value was excluded from analysis. The daily maximum and minimum 2 m air temperatures were the highest and lowest values recorded at the station in the 24 hours of the day beginning at 0000. Soil temperature observations were available at 23 of the 25 stations (measurements of soil temperature were unavailable from Finner and Macehead). At each station, soil temperature measurements were taken at 5, 10, 20, 30, 50 and 100 cm. The daily mean soil temperature at 5, 10 and 20 cm was derived from the mean of the observed temperature values for each depth at 0300, 0900, 1500 and 2100. Observations at 30, 50 and 100 cm were measured once daily at 0900. The total daily rainfall was the total precipitation (mm) recorded in the 24 hours beginning at 0000.

2.2.1.2 Forecasts

Daily forecast values for model grids corresponding to each weather station were obtained from the ECMWF Atmospheric Model high resolution 10-day 0000 forecast. This forecast had a horizontal grid resolution of approximately 25 km over Ireland until 26 January 2010, when it changed to approximately 16 km. The data extracted was mapped onto a latitude/longitude grid (0.125° x 0.125°) for ease of comparison of different model versions operating on different grid resolutions over the years. At some coastal stations, the grid box containing the station was describing a sea area in the model before the update, and a land area afterwards. It was found that this could lead to biases in the model. In order to overcome issues associated with the improvement of land-sea boundaries in the updated model, the forecast value used was that from the nearest land grid box output by the model. The ECMWF model predicts the minimum and maximum 2 m air temperatures in every consecutive 6-hour interval from +0 hours (0 hours after the model run) to +240 hours. The daily minimum and maximum temperatures were the lowest and highest, respectively, of the forecasts for the day in question. Each ECMWF model run gave 2 m air temperature and soil temperature forecasts at 3-hour intervals from +0 hours to +144 hours and at 6-hour intervals from +144 hours to +240 hours. The model forecast mean temperature was computed as the mean of the 2 m air temperature forecast values for the day. The day-1 to day-10 ECMWF forecasts for the four soil temperature forecast parameters (STL1 (0-7 cm), STL2 (7-28 cm), STL3 (28-100 cm) and STL4 (100-289 cm)) were obtained similarly. For example, the day-2 forecast for 2 January was the mean of the +24, 27, 30, 33, 36, 39, 42 and 45 hours forecasts from 1 January whereas the day-7 forecast for 7 January was the mean of the +150, 156, 162 and 168 hours forecasts from 1 January. Day-1 to day-10 rainfall forecasts were output directly by the ECMWF forecast model.

Monthly mean climatological values for rainfall and maximum, minimum and mean temperature between 1981 and 2010 were obtained at each station (Walsh 2012). The monthly mean climatological observation at each station was forecast for every day in the month to generate mean climatology forecasts. Persistence forecasts were generated by forecasting the observation of the day preceding the forecast generation (Joliffe and Stephenson 2011). For example, day-10 persistence forecasts were the observed weather conditions of 10 days ago.

2.2.2 Accuracy assessment of direct model output forecasts

A range of verification statistics were used to assess forecast accuracy and to identify biases for each weather variable. The statistics included Mean Systematic Bias (MSB

$= \frac{\sum_{i=1}^{n}(f_i - o_i)}{n}$), Root Mean Squared Error (RMSE = $\sqrt{\frac{\sum_{i=1}^{n}(f_i - o_i)^2}{n}}$), and Mean Absolute Error (MAE = $\frac{\sum_{i=1}^{n}|f_i - o_i|}{n}$), where o_i is the i^{th} observation, f_i is the i^{th} forecast and n is the total number of observations used in the calculation (Joliffe and Stephenson 2011). Standard deviation was used to assess the variability of the forecast and observed values. The forecasts were analysed at each individual forecast period from day-1 to day-10, or using all forecast periods combined if the statistic of interest did not vary substantially across forecast periods, to give a general description of forecast quality. The trends were examined by individual year, season, month and station, as well as across all seven years of data and twenty-five stations. When calculating yearly, monthly or seasonal statistics, limits were imposed on the number of missing values permitted: a minimum requirement of 183 values present in a year, 61 in a season and 22 in a month was set (yearly missing values at each station are shown in Appendix 2.1). To assess the skill of a forecast, its accuracy was compared with that of a mean climatology or a persistence forecast.

2.2.3 Bias correction

Seven bias correction (BC) techniques were tested: yearly, seasonal and monthly BC (1, 2 and 3, respectively), mean and variance (MAV) BC (4), regression model BC (5), regression model BC by station (6) and the composite (COM) post-processing method (7). Each BC method used a cross-validation approach; values from the period to be bias corrected were excluded from the data used to inform the bias corrections (Joliffe and Stephenson 2011). It was deemed reasonable to use data from before and after the target year (the year in which the forecasts were to be bias corrected) (Joliffe and Stephenson 2011). The details for each bias correction method are as follows:

1. *Year BC*: One year of forecasts and observations at one station was selected as the target year data. The MSB for all data at the station excluding the target year was calculated. This MSB was subtracted from ECMWF forecasts in the target year to obtain the yearly bias corrected forecasts (Joliffe and Stephenson 2011). This was repeated for each year at each station. Data from all forecast periods were used to do these bias corrections provided the MSB values did not vary much across forecast periods.
2. *Season BC*: Individually for each station, the MSB in the season of interest for all data outside the target year was subtracted from target year ECMWF forecasts for the particular season. The seasons were defined as spring (March, April, May), summer (June, July, August), autumn (September, October, November), and winter (December, January, February).
3. *Month BC*: Individually for each station, the MSB in the month of interest for all years outside the target year was subtracted from target year ECMWF forecasts for the particular month.
4. *Mean and variance (MAV) BC*: Separately for each forecast period, forecasts were bias corrected by scaling them to have the same mean and variance as the observations from the month being bias corrected (obs) using the cross-validation approach (Sweeney et al. 2011). The bias correction involved two steps:

   $$\mathrm{fBC1}_i = f_i * \frac{\sigma_{obs}}{\sigma_f}$$

   $$\mathrm{fBC2}_i = \mathrm{fBC1}_i + \mu_{obs} - \mu_{fBC1}$$

   Where f_i is the i^{th} forecast of interest, σ_{obs} and σ_f are the standard deviations of the observations and forecasts of interest, respectively,

$fBC1_i$ is the i^{th} forecast scaled for standard deviation, μ_{obs} and μ_{fBC1} are the means of the observations and forecasts (adjusted for standard deviation), respectively, and $fBC2_i$ is the i^{th} MAV bias corrected forecast.

5. *Model BC*: Using a linear regression approach (Acharya et al. 2013), separately for each forecast, forecasts were used as a predictor of observations:

   $$E[y] = \beta_0 + \beta_1 f + \delta_j + \alpha_k$$

   where y is the observation, f is the ECMWF forecast, δ_j is a categorical month specific term, $j = 1, \dots, 12$, α_k is a categorical station specific term, $k = 1, \dots, 25$. The coefficients of the model were estimated using the data outside the target year. The bias corrected forecasts for the target year were obtained by predicting from the estimated model.
6. *Model BC by station*: The model BC approach described in (5) was performed on each station separately with just the forecast and month as predictors.
7. *Composite (COM) post-processing*: Any subset (of size n) of the six bias-corrected forecasts and the ECMWF forecast can be combined to make a composite (COM) forecast (Sweeney and Lynch 2011). For each day, the cross-validation approach was employed separately on each station using data from the same month as that being bias corrected to obtain an MAE value (MAE_j) for each forecast (f_j) in the subset being used. The following equation was used to calculate the COM forecast for a particular day.

   $$\mathrm{COM} = \sum_{j=1}^{n} \frac{f_j}{\mathrm{MAE}_j \sum_{j=1}^{n} \frac{1}{\mathrm{MAE}_j}}$$

Any negative rainfall forecasts given by the bias correction methods were set to zero. All bias corrected forecasts were re-assessed for accuracy using the methods described in Section 2.2.2.

A moving window approach using cross validation, in which only data from the 31 days centred on the day to be bias corrected were used in the bias correction, was applied to methods 3, 4, 6 and 7 to assess if there was any improvement in forecast accuracy (Stensrud and Yussouf 2005). When using this approach, there was generally no overall improvement on the original bias correction methods. Since it was much more computationally expensive and therefore less useful for operational purposes, it has not been presented.

2.3 Results and discussion

2.3.1 Accuracy assessment of ECMWF forecasts

The accuracy of the ECMWF minimum, mean and maximum air temperature forecasts each decreased linearly as forecast period increased (Appendix 2.2); for example, the minimum temperature RMSE for data from all stations rose from 1.51 °C to 2.25 °C to 3.44 °C for forecast periods 1, 5 and 10, respectively. Comparing yearly RMSE values at each individual station and for each air temperature variable identified that persistence forecasts outperformed day-7 ECMWF forecasts over 75% of the time. This indicates that air temperature forecasts of forecast period longer than a week were not any more useful than low-skill forecasts. A number of stations showed considerable within year variability in maximum temperature RMSE (Figure 2.2).

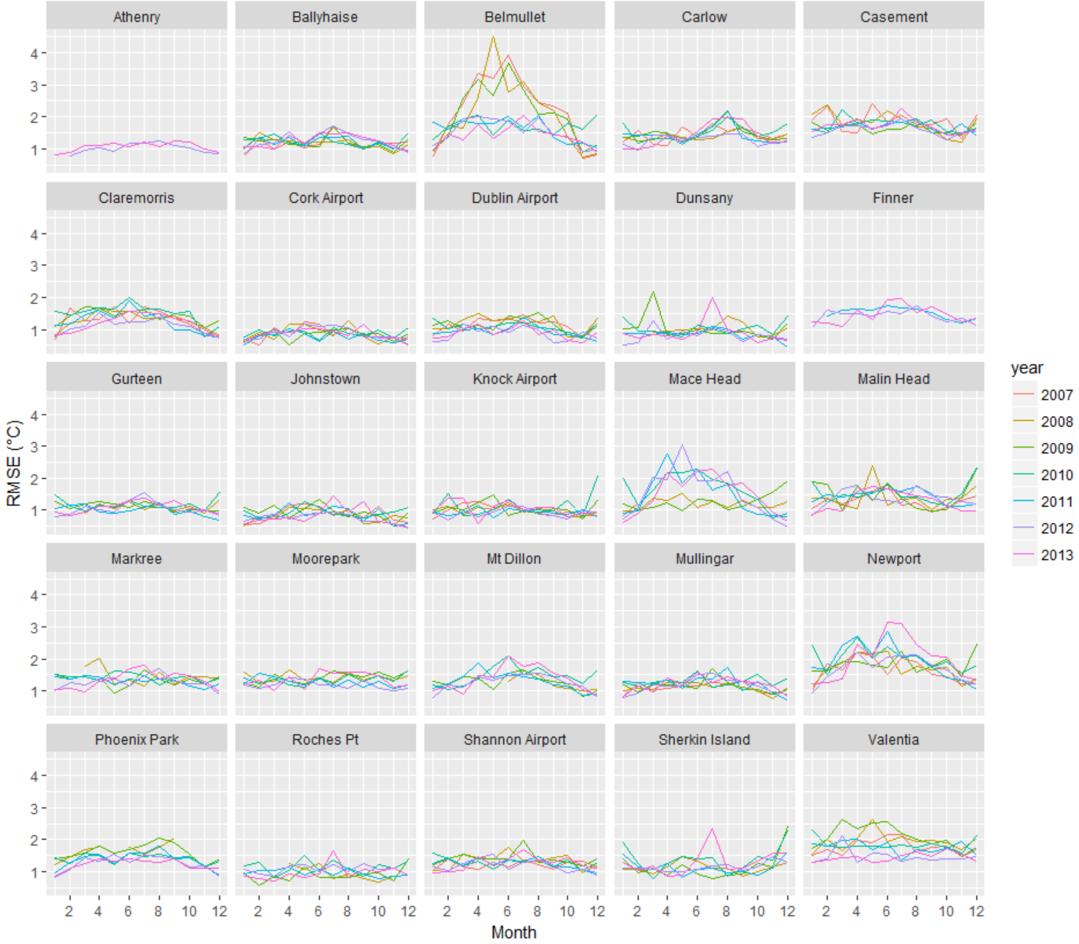


Figure 2.2. Monthly maximum temperature RMSE of day-1 ECMWF forecasts for each of the Met Éireann synoptic stations in each year from 2007 to 2013.

The highest yearly RMSE value for minimum temperature was recorded in 2010: 1.74 °C at forecast period 1 across all stations in comparison to a mean of 1.47 °C for the other six years of data, and this trend was evident in all forecast periods. In 2010, 20.1% of the daily minimum temperature observations were less than 0 °C compared to 7.0% on average across the other six years and these extremes were poorly predicted by the ECMWF model, explaining the poor forecasting performance in 2010. Poor forecasts of freezing temperatures are unlikely to hamper the prediction of grass growth provided the observations remain below 5 °C since this is the lower threshold for grass growth (Hopkins 2000). Minimum, maximum and mean 2 m air temperature MSB trends were generally constant across forecast period at each station indicating that the air temperature variables should respond positively to bias-correction (minimum temperature, Appendix 2.3). Appendix 2.3 shows yearly differences in MSB at Belmullet and Mace Head. This is likely due to the horizontal improvements in the model in 2010 better defining the land-sea boundary for these coastal stations. Otherwise, there were no noticeable differences in systematic biases before and after the ECMWF model update (Figure 2.3, Appendix 2.3, Appendix 2.4). Minimum temperature displayed a larger range of MSB values across stations than the other air temperature variables (between -1.86 °C and 1.00 °C for forecast period 1).

Soil temperature observations at each depth (5, 10, 20, 30, 50 and 100 cm) may not be most accurately forecasted by the forecast range into which they fall (STL1: 0-7, STL2: 7-28, STL3: 28-100 and STL4: 100-289 cm). When data from all stations and all forecast periods were included, the RMSE statistic identified that depths 5, 20, 50 and 100 cm were best forecast by their corresponding STL range but that 10 cm was best predicted by STL1 and 30 cm was best predicted by STL2; the most accurate range was used as the ECMWF forecast in each case. For all six soil temperature depths, the MSB was reasonably consistent across forecast periods at each station. However, there was a strong within year difference in MSB, with the forecast consistently under-estimating the observed in summer at all stations (shown for 5 cm, Figure 2.3). This could be attributed to the fact that the forecasts are not directly predicting the observation depths, rather they predict ranges, and therefore exhibit strong systematic biases. Bias correction should successfully reduce these biases (Joliffe and Stephenson 2011).

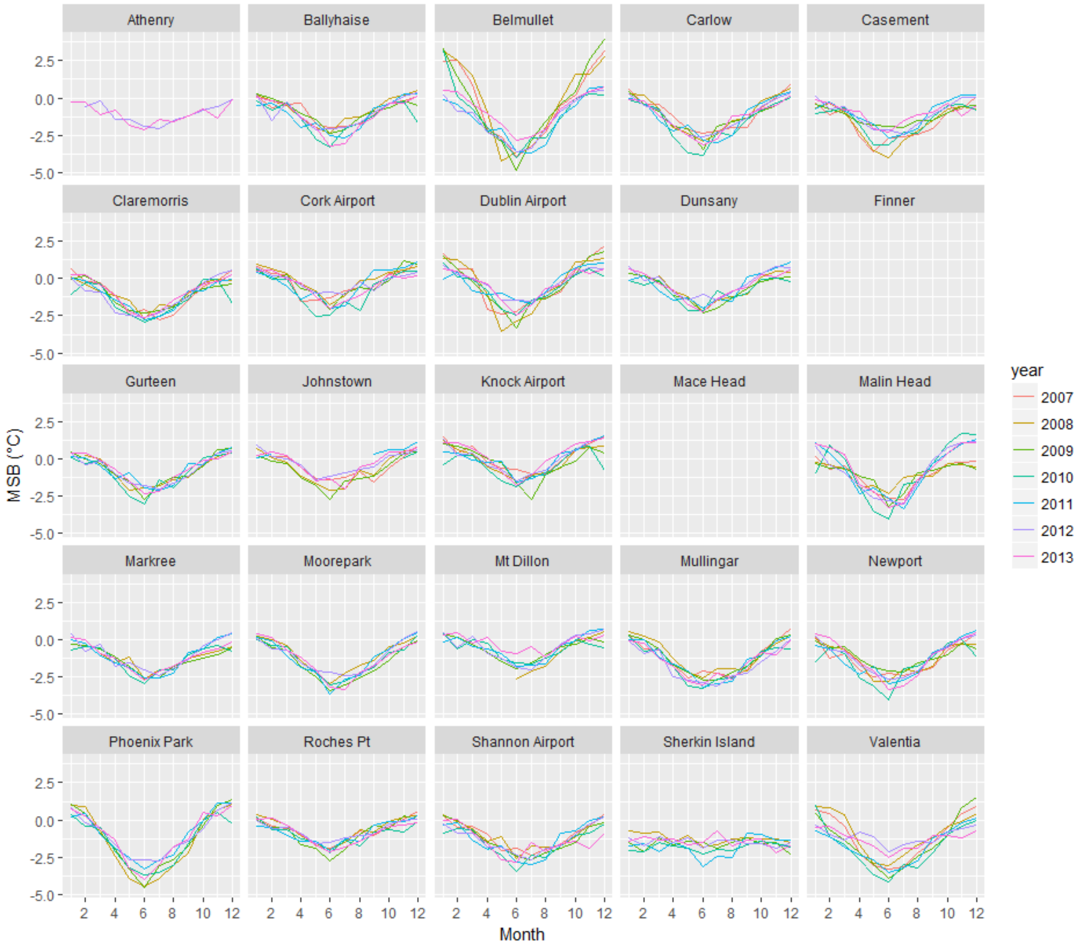


Figure 2.3. Monthly MSB of STL1 (0-7 cm) forecasts and 5cm soil temperature observations for day-1 forecasts in each year from 2007 to 2013 at 23 of the 25 Met Éireann synoptic stations. No soil observations were available at Finner or Mace Head.

Short-term ECMWF soil temperature forecasts performed well, with RMSE values below 1.71 °C at every depth at forecast period 1 when data from all stations were considered (Figure 2.4(a)). As expected, at all depths forecast accuracy declined as forecast period increased; for example, the RMSE for 10 cm soil temperatures across all stations was 1.62 °C at forecast period 1 and increased to

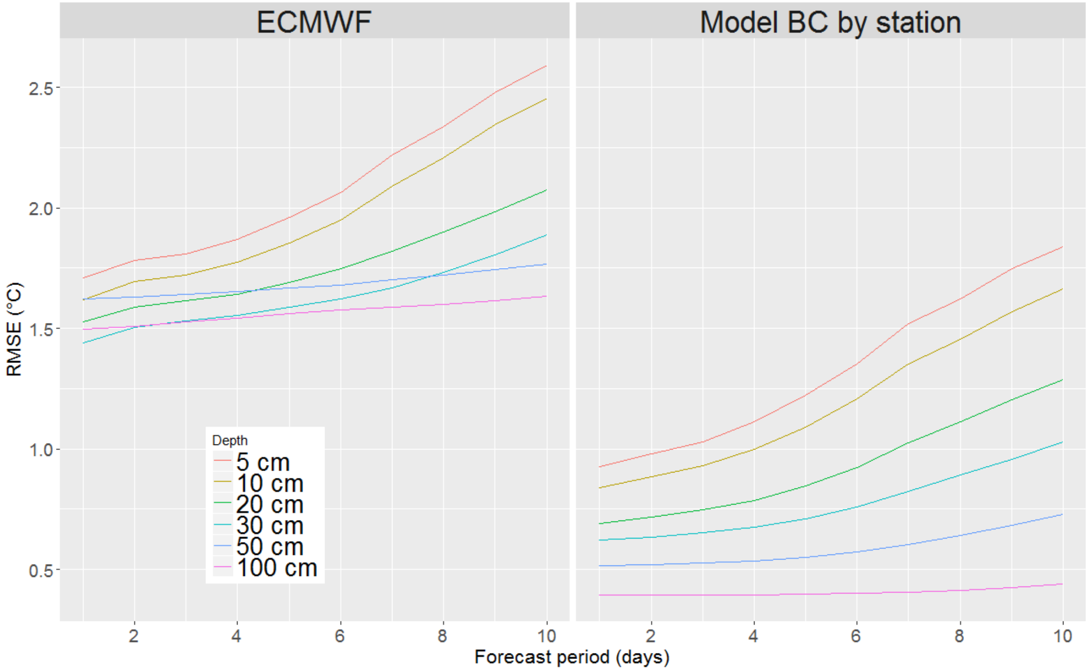


Figure 2.4. RMSE (°C) of (a) ECMWF and (b) model BC by station soil temperature forecasts and observations at six depths across all stations at forecast periods 1 to 10.

2.46 °C at forecast period 10 (Figure 2.4(a)). At greater depths, the forecasts were usually more accurate and the differences in RMSE values between forecast periods 1 and 10 were smaller (Figure 2.4(a)). This can be explained by the tendency of greater soil temperature depths to be less sensitive to changes in air temperature, making them more homogeneous over time, and therefore easier to predict. For example, at Dublin Airport between 1 June and 31 August 2012, the correlation between daily 5 cm soil temperature observations and corresponding 2 m mean air temperature observations was 0.83; the correlation between daily 100 cm soil temperature observations and corresponding 2 m mean air temperature observations was 0.52. There was a gradual overall warming effect at 100 cm, while there were persistent increases and decreases in 5 cm soil temperature over the same period

(Appendix 2.5). Since depths between 0 and 10 cm are most important for grass growth (Hurtado-Uria et al. 2013a), these results suggests that in the presence of accurate air temperature forecasts, soil temperature forecasts may not contribute appreciably further to the ability of a model to predict grass growth. However, ECMWF soil temperature forecasts have been shown for the first time to give accurate predictions of observations at a range of depths up to 100 cm; soil temperature forecasts have been verified only at 5 cm previously (Albergel et al. 2015). Although the soil temperature RMSE values were low, soil persistence forecasts generally outperformed ECMWF forecasts: 83.9% of yearly RMSE values for day-1 10 cm forecasts at individual stations were higher than equivalent persistence RMSE values. This rose to 100% for day-1 100 cm forecasts.

Day-5 climatology forecasts gave lower RMSE values than equivalent ECMWF forecasts in 64.9% of years at individual stations for daily rainfall, while day-7 climatology forecasts were better than ECMWF forecasts in all cases. Thus, after 7 days, ECMWF rainfall forecasts are no more useful in a grass growth model than mean climatology values. The yearly rainfall MSB values were often non-zero and were quite consistent within station across all forecast periods indicating that there were systematic biases in ECMWF rainfall forecasts that could possibly be eliminated by bias correction (Appendix 2.4). The MSB values in northwest Ireland at Belmullet, Finner and Mace Head were higher than the other stations: the ECMWF model over-predicted rainfall at these stations more than other stations. Perhaps due to their proximity to the Atlantic Ocean, high rainfall values were predicted here which were not reflected by the observations. It was also the case that the decrease in forecast accuracy of maximum temperature during the summer was most extreme at coastal stations such as Belmullet, Mace Head and Newport (Figure

2.2). The topography of the station could be different to the general area described by the ECMWF forecast. It could also be that the grid boxes used to forecast these stations are overly marine influenced. These biases should be reduced after bias correction.

As one might expect, high rainfall observations led to a decrease in forecast accuracy (Roberts and Lean 2008). For example, the MAE for day-5 forecasts from all stations with observations between 40 and 50 mm inclusive was 31.3 mm compared to an MAE of 2.2 mm for all observations between 0 and 10 mm inclusive. The MSB for the day-5 forecasts from all stations with observations between 40 and 50 mm inclusive was -31.3 mm, so these high rainfall events were all under-predicted at a lead time of 5 days. The MSB at Valentia Observatory in 2009 was lower than all other years at that station, particularly for day-5 forecasts and longer (Appendix 2.4). This was likely due to unusually high summer rainfall values causing the forecast to under-predict the observations; the total rainfall at Valentia in summer 2009 was 619.7 mm, compared to a mean total summer rainfall of 351.0 mm at the same station in the other 6 years. Inaccurate ECMWF forecasts of high rainfall events during the growing season can cause poor grazing management, resulting in reduced utilisation of available grass and subsequently impacting negatively on future grass growth. Due to poor utilisation, farms might be forced to supplement with concentrate, increasing variable costs. There are also risks of a decline in ground conditions and poor use of fertiliser (Shalloo et al. 2004).

2.3.2 Bias correction results

Model BC by station gave the greatest reductions in RMSE (compared to ECMWF forecasts) for almost all of the air and soil temperature variables for both forecast periods 1 and 10 (Table 2.1). It generally performed better than model BC since it

Table 2.1. The percentage reduction in RMSE resulting from the different bias correction methods compared to ECMWF forecasts across all stations for each of the rainfall and air and soil temperature forecasts at forecast periods 1 and 10. The COM BC included in this table is the composite of model BC and model BC by station since it performed best of all of the COM BC forecast combinations. The best bias correction technique, identified by the highest RMSE reduction, is highlighted by grey shading in each row.

	Year BC	Season BC	Month BC	MAV BC	Model BC	Model BC by station	COM BC
Min temp day-1	8.6	8.7	8.2	-42.7	9.5	12.7	12.3
Min temp day-10	1.6	1.6	1.5	-12.5	13.9	13.8	14.0
Max temp day- 1	26.8	27.5	27.4	-49.0	27.1	29.0	28.8
Max temp day-10	5.5	5.7	5.5	-11.7	15.3	15.6	15.7
Mean temp day-1	19.5	21.0	21.0	-102.4	21.0	24.8	24.5
Mean temp day-10	2.4	2.7	2.7	-14.8	13.0	13.2	13.3
5 cm day- 1	18.7	38.0	43.2	-2.6	42.1	45.9	45.6
5 cm day-10	11.2	16.7	17.7	-0.8	28.0	29.0	28.9
10 cm day-1	22.3	38.1	42.0	-4.2	43.7	48.1	47.8
10 cm day-10	13.0	17.3	17.9	1.3	31.1	32.3	32.2
20 cm day-1	26.2	43.8	48.4	1.1	48.3	54.6	54.2
20 cm day-10	19.4	25.4	26.4	3.7	35.8	38.0	37.8
30 cm day-1	27.5	34.2	35.0	0.9	48.1	56.8	55.9
30 cm day-10	21.8	24.6	24.9	8.2	42.1	45.6	45.2
50 cm day-1	33.5	56.8	64.0	22.5	57.4	68.1	67.2
50 cm day-10	34.4	51.1	55.6	20.8	51.5	58.7	58.0
100 cm day-1	47.0	54.4	55.5	30.6	59.4	73.7	72.4
100 cm day-10	47.3	54.6	56.0	35.2	61.0	73.0	71.8
Rainfall day-1	3.1	3.3	3.1	-12.7	7.8	9.0	8.9
Rainfall day-10	1.2	1.3	1.3	-1.5	23.6	23.5	23.6

does not assume common month coefficients for each station and was therefore more effective at eliminating monthly biases (such as those in Figure 2.3). Although COM BC gave similar and sometimes slightly higher reductions, model BC by station is recommended as it requires less computation time. Model BC, model BC by station and COM BC reduced the bias at every station to magnitudes of 0.05 °C or less for all air temperature variables at all forecast periods. Although MAV BC reduced bias in forecasts, it did not improve the RMSE in general and was generally the worst bias correction method (Table 2.1). The relatively high percentage reductions in RMSE for predictions of soil temperatures could be because there were often large systematic biases attributable to the fact that the forecasts predicted ranges rather than directly forecasting the observations. Model BC by station was also generally the most effective method at reducing the RMSE (Table 2.1) and eliminating bias in rainfall forecasts: across all forecast periods the MSB values at individual stations ranged from -0.00 mm to 0.06 mm. This was compared to methods such as month BC which gave forecasts that over-predicted the daily rainfall on average by 0.82 mm at Mace Head.

2.3.3 Accuracy assessment of best bias corrected forecasts

After model bias correction by station the day-1 MSB values were 0.002 °C, 0.000 °C and 0.001 °C for maximum, minimum and mean temperature, respectively, across all stations. In comparison, the ECMWF day-1 MSB values were -0.770 °C, 0.039 °C and -0.304 °C, respectively. Model BC by station reduced the range in MSB for day-1 maximum temperature ECMWF forecasts across stations from between -1.57 °C and 0.43 °C to between -0.004 °C and 0.006 °C. The RMSE values for the maximum temperature ECMWF forecast at forecast period 1 ranged from

0.86 °C to 1.96 °C at individual stations, with a station average of 1.32 °C when included in a mixed model with weather station included as a random variable to generalise the location. These values were reduced to range between 0.82 °C and 1.19 °C by model BC by station, with an average of 0.98 °C and forecast improvements at almost all stations. Day-1 air temperature forecasts experienced decreases in RMSE of up to 49% at individual stations. Yearly RMSE values for day-7 persistence forecast were preferable to day-7 model BC by station forecasts for maximum, minimum and mean temperature 47.1%, 48.4% and 58.0% of the time, respectively, a substantial improvement on the ECMWF forecasts. Thus, model BC by station improved the ECMWF air temperature forecasts, meaning they should be of practical use in a grass growth model up to approximately a week in advance. Model BC by station did not generally improve the accuracy of the imprecise ECMWF air temperature forecasts of the extreme low temperature observations in 2010. Regression model approaches are not usually good at predicting extremes (Allen and DeGaetano 2001, Zhai et al. 2005), and the forecasts used as predictors in the model were not accurate to begin with. However, as noted in Section 3.1, this may not be too important provided forecasts for extremely low observations are below 5 °C.

Model BC by station soil temperature forecasts had RMSE values below 0.93 °C for day-1 forecasts and below 1.84 °C for day-10 forecasts at all depths (Figure 2.4(b)). The forecast accuracy was generally worst at 5cm, improving at each subsequent depth (Figure 2.4(b)). Due to their homogeneity, the systematic biases were more constant at greater depths, meaning bias correction usually gave larger improvements at these depths. While ECMWF 30 cm soil temperature forecasts had lower RMSE values than all other depths at some forecast periods (Figure 2.4(a)), its systematic bias was not as large as other depths. As a result, it did not have the lowest RMSE values after model BC by station. The reductions in RMSE from the ECMWF forecasts were similar across all forecast periods but tended to be higher in summer when the MSB values were of higher magnitude (Figure 2.3). Model BC by station soil temperature forecasts for depths such as 5 and 10 cm were generally more useful than persistence at forecast periods of less than a week, while forecasts at depths such as 50 and 100 cm were not. For example, day-5 model BC by station forecasts for 10 cm gave lower RMSE values than persistence in 70.5% of years at individual stations (a large improvement on ECMWF forecasts), but day-1 100 cm persistence forecasts outperformed equivalent model BC by station forecasts in 99.3% of cases. The accuracy of forecasts at the 5 cm and 10 cm depths is useful since the soil temperature at shallow depths will be the most influential in determining grass growth.

Model BC, model BC by station and the COM BC approaches gave the highest RMSE reductions for rainfall at both forecast periods 1 and 10 (Table 2.1); and although day-7 model BC by station forecasts out-performed climatology forecasts in 81.6% of cases (a large improvement on ECMWF forecasts), they may not be useful in practice. Because rainfall is difficult to forecast and predictive rainfall accuracy diminishes rapidly with forecast period, these BC methods tend to forecast conservative estimates close to the mean rainfall. Although these methods usually resulted in lower RMSE values than the original forecast, they did not follow the trend of the observations as closely as other forecasts (see Ballyhaise for example, Figure 2.5). At Ballyhaise in January 2012, the RMSE values of the day-10 ECMWF, month BC and model BC by station forecasts were 5.56 mm, 5.60 mm and 4.40 mm, respectively. The standard deviation of the observations was 4.55 mm, and

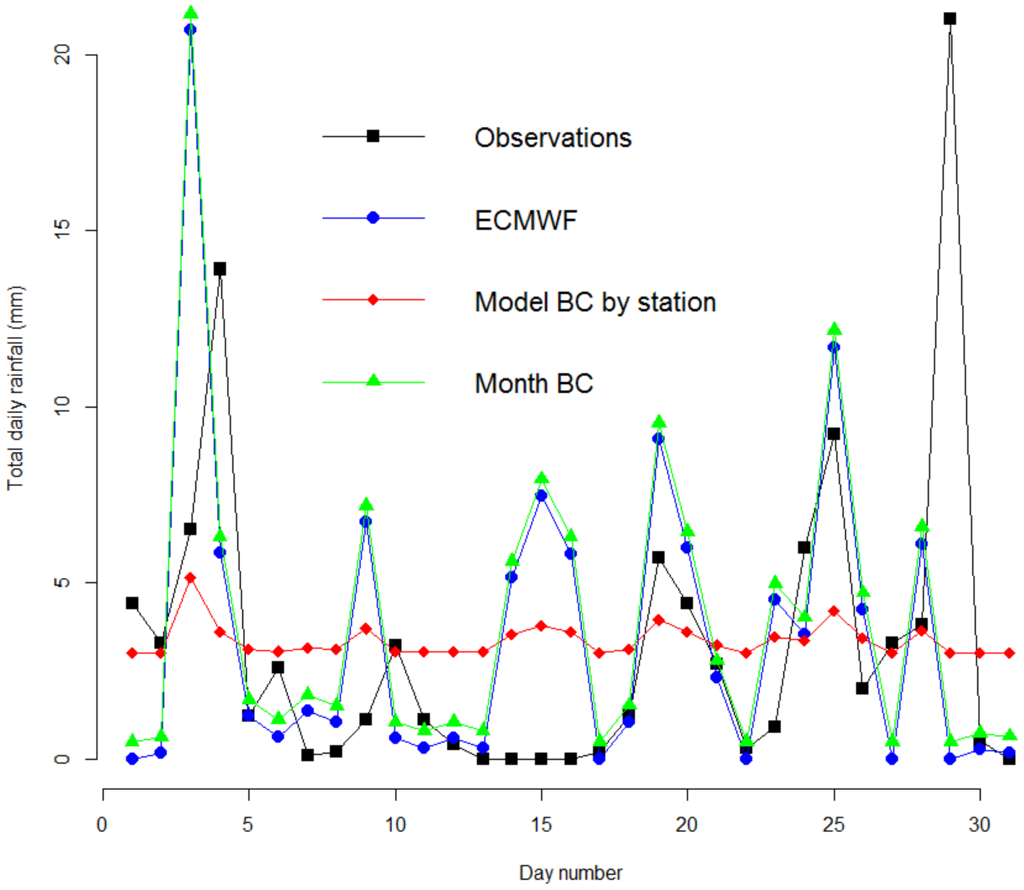


Figure 2.5. Daily observed rainfall (squares), day-10 ECMWF forecast (circles, RMSE = 5.56 mm), day-10 month bias corrected forecast (triangles, RMSE = 5.60 mm), day-10 model bias corrected by station forecast (diamonds, RMSE = 4.40 mm) at Ballyhaise in January 2012.

was 4.51 mm for both the ECMWF forecast and the monthly bias corrected forecasts. This fell to 0.47 mm for model BC by station forecast. Thus, bias correction approaches such as monthly BC were preferable even though the reductions in RMSE were minimal (Table 2.1). Although most BC methods gave reductions in RMSE when using all observations, the reductions usually only occurred for low rainfall observations. As highlighted in Section 3.1, high rainfall values were predicted poorly by ECMWF forecasts. When daily observations greater than 10 mm were considered, none of the methods gave substantial reductions in RMSE at any forecast period. A possible reason for this is that if the initial forecasts are generally imprecise, bias corrections cannot improve the forecasts. The inaccuracy of rainfall forecasts and their bias corrections is problematic since rainfall values greater than 10 mm are likely to be more influential on grass growth than lower rainfall values. Model BC by station air and soil temperature forecasts did not have the same problem with standard deviation as the rainfall forecasts, often giving standard deviation values closer to those of the observations than ECMWF forecasts.

2.3.4 Future work and implications for Irish agriculture

Future work will involve developing a grass growth model to simulate on-farm growth in Ireland that will incorporate weather forecast predictors. To date, grass growth prediction models have tested a wide range of predictors (for example soil type, fertiliser application, grazing events) but have only considered historical weather data (Hurtado-Uria et al. 2013b). The verification of the ECMWF forecasts suggest that they will accurately predict weather conditions in the coming five to seven days, and therefore will be valuable as predictors of grass growth. The best bias corrected forecasts were usually more accurate than the ECMWF forecasts. If they prove to give grass growth predictions preferable to the ECMWF forecasts, methods of applying bias corrections across the island at farm level could be useful. If the rainfall component of the grass growth model is not sufficient, the use of probabilistic rainfall forecasts could be explored (Hamill et al. 2004, Wilks 2009). Previous studies have shown that soil temperature forecasts at various European locations give accurate predictions at 5cm (Albergel et al. 2015). This study verified the accuracy of soil temperature forecasts at 23 Irish weather stations at six depths. The fact that soil temperature forecasts are useful will not only benefit grassland management, it also has positive implications for Irish crop production. Climatology and persistence forecasts both regularly out-performing ECMWF and bias-corrected forecasts after seven days suggests that a combination of these forecasts with recent observed weather and long term climate would work best when forecasting weather for on-farm grass growth prediction. Online resources allowing farmers to input local parameters and obtain grass growth predictions accounting for weather forecasts would increase grass utilisation and on-farm profitability.

2.4 Conclusions

Forecast verification was undertaken to determine how accurate ECMWF operational deterministic forecasts were at the Irish synoptic weather stations. Air temperature forecasts were accurate for all forecast periods but often gave higher RMSE values than persistence after 7 days. Previous studies on the bias correction of ECMWF 2 m air temperature forecasts illustrated the decrease on forecast quality as the forecast period increased for both ECMWF and bias corrected forecasts (Vannitsem and Hagedorn 2011). It has also been shown that bias correction methods work well up to five days in advance, but is difficult at longer lead times (Boi 2004). ECMWF soil temperature forecasts have not been extensively verified internationally but here we show for the first time in Ireland that they can predict observations well at a variety of depths, with prediction accuracy increasing as depth beneath the surface increases. This was due to the more conservative rate of change in temperatures at greater depths making them easier to model. However, this slow rate of change also meant that persistence forecasts gave lower RMSE values than

ECMWF forecasts in almost all years and at all forecast periods. ECMWF rainfall forecasts showed skill for forecast periods of six days or less. After this climatology forecasts tended to give lower RMSE values. Previous studies on rainfall forecasts in the United Kingdom (Harrison et al. 2000), and monsoon season rainfall forecasts in India (Acharya et al. 2013) have shown that bias correction can yield improvements in forecast accuracy. They also discuss the fact that rainfall forecasts often predict extreme rainfall events poorly.

Systematic biases were evident for all of the weather variables examined; their values varied in magnitude and orientation across locations. Thus, seven approaches to eliminate them were proposed. For air and soil temperatures, model BC by station was generally the most effective method of bias correction. Soil temperature forecasts were greatly improved by bias correction since the ECMWF forecasts exhibited strong systematic biases at many depths. Month BC was recommended for ECMWF rainfall forecasts, although none of the bias correction methods assessed gave large improvements because they did not improve forecast quality for observations greater than 10 mm, and some methods gave forecasts with much lower standard deviations than the observations.

Overall, we have shown that weather forecasts have the potential to contribute to grass growth prediction for up to one week in Ireland. This knowledge can contribute to better efficiency in on-farm management of grass resources and help improve primary productivity and profit for Irish agriculture.

References

Acharya, N., S. Chattopadhyay, U. Mohanty, S. Dash, and L. Sahoo. 2013. On the bias correction of general circulation model output for Indian summer monsoon. Meteorological Applications **20**:349-356.

Albergel, C., E. Dutra, J. Munoz-Sabater, T. Haiden, G. Balsamo, A. Beljaars, L. Isaksen, P. de Rosnay, I. Sandu, and N. Wedi. 2015. Soil temperature at ECMWF: An assessment using ground-based observations. Journal of Geophysical Research-Atmospheres **120**:1361-1373.

Allen, R., and A. DeGaetano. 2001. Estimating missing daily temperature extremes using an optimized regression approach. International Journal of Climatology **21**:1305-1319.

Auligne, T., A. McNally, and D. Dee. 2007. Adaptive bias correction for satellite data in a numerical weather prediction system. Quarterly Journal of the Royal Meteorological Society **133**:631-642.

Boi, P. 2004. A statistical method for forecasting extreme daily temperatures using ECMWF 2-m temperatures and ground station measurements. Meteorological Applications **11**:245-251.

Brereton, A. 1995. Regional and year-to-year variation in production. Pages 12-22. Royal Irish Academy, Irish grasslands - their biology and management.

Burke, J. I., A. J. Brereton, P. O'Kiely, and R. Schulte. 2004. Weather and Crop Production. In Climate, Weather and Irish Agriculture, 2nd. Edition. AGMET, Weather and Crop Production. In Climate, Weather and Irish Agriculture, 2nd. Edition.

DAFM. 2015. Food Wise 2025. Department of Agriculture Food and the Marine, http://www.agriculture.gov.ie/media/migration/agri-foodindustry/foodwise2025/report/FoodWise2025.pdf.

DAFM. 2016. Grass and Clover, Recommended List Varieties for Ireland 2016. Department of Agriculture, Food and the Marine, http://www.agriculture.gov.ie/media/migration/publications/2016/GrassWhiteCloverRecom2016080216.pdf.

Dillon, P. 2011. The Irish dairy industry-Planning for 2020. Pages 1-24, National Dairy Conference 2011, Cork, Ireland, http://www.teagasc.ie/publications/2011/1054/Dairy_Conference_Proceedings_2011.pdf.

Ebert, E., and J. McBride. 2000. Verification of precipitation in weather systems: determination of systematic errors. Journal of Hydrology **239**:179-202.

Fan, Y., and H. van den Dool. 2011. Bias Correction and Forecast Skill of NCEP GFS Ensemble Week-1 and Week-2 Precipitation, 2-m Surface Air Temperature, and Soil Moisture Forecasts. Weather and Forecasting **26**:355-370.

Finneran, E., P. Crosson, P. O'Kiely, L. Shalloo, P. D. Forristal, and M. Wallace. 2010. Simulation modelling of the cost of producing and utilizing feeds for ruminants on Irish farms. Journal of Farm Management **14**:95-116.

Frame, J. 1992. Improved Grassland Management. Farming Press Books, Wharfedale Road, Ipswich IP1 4 LG, United Kingdom.

Galanis, G., and M. Anadranistakis. 2002. A one-dimensional Kalman filter for the correction of near surface temperature forecasts. Meteorological Applications **9**:437-441.

Hamill, T., R. Hagedorn, and J. Whitaker. 2008. Probabilistic forecast calibration using ECMWF and GFS ensemble reforecasts. Part II: Precipitation. Monthly Weather Review **136**:2620-2632.

Hamill, T., J. Whitaker, and X. Wei. 2004. Ensemble reforecasting: Improving medium-range forecast skill using retrospective forecasts. Monthly Weather Review **132**:1434-1447.

Harrison, D., S. Driscoll, and M. Kitchen. 2000. Improving precipitation estimates from weather radar using quality control and correction techniques. Meteorological Applications **7**:135-144.

Hopkins, A. 2000. Grass : its production and utilization. 3rd edition. Published for the British Grassland Society by Blackwell Science, Oxford ; Malden, MA.

Huang, L., G. Isaac, and G. Sheng. 2012. Integrating NWP Forecasts and Observation Data to Improve Nowcasting Accuracy. Weather and Forecasting **27**:938-953.

Hurtado-Uria, C., D. Hennessy, L. Shalloo, D. O'Connor, and L. Delaby. 2013a. Relationships between meteorological data and grass growth over time in the south of Ireland. Pages 175-201, Irish Geography.

Hurtado-Uria, C., D. Hennessy, L. Shalloo, R. Schulte, L. Delaby, and D. O'Connor. 2013b. Evaluation of three grass growth models to predict grass growth in Ireland. Journal of Agricultural Science **151**:91-104.

Joliffe, I. T., and D. B. Stephenson. 2011. Forecast Verification: A Practitioner's Guide in Atmospheric Science, 2nd Edition. Wiley.

Louka, P., G. Galanis, N. Siebert, G. Kariniotaki, P. Katsafados, I. Pytharoulis, and G. Kallos. 2008. Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering. Journal of Wind Engineering and Industrial Aerodynamics **96**:2348-2362.

Mass, C., J. Baars, G. Wedam, E. Grimit, and R. Steed. 2008. Removal of systematic model bias on a model grid. Weather and Forecasting **23**:438-459.

O'Donovan, M. 2000. The relationship between the performance of dairy cows and grassland management on intensive dairy farms in Ireland. University College Dublin.

O'Donovan, M., E. Lewis, and P. O'Kiely. 2011. Requirements of future grass-based ruminant production systems in Ireland. Irish Journal of Agricultural and Food Research **50**:1-21.

Roberts, N. 2008. Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. Meteorological Applications **15**:163-169.

Roberts, N., and H. Lean. 2008. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. Monthly Weather Review **136**:78-97.

Shalloo, L. 2009. Milk production costs - Can we compete? Pages 19-38, Teagasc National Dairy Conference, Cork, Ireland, http://www.teagasc.ie/media/website/publications/2000/national_dairy_proceedings_2009.pdf.

Shalloo, L., P. Dillon, J. O'Loughlin, M. Rath, and M. Wallace. 2004. Comparison of a pasture-based system of milk production on a high rainfall, heavy-clay soil with that on a lower rainfall, free-draining soil. Grass and Forage Science **59**:157-168.

Stensrud, D., and N. Yussouf. 2005. Bias-corrected short-range ensemble forecasts of near surface variables. Meteorological Applications **12**:217-230.

Sun, B., L. Yu, and R. Weller. 2003. Comparisons of surface meteorology and turbulent heat fluxes over the Atlantic: NWP model analyses versus Moored buoy observations. Journal of Climate **16**:679-695.

Sweeney, C., and P. Lynch. 2011. Adaptive post-processing of short-term wind forecasts for energy applications. Wind Energy **14**:317-325.

Sweeney, C., P. Lynch, and P. Nolan. 2011. Reducing errors of wind speed forecasts by an optimal combination of post-processing methods. Meteorological Applications **20**:32-40.

Taylor, A., and L. Leslie. 2005. A single-station approach to model output statistics temperature forecast error assessment. Weather and Forecasting **20**:1006-1020.

Thorvaldsson, G., H. Bjornsson, and J. Hermannsson. 2004. The influence of weather on early growth rate of grasses. Icelandic Agricultural Sciences **16**:65-73.

Vannitsem, S., and R. Hagedorn. 2011. Ensemble forecast post-processing over Belgium: comparison of deterministic-like and ensemble regression methods. Meteorological Applications **18**:94-104.

Vrac, M., and P. Friederichs. 2015. Multivariate-intervariable, spatial, and temporal-bias correction. Journal of Climate **28**:218-237.

Walsh, S. 2012. A Summary of Climate Averages 1981-2010 for Ireland, Climatological Note No.14, Met Éireann, Dublin.

Wilks, D. 2009. Extending logistic regression to provide full-probability-distribution MOS forecasts. Meteorological Applications **16**:361-368.

Yussouf, N., and D. Stensrud. 2007. Bias-corrected short-range ensemble forecasts of near-surface variables during the 2005/06 cool season. Weather and Forecasting **22**:1274-1286.

Zhai, P., X. Zhang, H. Wan, and X. Pan. 2005. Trends in total precipitation and frequency of daily precipitation extremes over China. Journal of Climate **18**:1096-1108.

Chapter 3

Weather forecasts to enhance an Irish grass growth model

Collaborators on this paper were Caroline Brophy, Elodie Ruelle, Laurence Shalloo, Keith Lambkin and Deirdre Hennessy. An invited revision of the paper has been submitted to the European Journal of Agronomy.

ABSTRACT: Grass growth models have retrospectively predicted grass growth in Ireland using weather observations. However, to predict future grass growth to aid farm management, weather forecasts are necessary inputs. The Moorepark St. Gilles grass growth model (MoSt GGM) was developed to predict perennial ryegrass growth on any Irish farm. To date, it has used local farm information, (retrospective) weather data and management factors to predict daily paddock-level grass growth. Here, we assess the performance of the MoSt GGM using weather forecasts through two studies: daily grass growth predictions at four nitrogen fertiliser application levels using weather forecasts up to ten days in advance were compared with those using weather observations; and the GGM predictions for an Irish dairy farm using observed and forecast weather were compared with on-farm grass growth observations from 2013 to 2016. In the first study, all weather inputs captured the rise in grass growth predictions with higher fertiliser application. Based on the Root Mean Squared Error (RMSE), European Centre for Medium-Range Weather Forecasts (ECMWF) forecasts outperformed a low skill equivalent (based on climatological averages) as GGM inputs up to six days in advance, and up to ten days in advance after bias correction. In the second study, weather observations usually predicted total observed grass growth most accurately; however, weather forecasts sometimes gave lower RMSE values compared to observed weather when weather was overly influential in the model. Weather forecasts are useful inputs to the MoSt GGM, and yield accurate weekly predictions that could aid management decisions.

KEYWORDS: grass growth model, on-farm decision tools, grassland management, weather forecasts, *Lolium perenne* L.

3.1 Introduction

The UN predicts worldwide population to grow from 6.9 billion people in 2010 to 9.7 billion in 2050 (United Nations, 2015). A 1.1% growth per annum in worldwide consumption of agricultural products is projected between 2005 and 2050, giving rise to an approximate food demand growth of 60% in this period (Alexandratos and Bruinsma, 2012). Food Wise 2025 (DAFM, 2015) anticipates an 85% increase in Irish agri-food exports between 2015 and 2025 to capitalise on this extra demand. To meet these targets, Irish farmers must ensure their foodstuffs can be produced sustainably. Maximising economic growth by expanding and making best use of available resources must be coupled with environmental protection (DAFM, 2015). Livestock convert grass into human food such as milk and meat, often on land that is less suitable for crop production (van Zanten et al., 2016; Wilkinson, 2011). It is imperative that Irish dairy and beef farmers make best use of their grassland resources as grazed grass is the cheapest feed source available to them (Dillon et al., 2005; Finneran et al., 2012).

Management of grasslands on a short-term basis is essential to maximise grass growth and utilisation (Creighton et al., 2011). However, some factors that strongly affect farm management decisions are outside the farmer’s influence, for example weather conditions. Irish farmers have a competitive advantage over those from many other countries as the temperate Irish climate provides favourable conditions for high yields of grass dry matter (DM) from perennial ryegrass (*Lolium perenne* L.) over a long grazing season (O'Donovan et al., 2011). Perennial ryegrass growth begins at 5°C and ceases around 20-25°C (Frame, 1992; Hopkins, 2000), so the Irish growing season can last from early Spring to early Winter (Burke et al., 2004; Hopkins, 2000; Hurtado-Uria et al., 2013). Rainfall during the Irish growing season is often optimal for grass growth, although an excess of water can sometimes make grazing impossible or reduce grass growth (Burke et al., 2004). Solar radiation is essential for the conversion of carbon dioxide into biomass (Laidlaw and Frame, 2013). During a large part of the growing season, low solar radiation is more limiting than low temperature, and strongly affects growth in all seasons (Hurtado-Uria et al., 2013; Laidlaw and Frame, 2013).

A grass growth model (GGM) accounting for weather and local conditions, as well as factors that can be controlled by farmers, such as N fertiliser application, grazing rotation length and removal of excess herbage, would aid farm management decisions on feed supply and grassland management. GGMs accounting for some or all of these components have been developed for specific locations in countries such as England (Johnson and Thornley, 1983) and France (Jouven et al., 2006), as well as general models for multiple European sites in different countries (Schapendonk et al., 1998). These models can be mechanistic (for example Johnson and Thornley, 1983; Jouven et al., 2006; Schapendonk et al., 1998) or empirical (for example

Brereton et al., 1996). The model used in this paper is the Moorepark St Gilles (MoSt) GGM (Ruelle et al., 2018) , which is an Irish adaptation of the model developed by Jouven et al. (2006). Although Irish GGMs have been developed, none are being widely used in practice. An operational GGM would allow the farmer to describe their location, soil type and management practices initially. It would then account for these parameters and local weather conditions to predict on-farm growth over the next seven to ten days. Based on these predictions, farmers could make informed management decisions. For example, they could plan to supplement feed if a grass shortage occurs or remove excess herbage from paddocks when there is a surplus of grass on-farm.

Weather forecasts are potentially highly influential inputs for a GGM to predict future grass growth. To date, only predictions from the MoSt GGM using retrospective weather observations have been verified (Ruelle et al., 2018). The inclusion of forecasts will introduce an extra level of uncertainty to the model. McDonnell et al. (2018) assessed the accuracy of European Centre for Medium-Range Weather Forecasts (ECMWF) forecasts at 25 Irish weather stations, and applied bias correction techniques to improve forecast accuracy. Air temperatures were forecast accurately up to ten days in advance, with improvements after bias correction, and rainfall forecasts generally performed well up to five days in advance. However, high rainfall observations were often poorly predicted. Inaccurate rainfall forecasts could decrease the accuracy of grass growth predictions from the MoSt GGM due to the strong influence of rainfall on grass growth.

The objective of this paper is to compare weather observations and forecasts as predictors in the MoSt GGM, and to assess model predictions against on-farm grass growth observations. The practical benefits of the MoSt GGM as an on-farm management decision aid would be improved if weather forecasts can be identified as useful model inputs.

3.2 Materials and Methods

3.2.1 Outline of weather forecast assessment studies

Two assessments of the inclusion of weather forecasts in the Moorepark St Gilles grass growth model (MoSt GGM) were performed:

1. Fertiliser study: The predictions from the GGM when weather observations were used as inputs were compared with those using weather forecasts. Predictions were performed using four different fertiliser application levels, assuming other conditions for a single farm.
2. Observed grass growth study: GGM predictions using i) weather observations and ii) weather forecasts were verified against on-farm grass growth observations.

The GGM predictions, the weather data (observed and forecast) and the observed grass growth data are described in the following sections.

3.2.2 Grass growth model description

The MoSt GGM is a mechanistic grass growth prediction model developed in C++. It describes perennial ryegrass (*Lolium perenne* L.) growth in dairy production systems and is an adaptation of the Jouven model (2006), which was customised for local conditions (Hurtado-Uria, 2013). The MoSt GGM incorporates N, and soil and water sub-models added by Ruelle et al. (2018), which describe the availability of N to the plants and the movement of water through the soil (Fig. 3.1). The sub-models also interact to describe N leaching. The MoSt model is designed to be able to

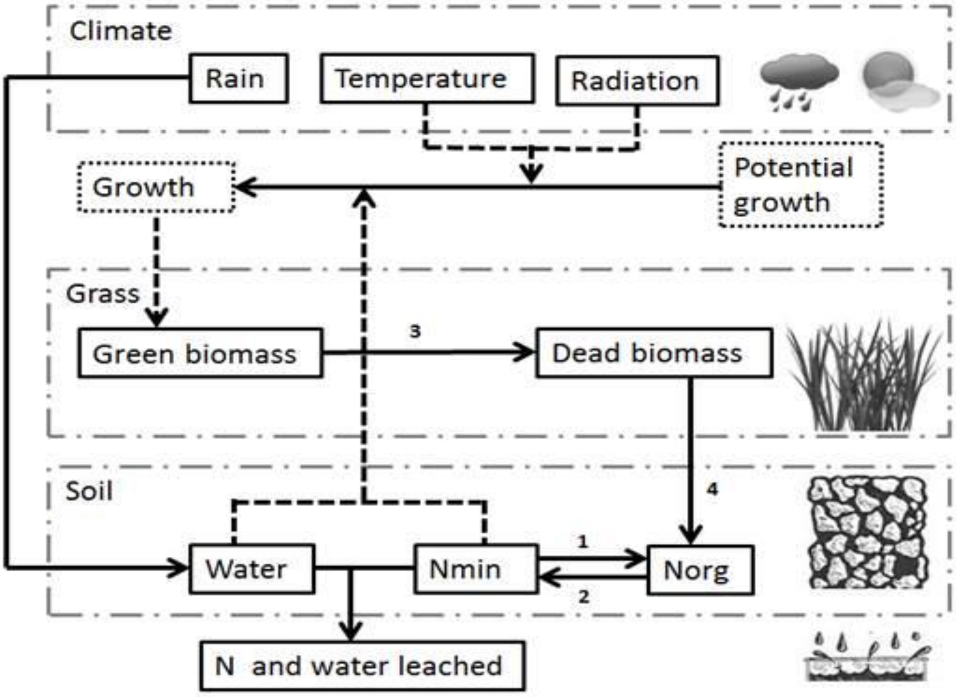


Fig. 3.1: Representation of the physical processes used to predict grass growth in the Moorepark St. Gilles Grass Growth Model (MoSt GGM). This figure is taken from Ruelle et al. (2018).

predict grass growth for any location in Ireland. To do this, it requires a number of inputs based on environmental factors (for example weather data (forecast or observed), N content, soil clay, sand and organic matter content) and management factors (for example grazing data, paddock size, N fertiliser application). Some of these inputs, such as soil type, are initial parameters required to enable the process model to run. Based on all of these factors, it updates the state of the systems controlling grass growth on a daily basis, such as the amount of water and N available to the plant (Fig. 3.1). The final output is a daily grass growth prediction at the paddock level, which can be summed over time. Predictions can be generated for multiple paddocks and aggregated to predict at farm level. The model is described in full detail in Ruelle et al. (2018).

3.2.3 Weather forecasts and observations

Six different weather inputs were examined in the GGM: 1) observed weather, 2) ECMWF forecasts, 3) monthly BC forecasts, 4) yearly BC forecasts, 5) model BC forecasts and 6) mean climatological forecasts, where BC stands for ‘bias corrected’. Daily observations of rainfall, solar radiation and maximum, minimum and mean 2 m air temperature were collected between 2008 and 2016 inclusive at the Met Éireann synoptic weather station at Teagasc, AGRIC, Moorepark, Fermoy, Co. Cork, Ireland (52.16N; 8.26W). For each daily observed weather value, corresponding forecasts from one day to ten days in advance were taken from the ECMWF Atmospheric Model high resolution deterministic 10-day 0000 forecast. Forecasts generated at 0000 on the day of the observation are denoted day-1, and similar notation was used up to day-10 forecast. McDonnell et al. (2018) applied various bias correction techniques (Joliffe and Stephenson, 2011) to the forecast rain and temperature variables to remove systematic biases in the forecasts. We also applied the bias correction techniques to solar radiation forecasts. Monthly and yearly BC use a leave one out method, and subtract the monthly or yearly mean difference between forecast and observed in the training set (i.e. with current year or month excluded) from the forecast being corrected. Model BC by station uses a regression model approach with the month and daily weather forecast as predictors of the observed weather. For each year of daily data, the regression model is calibrated using a leave one out method (i.e. current year excluded), and the resulting model is used to predict the observations. These regression model predictions are then used as the bias-corrected forecasts. Mean climatological forecasts for rainfall and air temperature were created using the Met Éireann mean climatological data for Ireland (Walsh, 2012). Solar radiation observations at Moorepark from between 2001 and 2016 were recorded, and monthly mean climatological forecasts were calculated using them: for example, the average of all values from January over the 16 years provides the forecast for every January day. In the MoSt GGM, missing weather observations were imputed by taking the mean of the observations from the days before and after the missing date. There was only a small number of missing observations: 14 across all weather variables, and only one (rainfall) from 2013 onwards.

3.2.4 Grass growth observations

A grazing experiment investigating the effect of calving date and stocking rate on animal performance was conducted at Teagasc Moorepark Curtins Research Farm, AGRIC, Moorepark, Fermoy, Co. Cork, Ireland (Coffey et al., 2018). Visual cover assessments of grass growth (Hanrahan et al., 2017) recorded for the 54 perennial ryegrass paddocks from 2013 to 2016 inclusive are used in this paper. The grazing season lasted from early February until late November each year and three stocking rates were studied in a randomised block design: 3.28, 2.91 and 2.51 cows/ha. Nitrogen fertiliser application rates were 250 kg N/ha per year for every paddock on every treatment. Full details of the experiment are described by Coffey et al. (2018). The visual assessments were regularly calibrated by cut and weigh measurements as described by O’Donovan et al. (2002a). The daily growth observations for each paddock were calculated using the visual cover estimates, previous growth rate, the number of days pre-grazing, the grazing residuals and the number of days since the last growth figure was obtained (Hanrahan et al., 2017). Although it is the most accurate method of pasture cover estimation (O'Donovan et al., 2002b), there are errors associated with visual assessment (Hanrahan et al., 2017).

3.2.5 Assessment of weather forecasts in MoSt GGM

3.2.5.1 Fertiliser study

Weather observations of rainfall, solar radiation and minimum, maximum and mean 2 m air temperature were employed as model inputs to give paddock-level daily grass growth predictions from the MoSt GGM between 2008 and 2016. These were compared with daily predictions from the GGM using day-1 to day-10 ECMWF and bias-corrected forecasts of rainfall, solar radiation and minimum, maximum and mean 2 m air temperature as inputs. Low-skill mean climatological forecasts were also used as weather inputs to compare with the GGM predictions employing more skilful forecasts. All of these predictions (daily predictions over nine years for each set of weather inputs) were performed at four fertiliser application levels: 0, 100, 200 and 300 kg N/ha, with the first fertilisation for each year on day 65 of the year, and the day after the end of each of the first four grazing events of the year. The first yearly grazing event happened when the paddock height reached 9 cm, and at 8 cm thereafter. There were 40 animals per grazing event in all model runs.

3.2.5.2 Observed grass growth study

Predictions from the MoSt GGM were performed for Teagasc Moorepark Curtins Research Farm, AGRIC, Moorepark, Fermoy, Co. Cork, Ireland (52.17N; 8.27W) using weather observations and forecasts, and were compared to the grass growth observations described in Section 2.2.4. To allow accurate comparisons between the experimental grass growth observations and predictions from the MoSt GGM, the farm management inputs in each of the 54 experimental paddocks were replicated in the model. For each grass growth observation, weather observations were the model inputs from the first day of the year in question until the day before the period of the grass growth observation to allow the updates of the MoSt GGM sub-models. Then the most recent weather forecasts available were used for the period of the grass growth observation. For example, the model run between March 4th and 10th 2013 used weather observations from January 1st to March 3rd, day-1 forecasts for March 4th, and day-7 forecasts for March 10th. The model predictions were also generated using weather observations for the period of the grass growth observation to allow comparisons. If the period of the grass growth observation was greater than ten days, forecasts were not available, and the period was not used for comparisons. This usually happened outside of the peak growing season (April to September). Thus, some periods at the beginning and end of the growing season (February to November) are not described in the study. The period of the grass growth observation is referred to as a ‘weekly’ observation but can be from four to ten days in length since it is the growth between pasture cover estimations. The grass growth was predicted for each paddock with available grass growth figures for the ‘week’, and the ‘weekly’ average paddock values were computed to describe average farm growth. The ‘weekly’ values were scaled to daily averages to ensure reasonable comparisons across weeks, and did not always contain the same number of paddocks.

3.2.6 Statistical criteria for comparisons

The criteria used to compare the MoSt model predictions with each other and with observed grass growth include Mean Systematic Bias (MSB = $\frac{\sum_{i=1}^{n}(p_i-o_i)}{n}$), Mean Squared Error (MSE =$\frac{\sum_{i=1}^{n}(p_i-o_i)^2}{n}$) and Root Mean Squared Error (RMSE =$\sqrt{MSE}$), where p_i and o_i are the i^{th} predicted and observed values respectively, and n is the number of predicted and observed values. Relative Prediction Error (RPE) is the RMSE divided by the mean of the observed values (Rook et al., 1990). The MSE can be partitioned into errors in central tendency (mean bias = $(\bar{p}-\bar{o})^2$), errors due to regression (slope bias = $\sigma_p^2(1-b)^2$) and errors due to unexplained random variation ($\sigma_o^2(1-R^2)$) (Dhanoa et al., 1999), i.e.

$$MSE = \frac{\sum_{i=1}^{n}(p_i-o_i)^2}{n} = (\bar{p}-\bar{o})^2 + \sigma_p^2(1-b)^2 + \sigma_o^2(1-R^2)$$

where σ_p^2 and σ_o^2 are the variances of the predictions and observations respectively, b is the slope of the regression line of observed on predicted and R^2 is the coefficient of determination from this regression. Ideally, the MSE (and therefore each of the three components) is close to zero, meaning the predictions and observations agree closely. A high mean bias component means the predictions are consistently over or under predicting the observed values. A high slope bias means the best fit regression line is not similar to the line of equality, meaning equality of predictions and observations does not describe the relationship. Errors due to unexplained random variation cannot be bias corrected by linear correction methods. A high value for errors due to unexplained random variation indicates that the points are scattered widely about the best-fit regression line. These comparison methods were also sometimes used to compare two sets of predictions from the MoSt GGM, those using weather forecasts with those using weather observations.

3.3 Results

3.3.1 Fertiliser study

Yearly total grass growth predictions increased as the amount of N fertiliser increased. For example, the predicted grass growth yearly totals in 2016 from model runs using weather observations with 0, 100, 200 and 300 kg N/ha were 7036, 9697, 10807 and 12726 kg DM/ha, respectively. However, across N application levels for each forecast period, the RPE of the predicted grass growth using ECMWF forecasts versus observed weather was similar, so there were no notable changes in the accuracy of the predictions with N application levels (Appendix 3.1).

The MoSt GGM predictions using forecasts generally followed those using weather observations closely for 200 kg N/ha of fertiliser (Fig. 3.2, Appendix 3.2). This shows that forecasts can be interchanged with observed weather with no serious change to the grass growth predictions, but the interchangeability decreases as the forecast period increases (Appendix 3.3). Of the weather forecast types examined, grass growth predictions using model BC forecasts gave the lowest RMSE values at all forecast periods (Table 3.1). It was also the most effective forecast in predicting grass growth in the short term (Fig. 3.2, Appendix 3.2), but did not capture the variations in daily growth well for forecast periods over five days (Appendix 3.3). The ECMWF forecasts yielded better grass growth predictions than the low-skill mean climatological forecasts up to six days in advance, but not for any longer forecast periods (Table 3.1).

For each forecast period in each year, most of the MSE from the GGM predictions using ECMWF forecasts was attributable to unexplained random variation (always above 60% for all N application levels). The remainder was predominantly due to regression. The error due to unexplained random variation was

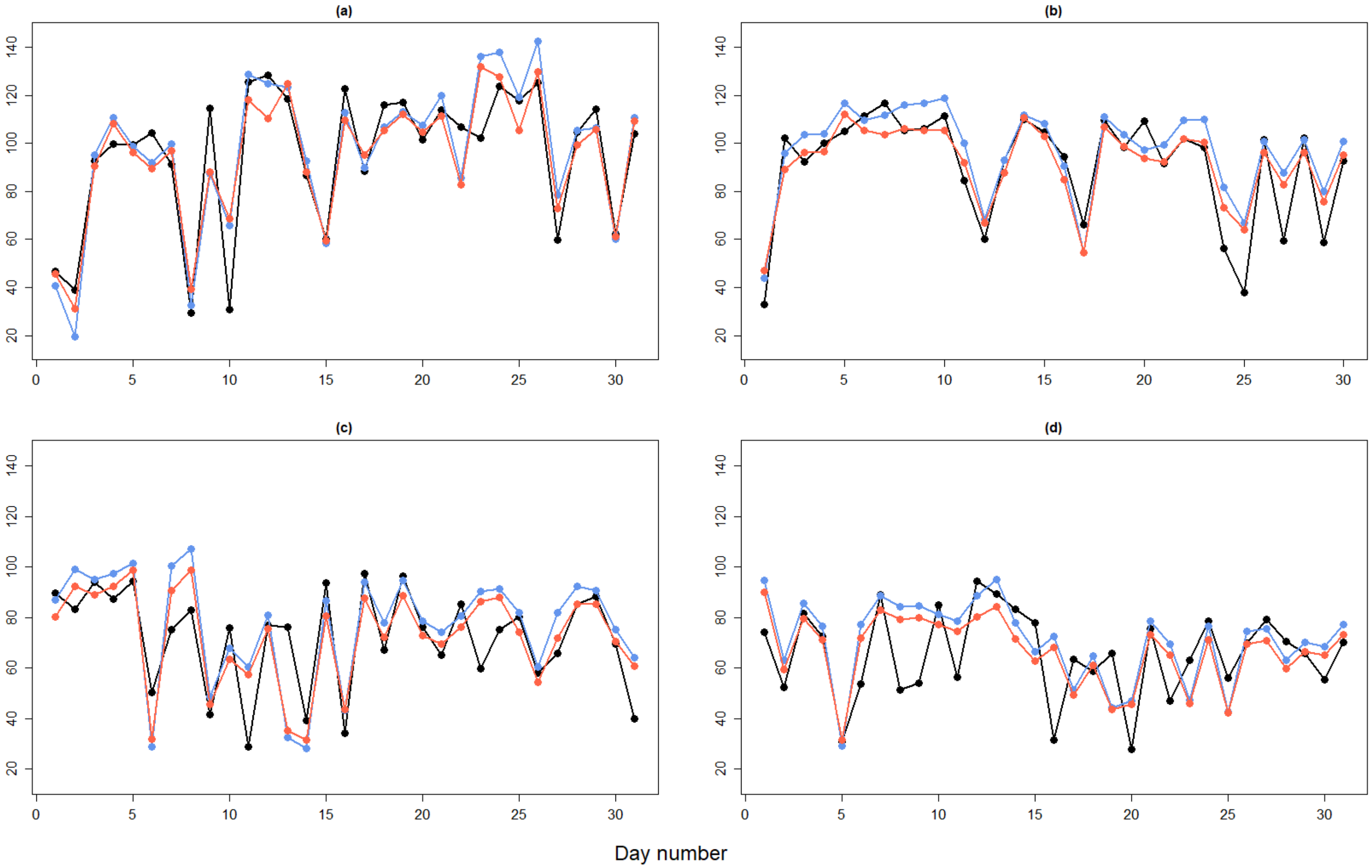


Fig. 3.2: Predicted daily grass growth in 2015 for (a) May, (b) June, (c) July and (d) August from the MoSt GGM using weather observations (black), day-2 ECMWF forecasts (blue) and day-2 model BC forecasts (red) with 200 kg N/ha.

	Weather forecast type				
Forecast period	ECMWF	Month BC	Year BC	Model BC	Mean climatology
1	9.1	9.2	9.3	8.6	16.6
2	10.2	10.2	10.3	9.8	17.0
3	11.7	11.8	11.9	10.9	17.0
4	13.6	13.9	13.9	12.6	17.0
5	15.3	16.0	15.7	13.8	17.0
6	16.3	17.3	17.1	14.8	17.0
7	18.2	19.0	18.7	15.6	17.0
8	18.4	19.0	19.0	15.9	17.0
9	20.2	20.7	20.8	16.6	17.0
10	20.1	20.8	20.8	16.9	17.0

Table 3.1: Forecast periods 1 to 10 RMSE values (kg DM/ha) comparing daily grass growth predictions for 2008 to 2016 from the MoSt grass growth model using observed weather, with grass growth predictions from the model using various weather forecasts. The fertiliser application level for these predictions is 200 kg N/ha. The most accurate predictions for each forecast period are highlighted.

only 63.6% of the total MSE at forecast period nine in 2012 with 300 kg N/ha. In 2012, the errors due to regression from the GGM predictions using ECMWF forecasts were higher at forecast periods of over five days than in the other years examined. This was due to ECMWF forecasts of solar radiation and rainfall over-predicting and under-predicting, respectively, the observations for many days in June 2012, leading to over-predictions of grass growth in the month. Model bias correction of the day-9 forecasts reduced the RMSE of the grass growth predictions in June 2012 from 41.2 kg DM/ha when ECMWF forecasts were used to 31.2 kg DM/ha.

3.3.2 Observed grass growth study

The MoSt GGM using the different weather inputs predicted the grass growth observations accurately in most weeks (Fig. 3.3). The grass growth observations suitable for comparisons in 2015 went from February 16th to November 2nd, the longest of the four years in the study (Fig. 3.3c). The weekly MoSt GGM predictions at the beginning and end of the period were lower than observed grass growth (Fig. 3.3c). This trend was apparent outside the peak growing season in all years and resulted in the under-prediction of yearly grass growth by all simulations, regardless of whether observed or forecast weather was used. To exclude the poor predictions at the beginning and end of the full growing season, the total grass growth over the peak growing season (sum of average paddock values on each measurement date from 1st April to 30th September) was examined. During this period, the predictions using ECMWF forecasts were within 25% of the weekly grass growth observation for two-thirds of all observations. MoSt GGM predictions using actual weather, ECMWF forecasts, model BC forecasts and mean climatology forecasts as weather inputs predicted the grass growth during the peak growing season accurately in all years from 2013 to 2016, inclusive (Fig. 3.4). As expected, the actual weather simulations gave the closest values to the observed grass growth in 2013, 2014 and 2015 (Fig. 3.4). However, in 2016, the peak growing season total was under-predicted by all simulations, with mean climatology simulations giving the best prediction (Fig. 3.4). The monthly observed grass growth totals were followed closely by monthly grass growth predictions for each of the weather inputs, but no weather input gave clearly better monthly totals (Table 3.2). However, in April 2013 there was a clear benefit of using weather inputs other than mean climatology, since the monthly prediction based on mean climatological weather was much higher than

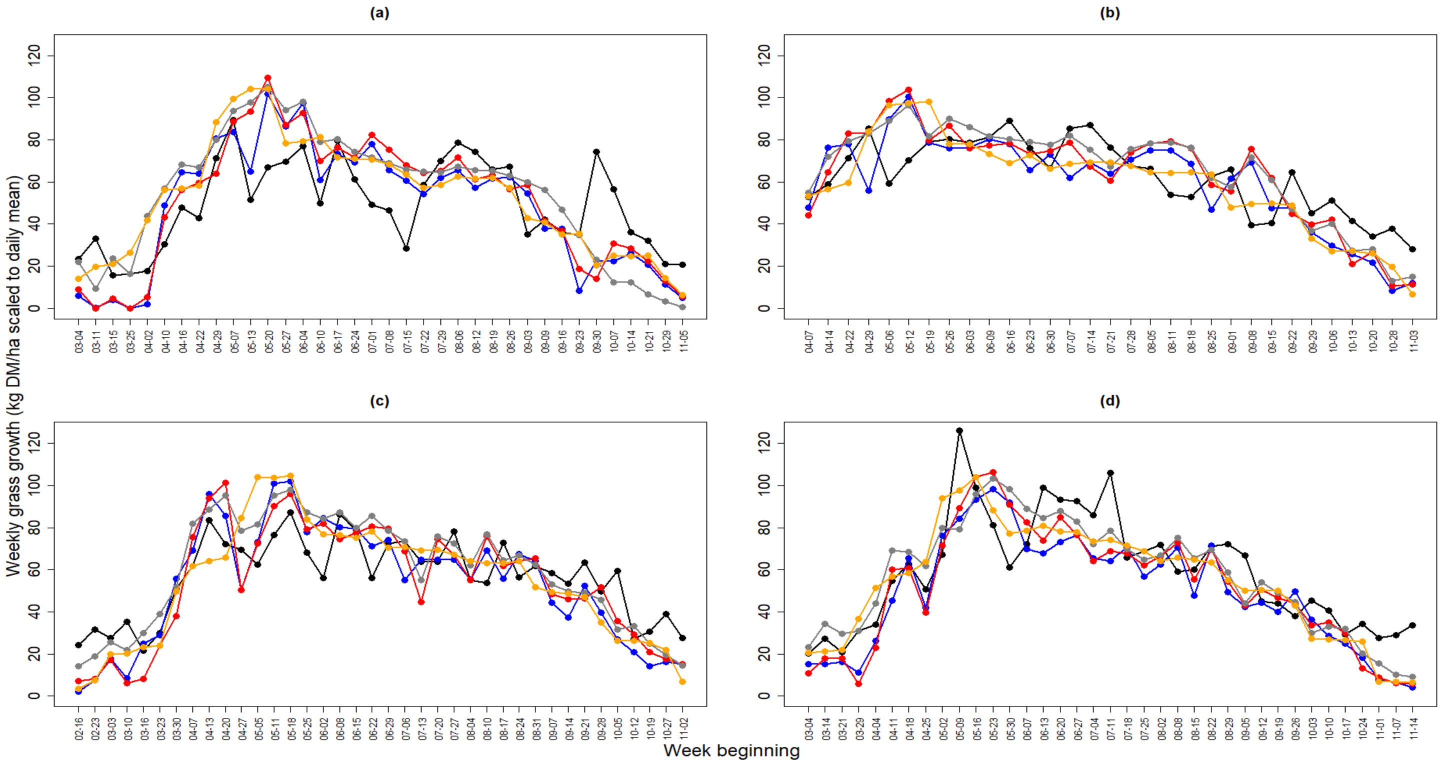


Fig. 3.3: Weekly grass growth (kg DM/ha scaled to daily mean) across all 54 paddocks (black), and corresponding predicted yearly grass growth from the MoSt GGM using observed weather (blue), ECMWF forecasts (red), model BC forecasts (grey), and mean climatology forecasts (orange) in (a) 2013, (b) 2014, (c) 2015, and (d) 2016. Grass growth is only shown for weeks in which it was recorded.

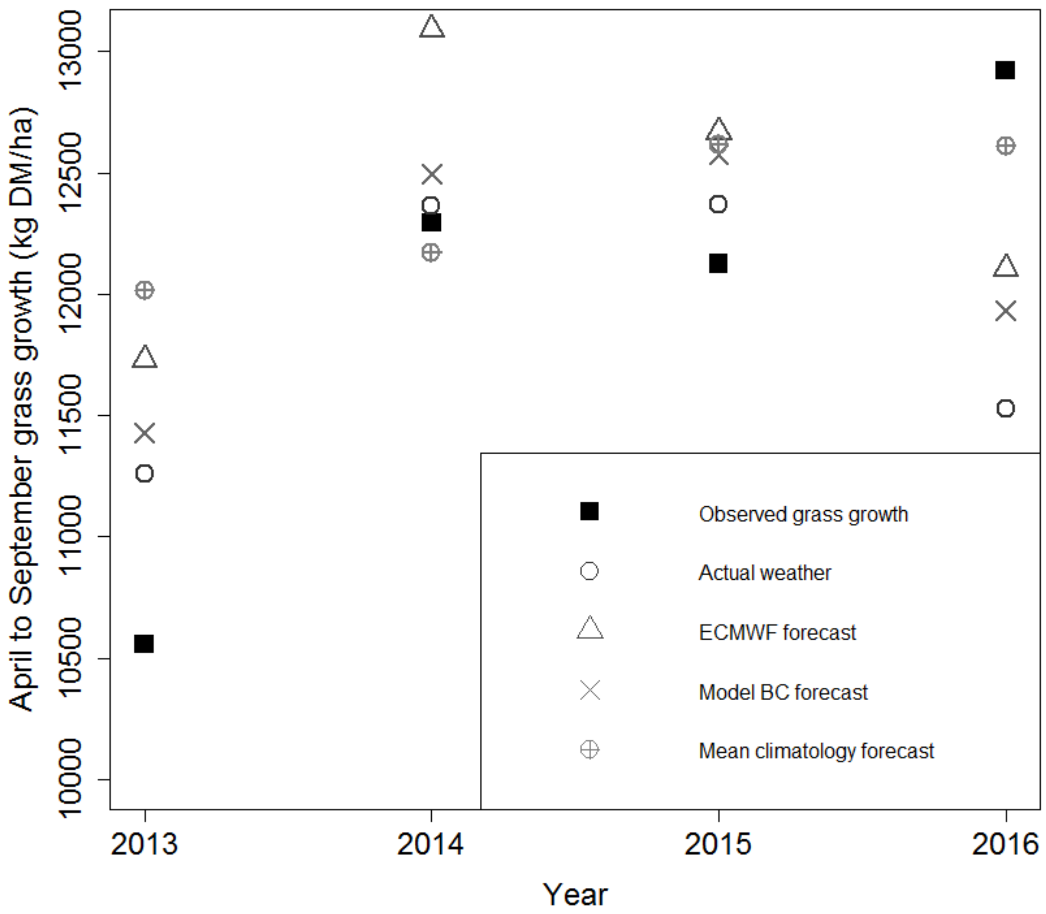


Fig. 3.4: Total grass growth from April to September inclusive, averaged across all 54 paddocks from 2013 to 2016 (filled squares), and corresponding predicted grass growth from the MoSt GGM using observed weather (empty circles), ECMWF forecasts (triangles), model BC forecasts (crosses) and mean climatology forecasts (crossed circles).

Table 3.2: Monthly grass growth observations and predictions (kg DM/ha) for 2013 to 2016.

			Weather input			
Month	Year	Observed grass growth	Actual	ECMWF	Model BC	Mean climatology
4	2013	1481	1790	1566	1763	2127
5	2013	1924	2358	2648	2616	2681
6	2013	1798	2008	2084	1992	2044
7	2013	1837	2306	2550	2285	2286
8	2013	1991	1723	1753	1670	1699
9	2013	1525	1074	1128	1102	1177
4	2014	1937	1881	1990	2027	1831
5	2014	2044	2400	2572	2490	2571
6	2014	2665	2537	2582	2478	2437
7	2014	2280	1929	2039	1975	1991
8	2014	1581	1782	1966	1808	1735
9	2014	1788	1835	1943	1716	1607
4	2015	2016	2088	2223	2274	1958
5	2015	2065	2485	2372	2522	2752
6	2015	2393	2644	2676	2591	2567
7	2015	2034	1812	1853	1848	2000
8	2015	2043	2127	2199	2074	2079
9	2015	1574	1216	1345	1266	1260
4	2016	1416	1252	1283	1406	1612
5	2016	3103	3197	3326	3259	3305
6	2016	2427	1941	2143	2083	2129
7	2016	2355	1853	1901	1895	2083
8	2016	2265	2049	2165	2075	2128
9	2016	1356	1236	1285	1214	1354

the observed grass growth (Table 3.2). The observed weather in April 2013 differed significantly from the climatology in April, and the MoSt GGM predictions using actual weather, and ECMWF and bias-corrected weather forecasts predicted the resulting low grass growth more accurately than mean climatology since weather observations and forecasts capture the variability that mean climatology does not (Table 3.2). In other months, the monthly grass growth total was not predicted well regardless of the weather input, for example in September 2013 (Table 3.2). This was likely due to the fact that the MoSt GGM must describe a number of physical processes to predict grass growth such as N leaching and the water content in the soil. These physical processes cannot be checked for accuracy using actual observations, and when they become inaccurate the grass growth predictions are less accurate.

The MoSt model run using model BC forecasts gave a lower RMSE value in 2015 than the runs using actual weather, and ECMWF and mean climatology forecasts (Table 3.3). In some of the years, and in some seasons within the years, actual weather had higher RMSE values than model BC and mean climatology forecasts. Splitting the MSE (Dhanoa et al., 1999) showed that there were high errors in central tendency in 2016, where they made up over 19% of the total MSE for actual weather, and ECMWF and model BC forecasts. This can be seen in the under-predictions across the peak growing season (Fig. 3.4). There were high slope biases in 2013, 2014 and 2015, with values of between 42% and 60% for actual weather, and ECMWF and model BC forecasts.

Table 3.3: RMSE values (kg DM/ha) comparing daily average of ‘weekly’ grass growth observations for 2013 to 2016 with grass growth predictions from the model using the weather observations and various weather forecasts.

	Weather input			
Year	**Actual**	**ECMWF**	**Model BC**	**Mean climatology**
2013	19.6	21.4	20.1	20.4
2014	16.8	17.2	15.4	14.5
2015	15.1	15.1	14.2	15.5
2016	17.9	16.8	16.6	14.3

3.4 Discussion

3.4.1 Fertiliser study

The results suggest that forecasts can be useful predictors in a GGM. The fertiliser study showed that grass growth predictions from the MoSt GGM using weather forecasts can give similar predictions to those using weather observations, particularly after bias corrections, although decrease in accuracy as forecast period increases (Fig. 3.2, Table 3.1). This is because weather forecasts predict weather observations less accurately as forecast period increases. The example from Section 3.3.1 in which poor ECMWF forecasts in June 2012 resulted in poor predictions from the MoSt GGM illustrates the influence that inaccurate weather forecasts can have on predictions from the MoSt GGM, and how bias corrections can improve prediction accuracy.

The predictions using model BC forecasts failing to capture variability was linked to the problem identified with model BC of rainfall forecasts in McDonnell et al. (2018): ECMWF forecasts (particularly of rainfall) were not accurate at longer forecast lead times, so model BC forecasts were close to the mean rainfall. This resulted in similarly conservative grass growth estimates from the MoSt GGM. When using the MoSt GGM in practice, if the intention is to get accurate daily grass growth predictions between 1 and 6 days in advance, it is generally best to use model BC forecasts (Fig. 3.2, Appendix 3.2). However, if the priority is to predict the daily fluctuations in grass growth for 7-10 days in advance, it would be recommendable not to use model BC forecasts or mean climatological forecasts, but to use ECMWF forecasts instead (Appendix 3.3).

3.4.2 Observed grass growth study

The observed grass growth simulations showed that grass growth could be modelled accurately during the peak growing season using weather observations, and predicted four to ten days in advance by weather forecasts. Since predictions using weather forecasts were within 25% of the observed grass growth on two out of three occasions, this could be a useful predictive tool to aid farm management. Model BC and mean climatological forecasts often gave the most accurate grass growth predictions, according to the RMSE values (Table 3.3). RMSE gives a high weight to high errors, and the more conservative weather forecasts yield grass growth predictions with lower RMSE values, suggesting that the MoSt GGM does not capture some extreme grass growth observations. Mean climatology can be a useful weather input, as it gives the expected grass growth performance of the farm based on the usual weather for the month in question. Thus, it is a good prediction to compare with predictions using other weather inputs which should describe the fluctuations in grass growth more accurately.

The high slope biases observed in 2013, 2014 and 2015 were due to the under-prediction of many of the lower grass growth observations, which occurred at the beginning and end of the growing season, (Fig. 3.3). The grass growth predictions in the shoulders of the distribution (at the beginning and end of the growing season) were generally lower than the observed grass growth for all weather inputs to the MoSt model. The model can be examined for possible causes, so this under-prediction can be corrected. This could also be due to the method of estimation for the grass growth observations. Visual observation may result in the estimation of higher covers on pastures than the actual values. Many of the high grass growth observations were also over-predicted, adding to the slope bias. The MoSt GGM under-predicted grass growth in the peak growing season in 2016 (Fig. 3.4). This appears to be due to an under-estimation of grass growth in a number of summer weeks, mainly weeks 10 (May 9th to May 15th), and 15 to 19 (June 13th to July 17th) (Fig. 3.3d). The weather input with the highest solar radiation values in those weeks gave the highest grass growth predictions in each case. However, all of the predictions were lower than observed grass growth. In all of these weeks, solar radiation was below mean climatology. It may be that the influence of low solar radiation is being overstated in the MoSt GGM in the summer months.

These predictions were conducted at one site which is a well-managed dairy research farm with free-draining soils and high grass growth, and easily accessible weather observations. However, if this GGM tool was to be used in practice by farmers as part of the Pasture Base Ireland framework (Hanrahan et al., 2017), it would have to give accurate predictions in many different locations on different soil types with varying levels of farm management (Ruelle et al., 2017). Although an increasing number of Irish farmers record daily on-farm weather observations, they will not be available at most sites, and the bias-correction methods may not give improvements as the nearest available weather stations may not describe the farm accurately enough. In these cases, it would make sense to use the ECMWF forecasts in the MoSt GGM.

3.4.3 Future work

The predictions presented in this paper are from the MoSt GGM described in Ruelle et al. (2018). As with any predictive model, it can be updated to improve the accuracy of the predictions. The MoSt GGM will be incorporated into the Pasture Base Ireland framework (Hanrahan et al., 2017), and use weather forecasts to give farmers grass growth predictions for their farm. Grass growth predictors such as the NZ Pasture Growth Forecaster (Dairy NZ, 2018) are being used in practice but do not use weather forecasts. Australian studies have assessed the potential usefulness of seasonal climate forecasts in agriculture, concluding that the lead time and potential managements benefits did not warrant widespread adoption by farmers (Ash et al., 2007). Another study included monthly to seasonal climatological hindcasts in a GGM (Harrison et al., 2017), and found that global climate models predicting pasture growth rates gave similar predictions to when historical climate data was included. However, none of these have used short-term weather forecasts. The MoSt GGM could be run using seasonal forecasts to allow farmers to plan for the coming season. The MoSt GGM currently predicts for perennial ryegrass systems (Ruelle et al., 2018). However, it could be adapted to include mixed-species swards, including those with white clover. White clover is being included in increasing numbers of grassland systems in Ireland because of its N fixation traits, and the resultant increase in dry matter yield (Guy et al., 2018), as well as increased animal performance associated with mixed perennial ryegrass white clover swards (Egan et al., 2018). Sowing species mixtures that include legumes such as white clover can help to stabilise yield output at different levels of N application (Enriquez-Hidalgo et al., 2016). For example, in the fertiliser study, the dry matter yields would probably have been more similar across the N fertiliser levels if white clover was included in the system.

The MoSt GGM under-predicted grass growth in the early and late growing season. Also, many of the extreme grass growth observations were not detected by the model for any of the weather inputs. It should be investigated whether these poor predictions happen at other locations, and if so, the causes of the problems identified and fixed. Weather forecasts are not usually good at predicting extremes accurately at a weekly scale, but the GGM model runs using weather observations could be updated to capture the extreme grass growth values. It is important for farmers to be able to prepare for extreme weather and grass growth conditions.

3.5 Conclusions

The MoSt GGM can utilise weather forecasts to predict short-term grass growth, and aid farmers with their daily management decisions. It has been shown to capture the variability in systems using different amounts of N fertiliser, and to accurately describe weekly on-farm grass growth observations. We have demonstrated that weather forecasts can be a useful input to a grass growth model and have the potential to enhance on-farm resource use efficiency, as pressure mounts on farms to increase outputs to meet extra food demands.

References

Alexandratos, N. and Bruinsma, J., 2012. World agriculture towards 2030/2050: the 2012 revision, ESA Working paper No. 12-03. Rome, FAO.

Ash, A., McIntosh, P., Cullen, B., Carberry, P. and Smith, M., 2007. Constraints and opportunities in applying seasonal climate forecasts in agriculture. Australian Journal of Agricultural Research, 58(10): 952-965.

Brereton, A.J., Danielov, S.A. and Scott, D., 1996. Agrometeorology of Grass and Grasslands for Middle Latitudes. World Meteorological Organisation, Technical Note No. 197. Geneva.

Burke, J.I., Brereton, A.J., O'Kiely, P. and Schulte, R., 2004. Weather and Crop Production. In Climate, Weather and Irish Agriculture, 2nd. Edition. AGMET, Weather and Crop Production. In Climate, Weather and Irish Agriculture, 2nd. Edition.

Coffey, E., Delaby, L., Fleming, C., Pierce, K. and Horan, B., 2018. Multi-year evaluation of stocking rate and animal genotype on milk production per hectare within intensive pasture-based production systems. Journal of Dairy Science, 101(3): 2448-2462.

Creighton, P., Kennedy, E., Shalloo, L., Boland, T. and O' Donovan, M., 2011. A survey analysis of grassland dairy farming in Ireland, investigating grassland management, technology adoption and sward renewal. Grass and Forage Science, 66(2): 251-264.

DAFM, 2015. Food Wise 2025. Department of Agriculture Food and the Marine, http://www.agriculture.gov.ie/media/migration/agri-foodindustry/foodwise2025/report/FoodWise2025.pdf.

Dhanoa, M., Lister, S., France, J. and Barnes, R., 1999. Use of mean square prediction error analysis and reproducibility measures to study near infrared calibration equation performance. Journal of Near Infrared Spectroscopy, 7(3): 133-143.

Dillon, P., Roche, J., Shalloo, L. and Horan, B., 2005. Optimising financial return from grazing in temperate pastures. In: M. JJ (Editor), Utilisation of grazed grass in temperate animal systems. Wageningen Academic, Cork, Ireland.

Egan, M., Galvin, N. and Hennessy, D., 2018. Incorporating white clover (Trifolium repens L.) into perennial ryegrass (Lolium perenne L.) swards receiving varying levels of nitrogen fertilizer: Effects on milk and herbage production. Journal of Dairy Science, 101(4): 3412-3427.

Enriquez-Hidalgo, D., Gilliland, T. and Hennessy, D., 2016. Herbage and nitrogen yields, fixation and transfer by white clover to companion grasses in grazed swards under different rates of nitrogen fertilization. Grass and Forage Science, 71(4): 559-574.

Finneran, E. et al., 2012. Stochastic simulation of the cost of home-produced feeds for ruminant livestock systems. Journal of Agricultural Science, 150: 123-139.

Frame, J., 1992. Improved Grassland Management. Farming Press Books, Wharfedale Road, Ipswich IP1 4 LG, United Kingdom.

Hanrahan, L. et al., 2017. PastureBase Ireland: A grassland decision support system and national database. Computers and Electronics in Agriculture, 136: 193-201.

Harrison, M., Christie, K. and Rawnsley, R., 2017. Assessing the reliability of dynamical and historical climate forecasts in simulating hindcast pasture growth rates. Animal Production Science, 57(7): 1525-1535.

Hopkins, A., 2000. Grass : its production and utilization. Published for the British Grassland Society by Blackwell Science, Oxford ; Malden, MA, xiii, 440

Hurtado-Uria, C., 2013. Evaluation, adaptation and validation of a model to predict grass growth in Ireland, Cork Institute of Technology.

Hurtado-Uria, C., Hennessy, D., Shalloo, L., O'Connor, D. and Delaby, L., 2013. Relationships between meteorological data and grass growth over time in the south of Ireland, Irish Geography, pp. 175-201.

Johnson, I. and Thornley, J., 1983. Vegetative crop growth-model incorporating leaf-area expansion and senescence, and applied to grass. Plant Cell and Environment, 6(9): 721-729.

Joliffe, I.T. and Stephenson, D.B., 2011. Forecast Verification: A Practitioner's Guide in Atmospheric Science, 2nd Edition. Wiley.

Jouven, M., Carrere, P. and Baumont, R., 2006. Model predicting dynamics of biomass, structure and digestibility of herbage in managed permanent pastures. 1. Model description. Grass and Forage Science, 61(2): 112-124.

Laidlaw, A. and Frame, J., 2013. Improved grassland management. Crowood, 2013, Ipswich, U.K. Alexandria Bay, NY, 352 pp.

McDonnell, J. et al., 2018. Verification and bias correction of ECMWF forecasts for Irish weather stations to evaluate their potential usefulness in grass growth modelling. Meteorological Applications, 25(2): 292-301.

O'Donovan, M., Connolly, I., Dillon, P., Rath, M. and Stakelum, G., 2002a. Visual assessment of herbage mass. Irish Journal of Agricultural and Food Research, 41(2): 201-211.

O'Donovan, M., Dillon, P., Rath, M. and Stakelum, G., 2002b. A comparison of four methods of herbage mass estimation. Irish Journal of Agricultural and Food Research, 41(1): 17-27.

O'Donovan, M., Lewis, E. and O'Kiely, P., 2011. Requirements of future grass-based ruminant production systems in Ireland. Irish Journal of Agricultural and Food Research, 50(1): 1-21.

Rook, A., Dhanoa, M. and Gill, M., 1990. Prediction of the voluntary intake of grass silages by beef-cattle. 3. Precision of alternative prediction models. Animal Production, 50: 455-466.

Ruelle, E., Delaby, L. and Hennessy, D., 2017. Usefulness of nitrogen application in heavy soils compared to more favourable land in Ireland - utilisation of the Moorepark Grass Growth model, European Grassland Federation 2017. Grassland Science in Europe, Alghero, Italy.

Ruelle, E., Delaby, L. and Hennessy, D., 2018. Development of the Moorepark St Gilles grass growth model (MoSt GG model): A predictive model for grass growth for pasture based systems. European Journal of Agronomy, 99(September 2018): 80-91.

Schapendonk, A., Stol, W., van Kraalingen, D. and Bouman, B., 1998. LINGRA, a sink/source model to simulate grassland productivity in Europe. European Journal of Agronomy, 9(2-3): 87-100.

United Nations, 2015. World population prospects: The 2015 revision, key findings and advance tables. Department of Economic and Social Affairs, Population Division, Working Paper No. ESA/P/WP.241.

van Zanten, H., Mollenhorst, H., Klootwijk, C., van Middelaar, C. and de Boer, I., 2016. Global food supply: land use efficiency of livestock systems. International Journal of Life Cycle Assessment, 21(5): 747-758.

Walsh, S., 2012. A Summary of Climate Averages 1981-2010 for Ireland, Climatological Note No.14, Met Éireann, Dublin.

Wilkinson, J., 2011. Re-defining efficiency of feed use by livestock. Animal, 5(7): 1014-1022.

Chapter 4

A mixed model for assessing the effect of many species interactions in complex grassland ecosystems

Collaborators on this paper were Thomas McKenna, Kathryn Yurkonis, Deirdre Hennessy, Rafael de Andrade Moral and Caroline Brophy. The paper has been formatted for the Journal of the American Statistical Association.

Abstract: Weed invasion in grasslands can reduce yields and forage quality. Increasing biodiversity or manipulating the spatial pattern of species in grasslands may mitigate weed invasion. However, when many species are present in an ecosystem, it can be difficult to model the numerous interactions among species. Motivated by the Species Pattern and Community Ecology (SPaCE) experiment that investigated the effect of species diversity and spatial planting pattern on weed invasion, we developed a Diversity-Interactions mixed model that can assess large numbers of species interactions and test the effects of spatial planting pattern over multiple years. The model includes a large number of pairwise interaction random effects, that are indexed by pairs of species rather than a plot-level factor as is typical in mixed models, but that only require one additional variance parameter per year. Applying our method, we identified species with functional traits (in particular nitrogen-fixing legumes) that reduced weed invasion when mixed with species that have other functional traits (e.g. grasses). We found that both spatial pattern and the proportion of legume in mixture influenced the species' interaction effects on weed invasion, but that the identities of species in the mixture determined whether overall weed invasion was increased or reduced.

KEYWORDS: weed invasion, Diversity-Interactions modelling, variance-covariance structure, random effects, species interactions, grassland

4.1 Introduction

Weed invasion has been shown to reduce yields and forage quality in grasslands, and can prevent restored grasslands from reaching their diversity goals (Levine et al. 2003, Corbin and D'Antonio 2012). With the increasing demand for food worldwide, the agricultural industry is under increasing pressure to maximize resource use (Foley et al. 2011). Weed invasion is a source of inefficiency and must be reduced to make grassland production systems more economically and environmentally sustainable. Herbicide use to control weeds is expensive (DiTomaso 2000) and can negatively impact the environment and human health (Freemark and Boutin 1995, Mahanty et al. 2017). Combining multiple species in grassland ecosystems offers a solution that is less economically and environmentally expensive, since it has been shown to reduce weed invasion by making better uses of resources (Dukes 2002, Hooper et al. 2005, Maron and Marler 2008, Sturludottir et al. 2014, Connolly et al. 2018). However, it can be challenging to model biodiversity effects in ecosystems with large numbers of species due to the numerous species interactions that may be jointly affecting ecosystem responses.

When sowing a multi-species mixture, species can be randomly dispersed, or a spatial pattern may be induced whereby individuals of the same species are sown in an aggregated pattern such that individuals are grouped together while maintaining density. Some studies have shown that spatially structured seeding approaches can be used to control species dominance and increase productivity (Rayburn and Schupp 2013, Zhang, Wang, and Yu 2014, McKenna and Yurkonis 2016, Seahra, Yurkonis, and Newman 2016). The impact of spatial pattern on grassland weed invasion is less well known, but it has been shown that the proximity of the individuals of the same species within plots can promote the persistence of sown species, thus promoting diversity and reducing weed invasion (Lamosova et al. 2010, Yurkonis, Wilsey, and Moloney 2012).

4.1.1 Background of statistical methods used in biodiversity and ecosystem function research

Ecosystem functions of grasslands are community-level responses such as plant biomass, weed biomass, and nitrogen yield (Hooper et al. 2005). Biodiversity can depend on species composition (identities of species), species richness (number of species) and functional group richness (the number of different plant groups that perform certain functions), as well as factors such as how equal the abundance of each species is (evenness). The biodiversity and ecosystem function (BEF) relationship has been studied in a wide range of ecosystem types including marine, forest and grassland ecosystems, and increasing biodiversity has been shown to consistently have positive effects on multiple ecosystem functions (Tilman et al. 2014). In grasslands, BEF studies have shown that biomass production and weed suppression increase with diversity as a result of increasing functional group diversity (Petermann et al. 2010, Cardinale et al. 2011, Isbell et al. 2017). However, many methods focus mainly on the number of species present (richness) as the driver of diversity effects in BEF relationships (Byrnes et al. 2014). These methods make it difficult to distinguish the effects of individual species identities, changing evenness, and species' interactions in mixtures. Diversity-Interactions (DI) models were introduced to facilitate the analysis of multi-species mixtures experiments using species proportions as predictors of ecosystem functions (Kirwan et al. 2007, Kirwan et al. 2009, Connolly et al. 2013, Dooley et al. 2015, Brophy et al. 2017). Diversity-Interactions models allow estimation of the contribution of each individual species and its interactions with other species, and the effects of evenness and richness on ecosystem function.

4.1.2 Overview of the SPaCE experiment

The Species Pattern and Community Ecology (SPaCE) experiment is a plot-based grassland BEF experiment designed to investigate the effects of biodiversity (richness and evenness) and species spatial planting pattern on weed and grass biomass (McKenna and Yurkonis 2016). Monoculture (single species) and mixture plots of up to eight species were constructed from a pool of sixteen species, ranging from grasses and legumes, to large herbs such as *Helianthus maximiliani* (maximilian sunflower). The spatial pattern treatment involved species in mixture plots being either randomly dispersed or aggregated in conspecific clusters within plots while maintaining plant density. Plots were hand weeded on a monthly basis during the growing season, and annual weed invasion was recorded over three growing seasons (2012-2014). The data presented three specific statistical challenges each to be addressed in a multi-year setting: 1) how to model the $\binom{16}{2} = 120$ pairwise species interactions in a biologically meaningful way; 2) how to assess the impact of spatial pattern, a manipulated treatment that was applied to mixture, but not monoculture, plots; and 3) given the variety of species types in the study, how to adjust for heterogeneous variation across plots within each year.

4.1.3 Overview of proposed methodology

We develop a new Diversity-Interactions model that will address the statistical challenges presented in the SPaCE experimental data. These include integrating the spatial pattern treatment into diversity effects, modelling the variance-covariance structure of the model to adjust for heterogeneous variance across plots, and introducing novel multi-year random pairwise interaction effects to handle the large numbers of possible species' interactions at play. The random effects included are highly unusual in that they are indexed by pairs of species (and year), rather than by a plot-level factor, which is more typical in mixed models. They are included to supplement a low degree of freedom fixed-effect description of the species interactions to acknowledge that there may be additional variation due to species interactions that cannot be captured by biologically meaningful fixed effects. For each year, a new random effect is introduced for each pair of species (120 pairs in each year, totaling 360 random effects), and they are assumed to be independent and to have the same variance, thus only introducing a single variance parameter for each year, i.e. 3 extra parameters in the model, not 360.

Section 4.2 contains a review of DI models, a detailed description of the SPaCE dataset, and a description of the new methods introduced in this paper. Section 4.3 contains details of model fitting and comparisons, as well as some simulation studies to illustrate the scope of the developed DI modelling methods. Section 4.4 is the results section, while section 4.5 provides discussion and conclusions drawn from the results.

4.2 Data and methods

4.2.1 Review of Diversity-Interactions models

Biodiversity and ecosystem function (BEF) experiments typically manipulate the number of species (richness) in a community to investigate how varying species diversity affects a community level response (the ecosystem function). It is common to keep the relative abundances of species equal across the richness levels (Roscher et al. 2005) although some studies also manipulated the species' relative abundances at given levels of richness (Kirwan et al. 2007, Wilsey and Stirling 2007). There is a long-standing history in BEF research to focus on species richness as the driver of ecosystem functions (Byrnes et al. 2014, Spehn et al. 2005), but evenness (how equally distributed the species' relative abundances are) may also be strongly influential (Kirwan et al. 2007). While it is common to find a positive and saturating relationship between ecosystem function and richness (Scherber et al. 2010, Mittelbach et al. 2001, Schwartz et al. 2000), variation around the line may be attributed to the identities and the relative abundances of the species in the community. Diversity-Interactions (DI) models (Kirwan et al. 2007, Kirwan et al. 2009) were developed to model the data from BEF experiments in which species diversity (richness and / or evenness) is manipulated, with the species proportions (simplex space) and their interactions used to model ecosystem functions (Cornell 2002). These models account for variation attributed to species' identities, species' relative abundances, interactions among species and evenness, *in addition to* species richness. If the experimental design provides good coverage around the simplex space, DI models can be used to predict for any community combination of relative abundances within the species pool, not just the exact communities that were included in the design, providing a major advantage over other modelling approaches used in BEF research. Since their original development (Kirwan et al. 2007, Kirwan et al. 2009), the family of DI models has grown to facilitate many of the complexities that arise with data from biodiversity experiments, such as multivariate responses (Dooley et al. 2015), non-linearity in the form of species interactions (Connolly et al. 2013), the modelling of a large numbers of species interactions in a single year (Brophy et al. 2017) and the modelling of interactions among phylogenetically diverse communities (Connolly et al. 2011), where phylogenetic diversity is a measure of how ancestrally related the species in the pool are. The approach has been applied to biodiversity experimental data from a wide range of ecosystem types including bacterial communities (Brophy et al. 2017, Piovia-Scott et al. 2017), plot-based grassland experiments (Frankow-Lindberg 2012, Hoekstra et al. 2015), dung fauna diversity studies (O'Hea, Kirwan, and Finn 2010) and tree species diversity experiments (De Groote et al. 2017, Sercu et al. 2017).

DI models generally take the form (Kirwan et al. 2009):

$$y = ID + DE + \varepsilon \qquad (1)$$

The community-level response y is some ecosystem function such as sown plant biomass in a grassland community or average daily respiration rate of bacterial communities. DI models use linear (mixed) models to estimate the identity (ID) and diversity effects (DE) of species on the ecosystem function. Identity effects measure the effects of individual species and can include block or treatment effects. Diversity effects are the additional effects due to the interaction between species in mixtures, and can take many forms. An example of a DI model is:

$$y_m = \sum_{i=1}^{s} \beta_i P_{im} + \sum_{\substack{i,j=1 \\ i<j}}^{s} \delta_{ij} P_{im} P_{jm} + \varepsilon_m \qquad (2)$$

where P_{im} is the proportion of species i in experimental unit m, s is the total number of species in the species pool, and $\varepsilon_m \sim N(0, \sigma^2)$. In monoculture (species i is the only species present), $P_i = 1$ and all other proportions are 0 and the expected response is $E[Y] = \beta_i$. The expected interaction between species i and j is δ_{ij}. $E[Y]$ in a mixture is a weighted sum of the monoculture performances of each of the species (ID) plus the combined pairwise interactions (DE). Model (2) is the full pairwise interactions model and it accounts for all pairwise interactions between species and their potential effects on the response of interest. If there is a small number of species in the species pool, model (2) is a reasonable model to fit. However, when the species pool is large, model (2) is generally not of interest: the number of pairwise interactions is either too large to be biologically informative, or it may not be possible to estimate all pairwise interactions due to the experimental design. For example, in a four-species system there are $\binom{4}{2}$ = 6 pairwise interactions, while in a sixteen-species system there are $\binom{16}{2}$ = 120 pairwise interactions. However, there are many biologically informative ways to simplify model (2) (Table 4.1); for example, there may be no diversity effect (the identity model), or it may be assumed that all of the pairwise interactions are equal (average pairwise model), or constraints among interactions may be introduced according to biological functional groupings (functional group model).

Table 4.1. Descriptions of the diversity effect (DE) in Diversity-Interactions models, for a pool of s species.

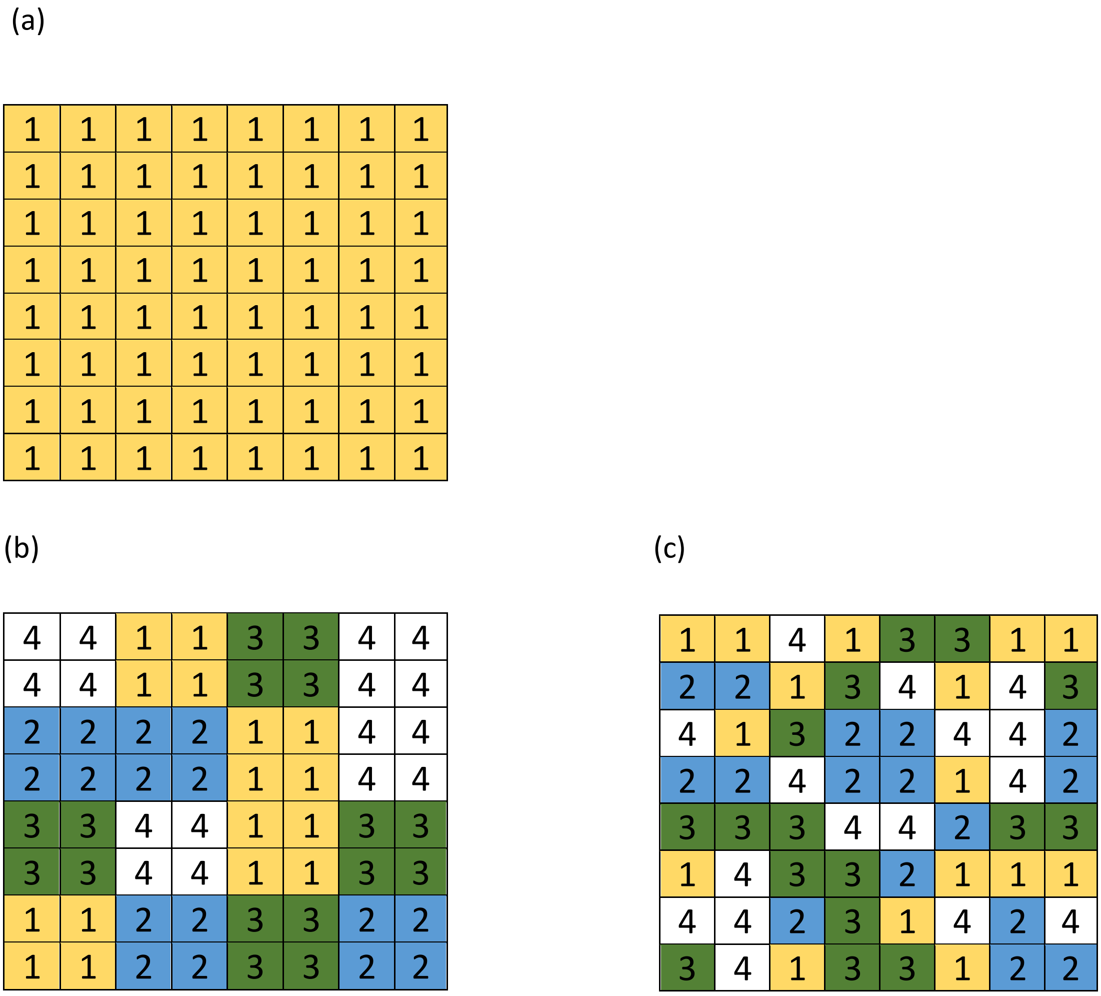
Model name	Model description	DE
Identity	No diversity effect.	
Average pairwise	All pairwise interactions are equal.	$\delta \sum_{\substack{i,j=1\\ i<j}}^{s} P_i P_j$
Functional group (FG)	Assume T functional groups (FG_1-FG_T), each with n_t species, where $t = 1, \dots, T$ and $\sum_{t=1}^{T} n_t = s$. The species in each functional group are numbered: $FG_1 = \{1,...,n_1\}$, $FG_2 = \{n_1+1,...,n_1+n_2\}$, … , $FG_T = \{1+\sum_{t=1}^{T-1} n_t,..., s\}$. The parameter ω_{qq} is the interaction between two species from functional group q; and ω_{qr} is the interaction between two species from different functional groups, i.e. where $q \neq r$.	$\sum_{q=1}^{T} \omega_{qq} \sum_{\substack{i,j\in FG_q\\ i<j}} P_i P_j + \sum_{\substack{q,r=1\\ q<r}}^{T} \omega_{qr} \sum_{i\in FG_q} \sum_{j\in FG_r} P_i P_j$
Additive species	The contribution of species i to a pairwise interaction (λ_i) is an additive constant, regardless of the species it is interacting with. The strength of the pairwise interaction between two species is the sum of the individual contributions of each species.	$\sum_{\substack{i,j=1\\ i<j}}^{s} (\lambda_i + \lambda_j) P_i P_j$
Full pairwise	All pairs of species interact uniquely.	$\sum_{\substack{i,j=1\\ i<j}}^{s} \delta_{ij} P_i P_j$

4.2.2 Experiment

The data used in this study was collected from the Species Pattern and Community Ecology (SPaCE) experiment at the University of North Dakota's Mekinock Field Station (Mekinock, ND, USA) from 2012 to 2014, inclusive. There were 170 plots of size 1 × 1 m in a randomized block design consisting of 5 blocks. Each plot was divided evenly into an 8 × 8 grid where each of the 64 cells was planted with a 16-week-old greenhouse grown transplant at the beginning of the growing season in 2012. Sixteen tallgrass prairie species were planted in the plots with either 1 (monoculture), 2, 4 or 8 (mixtures) species in each, with varying relative abundances at each richness level in mixtures, where richness is the number of species in a plot. A spatial treatment with two levels was manipulated across plots with more than one species (mixture plots): each individual was either planted randomly in the 8 × 8 grid (dispersed) or grouped with other individuals of the same species in 2 × 2 squares within the grid (aggregated) (Figure 4.1). The yearly aboveground plant biomass yield (g) was recorded in each plot. Non-focal species were removed monthly by hand and the weed biomass in each plot was collected, dried and weighed. Season total weed biomass removed (g) from each plot in each growing season was the response of interest in this analysis. The proportion of biomass of the planted species in each plot was recorded at the end of every growing season, providing 'realized' proportions in each year.

Plants with similar traits can be classified by their functional group (FG). In this experiment, there were four species from each of four functional groups: warm-season grasses, species 1 to 4: *Andropogon gerardii* (AG; big bluestem), *Schizachyrium scoparium* (SS; little bluestem), *Sorghastrum nutans* (SN; Indian grass), and *Panicum virgatum* (PV; switchgrass); cool-season grasses, species 5 to 8:

Figure 4.1: Hypothetical illustration of the spatial arrangement of experimental plots. Each plot was 1 m^2 and was split into a grid of 8 by 8 cells at planting. A single species was allocated to each of the 64 cells. Shown are examples of (a) a monoculture plot of species 1, (b) an aggregated plot with richness of 4 with species 1 to 4, and (c) a dispersed plot with richness of 4 with species 1 to 4.



(a)

1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1

(b)

4	4	1	1	3	3	4	4
4	4	1	1	3	3	4	4
2	2	2	2	1	1	4	4
2	2	2	2	1	1	4	4
3	3	4	4	1	1	3	3
3	3	4	4	1	1	3	3
1	1	2	2	3	3	2	2
1	1	2	2	3	3	2	2

(c)

1	1	4	1	3	3	1	1
2	2	1	3	4	1	4	3
4	1	3	2	2	4	4	2
2	2	4	2	2	1	4	2
3	3	3	4	4	2	3	3
1	4	3	3	2	1	1	1
4	4	2	3	1	4	2	4
3	4	1	3	3	1	2	2

Elymus canadensis (EC; Canada wildrye), *Elymus trachycaulus* (ET; slender wheatgrass), *Pascopyrum smithii* (PS; western wheatgrass), and *Nassella viridula* (NV; green needle grass); forbs, species 9 to 12: *Monarda fistulosa* (MF; wild bergamot), *Solidago rigida* (SR; stiff goldenrod), *Helianthus maximiliani* (HM;

maximilian sunflower), and *Ratibida columnifera* (RC; yellow coneflower) and legumes, species 13 to 16: *Desmodium canadense* (DC: showy tick trefoil), *Astragalus Canadensis* (AC; Canada milkvetch), *Dalea purpurea* (DP; purple prairie clover), and *Glycyrrhiza lepidota* (GL; American licorice). Two-species plots contained a grass and either a forb or a legume, four-species plots contained one species from each FG and eight-species plots contained two species from each FG. The species were randomly selected from the FGs for each plot according to these constraints.

4.2.3 Statistical innovations

In this paper, we develop a Diversity-Interactions (DI) model for the SPaCE data. While we build on previous versions of DI models, the statistically innovative aspects include: the testing of a treatment (spatial pattern) that only applies to mixture (and not monoculture) communities, the modelling of inhomogeneous variation across communities within each year and the inclusion of novel random effects to facilitate modelling pairwise interactions when there are a high number of species pairwise interactions over multiple years. We also present a simulation study to test the limits of DI models in a species rich setting to identify conditions when the random effects approach to modelling pairwise interactions will break down.

4.2.4 Description of new methods

In the SPaCE data, there are 16 species and $\binom{16}{2}$ = 120 pairwise interactions.

Estimating all 120 pairwise interactions is not of interest here since their large number would be devoid of biological meaning, and not possible given the design of the experiment (there is partial confounding among pairwise interactions). We aim to describe the diversity effect (DE) using fixed effects as parsimoniously as possible, and to test the inclusion of random pairwise interactions to identify if any variation due to pairwise interactions remains unexplained. In this section, we assume that all pairwise interactions are equal (Table 4.1, average pairwise model; a suitable starting-point candidate model), while in sections 4.3 and 4.4 we describe the full model fitting process and final choice of model, respectively. The model for the SPaCE data (3 years of repeated measurements on 170 plots) can be written generally as:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Zu} + \boldsymbol{\varepsilon} \qquad (3)$$

The predictors in the $\boldsymbol{X}$ matrix include block effects, species proportions (P_i), and the sum of all pairwise interactions, and interactions with other explanatory variables. The error term $\boldsymbol{\varepsilon} \sim N(\boldsymbol{0}, \boldsymbol{R})$, where $\boldsymbol{R}$ is a (510 × 510) block diagonal matrix with 3 × 3 blocks for the repeated measurements on each plot. The blocks can be the same across all plots or can differ based on plot characteristics. The species in the SPaCE experiment range from grasses to large forbs and vary significantly in their ability to resist weed invasion and as such, a homogeneous response to varying within-plot diversity was not expected. Preliminary analyses identified that in the second and third years of the experiment, the mean and variance of weed biomass for legume monocultures (FG 4) was considerably higher than all other types of communities (Figure 4.2). Heterogeneity was also observed in year 1, but patterns in mean and variance were not consistent with functional groups. This confirms that homogeneity across all communities within each year is unlikely to be a valid assumption.

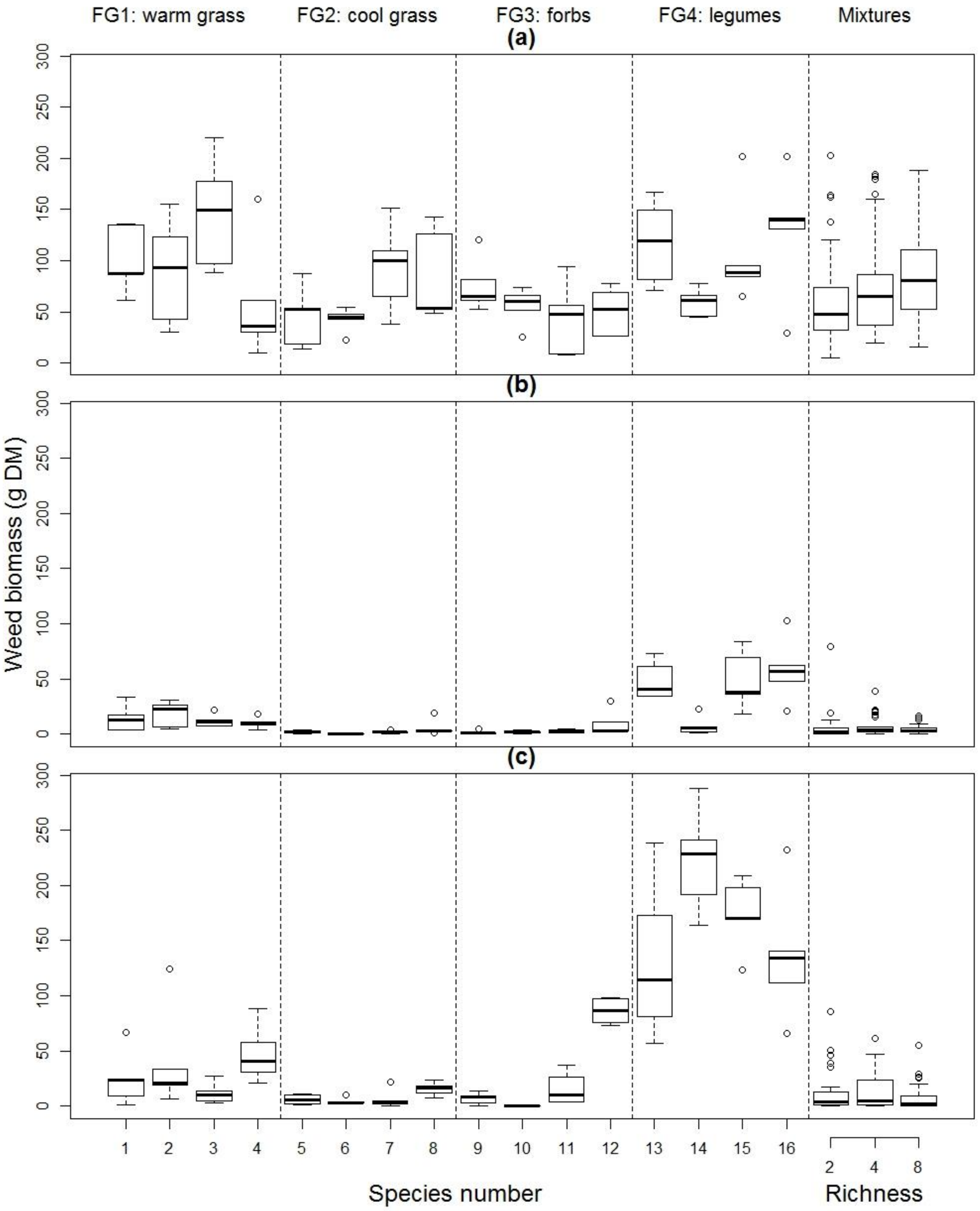


Figure 4.2. Total plot weed biomass (g DM in 1 m^2 plot) in (a) 2012, (b) 2013, (c) 2014 for each monoculture (1 to 16) and each level of richness in mixture (2, 4 or 8 species). The 4 functional groups and the mixtures are separated by the dotted lines.

For species i=1,…,15, j=2,…,16, $i < j$, plot m=1,…,170, and year k=1,2,3, the $\boldsymbol{Z}$ and $\boldsymbol{u}$ matrices are

$$[510 \times 360;\ PP_{ijmk} = P_{imk}P_{jmk}] \quad \times \quad [360 \times 1;\ d_{ijk}]$$

$$\boldsymbol{Zu} = \begin{pmatrix} PP_{1,2,1,1} & 0 & 0 & & PP_{15,16,1,1} & 0 & 0 \\ 0 & PP_{1,2,1,2} & 0 & \cdots & 0 & PP_{15,16,1,2} & 0 \\ 0 & 0 & PP_{1,2,1,3} & & 0 & 0 & PP_{15,16,1,3} \\ & \vdots & & \ddots & & \vdots & \\ PP_{1,2,170,1} & 0 & 0 & & PP_{15,16,170,1} & 0 & 0 \\ 0 & PP_{1,2,170,2} & 0 & \cdots & 0 & PP_{15,16,170,2} & 0 \\ 0 & 0 & PP_{1,2,170,3} & & 0 & 0 & PP_{15,16,170,3} \end{pmatrix} \begin{pmatrix} d_{1,2,1} \\ d_{1,2,2} \\ d_{1,2,3} \\ \vdots \\ d_{15,16,1} \\ d_{15,16,2} \\ d_{15,16,3} \end{pmatrix}$$

i.e. $\boldsymbol{Z}$ is a 510 by 360 matrix which contains the 120 pairwise interactions (P_iP_j) separated out into yearly columns (with zeros outside of the current year). It is assumed that $\boldsymbol{u} \sim N(\boldsymbol{0}, \boldsymbol{G})$, where $\boldsymbol{G} = \begin{pmatrix} \boldsymbol{M_{1,1}} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{M_{15,16}} \end{pmatrix}$ is a 360 x 360 block matrix and each $\boldsymbol{M_{ij}} = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & 0 \\ 0 & 0 & \sigma_3^2 \end{pmatrix}$ is a 3 × 3 block solely indexed by year. I.e., there are 120 random effects included in each year, but they are constrained to have equal variance, thus only one variance parameter per year is added. The purpose of this variance parameter in each year is to test if there is variability due to pairwise interactions additional to the fixed diversity effect terms: for example, for the average pairwise interaction model ($\delta_{ijk} = \delta_k$), σ_k^2 measures variation in the true δ_{ijk} around δ_k, if it exists. Indexing the random effects d_{ijk} by species pair i, j (and year k) is highly unusual, since random effects are typically indexed by a plot-level factor such as a block. It is usual that the experimental or sampling design generates observations grouped according to one or more factors (which may be crossed or nested), yielding a hierarchical modelling structure in which observations within the same group are correlated. Hence, in our modelling framework the $\boldsymbol{Z}$ matrix differs to that expected by standard mixed model fitting software, such as nlme (Pinheiro et al. 2018) and lme4 (Bates et al. 2015) in R (R Core Team 2018) for example.

The spatial pattern treatment (aggregated or dispersed, Figure 4.1) of the SPaCE experiment presents statistical challenges that have not yet been addressed within the DI modelling framework. By definition, the treatment in the SPaCE experiment is applied only to mixture plots. The spatial pattern treatment can interact with the diversity effect terms, allowing the fixed average pairwise interaction of all pairs of species to differ for aggregated and dispersed plots, *i.e.* pairs of species interact differently depending on the spatial planting pattern. This could be due to higher intraspecific interactions (between individuals of the same species) and lower interspecific interactions (between individuals of different species) in aggregated plots than dispersed plots (Murrell, Purves, and Law 2001, Stoll and Prati 2001). Incorporating the spatial pattern treatment, model (3) can be written as

$$\text{Response} = \text{ID} + \text{DE}_{\text{fixed}} + \text{DE}_{\text{random}} + \text{Error} \qquad (4)$$

$$y_{klmn} = \overbrace{\alpha_{kl} + \sum_{i=1}^{s} \beta_{ik}P_{ikm}} + \overbrace{\delta_{kn} \sum_{\substack{i,j=1 \\ i<j}}^{s} P_{ikm}P_{jkm}} + \overbrace{\sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ijk}P_{ikm}P_{jkm}} + \varepsilon_{klmn} \qquad (5)$$

where α_{kl} is block effect l in year k, P_{ikm} is the proportion of species i in plot m in year k, and $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{R})$ as in equation (3). The β_{ik} terms can be interpreted as the expected weed biomass for a monoculture of species i in year k (all $P_iP_j = 0$ in monoculture). δ_{kn} is the fixed average pairwise interaction between species i and j in year k for spatial pattern n, where n can be either 1=aggregated or 2=dispersed. Suppose that $\text{DE}_{\text{random}}$ is needed in year k, for example: the DE_{fixed} term assumes that all 120 δ_{ijkn} pairwise interactions are equal to δ_{kn}, but this may not be sufficient and including the $\text{DE}_{\text{random}}$ term acknowledges there is variation around δ_{kn} across all pairwise interactions. Extra terms may be needed in the $\mathrm{DE}_{\mathrm{fixed}}$ component (see Section 4.3) but the inclusion of random effects acknowledges there may be additional variation, unexplained by $\mathrm{DE}_{\mathrm{fixed}}$, and the random effect variance terms will be incorporated into fixed effects standard errors, improving inference.

Fitting 360 random effects with a constrained variance in each year is not a trivial coding challenge. We used SAS software (SAS Institute, Cary, North Carolina, USA) to fit our models, utilizing the LIN covariance structure in proc mixed, which allows user-defined variance-covariance matrix structures for random effects. SAS University Edition is currently freely available for academic and non-commercial use and our models can be fitted using it. It is not possible to fit the models in the standard mixed model packages in R (e.g. LME4 (Bates et al. 2015)) since random effects in these packages are always indexed by a plot level factor. However, we succeeded in fitting simple versions of a multi-year model using optimx in R (Nash and Varadhan 2011, R Core Team 2018). Appendix 4.1 provides a tutorial style guide to fitting simple versions of the models in SAS and R for comparison.

4.3 Model fitting and simulations

4.3.1 Model fitting for the SPaCE experiment

Diversity-Interactions models were fitted to the SPaCE data from all three years with total yearly weed biomass as the response. A log transformation of the yearly weed biomass was also modelled as a response to try to account for non-constant variance. One was added to each of the weed biomass values before taking the log to avoid problems with zero values. Proportions planted were used as predictors in year 1 and ‘realized’ proportions in the preceding year were used in years 2 and 3. The model selection process involved:

1. Preliminary fixed effects selection: The identity, average pairwise, functional group and additive species models (Table 4.1) were fitted by maximum likelihood (ML) to allow comparisons of different fixed effects. Likelihood ratio tests (LRTs) were conducted to select the best model.
2. Best variance-covariance structure: The fixed effects model identified in step 1 was used as a ‘baseline’ model: the fixed effects in the model stayed the same while different variance-covariance structures were fitted by restricted maximum likelihood (REML) to account for repeated measures over years: compound symmetry, first-order auto-regressive and unstructured (Littell et al. 2006). REML fitting was necessary in this case because the fixed effects were constant and the variance-covariance structures were changing. To test for heterogeneous variance across plots, the $\boldsymbol{A_i}$ matrices within the $\boldsymbol{R}$ matrix were first fitted as constant across all 170 plots, and then allowed to differ according to: (a) monocultures and mixtures, (b) all FGs in monocultures, and mixtures. These models were also compared using LRTs.
3. Best fixed effects model with the best variance-covariance structure: If the variance-covariance structure was changed in step 2, the selection of fixed effects was carried out again as described in step 1, using the new error structure. In addition, the diversity effect was tested for interaction with spatial pattern and with legume percentage. Models were fitted by ML and compared using LRTs.
4. Include random pairwise interactions: The best fixed effects model with the best variance-covariance structure was fitted using REML (because the fixed

effects did not change) with and without random pairwise interactions in each year individually and compared using LRTs. The purpose of including random pairwise interactions in a given year is twofold: 1) if they are not needed, this indicates no evidence of lack of fit in the diversity effect explanation, 2) if they are needed, the extra variance parameter acknowledges that there is additional variability due to pairwise interactions and incorporates this extra uncertainty into standard errors, without the need for many additional fixed terms (Brophy et al. 2017). To counteract the boundary space problem (Self and Liang 1987, Shapiro 1988) when testing the inclusion of random effects in each year, the P-values of the LRTs were halved (Littell et al. 2006).

4.3.2 Simulation studies for model evaluation

We conducted simulation studies to evaluate the performance of the random effects approach to modelling pairwise interactions over multiple years. An experimental design was generated with $m = 110$ plots, a pool of ten species categorized according to two functional groups (each of size five), and plots of species richness 1, 2, 5, 9 and 10. Each species was in monoculture in two plots; each pair of species appeared once in a plot of richness two in equal proportion; and there were 30, 10 and 5 plots of richness 5, 9 and 10 respectively, with all species in equal proportion in each plot. Response values were simulated assuming a functional group model (Table 4.1) over $k = 3$ experimental years. In addition to functional group interaction effects, extra variation due to pairwise interactions was included in the simulated responses. For year k, plot m, species i, j, and $FG_1 = \{1,2,3,4,5\}$ and $FG_2 = \{6,7,8,9,10\}$, the simulated model was

$$y_{km} = \sum_{i=1}^{10} \beta_{ik} P_{ikm} + \sum_{q=1}^{2} \omega_{qqk} \sum_{\substack{i,j \in FG_q \\ i<j}} P_{ikm} P_{jkm} + \omega_{12k} \sum_{i \in FG_1} \sum_{j \in FG_2} P_{ikm} P_{jkm} + \sum_{\substack{i,j=1 \\ i<j}}^{10} d_{ijk} P_{ikm} P_{jkm} + \varepsilon_{klm} \qquad (6)$$

where the identity effects values and the within and between functional group interaction values in each year are given in Table 4.2. The random effects were assumed to be normally distributed in each year; the standard deviation values simulated (assumed the same in each year) were 0, 500, 1000, 1500, 2000 and 2500. The residual errors, also assumed normally distributed, were simulated assuming homogeneity across plots within each year, and for simplicity, assuming zero covariance across years; the simulated standard deviation values for each year were 100, 200 and 300. Each combination of random effects variance by error variance values gave rise to 18 sets of simulations, each of which contained 1000 datasets. The code to perform the simulation study is provided in Appendix 4.2. The experimental design dataset is provided in the supporting electronic file EF4: Experimental_design.csv.

Table 4.2. Values of the identity effects and the within and between functional group interactions in each year from the simulation study for model evaluation.

	Year		
Effect	1	2	3
β_1	700	1477	1103
β_2	472	1458	1411
β_3	512	1636	1601
β_4	650	1267	604
β_5	440	879	753
β_6	343	759	792
β_7	630	931	627
β_8	709	874	637
β_9	699	1100	752
β_{10}	584	845	748
ω_{11}	-2400	-2345	4592
ω_{22}	1753	-2151	-3954
ω_{12}	841	-128	524

4.4 Results

4.4.1 Data overview

Weed biomass was similar for all levels of species richness in 2012 (Figure 4.3a). In 2013 and 2014, monocultures showed both a higher median and range in weed biomass than mixtures (Figure 4.3b,c), suggesting that diversity suppressed weed invasion over time. Weed biomass in 2013 and 2014 did not vary much among richness levels in mixture plots, although the values appeared to be lower at all richness levels in 2013 than in any other year (Figure 4.3). This was most likely because 2012 was an establishment year and biomass production of most planted species was greatest in 2013 (Appendix 4.3), leaving more space, light and nutrients for weeds to grow in (McKenna and Yurkonis 2016). The medians and the variances of the weed biomass across the 16 monocultures varied in all years (Figure 4.2). Some species from FGs 2 and 3 (cool-season grasses and forbs) had lower medians in monoculture than the mixtures in all years. For example, all five monocultures of *Solidago rigida* (species 10) (SR) completely suppressed weeds in 2014. This was most likely due to the dense basal rosette form of this species making it difficult for invaders to grow. *S. rigida* monocultures also had some of the highest biomass yields in 2014 (Appendix 4.3). Legume (FG4) monocultures had high weed biomass values in each year (Figure 4.2). Mixtures containing legumes were usually far superior at resisting weed invasion compared to their monocultures (e.g. Figure 4.2) and thus legume proportion in the mixture may be strongly influential on the diversity effect.

4.4.2 Model selection

To find the best model, the steps outlined in Section 4.3.1 were followed. The additive species model (Table 4.1) was the best model identified in step 1 (which assumed homogeneous variance across plots within each year). In step 2, the best variance-covariance structure had unstructured blocks in the ***R*** matrix (i.e. a unique variance for each year and a unique covariance for each pair of years) with different blocks for each of the four functional groups in monoculture, and another for mixtures (Appendix 4.4). In step 3, the best fixed effects model identified (using the new variance-covariance structure) was the functional group model where spatial planting pattern interacted with the within functional group and between functional group interactions, and proportion of FG4 (legumes) interacted with additive species pairwise interactions (Table 4.3). Models with more parameters than model 4 in Table 4.3 that were not nested with model 4 were compared using AIC values, but were not preferable. Additional random pairwise interactions were needed in 2014 only (Table 4.3).

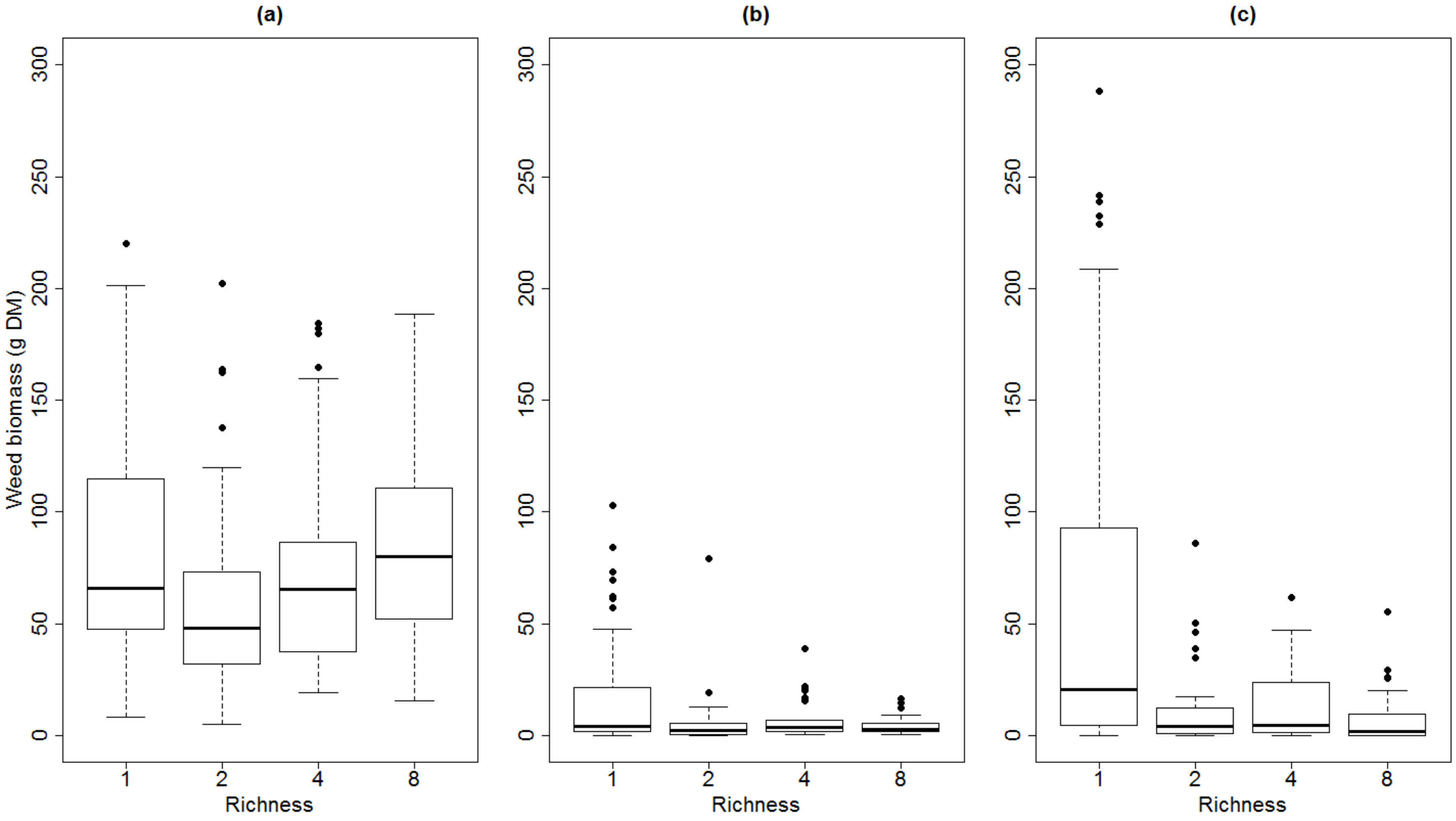


Figure 4.3. Total plot weed biomass (g DM in 1 m^2 plot) in (a) 2012, (b) 2013, and (c) 2014 by number of species in plot (richness).

Table 4.3. Details of tests conducted to determine the best linear predictor for the fixed effects (step 3) and tests for the inclusion of random pairwise interactions (step 4). Likelihood ratio test statistics (LRT), degrees of freedom (df = difference between the compared models) and P-values are shown. Model 1 includes block effects, identity effects for each species, and an average pairwise interaction term. Models 1-3 were fitted using maximum likelihood (ML). Model 4 was fitted by ML for comparisons of fixed effects and REML for comparisons of random effects. Models 5-7 were fitted using restricted maximum likelihood (REML). The P-values for the random pairwise interactions variance components are halved due to the boundary space issue.

Model number	Model description	Test	LRT	df	P-value
	Fixed effects model comparisons				
1	Average pairwise model				
2	Functional group model	1 v 2	62.8	27	<0.001
3	With functional group interactions * Spatial pattern	2 v 3	52.1	30	0.007
4	With additive species interactions * legume %	3 v 4	210.9	48	<0.001
	Random effects model comparisons				
5	With random pairwise interactions in 2012	4 v 5	0	1	0.500
6	With random pairwise interactions in 2013	4 v 6	0.4	1	0.264
7	With random pairwise interactions in 2014	4 v 7	5.6	1	0.009

For year k, block l, plot m, spatial pattern n, species i, j, and the warm-season grasses FG$_1$ = {1,2,3,4}, the cool-season grasses FG$_2$ = {5,6,7,8}, the forbs FG$_3$ = {9,10,11,12} and the legumes FG$_4$ = {13,14,15,16}, the final chosen model was

$$
\begin{aligned}
y_{klmn} &= \alpha_{kl} + \sum_{i=1}^{s} \beta_{ik} P_{ikm} && \text{ID (including block effects)} \\
&+ \sum_{q=1}^{4} \omega_{qqkn} \sum_{\substack{i,j \in FG_q \\ i<j}} P_{ikm} P_{jkm} && \text{DE}_{\text{fixed}}\text{(within functional groups by spatial pattern)} \\
&+ \sum_{\substack{q,r=1 \\ q<r}}^{4} \omega_{qrkn} \sum_{i \in FG_q} \sum_{j \in FG_r} P_{ikm} P_{jkm} && \text{DE}_{\text{fixed}}\text{(between functional group by spatial pattern)} \\
&+ \sum_{\substack{i,j=1 \\ i<j}}^{s} (\lambda_{ik} + \lambda_{jk}) P_{ikm} P_{jkm} \sum_{i=13}^{16} P_{ikm} && \text{DE}_{\text{fixed}}\text{(legume interactions)} \\
&+ \sum_{\substack{i,j=1 \\ i<j}}^{s} d_{ij3} P_{i3m} P_{j3m} && \text{DE}_{\text{random}}\text{(year } k = 3 \text{ only)} \\
&+ \varepsilon_{klmn} && (7)
\end{aligned}
$$

where $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{R})$. $\boldsymbol{R}$ contains different blocks for each functional group monoculture, and another for mixtures (i.e. five repeated measures variance-covariance blocks across the 170 plots), $d_{ij3} \sim N(0, \sigma_3^2)$ independent of $\boldsymbol{\varepsilon}$. The parameter α_{kl} is the effect of block l in year k, β_{ik} is the expected weed biomass of a monoculture of species i in year k, and λ_{ik} and λ_{jk} are the fixed additive species interaction effects of species' i and j in year k. $\sum_{i=13}^{16} P_{ikm}$ is the total legume proportion for year k, plot m. ω_{qqkn} is the coefficient of the total pairwise interactions between species from functional group q for spatial pattern n in year k. ω_{qrkn} is the coefficient of the total pairwise interactions between species from functional group q and r for spatial pattern n in year k. The DE$_{\text{fixed}}$ (functional group) terms mean that a pair of species interacts in the same way as any pair of species from the same functional group(s). However, these interactions changed depending on whether the spatial planting pattern was aggregated or dispersed (n). The remainder of $\mathrm{DE}_{\mathrm{fixed}}$ is the legume interaction terms; these interaction terms indicate that in addition to functional group effects, there is a species-specific effect that contributes towards the DE and depends on legumes percentage in the mixture. The inclusion of $\mathrm{DE}_{\mathrm{random}}$ in the final model means that in the third year of the experiment, there was additional variability due to the individual pairwise interactions that was not picked up by $\mathrm{DE}_{\mathrm{fixed}}$, and was incorporated into the fixed effects standard errors. The estimated variance of the random pairwise interactions ($\hat{\sigma}_3^2$) was 3224.3. The estimates of the fixed effects of the final model are in Appendix 4.5 and the variance components estimates are in Appendix 4.6. A graphical assessment of model assumptions is included in Appendix 4.7.

Figure 4.2 shows that the weed biomass medians across the 16 monocultures varied in all years. The monoculture estimates in 2012 were all significantly higher than those in 2013 at the average block level (P<0.001 in most cases; all P<0.02), and in 2014 except for species 4, 11, 13 and 16 (*Panicum virgatum* (PV), *Helianthus maximiliani* (HM), *Desmodium canadense* (DC), and *Glycyrrhiza lepidota* (GL) respectively). Two-species community predictions with both species in equal proportions (centroid communities) at the average block level were conducted for all communities which contained either one warm-season or cool-season grass species and either one forb or legume species. In 2012, the aggregated planting pattern generally yielded lower predictions than dispersed, except for communities mixing a species from FGs 2 and 4 (Figure 4.4). This could be due to aggregation allowing subordinate species to establish more easily due to decreased interspecific

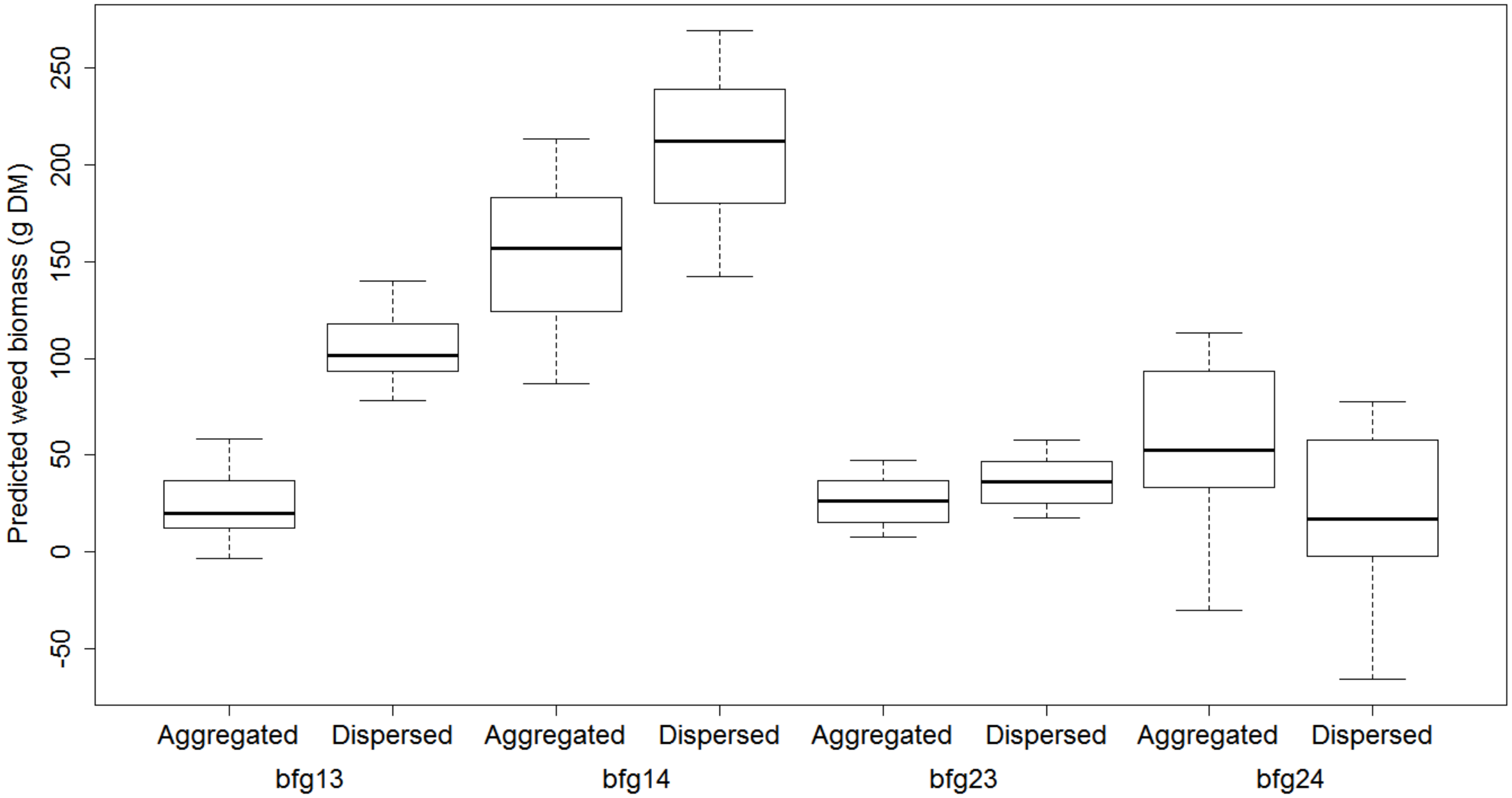


Figure 4.4. Predicted weed biomass (g DM in 1 m^2 plot) in 2012 in two-species centroid communities containing one species from either FG1 or FG2 and one species from either FG3 or FG4 split by spatial pattern (either aggregated or dispersed).

interactions, leaving the community less susceptible to invasion at establishment. To examine the effect of altering legume percentage, four-species predictions containing one species from each functional group were calculated from the final model at the average block level, with legume percentage taking six values between 0 and 0.25 to stay within the bounds of the experimental values. The remaining proportion was divided between the other three chosen species. In 2012, the predicted weed biomass changed more with legume percentage in dispersed plots than in aggregated plots for the chosen communities: the predictions for communities with no legume present were much lower for dispersed, but at the centroid community ($\sum_{i=13}^{16} P_i$=0.25) the predictions were similar for both spatial planting patterns, with dispersed still slightly lower (Appendix 4.8a). In 2013 and 2014, the predictions were similar for both spatial patterns in the communities with no legumes present, and as the legume percentage increased, the predicted weed biomass changed similarly with legume percentage with the aggregated community predictions becoming higher than the corresponding dispersed predictions (Appendix 4.8 b,c). This could be due to the aggregation of legumes leaving niche space for invaders to grow. The orientation and strength of the relationship between legume percentage and predicted weed biomass depended on species present, spatial pattern and year. Some predictions were negative, but this was to be expected since there were a number of weed biomass observations close to and equal to zero in the SPaCE data. These negative predictions can be thought of as predicting that weeds were completely suppressed.

4.4.3 Log transformation model selection

The procedure in Section 4.3.1 was followed for the log transformed data. As for the untransformed data, the additive species model was the best model in step 1 (which assumed homogeneous variance across plots within each year). The best ***R*** matrix with the additive species model had a single unstructured block that was the same for each plot (i.e. a unique variance for each year and a unique covariance for each pair of years). In step 3, the spatial planting pattern was accounted for in the model along with the fixed effects models from step 1. The chosen model was the additive species model in which spatial planting pattern interacted with the species-specific interactions, and proportion of FG4 (legumes) interacted with functional group pairwise interactions. Random pairwise interactions in any year did not improve the best model. The best model for the log-transformed data is in Appendix 4.9, and a graphical assessment of the model assumptions is in Appendix 4.10.

4.4.4 Simulation results

In our simulation study, the ability of the modelling approach to detect the random pairwise interactions was assessed across various combinations of random effect variance values by residual error variance. When the random pairwise interactions standard deviation was set to 0, DI models found significant variation due to random pairwise interactions in one of the years on just a small number of occasions (Table 4.4). However, many of the models including random pairwise interactions were stopped due to infinite likelihood when no random pairwise interactions were present, or their variance was small compared to residual variance (Table 4.4). Diversity Interactions modelling methods detected random pairwise interactions at a higher rate when the variance of the random pairwise interactions was large in comparison to the residual variance (Table 4.4). To be distinguishable from residual

Table 4.4. The number of simulations from 1000 that gave 0, 1, 2 or 3 significant yearly random pairwise interactions at the 0.05 significance level, or number of times the model was stopped because of too many likelihood evaluations (NA) in each of three years for 18 different combinations of residual standard deviation by random pairwise interactions standard deviation.

		Number of significant yearly random pairwise interactions				
Residual standard deviation	Random pairwise interactions standard deviation	0	1	2	3	NA
100	0	555	29	1	0	415
100	500	5	83	379	533	0
100	1000	0	0	0	1000	0
100	1500	0	0	0	1000	0
100	2000	0	0	0	1000	0
100	2500	0	0	0	1000	0
200	0	554	31	1	0	414
200	500	600	284	51	7	58
200	1000	6	74	369	551	0
200	1500	0	0	9	991	0
200	2000	0	0	0	1000	0
200	2500	0	0	0	1000	0
300	0	600	19	0	0	381
300	500	720	102	9	0	169
300	1000	308	417	234	32	9
300	1500	8	67	393	531	1
300	2000	0	2	44	954	0
300	2500	0	0	4	996	0

variance, random pairwise interactions must usually have sufficiently large variance in comparison to the residual variance.

4.5 Discussion and conclusions

Diversity-Interactions models are a valuable tool in BEF research: they can simultaneously assess the impact of individual species, species interactions, species' proportions and richness on ecosystem function (Kirwan et al. 2009). The ability to predict around the simplex space of the species pool provided a superior level of information compared to ANOVA or richness only based approaches to estimating BEF relationships. We identified varying abilities of species to resist weed invasion, with legumes in particular proving very susceptible to invasion in monoculture. We also found that diversity effects were driven by functional group membership and that the functional group effects varied according to species spatial pattern. Thus, the spatial pattern effects depended on which species were in the mixture and could not be generalized to any mixture. Examining the raw data, when plots were grouped by spatial pattern within year, no apparent effect of spatial pattern treatment was observed in any of the years, except for a slightly inferior weed suppression in aggregated plots in 2013 and 2014 (Appendix 4.11), highlighting the benefit of our approach over traditional ANOVA based methods for analyzing this type of data. Previous studies have found that aggregation of plant species helped maintain diversity and reduce weed invasion by allowing less competitive species to persist (Wassmuth et al. 2009). However, model prediction here suggested that aggregated plots were more prone to invasion outside the year of establishment which is consistent with Yurkonis, Wilson and Moloney (2012). We also identified a species-specific effect of the proportion of legume in mixtures on the diversity effects. Model predictions illustrated this, with predicted weed biomass in dispersed spatial pattern communities in particular changing with increasing proportions of legumes. The results supported the argument that increasing the number of species with different functional traits has a positive impact on weed suppression (Hector et al. 2001, Pokorny et al. 2005).

The inclusion of the unusual random effects (indexed by pairwise species rather than by a plot-level factor) provides a novel and parsimonious way to model a large number of pairwise interactions over multiple years, extending the methods from Brophy et al. (2017) for fitting the random effects in a single year. As shown by the simulation studies, the ability of the modelling approach to pick up the random effects is determined by the relative size of the random effects variance to the residual error variance, but when the random effects variance is sufficiently (relatively) large, the model performs well at detecting the need for the additional variance components. Our simulation studies also showed that when the random pairwise variance does not exist, the model will rarely detect that is required, but it may encounter convergence issues. Thus, when there were problems including random pairwise interactions in a model that otherwise has no convergence issues, we interpreted it as a lack of need for the variance component. The analysis in Connolly et al. (2013) compared DI models to a reference model including a coefficient for each distinct community and blocking structure or treatment (where there was replication at this level). DI models fitted similarly to the reference model in six experiments out of seven, meaning that pairwise interactions were sufficient to describe the data structure. A similar analysis could not be done on the SPaCE dataset because individual community types were not replicated (except in monoculture).

This work provides a useful extension to DI models that accounts for spatial pattern treatment in the experiment. At establishment, aggregated plots had higher mean total intraspecific interactions than dispersed plots at each richness level. The difference became larger with increasing richness (Appendix 4.12). Future work could focus on making bigger same-species clusters in the aggregated plots to increase the intraspecific interactions and make them more different from the dispersed plots. To ensure the dispersed plots lead to genuinely different groupings of planted species, a ‘dispersed control’ treatment could be added. This would include plots in which the dispersed plots are manipulated to reflect varying levels of interspecific interactions. Not only will the inclusion of spatial pattern in DI models benefit the modelling of grassland plot experiments, DI models are being used to model the interactions between species in forests (De Groote et al. 2017). The spatial pattern in forests could play a very important role in ecosystem functioning due to light interception, and the ability of species to interact with distant species could be diminished on a larger scale than in the SPaCE experiment.

A log transformation of weed biomass is a possible alternative method of ensuring that the linear mixed model, as proposed here, provides a good fit to the data. The analysis conducted on the log transformed data (with one added to each weed biomass value to avoid zero values), satisfied linear mixed model assumptions (Appendix 4.10). Similar to the model for the untransformed data, the best DI model for the log transformed data detected the functional grouping structure and the legume effects present in the data (Appendix 4.9). The model on the log-scale had a simpler R matrix structure, because the log transformation resolved the non-constant variance issues evident on the original scale. However, a log transformation makes the interpretation of individual coefficients less straightforward. For example, on the untransformed data, the estimated coefficients of the species 1 identity effects are the estimated weed biomass values for species 1 monocultures in each year. For the log transformed data, a back-transformation of the model identity coefficients gives the estimated median weed biomass value instead of the mean, which is still directly interpretable. However, for the log transformed data, the diversity effect terms are more complex to understand since when back-transformed, the effects are multiplicative rather than additive. Thus, when modelling on the original scale of the data, it is easier to attribute specific parameter estimates to the underlying biological processes than when working on a transformed scale.

The SPaCE experimental design in this study partially confounded species richness and functional group richness: all two-species plots contained two functional groups (always a warm-season or cool-season grass with a legume or a forb), and all four-species and all eight-species plots contained four functional groups. The design was suitable for the aim of the study to investigate the interactions among different functional groups as they relate to a reconstructed grassland. However, there are implications of the restricted design space for Diversity-Interactions modelling predictions: for our estimated model, it was not reasonable to predict for richness=2 plots of species from the same functional group, or for two grasses (warm and cool season) together, or a legume and a forb together. Generally, if it is desirable to predict for any combination of any species at any richness or evenness levels, care must be taken to have adequate representation around the simplex space when designing the experiment.

Many of the species in this reconstructed grassland system did not persist very well, as is common in reconstructed grasslands (Martin, Moloney, and Wilsey 2005). For example, the highest proportion of *Glycyrrhiza lepidota* (species 16) (GL) in any of the plots in 2014 was 0.12, from a species-high of 0.5 in 2012. Because of this, it was not reasonable to predict at high values of species 16 in 2014. Similar declines occurred for some of the other species. In grassland systems, producers can manage their grasslands to help legumes persist, for example by lowering soil nitrogen, and through alternate harvesting for silage between and within years (Humphreys et al. 2017). However, this is not possible in the restored grassland system used in the SPaCE experiment, in which the only management was monthly weeding and annual harvesting. In such systems, the initial management of the experiment, for example the seed density composition and site age, can be influential (Grman, Bassett, and Brudvig 2013). Varying the establishment times of different species can also affect diversity and species persistence (Martin and Wilsey 2012).

References

Bates, D, M Machler, BM Bolker, and SC Walker. 2015. "Fitting linear mixed-effects models using lme4." *Journal of Statistical Software* no. 67 (1):1-48.

Brophy, C, A Dooley, L Kirwan, JA Finn, J McDonnell, T Bell, MW Cadotte, and J Connolly. 2017. "Biodiversity and ecosystem function: making sense of numerous species interactions in multi-species communities." *Ecology* no. 98 (7):1771-1778. doi: 10.1002/ecy.1872.

Byrnes, JEK, L Gamfeldt, F Isbell, JS Lefcheck, JN Griffin, A Hector, BJ Cardinale, DU Hooper, LE Dee, and JE Duffy. 2014. "Investigating the relationship between biodiversity and ecosystem multifunctionality: challenges and solutions." *Methods in Ecology and Evolution* no. 5 (2):111-124. doi: 10.1111/2041-210X.12143.

Cardinale, BJ, KL Matulich, DU Hooper, JE Byrnes, E Duffy, L Gamfeldt, P Balvanera, MI O'Connor, and A Gonzalez. 2011. "The functional role of producer diversity in ecosystems." *American Journal of Botany* no. 98 (3):572-592. doi: 10.3732/ajb.1000364.

Connolly, J, T Bell, T Bolger, C Brophy, T Carnus, JA Finn, L Kirwan, F Isbell, J Levine, A Luscher, V Picasso, C Roscher, MT Sebastia, M Suter, and A Weigelt. 2013. "An improved model to predict the effects of changing biodiversity levels on ecosystem function." *Journal of Ecology* no. 101 (2):344-355. doi: 10.1111/1365-2745.12052.

Connolly, J, MW Cadotte, C Brophy, A Dooley, J Finn, L Kirwan, C Roscher, and A Weigelt. 2011. "Phylogenetically diverse grasslands are associated with pairwise interspecific processes that increase biomass." *Ecology* no. 92 (7):1385-1392. doi: 10.1890/10-2270.1.

Connolly, J, MT Sebastia, L Kirwan, JA Finn, R Llurba, M Suter, RP Collins, C Porqueddu, A Helgadottir, OH Baadshaug, G Belanger, A Black, C Brophy, J Cop, S Dalmannsdottir, I Delgado, A Elgersma, M Fothergill, BE Frankow-Lindberg, A Ghesquiere, P Golinski, P Grieu, AM Gustavsson, M Hoglind, O Huguenin-Elie, M Jorgensen, Z Kadziuliene, T Lunnan, P Nykanen-Kurki, A Ribas, F Taube, U Thumm, A De Vliegher, and A Luscher. 2018. "Weed suppression greatly increased by plant diversity in intensively managed grasslands: A continental-scale experiment." *Journal of Applied Ecology* no. 55 (2):852-862. doi: 10.1111/1365-2664.12991.

Corbin, JD, and CM D'Antonio. 2012. "Gone but not forgotten? Invasive plants' legacies on community and ecosystem properties." *Invasive Plant Science and Management* no. 5 (1):117-124. doi: 10.1614/IPSM-D-11-00005.1.

Cornell, John A. 2002. *Experiments with mixtures: designs, models, and the analysis of mixture data*. 3rd ed, *Wiley series in probability and statistics*. New York: Wiley.

De Groote, SRE, IMV Lantman, BK Sercu, D Dekeukeleire, R Boonyaritthichaikij, HK Smith, R De Beelde, K Ceunen, P Vantieghem, H Matheve, L De Neve, M Vanhellemont, L Baeten, E de la Pena, D Bonte, A Martel, K Verheyen, and L Lens. 2017. "Tree species identity outweighs the effects of tree species diversity and forest fragmentation on understorey diversity and composition." *Plant Ecology and Evolution* no. 150 (3):229-239. doi: 10.5091/plecevo.2017.1331.

DiTomaso, JM. 2000. "Invasive weeds in rangelands: Species, impacts, and management." *Weed Science* no. 48 (2):255-265. doi: 10.1614/0043-1745(2000)048[0255:IWIRSI]2.0.CO;2.

Dooley, A, F Isbell, L Kirwan, J Connolly, JA Finn, and C Brophy. 2015. "Testing the effects of diversity on ecosystem multifunctionality using a multivariate model." *Ecology Letters* no. 18 (11):1242-1251. doi: 10.1111/ele.12504.

Dukes, JS. 2002. "Species composition and diversity affect grassland susceptibility and response to invasion." *Ecological Applications* no. 12 (2):602-617. doi: 10.1890/1051-0761(2002)012[0602:SCADAG]2.0.CO;2.

Foley, JA, N Ramankutty, KA Brauman, ES Cassidy, JS Gerber, M Johnston, ND Mueller, C O'Connell, DK Ray, PC West, C Balzer, EM Bennett, SR Carpenter, J Hill, C Monfreda, S Polasky, J Rockstrom, J Sheehan, S Siebert, D Tilman, and DPM Zaks. 2011. "Solutions for a cultivated planet." *Nature* no. 478 (7369):337-342. doi: 10.1038/nature10452.

Frankow-Lindberg, BE. 2012. "Grassland plant species diversity decreases invasion by increasing resource use." *Oecologia* no. 169 (3):793-802. doi: 10.1007/s00442-011-2230-7.

Freemark, K, and C Boutin. 1995. "Imapcts of agricultural herbicide use on terrestrail wildlife in temeprate landscapes - a review with special reference to North America." *Agriculture Ecosystems & Environment* no. 52 (2-3):67-91. doi: 10.1016/0167-8809(94)00534-L.

Grman, E, T Bassett, and LA Brudvig. 2013. "Confronting contingency in restoration: management and site history determine outcomes of assembling prairies, but site characteristics and landscape context have little effect." *Journal of Applied Ecology* no. 50 (5):1234-1243. doi: 10.1111/1365-2664.12135.

Hector, A, K Dobson, A Minns, E Bazeley-White, and JH Lawton. 2001. "Community diversity and invasion resistance: An experimental test in a grassland ecosystem and a review of comparable studies." *Ecological Research* no. 16 (5):819-831. doi: 10.1046/j.1440-1703.2001.00443.x.

Hoekstra, NJ, M Suter, JA Finn, S Husse, and A Luscher. 2015. "Do belowground vertical niche differences between deep- and shallow-rooted species enhance resource uptake and drought resistance in grassland mixtures?" *Plant and Soil* no. 394 (1-2):21-34. doi: 10.1007/s11104-014-2352-x.

Hooper, DU, FS Chapin, JJ Ewel, A Hector, P Inchausti, S Lavorel, JH Lawton, DM Lodge, M Loreau, S Naeem, B Schmid, H Setala, AJ Symstad, J Vandermeer, and DA Wardle. 2005. "Effects of biodiversity on ecosystem functioning: A consensus of current knowledge." *Ecological Monographs* no. 75 (1):3-35. doi: 10.1890/04-0922.

Humphreys, J, P Phelan, DJ Li, W Burchill, J Eriksen, I Casey, D Enriquez-Hidalgo, K Soegaard, D MurphyBokern, FL Stoddard, and CA Watson. 2017. "White Clover Supported Pasture-based Systems in North-west Europe." *Legumes in Cropping Systems*:139-156.

Isbell, F, PR Adler, N Eisenhauer, D Fornara, K Kimmel, C Kremen, DK Letourneau, M Liebman, HW Polley, S Quijas, and M Scherer-Lorenzen. 2017. "Benefits of increasing plant diversity in sustainable agroecosystems." *Journal of Ecology* no. 105 (4):871-879. doi: 10.1111/1365-2745.12789.

Kirwan, L, J Connolly, JA Finn, C Brophy, A Luscher, D Nyfeler, and MT Sebastia. 2009. "Diversity-interaction modeling: estimating contributions of species identities and interactions to ecosystem function." *Ecology* no. 90 (8):2032-2038. doi: 10.1890/08-1684.1.

Kirwan, L, A Luescher, MT Sebastia, JA Finn, RP Collins, C Porqueddu, A Helgadottir, OH Baadshaug, C Brophy, C Coran, S Dalmannsdottir, I Delgado, A Elgersma, M Fothergill, BE Frankow-Lindberg, P Golinski, P Grieu, AM Gustavsson, M Hoglind, O Huguenin-Elie, C Iliadis, M Jorgensen, Z Kadziuliene, T Karyotis, T Lunnan, M Malengier, S Maltoni, V Meyer, D Nyfeler, P Nykanen-Kurki, J Parente, HJ Smit, U Thumm, and J Connolly. 2007. "Evenness drives consistent diversity effects in intensive grassland systems across 28 European sites." *Journal of Ecology* no. 95 (3):530-539. doi: 10.1111/j.1365-2745.2007.01225.x.

Lamosova, T, J Dolezal, V Lanta, and J Leps. 2010. "Spatial pattern affects diversity-productivity relationships in experimental meadow communities." *Acta Oecologica* no. 36 (3):325-332. doi: 10.1016/j.actao.2010.02.005.

Levine, JM, M Vila, CM D'Antonio, JS Dukes, K Grigulis, and S Lavorel. 2003. "Mechanisms underlying the impacts of exotic plant invasions." *Proceedings of the Royal Society B-Biological Sciences* no. 270 (1517):775-781. doi: 10.1098/rspb.2003.2327.

Littell, Ramon C., George A. Milliken, Walter W. Stroup, Russell D. Wolfinger, and Oliver Schabenberger. 2006. *SAS for mixed models*. 2nd ed. Cary, NC, USA: SAS Institute.

Mahanty, T, S Bhattacharjee, M Goswami, P Bhattacharyya, B Das, A Ghosh, and P Tribedi. 2017. "Biofertilizers: a potential approach for sustainable agriculture development." *Environmental Science and Pollution Research* no. 24 (4):3315-3335. doi: 10.1007/s11356-016-8104-0.

Maron, JL, and M Marler. 2008. "Effects of native species diversity and resource additions on invader impact." *American Naturalist* no. 172:S18-S33. doi: 10.1086/588303.

Martin, LM, KA Moloney, and BJ Wilsey. 2005. "An assessment of grassland restoration success using species diversity components." *Journal of Applied Ecology* no. 42 (2):327-336. doi: 10.1111/j.1365-2664.2005.01019.x.

Martin, LM, and BJ Wilsey. 2012. "Assembly history alters alpha and beta diversity, exotic-native proportions and functioning of restored prairie plant communities." *Journal of Applied Ecology* no. 49 (6):1436-1445. doi: 10.1111/j.1365-2664.2012.02202.x.

McKenna, TP, and KA Yurkonis. 2016. "Across species-pool aggregation alters grassland productivity and diversity." *Ecology and Evolution* no. 6 (16):5788-5795. doi: 10.1002/ece3.2325.

Mittelbach, GG, CF Steiner, SM Scheiner, KL Gross, HL Reynolds, RB Waide, MR Willig, SI Dodson, and L Gough. 2001. "What is the observed relationship between species richness and productivity?" *Ecology* no. 82 (9):2381-2396.

Murrell, DJ, DW Purves, and R Law. 2001. "Uniting pattern and process in plant ecology." *Trends in Ecology & Evolution* no. 16 (10):529-530. doi: 10.1016/S0169-5347(01)02292-3.

Nash, JC, and R Varadhan. 2011. "Unifying optimization algorithms to aid software system users: optimx for R." *Journal of Statistical Software* no. 43 (9):1-14.

O'Hea, NM, L Kirwan, and JA Finn. 2010. "Experimental mixtures of dung fauna affect dung decomposition through complex effects of species interactions." *Oikos* no. 119 (7):1081-1088. doi: 10.1111/j.1600-0706.2009.18116.x.

Petermann, JS, AJF Fergus, C Roscher, LA Turnbull, A Weigelt, and B Schmid. 2010. "Biology, chance, or history? The predictable reassembly of temperate grassland communities." *Ecology* no. 91 (2):408-421. doi: 10.1890/08-2304.1.

Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and R Core Team. 2018. "nlme: Linear and nonlinear mixed effects models." no. R package version 3.1-137.

Piovia-Scott, J, D Rejmanek, DC Woodhams, SJ Worth, H Kenny, V McKenzie, SP Lawler, and JE Foley. 2017. "Greater species richness of bacterial skin symbionts better suppresses the amphibian fungal pathogen Batrachochytrium Dendrobatidis." *Microbial Ecology* no. 74 (1):217-226. doi: 10.1007/s00248-016-0916-4.

Pokorny, ML, RL Sheley, CA Zabinski, RE Engel, TJ Svejcar, and JJ Borkowski. 2005. "Plant functional group diversity as a mechanism for invasion resistance." *Restoration Ecology* no. 13 (3):448-459. doi: 10.1111/j.1526-100X.2005.00056.x.

R Core Team. 2018. "R: A language and environment for statistical computing." *R Foundation for Statistical Computing*.

Rayburn, AP, and EW Schupp. 2013. "Effects of community- and neighborhood-scale spatial patterns on semi-arid perennial grassland community dynamics." *Oecologia* no. 172 (4):1137-1145. doi: 10.1007/s00442-012-2567-6.

Roscher, C, VM Temperton, M Scherer-Lorenzen, M Schmitz, J Schumacher, B Schmid, N Buchmann, WW Weisser, and ED Schulze. 2005. "Overyielding in experimental grassland communities - irrespective of species pool or spatial scale." *Ecology Letters* no. 8 (4):419-429. doi: 10.1111/j.1461-0248.2005.00736.x.

Scherber, C, N Eisenhauer, WW Weisser, B Schmid, W Voigt, M Fischer, ED Schulze, C Roscher, A Weigelt, E Allan, H Bessler, M Bonkowski, N Buchmann, F Buscot, LW Clement, A Ebeling, C Engels, S Halle, I Kertscher, AM Klein, R Koller, S Konig, E Kowalski, V Kummer, A Kuu, M Lange, D Lauterbach, C Middelhoff, VD Migunova, A Milcu, R Muller, S Partsch, JS Petermann, C Renker, T Rottstock, A Sabais, S Scheu, J Schumacher, VM Temperton, and T Tscharntke. 2010. "Bottom-up effects of plant diversity on multitrophic interactions in a biodiversity experiment." *Nature* no. 468 (7323):553-556. doi: 10.1038/nature09492.

Schwartz, MW, CA Brigham, JD Hoeksema, KG Lyons, MH Mills, and PJ van Mantgem. 2000. "Linking biodiversity to ecosystem function: implications for conservation ecology." *Oecologia* no. 122 (3):297-305. doi: 10.1007/s004420050035.

Seahra, SE, KA Yurkonis, and JA Newman. 2016. "Species patch size at seeding affects diversity and productivity responses in establishing grasslands." *Journal of Ecology* no. 104 (2):479-486. doi: 10.1111/1365-2745.12514.

Self, SG, and KY Liang. 1987. "Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions." *Journal of the American Statistical Association* no. 82 (398):605-610. doi: 10.2307/2289471.

Sercu, Bram, Lander Baeten, Frieke van Coillie, An Martel, Luc Lens, Kris Verheyen, and Dries Bonte. 2017. "How tree species identity and diversity affect light transmittance to the understory in mature temperate forests." *Ecology and Evolution* no. 7 (24):10861-10870. doi: 10.1002/ece3.3528.

Shapiro, A. 1988. "Towards a unified theory of inequality constrained testing in multivariate analysis." *International Statistical Review* no. 56 (1):49-62. doi: 10.2307/1403361.

Spehn, EM, A Hector, J Joshi, M Scherer-Lorenzen, B Schmid, E Bazeley-White, C Beierkuhnlein, MC Caldeira, M Diemer, PG Dimitrakopoulos, JA Finn, H Freitas, PS Giller, J Good, R Harris, P Hogberg, K Huss-Danell, A Jumpponen, J Koricheva, PW Leadley, M Loreau, A Minns, CPH Mulder, G O'Donovan, SJ Otway, C Palmborg, JS Pereira, AB Pfisterer, A Prinz, DJ Read, ED Schulze, ASD Siamantziouras, AC Terry, AY Troumbis, FI Woodward, S Yachi, and JH Lawton. 2005. "Ecosystem effects of biodiversity manipulations in European grasslands." *Ecological Monographs* no. 75 (1):37-63. doi: 10.1890/03-4101.

Stoll, P, and D Prati. 2001. "Intraspecific aggregation alters competitive interactions in experimental plant communities." *Ecology* no. 82 (2):319-327.

Sturludottir, E, C Brophy, G Belanger, AM Gustavsson, M Jorgensen, T Lunnan, and A Helgadottir. 2014. "Benefits of mixing grasses and legumes for herbage yield and nutritive value in Northern Europe and Canada." *Grass and Forage Science* no. 69 (2):229-240. doi: 10.1111/gfs.12037.

Tilman, D, F Isbell, JM Cowles, and DJ Futuyma. 2014. "Biodiversity and Ecosystem Functioning." *Annual Review of Ecology, Evolution, and Systematics, Vol 45* no. 45:471-493. doi: 10.1146/annurev-ecolsys-120213-091917.

Wassmuth, BE, P Stoll, T Tscharntke, and C Thies. 2009. "Spatial aggregation facilitates coexistence and diversity of wild plant species in field margins." *Perspectives in Plant Ecology Evolution and Systematics* no. 11 (2):127-135. doi: 10.1016/j.ppees.2009.02.001.

Wilsey, B, and G Stirling. 2007. "Species richness and evenness respond in a different manner to propagule density in developing prairie microcosm communities." *Plant Ecology* no. 190 (2):259-273. doi: 10.1007/s11258-006-9206-4.

Yurkonis, KA, BJ Wilsey, and KA Moloney. 2012. "Initial species pattern affects invasion resistance in experimental grassland plots." *Journal of Vegetation Science* no. 23 (1):4-12. doi: 10.1111/j.1654-1103.2011.01331.x.

Zhang, YH, YF Wang, and SX Yu. 2014. "Interspecific neighbor interactions promote the positive diversity-productivity relationship in experimental grassland communities." *Plos One* no. 9 (10). doi: 10.1371/journal.pone.0111434.

Chapter 5

Conclusion

My research has focused on assessing the accuracy of weather forecasts, and incorporating them into an Irish grass growth model, as well as extending the DI modelling framework for the modelling of species-diverse grassland experiments. Below I will summarise the main conclusions from each chapter.

Chapter 2

Chapter 2 focused on the first aim of the thesis, which was to quantify the accuracy of weather forecasts in Ireland, and if possible to improve their accuracy. It was shown that ECMWF weather forecasts of all weather variables examined declined in accuracy as forecast period increased, and were accurate up to at least five days in advance. The best bias-correction method used a regression approach, and improved the accuracy of the ECMWF air and soil temperature forecasts, but rainfall forecasts were not greatly improved. Previous studies on the verification and bias correction of rainfall forecasts showed similar results: improvements in rainfall forecast accuracy but difficulty in predicting extreme events and for longer lead times (Acharya et al., 2013; Harrison et al., 2000). Although the bias corrections showed improvements in forecast accuracy, recently observed weather data are required to use these bias correction methods. Farms that are not beside weather stations and that do not record their daily weather would have less accurate bias-correction options.

Although this work was done primarily to aid with the inclusion of a grass growth model, it could be used for many other applications. For example, ECMWF soil temperature forecasts had not been verified previously in Ireland at all of these depths. Albergel et al. (2015) verified the ECMWF 5 cm soil temperature at a number of European sites, including at some Irish locations. As well as being a strong predictor of grass growth, soil temperature is important in the growth of many Irish crops, including potatoes. Up until now, these forecasts have been used primarily to inform the interaction between the air and the earth surface in the weather forecasting models produced by the ECMWF. Met Éireann can now advise users of soil temperature forecasts that they are reasonably accurate and can be used in practice. They also contain seasonal biases that can be corrected if observations are available.

Chapter 3

Chapter 3 investigated whether weather forecasts were reasonable inputs for a grass growth model at an Irish dairy farm. Previous work by Ruelle et al. (2018) and Hurtado-Uria et al. (2013) showed that Irish grass growth observations could be predicted accurately using retrospective weather observations. Internationally, grass growth modelling had focused on long-term seasonal hindcasting, but the potential benefits were not sufficient to promote adoption by farmers in practice (Ash et al., 2007; Harrison et al., 2017). Here, it was shown that they could be predicted with reasonable accuracy up to a week in advance using weather forecasts. The weekly, monthly and yearly totals from the MoSt GGM predictions generally followed the grass growth observations closely for both forecast and observed weather, showing that the MoSt GGM including weather forecasts can be a useful management tool for farmers. The grass growth observations outside of the peak growing season were generally under-predicted by the MoSt GGM, regardless of the weather input. It will be important to identify the source of this error because it can be important for farmers to know their grass growth early in the growing season to make grazing decisions that can have a strong influence on and grass utilisation and quality across the growing season (Kennedy et al., 2006; O'Donovan and Delaby, 2008). Some circumstances in which certain weather conditions were overly influential during the peak growing season were identified, in particular low solar radiation during the peak growing season resulted in grass growth predictions lower than observed.

Chapter 4

Chapter 4 aimed to enhance existing methods for modelling multi-species experiments. DI models have been used previously to describe numerous different types of experiments, including tree species interactions, weed biomass in grasslands, and the effect of dung fauna on dung decomposition (Connolly et al., 2018; De Groote et al., 2017; O'Hea et al., 2010). They also incorporated many treatments, but a spatial planting pattern treatment had not previously been studied. In the year of establishment there was no difference between weed biomass in mixture and monoculture plots. Mean weed biomass was lower in mixture plots than in monoculture plots in the second and third years of the study, showing that increasing species richness reduced weed invasion. The mean weed biomass increased from year two to year three of the study. This could be because the plots were in a natural system, and soil nutrients may have been depleted. Although some species suppressed weeds effectively in monoculture, there are other benefits in species mixtures in grasslands such as increased grass yields, and environmental enrichment.

A number of DI modelling methods were developed, as a complex covariance structure in a repeated measures setting was implemented, and random pairwise interactions were fitted in multiple years. In the final model, there were different blocks for monocultures from each functional group, and another for mixtures in the ***R*** matrix, while random pairwise interactions were included in the model in 2014. The chosen model had a functional group structure which interacted with spatial planting pattern, meaning each pair of species interacted depending on their functional group membership and the spatial planting pattern. Legumes influenced the interactions: each pair of species interacted based on the sum of two species-specific (additive species) interaction effects, scaled by their proportions as well as by the proportion of legumes present. It was shown that the random pairwise interactions methods described almost always detected random pairwise interactions when their common variance was sufficiently large in relation to the error variance. However, when this was not the case, the models including random pairwise interactions were often stopped due to infinite likelihood. Since spatial planting pattern appeared in the final model, predictions allowed us to explore how the two spatial patterns differed. For two-species mixtures in 2012, aggregation suppressed weed invasion better possibly due to it allowing less competitive species to establish better. However, in 2013 and 2014, dispersed communities generally gave lower weed biomass predictions in four-species mixtures. Tools and code in SAS and R for analysing and fitting DI models to multi-year and species-rich biodiversity experiments were also provided.

Future work

There are plenty of opportunities for further research in grassland modelling with weather forecasts. Solar radiation and wind speed and direction are two weather variables that influence grass growth that have not been extensively verified here. An increased understanding of the accuracy of their forecasts, as well as their relationship with grass growth could help to develop the MoSt GGM. For example, using mean climatological solar radiation forecasts in conjunction with ECMWF or bias-corrected temperature and rainfall forecasts could yield accurate predictions if the solar radiation forecasts are not accurate a week in advance. Climate change will bring challenges to the agricultural industry, and studies investigating the effects of weather on areas such as grass and crop growth will be important in ensuring food sustainability in the future. As drought and flood conditions become more common in Ireland, research into their effects on agricultural production is imperative. There are also opportunities to collaborate with agricultural researchers as a statistician. Many potentially information-rich experiments are not being utilised to their fullest.

Although the MoSt GGM has been shown to be producing accurate predictions, a number of areas in which its predictive ability could be improved have been highlighted. The MoSt GGM study was carried out for a well-drained farm, and the model is designed to be able to predict grass growth on any Irish dairy farm. It would be useful to test the model outputs at other sites with different climatic conditions and soil types, for example on heavy soils. The MoSt GGM must also be adapted to include multi-species swards, as Irish farmers are incorporating species such as white clover to improve grass and milk yield (Egan et al., 2018; Guy et al., 2018), and to promote biodiversity to enhance the ecosystem in an increasingly environmentally aware industry (Tilman et al., 2006).

Grass growth models have been developed in other countries. New Zealand’s Pasture Growth Forecaster appears to be the most advanced of these models, and is used in practice by farmers to make informed grassland management decisions (NZ, 2018). It is similar to the PastureBase Ireland tools: it is split into regions, and in each region daily grass growth predictions for the next seven days are available as well as the typical growth in the region based on historical climate and grass growth data. However, there are some interesting features in this model that could be useful to incorporate into PastureBase Ireland, in conjunction with the MoSt GGM. For example, it uses climatic data and the starting soil and water conditions in the system to forecast the expected growth for the next three months and the coming year along with confidence bounds of 50% and 90%. When the PastureBase Ireland dataset accumulates sufficient historical grass growth data, the MoSt GGM could be used in conjunction with the Met Éireann climatology data (Walsh, 2012) to produce similar estimates.

I hope to further develop DI modelling methods and help with the dissemination of the work to researchers working with appropriate species mixture data. DI modelling methods can be applied to a variety of experiments that mix species other than grassland plot experiments, including bacterial and tree species studies. Because DI models can be used in so many areas, there is potential for collaborative work with a growing number of researchers from a wide variety of backgrounds. Publishing material about DI modelling in the statistical literature would help to bring a wider audience to the methods.

References

Acharya, N., Chattopadhyay, S., Mohanty, U., Dash, S. and Sahoo, L., 2013. On the bias correction of general circulation model output for Indian summer monsoon. Meteorological Applications, 20(3): 349-356.

Albergel, C. et al., 2015. Soil temperature at ECMWF: An assessment using ground-based observations. Journal of Geophysical Research-Atmospheres, 120(4): 1361-1373.

Ash, A., McIntosh, P., Cullen, B., Carberry, P. and Smith, M., 2007. Constraints and opportunities in applying seasonal climate forecasts in agriculture. Australian Journal of Agricultural Research, 58(10): 952-965.

Connolly, J. et al., 2018. Weed suppression greatly increased by plant diversity in intensively managed grasslands: A continental-scale experiment. Journal of Applied Ecology, 55(2): 852-862.

De Groote, S. et al., 2017. Tree species identity outweighs the effects of tree species diversity and forest fragmentation on understorey diversity and composition. Plant Ecology and Evolution, 150(3): 229-239.

Egan, M., Galvin, N. and Hennessy, D., 2018. Incorporating white clover (Trifolium repens L.) into perennial ryegrass (Lolium perenne L.) swards receiving varying levels of nitrogen fertilizer: Effects on milk and herbage production. Journal of Dairy Science, 101(4): 3412-3427.

Guy, C., Hennessy, D., Gilliland, T., Coughlan, F. and McCarthy, B., 2018. Growth, morphology and biological nitrogen fixation potential of perennial ryegrass-white clover swards throughout the grazing season. Journal of Agricultural Science, 156(2): 188-199.

Harrison, D., Driscoll, S. and Kitchen, M., 2000. Improving precipitation estimates from weather radar using quality control and correction techniques. Meteorological Applications, 7(2): 135-144.

Harrison, M., Christie, K. and Rawnsley, R., 2017. Assessing the reliability of dynamical and historical climate forecasts in simulating hindcast pasture growth rates. Animal Production Science, 57(7): 1525-1535.

Hurtado-Uria, C., 2013. Evaluation, adaptation and validation of a model to predict grass growth in Ireland, Cork Institute of Technology.

Kennedy, E., O'Donovan, M., Murphy, J., O'Mara, F. and Delaby, L., 2006. The effect of initial spring grazing date and subsequent stocking rate on the grazing management, grass dry matter intake and milk production of dairy cows in summer. Grass and Forage Science, 61(4): 375-384.

NZ, D., 2018. Pasture Growth Forecaster, http://pasture-growth-forecaster.dairynz.co.nz/Home/About.

O'Donovan, M. and Delaby, L., 2008. Sward characteristics, grass dry matter intake and milk production performance is affected by timing of spring grazing and subsequent stocking rate. Livestock Science, 115(2-3): 158-168.

O'Hea, N., Kirwan, L. and Finn, J., 2010. Experimental mixtures of dung fauna affect dung decomposition through complex effects of species interactions. Oikos, 119(7): 1081-1088.

Ruelle, E., Delaby, L. and Hennessy, D., 2018. Development of the Moorepark St Gilles grass growth model (MoSt GG model): A predictive model for grass growth for pasture based systems. European Journal of Agronomy, 99(September 2018): 80-91.

Tilman, D., Reich, P. and Knops, J., 2006. Biodiversity and ecosystem stability in a decade-long grassland experiment. Nature, 441(7093): 629-632.

Walsh, S., 2012. A Summary of Climate Averages 1981-2010 for Ireland, Climatological Note No.14, Met Éireann, Dublin.

Appendices

Appendix 2.1: Number of missing values at each station in each year for the ten weather variables examined: rainfall (rain), minimum temperature (min), maximum temperature (max), mean temperature (mean), 5 cm soil temperature (5cm), 10 cm soil temperature (10cm), 20 cm soil temperature (20cm), 30 cm soil temperature (30cm), 50 cm soil temperature (50cm), 100 cm soil temperature (100cm). Some of the stations became operational after 1 January 2007. Other missing values are usually due to servicing, calibrations or instrument outages.

variable	station	2007	2008	2009	2010	2011	2012	2013
rain	Athenry	365	366	365	365	258	11	3
min	Athenry	365	366	365	365	258	11	3
max	Athenry	365	366	365	365	258	11	3
mean	Athenry	365	366	365	365	258	10	0
5cm	Athenry	365	366	365	365	279	10	0
10cm	Athenry	365	366	365	365	266	11	0
20cm	Athenry	365	366	365	365	257	11	0
30cm	Athenry	365	366	365	365	273	11	1
50cm	Athenry	365	366	365	365	252	11	1
100cm	Athenry	365	366	365	365	254	11	1
rain	Ballyhaise	6	7	2	3	4	0	1
min	Ballyhaise	6	4	0	0	0	0	0
max	Ballyhaise	6	4	0	0	0	0	0
mean	Ballyhaise	2	0	0	0	0	0	0
5cm	Ballyhaise	2	0	0	0	3	1	0
10cm	Ballyhaise	2	0	0	0	0	0	0
20cm	Ballyhaise	2	0	0	0	0	0	0
30cm	Ballyhaise	1	0	0	0	1	0	0
50cm	Ballyhaise	0	0	0	0	0	3	0
100cm	Ballyhaise	0	0	0	0	13	3	0
rain	Belmullet	0	0	0	0	0	2	0
min	Belmullet	0	0	0	0	0	0	0
max	Belmullet	0	0	0	0	0	0	0
mean	Belmullet	0	0	0	0	0	0	0
5cm	Belmullet	0	0	0	0	0	0	0
10cm	Belmullet	0	0	0	0	0	0	0
20cm	Belmullet	0	0	0	0	0	0	0
30cm	Belmullet	0	0	0	0	0	0	0
50cm	Belmullet	0	0	0	0	0	0	0
100cm	Belmullet	0	0	0	0	0	0	0
rain	Carlow	24	2	0	0	1	0	3

min	Carlow	0	0	0	0	1	0	0
max	Carlow	0	0	0	0	1	0	0
mean	Carlow	0	2	2	0	0	0	0
5cm	Carlow	0	0	2	0	0	0	0
10cm	Carlow	0	0	2	0	0	0	0
20cm	Carlow	0	0	1	0	0	0	0
30cm	Carlow	0	0	1	0	0	0	0
50cm	Carlow	0	0	0	0	0	0	0
100cm	Carlow	0	0	0	0	0	0	1
rain	Casement	0	0	0	0	0	0	0
min	Casement	0	0	0	0	0	0	0
max	Casement	0	0	0	0	0	0	0
mean	Casement	0	0	0	0	0	0	0
5cm	Casement	0	0	0	0	0	0	0
10cm	Casement	0	0	0	0	0	0	0
20cm	Casement	0	0	0	0	0	0	0
30cm	Casement	0	0	0	0	0	0	0
50cm	Casement	0	0	0	0	0	0	0
100cm	Casement	0	0	0	0	0	0	0
rain	Claremorris	0	0	2	0	1	0	1
min	Claremorris	0	0	3	0	1	0	0
max	Claremorris	0	0	3	0	1	0	0
mean	Claremorris	0	0	0	0	23	0	0
5cm	Claremorris	0	0	0	0	1	1	0
10cm	Claremorris	0	0	0	0	1	2	0
20cm	Claremorris	0	0	0	0	1	0	0
30cm	Claremorris	0	0	0	0	1	0	0
50cm	Claremorris	0	0	0	0	1	0	0
100cm	Claremorris	0	1	0	0	1	0	0
rain	Cork Airport	0	0	0	0	0	0	0
min	Cork Airport	0	0	0	0	0	0	0
max	Cork Airport	0	0	0	0	0	0	0
mean	Cork Airport	0	0	0	0	0	0	0
5cm	Cork Airport	0	0	0	0	0	0	0
10cm	Cork Airport	0	0	0	0	0	0	0
20cm	Cork Airport	0	0	0	0	0	0	0
30cm	Cork Airport	0	0	0	0	0	0	0
50cm	Cork Airport	0	0	0	0	0	0	0
100cm	Cork Airport	0	0	0	0	0	0	0
rain	Dublin Airport	0	0	0	0	0	0	0
min	Dublin Airport	0	0	0	0	0	0	0
max	Dublin Airport	0	0	0	0	0	0	0
mean	Dublin Airport	0	0	0	0	0	0	0
5cm	Dublin Airport	0	0	0	0	0	0	0
10cm	Dublin Airport	0	0	0	0	0	0	0
20cm	Dublin Airport	0	0	0	0	0	0	0
30cm	Dublin Airport	1	0	0	0	0	0	0
50cm	Dublin Airport	1	0	0	0	0	0	0
100cm	Dublin Airport	1	0	0	0	0	0	0

rain	Dunsany	314	61	0	1	1	2	1
min	Dunsany	315	61	0	1	1	1	1
max	Dunsany	315	61	0	1	1	1	1
mean	Dunsany	314	59	0	0	0	0	0
5cm	Dunsany	314	59	0	0	0	0	0
10cm	Dunsany	314	59	0	0	0	0	0
20cm	Dunsany	314	59	0	0	0	0	0
30cm	Dunsany	314	59	0	0	0	0	1
50cm	Dunsany	314	59	0	0	0	0	0
100cm	Dunsany	314	59	0	0	0	0	0
rain	Finner	365	366	365	365	26	2	5
min	Finner	365	366	365	365	26	2	8
max	Finner	365	366	365	365	26	2	8
mean	Finner	27	2	8	29	6	0	5
5cm	Finner	365	366	365	365	365	366	365
10cm	Finner	365	366	365	365	365	366	365
20cm	Finner	365	366	365	365	365	366	365
30cm	Finner	365	366	365	365	365	366	365
50cm	Finner	365	366	365	365	365	366	365
100cm	Finner	365	366	365	365	365	366	365
rain	Gurteen	365	40	0	2	1	0	2
min	Gurteen	365	40	0	2	0	0	0
max	Gurteen	365	40	0	2	0	0	0
mean	Gurteen	365	39	0	0	0	0	0
5cm	Gurteen	365	38	0	0	0	1	0
10cm	Gurteen	365	38	0	0	0	0	0
20cm	Gurteen	365	38	0	0	0	0	0
30cm	Gurteen	365	39	0	0	0	0	0
50cm	Gurteen	365	39	0	0	0	1	0
100cm	Gurteen	365	39	0	0	0	0	0
rain	Johnstown	13	2	0	4	1	2	0
min	Johnstown	0	0	0	0	0	0	0
max	Johnstown	0	0	0	0	0	0	0
mean	Johnstown	0	0	0	45	6	0	0
5cm	Johnstown	0	1	5	219	124	0	0
10cm	Johnstown	0	0	3	55	15	0	0
20cm	Johnstown	0	1	3	7	2	0	0
30cm	Johnstown	0	1	3	17	0	0	0
50cm	Johnstown	0	0	3	12	0	0	0
100cm	Johnstown	0	0	4	13	0	0	0
rain	Knock Airport	0	0	0	0	0	0	0
min	Knock Airport	0	0	0	0	0	0	0
max	Knock Airport	0	0	0	0	0	0	0
mean	Knock Airport	0	0	0	0	0	0	0
5cm	Knock Airport	0	0	0	0	0	0	0
10cm	Knock Airport	0	0	0	0	0	0	0
20cm	Knock Airport	0	0	0	0	0	0	0
30cm	Knock Airport	0	0	0	0	0	0	0
50cm	Knock Airport	0	0	0	0	0	0	0

100cm	Knock Airport	0	0	0	0	0	0	0
rain	Mace Head	220	13	7	1	1	0	4
min	Mace Head	229	22	20	1	1	0	4
max	Mace Head	229	22	20	1	1	0	4
mean	Mace Head	209	1	0	0	0	0	1
5cm	Mace Head	365	366	365	365	365	366	365
10cm	Mace Head	365	366	365	365	365	366	365
20cm	Mace Head	365	366	365	365	365	366	365
30cm	Mace Head	365	366	365	365	365	366	365
50cm	Mace Head	365	366	365	365	365	366	365
100cm	Mace Head	365	366	365	365	365	366	365
rain	Malin Head	0	0	0	0	0	0	0
min	Malin Head	0	0	0	1	0	0	0
max	Malin Head	0	0	0	1	0	0	0
mean	Malin Head	0	0	0	1	0	0	0
5cm	Malin Head	0	0	0	0	0	1	0
10cm	Malin Head	0	0	0	0	0	0	0
20cm	Malin Head	0	0	0	0	0	0	0
30cm	Malin Head	0	0	0	0	0	0	0
50cm	Malin Head	0	0	0	0	0	0	0
100cm	Malin Head	0	0	0	0	0	0	0
rain	Markree	312	60	3	1	2	1	1
min	Markree	312	60	4	1	0	0	0
max	Markree	312	60	4	1	0	0	0
mean	Markree	312	60	0	0	0	0	0
5cm	Markree	312	60	0	0	0	0	0
10cm	Markree	312	60	0	0	0	0	0
20cm	Markree	312	60	0	0	0	0	0
30cm	Markree	312	60	0	0	0	0	1
50cm	Markree	312	60	0	0	0	0	0
100cm	Markree	312	60	0	0	0	0	0
rain	Moorepark	268	2	1	2	1	0	1
min	Moorepark	269	1	1	0	1	0	0
max	Moorepark	269	1	1	0	1	0	0
mean	Moorepark	268	0	0	0	0	0	0
5cm	Moorepark	268	0	0	0	0	0	0
10cm	Moorepark	268	0	0	0	0	0	0
20cm	Moorepark	268	0	0	0	0	0	0
30cm	Moorepark	268	0	0	0	0	0	0
50cm	Moorepark	268	0	0	0	0	0	0
100cm	Moorepark	268	0	0	0	0	0	0
rain	Mt Dillon	321	148	11	3	1	0	1
min	Mt Dillon	314	146	12	7	1	1	0
max	Mt Dillon	314	146	12	7	1	1	0
mean	Mt Dillon	313	144	6	0	0	0	0
5cm	Mt Dillon	313	144	5	0	0	0	0
10cm	Mt Dillon	313	144	5	0	0	0	0
20cm	Mt Dillon	313	144	5	0	0	0	0
30cm	Mt Dillon	313	144	6	0	0	0	0

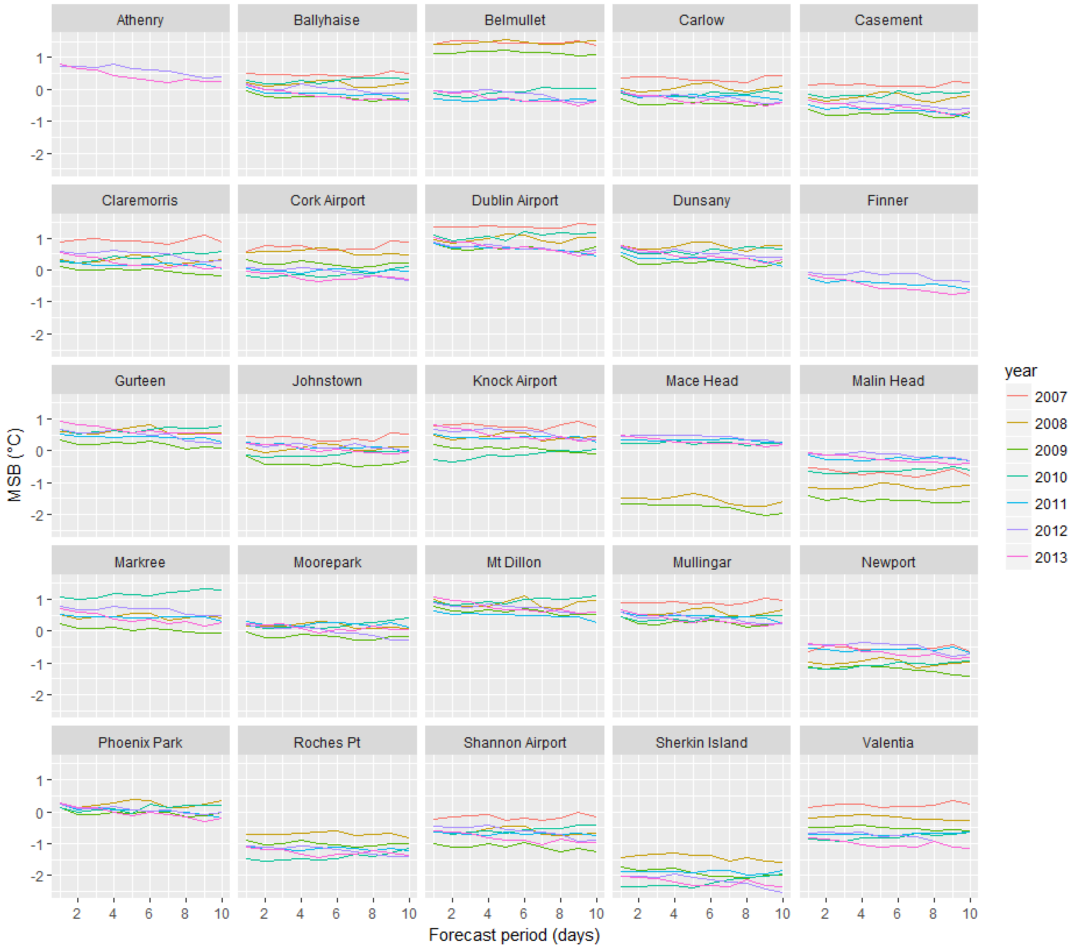
50cm	Mt Dillon	313	144	6	0	1	0	0
100cm	Mt Dillon	313	144	6	0	0	0	0
rain	Mullingar	0	0	0	1	1	1	1
min	Mullingar	0	1	0	0	0	1	0
max	Mullingar	0	1	0	0	0	1	0
mean	Mullingar	0	0	0	0	0	0	0
5cm	Mullingar	0	0	0	0	0	0	0
10cm	Mullingar	0	0	0	0	0	0	0
20cm	Mullingar	0	0	0	0	0	0	0
30cm	Mullingar	0	0	0	0	0	0	1
50cm	Mullingar	0	0	0	0	0	0	1
100cm	Mullingar	0	0	0	0	0	2	1
rain	Newport	5	3	0	1	4	3	1
min	Newport	14	3	0	1	1	3	2
max	Newport	14	3	0	1	1	3	2
mean	Newport	1	0	0	0	0	0	0
5cm	Newport	1	0	0	0	0	0	0
10cm	Newport	1	0	0	0	0	0	0
20cm	Newport	1	0	0	0	0	0	0
30cm	Newport	1	0	0	0	0	1	0
50cm	Newport	1	0	0	0	0	1	1
100cm	Newport	1	0	0	0	0	1	1
rain	Phoenix Park	239	30	2	2	1	1	1
min	Phoenix Park	238	58	2	1	0	0	0
max	Phoenix Park	238	57	2	1	0	0	0
mean	Phoenix Park	235	2	0	0	0	0	0
5cm	Phoenix Park	235	0	0	0	0	0	0
10cm	Phoenix Park	235	0	0	0	0	0	0
20cm	Phoenix Park	235	0	0	0	0	0	0
30cm	Phoenix Park	235	1	0	0	0	0	0
50cm	Phoenix Park	235	1	0	0	0	0	0
100cm	Phoenix Park	235	1	0	0	0	0	0
rain	Roches Pt	365	97	0	6	1	0	1
min	Roches Pt	365	98	0	1	1	0	0
max	Roches Pt	365	98	0	1	1	0	0
mean	Roches Pt	0	2	0	0	0	0	0
5cm	Roches Pt	0	0	0	0	0	0	0
10cm	Roches Pt	0	0	0	0	0	0	0
20cm	Roches Pt	0	0	0	0	0	0	0
30cm	Roches Pt	0	0	0	1	0	0	0
50cm	Roches Pt	0	0	0	1	0	0	1
100cm	Roches Pt	0	0	0	1	0	0	1
rain	Shannon Airport	0	0	0	0	0	0	0
min	Shannon Airport	0	0	0	0	0	0	0
max	Shannon Airport	0	0	0	0	0	0	0
mean	Shannon Airport	0	0	0	0	0	0	0
5cm	Shannon Airport	0	0	0	0	0	0	0
10cm	Shannon Airport	0	0	0	0	0	0	0
20cm	Shannon Airport	0	0	0	0	0	0	0

30cm	Shannon Airport	0	0	0	0	0	0	0
50cm	Shannon Airport	0	0	0	0	0	0	0
100cm	Shannon Airport	0	0	0	0	0	0	0
rain	Sherkin Island	268	22	6	1	2	2	0
min	Sherkin Island	270	17	15	1	2	2	1
max	Sherkin Island	270	16	15	1	2	2	1
mean	Sherkin Island	268	9	3	0	0	0	0
5cm	Sherkin Island	268	4	3	0	0	0	0
10cm	Sherkin Island	268	4	3	0	0	0	0
20cm	Sherkin Island	268	4	3	0	0	0	0
30cm	Sherkin Island	268	8	3	0	1	0	0
50cm	Sherkin Island	268	8	3	0	1	0	0
100cm	Sherkin Island	268	8	3	0	1	0	0
rain	Valentia	0	0	0	0	0	2	0
min	Valentia	0	0	0	0	0	1	0
max	Valentia	0	0	0	0	0	1	0
mean	Valentia	0	0	0	0	0	4	0
5cm	Valentia	0	0	0	0	0	0	0
10cm	Valentia	0	0	0	0	0	0	0
20cm	Valentia	0	0	0	0	0	1	0
30cm	Valentia	0	0	0	0	0	4	0
50cm	Valentia	0	0	0	0	0	2	0
100cm	Valentia	0	0	0	0	0	2	0

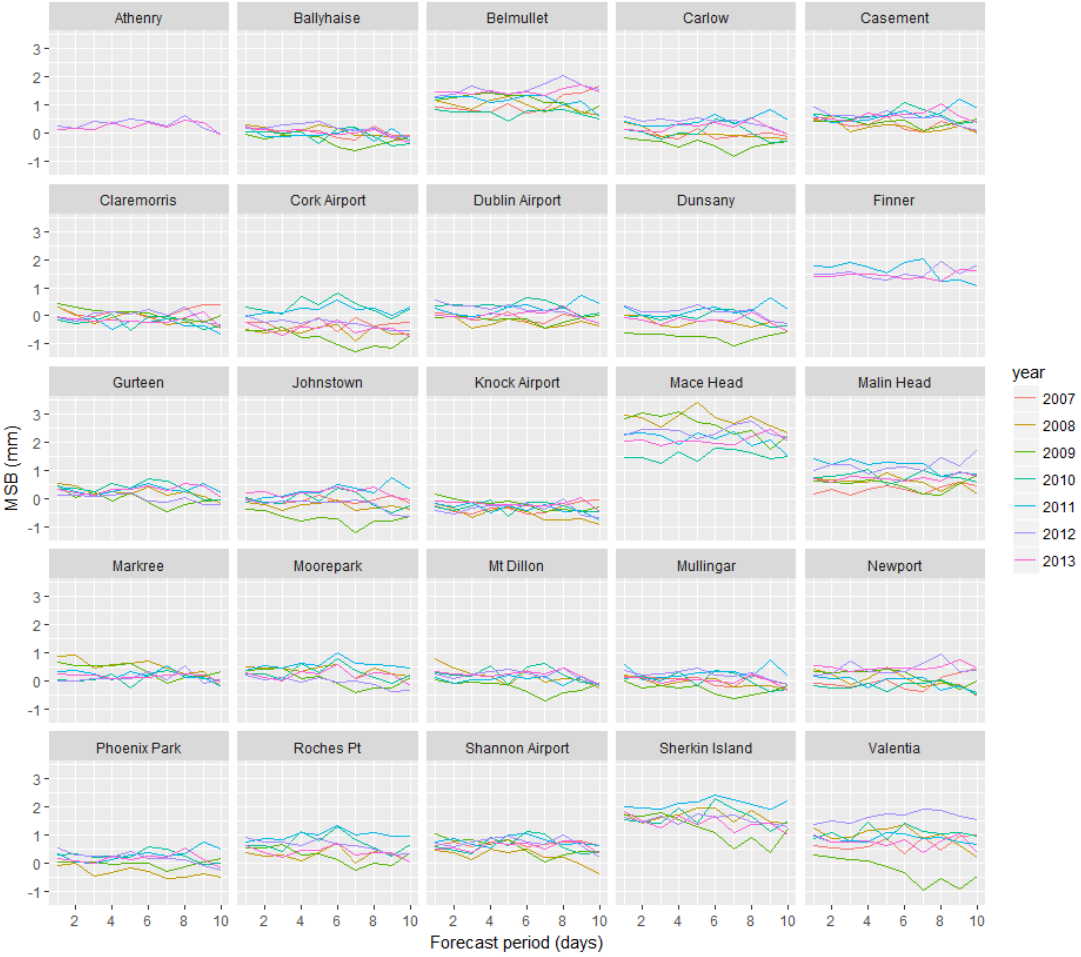
Appendix 2.2: RMSE values (°C) for minimum temperature (min), maximum temperature (max) and mean temperature (mean) at forecast periods 1 to 10 for all years at all stations.

Forecast period	1	2	3	4	5	6	7	8	9	10
min	1.51	1.63	1.79	1.98	2.25	2.53	2.82	3.03	3.25	3.44
max	1.36	1.45	1.55	1.70	1.89	2.10	2.31	2.52	2.75	2.93
mean	0.89	0.99	1.09	1.25	1.48	1.73	2.02	2.23	2.49	2.70

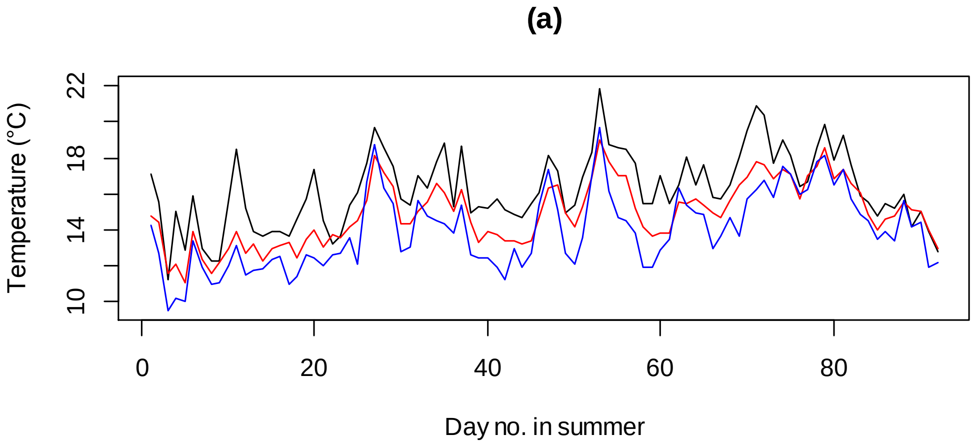
Appendix 2.3: Yearly minimum temperature MSB (°C) of ECMWF forecasts for each forecast period from day 1 to day 10 for each of the Met Éireann synoptic stations.

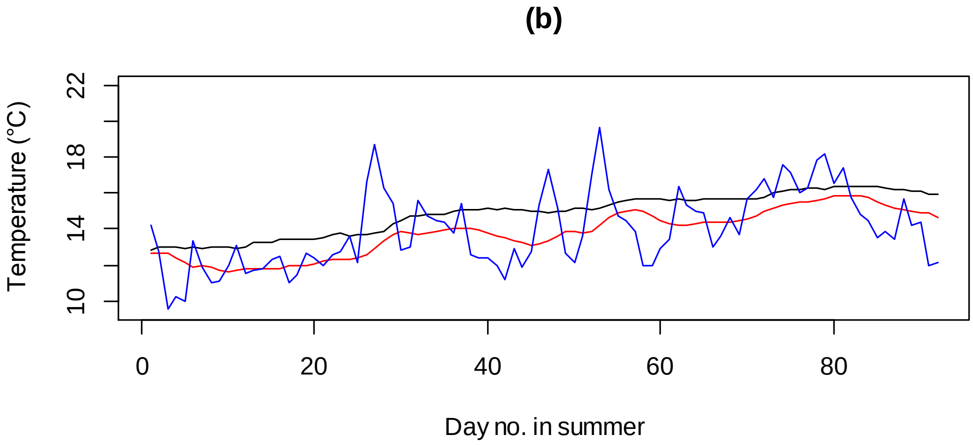


Appendix 2.4: Yearly total daily rainfall MSB (mm) of ECMWF forecasts for each forecast period from day 1 to day 10 for each of the Met Éireann synoptic stations.



Appendix 2.5:. June 1st to August 31st 2012 daily soil temperature observations (black) at (a) 5cm and (b) 100cm and forecasts (red) for (a) STL1 and (b) STL3 at Dublin Airport. The blue line shows the daily mean 2m air temperature observations for the same time period.

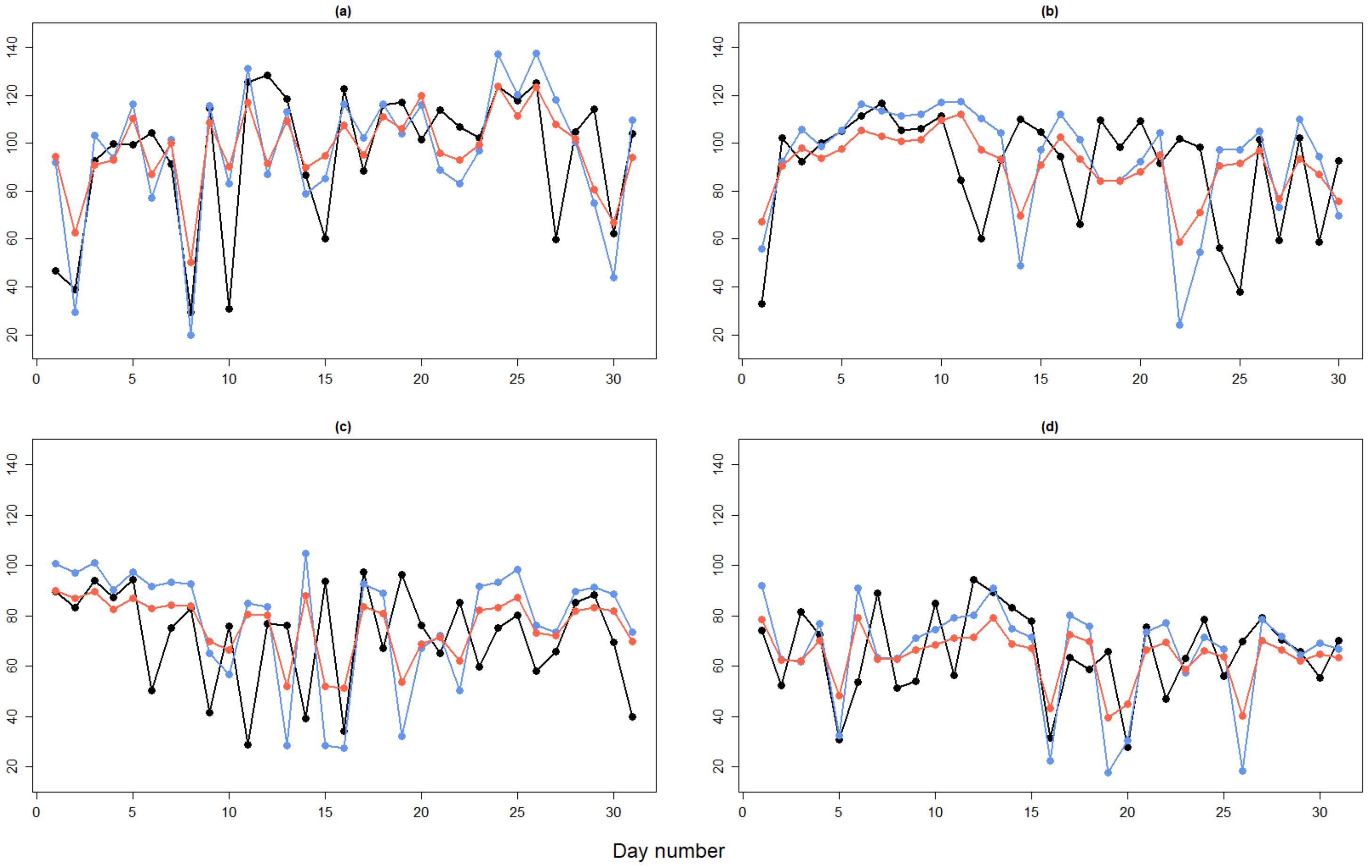




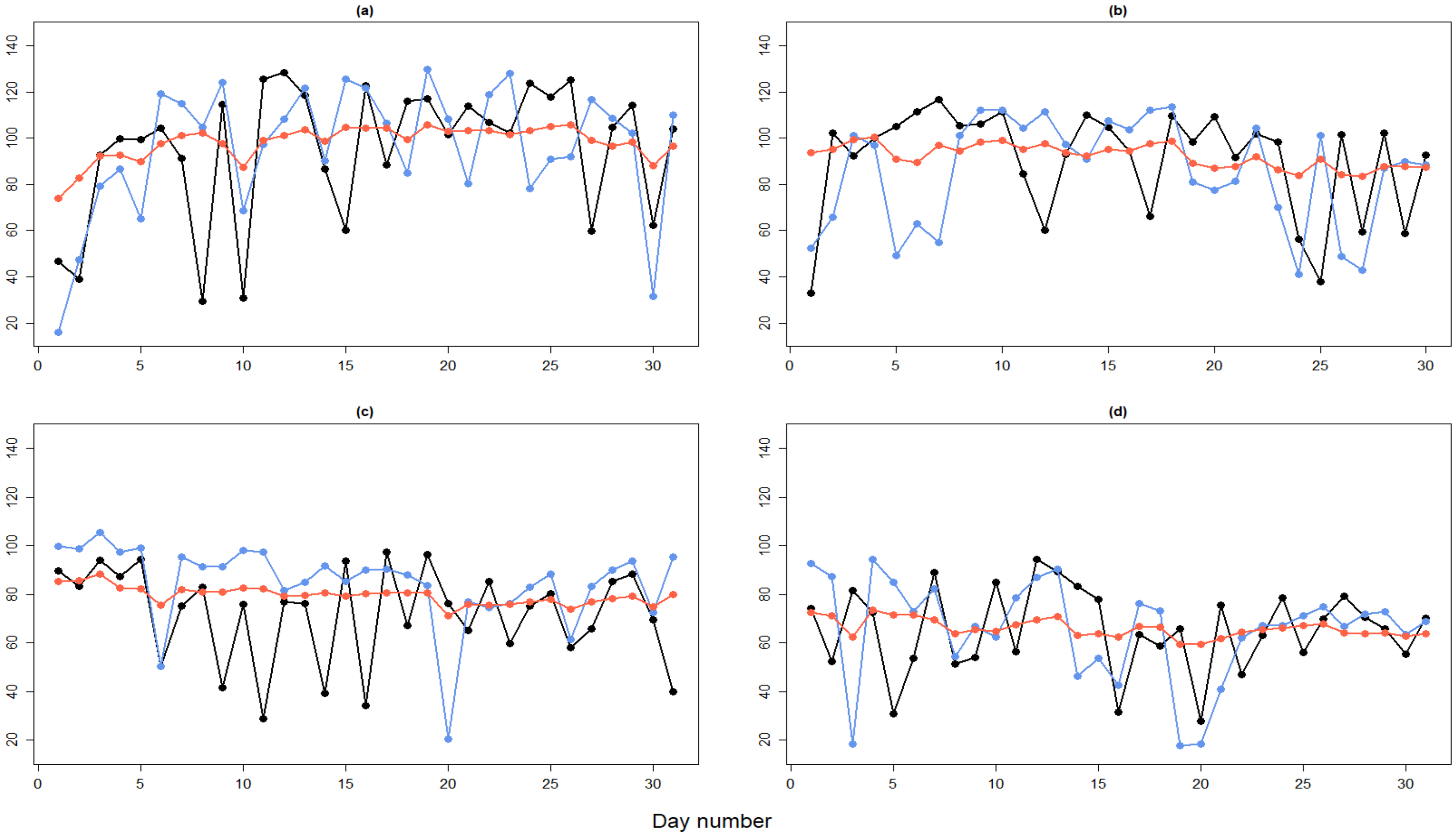
Appendix 3.1: Forecast periods 1 to 10 RPE values (%) for daily grass growth predictions from 2008 to 2016 from the MoSt model using ECMWF forecasts as weather inputs. The RMSE of these predictions with the daily grass growth predictions using weather observations is expressed as a percentage of the mean of the daily grass growth predictions from 2008 to 2016 from the MoSt model using weather observations. Numbers in the top row indicate the fertiliser application level (kg N/ha) used for the predictions.

Forecast period	N application (kg/ha)			
	0	100	200	300
1	28.5	27.4	27.3	27.4
2	32.5	30.8	30.5	30.5
3	36.4	35.1	34.9	34.8
4	43.0	40.9	40.5	40.3
5	47.8	45.9	45.5	45.4
6	51.8	49.8	48.5	48.8
7	57.8	55.1	54.3	54.0
8	57.7	55.7	54.7	55.1
9	63.3	61.0	60.3	60.1
10	63.5	60.6	60.1	59.6

Appendix 3.2: Predicted daily grass growth in 2015 for (a) May, (b) June, (c) July and (d) August from the MoSt GGM using weather observations (black), day-5 ECMWF forecasts (blue) and day-5 model BC forecasts (red) with 200 kg N/ha.



Appendix 3.3: Predicted daily grass growth in 2015 for (a) May, (b) June, (c) July and (d) August from the MoSt GGM using weather observations (black), day-9 ECMWF forecasts (blue) and day-9 model BC forecasts (red) with 200 kg N/ha.



Appendix 4.1

DI model fitting in SAS and R for a simulated dataset with an identity effect structure, simple covariance structure (no covariances between year) and random pairwise interactions in 3 years.

Note: Collaborator R. de Andrade Moral contributed significantly to the R code presented in this section.

The following SAS code can be used to estimate the DI model for a simulated data set (uploaded as EF2: sim_data_R_SAS.csv) representing an experiment with 10 species, 110 plots and 3 years, totalling 330 observations. The code uses similar methods to that in Appendix 4.11, where the methods are described in detail.

```
/*IMPORT THE DATASET*/
PROC IMPORT OUT=sim_data_R_SAS
     DATAFILE="'/folders/myfolders/sim_data_R_SAS.csv"
     DBMS=CSV REPLACE;
     GETNAMES=YES;
RUN;

/* CREATING A DATA SET NEEDED FOR THE RANDOM EFFECTS
SPECIFICATION*/
data re;
      do i=1 to 45;
           parm=1;
           row=i;
           col=i;
           value=1;
           output;
      end;
      drop i;
run;
```

```
/* Identity model with random pairwise interactions REML */
proc mixed data=sim_data_R_SAS method=reml;
      class year plot;
      model y = P1*year P2*year P3*year P4*year P5*year
P6*year P7*year P8*year P9*year P10*year / noint solution;
      repeated year / subject=plot type=vc;
      random Y1PP1-Y1PP45/ type=lin(1) ldata=re solution;
      random Y2PP1-Y2PP45/ type=lin(1) ldata=re solution;
      random Y3PP1-Y3PP45/ type=lin(1) ldata=re solution;
run;
```

```
/* Identity model with random pairwise interactions ML */
proc mixed data=sim_data_R_SAS method=ml;
      class year plot;
      model y = P1*year P2*year P3*year P4*year P5*year
P6*year P7*year P8*year P9*year P10*year / noint solution;
      repeated year / subject=plot type=vc;
      random Y1PP1-Y1PP45/ type=lin(1) ldata=re solution;
      random Y2PP1-Y2PP45/ type=lin(1) ldata=re solution;
      random Y3PP1-Y3PP45/ type=lin(1) ldata=re solution;
run;
```

The following R code can be used to estimate the DI model for the simulated data set (uploaded as EF2: sim_data_R_SAS.csv) representing an experiment with 10 species, 110 plots and 3 years, totalling 330 observations. We include only identity effects by year as fixed, and random pairwise interactions with a different variance per year as the linear predictor. Note that the open-source, free R packages lme4 and nlme cannot be used here, since their syntax does not allow for the specification of the random effects without a grouping factor. Hence, we write code that optimizes the profile (restricted) log-likelihood function for the variance components.

First, we read the data into R and import packages “optimx” (Nash and Varadhan, 2011) and “bbmle” (Boler and R Development Core Team, 2016), which will be used to perform optimization:

```
sdata <- read.csv("sim_data_R_SAS.csv", h = T)
sdata$Year <- as.factor(sdata$Year)
sdata$Plot <- as.factor(sdata$Plot)
sdata$PlotYear <- with(sdata, Plot:Year)

sdata.x <- sdata
rownames(sdata.x) <- sdata.x$PlotYear
sdata <- sdata.x[levels(sdata$PlotYear),]

require(optimx)
require(bbmle)
```

We now code the X and Z matrices:

```
nPlots <- 110
nSpecies <- 10
nPairwise <- choose(nSpecies, 2)
fmlaX <- as.formula(paste("~ 0 + (", paste("P", 1:nSpecies,
                                            sep = "", collapse =
"+"),
                          "):Year"))
fmlaZ <- as.formula(paste("~ 0 + ((", paste("P", 1:nSpecies,
                                             sep = "", collapse
= "+"),
                          ")^2):Year"))

X <- model.matrix(fmlaX, data = sdata)
Z <- model.matrix(fmlaZ, data = sdata)[,-c(1:(nSpecies*3))]
```

The Z matrix assumes non-standard form for mixed models, since there are no grouping factors. We may visualise it by executing

```
require(Matrix)
image(Matrix(Z))
```

We then write functions that generate the R, G, and V matrices, with V = ZGZ’ + R the marginal variance-covariance matrix of Y:

```
G <- function(sigma.vector, Z) {
  nZ <- ncol(Z)/3
  Gmat <- diag(rep(sigma.vector, nZ))
  return(Gmat)
}

R <- function(sigmaR, y) {
  n <- length(y)
  Rmat <- diag(rep(sigmaR, n))
  return(Rmat)
}

V <- function(sigma.vector, Z, y) {
  Gmat <- G(sigma.vector[1:3], Z = Z)
  Rmat <- R(sigma.vector[4], y = y)
  Vmat <- Z%*%Gmat%*%t(Z) + Rmat
  return(Vmat)
}
```

Given V, we may estimate the fixed effects using weighted least squares estimation, with the inverse of V as the matrix of weights, coded below:

```
beta.hat <- function(y, X, Vinv) {
  XVX <- t(X)%*%Vinv%*%X
  XVXinv <- solve(XVX)
  XVy <- t(X)%*%Vinv%*%y
  bh <- XVXinv%*%XVy
  return(bh)
}
```

We then write the profile (restricted) log-likelihood function to be maximized to obtain the estimates for the variance components (note that we multiply the function by -1, because the algorithms used are minimizers). We include the argument REML to choose between ML (REML = FALSE) and REML (REML = TRUE) estimation. Also note that we apply exponentials to the variance components, to bound the parameter space (repar = TRUE, the default in the function):

```
profloglik <- function(sigma2G1, sigma2G2, sigma2G3, sigma2R,
                       y, X, Z, REML = FALSE, repar = TRUE) {
  if(repar) {
    sig.vec <- exp(c(sigma2G1, sigma2G2, sigma2G3, sigma2R))
  } else {
    sig.vec <- c(sigma2G1, sigma2G2, sigma2G3, sigma2R)
```

```
  }
  V.hat <- V(sigma.vector = sig.vec, Z = Z, y = y)
  V.hat.inv <- solve(V.hat)
  betas <- beta.hat(y = y, X = X, Vinv = V.hat.inv)
  raw.res <- y - X%*%betas
  llik <- -.5*(determinant(V.hat)$modulus +
t(raw.res)%*%V.hat.inv%*%raw.res)
  if(REML) llik <- llik -
.5*determinant(t(X)%*%V.hat.inv%*%X)$modulus
  minusloglik <- as.numeric(-llik)
  return(minusloglik)
}
```

We use exp(10) as an initial value for all variance components and the bobyqa method, since we do not provide a gradient function, and because we have reparameterized the variance components, estimation is now unbounded. Note that depending on the initial values and the number of fixed effects that are included in the model, the algorithm may crash, and this is due to numerical instability that may require further investigation. We estimate the model using ML (fitML) and REML (fitREML):

```
init <- c(10,10,10,10)
names(init) <- c("sigma2G1","sigma2G2","sigma2G3","sigma2R")

fitML <- mle2(profloglik, start = as.list(init),
              optimizer = "optimx", method = "bobyqa",
              data = list(y = sdata$y,
                          X = X, Z = Z, REML = FALSE))

fitREML <- mle2(profloglik, start = as.list(init),
                optimizer = "optimx", method = "bobyqa",
                data = list(y = sdata$y,
                            X = X, Z = Z, REML = TRUE))
```

On a 4GB RAM with an Intel Core i5 processor machine, these model fitting procedures take less than 15 seconds each. We now may compile the estimates for the fixed effects, variance components, and predictions for the random effects for further inspection:

```
# Estimated variance components
sigML <- exp(fitML@coef)
sigREML <- exp(fitREML@coef)

# Estimated fixed effects
betasML <- beta.hat(y = sdata$y, X = X,
                    Vinv = solve(V(sigma.vector = sigML,
                                   Z = Z, y = sdata$y)))
betasREML <- beta.hat(y = sdata$y, X = X,
                      Vinv = solve(V(sigma.vector = sigREML,
                                     Z = Z, y = sdata$y)))

# Estimated marginal variance-covariance matrix (V)
G.hatML <- G(sigma.vector = sigML[1:3], Z = Z)
G.hatREML <- G(sigma.vector = sigREML[1:3], Z = Z)
V.hatML <- V(sigma.vector = sigML, Z = Z, y = sdata$y)
V.hatREML <- V(sigma.vector = sigREML, Z = Z, y = sdata$y)

# Predicted random effects
ranefsML <- G.hatML%*%t(Z) %*% solve(V.hatML) %*% (sdata$y -
X%*%betasML)
ranefsREML <- G.hatREML%*%t(Z) %*% solve(V.hatREML) %*%
(sdata$y - X%*%betasREML)

# Compilation of estimates and predictions
ranefs.rnames <- NULL
for(i in 1:45){
  ranefs.rnames <- c(ranefs.rnames, paste("PP", i, ":year1",
sep = ""),
                     paste("PP", i, ":year2", sep = ""),
                     paste("PP", i, ":year3", sep = ""))
}
ranefs <- data.frame(ranefsML, ranefsREML,
                     row.names = ranefs.rnames)
estimates <- data.frame("ML" = c(betasML, sigML),
                        "REML" = c(betasREML, sigREML),
                        row.names = c(paste("P", 1:10,
":year1", sep = ""),
                                      paste("P", 1:10,
":year2", sep = ""),
                                      paste("P", 1:10,
":year3", sep = ""),

"sigma2G1","sigma2G2","sigma2G3","sigma2R"))
round(ranefs, 2)
round(estimates, 2)
```

Object `ranefs` compiles the predicted random effects for each of the 45 pairwise interactions for each year. Object `estimates` compiles the estimated fixed effects (for this example only the identity effects by year).

References

Nash, J.C. and Varadhan, R. (2011) Unifying optimization algorithms to aid software system users: optimx for R. *Journal of Statistical Software* 43(9), 1-14.

Bolker, B. and R Development Core Team (2016) bbmle: Tools for general maximum likelihood estimation. R package version 1.0.18. https://CRAN.R-project.org/package=bbmle

Appendix 4.2

SAS code to conduct the simulation study to demonstrate detection of random pairwise interactions. The code to be changed to change the error standard deviation (100 in this case) and random pairwise interactions standard deviation (2500 in this case) are in bold. We conduct each of the 18 sets of simulations individually due to memory constraints with SAS University Edition.

```
/*IMPORT THE BASE EXPERIMENTAL DESIGN*/
PROC IMPORT OUT= data
DATAFILE="'/folders/myfolders/Experimental_design.csv"
     DBMS=CSV REPLACE;
     GETNAMES=YES;
RUN;

/*REPEAT THE EXPERIMENTAL DESIGN OVER THREE YEARS*/
data data;
      set data (in=INA) data (in=INB) data (in=INC);
      if INA then Year=1;
      if INB then Year=2;
      if INC then Year=3;
run;

/* REPEAT THE DATASET 1000 TIMES */
data data_1000;
      set data;
      do Count = 1 to 1000;
                output;
      end;
run;
```

```
/*SET UP IDENTITY EFFECTS*/
data id;
      input year b1 b2 b3 b4 b5 b6 b7 b8 b9 b10;
      cards;
1 700 472 512 650 440 343 630 709 699 584
2 1477 1458 1636 1267 879 759 931 874 1100 845
3 1103 1411 1601 604 753 792 627 637 752 748
run;

/* REPEAT THE IDENTITY EFFECTS 1000 TIMES */
data id_1000;
      set id;
      do Count = 1 to 1000;
                  output;
      end;
run;

/* GENERATE THE FUNCTIONAL GROUP AND RANDOM PAIRWISE
INTERACTIONS IN EACH YEAR */
data interactions1(keep= year int_total1-int_total45);
            array int_Y1 {45} int_total1-int_total45;
                  int_wfg1=-2400;
                  int_wfg2=1753;
                  int_bfg12=841;
      do l = 1 to 110;
      do i=1 to 4, 10 to 12, 18 to 19, 25;
                        int_Y1{i} = int_wfg1 ;
                  end;
      do i=5 to 9, 13 to 17, 20 to 24, 26 to 35;
                        int_Y1{i} = int_bfg12 ;
                  end;
      do i=36 to 45;
                        int_Y1{i} = int_wfg2 ;
```

```
                  end;

             Year=1;
             output;
             end;
drop i l;
run;

data interactions2(keep= year int_total1-int_total45);
             array int_Y1 {45} int_total1-int_total45;
                  int_wfg1=-2345;
                  int_wfg2=-2151;
                  int_bfg12=-128;
      do l = 1 to 110;
      do i=1 to 4, 10 to 12, 18 to 19, 25;
                       int_Y1{i} = int_wfg1 ;
                  end;
      do i=5 to 9, 13 to 17, 20 to 24, 26 to 35;
                       int_Y1{i} = int_bfg12 ;
                  end;
      do i=36 to 45;
                       int_Y1{i} = int_wfg2 ;
                  end;

             Year=2;
             output;
             end;
drop i l;
run;



data interactions3(keep= year int_total1-int_total45);
             array int_Y1 {45} int_total1-int_total45;
                  int_wfg1=4592;
```

```
                int_wfg2=-3954;
                int_bfg12=524;
      do l = 1 to 110;
      do i=1 to 4, 10 to 12, 18 to 19, 25;
                      int_Y1{i} = int_wfg1 ;
                end;
      do i=5 to 9, 13 to 17, 20 to 24, 26 to 35;
                      int_Y1{i} = int_bfg12 ;
                end;
      do i=36 to 45;
                      int_Y1{i} = int_wfg2 ;
                end;

           Year=3;
           output;
           end;
drop i l;
run;



/* PUT THE THREE DATASETS CONTAINING THE THREE YEARS OF
INTERACTIONS TOGETHER */
Data interactions;
      set interactions1 interactions2 interactions3;
      by Year;
run;

data interactions_1000;
      set interactions;
      do Count = 1 to 1000;
                output;
      end;
run;
```

```
/* GENERATE THE FUNCTIONAL GROUP AND RANDOM PAIRWISE
INTERACTIONS IN EACH YEAR */

data interactions1_random_2500(keep= count year random_int1-
random_int45);

           array int_Y1 {45} random_int1-random_int45;

           do j=1 to 1000;

      do i=1 to 4, 10 to 12, 18 to 19, 25;

                      x = rand("Normal");

                      int_Y1{i} = x*2500;

                end;

      do i=5 to 9, 13 to 17, 20 to 24, 26 to 35;

                           y = rand("Normal");

                      int_Y1{i} = y*2500;

                end;

      do i=36 to 45;

                           z = rand("Normal");

                      int_Y1{i} = z*2500;

                end;

           Year=1;

           Count=j;

           output;

           end;

drop i j;

run;


/* GENERATE THE FUNCTIONAL GROUP AND RANDOM PAIRWISE
INTERACTIONS IN EACH YEAR */

data interactions2_random_2500(keep= count year random_int1-
random_int45);

           array int_Y1 {45} random_int1-random_int45;

           do j=1 to 1000;

      do i=1 to 4, 10 to 12, 18 to 19, 25;

                      x = rand("Normal");

                      int_Y1{i} = x*2500;

                end;

      do i=5 to 9, 13 to 17, 20 to 24, 26 to 35;
```

```
                           y = rand("Normal");
                      int_Y1{i} = y*2500;
                 end;
      do i=36 to 45;
                           z = rand("Normal");
                      int_Y1{i} = z*2500;
                 end;
           Year=2;
           Count=j;
           output;
           end;
drop i j;
run;


/* GENERATE THE FUNCTIONAL GROUP AND RANDOM PAIRWISE
INTERACTIONS IN EACH YEAR */
data interactions3_random_2500(keep= count year random_int1-
random_int45);
           array int_Y1 {45} random_int1-random_int45;
           do j=1 to 1000;
      do i=1 to 4, 10 to 12, 18 to 19, 25;
                      x = rand("Normal");
                      int_Y1{i} = x*2500;
                 end;
      do i=5 to 9, 13 to 17, 20 to 24, 26 to 35;
                           y = rand("Normal");
                      int_Y1{i} = y*2500;
                 end;
      do i=36 to 45;
                           z = rand("Normal");
                      int_Y1{i} = z*2500;
                 end;
           Year=3;
           Count=j;
           output;
```

```
            end;
drop i j;
run;


/* PUT THE THREE DATASETS CONTAINING THE THREE YEARS OF RANDOM
INTERACTIONS TOGETHER */
data interactions_random_2500;
      set interactions1_random_2500 interactions2_random_2500
interactions3_random_2500;
      by Year;
run;


/* REPEAT THE RANDOM INTERACTIONS AT EACH PLOT WITHIN EACH OF
THE 1000 SIMULATIONS */
data interactions_random_1000_2500;
      set interactions_random_2500;
      do Rep = 1 to 110;
                  output;
      end;
run;


/* SORT THE RELEVANT DATASETS */
proc sort data=data_1000;
      by Count Year;
run;
proc sort data=id_1000;
      by Count Year;
run;
proc sort data=interactions_1000;
      by Count Year;
run;
proc sort data=interactions_random_1000_2500;
      by Count Year;
run;
```

```
/* MERGE THE DATASETS TO ALLOW COMPUTING THE RESPONSE
VARIABLE*/
data data_1000_2500;
      merge data_1000 id_1000 interactions_1000
interactions_random_1000_2500;
      by Count Year;
run;

/* COMPUTE THE PIPJ PAIRWISE INTERACTION TERMS */
data data_1000_2500;
      set data_1000_2500;
      array p {10} p1-p10;
      array pp {45};
      l=0;
            do i=1 to 9;
                  do j=(i+1) to 10;
                  l=l+1;
                  pp{l}=p{i}*p{j};
                  end;
            end;
      drop i j l;
run;

/* GENERATE THE IDENTITY EFFECT FOR EACH PLOT */
data data_1000_2500;
      set data_1000_2500;
      array b {10} b1-b10;
      array p {10} p1-p10;
      id=0;
            do i = 1 to 10;
                  id = id + b{i}*p{i};
            end;
      drop i;
run;
```

```
/* GENERATE THE DIVERSITY EFFECT FOR EACH PLOT AND COMPUTE THE
SUM OF THE PIPJ TERMS*/

data data_1000_2500;

      set data_1000_2500;

      array int {45} int_total1-int_total45;

      array r_int {45} random_int1-random_int45;

      array pp {45} pp1-pp45;

      DE=0;

      PiPj_sum=0;

            do i = 1 to 45;

                  DE = DE + int{i}*pp{i} + r_int{i}*pp{i};

                  PiPj_sum = PiPj_sum + pp{i};

            end;

      drop i;

run;



/* CREATING A DATA SET NEEDED FOR THE RANDOM EFFECTS
SPECIFICATION*/

data re;

      do i=1 to 45;

            parm=1;

            row=i;

            col=i;

            value=1;

            output;

      end;

      drop i;

run;


/* REPLICATING AND SORTING THE DATA SET NEEDED FOR THE RANDOM
EFFECTS SPECIFICATION*/

data re_1000;

      set re;

      do Count = 1 to 1000;
```

```
                  output;
      end;
run;
proc sort data=re_1000;
      by Count;
run;

/* ADD IN THE RESIDUALS */
data data_1000_2500_100;
      set data_1000_2500;
      call streaminit(320567);    /* set seed */
      e1 = rand("Normal");        /* e1 ~ N(0,1) */
      e2 = rand("Normal");
      e3 = rand("Normal");
      if year=1 then   y = ID + DE + e1*100;
      if year=2 then   y = ID + DE + e2*100;
      if year=3 then   y = ID + DE + e3*100;
run;

/*************************************************************/
/* SET UP THE DATASET SO THAT THE DI MODELS CAN BE FITTED */
Data data_1000_2500_100;
      SET data_1000_2500_100;
            /*    ARRAY 1: CONTAINS THE SPECIES PROPORTIONS P1
TO P10 */
            array A1 (10) P1-P10;
            /*    ARRAY 2: CONTAINS THE PAIRWISE SPECIES
INTERACTIONS PP1-PP45
                  PPsum IS THE SUM OF PP1-PP45 */
            k=0;
            array A2 (45) PP1 - PP45;
            PPsum=0;
            do i=1 to 9;
                  do j=(i+1) to 10;
                        k=k+1;
```

```
                        A2{k}=A1{i}*A1{j};
                        PPsum = PPsum + A2{k};
                  end;
            end;

            /* CREATING THE FUNCTIONAL GROUP SUMMED PiPj VALUES
*/
            /*SPECIES 1-5 ARE FG1, 6-10 ARE FG2*/
            PPwfg1=0;
            do i=1 to 4;
                  do j=(i+1) to 5;
                        PPwfg1 = PPwfg1 + A1{i}*A1{j};
                  end;
            end;
            PPwfg2=0;
            do i=6 to 9;
                  do j=(i+1) to 10;
                        PPwfg2 = PPwfg2 + A1{i}*A1{j};
                  end;
            end;
            PPbfg12=0;
            do i=1 to 5;
                  do j=6 to 10;
                        PPbfg12 = PPbfg12 + A1{i}*A1{j};
                  end;
            end;
      drop i j k;
run;

/* SETTING UP SEPARATE PAIRWISE INTERACTIONS FOR EACH YEAR
WITH ZEROS FOR OTHER YEARS*/
data data_1000_2500_100;
      set data_1000_2500_100;
            array PP {45} PP1 - PP45;
            array PP_y1 {45} Y1PP1 - Y1PP45;
```

```
            array PP_y2 {45} Y2PP1 - Y2PP45;
            array PP_y3 {45} Y3PP1 - Y3PP45;
                  if year=1
                        then do i=1 to 45;
                        PP_y1{i}=PP{i};
                        PP_y2{i}=0;
                        PP_y3{i}=0;
                        end;
                  if year=2
                        then do i=1 to 45;
                        PP_y1{i}=0;
                        PP_y2{i}=PP{i};
                        PP_y3{i}=0;
                        end;
                  if year=3
                        then do i=1 to 45;
                        PP_y1{i}=0;
                        PP_y2{i}=0;
                        PP_y3{i}=PP{i};
                        end;
      drop i;
run;

/*************************************************************/
               /* FIT THE DI MODELS TO THE DATA */
/*************************************************************/


/* NOW FIND THE BEST BASELINE MODEL AND SEE IF THE RANDOM
PAIRWISE INTERACTIONS ARE NEEDED IN THIS BASELINE MODEL */


/* Set up datasets to store output from models */
/* Want to store the log likelihood and AIC values */
data fits_1000_2500_100;
input Value Model $;
```

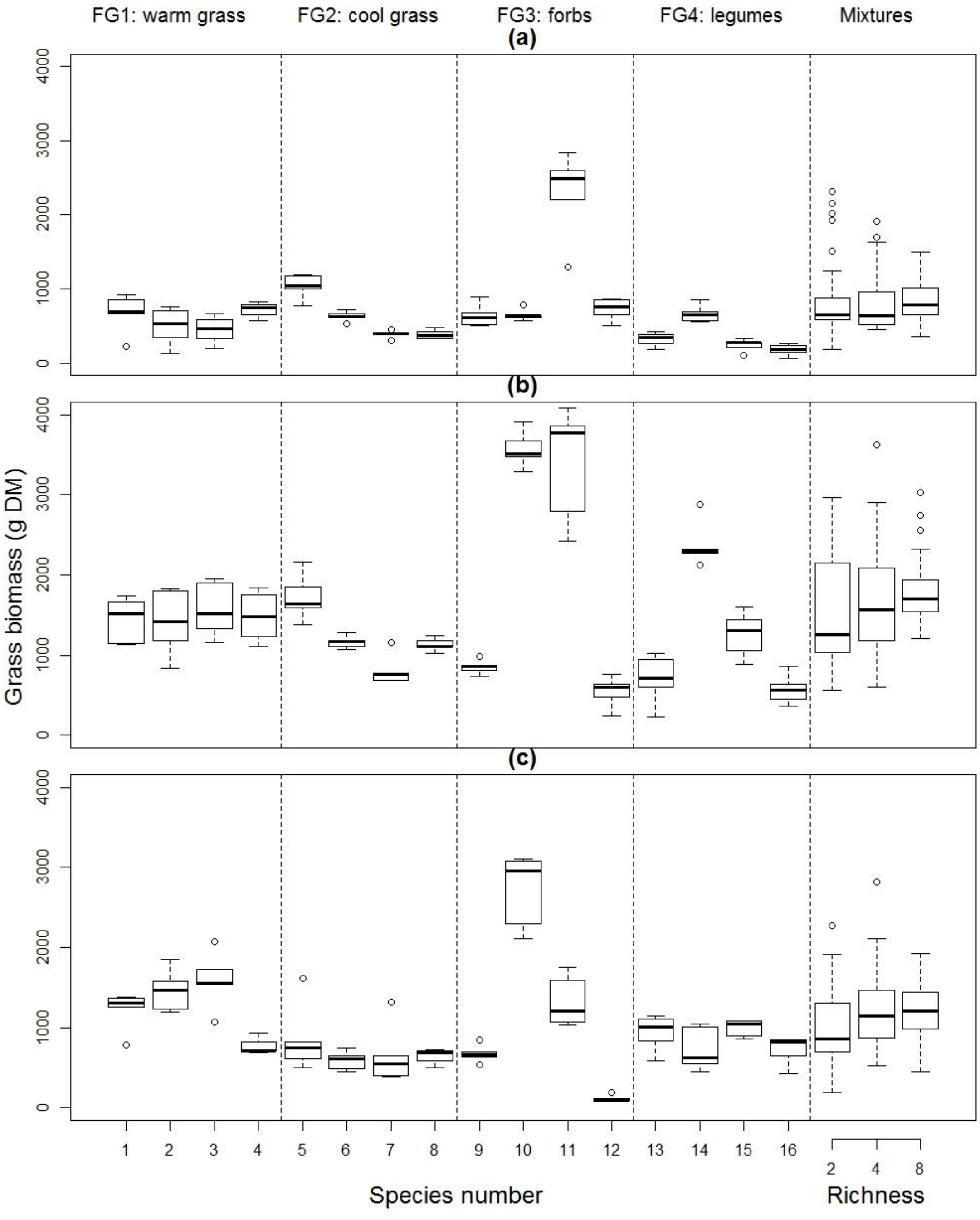
```
cards;

run;


/* Want to store the covariance and lin(1) (variance of the
random pairwise interactions) values */

data covs_1000_2500_100;

input Value Model $;

cards;

run;



/* Functional Group Model without RE */

proc mixed data=data_1000_2500_100 method=reml;

where count <= 1000;

      by count;

      class year plot;

      model Y = P1*year P2*year P3*year P4*year P5*year
P6*year P7*year P8*year P9*year P10*year PPwfg1*year
PPwfg2*year PPbfg12*year / noint solution;

      repeated year / subject=plot type=un;

      ods output FitStatistics=fits;

      ods output CovParms=covs;

      ods output SolutionF=sol_1000_2500_100_fgwo;

run;


data fits;

set fits;

Model='FGWO';

run;

data covs;

set covs;

Model='FGWO';

run;


data fits_1000_2500_100;
```

```
set fits_1000_2500_100 fits;
run;
data covs_1000_2500_100;
set covs_1000_2500_100 covs;
run;

/* Functional Group Model with RE in Y1 */
proc mixed data=data_1000_2500_100 method=reml;
where count <= 1000;
      by count;
      class year plot;
      model Y = P1*year P2*year P3*year P4*year P5*year
P6*year P7*year P8*year P9*year P10*year PPwfg1*year
PPwfg2*year PPbfg12*year / noint solution;
      repeated year / subject=plot type=un;
      random Y1PP1-Y1PP45/ type=lin(1) ldata=re_1000 solution;
      ods output FitStatistics=fits;
      ods output CovParms=covs;
      ods output SolutionF=sol_1000_2500_100_fgy1;
run;

data fits;
set fits;
Model='FGY1';
run;
data covs;
set covs;
Model='FGY1';
run;

data fits_1000_2500_100;
set fits_1000_2500_100 fits;
run;
data covs_1000_2500_100;
set covs_1000_2500_100 covs;
```

```
run;

/* Functional Group Model with RE in Y2 */
proc mixed data=data_1000_2500_100 method=reml;
where count <= 1000;
      by count;
      class year plot;
      model Y = P1*year P2*year P3*year P4*year P5*year
P6*year P7*year P8*year P9*year P10*year PPwfg1*year
PPwfg2*year PPbfg12*year / noint solution;
      repeated year / subject=plot type=un;
      random Y2PP1-Y2PP45/ type=lin(1) ldata=re_1000 solution;
      ods output FitStatistics=fits;
      ods output CovParms=covs;
      ods output SolutionF=sol_1000_2500_100_fgy2;
run;

data fits;
set fits;
Model='FGY2';
run;
data covs;
set covs;
Model='FGY2';
run;

data fits_1000_2500_100;
set fits_1000_2500_100 fits;
run;
data covs_1000_2500_100;
set covs_1000_2500_100 covs;
run;

/* Functional Group Model with RE in Y3 */
proc mixed data=data_1000_2500_100 method=reml;
```

```
where count <= 1000;
      by count;
      class year plot;
      model Y = P1*year P2*year P3*year P4*year P5*year
P6*year P7*year P8*year P9*year P10*year PPwfg1*year
PPwfg2*year PPbfg12*year / noint solution;
      repeated year / subject=plot type=un;
      random Y3PP1-Y3PP45/ type=lin(1) ldata=re_1000 solution;
      ods output FitStatistics=fits;
      ods output CovParms=covs;
      ods output SolutionF=sol_1000_2500_100_fgy3;
run;

data fits;
set fits;
Model='FGY3';
run;
data covs;
set covs;
Model='FGY3';
run;

data fits_1000_2500_100;
set fits_1000_2500_100 fits;
run;
data covs_1000_2500_100;
set covs_1000_2500_100 covs;
run;
```

Appendix 4.3: Total plot grass biomass (g DM in 1 m^2 plot) in (a) 2012, (b) 2013, (c) 2014 for each monoculture (1 to 16) and each level of richness in mixture (2, 4 or 8 species). The 4 functional groups and the mixtures are separated by the dotted lines.



Appendix 4.4: Details of step 2 of the model fitting process. Various structures to account for repeated measures and the inclusion of variance-covariance blocks to account for inhomogeneous variance across plots are tested. These models were fitted using restricted maximum likelihood (REML). The fixed effects model selected in step 1 was the additive species model. Differences in AIC are from model 9, the best overall model. The likelihood ratio test (LRT) statistic and p-value are also presented.

Model	Repeated measures	Variance-covariance block structure	No. parameters	Δ AIC	Model comparison	LRT stat	P-value
1	CS	Constant across all plots	114	377.9			
2	CS	Different blocks for mixtures and monocultures	116	378.4	2 v 1	3.5	0.174
3	CS	Different blocks for mixtures and for each FG in monoculture	122	338.8	3 v 1	55.1	<0.001
4	AR(1)	Constant across all plots	114	377.2			
5	AR(1)	Different blocks for mixtures and monocultures	116	378.7	5 v 4	2.5	0.287
6	AR(1)	Different blocks for mixtures and for each FG in monoculture	122	339.7	6 v 4	53.5	<0.001
7	UN	Constant across all plots	118	84.0			
8	UN	Different blocks for mixtures and monocultures	124	108.0	8 v 7	71	<0.001
9	UN	Different blocks for mixtures and for each FG in monoculture	142	0	9 v 8	144	<0.001

Appendix 4.5: Estimates (Est) of the coefficients for the fixed effects of the final model with standard errors (SE) and P-values. The model was fitted with REML.

	Year								
	2012			2013			2014		
Effect	Est	SE	P-value	Est	SE	P-value	Est	SE	P-value
P1	68.0	19.26	0.0005	10.6	3.37	0.0018	8.9	8.89	0.3165
P2	65.0	19.18	0.0008	12.2	3.50	0.0006	19.3	7.61	0.0117
P3	127.4	21.38	<.0001	11.8	4.25	0.0060	14.4	12.86	0.2646
P4	32.7	19.53	0.0947	9.1	3.95	0.0224	50.0	11.68	<.0001
P5	26.3	14.09	0.0627	0.6	1.98	0.7452	4.8	3.65	0.1924
P6	21.9	14.08	0.1200	-0.3	1.93	0.8781	4.8	3.43	0.1635
P7	72.9	14.71	<.0001	0.6	2.11	0.7661	4.8	3.89	0.2185
P8	62.9	14.20	<.0001	5.4	2.06	0.0092	16.1	3.82	<.0001
P9	53.1	13.25	<.0001	0.7	2.51	0.7898	6.9	5.07	0.1732
P10	43.9	12.40	0.0005	-0.5	2.48	0.8464	5.1	3.87	0.1899
P11	24.0	12.84	0.0627	2.2	1.98	0.2702	23.3	4.00	<.0001
P12	29.6	12.39	0.0173	4.9	2.11	0.0205	85.5	4.30	<.0001
P13	96.3	19.32	<.0001	47.9	9.98	<.0001	130.3	24.62	<.0001
P14	39.9	19.33	0.0398	3.7	9.97	0.7114	226.7	24.39	<.0001
P15	85.2	19.31	<.0001	51.4	9.92	<.0001	169.9	24.19	<.0001
P16	107.7	19.27	<.0001	56.5	9.97	<.0001	135.5	24.64	<.0001
Block 1	41.7	9.31	<.0001	0.8	1.47	0.5720	2.4	2.57	0.3478
Block 2	32.9	9.35	0.0005	1.1	1.45	0.4330	5.7	2.63	0.0320

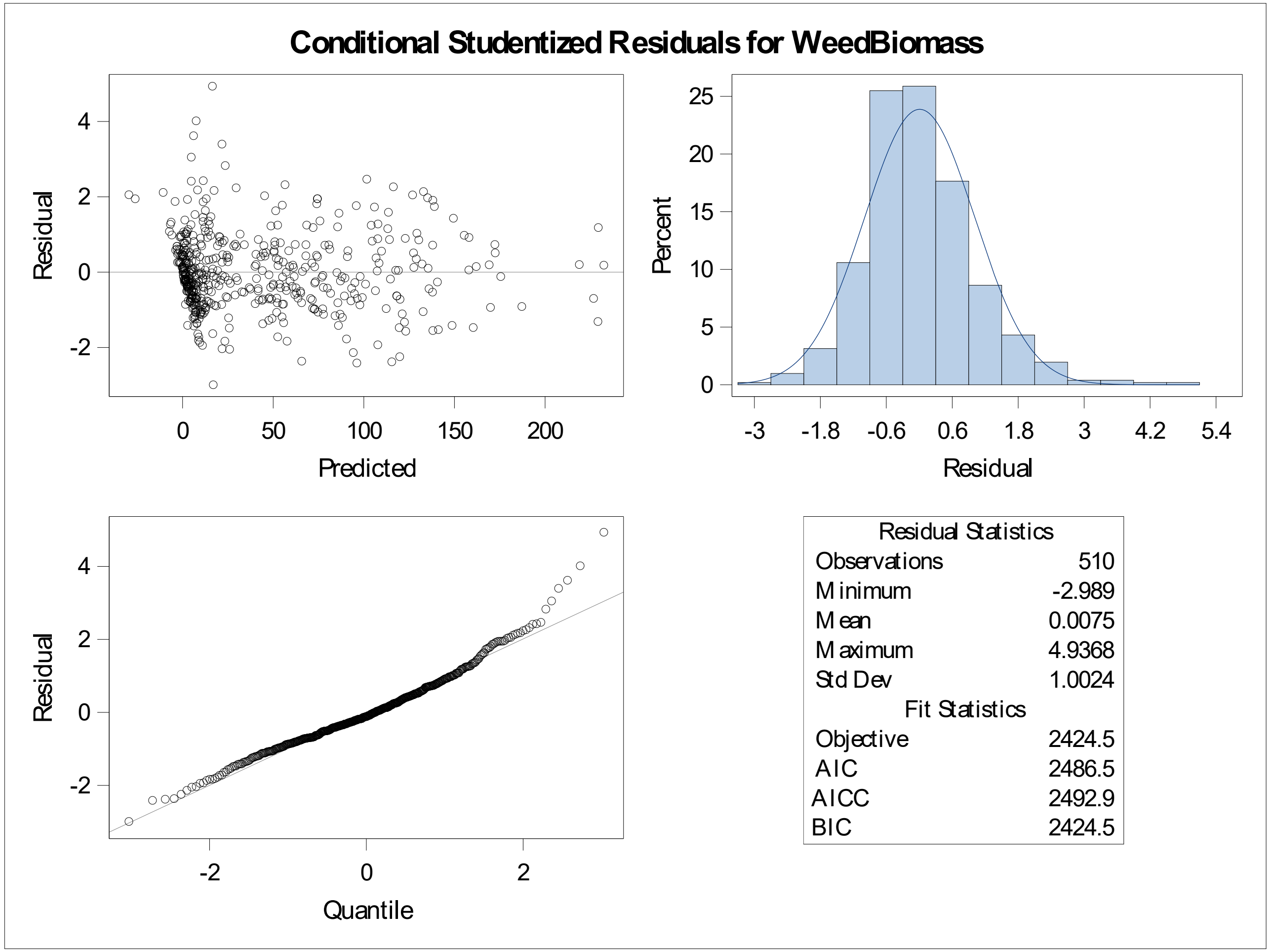
	Year								
	2012			2013			2014		
Effect	Est	SE	P-value	Est	SE	P-value	Est	SE	P-value
Block 3	21.3	9.45	0.0252	2.6	1.47	0.0817	2.6	2.68	0.3308
Block 4	7.7	9.50	0.4155	0.7	1.49	0.6554	-7.9	2.73	0.0042
Block 5	0.0	.	.	0.0	.	.	0.0	.	.
PPwfg1 Agg	5494	2813.6	0.0518	403	444.8	0.3659	523	543.6	0.3367
PPwfg1 Disp	2639	1756.6	0.1341	201	582.3	0.7297	958	467.3	0.0413
PPwfg2 Agg	-2948	1536.6	0.0560	247	362.2	0.4959	-637	540.5	0.2394
PPwfg2 Disp	-1974	1469.1	0.1802	-25	201.8	0.9017	-93	191.3	0.6260
PPwfg3 Agg	1109	1376.8	0.4214	-43	46.7	0.3521	14	93.6	0.8813
PPwfg3 Disp	1979	1449.7	0.1733	-110	63.5	0.0852	102	75.8	0.1798
PPwfg4 Agg	-1724	2676.6	0.5199	-23	1464.2	0.9874	919	1168.2	0.4322
PPwfg4 Disp	-485	3133.5	0.8770	-1526	2887.8	0.5977	-6847	2763.0	0.0138
PPbfg12 Agg	39	377.6	0.9180	-46	71.4	0.5190	-123	84.6	0.1475
PPbfg12 Disp	-890	374.5	0.0181	3	47.0	0.9528	-13	70.4	0.8496
PPbfg13 Agg	-209	183.8	0.2555	-5	15.6	0.7624	-22	38.9	0.5672
PPbfg13 Disp	115	72.2	0.1106	1	14.8	0.9611	-77	36.1	0.0328
PPbfg14 Agg	2744	1139.7	0.0167	137	171.6	0.4268	-1271	489.8	0.0099
PPbfg14 Disp	2967	1163.5	0.0113	-336	448.1	0.4539	-1802	562.0	0.0015
PPbfg23 Agg	-145	95.4	0.1297	12	12.5	0.3425	-62	26.9	0.0211
PPbfg23 Disp	-105	80.8	0.1963	-3	10.6	0.7427	-88	23.1	0.0002
PPbfg24 Agg	-6128	1910.7	0.0015	-249	212.7	0.2421	-536	149.1	0.0004
PPbfg24 Disp	-6270	2040.3	0.0023	-401	162.6	0.0143	-57	161.7	0.7242
PPbfg34 Agg	1803	2028.0	0.3747	82	75.2	0.2743	-24	108.2	0.8263
PPbfg34 Disp	2004	2089.5	0.3383	262	105.7	0.0136	-152	160.0	0.3419

	Year								
	2012			2013			2014		
Effect	Est	SE	P-value	Est	SE	P-value	Est	SE	P-value
PQ1*PLeg	-5550	2348.9	0.0188	-483	381.5	0.2062	1765	780.9	0.0245
PQ2*Pleg	-5710	2385.9	0.0173	-199	390.8	0.6109	1439	631.1	0.0233
PQ3*Pleg	-5709	2440.7	0.0200	-768	554.6	0.1674	721	661.5	0.2764
PQ4*Pleg	-5154	2369.2	0.0304	-154	407.7	0.7065	1006	643.1	0.1189
PQ5*Pleg	11254	3234.6	0.0006	447	283.3	0.1153	304	286.7	0.2894
PQ6*Pleg	11572	3297.1	0.0005	596	343.9	0.0840	294	318.8	0.3565
PQ7*Pleg	11624	3293.4	0.0005	437	302.1	0.1493	450	277.0	0.1050
PQ8*Pleg	11661	3301.7	0.0005	636	295.4	0.0321	464	261.2	0.0770
PQ9*Pleg	-4072	3540.0	0.251	-296	240.6	0.2201	137	305.9	0.6550
PQ10*Pleg	-3977	3625.2	0.2735	452	455.9	0.3227	-80	297.3	0.7875
PQ11*Pleg	-3994	3627.2	0.2717	-178	191.3	0.3519	-292	257.4	0.2569
PQ12*Pleg	-4211	3579.1	0.2403	-331	203.5	0.1049	-85	283.3	0.7638
PQ13*Pleg	637	934.0	0.4956	105	429.4	0.8073	1113	592.3	0.0612
PQ14*Pleg	330	870.5	0.7048	-18	162.9	0.9132	-539	181.2	0.0032
PQ15*Pleg	332	886.9	0.7081	113	404.9	0.7805	237	311.0	0.4465
PQ16*Pleg	648	929.6	0.4865	1718	572.7	0.0029	3100	1109.1	0.0055

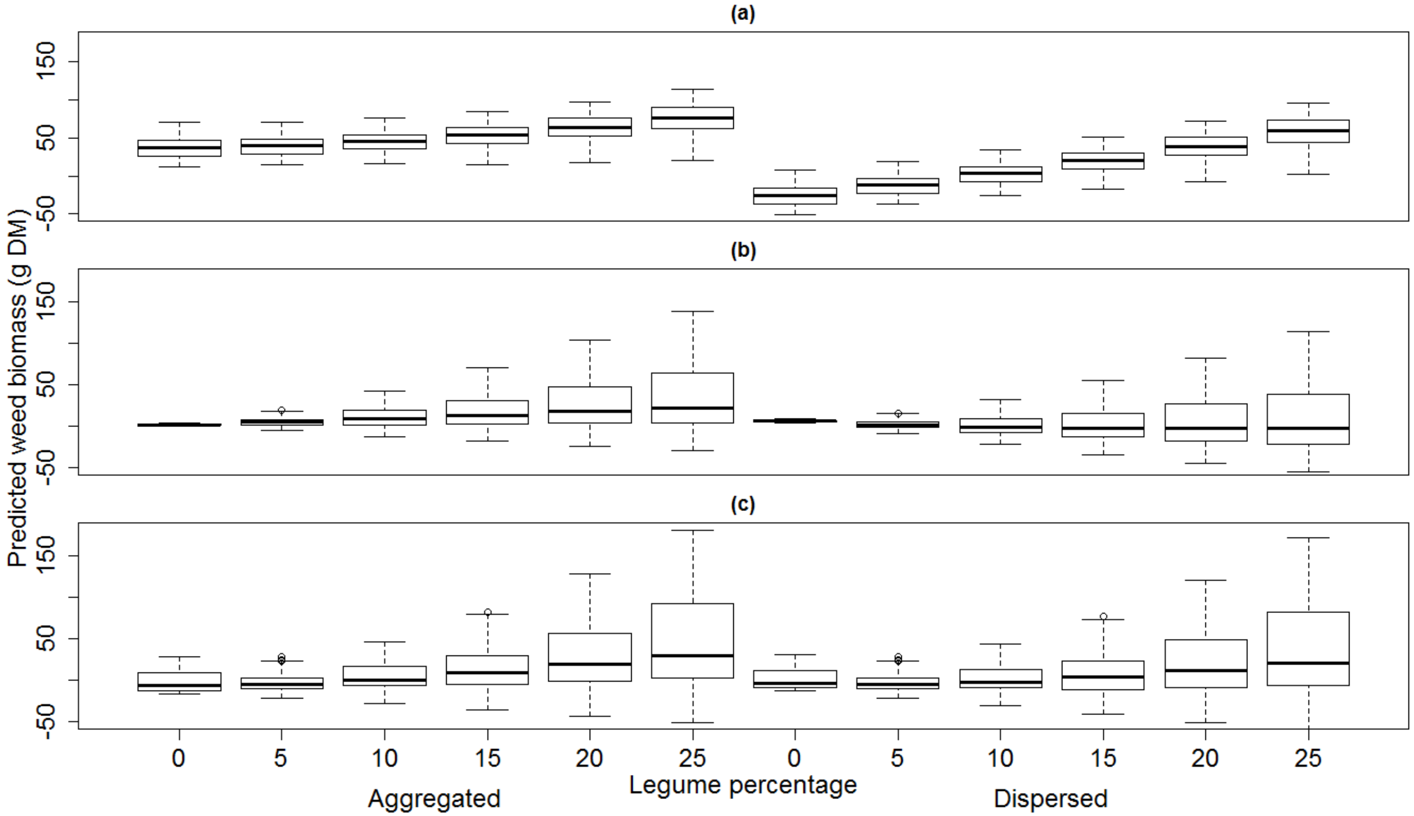
Appendix 4.6: Estimates of the variance components of the final model. Model fitted with REML. Variances are in bold. The variances of legume (FG4) monocultures is considerably higher than other groups in years 2 and 3.

		Block				
		FG1	FG2	FG3	FG4	Mixture
Row	Column	Estimate				
1	1	**2311.4**	**915.2**	**747.8**	**1690.3**	**1811.2**
2	1	174.6	-8.4	8.6	-50.2	9.6
2	2	**93.9**	**18.1**	**34.8**	**499.1**	**31.4**
3	1	435.5	-68.0	56.4	124.9	-269.9
3	2	183.0	22.4	-18.1	186.9	6.3
3	3	**974.3**	**63.3**	**130.7**	**3062.9**	**80.2**

Appendix 4.7: Plots of conditional studentized residuals for weed biomass in the final selected model.



Appendix 4.8: Predicted weed biomass (g DM in 1 m^2 plot) in (a) 2012, (b) 2013, (c) 2014 in all four-species communities for both spatial patterns (aggregated and dispersed) containing one species from each functional group at the average block level, with legume percentage taking six values between 0% and 25% with the remaining percentage divided between the other three species.



Appendix 4.9: The final chosen model for the log-transformed weed biomass.

For year k, block l, plot m, spatial pattern n, species i, j, and the warm-season grasses FG_1 = {1,2,3,4}, the cool-season grasses FG2 = {5,6,7,8}, the forbs FG3= {9,10,11,12} and the legumes FG4 = {13,14,15,16}, the final chosen model was

$$\log(y_{klmn} + 1) = \alpha_{kl} + \sum_{i=1}^{s} \beta_{ik} P_{ikm} \qquad \text{ID (including block effects)}$$

$$+ \sum_{q=1}^{4} \omega_{qqk} \sum_{\substack{i,j \in FG_q \\ i<j}} P_{ikm} P_{jkm} \sum_{i=13}^{16} P_{ikm} \qquad \text{DE}_{\text{fixed}}\text{(within functional groups, legume interaction)}$$

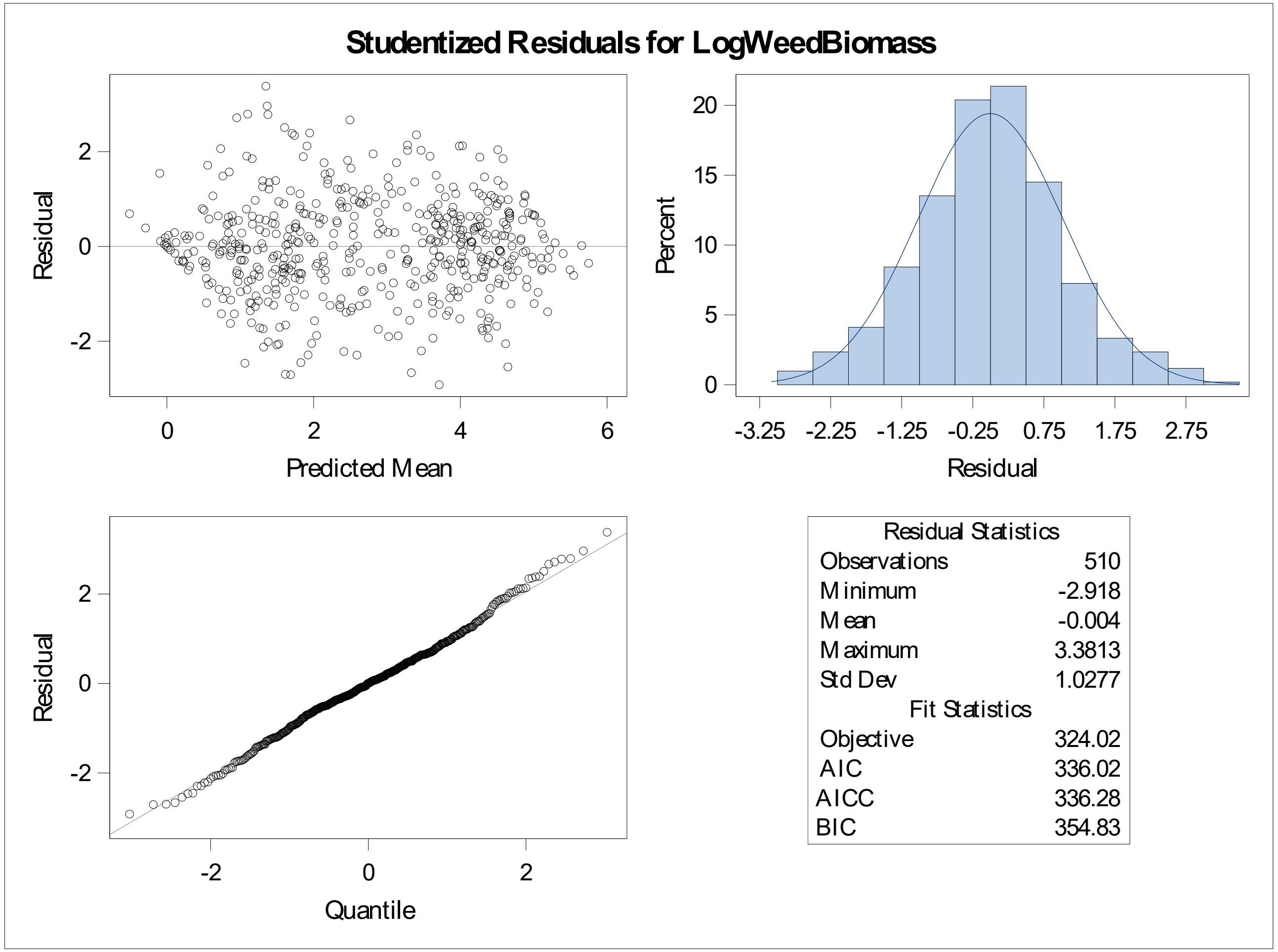
$$+ \sum_{\substack{q,r=1 \\ q<r}}^{4} \omega_{qrk} \sum_{i \in FG_q} \sum_{j \in FG_r} P_{ikm} P_{jkm} \sum_{i=13}^{16} P_{ikm} \qquad \text{DE}_{\text{fixed}}\text{(between functional group, legume interaction)}$$

$$+ \sum_{\substack{i,j=1 \\ i<j}}^{s} (\lambda_{ikn} + \lambda_{jkn}) P_{ikm} P_{jkm} \qquad \text{DE}_{\text{fixed}}\text{(additive species by spatial pattern)}$$

$$+ \varepsilon_{klmn} \qquad (7)$$

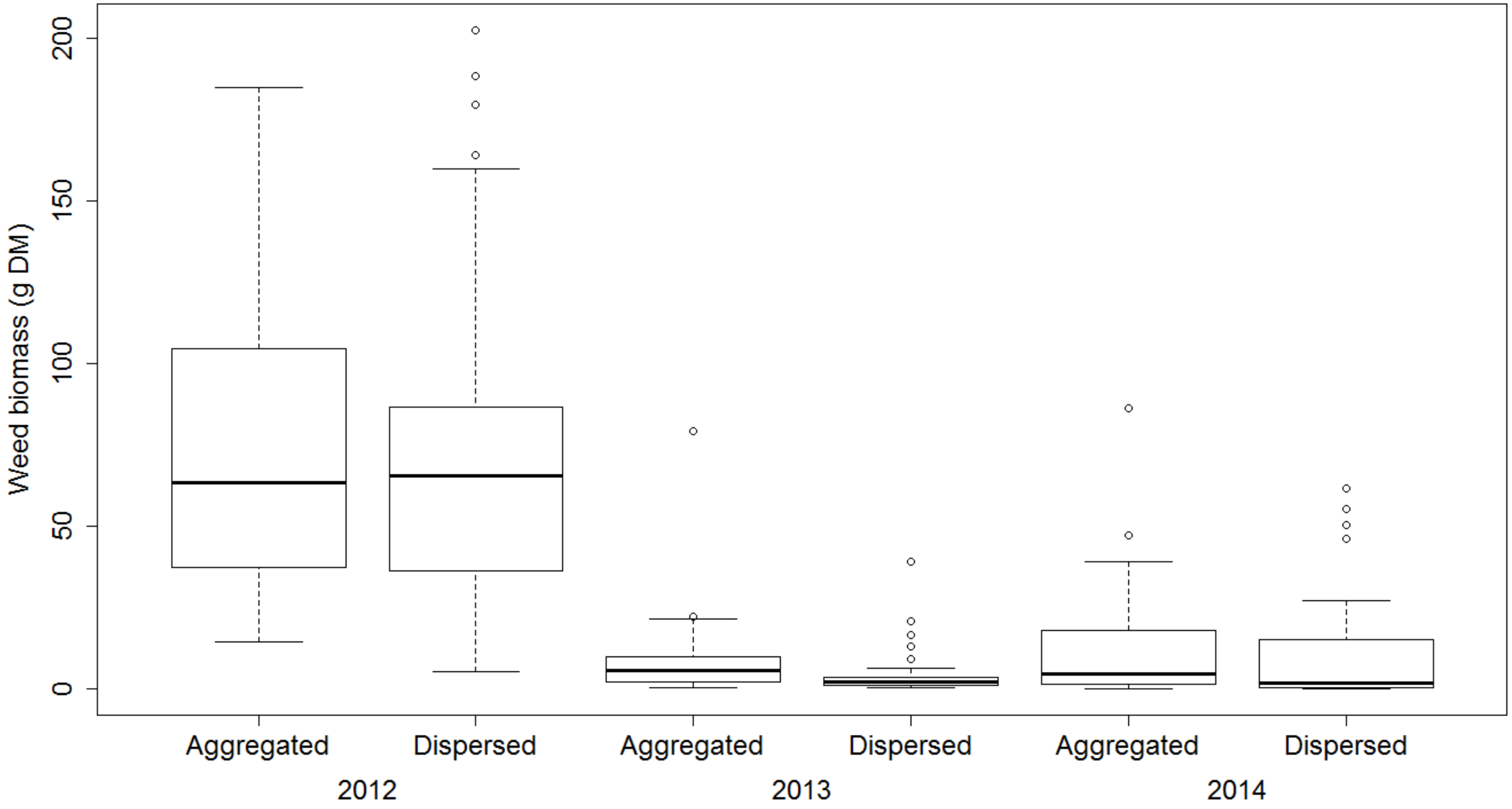
where $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{R})$. $\boldsymbol{R}$ contains the same block for each plot. The parameter α_{kl} is the effect of block l in year k, β_{ik} is the expected weed biomass of a monoculture of species i in year k, and λ_{ikn} and λ_{jkn} are the fixed additive species interaction effects of species' i and j in year k for spatial pattern n. $\sum_{i=13}^{16} P_{ikm}$ is the total legume proportion for year k, plot m . ω_{qqk} is the coefficient of the total pairwise interactions between species from functional group q in year k. ω_{qrk} is the coefficient of the total pairwise interactions between species from functional group q and r in year k. A graphical assessment of model assumptions is included in Appendix 4.10.

Appendix 4.10: Plots of studentized residuals for weed biomass in the final selected model with log weed biomass as the response.

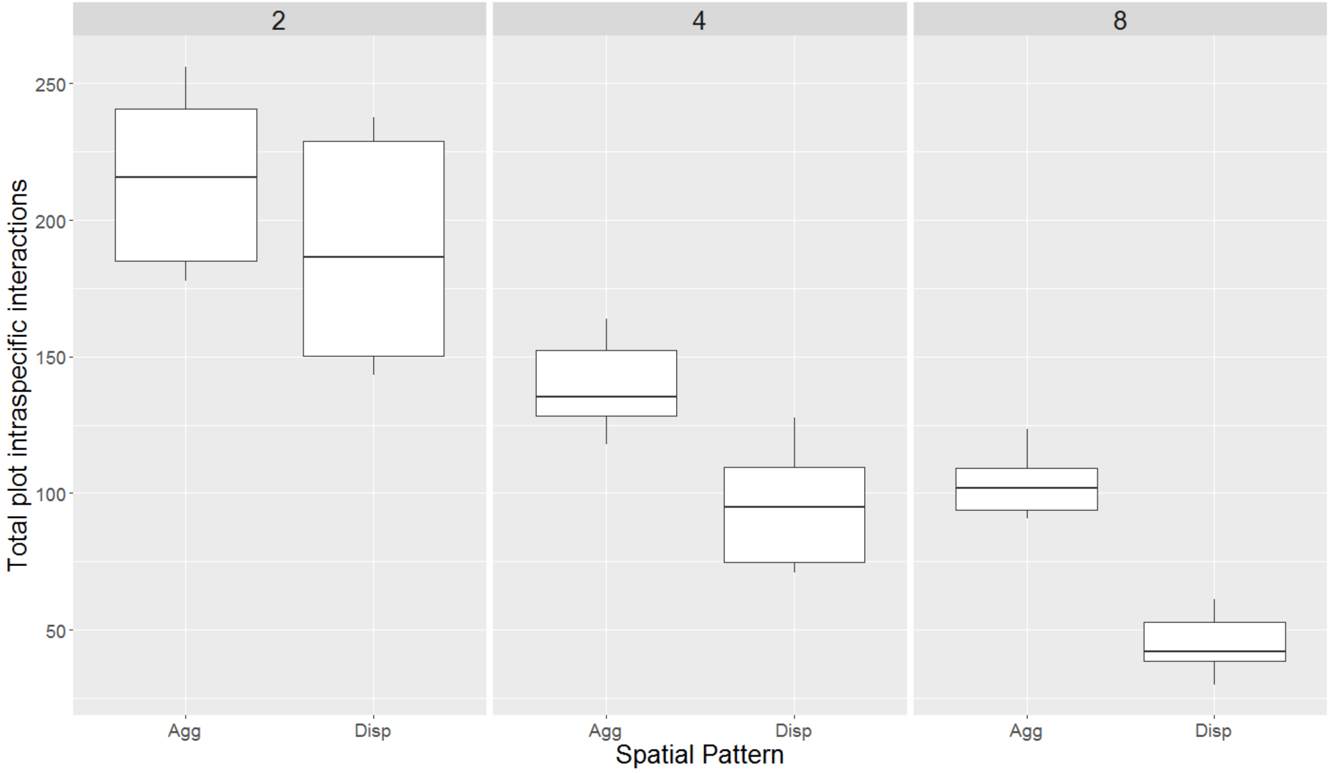


Residual Statistics	
Observations	510
Minimum	-2.918
Mean	-0.004
Maximum	3.3813
Std Dev	1.0277
Fit Statistics	
Objective	324.02
AIC	336.02
AICC	336.28
BIC	354.83

Appendix 4.11: Total plot weed biomass (g DM in 1 m^2 plot) in each year in mixture plots (plots of richness greater than one) by spatial pattern (either aggregated or dispersed).



Appendix 4.12: Total plot intraspecific interactions split by spatial pattern (aggregated or dispersed) for each level of richness (2, 4, 8). The intraspecific interaction for each species in the plot is the sum of the inverse squared distance between each of the cells containing the species, where adjacent species are a distance of 1 apart. The total plot intraspecific interaction is the sum of the intraspecific interactions for all species in the plot.



Appendix 4.13

DI model coding in SAS for a simulated dataset with a covariance structure (different blocks for monocultures and mixtures) and random pairwise interactions in 3 years.

Data were simulated to have a functional group structure with random pairwise interactions where the $\boldsymbol{A_i}$ matrices within the $\boldsymbol{R}$ matrix differed for monocultures and mixtures (monomix) in each year. The same experimental design as described in Section 4.3.2 was used, and similar identity and functional group interaction effects to those described in Table 4.2 were used. The data is in EF1: sim_data_FG_monomix.csv.

```
/*IMPORT THE DATASET*/
PROC IMPORT OUT=sim_data_FG_monomix
     DATAFILE="'/folders/myfolders/sim_data_FG_monomix.csv"
     DBMS=CSV REPLACE;
     GETNAMES=YES;
RUN;

/* CREATING A DATA SET NEEDED FOR THE RANDOM EFFECTS
SPECIFICATION (LDATA) */
data re;
     do i=1 to 45;
          parm=1;
          row=i;
          col=i;
          value=1;
          output;
     end;
```

```
      drop i;
run;


/* Functional group model with random pairwise interactions
and variance split by monocultures and mixtures */

proc mixed data=sim_data_FG_monomix method=reml;

      class year plot monomix;

      model y = P1*year P2*year P3*year P4*year P5*year
P6*year P7*year P8*year P9*year P10*year PPwfg1*year
PPwfg2*year PPbfg12*year / noint solution;

      repeated year / subject=plot type=un group=monomix;

      random Y1PP1-Y1PP45/ type=lin(1) ldata=re solution;

      random Y2PP1-Y2PP45/ type=lin(1) ldata=re solution;

      random Y3PP1-Y3PP45/ type=lin(1) ldata=re solution;

run;
```

- The class statement states which of the variables are categorical.
- The model statement specifies the response (y) and the explanatory variables, in this case a functional group model in which every term is starred with year to obtain an estimate for each identity and diversity effect in each year.
- The repeated statement gives the variable that indexes the repeated measurements (year=1,2,3) and the subject in which the repeated measurements are recorded (plot=1,...,110). Type specifies the covariance model used (in this case unstructured ie. within-subject errors for each pair of years have a unique covariance). Group specifies the groups that form the blocks of the ***R*** matrix. In this case the blocks of the ***R*** matrix differed for monocultures and mixtures in each year, but this can take other forms.
- Each of Y1PP1 - Y1PP45 is a column containing the values of the 45 pairwise interactions in year 1 at every plot in every year, meaning that in years 2 and 3 all of the Y1PP1 - Y1PP45 values are zero. Y2PP1 – Y2PP45

and Y3PP1 – Y3PP45 are similar. The ldata re is an identity matrix with 45 rows. `Type=lin(1)` allows us to fit the unusual random effects which share a common variance when we specify `ldata = re`.

The code below shows how to specify the three years of random pairwise interactions in one statement, using type=lin(3), different ldata (re2) and a different ordering of the columns containing the values of the 45 pairwise interactions in each year. However, this specification does not allow the random effects to be fitted in individual years as easily.

```
/* CREATING A DATA SET NEEDED FOR THE SPECIFICATION OF RANDOM
EFFECTS IN ONE STATEMENT */
data re2;
      do i=1 to 45;
      do j=1 to 3;
            parm=j;
            row=3*(i-1) +j;
            col=3*(i-1)+j;
            value=1;
            output;
      end;
      end;
      drop i j;
run;

/* Functional group model with random pairwise interactions
and variance split by monocultures and mixtures */
proc mixed data=sim_data_FG_monomix method=reml;
      class year plot monomix;
      model y = P1*year P2*year P3*year P4*year P5*year
P6*year P7*year P8*year P9*year P10*year PPwfg1*year
PPwfg2*year PPbfg12*year / noint solution;
```

```
     repeated year / subject=plot type=un group=monomix;

     random Y1PP1 Y2PP1 Y3PP1 Y1PP2 Y2PP2 Y3PP2 Y1PP3 Y2PP3
Y3PP3 Y1PP4 Y2PP4 Y3PP4 Y1PP5 Y2PP5 Y3PP5 Y1PP6 Y2PP6 Y3PP6
Y1PP7 Y2PP7 Y3PP7 Y1PP8 Y2PP8 Y3PP8 Y1PP9 Y2PP9 Y3PP9 Y1PP10
Y2PP10 Y3PP10 Y1PP11 Y2PP11 Y3PP11 Y1PP12 Y2PP12 Y3PP12 Y1PP13
Y2PP13 Y3PP13 Y1PP14 Y2PP14 Y3PP14 Y1PP15 Y2PP15 Y3PP15 Y1PP16
Y2PP16 Y3PP16 Y1PP17 Y2PP17 Y3PP17 Y1PP18 Y2PP18 Y3PP18 Y1PP19
Y2PP19 Y3PP19 Y1PP20 Y2PP20 Y3PP20 Y1PP21 Y2PP21 Y3PP21 Y1PP22
Y2PP22 Y3PP22 Y1PP23 Y2PP23 Y3PP23 Y1PP24 Y2PP24 Y3PP24 Y1PP25
Y2PP25 Y3PP25 Y1PP26 Y2PP26 Y3PP26 Y1PP27 Y2PP27 Y3PP27 Y1PP28
Y2PP28 Y3PP28 Y1PP29 Y2PP29 Y3PP29 Y1PP30 Y2PP30 Y3PP30 Y1PP31
Y2PP31 Y3PP31 Y1PP32 Y2PP32 Y3PP32 Y1PP33 Y2PP33 Y3PP33 Y1PP34
Y2PP34 Y3PP34 Y1PP35 Y2PP35 Y3PP35 Y1PP36 Y2PP36 Y3PP36 Y1PP37
Y2PP37 Y3PP37 Y1PP38 Y2PP38 Y3PP38 Y1PP39 Y2PP39 Y3PP39 Y1PP40
Y2PP40 Y3PP40 Y1PP41 Y2PP41 Y3PP41 Y1PP42 Y2PP42 Y3PP42 Y1PP43
Y2PP43 Y3PP43 Y1PP44 Y2PP44 Y3PP44 Y1PP45 Y2PP45 Y3PP45 /
type=lin(3) ldata=re2 solution;

run;
```

Table A4.13.1. Details of selected models fitted and P-values of the likelihood ratio test statistic (LRT), (with degrees of freedom (df = difference between the compared models), conducted to test for the inclusion of random pairwise interactions and a covariance structure with different blocks for monocultures and mixtures (monomix). Model 1 includes identity effects for each species, and functional group interaction terms. Models were fitted using restricted maximum likelihood (REML).

Model number	Model description	Test	LRT	df	P-value
1	Functional group interactions				
2	With monomix covariance structure	1 v 2	10.6	6	0.102
3	With random pairwise interactions	1 v 3	81.1	3	<0.001
4	With monomix covariance structure and random pairwise interactions	3 v 4	51.6	6	<0.001

Appendix 4.14

Code from SAS to fit the DI models to the SPaCE data as in Section 4 of manuscript.

```
/*IMPORT THE DATASET*/
PROC IMPORT OUT=space_data
     DATAFILE="'/folders/myfolders/space_data.csv"
     DBMS=CSV REPLACE;
     GETNAMES=YES;
RUN;
```

The best model with basic covariance structure fitted in ML (step 1) was the additive species model.

```
/* Additive species model */
proc mixed data=space_data method=ml;
      class block year plot;
      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*pq1 year*pq2 year*pq3 year*pq4 year*pq5 year*pq6 year*pq7
year*pq8 year*pq9 year*pq10 year*pq11 year*pq12 year*pq13
year*pq14 year*pq15 year*pq16 / noint solution;
      repeated year / subject=plot type=un;
run;
```

Fit the best model with the different variance structures: monomix (different blocks for monocultures and mixtures) and comp2 (different blocks for monocultures in each FG and mixtures). Do this using CS, AR1 and UN covariance structures (step 2).

```
/* First using CS (Compound Symmetry) */


/* Additive species model */

proc mixed data=space_data method=reml;

      class block year plot;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*pq1 year*pq2 year*pq3 year*pq4 year*pq5 year*pq6 year*pq7
year*pq8 year*pq9 year*pq10 year*pq11 year*pq12 year*pq13
year*pq14 year*pq15 year*pq16 / noint solution;

      repeated year / subject=plot type=cs;

run;


/* Additive species model with monomix var structure */

/* Different blocks for monocultures and mixtures */

proc mixed data=space_data method=reml;

      class block year plot monomix;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*pq1 year*pq2 year*pq3 year*pq4 year*pq5 year*pq6 year*pq7
year*pq8 year*pq9 year*pq10 year*pq11 year*pq12 year*pq13
year*pq14 year*pq15 year*pq16 / noint solution;

      repeated year / subject=plot type=cs group=monomix;

run;


/* Additive species model with comp2 var structure */

/* Different blocks for monocultures in each FG and mixtures
*/

proc mixed data=space_data method=reml;

      class block year plot comp2;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*pq1 year*pq2 year*pq3 year*pq4 year*pq5 year*pq6 year*pq7
year*pq8 year*pq9 year*pq10 year*pq11 year*pq12 year*pq13
year*pq14 year*pq15 year*pq16 / noint solution;

      repeated year / subject=plot type=cs group=comp2;

run;
```

```
/* Next using AR(1), first-order auto-regressive */

/* Additive species model */

proc mixed data=space_data method=reml;

      class block year plot;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*pq1 year*pq2 year*pq3 year*pq4 year*pq5 year*pq6 year*pq7
year*pq8 year*pq9 year*pq10 year*pq11 year*pq12 year*pq13
year*pq14 year*pq15 year*pq16 / noint solution;

      repeated year / subject=plot type=ar(1);

run;


/* Additive species model with monomix var structure */

/* Different blocks for monocultures and mixtures */

proc mixed data=space_data method=reml;

      class block year plot monomix;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*pq1 year*pq2 year*pq3 year*pq4 year*pq5 year*pq6 year*pq7
year*pq8 year*pq9 year*pq10 year*pq11 year*pq12 year*pq13
year*pq14 year*pq15 year*pq16 / noint solution;

      repeated year / subject=plot type=ar(1) group=monomix;

run;


/* Additive species model with comp2 var structure */

/* Different blocks for monocultures in each FG and mixtures
*/

proc mixed data=space_data method=reml;

      class block year plot comp2;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*pq1 year*pq2 year*pq3 year*pq4 year*pq5 year*pq6 year*pq7
year*pq8 year*pq9 year*pq10 year*pq11 year*pq12 year*pq13
year*pq14 year*pq15 year*pq16 / noint solution;

      repeated year / subject=plot type=ar(1) group=comp2;

run;
```

```
/* Next using UN, unstructured */

/* Additive species model */

proc mixed data=space_data method=reml;

      class block year plot;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*pq1 year*pq2 year*pq3 year*pq4 year*pq5 year*pq6 year*pq7
year*pq8 year*pq9 year*pq10 year*pq11 year*pq12 year*pq13
year*pq14 year*pq15 year*pq16 / noint solution;

      repeated year / subject=plot type=un;

run;


/* Additive species model with monomix var structure */

/* Different blocks for monocultures and mixtures */

proc mixed data=space_data method=reml;

      class block year plot monomix;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*pq1 year*pq2 year*pq3 year*pq4 year*pq5 year*pq6 year*pq7
year*pq8 year*pq9 year*pq10 year*pq11 year*pq12 year*pq13
year*pq14 year*pq15 year*pq16 / noint solution;

      repeated year / subject=plot type=un group=monomix;

run;


/* Additive species model with comp2 var structure */

/* Different blocks for monocultures in each FG and mixtures
*/

proc mixed data=space_data method=reml;

      class block year plot comp2;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*pq1 year*pq2 year*pq3 year*pq4 year*pq5 year*pq6 year*pq7
year*pq8 year*pq9 year*pq10 year*pq11 year*pq12 year*pq13
year*pq14 year*pq15 year*pq16 / noint solution;

      repeated year / subject=plot type=un group=comp2;

run;
```

Fit fixed effects models with the best covariance structure (comp2) (step3).

```
/* Average pairwise model & comp2 */

proc mixed data=space_data method=ml;

      class block year plot spatialpattern comp2;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*PPsum / noint solution;

      repeated year / subject=plot type=un group=comp2;

run;


/* FG model & comp2 */

proc mixed data=space_data method=ml;

      class block year plot spatialpattern comp2;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*PPwfg1 year*PPwfg2 year*PPwfg3 year*PPwfg4 year*PPbfg12
year*PPbfg13 year*PPbfg14 year*PPbfg23 year*PPbfg24
year*PPbfg34 / noint solution;

      repeated year / subject=plot type=un group=comp2;

run;


/* FG model with functional group interactions * spatial
pattern & comp2 */

proc mixed data=space_data method=ml;

      class block year plot spatialpattern comp2;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*PPwfg1*spatialpattern year*PPwfg2*spatialpattern
year*PPwfg3*spatialpattern year*PPwfg4*spatialpattern
year*PPbfg12*spatialpattern year*PPbfg13*spatialpattern
year*PPbfg14*spatialpattern year*PPbfg23*spatialpattern
year*PPbfg24*spatialpattern year*PPbfg34*spatialpattern /
noint solution;

      repeated year / subject=plot type=un group=comp2;
```

```
run;


/* FG model with functional group interactions * spatial
pattern & additive species interactions * legume % & comp2 */

proc mixed data=space_data method=ml;

      class block year plot spatialpattern comp2;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*PPwfg1*spatialpattern year*PPwfg2*spatialpattern
year*PPwfg3*spatialpattern year*PPwfg4*spatialpattern
year*PPbfg12*spatialpattern year*PPbfg13*spatialpattern
year*PPbfg14*spatialpattern year*PPbfg23*spatialpattern
year*PPbfg24*spatialpattern year*PPbfg34*spatialpattern
year*pq1*legumepc year*pq2*legumepc year*pq3*legumepc
year*pq4*legumepc year*pq5*legumepc year*pq6*legumepc
year*pq7*legumepc year*pq8*legumepc year*pq9*legumepc
year*pq10*legumepc year*pq11*legumepc year*pq12*legumepc
year*pq13*legumepc year*pq14*legumepc year*pq15*legumepc
year*pq16*legumepc / noint solution;

      repeated year / subject=plot type=un group=comp2;

run;



/* Include random effects in the best model */


/* First fit best model without random pairwise interactions
using REML to allow comparisons */

/* FG model with DE*Space & legumepc*lambdas & comp2 */

proc mixed data=space_data method=reml;

      class block year plot spatialpattern comp2;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*PPwfg1*spatialpattern year*PPwfg2*spatialpattern
year*PPwfg3*spatialpattern year*PPwfg4*spatialpattern
year*PPbfg12*spatialpattern year*PPbfg13*spatialpattern
year*PPbfg14*spatialpattern year*PPbfg23*spatialpattern
year*PPbfg24*spatialpattern year*PPbfg34*spatialpattern
year*pq1*legumepc year*pq2*legumepc year*pq3*legumepc
year*pq4*legumepc year*pq5*legumepc year*pq6*legumepc
year*pq7*legumepc year*pq8*legumepc year*pq9*legumepc
year*pq10*legumepc year*pq11*legumepc year*pq12*legumepc
year*pq13*legumepc year*pq14*legumepc year*pq15*legumepc
year*pq16*legumepc / noint solution;
```

```
      repeated year / subject=plot type=un group=comp2;
run;


/* Include random effects in individual years */


/* FG model with DE*Space & legumepc*lambdas & comp2 & Y1 REs
*/
proc mixed data=space_data method=reml;
      class block year plot spatialpattern comp2;
      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*PPwfg1*spatialpattern year*PPwfg2*spatialpattern
year*PPwfg3*spatialpattern year*PPwfg4*spatialpattern
year*PPbfg12*spatialpattern year*PPbfg13*spatialpattern
year*PPbfg14*spatialpattern year*PPbfg23*spatialpattern
year*PPbfg24*spatialpattern year*PPbfg34*spatialpattern
year*pq1*legumepc year*pq2*legumepc year*pq3*legumepc
year*pq4*legumepc year*pq5*legumepc year*pq6*legumepc
year*pq7*legumepc year*pq8*legumepc year*pq9*legumepc
year*pq10*legumepc year*pq11*legumepc year*pq12*legumepc
year*pq13*legumepc year*pq14*legumepc year*pq15*legumepc
year*pq16*legumepc / noint solution;
      repeated year / subject=plot type=un group=comp2;
      random Y1PP1-Y1PP120/ type=lin(1) ldata=random_pairwise
solution;
run;


/* FG model with DE*Space & legumepc*lambdas & comp2 & Y2 REs
*/
proc mixed data=space_data method=reml;
      class block year plot spatialpattern comp2;
      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*PPwfg1*spatialpattern year*PPwfg2*spatialpattern
year*PPwfg3*spatialpattern year*PPwfg4*spatialpattern
year*PPbfg12*spatialpattern year*PPbfg13*spatialpattern
year*PPbfg14*spatialpattern year*PPbfg23*spatialpattern
year*PPbfg24*spatialpattern year*PPbfg34*spatialpattern
year*pq1*legumepc year*pq2*legumepc year*pq3*legumepc
year*pq4*legumepc year*pq5*legumepc year*pq6*legumepc
year*pq7*legumepc year*pq8*legumepc year*pq9*legumepc
year*pq10*legumepc year*pq11*legumepc year*pq12*legumepc
```

```
year*pq13*legumepc year*pq14*legumepc year*pq15*legumepc
year*pq16*legumepc / noint solution;

      repeated year / subject=plot type=un group=comp2;

      random Y2PP1-Y2PP120/ type=lin(1) ldata=random_pairwise
solution;

run;


/* FG model with DE*Space & legumepc*lambdas & comp2 & Y3 REs
*/

proc mixed data=space_data method=reml;

      class block year plot spatialpattern comp2;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*PPwfg1*spatialpattern year*PPwfg2*spatialpattern
year*PPwfg3*spatialpattern year*PPwfg4*spatialpattern
year*PPbfg12*spatialpattern year*PPbfg13*spatialpattern
year*PPbfg14*spatialpattern year*PPbfg23*spatialpattern
year*PPbfg24*spatialpattern year*PPbfg34*spatialpattern
year*pq1*legumepc year*pq2*legumepc year*pq3*legumepc
year*pq4*legumepc year*pq5*legumepc year*pq6*legumepc
year*pq7*legumepc year*pq8*legumepc year*pq9*legumepc
year*pq10*legumepc year*pq11*legumepc year*pq12*legumepc
year*pq13*legumepc year*pq14*legumepc year*pq15*legumepc
year*pq16*legumepc / noint solution;

      repeated year / subject=plot type=un group=comp2;

      random Y3PP1-Y3PP120/ type=lin(1) ldata=random_pairwise
solution;

run;
```

Run the best model using REML to assess assumptions and model parameter estimates. The estimate statement shows how to predict from the model for a four-species centroid community.

```
/* FG model with DE*Space & legumepc*lambdas & comp2 & Y3 REs
*/

proc mixed data=space_data method=reml plots=all;

      class block year plot spatialpattern comp2;

      model WeedBiomass = year year*p1 year*p2 year*p3 year*p4
year*p5 year*p6 year*p7 year*p8  year*p9 year*p10 year*p11
```

```
year*p12 year*p13 year*p14 year*p15 /* year*p16 */ year*block
year*PPwfg1*spatialpattern year*PPwfg2*spatialpattern
year*PPwfg3*spatialpattern year*PPwfg4*spatialpattern
year*PPbfg12*spatialpattern year*PPbfg13*spatialpattern
year*PPbfg14*spatialpattern year*PPbfg23*spatialpattern
year*PPbfg24*spatialpattern year*PPbfg34*spatialpattern
year*pq1*legumepc year*pq2*legumepc year*pq3*legumepc
year*pq4*legumepc year*pq5*legumepc year*pq6*legumepc
year*pq7*legumepc year*pq8*legumepc year*pq9*legumepc
year*pq10*legumepc year*pq11*legumepc year*pq12*legumepc
year*pq13*legumepc year*pq14*legumepc year*pq15*legumepc
year*pq16*legumepc / noint solution;

     repeated year / subject=plot type=un group=comp2;

     random Y3PP1-Y3PP120/ type=lin(1) ldata=random_pairwise
solution;

     estimate ‘2014 disp species 1 5 9 15 avg block’ year 0 0
     1 year*p1 0 0 0.25 year*p2 0 0 0  year*p3 0 0 0 year*p4
     0 0 0 year*p5 0 0 0.25 year*p6 0 0 0 year*p7 0 0 0
     year*p8 0 0 0 year*p9 0 0 0.25 year*p10 0 0 0 year*p11 0
     0 0 year*p12 0 0 0 year*p13 0 0 0 year*p14 0 0 0
     year*p15 0 0 0.25 year*block 0 0 0.2 0 0 0.2 0 0 0.2 0 0
     0.2 0 0 0.2 year*PPwfg1*spatialpattern 0 0 0 0 0 0 0 0 0
     year*PPwfg2*spatialpattern 0 0 0 0 0 0 0 0 0
     year*PPwfg3*spatialpattern 0 0 0 0 0 0 0 0 0
     year*PPwfg4*spatialpattern 0 0 0 0 0 0 0 0 0
     year*PPbfg12*spatialpattern 0 0 0 0 0 0 0 0.0625 0
     year*PPbfg13*spatialpattern 0 0 0 0 0 0 0 0.0625 0
     year*PPbfg14*spatialpattern 0 0 0 0 0 0 0 0.0625 0
     year*PPbfg23*spatialpattern 0 0 0 0 0 0 0 0.0625 0
     year*PPbfg24*spatialpattern 0 0 0 0 0 0 0 0.0625 0
     year*PPbfg34*spatialpattern 0 0 0 0 0 0 0 0.0625 0
     year*pq1*legumepc 0 0 0.046875 year*pq2*legumepc 0 0 0
     year*pq3*legumepc 0 0 0 year*pq4*legumepc 0 0 0
     year*pq5*legumepc 0 0 0.046875 year*pq6*legumepc 0 0 0
     year*pq7*legumepc 0 0 0 year*pq8*legumepc 0 0 0
     year*pq9*legumepc 0 0 0.046875 year*pq10*legumepc 0 0 0
     year*pq11*legumepc 0 0 0 year*pq12*legumepc 0 0 0
     year*pq13*legumepc 0 0 0 year*pq14*legumepc 0 0 0
     year*pq15*legumepc 0 0 0.046875 year*pq16*legumepc 0 0
     0;

run;
```